

DOI: <https://doi.org/10.17721/1812-5409.2025/1>

Наведено результати теоретичних, прикладних і науково-експериментальних досліджень з актуальних питань фізики та математики.

Для науковців, викладачів, студентів та аспірантів.

ГОЛОВНИЙ РЕДАКТОР РЕДАКЦІЙНА КОЛЕГІЯ	Розора Ірина, д-р фіз.-мат. наук, доц. (Україна) Анікушин Андрій, канд. фіз.-мат. наук, доц. (Україна); Анісімов Анатолій, акад. НАН України, д-р фіз.-мат. наук, проф. (Україна); Анісімов Ігор, д-р фіз.-мат. наук, проф. (Україна); Бавула Володимир, проф. (Велика Британія); Бартлова Мілада, PhD (Чеська Республіка); Бегін Луїза, проф. (Італія); Безущак Оксана, д-р фіз.-мат. наук, проф. (Україна); Гаврюшенко Дмитро, д-р фіз.-мат. наук, проф. (Україна); Горячко Андрій, д-р фіз.-мат. наук, доц. (Україна); Джуліано Ріта, проф. (Італія); Дмитрук Ігор, д-р фіз.-мат. наук, проф. (Україна); Жук Ярослав, д-р фіз.-мат. наук, проф. (Україна); Заславський Володимир, д-р техн. наук, проф. (Україна); Затула Дмитро, канд. фіз.-мат. наук, доц. (заст. голов. ред.) (Україна); Капустян Олексій, д-р фіз.-мат. наук, проф. (Україна); Капустян Олена, д-р фіз.-мат. наук, ст. наук. співроб. (Україна); Кашпур Олена, д-р фіз.-мат. наук, доц. (Україна); Коган Ніколас, проф. (США); Кондратенко Сергій, д-р фіз.-мат. наук, проф. (Україна); Кудін Володимир, д-р техн. наук, ст. наук. співроб. (Україна); Леоненко Микола, проф. (Велика Британія); Львов Віктор, д-р фіз.-мат. наук, проф. (Україна); Майборода Ростислав, д-р фіз.-мат. наук, проф. (Україна); Макарець Микола, д-р фіз.-мат. наук, проф. (Україна); Медвідс Артурс, д-р наук, проф. (Латвія); Мішура Юлія, д-р фіз.-мат. наук, проф. (Україна); Оленко Андрій, проф. (Австралія); Орсінгер Енцо, проф. (Італія); Пашко Анатолій, д-р фіз.-мат. наук, проф. (Україна); Петравчук Анатолій, д-р фіз.-мат. наук, проф. (Україна); Погані Тібор, проф. (Хорватія); Погорілий Сергій, д-р техн. наук, проф. (Україна); Ронто Мікрос, д-р наук, проф. (Угорщина); Савенков Сергій, д-р фіз.-мат. наук, доц. (Україна); Сільвестров Дмитро, д-р наук, проф. (Швеція); Скришевський Валерій, д-р фіз.-мат. наук, проф. (Україна); Сливка-Тилищак Ганна, д-р фіз.-мат. наук, проф. (Україна); Соттінен Томмі, проф. (Фінляндія); Тору Аокі, проф. (Японія); Трофімчук Сергій, проф. (Чилі); Футорний Вячеслав, проф. (Бразилія); Худак Стефан, д-р наук, проф. (Словацька Республіка); Хусаїнов Денис, д-р фіз.-мат. наук, проф. (Україна); Яневич Тетяна, канд. фіз.-мат. наук, ст. наук. співроб. (відп. ред.), (Україна)
Адреса редколегії	Факультет комп'ютерних наук та кібернетики просп. Академіка Глушкова, 4д, м. Київ, 03022 ☎ (38044) 259 01 49 e-mail: physmath.bulletin@knu.ua web: https://bphm.knu.ua
Затверджено	вченою радою факультету комп'ютерних наук та кібернетики 18.03.25 (протокол № 11)
Зареєстровано	Національною радою України з питань телебачення і радіомовлення Рішення № 1089 від 28.03.24 Ідентифікатор друкованого медіа: R30-03805
Атестовано	Міністерством освіти і науки України (категорія А) Наказ № 641 від 28.04.25
Індексування	Scopus, Zentralblatt MATH (zbMATH), JournalTOCs, Crossref, ICI World of Journals, Road, Google Scholar
Засновник та видавець	Київський національний університет імені Тараса Шевченка Видавничо-поліграфічний центр "Київський університет" Свідоцтво про внесення до Державного реєстру ДК № 1103 від 31.10.02
Адреса видавця	ВПЦ "Київський університет" б-р Тараса Шевченка, 14, м. Київ, 01601 ☎ (38044) 239 32 22, 239 31 58, 239 31 28 e-mail: vpc@knu.ua

BULLETIN

TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV

ISSN 1812-5409 (Print), ISSN 2218-2055 (Online)

PHYSICS AND MATHEMATICS

1(80)/2025

Established in 2005

DOI: <https://doi.org/10.17721/1812-5409.2025/1>

The Bulletin highlights the results of theoretical, scientific, applied and scientific-experimental researches on the current issues of Physics and Mathematics.

For scientists, researchers, lecturers, PhD students and students.

EDITOR-IN-CHIEF	Rozora Iryna, DSc (Phys. & Math.), Assoc. Prof. (Ukraine)
EDITORIAL BOARD	Anikushyn Andrii, PhD (Phys. & Math.), Assoc. Prof. (Ukraine); Anisimov Anatoliy, Active Member of the National Academy of Sciences of Ukraine, DSc (Phys. & Math.), Prof. (Ukraine); Anisimov Igor, DSc (Phys. & Math.), Prof. (Ukraine); Bartlova Milada, PhD (Czech Republic); Bavula Vladimir, Prof. (Great Britain); Beghin Luisa, Prof. (Italy); Bezusach Oksana, DSc (Phys. & Math.), Prof. (Ukraine); Cogan Nicholas, Prof. (USA); Dmitruk Igor, DSc (Phys. & Math.), Prof. (Ukraine); Futorny Vyacheslav, Prof. (Brazil); Gavryushenko Dmytro, DSc (Phys. & Math.), Prof. (Ukraine); Giuliano Rita, Prof. (Italy); Goriachko Andrii, DSc (Phys. & Math.), Assoc. Prof. (Ukraine); Hudak Stefan, DSc, Prof. (Slovak Republic); Husainov Denys, DSc (Phys. & Math.), Prof. (Ukraine); Ianevych Tetiana, PhD (Phys. & Math.), Senior Researcher (Managing Editor) (Ukraine); Kashpur Olena, DSc (Phys. & Math.), Assoc. Prof. (Ukraine); Kapustian Olena, DSc (Phys. & Math.), Senior Researcher (Ukraine); Kapustyan Oleksiy, DSc (Phys. & Math.), Prof. (Ukraine); Kondratenko Serhiy, DSc (Phys. & Math.), Prof. (Ukraine); Kudin Volodymyr, DSc, (Engin.), Senior Researcher (Ukraine); Leonenko Nikolay, Prof. (Great Britain); Lvov Victor, DSc (Phys. & Math.), Prof. (Ukraine); Maiboroda Rostyslav, DSc (Phys. & Math.), Prof. (Ukraine); Makarets' Nikolay, DSc (Phys. & Math.), Prof. (Ukraine); Medvids Arturs, DSc (habil.), Prof. (Latvia); Mishura Yuliya, DSc (Phys. & Math.), Prof. (Ukraine); Olenko Andriy, Prof. (Australia); Orsingher Enzo, Prof. (Italy); Pashko Anatolij, DSc (Phys. & Math.), Prof. (Ukraine); Petravchuk Anatolij, DSc (Phys. & Math.), Prof. (Ukraine); Pogány Tibor, Prof. (Croatia); Pogorily Serhiy, DSc (Engin.), Prof. (Ukraine); Rontó Miklós, DSc, Prof. (Hungary); Savenkov Serhij, DSc (Phys. & Math.), Assoc. Prof. (Ukraine); Silvestrov Dmitrii, DSc, Prof. (Sweden); Skryshevsky Valery, DSc (Phys. & Math.), Prof. (Ukraine); Slyvka-Tylyshchak Anna, DSc (Phys. & Math.), Prof. (Ukraine); Sottinen Tommi, Prof. (Finland); Toru Aoki, Prof. (Japan); Trofimchuk Sergey, Prof. (Chile); Zaslav'skyj Volodymyr, DSc (Engin.), Prof. (Ukraine); Zatula Dmytro, PhD (Phys. & Math.), Assoc. Prof. (Deputy Editor-in-Chief) (Ukraine); Zhuk Yaroslav, DSc (Phys. & Math.), Prof. (Ukraine)
Address	Faculty of Computer Science and Cybernetics 4d, Hlushkov ave., Kyiv, 03022 ☎ (38044) 259 01 49 e-mail: physmath.bulletin@knu.ua web: https://bphm.knu.ua
Approved by	the Academic Council of the Faculty of Computer Science and Cybernetics 18.03.25 (protocol № 11)
Registered by	the National Council of Television and Radio Broadcasting of Ukraine Decision № 1089 of 28.03.24 Identifier of printed media: R30-03805
Certified by	the Ministry of Education and Science of Ukraine (category A) Order № 641 dated 28.04.25
Indexing	Scopus, Zentralblatt MATH (zbMATH), JournalTOCs, Crossref, ICI World of Journals, Road, Google Scholar
Founder and publisher	Taras Shevchenko National University of Kyiv Publishing and Polygraphic Center "Kyiv University" Certificate of entry into the State Register ДК № 1103 dated 31.10.02
Address	PPC "Kyiv University" 14, Taras Shevchenko blvd., Kyiv, 01601 ☎ (38044) 239 32 22, 239 31 58, 239 31 28 e-mail: vpc@knu.ua

ЗМІСТ

БЕЗУЩАК Оксана, БОДНАРЧУК Ірина, ГОРОДНІЙ Михайло, КУКУШ Олександр, МІШУРА Юлія, НЕСТЕРЕНКО Олексій, СТАНЖИЦЬКИЙ Олександр, ЧАЙКОВСЬКИЙ Андрій, ШЕВЧУК Ігор, ЯМНЕНКО Ростислав До 60-річчя від дня народження Вадима Миколайовича Радченка.....	5
--	---

АЛГЕБРА, ГЕОМЕТРІЯ ТА ТЕОРІЯ ЙМОВІРНОСТЕЙ

БОРИСЕНКО Олександр, БОРИСЕНКО Ольга Стохастична SIR модель із пуассонівськими шумами.....	8
ВАСИЛИК Ольга, ШУРУБУРА Костянтин Оптимізація моделі ф-субгауссового узагальненого дробового броунівського руху.....	17
ЗУБЧЕНКО Володимир, ЗЕЛЕНЬКО Петро Тестування ентропійної міри ризику EVaR на реальних даних.....	29
ЛЕВЧЕНКО Володимир, ЯМНЕНКО Ростислав Оцінювання ймовірності банкрутства для випадкової суми негативно біноміально розподіленого числа субгауссових позовів.....	40
МАКАРЧУК Олег, ХАЛЕЦЬКИЙ Богдан Структура розподілу одного класу узагальнених згорток Бернуллі.....	47
ХУАНГ Хенінг Теорія інформованості: новий підхід до теорії ймовірностей.....	53

ДИФЕРЕНЦІАЛЬНІ РІВНЯННЯ, МАТЕМАТИЧНА ФІЗИКА ТА МЕХАНІКА

КІПНІС Олександр Стиск шаруватого композиту з компонентами, що проковзують, вздовж двох паралельних міжфазних структурних тріщин.....	60
КУБЕНКО Веніамін, ЯНЧЕВСЬКИЙ Ігор, ОСТОС Олександр Динамічні характеристики ідеальної стисливої рідини, що збуджується сферичним сегментом в циліндричній порожнині.....	64
САНДРАКОВ Геннадій, КИРИЧЕК Микола Моделювання і гомогенізація процесів у циліндричних шаруватих композитах таких як конденсатори та трансформатори.....	71
ХОРОШЕВ Костянтин, ДУЧЕНКО Кирил, КИКОТЬ Сергій Кінетостатичний аналіз діад Ассура.....	83

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

АЙЗЕНБЕРГ Ігор, ВАСЬКО Олександр Порівняльний аналіз згорток отриманих мережами CNNMVN та MLMVN як згорткової мережі в частотній області.....	89
АТОЄВ Констянтин, КНОПОВ Павло Шестисекторна модель Лоренца для управління ризиками критичної інфраструктури.....	97
БАЛАК Юрій Інформаційна система на основі комплексної моделі з використанням машинного навчання для спектрального аналізу.....	104
ДОЛЕНКО Галина, МАНОВИЦЬКА Дарія, СУББОТІНА Олена, ЧАПЛІНСЬКИЙ Юрій, БУДЯНСЬКИЙ Ігор Розробка чат-боту служби підтримки клієнтів мережі зарядних станцій електромобілів.....	115
КЛЮШИН Дмитро, МАЙСТРЕНКО Олександр Використання нейронів Іжикевича в моделях Гопфілда.....	122
МЕЛЬНИК Ігор, ПОЧИНОК Аліна, СКРИПКА Михайло, ДЕМ'ЯНЧЕНКО Ольга Метод інтерполяції з використанням коренево-дробово-раціональних функцій різного порядку.....	130
ОСЬМАК Євген Розроблення моделі прийняття рішень щодо пріоритетів виконання робіт в умовах обмежених ресурсів для випадку сервісної компанії.....	139
ПІКУЛЬСЬКИЙ Ростислав Аналіз та порівняння методологій відкладеного рендерингу.....	144
ПОЛІЩУК Володимир, МЛАВЕЦЬ Юрій, ДАНКО Лукаш, ВЛОХ Івона Технологія оцінювання та вибору коворкінг-просторів відповідно до цільових потреб клієнтів.....	148
СКЛЯРЕНКО Олена, ПАШКО Анатолій, ЛИТВИНЕНКО Леонід, КОЛОДІНСЬКА Яніна Моделі глибокого навчання для вирішення задачі порівняння неструктурованої текстової інформації.....	157
ФРАНКІВ Олександр Машинне навчання в покращенні візуалізації просторової моделі архітектури програми.....	164
ШКІЛЬНЯК Оксана Першопорядкові транзитивні модальні логіки квазіарних предикатів з рівністю.....	174

СУЧАСНА ФІЗИКА

ЧОЛІЙ Ярослав, МАКАРЕЦЬ Микола Вплив моделі пружних іон-атомних зіткнень на кумулянти просторового розподілу імплантованих іонів.....	182
--	-----

CONTENTS

BEZUSHCHAK Oksana, BODNARCHUK Iryna, HORODNII Mykhailo, KUKUSH Alexander, MISHURA Yuliya, NESTERENKO Oleksii, STANZHYTSKYI Oleksandr, CHAIKOVSKYI Andrii, SHEVCHUK Igor, YAMNENKO Rostyslav To the 60th anniversary of the birth of Vadym Mykolayovych Radchenko	5
--	---

ALGEBRA, GEOMETRY AND PROBABILITY THEORY

BORYSENKO Oleksandr, BORYSENKO Olga Stochastic SIR model with poisson noises	8
VASYLYK Olga, SHURUBURA Kostiantyn Optimization of the model of ϕ -sub-Gaussian generalized fractional Brownian motion	17
ZUBCHENKO Volodymyr, ZELENKO Petro Entropic risk measure EVaR real data testing	29
LEVCHENKO Volodymyr, YAMNENKO Rostyslav Estimation of ruin probability for random sum of negative binomially distributed number of sub-Gaussian claims	40
MAKARCHUK Oleg, KHALETSKY Bohdan Structure of the distribution of one class of generalized Bernoullie convolutions	47
HUANG Hening The theory of informity: a novel probability framework	53

DIFFERENTIAL EQUATIONS, MATHEMATICAL PHYSICS AND MECHANICS

KIPNIS Alexander Compression of a layered composite with sliding components along two parallel interface structural cracks	60
KUBENKO Veniamin, YANCHEVSKYI Ihor, OSTOS Oleksandr Dynamic characteristics of an ideal compressible fluid excited by a spherical segment in a cylindrical cavity	64
SANDRAKOV Gennadiy, KYRYCHEK Mykola Modeling and homogenization of processes in cylindrical layered composites like capacitors and transformers	71
KHOROSHEV Kostiantyn, DUCHENKO Kyryl, KYKOT Serhii Kinetostatic analysis of Assur dyads	83

COMPUTER SCIENCES AND INFORMATICS

AIZENBERG Igor, VASKO Alexander Comparative analysis of CNNMVN and MLMVN as frequency domain CNN convolutions	89
ATOYEV Konstantin, KNOPOV Pavlo Six-sector Lorenz model for critical infrastructure risk management	97
BILAK Yuri Information system based on a complex model using machine learning for spectral analysis	104
DOLENKO Galyna, MANOVYTSKA Dariia, SUBBOTINA Olena, CHAPLINSKYI Yuriy, BUDYANSKY Igor Chatbot for customer support of a company which operates charging stations for electric cars	115
KLYUSHIN Dmitriy, MAISTRENKO Oleksandr Use of Izhikevich neurons in Hopfield models	122
MELNYK Igor, POCHYNOK Alina, SKRYPKA Mykhailo, DEMYANCHENKO Olga Method of interpolation using root-fractional-rational functions of different orders	130
OSMAK Yevhen Modelling of the decision-making process of jobs priorities definition under conditions of limited resources: case of a service company	139
PIKULSKY Rostyslav Deferred rendering methodologies analysis and comparison	144
POLISHCHUK Volodymyr, MLAVETS Yuri, DANKO Lukáš, WŁOCH Iwona Technology for evaluating and selecting coworking spaces according to clients' target needs	148
SKLIARENKO Olena, PASHKO Anatolii, LYTVYNNENKO Leonid, KOLODINSKA Yanina Deep learning models for solving the problem of comparing unstructured textual information	157
FRANKIV Oleksandr Machine learning in enhancing visualization of the spatial software architecture model	164
SHKILNIAK Oksana First-order transitional modal logics of quasiary predicates with equality	174

MODERN PHYSICS

CHOLIY Yaroslav, MAKARETS Mykola, Influence of the model of elastic ion-atomic collisions on the cumulants of the spatial distribution of implanted ions	182
---	-----

УДК 519.21

DOI: <https://doi.org/10.17721/1812-5409.2025/1.1>

Оксана БЕЗУЩАК¹, д-р фіз.-мат. наук, проф.
ORCID ID: 0000-0003-3654-6753
e-mail: bezushchak@knu.ua

Ірина БОДНАРЧУК¹, канд. фіз.-мат. наук, асист.
ORCID ID: 0000-0002-4713-5274
e-mail: ibodnarchuk@knu.ua

Михайло ГОРОДНІЙ¹, д-р фіз.-мат. наук, проф.
ORCID ID: 0000-0002-9991-910X
e-mail: horodnii@knu.ua

Олександр КУКУШ¹, д-р фіз.-мат. наук, проф.
ORCID ID: 0000-0002-4143-8928
e-mail: oleksandr.kukush@knu.ua

Юлія МІШУРА¹, д-р фіз.-мат. наук, проф.
ORCID ID: 0000-0002-6877-1800
e-mail: yuliyamishura@knu.ua

Олексій НЕСТЕРЕНКО¹, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0003-4081-1206
e-mail: oleksiinesterenko@knu.ua

Олександр СТАНЖИЦЬКИЙ¹, д-р фіз.-мат. наук, проф.
ORCID ID: 0000-0002-1456-729X
e-mail: stanzhytskyi@knu.ua

Андрій ЧАЙКОВСЬКИЙ¹, д-р фіз.-мат. наук, доц.
ORCID ID: 0000-0002-3357-8744
e-mail: andriichaikovskiy@knu.ua

Ігор ШЕВЧУК¹, д-р фіз.-мат. наук, акад. НАН України
ORCID ID: 0000-0003-1140-373X
e-mail: shevchukio@knu.ua

Ростислав ЯМНЕНКО¹, д-р фіз.-мат. наук, доц.
ORCID ID: 0000-0002-9612-7959
e-mail: rostyslav.yamnenko@knu.ua

¹Київський національний університет імені Тараса Шевченка, Київ, Україна

ДО 60-РІЧЧЯ ВІД ДНЯ НАРОДЖЕННЯ ВАДИМА МИКОЛАЙОВИЧА РАДЧЕНКА

Присвячено 60-річчю відомого українського математика, професора Київського національного університету імені Тараса Шевченка Вадима Миколайовича Радченка. Висвітлено його творчий шлях, наукову та педагогічну діяльність, а також вагомий внесок у розвиток математичної освіти в Україні, зокрема в олімпіадний рух. Особливу увагу приділено багаторічній роботі професора Радченка з обдарованими школярами та значній ролі у підготовці українських команд до міжнародних математичних олімпіад.

Ключові слова: Вадим Радченко, математичні олімпіади, стохастична міра, стохастичний аналіз, стохастичні диференціальні рівняння з частинними похідними.

Класифікація відповідно до AMS 2020: 60G57, 60H05, 60H15.

2 грудня 2024 р. відсвяткував своє 60-річчя відомий український математик, знаний фахівець зі стохастичного аналізу, заслужений вчитель України, доктор фізико-математичних наук, професор Вадим Миколайович Радченко.

Вадим Миколайович народився 2 грудня 1964 р. в Запоріжжі у родині інженерів. У 1977 р. він закінчив 7-й клас середньої школи № 91 м. Запоріжжя і вступив до 8-го класу Республіканської спеціалізованої школи-інтернату фізико-математичного профілю при Київському державному університеті імені Т. Г. Шевченка. Під час навчання у школі Вадим Миколайович активно брав участь у багатьох олімпіадах з математики й неодноразово здобував на них призові місця, зокрема, тричі був призером Всесоюзної математичної олімпіади. У 1980 р. він закінчив школу, а в 1985 р. – механіко-математичний факультет Київського державного університету імені Т. Г. Шевченка, отримавши диплом з відзнакою. Під час навчання в університеті йому пощастило слухати лекції видатних учених і педагогів: М. Й. Ядренка (теорія ймовірностей та математична статистика), А. М. Самойленка (диференціальні рівняння), М. Л. Горбачука (функціональний аналіз). Однак визначальний вплив на студента Радченка здійснив курс теорії випадкових процесів, який викладав академік НАН України А. В. Скороход. Під керівництвом Анатолія Володимировича В. М. Радченко писав дипломну роботу, навчався з 1985 р. по 1988 р. в аспірантурі, захистив у 1989 р. кандидатську дисертацію "Випадкові міри та випадкові лінійні функціонали" та у 2007 р. докторську дисертацію "Інтеграли за загальними випадковими мірами".



© Безущак Оксана, Боднарчук Ірина, Городній Михайло, Кукуш Олександр, Мішура Юлія, Нестеренко Олексій, Станжицький Олександр, Чайковський Андрій, Шевчук Ігор, Ямненко Ростислав, 2025

У 1989–1994 рр. Вадим Миколайович працював у Запорізькому державному університеті на посадах наукового співробітника (1989–1990 рр.), асистента (1990–1991 рр.), старшого викладача (1991–1992 рр.) та доцента (1992–1994 рр.). З 1994 р. науково-педагогічна діяльність ювіляра нерозривно пов'язана з механіко-математичним факультетом Київського національного університету імені Тараса Шевченка: у 1994–1997 рр. – докторант механіко-математичного факультету, у 1997–1999 рр. – доцент кафедри загальної математики, у 1999–2008 рр. – доцент, а із 2008 р. – професор кафедри математичного аналізу. У 2001 р. він отримав вчене звання доцента, а у 2013 р. – професора.

Наукові дослідження В. М. Радченко розпочав під керівництвом О. І. Клесова: результати курсової роботи за третій курс були опубліковані в 1985 р. в журналі "Теорія ймовірностей та математична статистика". Наукові праці Вадима Миколайовича присвячені різноманітним питанням теорії випадкових процесів. Основними напрямками його наукової діяльності є дослідження загальних стохастичних мір, рівнянь з такими мірами, властивості траєкторій випадкових процесів. Зокрема, він довів, що неперервні траєкторії стохастичних мір належать просторам Бессова, можуть бути зображені випадковим рядом Фур'є. Разом з учнями показав, що рівняння теплопровідності та хвильове рівняння, керовані стохастичними мірами, мають єдині розв'язки, що задовольняють умову Гельдера; було також отримано принцип усереднення для цих рівнянь. В. М. Радченко запропонував конструкцію симетричного стохастичного інтеграла від випадкової функції за загальною стохастичною мірою та досліджено рівняння з такими інтегралами.

В. М. Радченко – автор двох монографій (одна з яких видана за кордоном), понад 60 наукових праць, з яких більше 40 проіндексовані в наукометричній базі Scopus.

Він був стипендіатом престижного фонду Александра фон Гумбольдта (Німеччина), є членом Київського математичного товариства, членом спеціалізованої вченої ради із захисту докторських дисертацій.

Непересічною, багатогранною і плідною є й педагогічна діяльність професора Радченка. Протягом понад двадцяти років він щорічно бездоганно, на найвищому науково-методичному рівні, читає нормативний курс теорії міри та інтеграла, а також низку нормативних і спеціальних курсів як теоретичного, так і прикладного характеру. Керує він і написанням курсових та кваліфікаційних робіт, підготував трьох кандидатів наук. Автор підручника з теорії міри, співавтор збірника задач із фінансової математики та багатьох чудових методичних розробок, які широко використовують у навчальному процесі. Уміло поєднуючи високий науковий рівень лекцій із чіткістю, ясністю та доступністю викладу, високу вимогливість і принциповість в оцінюванні з чуйним і турботливим ставленням до студентів, Вадим Миколайович є взірцем для наслідування молодшими колегами.

В. М. Радченко є, без перебільшення, видатним діячем математичного олімпіадного руху в Україні. Маючи успішний досвід учасника шкільних і студентських олімпіад з математики, керівника гуртків, набуваючи організаційного досвіду під опікою корифея олімпіадного руху члена-кореспондента НАН України професора М. Й. Ядренка, Вадим Миколайович разом із В. М. Лейфуру, І. М. Мітельманом, В. А. Ясінським та іншими сформував творчий колектив, який у перші 15 років після проголошення незалежності України відповідав за укладання завдань заключного етапу Всеукраїнської учнівської олімпіади з математики та формування команди України на Міжнародні математичні олімпіади. Вадим Миколайович був (історично першим!) науковим керівником команд України на Міжнародних математичних олімпіадах у 1992–2002 рр., членом і заступником голови журі численних всеукраїнських і регіональних математичних змагань для школярів. Він є співавтором видань енциклопедичного характеру "Математичні олімпіади школярів України. 1991–2000 рр.", "Математичні олімпіади школярів України. 2001–2006 рр.", а також багатьох інших збірників олімпіадних завдань і посібників для підготовки до олімпіад. І нині він не полишає працювати з обдарованими школярами: веде гуртки в різних школах, зокрема, й у своїй рідній – Українському фізико-математичному ліцеї. Вихованці цих гуртків займають призові місця на всеукраїнських та міжнародних математичних олімпіадах, а також інших престижних математичних змаганнях.

За свою багаторічну працю В. М. Радченко нагороджений Грамотою (2014) і Почесною грамотою (2020) Київського національного університету імені Тараса Шевченка, нагрудними знаками МОН України "Відмінник освіти України" (2000) та "За наукові та освітні досягнення" (2024), удостоєний почесного звання "Заслужений вчитель України" (2006).

Колектив механіко-математичного факультету щиро вітає Вадима Миколайовича зі славним ювілеєм і зичить йому міцного здоров'я, наснаги та нових творчих успіхів.

Вибрані роботи В. М. Радченка

- Борисенко, О. Д., Мішура, Ю. С., Радченко, В. М., & Шевченко, Г. М. (2007). *Збірник задач з фінансової математики*. Техніка.
- Конет, І. М., Радченко, В. М., & Теплінський, Ю. В. (2010). *Обласні олімпіади з математики*. Абетка.
- Лейфура, В. М., Мітельман, І. М., Радченко, В. М., & Ясінський, В. М. (2003). *Математичні олімпіади школярів України: 1991–2000 рр.* Техніка.
- Лейфура, В. М., Мітельман, І. М., Радченко, В. М., & Ясінський, В. А. (2008). *Математичні олімпіади школярів України: 2001–2006 рік*. Каменяр.
- Radchenko, V. M. (2007). Besov regularity of stochastic measures. *Statistics & Probability Letters*, 77(8), 822–825.
- Radchenko, V. (2009). Mild solution of the heat equation with a general stochastic measure. *Studia Mathematica*, 194, 231–251.
- Radchenko, V. (2016). Stratonovich-type integral with respect to a general stochastic measure. *Stochastics*, 88(7), 1060–1072.
- Radchenko, V. M. (2019). Averaging principle for equation driven by a stochastic measure. *Stochastics*, 91(6), 905–915.
- Radchenko, V. M. (2022). *General stochastic measures: integration, path properties and equations*. John Wiley & Sons.

Отримано редакцією журналу / Received: 05.02.25

Прорецензовано / Revised: 01.03.25

Схвалено до друку / Accepted: 06.03.25

Oksana BEZUSHCHAK¹, DSc (Phys. & Math.), Prof.
ORCID ID: 0000-0003-3654-6753
e-mail: bezushchak@knu.ua

Iryna BODNARCHUK¹, PhD (Phys. & Math.), Assist. Prof.
ORCID ID: 0000-0002-4713-5274
e-mail: ibodnarchuk@knu.ua

Mykhailo HORODNII¹, DSc (Phys. & Math.), Prof.
ORCID ID: 0000-0002-9991-910X
e-mail: horodnii@knu.ua

Alexander KUKUSH¹, DSc (Phys. & Math.), Prof.
ORCID ID: 0000-0002-4143-8928
e-mail: oleksandr.kukush@knu.ua

Yuliya MISHURA¹, DSc (Phys. & Math.), Prof.
ORCID ID: 0000-0002-6877-1800
e-mail: yuliyamishura@knu.ua

Oleksii NESTERENKO¹, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0003-4081-1206
e-mail: oleksiinesterenko@knu.ua

Oleksandr STANZHYTSKYI¹, DSc (Phys. & Math.), Prof.
ORCID ID: 0000-0002-1456-729X
e-mail: stanzhytskyi@knu.ua

Andrii CHAIKOVSKYI¹, DSc (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0002-3357-8744
e-mail: andriichaikovskiyi@knu.ua

Igor SHEVCHUK¹, DSc (Phys. & Math.), Active Member of the National Academy of Sciences of Ukraine
ORCID ID: 0000-0003-1140-373X
e-mail: shevchukio@knu.ua

Rostyslav YAMNENKO¹, DSc (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0002-9612-7959
e-mail: rostyslav.yamnenko@knu.ua

¹Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

TO THE 60TH ANNIVERSARY OF THE BIRTH OF VADYM MYKOLAYOVYCH RADCHENKO

The article is dedicated to the 60th anniversary of the birth of Vadym Mykolaiovych Radchenko, a famous Ukrainian mathematician, professor at Taras Shevchenko National University of Kyiv. The article highlights his academic career, scientific and pedagogical activities, as well as his significant contribution to the development of mathematics education in Ukraine, in particular, the Olympiad movement. Special attention is paid to professor Radchenko's longstanding work with gifted students and his significant role in preparing Ukrainian teams for the International Mathematical Olympiads.

K e y w o r d s : *Vadym Radchenko, mathematical olympiads, stochastic measure, stochastic analysis, stochastic partial differential equations.*

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

АЛГЕБРА, ГЕОМЕТРІЯ ТА ТЕОРІЯ ЙМОВІРНОСТЕЙ

UDC 519.2

DOI: <https://doi.org/10.17721/1812-5409.2025/1.2>

Oleksandr BORYSENKO, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0002-6670-8605
e-mail: borysenko_oleksandr@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Olga BORYSENKO, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0002-2099-2911
e-mail: oborisenko1373@gmail.com

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

STOCHASTIC SIR MODEL WITH POISSON NOISES

Due to the negative impact of infectious diseases on population growth, it is important to understand the dynamic behavior of such diseases. Mathematical deterministic SIR models are widely used to study the spread of infectious diseases. In real life, there is a lot of randomness and stochasticity such as environmental noise, so using stochastic models is more suitable. Suppose we want to take into account abrupt environmental perturbations, such as epidemics, fires, earthquakes, etc. in the considered models. In that case, we must introduce Poisson noises into the population models for describing such discontinuous systems. The existence and uniqueness of the global positive solution are proved for the system of stochastic differential equations describing a non-autonomous SIR model disturbed by white noise, centered and non-centered Poisson noises. In the deterministic case there is a threshold of the system for an epidemic to occur, so called the basic reproduction number. Depending on the value of the reproduction number there is the disease-free equilibrium, or there is an endemic equilibrium, which implies the disease always remains. In the case of the autonomous stochastic SIR model, we study the asymptotic behavior of the solution to the corresponding system of stochastic differential equations around these points of equilibrium.

Keywords: stochastic SIR model, environmental noise, centered and non-centered Poisson noises, global positive solution, disease-free equilibrium, endemic equilibrium, asymptotic behavior.

AMS 2020 classification: 60H10, 92D30.

Introduction

Due to the negative impact of infectious diseases on population growth, it is important to understand the dynamic behavior of such diseases. Mathematical deterministic SIR models are widely used to study the spread of infectious diseases. Let $S(t)$ denote the number of individuals susceptible to the disease, $I(t)$ denote the number of infected members and $R(t)$ denote the recovered members, who have been removed from the possibility of infection through full immunity. One of the basic deterministic SIR models of Kermack and McKendrick is presented and studied in (Anderson, & May, 1979). This deterministic SIR model has a form

$$\begin{aligned}\frac{dS(t)}{dt} &= (\Lambda - \beta S(t)I(t) - \mu S(t)), \\ \frac{dI(t)}{dt} &= \beta S(t)I(t) - (\mu + \varepsilon + \rho)I(t), \\ \frac{dR(t)}{dt} &= \rho I(t) - \mu R(t),\end{aligned}\tag{1}$$

where all parameters are positive constants, Λ is the influx of individuals into the susceptibles, μ is the natural death rate, β is the average number of contacts per infective per day (so-called "transmission coefficient"), ε is the death rate of infective caused by the disease, ρ is the rate of recovery from infection.

If the basic reproduction number $R_0 = \frac{\beta\Lambda}{\mu(\mu+\varepsilon+\rho)}$ we have $R_0 \leq 1$, then model (1) has only equilibrium $E_0 = (\frac{\Lambda}{\mu}, 0, 0)$, which is globally asymptotically stable. In the case $R_0 > 1$ there exists a global asymptotically stable endemic equilibrium $E^* = (\frac{\mu+\varepsilon+\rho}{\beta}, \frac{\Lambda}{\mu+\varepsilon+\rho} - \frac{\mu}{\beta}, \frac{\Lambda\rho}{\beta(\mu+\varepsilon+\rho)} - \frac{\rho}{\beta})$. This means that disease always remains.

In (Jiang et al., 2011) the autonomous stochastic SIR model was studied. This model is driven by a system of stochastic differential equations

$$\begin{aligned}dS(t) &= (\Lambda - \beta S(t)I(t) - \mu S(t))dt + \sigma_1 S(t)dw_1(t), \\ dI(t) &= (\beta S(t)I(t) - (\mu + \varepsilon + \rho)I(t))dt + \sigma_2 I(t)dw_2(t), \\ dR(t) &= (\rho I(t) - \mu R(t))dt + \sigma_3 R(t)dw_3(t),\end{aligned}\tag{2}$$

where $w_i(t)$, $i = 1, 2, 3$ are independent one-dimensional Wiener processes. The existence and uniqueness of a global positive solution of model (2) are proved. The authors study the asymptotic behavior of the solution to the system (2) around E_0 and E^* respectively. In the paper by (Zhang, & Wang, 2013) the analogous results obtained for the autonomous stochastic SIR model with jumps, generated by centered Poisson measure

$$\begin{aligned}dS(t) &= (\Lambda - \beta S(t)I(t) - \mu S(t))dt + \sigma_1 S(t)dw_1(t) + \int_{\mathbb{Z}} C_1(z)S(t^-)\tilde{N}(dt, dz), \\ dI(t) &= (\beta S(t)I(t) - (\mu + \varepsilon + \rho)I(t))dt + \sigma_2 I(t)dw_2(t) + \int_{\mathbb{Z}} C_2(z)I(t^-)\tilde{N}(dt, dz), \\ dR(t) &= (\rho I(t) - \mu R(t))dt + \sigma_3 R(t)dw_3(t) + \int_{\mathbb{Z}} C_3(z)R(t^-)\tilde{N}(dt, dz),\end{aligned}$$

where $X(t^-)$ means the left limit of $X(t)$, $\tilde{N}(dt, dz)$ is the centered Poisson measure independent on Wiener processes $w_i(t)$, $i = 1, 2, 3$, $Z \subseteq [0, +\infty)$ is a measurable subset. So considered model takes into account only "small" jumps corresponding to the integral with respect to the centered Poisson measure.

The impact of centered and non-centered Poisson noises to the stochastic non-autonomous stochastic predator-prey model that depends on the population density of the predator is studied in the paper (Borysenko, O. D., & Borysenko, O. V., 2022). The authors proved the existence and uniqueness theorem for the positive, global (no explosions in the finite time) solution of the considered system is proved. It is obtained sufficient conditions of stochastic ultimate boundedness and stochastic permanence in the considered stochastic predator-prey model.

In this paper, we consider the non-autonomous SIR model, disturbed by white noise and jumps generated by centered and non-centered Poisson measures. So, the novelty of the considered model is the following: we investigate the non-autonomous stochastic SIR model and we take into account not only "small" jumps, corresponding to the centered Poisson measure but also the "large" jumps, corresponding to the non-centered Poisson measure.

This model is driven by the system of stochastic differential equations

$$\begin{aligned} dS(t) &= (\Lambda(t) - \beta(t)S(t)I(t) - \mu(t)S(t))dt + \sigma_1(t)S(t)dw_1(t) + \\ &+ \int_{\mathbb{R}} \gamma_1(t, z)S(t^-)\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_1(t, z)S(t^-)v_2(dt, dz), \\ dI(t) &= (\beta(t)S(t)I(t) - (\mu(t) + \varepsilon(t) + \rho(t))I(t))dt + \sigma_2(t)I(t)dw_2(t) + \\ &+ \int_{\mathbb{R}} \gamma_2(t, z)I(t^-)\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_2(t, z)I(t^-)v_2(dt, dz), \\ dR(t) &= (\rho(t)I(t) - \mu(t)R(t))dt + \sigma_3(t)R(t)dw_3(t) + \\ &+ \int_{\mathbb{R}} \gamma_3(t, z)R(t^-)\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_3(t, z)R(t^-)v_2(dt, dz), \end{aligned} \quad (3)$$

where $v_i(t, A)$, $i = 1, 2$ are independent Poisson measures, which are independent on $w_i(t)$, $i = 1, 2, 3$, $\tilde{v}_1(t, A) = v_1(t, A) - t\Pi_1(A)$, $E[v_i(t, A)] = t\Pi_i(A)$, $i = 1, 2$, $\Pi_i(A)$, $i = 1, 2$ are finite measures on the Borel sets A in $\mathbb{R}$.

We prove the existence and uniqueness of the global positive solution to the non-autonomous system (3). For the autonomous system

$$\begin{aligned} dS(t) &= (\Lambda - \beta S(t)I(t) - \mu S(t))dt + \sigma_1 S(t)dw_1(t) + \int_{\mathbb{R}} \gamma_1(z)S(t^-)\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_1(z)S(t^-)v_2(dt, dz), \\ dI(t) &= (\beta S(t)I(t) - (\mu + \varepsilon + \rho)I(t))dt + \sigma_2 I(t)dw_2(t) + \int_{\mathbb{R}} \gamma_2(z)I(t^-)\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_2(z)I(t^-)v_2(dt, dz), \\ dR(t) &= (\rho I(t) - \mu R(t))dt + \sigma_3 R(t)dw_3(t) + \int_{\mathbb{R}} \gamma_3(z)R(t^-)\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_3(z)R(t^-)v_2(dt, dz) \end{aligned} \quad (4)$$

we study the asymptotic behavior of the solution to the system (4) around E_0 and E^* respectively.

1. Main results

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $w_i(t)$, $i = 1, 2, 3$, $t \geq 0$, are mutually independent standard one-dimensional Wiener processes on $(\Omega, \mathcal{F}, P)$, and $v_i(t, A)$, $i = 1, 2$, are mutually independent Poisson measures defined on $(\Omega, \mathcal{F}, P)$ independent on $w_i(t)$, $i = 1, 2, 3$, $t \geq 0$. Here $E[v_i(t, A)] = t\Pi_i(A)$, $i = 1, 2$, $\tilde{v}_i(t, A) = v_i(t, A) - t\Pi_i(A)$, $i = 1, 2$, $\Pi_i(A)$, $i = 1, 2$ are finite measures on the Borel sets in $\mathbb{R}$. On the probability space $(\Omega, \mathcal{F}, P)$ we consider an increasing, right continuous family of complete sub- σ -algebras $\{\mathcal{F}_t\}_{t \geq 0}$, where $\mathcal{F}_t = \sigma\{w_i(s), v_j(s, A), s \leq t, i = 1, 2, 3, j = 1, 2\}$.

1.1. Existence and uniqueness of a global positive solution

Assumption 1. It is assumed, that $\Lambda(t), \beta(t), \mu(t), \varepsilon(t), \rho(t)$ are positive, bounded, continuous functions, $\sigma_i(t)$, $i = 1, 2, 3$, $\gamma_j(t, z), \delta_j(t, z)$, $j = 1, 2$ are continuous on t functions, $\sigma_i(t)$, $i = 1, 2, 3$, $\ln(1 + \gamma_j(t, z))$, $\ln(1 + \delta_j(t, z))$, $j = 1, 2, 3$ are bounded, $\inf_{t \geq 0} (\mu(t) + \varepsilon(t) + \rho(t)) > 0$, $\Pi_i(\mathbb{R}) < \infty$, $i = 1, 2$.

Theorem 1. If Assumption 1 holds, then there exists a unique global solution $X(t) = (S(t), I(t), R(t))$ to the system (3) for any initial value $X(0) \in \mathbb{R}_+^3 = \{x \in \mathbb{R}^3 \mid x_i > 0, i = 1, 2, 3\}$, and $P\{X(t) \in \mathbb{R}_+^3, t \geq 0\} = 1$.

Proof. The coefficients of the system (3) are local Lipschitz continuous. Therefore, for any initial value $(S(0), I(0), R(0)) \in \mathbb{R}_+^3$ there exists a unique local solution $X(t) = (S(t), I(t), R(t))$ on $[0, \tau_e)$, where $\sup_{t < \tau_e} |X(t)| = +\infty$ (cf. Gikhman, & Skorokhod, 1982, p. 246). To show this solution is global, we need to show that $\tau_e = +\infty$ a.s. At first, we prove that $S(t)$ and $I(t)$ do not explode to infinity in a finite time. Let $n_0 \in \mathbb{N}$ be sufficiently large for $S(0) \in [\frac{1}{n_0}, n_0]$ and $I(0) \in [\frac{1}{n_0}, n_0]$. For any $n \geq n_0$ we define the stopping time $\tau_n = \inf\{t \in [0, \tau_e) : S(t) \notin (\frac{1}{n}, n) \text{ or } I(t) \notin (\frac{1}{n}, n)\}$.

Clearly, that τ_n is increasing as $n \rightarrow +\infty$. Set $\tau_\infty = \lim_{n \rightarrow \infty} \tau_n$, whence $\tau_\infty \leq \tau_e$ a.s. If we prove that $\tau_\infty = \infty$ a.s., then $\tau_e = \infty$ a.s. and $(S(t), I(t)) \in \mathbb{R}_+^2$, $t \geq 0$ a.s. If this statement is false, then there are constants $T > 0$ and $\varepsilon \in (0, 1)$, such that $P\{\tau_\infty < T\} > \varepsilon$. So, there is $n_1 \geq n_0$ such that

$$P\{\tau_n < T\} > \varepsilon, \forall n \geq n_1. \quad (5)$$

Let us consider the non-negative function $V(S, I) = (S - a - a \ln \frac{S}{a}) + (I - 1 - \ln I)$, $S > 0, I > 0$, where $a > 0$ is a constant, which will be defined later. If we denote

$$\kappa_i(t) = \frac{\sigma_i^2(t)}{2} + \int_{\mathbb{R}} [\gamma_i(t, z) - \ln(1 + \gamma_i(t, z))]\Pi_1(dz) - \int_{\mathbb{R}} \ln(1 + \delta_i(t, z))\Pi_2(dz), \quad i = 1, 2,$$

$$\Phi(t) = \Lambda(t) + a\mu(t) + (\mu(t) + \varepsilon(t) + \rho(t)) + a\kappa_1(t) + \kappa_2(t) - (\mu(t) + \beta(t))S(t) - a\frac{\Lambda(t)}{S(t)} + \\ + [a\beta(t) - (\mu(t) + \varepsilon(t) + \rho(t))]I(t),$$

then by the Itô formula and the definition of a stochastic integral with respect to the non-centered Poisson measure $\nu_2(du, dz)$ we derive

$$\begin{aligned} V(S(T \wedge \tau_n), I(T \wedge \tau_n)) = \\ = V(S(0), I(0)) + \int_0^{T \wedge \tau_n} \Phi(u)du + \int_0^{T \wedge \tau_n} (S(u) - a)\sigma_1(u)dw_1(u) + \int_0^{T \wedge \tau_n} (I(u) - 1)\sigma_2(u)dw_2(u) + \\ + \int_0^{T \wedge \tau_n} \int_{\mathbb{R}} [\gamma_1(u, z)S(u^-) - a \ln(1 + \gamma_1(u, z)) + \gamma_2(u, z)I(u^-) - \ln(1 + \gamma_2(u, z))] \tilde{\nu}_1(du, dz) + \\ + \int_0^{T \wedge \tau_n} \int_{\mathbb{R}} [\delta_1(u, z)S(u^-) - a \ln(1 + \delta_1(u, z)) + \delta_2(u, z)I(u^-) - \ln(1 + \delta_2(u, z))] \tilde{\nu}_2(du, dz). \end{aligned} \quad (6)$$

From the definition of stopping time τ_n it follows that for $0 \leq u \leq T \wedge \tau_n$ we have $\frac{1}{n} < S(u) < n, \frac{1}{n} < I(u) < n$. Hence under condition on the coefficients of the system (3), there is a constant $K > 0$ such that for $0 \leq u \leq T \wedge \tau_n$ we obtain

$$\begin{aligned} \Phi(u) \leq \sup_{t \geq 0} \{ \Lambda(t) + a\mu(t) + (\mu(t) + \varepsilon(t) + \rho(t)) + a\kappa_1(t) + \kappa_2(t) \} + \\ + \left[a \sup_{t \geq 0} \beta(t) - \inf_{t \geq 0} (\mu(t) + \varepsilon(t) + \rho(t)) \right] I(u) \leq K. \end{aligned}$$

If we choose $a = \frac{\inf_{t \geq 0} (\mu(t) + \varepsilon(t) + \rho(t))}{\sup_{t \geq 0} \beta(t)}$, then taking expectations, we derive from (6)

$$E[V(S(T \wedge \tau_n), I(T \wedge \tau_n))] \leq V(S(0), I(0)) + KT.$$

Let $\Omega_n = \{\omega \in \Omega \mid \tau_n(\omega) \leq T\}$, for $n \geq n_1$. Then by (5) we have $P(\Omega_n) > \varepsilon$, for $n \geq n_1$. Note that for every $\omega \in \Omega_n$ at least one of $S(\tau_n, \omega)$ and $I(\tau_n, \omega)$ equals either n or $1/n$. So

$$V(S(\tau_n, \omega), I(\tau_n, \omega)) \geq \min\left(n - 1 - \ln n, \frac{1}{n} - 1 + \ln n\right).$$

Therefore,

$$V(S(0), I(0)) + KT \geq E[\mathbb{I}_{\Omega_n} V(S(\tau_n), I(\tau_n))] \geq \varepsilon \min\left(n - 1 - \ln n, \frac{1}{n} - 1 + \ln n\right),$$

where $\mathbb{I}_{\Omega_n}$ is the indicator function of Ω_n .

Letting $n \rightarrow \infty$ leads to the contradiction $\infty > V(S(0), I(0)) + KT = \infty$. Therefore, it implies $S(t)$ and $I(t)$ will not explode in a finite time almost surely.

For the process $\zeta(t) = R_0(t)R(t)$ we have by the Itô formula $d\zeta(t) = d(R_0(t)R(t)) = R_0(t)\rho(t)I(t)dt$, where

$$\begin{aligned} R_0(t) \\ = \exp \left\{ \int_0^t \left[\mu(s) + \frac{\sigma_3^2(s)}{2} + \int_{\mathbb{R}} [\gamma_3(s, z) - \ln(1 + \gamma_3(s, z))] \Pi_1(dz) \right] ds - \int_0^t \sigma_3(s)dw_3(s) - \right. \\ \left. - \int_0^t \int_{\mathbb{R}} \ln(1 + \gamma_3(s, z)) \tilde{\nu}_1(ds, dz) - \int_0^t \int_{\mathbb{R}} \ln(1 + \delta_3(s, z)) \nu_2(ds, dz) \right\}, \end{aligned}$$

$R(t)$ satisfies the third equation in (3). Therefore

$$R(t) = \frac{R(0) + \int_0^t R_0(s)\rho(s)I(s)ds}{R_0(t)}.$$

Since $I(t)$ has been proven to be global and positive, $R(t)$ is also global and positive. This completes the proof of the theorem.

1.2. Asymptotic behavior around the disease-free equilibrium of the autonomous deterministic model.

As we mentioned in the Introduction if $R_0 \leq 1$, where $R_0 = \frac{\beta\Lambda}{\mu(\mu+\varepsilon+\rho)}$, then model (1) has only the disease-free equilibrium

$E_0 = \left(\frac{\Lambda}{\mu}, 0, 0\right)$, which is globally asymptotically stable. While for the autonomous stochastic system (4), E_0 is no longer the disease-free equilibrium. In this section, we will study the asymptotic behavior of the solution to the system (4) around E_0 .

Assumption 2. It is assumed, that $\Lambda, \beta, \mu, \varepsilon, \rho$ are positive constants, $\ln(1 + \gamma_i(z)), \ln(1 + \delta_i(z)), i = 1, 2, 3$ are bounded, $\Pi_i(\mathbb{R}) < \infty, i = 1, 2$.

Theorem 2. Let Assumption 2 be fulfilled and $(S(t), I(t), R(t))$ be the solution of system (4) with any initial value $(S(0), I(0), R(0)) \in \mathbb{R}_+^3$. If

$$R_0 = \frac{\beta\Lambda}{\mu(\mu + \varepsilon + \rho)} \leq 1,$$

$$2\mu > \max \left\{ \left[2\sigma_1^2 + 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} [3\delta_1^2(z) + 4|\delta_1(z)| + |\delta_2(z)|] \Pi_2(dz) \right], 2 \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) \right\},$$

$$2(\mu + \varepsilon + \rho) > \sigma_2^2 + 3 \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) + \int_{\mathbb{R}} [3\delta_2^2(z) + 4|\delta_2(z)| + 2|\delta_1(z)|] \Pi_2(dz),$$

then

$$\limsup_{t \rightarrow \infty} \frac{1}{t} E \left[\int_0^t \left(S(v) - \frac{\Lambda}{\mu} \right)^2 + I^2(v) + R(v) dv \right] \leq \\ \leq \frac{1}{K_1} \left(\frac{\Lambda}{\mu} \right)^2 \left(2\sigma_1^2 + 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} \left[3\delta_1^2(z) + 4|\delta_1(z)| + \frac{(\mu c_1)^2}{4\Lambda^2} |\delta_2(z)| \right] \Pi_2(dz) \right),$$

where

$$K_1 = \min \left\{ \left[2\mu - 2\sigma_1^2 - 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [3\delta_1^2(z) + 4|\delta_1(z)| + |\delta_2(z)|] \Pi_2(dz) \right], \right. \\ \left. \left[2(\mu + \varepsilon + \rho) - \sigma_2^2 - 3 \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [3\delta_2^2(z) + 4|\delta_2(z)| + 2|\delta_1(z)|] \Pi_2(dz) \right], c_2 \left(\mu - \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) \right) \right\}, \\ c_1 = \frac{2}{\beta} (2\mu + \varepsilon + \rho), \quad c_2 = c_1 \frac{\Lambda \beta}{\mu \rho} \left(\frac{1 - R_0}{R_0} \right).$$

Proof. If $(t) = S(t) - \frac{\Lambda}{\mu}$, then from the system (4), we derive

$$dU(t) = \left(-\mu U(t) - \beta U(t) I(t) - \beta \frac{\Lambda}{\mu} I(t) \right) dt + \sigma_1 \left(U(t) + \frac{\Lambda}{\mu} \right) dw_1(t) + \\ + \int_{\mathbb{R}} \gamma_1(z) \left(U(t^-) + \frac{\Lambda}{\mu} \right) \tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_1(z) \left(U(t^-) + \frac{\Lambda}{\mu} \right) v_2(dt, dz), \\ dI(t) = \left(\beta \left(U(t) + \frac{\Lambda}{\mu} \right) I(t) - (\mu + \varepsilon + \rho) I(t) \right) dt + \sigma_2 I(t) dw_2(t) + \\ + \int_{\mathbb{R}} \gamma_2(z) I(t^-) \tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_2(z) I(t^-) v_2(dt, dz),$$

$$dR(t) = (\rho I(t) - \mu R(t)) dt + \sigma_3 R(t) dw_3(t) + \int_{\mathbb{R}} \gamma_3(z) R(t^-) \tilde{v}_1(dt, dz) + \int_{\mathbb{R}} \delta_3(z) R(t^-) v_2(dt, dz).$$

Let us consider the non-negative function $V(U, I, R) = (U + I)^2 + c_1 I + c_2 R$, $U \in \mathbb{R}$, $I > 0$, $R > 0$, where constants $c_i > 0, i = 1, 2$ will be defined later.

Using the Itô formula, system (7), and the definition of a stochastic integral with respect to the non-centered Poisson measure $v_2(dt, dz)$ we derive

$$dV(U(t), I(t), R(t)) = \{ 2(U(t) + I(t)) [-\mu U(t) - (\mu + \varepsilon + \rho) I(t)] + \sigma_1^2 \left(U(t) + \frac{\Lambda}{\mu} \right)^2 + \sigma_2^2 I^2(t) + \\ + c_1 \left[\beta U(t) I(t) - \left(\mu + \varepsilon + \rho - \beta \frac{\Lambda}{\mu} \right) I(t) \right] + c_2 (\rho I(t) - \mu R(t) + \int_{\mathbb{R}} [\gamma_1(z) \left(U(t) + \frac{\Lambda}{\mu} \right) + \gamma_2(z) I(t)]^2 \Pi_1(dz) + \\ + \int_{\mathbb{R}} \left\{ \left[\delta_1(z) \left(U(t) + \frac{\Lambda}{\mu} \right) + \delta_2(z) I(t) \right]^2 + 2(U(t) + I(t)) \left[\delta_1(z) \left(U(t) + \frac{\Lambda}{\mu} \right) + \delta_2(z) I(t) \right] + \right. \right. \\ \left. \left. + c_1 \delta_2(z) I(t) + c_2 \delta_3(z) R(t) \right\} \Pi_2(dz) \right\} dt + dJ_{stoch}(t) = \\ = \left\{ -U^2(t) \left[2\mu - \sigma_1^2 - \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [\delta_1^2(z) + 2\delta_1(z)] \Pi_2(dz) \right] - \right. \\ \left. - I^2(t) \left[2(\mu + \varepsilon + \rho) - \sigma_2^2 - \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [\delta_2^2(z) + 2\delta_2(z)] \Pi_2(dz) \right] - c_2 \mu R(t) + \right. \\ \left. + U(t) I(t) \left[-2(2\mu + \varepsilon + \rho) + c_1 \beta + 2 \int_{\mathbb{R}} \gamma_1(z) \gamma_2(z) \Pi_1(dz) + 2 \int_{\mathbb{R}} [\delta_1(z) \delta_2(z) + \delta_1(z) + \delta_2(z)] \Pi_2(dz) \right] + \right. \\ \left. + I(t) \left[-c_1 \beta \frac{\Lambda}{\mu} \left(\frac{1 - R_0}{R_0} \right) + c_2 \rho + 2 \frac{\Lambda}{\mu} \left[\int_{\mathbb{R}} \gamma_1(z) \gamma_2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_1(z) \delta_2(z) + \delta_1(z)) \Pi_2(dz) \right] + \right. \right. \\ \left. \left. + c_1 \int_{\mathbb{R}} \delta_2(z) \Pi_2(dz) \right] + 2 \frac{\Lambda}{\mu} U(t) \left[\sigma_1^2 + \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_1^2(z) + \delta_1(z)) \Pi_2(dz) \right] + \right. \\ \left. + c_2 R(t) \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) + \left(\frac{\Lambda}{\mu} \right)^2 \left(\sigma_1^2 + \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} \delta_1^2(z) \Pi_2(dz) \right) \right\} dt + dJ_{stoch}(t) = \\ = LF(U(t), I(t), R(t)) dt + dJ_{stoch}(t),$$

where

$$dJ_{stoch}(t) = 2(U(t) + I(t)) \sigma_1 \left(U(t) + \frac{\Lambda}{\mu} \right) dw_1(t) + (2(U(t) + I(t)) + c_1) \sigma_2 I(t) dw_2(t) + c_2 \sigma_3 R(t) dw_3(t) + \\ + \int_{\mathbb{R}} \left\{ \left[\gamma_1(z) \left(U(t^-) + \frac{\Lambda}{\mu} \right) + \gamma_2(z) I(t^-) \right]^2 + 2(U(t^-) + I(t^-)) \left[\gamma_1(z) \left(U(t^-) + \frac{\Lambda}{\mu} \right) + \gamma_2(z) I(t^-) \right] + \right. \\ \left. + c_1 \gamma_2(z) I(t^-) + c_2 \gamma_3(z) R(t^-) \right\} \tilde{v}_1(dt, dz) + \\ + \int_{\mathbb{R}} \left\{ \left[\delta_1(z) \left(U(t^-) + \frac{\Lambda}{\mu} \right) + \delta_2(z) I(t^-) \right]^2 + 2(U(t^-) + I(t^-)) \left[\delta_1(z) \left(U(t^-) + \frac{\Lambda}{\mu} \right) + \delta_2(z) I(t^-) \right] + \right. \\ \left. + c_1 \delta_2(z) I(t^-) + c_2 \delta_3(z) R(t^-) \right\} v_2(dt, dz).$$

If we choose $c_1 = \frac{2}{\beta}(2\mu + \varepsilon + \rho) > 0$, $c_2 = c_1 \frac{\Lambda\beta}{\mu\rho} \left(\frac{1-R_0}{R_0}\right) > 0$, then using inequality $a^2 + b^2 \geq 2ab$ we obtain the estimate

$$\begin{aligned} LF(U, I, R) = & -U^2 \left[2\mu - \sigma_1^2 - \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [\delta_1^2(z) + 2\delta_1(z)] \Pi_2(dz) \right] - \\ & -I^2 \left[2(\mu + \varepsilon + \rho) - \sigma_2^2 - \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [\delta_2^2(z) + 2\delta_2(z)] \Pi_2(dz) \right] - c_2 R \left(\mu - \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) \right) + \\ & + 2UI \left[\int_{\mathbb{R}} \gamma_1(z) \gamma_2(z) \Pi_1(dz) + \int_{\mathbb{R}} [\delta_1(z) \delta_2(z) + \delta_1(z) + \delta_2(z)] \Pi_2(dz) \right] + \\ & + 2I \left[\frac{\Lambda}{\mu} \left(\int_{\mathbb{R}} \gamma_1(z) \gamma_2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_1(z) \delta_2(z) + \delta_1(z)) \Pi_2(dz) \right) + \frac{(2\mu + \varepsilon + \rho)}{\beta} \int_{\mathbb{R}} \delta_2(z) \Pi_2(dz) \right] + \\ & + 2 \frac{\Lambda}{\mu} U \left[\sigma_1^2 + \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_1^2(z) + \delta_1(z)) \Pi_2(dz) \right] + \left(\frac{\Lambda}{\mu} \right)^2 \left(\sigma_1^2 + \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} \delta_1^2(z) \Pi_2(dz) \right) \leq \\ & \leq -U^2 \left[2\mu - 2\sigma_1^2 - 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [3\delta_1^2(z) + 4|\delta_1(z)| + |\delta_2(z)|] \Pi_2(dz) \right] - \\ & -I^2 \left[2(\mu + \varepsilon + \rho) - \sigma_2^2 - 3 \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [3\delta_2^2(z) + 4|\delta_2(z)| + 2|\delta_1(z)|] \Pi_2(dz) \right] - \\ & -c_2 R \left(\mu - \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) \right) + \left(\frac{\Lambda}{\mu} \right)^2 \left(2\sigma_1^2 + 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} [3\delta_1^2(z) + 4|\delta_1(z)| + \frac{(\mu c_1)^2}{4\Lambda^2} |\delta_2(z)|] \Pi_2(dz) \right). \end{aligned} \quad (9)$$

Let $n_0 \in \mathbb{N}$ be sufficiently large for $U(0) + \frac{\Lambda}{\mu} \in \left[\frac{1}{n_0}, n_0\right]$, $I(0) \in \left[\frac{1}{n_0}, n_0\right]$ and $R(0) \in \left[\frac{1}{n_0}, n_0\right]$. For any $n \geq n_0$ we define the stopping time $\tau_n = \inf \left\{ t \geq 0 : X(t) \notin \left(\frac{1}{n}, n\right)^3 \right\}$, where $X(t) = (U(t) + \frac{\Lambda}{\mu}, I(t), R(t))$, $(U(t), I(t), R(t))$ is the solution to the system (7). According to Theorem 1, the system (4) has a unique global (no explosion in the finite time interval) solution $(S(t), I(t), R(t))$, therefore the solution to the system (7) is also global and $\lim_{n \rightarrow \infty} \tau_n = +\infty$ a.s. From (8), applying the properties of stochastic integrals with respect to Wiener processes and with respect to centered Poisson measures, we derive

$$0 \leq E[V(U(t \wedge \tau_n), I(t \wedge \tau_n), R(t \wedge \tau_n))] = V(U(0), I(0), R(0)) + E \left[\int_0^{t \wedge \tau_n} LF(U(v), I(v), R(v)) dv \right]. \quad (10)$$

If $n \rightarrow \infty$, then from (10) we have

$$0 \leq E[V(U(t), I(t), R(t))] = V(U(0), I(0), R(0)) + E \left[\int_0^t LF(U(v), I(v), R(v)) dv \right]. \quad (11)$$

Therefore, applying the estimate (9), we obtain from (11)

$$\begin{aligned} K_1 E \left[\int_0^t (U^2(v) + I^2(v) + R(v)) dv \right] & \leq V(U(0), I(0), R(0)) + \\ & + \left(\frac{\Lambda}{\mu} \right)^2 \left(2\sigma_1^2 + 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} \left[3\delta_1^2(z) + 4|\delta_1(z)| + \frac{(\mu c_1)^2}{4\Lambda^2} |\delta_2(z)| \right] \Pi_2(dz) \right) t, \end{aligned}$$

where

$$\begin{aligned} K_1 = \min \left\{ \left[2\mu - 2\sigma_1^2 - 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [3\delta_1^2(z) + 4|\delta_1(z)| + |\delta_2(z)|] \Pi_2(dz) \right], \right. \\ \left. \left[2(\mu + \varepsilon + \rho) - \sigma_2^2 - 3 \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \int_{\mathbb{R}} [3\delta_2^2(z) + 4|\delta_2(z)| + 2|\delta_1(z)|] \Pi_2(dz) \right], c_2 \left(\mu - \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) \right) \right\}. \end{aligned}$$

Therefore,

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{1}{t} E \left[\int_0^t \left(\left(S(v) - \frac{\Lambda}{\mu} \right)^2 + I^2(v) + R(v) \right) dv \right] & \leq \\ & \leq \frac{1}{K_1} \left(\frac{\Lambda}{\mu} \right)^2 \left(2\sigma_1^2 + 3 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} \left[3\delta_1^2(z) + 4|\delta_1(z)| + \frac{(\mu c_1)^2}{4\Lambda^2} |\delta_2(z)| \right] \Pi_2(dz) \right). \end{aligned}$$

The proof of the theorem is completed.

Remark 1. If $\gamma_i(z) \equiv 0$, $\delta_i(z) \equiv 0$, $i = 1, 2, 3$, then Theorem 2 is Theorem 3.1 from (Jiang et al., 2011).

1.3. Asymptotic behavior around the endemic equilibrium of the autonomous deterministic model

As we mentioned in the Introduction, if $R_0 = \frac{\beta\Lambda}{\mu(\mu+\varepsilon+\rho)} > 1$, then there exists a global asymptotically stable endemic equilibrium $E^* = \left(\frac{\mu+\varepsilon+\rho}{\beta}, \frac{\Lambda}{\mu+\varepsilon+\rho} - \frac{\mu}{\beta}, \frac{\Lambda\rho}{\mu(\mu+\varepsilon+\rho)} - \frac{\rho}{\beta} \right)$ in the deterministic model (1). While for the autonomous stochastic system (4), E^* is no longer the endemic equilibrium. In this section, we will study the asymptotic behavior of the solution to the system (4) around E^* .

Theorem 3. Let Assumption 2 be fulfilled and $(S(t), I(t), R(t))$ be the solution of system (4) with any initial value $(S(0), I(0), R(0)) \in \mathbb{R}_+^3$. If

$$\begin{aligned} R_0 &= \frac{\beta \Lambda}{\mu(\mu + \varepsilon + \rho)} > 1, \\ 2\mu &> \sigma_1^2 + 2 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (2\delta_1^2(z) + 3|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz), \\ 2(\mu + \varepsilon + \rho) &> \sigma_2^2 + 2 \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) + \int_{\mathbb{R}} [2\delta_2^2(z) + (3 + a)|\delta_2(z)| + |\delta_1(z)|] \Pi_2(dz), \\ \mu &> \sigma_3^2 + \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_3^2(z) + 2|\delta_3(z)|) \Pi_2(dz), \end{aligned} \quad (12)$$

then

$$\limsup_{t \rightarrow \infty} \frac{1}{t} E \left[\int_0^t ((S(v) - m_1 S^*)^2 + (I(v) - m_2 I^*)^2 + (R(v) - m_3 R^*)^2) dv \right] \leq \frac{M(\sigma, \gamma, \delta)}{K(\sigma, \gamma, \delta)},$$

where

$$\begin{aligned} K(\sigma, \gamma, \delta) &= \min \left\{ \frac{k_1}{2}, \frac{k_2}{2\mu}, \frac{bk_3}{2} \right\}, \quad a = \frac{(2\mu + \varepsilon + \rho)}{\beta}, \\ 0 < b &< \frac{\mu}{\rho^2} \left(2(\mu + \varepsilon + \rho) - \sigma_1^2 - 2 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} (2\delta_1^2(z) + 3|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz) \right), \\ k_1 &= 2\mu - \sigma_1^2 - 2 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} (2\delta_1^2(z) + 3|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz), \\ m_1 &= \frac{1}{k_1} \left(2\mu - \int_{\mathbb{R}} (|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz) \right), \\ k_2 &= 2\mu(\mu + \varepsilon + \rho) - \mu\sigma_2^2 - 2\mu \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \mu \int_{\mathbb{R}} (2\delta_2^2(z) + (3 + a)|\delta_2(z)| + |\delta_1(z)|) \Pi_2(dz) - b\rho^2, \\ m_2 &= \frac{2\mu}{k_2} \left(\mu + \varepsilon + \rho - \frac{1}{2} \int_{\mathbb{R}} (|\delta_1(z)| + (1 + a)|\delta_2(z)|) \Pi_2(dz) - \frac{b\rho^2}{2\mu} \right), \\ k_3 &= \mu - \sigma_3^2 - \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) - \int_{\mathbb{R}} (\delta_3^2(z) + 2|\delta_3(z)|) \Pi_2(dz), \quad m_3 = \frac{1}{k_3} \left(\mu - \int_{\mathbb{R}} |\delta_3(z)| \Pi_2(dz) \right), \\ M(\sigma, \gamma, \delta) &= \frac{k_1}{2} m_1 (m_1 - 1) S^{*2} + \frac{k_2}{2\mu} m_2 (m_2 - 1) I^{*2} + \frac{bk_3}{2} m_1 (m_1 - 1) R^{*2} + \\ &+ a \left(\frac{1}{2} \sigma_2^2 + \int_{\mathbb{R}} (\gamma_2(z) - \ln(1 + \gamma_2(z))) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_2(z) - \ln(1 + \delta_2(z))) \Pi_2(dz) \right) I^* + \frac{a}{2} \int_{\mathbb{R}} |\delta_2(z)| \Pi_2(dz). \end{aligned}$$

Proof. Let us consider the function $F(S, I, R) = F_1(S, I) + F_2(R) \geq 0$, $(S, I, R) \in \mathbb{R}_+^3$, where

$$F_1(S, I) = \frac{1}{2} (S - S^* + I - I^*)^2 + a \left(I - I^* - I^* \ln \frac{I}{I^*} \right), \quad F_2(R) = \frac{b}{2} (R - R^*)^2.$$

Here constants $a > 0$, $b > 0$ to be defined later.

By the Itô formula, system (4), and the definition of a stochastic integral with respect to the non-centered Poisson measure $\nu_2(dt, dz)$ we derive

$$\begin{aligned} dF_1(S(t), I(t)) &= LF_1(S(t), I(t))dt + (S(t) - S^* + I(t) - I^*)\sigma_1 S(t)dw_1(t) + \\ &+ (S(t) - S^* + I(t) - I^*)\sigma_2 I(t)dw_2(t) + a(I(t) - I^*)\sigma_2 dw_2(t) + \\ &+ \int_{\mathbb{R}} \left[(S(t^-) - S^* + I(t^-) - I^*)(\gamma_1(z)S(t^-) + \gamma_2(z)I(t^-)) + \frac{1}{2}(\gamma_1(z)S(t^-) + \gamma_2(z)I(t^-))^2 + \right. \\ &+ a(\gamma_2(z)I(t^-) - I^* \ln(1 + \gamma_2(z))) \tilde{\nu}_1(dt, dz) + \int_{\mathbb{R}} [(S(t^-) - S^* + I(t^-) - I^*)(\delta_1(z)S(t^-) + \delta_2(z)I(t^-)) + \\ &+ \frac{1}{2}(\delta_1(z)S(t^-) + \delta_2(z)I(t^-))^2 + a(\delta_2(z)I(t^-) - I^* \ln(1 + \delta_2(z)))] \tilde{\nu}_2(dt, dz). \\ dF_2(R(t)) &= LF_2(R(t), I(t))dt + b(R(t) - R^*)\sigma_3 R(t)dw_3(t) + \\ &+ \int_{\mathbb{R}} \left[b(R(t^-) - R^*)\gamma_3(z)R(t^-) + \frac{1}{2}b\gamma_3^2(z)R^2(t^-) \right] \tilde{\nu}_1(dt, dz) + \\ &+ \int_{\mathbb{R}} \left[b(R(t^-) - R^*)\delta_3(z)R(t^-) + \frac{1}{2}b\delta_3^2(z)R^2(t^-) \right] \tilde{\nu}_2(dt, dz), \end{aligned} \quad (13)$$

where

$$\begin{aligned} LF_1(S, I) &= (S - S^* + I - I^*)(\Lambda - \mu S - (\mu + \varepsilon + \rho)I) + a(I - I^*)(\beta S - (\mu + \varepsilon + \rho)) + \\ &+ \frac{1}{2}(\sigma_1^2 S^2 + \sigma_2^2 I^2) + \frac{a}{2}\sigma_2^2 I^* + \int_{\mathbb{R}} \left[\frac{1}{2}(\gamma_1(z)S + \gamma_2(z)I)^2 + a(\gamma_2(z) - \ln(1 + \gamma_2(z)))I^* \right] \Pi_1(dz) + \\ &+ \int_{\mathbb{R}} \left[(S - S^* + I - I^*)(\delta_1(z)S + \delta_2(z)I) + \frac{1}{2}(\delta_1(z)S + \delta_2(z)I)^2 + a(\delta_2(z)I - I^* \ln(1 + \delta_2(z))) \right] \Pi_2(dz) \\ &\leq -\mu(S - S^*)^2 - (\mu + \varepsilon + \rho)(I - I^*)^2 + \end{aligned} \quad (14)$$

$$\begin{aligned}
 & + (a\beta - (2\mu + \varepsilon + \rho))(S - S^*)(I - I^*) + \frac{1}{2}(\sigma_1^2 S^2 + \sigma_2^2 I^2) + \frac{a}{2}\sigma_2^2 I^* + \\
 & + S^2 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + I^2 \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) + aI^* \int_{\mathbb{R}} (\gamma_2(z) - \ln(1 + \gamma_2(z))) \Pi_1(dz) + \\
 & + \frac{1}{2}(S - S^*)^2 \int_{\mathbb{R}} (|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz) + \frac{1}{2}(I - I^*)^2 \int_{\mathbb{R}} (|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz) + \\
 & + S^2 \int_{\mathbb{R}} (\delta_1^2(z) + |\delta_1(z)|) \Pi_2(dz) + I^2 \int_{\mathbb{R}} (\delta_2^2(z) + |\delta_2(z)|) \Pi_2(dz) + \frac{a}{2} \int_{\mathbb{R}} |\delta_2(z)| \Pi_2(dz) + \\
 & + \frac{a}{2}(I - I^*)^2 \int_{\mathbb{R}} |\delta_2(z)| \Pi_2(dz) + aI^* \int_{\mathbb{R}} (\delta_2(z) - \ln(1 + \delta_2(z))) \Pi_2(dz).
 \end{aligned}$$

Here we use equalities $\Lambda = I^*(\mu + \varepsilon + \rho) + \mu S^*$, $\beta S^* = \mu + \varepsilon + \rho$, and basic estimates $2cd \leq c^2 + d^2$, $(c + d)^2 \leq 2(c^2 + d^2)$. Applying the equality $\rho I^* = \mu R^*$ and estimate $2cd \leq c^2 + d^2$ we obtain

$$\begin{aligned}
 LF_2(R, I) & = b(R - R^*)(\rho I - \mu R) + \frac{1}{2}b\sigma_3^2 R^2 + \frac{b}{2}R^2 \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) + b(R - R^*)R \int_{\mathbb{R}} \delta_3(z) \Pi_2(dz) + \\
 & + \frac{b}{2}R^2 \int_{\mathbb{R}} \delta_3^2(z) \Pi_2(dz) \leq b(R - R^*)\rho(I - I^*) - b\mu(R - R^*)^2 + \frac{b}{2}\left(\sigma_3^2 + \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) + \right. \\
 & \left. + \int_{\mathbb{R}} (\delta_3^2(z) + |\delta_3(z)|) \Pi_2(dz)\right) R^2 + \frac{b}{2}(R - R^*)^2 \int_{\mathbb{R}} |\delta_3(z)| \Pi_2(dz) \leq \frac{b}{2}\left[\frac{\rho^2}{\mu}(I - I^*)^2 - \right. \\
 & \left. - \left(\mu - \int_{\mathbb{R}} |\delta_3(z)| \Pi_2(dz)\right)(R - R^*)^2 + \left(\sigma_3^2 + \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_3^2(z) + |\delta_3(z)|) \Pi_2(dz)\right) R^2\right].
 \end{aligned} \tag{15}$$

Choose $a = \frac{(2\mu + \varepsilon + \rho)}{\beta}$ and taking together (14) and (15), we get

$$\begin{aligned}
 LF(S, I, R) & = LF_1(S, I) + LF_2(R, I) \leq -\left(\mu - \frac{1}{2} \int_{\mathbb{R}} (|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz)\right)(S - S^*)^2 - \\
 & - \left(\mu + \varepsilon + \rho - \frac{1}{2} \int_{\mathbb{R}} (|\delta_1(z)| + (1 + a)|\delta_2(z)|) \Pi_2(dz) - \frac{b\rho^2}{2\mu}\right)(I - I^*)^2 - \frac{b}{2}\left(\mu - \int_{\mathbb{R}} |\delta_3(z)| \Pi_2(dz)\right)(R - R^*)^2 \\
 & + \left(\frac{1}{2}\sigma_1^2 + \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_1^2(z) + |\delta_1(z)|) \Pi_2(dz)\right)S^2 + \left(\frac{1}{2}\sigma_2^2 + \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) + \right. \\
 & \left. + \int_{\mathbb{R}} (\delta_2^2(z) + |\delta_2(z)|) \Pi_2(dz)\right)I^2 + \frac{b}{2}\left(\sigma_3^2 + \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_3^2(z) + |\delta_3(z)|) \Pi_2(dz)\right)R^2 + \\
 & + a\left(\frac{1}{2}\sigma_2^2 + \int_{\mathbb{R}} (\gamma_2(z) - \ln(1 + \gamma_2(z))) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_2(z) - \ln(1 + \delta_2(z))) \Pi_2(dz)\right)I^* + \frac{a}{2} \int_{\mathbb{R}} |\delta_2(z)| \Pi_2(dz) = \\
 & = -\frac{k_1}{2}(S - m_1 S^*)^2 - \frac{k_2}{2\mu}(I - m_2 I^*)^2 - \frac{bk_3}{2}(R - m_2 R^*)^2 + M(\sigma, \gamma, \delta),
 \end{aligned} \tag{16}$$

where

$$\begin{aligned}
 k_1 & = 2\mu - \sigma_1^2 - 2 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} (2\delta_1^2(z) + 3|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz), \\
 m_1 & = \frac{1}{k_1}\left(2\mu - \int_{\mathbb{R}} (|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz)\right), \\
 k_2 & = 2\mu(\mu + \varepsilon + \rho) - \mu\sigma_2^2 - 2\mu \int_{\mathbb{R}} \gamma_2^2(z) \Pi_1(dz) - \mu \int_{\mathbb{R}} (2\delta_2^2(z) + (3 + a)|\delta_2(z)| + |\delta_1(z)|) \Pi_2(dz) - b\rho^2, \\
 m_2 & = \frac{2\mu}{k_2}\left(\mu + \varepsilon + \rho - \frac{1}{2} \int_{\mathbb{R}} (|\delta_1(z)| + (1 + a)|\delta_2(z)|) \Pi_2(dz) - \frac{b\rho^2}{2\mu}\right), \\
 k_3 & = \mu - \sigma_3^2 - \int_{\mathbb{R}} \gamma_3^2(z) \Pi_1(dz) - \int_{\mathbb{R}} (\delta_3^2(z) + 2|\delta_3(z)|) \Pi_2(dz), \quad m_3 = \frac{1}{k_3}\left(\mu - \int_{\mathbb{R}} |\delta_3(z)| \Pi_2(dz)\right), \\
 M(\sigma, \gamma, \delta) & = \frac{k_1}{2}m_1(m_1 - 1)S^{*2} + \frac{k_2}{2\mu}m_2(m_2 - 1)I^{*2} + \frac{bk_3}{2}m_1(m_1 - 1)R^{*2} + \\
 & + a\left(\frac{1}{2}\sigma_2^2 + \int_{\mathbb{R}} (\gamma_2(z) - \ln(1 + \gamma_2(z))) \Pi_1(dz) + \int_{\mathbb{R}} (\delta_2(z) - \ln(1 + \delta_2(z))) \Pi_2(dz)\right)I^* + \frac{a}{2} \int_{\mathbb{R}} |\delta_2(z)| \Pi_2(dz).
 \end{aligned}$$

Due to conditions (12) we can choose

$$0 < b < \frac{\mu}{\rho^2}\left(2(\mu + \varepsilon + \rho) - \sigma_1^2 - 2 \int_{\mathbb{R}} \gamma_1^2(z) \Pi_1(dz) - \int_{\mathbb{R}} (2\delta_1^2(z) + 3|\delta_1(z)| + |\delta_2(z)|) \Pi_2(dz)\right),$$

and we obtain that $k_i > 0, m_i > 1, i = 1, 2, 3$.

From (13) we derive

$$dF(S(t), I(t), R(t)) = dF_1(S(t), I(t)) + dF_2(R(t)) = LF(S(t), I(t), R(t))dt + dJ_{stoch}(t), \tag{17}$$

where

$$\begin{aligned} dJ_{stoch}(t) = & (S(t) - S^* + I(t) - I^*)\sigma_1 S(t)dw_1(t) + (S(t) - S^* + I(t) - I^*)\sigma_2 I(t)dw_2(t) + a(I(t) - I^*)\sigma_2 dw_2(t) + \\ & + \int_{\mathbb{R}} \left[(S(t^-) - S^* + I(t^-) - I^*)(\gamma_1(z)S(t^-) + \gamma_2(z)I(t^-)) + \frac{1}{2}(\gamma_1(z)S(t^-) + \gamma_2(z)I(t^-))^2 + \right. \\ & + a(\gamma_2(z)I(t^-) - I^* \ln(1 + \gamma_2(z)))\tilde{v}_1(dt, dz) + \int_{\mathbb{R}} [(S(t^-) - S^* + I(t^-) - I^*)(\delta_1(z)S(t^-) + \delta_2(z)I(t^-)) + \\ & + \frac{1}{2}(\delta_1(z)S(t^-) + \delta_2(z)I(t^-))^2 + a(\delta_2(z)I(t^-) - I^* \ln(1 + \delta_2(z)))\tilde{v}_2(dt, dz) + b(R(t) - R^*)\sigma_3 R(t)dw_3(t) + \\ & + \int_{\mathbb{R}} \left[b(R(t^-) - R^*)\gamma_3(z)R(t^-) + \frac{1}{2}b\gamma_3^2(z)R^2(t^-) \right] \tilde{v}_1(dt, dz) + \\ & + \int_{\mathbb{R}} \left[b(R(t^-) - R^*)\delta_3(z)R(t^-) + \frac{1}{2}b\delta_3^2(z)R^2(t^-) \right] \tilde{v}_2(dt, dz). \end{aligned}$$

Let $n_0 \in \mathbb{N}$ be sufficiently large for $S(0) \in [\frac{1}{n_0}, n_0]$, $I(0) \in [\frac{1}{n_0}, n_0]$ and $R(0) \in [\frac{1}{n_0}, n_0]$. For any $n \geq n_0$ we define the stopping time $\tau_n = \inf\{t \geq 0: X(t) \notin (\frac{1}{n}, n)^3\}$, where $X(t) = (S(t), I(t), R(t))$, $(S(t), I(t), R(t))$ is the solution to the system (4). According to Theorem 1, the system (4) has a unique global (no explosion in the finite time interval) solution $(S(t), I(t), R(t))$, and $\lim_{n \rightarrow \infty} \tau_n = +\infty$ a.s. Integrating both sides of (17) from 0 to $t \wedge \tau_n$, then taking expectations, and applying inequality (16) we obtain

$$\begin{aligned} 0 \leq F(S(t \wedge \tau_n), I(t \wedge \tau_n), R(t \wedge \tau_n)) & \leq F(S(0), I(0), R(0)) - \\ - E \left[\int_0^{t \wedge \tau_n} \left(\frac{k_1}{2} (S(v) - m_1 S^*)^2 + \frac{k_2}{2\mu} (I(v) - m_2 I^*)^2 + \frac{bk_3}{2} (R - m_2 R^*)^2 \right) dv \right] + M(\sigma, \gamma, \delta)t & \leq \\ \leq F(S(0), I(0), R(0)) - K(\sigma, \gamma, \delta) E \left[\int_0^{t \wedge \tau_n} ((S(v) - m_1 S^*)^2 + (I(v) - m_2 I^*)^2 + (R - m_2 R^*)^2) dv \right] + M(\sigma, \gamma, \delta)t & \end{aligned} \quad (18)$$

where $K(\sigma, \gamma, \delta) = \min\{\frac{k_1}{2}, \frac{k_2}{2\mu}, \frac{bk_3}{2}\}$.

If $n \rightarrow \infty$, then from (18) we have

$$\frac{1}{t} E \left[\int_0^t ((S(v) - m_1 S^*)^2 + (I(v) - m_2 I^*)^2 + (R - m_2 R^*)^2) dv \right] \leq \frac{1}{K(\sigma, \gamma, \delta)t} [F(S(0), I(0), R(0)) + M(\sigma, \gamma, \delta)t]. \quad (19)$$

If $t \rightarrow \infty$, then we have from (19)

$$\limsup_{t \rightarrow \infty} \frac{1}{t} E \left[\int_0^t ((S(v) - m_1 S^*)^2 + (I(v) - m_2 I^*)^2 + (R(v) - m_3 R^*)^2) dv \right] \leq \frac{M(\sigma, \gamma, \delta)}{K(\sigma, \gamma, \delta)}.$$

The proof of the theorem is completed.

Remark 2. If $\gamma_i(z) \equiv 0, \delta_i(z) \equiv 0, i = 1, 2, 3$, then Theorem 3 is Theorem 4.1 from (Jiang et al., 2011).

Discussion and conclusions

The existence and uniqueness of the global positive solution are proved for the system of stochastic differential equations describing a non-autonomous SIR model disturbed by white noise, centered and non-centered Poisson noises. In the deterministic case there is a threshold of the system for an epidemic to occur, so called the basic reproduction number. Depending on the value of the reproduction number there is the disease-free equilibrium, or there is an endemic equilibrium, which implies the disease always remains. In the case of the autonomous stochastic SIR model, we study the asymptotic behavior of the solution to the corresponding system of stochastic differential equations around these points of equilibrium.

Authors' contribution: Oleksandr Borysenko – conceptualization, writing and editing; Olga Borysenko – methodology, writing-viewing and editing.

References

- Anderson, R. M., & May, R. M. (1979). Population biology of infectious diseases. *Nature*, 280 (1), 361–367.
- Jiang, D., Yua, J., Ji, Ch., & Shi, N. (2011). Asymptotic behavior of global positive solution to a stochastic SIR model. *Mathematical and Computer Modelling*, 54, 221–232. <https://doi.org/10.1016/j.mcm.2011.02.004>
- Zhang, X., & Wang, Ke. (2013). Stochastic SIR model with jumps. *Applied Mathematics Letters*, 26, 867–874. <https://doi.org/10.1016/j.aml.2013.03.013>
- Borysenko, O. D., & Borysenko, O. V. (2022). A stochastic predator-prey model that depends on the population density of the predator. *Bulletin of Taras Shevchenko National University of Kyiv, Series: Physics & Mathematics*, 4, 11–17. <https://doi.org/10.17721/1812-5409.2022/4.1>
- Gikhman, I. I., & Skorokhod, A. V. (1982). *Stochastic Differential Equations and its Applications*. Naukova Dumka [in Russian]. [Гихман, И. И., & Скороход А. В. (1982). *Стохастические дифференциальные уравнения и их приложения*. Наукова думка].

Отримано редакцією журналу / Received: 17.12.24

Прорецензовано / Revised: 10.03.25

Схвалено до друку / Accepted: 14.03.25

Олександр БОРИСЕНКО, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0002-6670-8605
e-mail: borysenko_oleksandr@knu.ua
Київський національний університет імені Тараса Шевченка, Київ, Україна

Ольга БОРИСЕНКО, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0002-2099-2911
e-mail: oborisenko1373@gmail.com
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

СТОХАСТИЧНА SIR-МОДЕЛЬ ІЗ ПУАССОНІВСЬКИМИ ШУМАМИ

Ураховуючи негативний вплив інфекційних захворювань на зростання популяції, важливо зрозуміти динаміку поведінки таких захворювань. У вивченні розповсюдження інфекційних захворювань широко використовують детерміновані математичні SIR-моделі. У реальному світі є багато випадкових чинників, що впливають на динаміку поширення інфекційних захворювань, тому використання стохастичних моделей є більш виправданим. Якщо під час побудови математичної моделі ми хочемо брати до уваги такі збурення, яким піддається довкілля, як пожежі, землетруси тощо, то ми маємо вводити в популяційну модель пуассонівські шуми, щоб описувати такі системи зі стрибками. Ми довели існування і єдиність глобального додатного розв'язку системи стохастичних диференціальних рівнянь, що описує неавтономну SIR-модель, збурену білим шумом, центрованим і нецентрованим пуассонівськими шумами. У детермінованому автономному випадку існує поріг системи, так зване основне репродуктивне число, за якого відбувається епідемія. Залежно від величини репродуктивного числа існує рівновага, вільна від захворювання, або ендемічна рівновага, яка означає, що захворювання залишається назавжди. Щодо автономної стохастичної SIR-моделі зі стрибками, ми вивчили асимптотичну поведінку розв'язку відповідної системи стохастичних диференціальних рівнянь навколо цих точок рівноваги.

Ключові слова: стохастична SIR-модель, шум довкілля, центрований і нецентрований пуассонівські шуми, глобальний додатний розв'язок, рівновага вільна від захворювання, ендемічна рівновага, асимптотична поведінка.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.2

DOI: DOI: <https://doi.org/10.17721/1812-5409.2025/1.3>

Olga VASYLYK, DSc (Phys. & Math.), Assoc. Prof.

ORCID ID: 0000-0002-0880-3751

e-mail: vasylyk@matan.kpi.ua

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Kostiantyn SHURUBURA, PhD Student

ORCID ID: 0009-0003-5535-2644

e-mail: shurubura@hotmail.com

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

OPTIMIZATION OF THE MODEL OF φ -SUB-GAUSSIAN GENERALIZED FRACTIONAL BROWNIAN MOTION

Simulation methods for stochastic processes are used in many scientific areas, and special attention is given to the fractional Brownian motion. There are many results on simulation of Gaussian fractional Brownian motion. In this paper, we consider simulation of processes from more general class, namely, simulation of a strictly φ -sub-Gaussian generalized fractional Brownian motion defined on the finite domain $[0, 1]$. In the beginning of the article, we presented preliminary information which is necessary for understanding the results of our study. In the main part of our paper, for the strictly φ -sub-Gaussian generalized fractional Brownian motion with some given function φ optimal values of the parameters necessary for simulation with given reliability and accuracy in the space $C([0; 1])$ are found. This gives us a significant reduction in the number of terms needed to ensure the reliability and accuracy of our model. Also, in the paper some results of calculations and simulation are given. As expected, in our case the obtained values of terms in the model are significantly smaller than the corresponding numbers for the models with the values of parameters that are not optimal.

Key words: φ -sub-Gaussian generalized fractional Brownian motion, optimization of a model, reliability and accuracy of a model, simulation, stochastic process.

AMS 2020 classification: 60G15, 60G22, 60G51, 68U20

Introduction

Simulation methods for stochastic processes are used in many scientific areas, and special attention is given to the fractional Brownian motion, see, for example, papers (Banna et al., 2019; Pashko et al., 2019), as well as the works (Cai, Huan, & Zhu, 2022; Kozachenko & Rozora, 2019; Linda & Allen, 2010; Mode, 2011; Molz & Liu, 1997; Pashko & Rozora, 2018; Pashko, Rozora, & Synyavska, 2021).

In the article (Kozachenko, Sottinen, & Vasylyk, 2005), there was proposed an algorithm for simulation of a strictly φ -sub-Gaussian generalized fractional Brownian motion (φ -GFBM) with given reliability and accuracy in the space $C([0; 1])$. The simulation was based on a series expansion of the fractional Brownian motion. The authors proved an estimate of the accuracy of the simulation in the space $C([0; 1])$ of continuous functions equipped with the usual sup-norm.

In the article (Kozachenko & Vasylyk, 2006), this study was continued and there was considered the case of simulation of sub-Gaussian generalized fractional Brownian motion, that is, the case when $\varphi(x) = \frac{x^2}{2}, x \in \mathbb{R}$.

In the paper (Vasylyk & Lovytska, 2021), there were derived conditions, under which the model based on a series representation approximates a strictly φ -sub-Gaussian generalized fractional Brownian motion with given reliability and accuracy in the space $C([0; 1])$ in the case, when $\varphi(x) = \frac{|x|^p}{p}, |x| \geq 1, p > 1$. Also, necessary simulation parameters were calculated and models of sample paths of corresponding processes were constructed for various values of the Hurst parameter H and for given reliability and accuracy. Then in (Vasylyk et al., 2021), the authors derived conditions, under which the model based on a series representation approximates a strictly φ -sub-Gaussian generalized fractional Brownian motion with given reliability and accuracy in the space $C([0; 1])$ in the case, when $\varphi(x) = \exp |x| - |x| - 1, x \in \mathbb{R}$.

But in all the above-mentioned articles the parameters necessary for construction of the model were chosen from acceptable intervals for reasons of computational convenience, and the problem of optimization of the model was not considered.

In this paper, we consider simulation of the strictly φ -sub-Gaussian generalized fractional Brownian motion on the finite domain $[0, 1]$ with covariance function

$$R_H(t, s) = EZ(s)Z(t) = \frac{1}{2}(t^{2H} + s^{2H} - |s - t|^{2H})$$

and the Hurst index H , which takes values in the interval $(0, 1)$. For the strictly φ -sub-Gaussian generalized fractional Brownian motion with the function $\varphi(x) = \frac{x^2}{\omega} \mathbb{I}_{\{|x| < 1\}} + \frac{|x|^\omega}{\omega} \mathbb{I}_{\{|x| \geq 1\}}, \omega \geq 2$, optimal values of the parameters necessary for simulation in the space $C([0; 1])$ are found. This gives us a significant gain in the number of terms needed to ensure the reliability and accuracy of our model.

The plan of the paper is as follows.

In Sections 1 and 2 we present preliminary information which is necessary for understanding the results of our study. Section 3 contains our main results. And Section 4 presents some results of calculations and simulation, as well as comparison of our results with the results obtained in the paper (Vasylyk & Lovytska, 2021).

1. φ -sub-Gaussian random variables and processes

Let us recall some facts about φ -sub-Gaussian random variables and processes.

Definition 1. (Buldygin & Kozachenko, 2000) A continuous even convex function φ is called an *Orlicz N -function*, if $\varphi(0) = 0$, $\varphi(x) > 0$ as $x \neq 0$ and the following conditions hold: $\lim_{x \rightarrow 0} \frac{\varphi(x)}{x} = 0$, $\lim_{x \rightarrow \infty} \frac{\varphi(x)}{x} = \infty$.

The function φ^* defined by $\varphi^*(x) = \sup_{y \in \mathbb{R}} (xy - \varphi(y))$ is called *the Young-Fenchel transform (or convex conjugate)* of the function φ .

We say that φ satisfies the *Condition Q* , if φ is such an Orlicz N -function that $\liminf_{x \rightarrow 0} \frac{\varphi(x)}{x^2} = c > 0$, where the case $c = \infty$ is possible (Giuliano, Kozachenko, & Nikitina, 2003).

In what follows we will always deal with Orlicz N -functions for which condition Q holds.

Examples of Orlicz N -functions, for which the condition Q is satisfied:

$$\begin{aligned} \varphi(x) &= \frac{|x|^\alpha}{\alpha}, \quad 1 < \alpha \leq 2, \\ \varphi(x) &= \begin{cases} \frac{|x|^\alpha}{\alpha}, & |x| \geq 1; \\ \frac{|x|^2}{\alpha}, & |x| < 1, \end{cases} \quad \alpha > 2. \end{aligned} \quad (1)$$

Definition 2. (Buldygin & Kozachenko, 2000) Let $\{\Omega, \mathcal{F}, \mathbf{P}\}$ be a standard probability space. The random variable ζ is said to be φ -sub-Gaussian (or belongs to the space $\text{Sub}_\varphi(\Omega)$), if $E\zeta = 0$, $E \exp\{\lambda\zeta\}$ exists for all $\lambda \in \mathbb{R}$ and there exists a constant $a > 0$ such that for all $\lambda \in \mathbb{R}$ the following inequality holds: $E \exp\{\lambda\zeta\} \leq \exp\{\varphi(a\lambda)\}$.

The space $\text{Sub}_\varphi(\Omega)$ is a Banach space with respect to the norm (Buldygin & Kozachenko, 2000)

$$\tau_\varphi(\zeta) = \inf\{a > 0 : E \exp\{\lambda\zeta\} \leq \exp\{\varphi(a\lambda)\},$$

which is called the φ -sub-Gaussian standard of the random variable ζ .

Centered Gaussian random variables $\zeta \sim N(0, \sigma^2)$ are φ -sub-Gaussian with $\varphi(x) = \frac{x^2}{2}$ and $\tau_\varphi^2(\zeta) = E\zeta^2 = \sigma^2$. In the case $\varphi(x) = \frac{x^2}{2}$, φ -sub-Gaussian random variables are simply sub-Gaussian.

The important property of φ -sub-Gaussian random variables is the exponential estimate for their tail probabilities. If ζ is a φ -sub-Gaussian random variable, then for all $u > 0$

$$P\{|\zeta| > u\} \leq 2 \exp \left\{ -\varphi^* \left(\frac{u}{\tau_\varphi(\zeta)} \right) \right\}. \quad (2)$$

Definition 3. A family Δ of random variables $\zeta \in \text{Sub}_\varphi(\Omega)$ is called strictly φ -sub-Gaussian (see (Buldygin & Kozachenko, 2000)), if there exists a constant C_Δ such that for all countable sets I of random variables $\zeta_i \in \Delta$, $i \in I$, the following inequality holds:

$$\tau_\varphi \left(\sum_{i \in I} \lambda_i \zeta_i \right) \leq C_\Delta \left(E \left(\sum_{i \in I} \lambda_i \zeta_i \right)^2 \right)^{1/2}. \quad (3)$$

The constant C_Δ is called the *determining* constant of the family Δ .

The linear closure of a strictly φ -sub-Gaussian family Δ in the space $L_2(\Omega)$ is the strictly φ -sub-Gaussian with the same determining constant.

Definition 4. The random process $X = \{X(t), t \in T\}$ is called (strictly) φ -sub-Gaussian if the family of random variables $\{X(t), t \in T\}$ is (strictly) φ -sub-Gaussian.

Example. (Kozachenko & Koval'chuk, 1998) Let a family of random variables $\{\xi_k, k = \overline{1, \infty}\}$ be a strictly φ -sub-Gaussian with determining constant C_ξ , and let $X(t) = \sum_{k=1}^{\infty} \xi_k f_k(t)$, where $\{f_k(t), t \in T, k \geq 1\}$ is a sequence of nonrandom functions and the series converges in mean square. Then the random process $X = \{X(t), t \in T\}$ is strictly φ -sub-Gaussian with determining constant C_ξ .

2. Generalized fractional Brownian motion

Let $\{B(t), t \in [0, 1]\}$ be a fractional Brownian motion with Hurst parameter $H \in (0, 1)$. This means that $B(t)$ is a centered Gaussian process with covariance function of the form

$$EB(s)B(t) = \frac{1}{2}(t^{2H} + s^{2H} - |s - t|^{2H}).$$

In the paper (Dzhaparidze & Zanten, 2002), it is shown that $B(t)$ can be represented in the following way:

$$B(t) = \sum_{n=1}^{\infty} \frac{\sin x_n t}{x_n} X_n + \sum_{n=1}^{\infty} \frac{1 - \cos y_n t}{y_n} Y_n, \quad (4)$$

where the series in (4) are mean square convergent, $\{X_n, n = 1, 2, \dots\}$ and $\{Y_n, n = 1, 2, \dots\}$ are independent Gaussian variables such that $EX_n = EY_n = 0$,

$$\text{Var}X_n = 2C_H^2 x_n^{-2H} J_{1-H}^{-2}(x_n), \quad \text{Var}Y_n = 2C_H^2 y_n^{-2H} J_{-H}^{-2}(y_n), \quad C_H^2 = \pi^{-1} \Gamma(1+2H) \sin(\pi H),$$

$x_1 < x_2 < \dots$, are positive real zeroes of the Bessel function J_{-H} , and $y_1 < y_2 < \dots$, are positive real zeroes of the Bessel function J_{1-H} .

The representation (4) can be rewritten in the following form:

$$B(t) = \sum_{n=1}^{\infty} \tilde{X}_n c_n \sin x_n t + \sum_{n=1}^{\infty} \tilde{Y}_n d_n (1 - \cos y_n t),$$

where $\{\tilde{X}_n, n = 1, 2, \dots\}$ and $\{\tilde{Y}_n, n = 1, 2, \dots\}$ are independent Gaussian centered random variables such that $E\tilde{X}_n^2 = E\tilde{Y}_n^2 = 1$,

$$c_n = \frac{\pi^H \sqrt{2c}}{x_n^{H+1} J_{1-H}(x_n)}, \quad n = 1, 2, \dots, \quad (5)$$

$$d_n = \frac{\pi^H \sqrt{2c}}{y_n^{H+1} J_{-H}(y_n)}, \quad n = 1, 2, \dots, \quad (6)$$

$$c = \frac{\Gamma(2H+1) \sin(\pi H)}{\pi^{2H+1}}. \quad (7)$$

Definition 5. A stochastic process $Z = \{Z(t), t \in [0, 1]\}$ is called generalized fractional Brownian motion (GFBM) with a Hurst parameter $H \in (0, 1)$, if $EZ(t) = 0$ and

$$R_H(t, s) = EZ(s)Z(t) = \frac{1}{2}(t^{2H} + s^{2H} - |s - t|^{2H}).$$

If, in addition, the process $Z = \{Z(t), t \in [0, 1]\}$ belongs to a space $\text{Sub}_\varphi(\Omega)$, then it is said to be φ -sub-Gaussian generalized fractional Brownian motion (φ -GFBM).

Let

$$Z(t) = \sum_{n=1}^{\infty} \xi_n c_n \sin x_n t + \sum_{n=1}^{\infty} \eta_n d_n (1 - \cos y_n t), \quad (8)$$

where $t \in [0, 1]$, c_n and d_n are given in (5) and (6), ξ_n, η_n are independent centered random variables such that $E\xi_n^2 = E\eta_n^2 = 1, n = 1, 2, \dots$. Then the following proposition holds.

Proposition 1. (Kozachenko, Sottinen, & Vasylyk, 2005) *The series in (8) converge in mean square and the covariance function of the process Z is:*

$$EZ(s)Z(t) = \frac{1}{2}(t^{2H} + s^{2H} - |s - t|^{2H}).$$

According to Definition 5, the process Z , defined by (8), is a generalized fractional Brownian motion.

Theorem 1. (Kozachenko, Sottinen, & Vasylyk, 2005) *The series in (8) converge uniformly with probability one to the process $Z(t)$, and the process $Z(t)$ is uniformly continuous on $[0, 1]$ with probability one.*

Let $\xi_n, \eta_n, n = 1, 2, \dots$, be independent identically distributed strictly φ -sub-Gaussian random variables, such that $E\xi_n^2 = E\eta_n^2 = 1, n = 1, 2, \dots$.

Then, according to Example 1, Proposition 1 and Definition 5, the process $Z = \{Z(t), t \in [0, 1]\}$ defined by the representation

$$Z(t) = \sum_{n=1}^{\infty} \xi_n c_n \sin x_n t + \sum_{n=1}^{\infty} \eta_n d_n (1 - \cos y_n t), \quad (9)$$

where c_n and d_n are given in (5) and (6), is a strictly φ -sub-Gaussian generalized fractional Brownian motion.

In what follows, we assume that the function $\varphi(\sqrt{\cdot})$ is convex.

Definition 6. The model $\tilde{Z}$ approximates the process Z with given *reliability* $1 - \nu, 0 < \nu < 1$, and *accuracy* $\delta > 0$ in $C([0, 1])$, if

$$\mathbf{P} \left(\sup_{t \in [0, 1]} |Z(t) - \tilde{Z}(t)| > \delta \right) \leq \nu.$$

A natural model for Z , defined by the expansion (9), would be the truncated series

$$\sum_{n=1}^N (c_n \sin(x_n t) \xi_n + d_n (1 - \cos(y_n t)) \eta_n).$$

As in (Kozachenko, Sottinen, & Vasylyk, 2005), we assume that the constants c_n and d_n and the zeros x_n, y_n are only calculated approximately, since there are fast-to-compute asymptotic formulas for the zeros x_n and y_n (Watson, 1944, p. 506).

Let $\tilde{c}_n$ and $\tilde{d}_n$ be the approximated values of the c_n and d_n , respectively. Let

$$|\tilde{c}_n - c_n| \leq \gamma_n^c, \quad |\tilde{d}_n - d_n| \leq \gamma_n^d,$$

$n = 1, \dots, N$. The errors γ_n^c and γ_n^d are assumed to be known. Let $\tilde{x}_n$ and $\tilde{y}_n$ be approximations of the corresponding zeros x_n and y_n with error bounds

$$|\tilde{x}_n - x_n| \leq \gamma_n^x, \quad |\tilde{y}_n - y_n| \leq \gamma_n^y.$$

The error bounds γ_n^x and γ_n^y are also assumed to be known.

Then, the model of the process Z is

$$\tilde{Z}(t) = \sum_{n=1}^N \left(\tilde{c}_n \sin(\tilde{x}_n t) \xi_n + \tilde{d}_n (1 - \cos(\tilde{y}_n t)) \eta_n \right), \quad (10)$$

and simulation error has the form

$$\begin{aligned} \Delta_t := Z(t) - \tilde{Z}(t) &= \sum_{n=1}^N \left\{ \left(c_n \sin(x_n t) - \tilde{c}_n \sin(\tilde{x}_n t) \right) \xi_n + \left(d_n (1 - \cos(y_n t)) - \tilde{d}_n (1 - \cos(\tilde{y}_n t)) \right) \eta_n \right\} + \\ &+ \sum_{n=N+1}^{\infty} \left\{ c_n \sin(x_n t) \xi_n + d_n (1 - \cos(y_n t)) \eta_n \right\}. \end{aligned}$$

In order to estimate Δ in $C([0, 1])$, it was necessary to estimate $\tau_\varphi(\Delta_t)$ and $\tau_\varphi(\Delta_t - \Delta_s)$ for all $s, t \in [0, 1]$, and it was done in (Kozachenko, Sottinen, & Vasylyk, 2005). Then on the basis of the obtained estimates for the simulation error and using the corresponding estimates of the distributions of supremums of φ -sub-Gaussian random processes, the following theorem on the simulation of φ -sub-Gaussian generalized fractional Brownian motion with given reliability and accuracy in the space $C([0, 1])$ was proved in (Kozachenko, Sottinen, & Vasylyk, 2005, Theorem 4.3).

Theorem 2. (Kozachenko, Sottinen, & Vasylyk, 2005) *Let b and α be such that $0 < b < \alpha < H$. The model $\tilde{Z}$, defined by (10), approximates the process Z , defined by (9), with given reliability $1 - \nu$, $0 < \nu < 1$, and accuracy $\delta > 0$ in $C([0, 1])$, if the following three inequalities are satisfied:*

$$\gamma_0 < \delta, \quad \frac{\beta \gamma_0}{\gamma_\alpha} < \frac{\delta}{2^\alpha (\exp\{\varphi(1)\} - 1)^\alpha}, \quad \nu \geq 2 \exp \left\{ -\varphi^* \left(\frac{\delta}{\gamma_0} - 1 \right) \right\} \left(\frac{1}{2^{b(1-\frac{b}{\alpha})}} \left(\frac{\gamma_\alpha \delta}{\beta \gamma_0} \right)^{\frac{b}{\alpha}} l^{(-1)} \left(\frac{\delta}{\gamma_0} - 1 \right) + 1 \right)^{\frac{2}{b}},$$

where the parameters $\gamma_0 = \gamma_0(N)$, $\gamma_\alpha = \gamma_\alpha(N)$ and $\beta = \min\{\gamma_0, \frac{\gamma_\alpha}{2^\alpha}\}$ are expressed through the estimates derived for the simulation error, and $l^{(-1)}$ is the generalised inverse function of the density function l of the function φ .

Remark 1. The density function l of the Orlicz N-function φ is nondecreasing right continuous function such that $l(u) > 0$ as $u > 0$, $l(0) = 0$, $l(u) \rightarrow \infty$ as $u \rightarrow \infty$ and $\varphi(x) = \int_0^{|x|} l(u) du$, $x \in \mathbb{R}$. The function $l^{(-1)}$ is the generalised inverse function of l , i.e. $l^{(-1)}(x) := \sup\{v \geq 0 : l(v) \leq x\}$.

Also we need the following Corollary (Kozachenko, Sottinen, & Vasylyk, 2005, Corollary 4.7). Assume that the constants c_n and d_n , as well as the zeros x_n and y_n are calculated exactly.

Denote $a_\varphi := \tau_\varphi(\xi_n) = \tau_\varphi(\eta_n)$.

Corollary 1. (Kozachenko, Sottinen, & Vasylyk, 2005) *Suppose that there is no approximation error, i.e. $\gamma_n^c = \gamma_n^d = \gamma_n^x = \gamma_n^y = 0$. Then the conditions of the Theorem 2 are satisfied if*

$$N \geq \max \left\{ \left(\frac{A_0}{\delta} \right)^{\frac{1}{H}} + 1; \left(\frac{A_0 (\exp\{\varphi(1)\} - 1)^\alpha}{\delta} \right)^{\frac{1}{H}} + 1; 2 \left(\frac{A_0}{A_\alpha} \right)^{\frac{1}{\alpha}} \right\} \quad (11)$$

and

$$\nu \geq 2 \exp \left\{ -\varphi^* \left(\frac{\delta N^H}{A_0} - 1 \right) \right\} \times \left(\frac{(\delta A_\alpha)^{\frac{b}{\alpha}} (N+1)^{2Hb/\alpha}}{2^b \left(1 - \frac{b}{\alpha} \right) A_0^{\frac{2b}{\alpha}} N^{\frac{(H-\alpha)b}{\alpha}}} l^{-1} \left(\frac{\delta (N+1)^H}{A_0} - 1 \right) + 1 \right)^{\frac{2}{b}}, \quad (12)$$

where

$$A_0 = a_\varphi \sqrt{\frac{5c}{2H}}, \quad A_\alpha = 2^{1-\alpha} a_\varphi \pi^\alpha \sqrt{\frac{c}{H-\alpha}}.$$

3. Main results

Based on the above results, we derive the conditions under which the model of the form (10) constructed for the optimal values of the parameters b and α approximates the process of a strictly φ -sub-Gaussian generalized fractional Brownian motion with given reliability and accuracy in the space $C([0, 1])$ in the case when

$$\varphi(x) = \begin{cases} \frac{|x|^\omega}{\omega}, & |x| \geq 1; \\ \frac{x^2}{\omega}, & |x| < 1, \end{cases} \quad \omega \geq 2.$$

The next theorem presents the main results of our study.

Theorem 3. Let Z be a strictly φ -sub-Gaussian generalized fractional Brownian motion with Hurst parameter $H \in (0, 1)$ and the function $\varphi(x) = \begin{cases} \frac{|x|^\omega}{\omega}, & |x| \geq 1; \\ \frac{x^2}{\omega}, & |x| < 1, \end{cases} \quad \omega \geq 2$. The model (10) approximates the process Z with given reliability $1 - \nu$, $0 < \nu < 1$, and accuracy $\delta > 0$ in the space $C([0, 1])$, if the following conditions hold:

$$N \geq \left(\frac{a_\varphi}{\delta} \sqrt{\frac{5c}{2H}} \right)^{\frac{1}{H}} + 1 \quad \text{and} \quad \frac{\pi^2 N^2}{8} \exp \left[\frac{1}{H(1-x_G)} - \frac{1}{q} (\sqrt{G} - 1)^q \right] \leq \nu, \quad (13)$$

where

$$G = \frac{N^{2H} \delta^{2H}}{5c a_\varphi^2}, \quad x_G = R \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y_G)^2} \right)^{\frac{1}{y_G}} \right), \quad y_G = R \left(G^{\frac{1}{2(\omega-1)}} \right), \quad (14)$$

$c = \frac{\Gamma(2H+1) \sin(\pi H)}{\pi^{2H+1}}$, $q > 0 : \frac{1}{\omega} + \frac{1}{q} = 1$, and $R(p)$ is a solution of the following equation with respect to z :

$$\frac{1}{1-z} + \ln(1-z) - \ln p - 1 = 0, \quad \text{for } z \in (0, 1), \quad p > 1. \quad (15)$$

In addition, the resulting values of the parameters $\alpha = H x_G$ and $b = \alpha y_G$ are optimal, that is, they make it possible to determine the minimum required number of terms N in the model (10) comparing with other possible values $0 < b < \alpha < H$.

Proof. Consider conditions of Theorem 2 and Corollary 12 in the case when $\varphi(x) = \frac{x^2}{\omega} \mathbb{I}_{\{|x|<1\}} + \frac{|x|^\omega}{\omega} \mathbb{I}_{\{|x|\geq 1\}}$, $\omega \geq 2$. Then

$$\varphi^*(x) = \frac{x^q}{q}, \quad q > 0 : \frac{1}{\omega} + \frac{1}{q} = 1, \quad \text{and} \quad l(x) = \varphi'(x) = x^{\omega-1}, \quad l^{-1}(x) = x^{\frac{1}{\omega-1}}, \quad x \geq 1.$$

Thus,

$$N \geq \max \left\{ \left(\frac{A_0}{\delta} \right)^{\frac{1}{H}} + 1; \quad 2 \left(\frac{A_0}{A_\alpha} \right)^{\frac{1}{\alpha}} \right\} \quad (16)$$

and

$$\nu \geq 2 \exp \left\{ -\frac{1}{q} \left(\frac{\delta N^H}{A_0} - 1 \right)^q \right\} \left(\frac{(\delta A_\alpha)^{\frac{b}{\alpha}} (N+1)^{\frac{2Hb}{\alpha}}}{2^b (1-\frac{b}{\alpha}) A_0^{\frac{2b}{\alpha}} N^{\frac{(H-\alpha)b}{\alpha}}} \left(\frac{\delta (N+1)^H}{A_0} - 1 \right)^{\frac{1}{\omega-1}} + 1 \right)^{\frac{2}{b}}. \quad (17)$$

Recall that $A_0 = a_\varphi \sqrt{\frac{5c}{2H}}$ and $A_\alpha = 2^{1-\alpha} a_\varphi \pi^\alpha \sqrt{\frac{c}{H-\alpha}}$.

Let us consider the second inequality in (16)

$$\begin{aligned} 2 \left(\frac{A_0}{A_\alpha} \right)^{\frac{1}{\alpha}} &= 2 \left(\frac{a_\varphi \sqrt{\frac{5c}{2H}}}{2^{(1-\alpha)} a_\varphi \pi^\alpha \sqrt{\frac{c}{H-\alpha}}} \right)^{\frac{1}{\alpha}} = \frac{2}{\pi} \left(\frac{\frac{5c}{2H}}{4^{(1-\alpha)} \frac{c}{H-\alpha}} \right)^{\frac{1}{2\alpha}} = \\ &= \frac{4}{\pi} \left(\frac{5(H-\alpha)}{8H} \right)^{\frac{1}{2\alpha}} \leq \frac{4}{\pi} \left(\frac{5}{8} \right)^{\frac{1}{2\alpha}} \leq \frac{4}{\pi} \left(\frac{5}{8} \right)^{\frac{1}{2H}} \leq \frac{4}{\pi} \left(\frac{5}{8} \right)^{\frac{1}{2}} \approx 1,00658. \end{aligned} \quad (18)$$

Therefore from (18) follows that

$$N \geq \left(\frac{a_\varphi}{\delta} \sqrt{\frac{5c}{2H}} \right)^{\frac{1}{H}} + 1. \quad (19)$$

Since N is a large number, then in (17) we have

$$\left(\frac{(\delta A_\alpha)^{\frac{b}{\alpha}} (N+1)^{\frac{2Hb}{\alpha}}}{2^b (1-\frac{b}{\alpha}) A_0^{\frac{2b}{\alpha}} N^{\frac{(H-\alpha)b}{\alpha}}} \left(\frac{\delta (N+1)^H}{A_0} - 1 \right)^{\frac{1}{\omega-1}} + 1 \right)^{\frac{2}{b}} \approx$$

$$\approx \left(\frac{(\delta A_\alpha)^{\frac{b}{\alpha}} N^{\frac{2Hb}{\alpha}}}{2^b \left(1 - \frac{b}{\alpha}\right) A_0^{\frac{2b}{\alpha}} N^{\frac{(H-\alpha)b}{\alpha}}} \left(\frac{\delta N^H}{A_0} \right)^{\frac{1}{\omega-1}} \right)^{\frac{2}{b}}.$$

Rewrite the obtained expression in details:

$$\begin{aligned} & \left(\frac{(\delta A_\alpha)^{\frac{b}{\alpha}} N^{\frac{2Hb}{\alpha}}}{2^b \left(1 - \frac{b}{\alpha}\right) A_0^{\frac{2b}{\alpha}} N^{\frac{(H-\alpha)b}{\alpha}}} \left(\frac{\delta N^H}{A_0} \right)^{\frac{1}{\omega-1}} \right)^{\frac{2}{b}} = \\ & = \left(\frac{\left(\delta 2^{1-\alpha} a_\varphi \pi^\alpha \sqrt{\frac{c}{H-\alpha}} \right)^{\frac{b}{\alpha}} N^{\frac{2Hb}{\alpha}}}{2^b \left(1 - \frac{b}{\alpha}\right) \left(a_\varphi \sqrt{\frac{5c}{2H}} \right)^{\frac{2b}{\alpha}} N^{\frac{(H-\alpha)b}{\alpha}}} \left(\frac{\delta N^H}{a_\varphi \sqrt{\frac{5c}{2H}}} \right)^{\frac{1}{\omega-1}} \right)^{\frac{2}{b}} = \\ & = \pi^2 \left(\frac{H}{5} \right)^{\frac{2}{\alpha} + \frac{1}{b(\omega-1)}} 2^{\frac{4}{\alpha} + \frac{1}{b(\omega-1)} - 4} (H - \alpha)^{-\frac{1}{\alpha}} \left(\frac{\delta^2}{a_\varphi^2 c} \right)^{\frac{1}{\alpha} + \frac{1}{b(\omega-1)}} \left(1 - \frac{b}{\alpha} \right)^{-\frac{2}{b}} N^{2H \left(\frac{1}{\alpha} + \frac{1}{b(\omega-1)} \right) + 2} = \\ & = \pi^2 \left(\frac{H}{5} \right)^{\frac{1}{\alpha}} 2^{\left(\frac{3}{\alpha} - 4 \right)} (H - \alpha)^{-\frac{1}{\alpha}} \left(1 - \frac{b}{\alpha} \right)^{-\frac{2}{b}} N^2 \left(\frac{N^{2H} \delta^2 2H}{5 c a_\varphi^2} \right)^{\left(\frac{1}{\alpha} + \frac{1}{b(\omega-1)} \right)} = \\ & = \frac{\pi^2 N^2}{16} \left(\frac{8H}{5(H - \alpha)} \right)^{\frac{1}{\alpha}} \left(1 - \frac{b}{\alpha} \right)^{-\frac{2}{b}} \left(\frac{N^{2H} \delta^2 2H}{5 c a_\varphi^2} \right)^{\left(\frac{1}{\alpha} + \frac{1}{b(\omega-1)} \right)}. \end{aligned} \quad (20)$$

For the sake of simplicity we put $x := \frac{\alpha}{H}$, and $y := \frac{b}{\alpha} = \frac{b}{xH}$. Now we consider new unknown parameters $(x, y) \in (0, 1)^2$. The expression (20) in new notations takes the form

$$\left(\frac{\pi N}{4} \right)^2 \left(\frac{8}{5(1-x)} \right)^{\frac{1}{xH}} (1-y)^{-\frac{2}{xyH}} \left(\frac{N^{2H} \delta^2 2H}{5 c a_\varphi^2} \right)^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{xH}}.$$

Let us use (14), in particular $G = \frac{N^{2H} \delta^2 2H}{5 c a_\varphi^2}$. From (19) follows that $G > 1$. Then we get

$$\left(\frac{\pi N}{4} \right)^2 \left(\left(\frac{8}{5(1-x)} \right)^{\frac{1}{x}} (1-y)^{-\frac{2}{xy}} G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} \right)^{\frac{1}{H}}. \quad (21)$$

Thus, now the main task consists in finding such values of the parameters x and y , which minimize the function

$$S^G(x, y) := \left(\frac{8}{5(1-x)} \right)^{\frac{1}{x}} (1-y)^{-\frac{2}{xy}} G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} = \left(\frac{8}{5} \frac{G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} (1-y)^{-\frac{2}{y}}}{(1-x)} \right)^{\frac{1}{x}}. \quad (22)$$

In order to do that we do the following:

- find partial derivatives $\frac{\partial S^G}{\partial x}$, $\frac{\partial S^G}{\partial y}$ and solve corresponding equations $\frac{\partial S^G}{\partial x} = 0$, $\frac{\partial S^G}{\partial y} = 0$;
- check the correctness of the found solutions;
- check the matrix of second derivatives for positive definiteness.

Let us take the partial derivative on x :

$$\begin{aligned} \frac{\partial S^G}{\partial x} &= - \frac{\left(\frac{8}{5} \right)^{\frac{1}{x}} \left(\frac{G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} (1-y)^{-\frac{2}{y}}}{(1-x)} \right)^{\frac{1}{x}} \ln \frac{8}{5}}{x^2} + \\ &+ \left(\frac{8}{5} \right)^{\frac{1}{x}} \left(\frac{G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} (1-y)^{-\frac{2}{y}}}{(1-x)} \right)^{\frac{1}{x}} \left(\frac{1}{(1-x)x} - \frac{\ln \left(\frac{G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} (1-y)^{-\frac{2}{y}}}{(1-x)} \right)}{x^2} \right) = \\ &= - \frac{S^G(x, y) \ln \frac{8}{5}}{x^2} + S^G(x, y) \left(\frac{1}{(1-x)x} - \frac{\ln \left(\frac{G^{\left(\frac{1}{y(\omega-1)} + 1 \right) \frac{1}{x}} (1-y)^{-\frac{2}{y}}}{(1-x)} \right)}{x^2} \right) = \end{aligned}$$

$$\begin{aligned}
 &= \frac{S^G(x, y)}{x^2} \left(\frac{x}{(1-x)} - \ln \left(\frac{8}{5} \frac{G^{(1+\frac{1}{y(\omega-1)})} (1-y)^{-\frac{2}{y}}}{1-x} \right) \right) = \\
 &= \frac{S^G(x, y)}{x^2} \left(\frac{1}{1-x} + \ln(1-x) - \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y}} \right) - 1 \right). \quad (23)
 \end{aligned}$$

Now we take the partial derivative on y :

$$\begin{aligned}
 \frac{\partial S^G}{\partial y} &= \frac{\partial}{\partial y} \left(\frac{8}{5} \frac{G^{(\frac{1}{y(\omega-1)}+1)} (1-y)^{-\frac{2}{y}}}{(1-x)} \right)^{\frac{1}{x}} = \\
 &= \left(\frac{8}{5(1-x)} \right)^{\frac{1}{x}} \frac{1}{x} \left(G^{(\frac{1}{y(\omega-1)}+1)} (1-y)^{-\frac{2}{y}} \right)^{(\frac{1}{x}-1)} \times \\
 &\times \left(-\frac{G^{(\frac{1}{y(\omega-1)}+1)} \ln G (1-y)^{-\frac{2}{y}}}{y^2 (\omega-1)} + G^{(\frac{1}{y(\omega-1)}+1)} (1-y)^{-\frac{2}{y}} \left(\frac{2}{(1-y)y} + \frac{2 \ln(1-y)}{y^2} \right) \right) = \\
 &= \frac{1}{x} S^G(x, y) \left(-\frac{\ln G}{y^2 (\omega-1)} + \frac{2}{(1-y)y} + \frac{2 \ln(1-y)}{y^2} \right) = \\
 &= \frac{2}{xy^2} S^G(x, y) \left(-\frac{\ln G}{2(\omega-1)} + \frac{1}{1-y} - 1 + \ln(1-y) \right). \quad (24)
 \end{aligned}$$

Next we solve the equations $\frac{\partial S^G}{\partial y} = 0$ and $\frac{\partial S^G}{\partial x} = 0$.

We use the facts that $(x, y) \in (0, 1)^2$ and $S^G(x, y) > 0$. From (23) and (24) we get a corresponding equivalent system of equations:

$$\begin{cases} \frac{1}{1-x} + \ln(1-x) - \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y}} \right) - 1 = 0; \\ \frac{1}{1-y} + \ln(1-y) - \ln G^{\frac{1}{2(\omega-1)}} - 1 = 0. \end{cases} \quad (25)$$

Consider (15) with respect to $z \in (0, 1)$, with the parameter $p > 1$: $\frac{1}{1-z} + \ln(1-z) - \ln p - 1 = 0$. Denote by $R(p)$ the solution of this equation. It exists and is unique for $p > 1$. Really, the function $\frac{1}{1-z} + \ln(1-z) - \ln p - 1$ is strictly increasing on $z \in (0, 1)$, because

$$\left(\frac{1}{1-z} + \ln(1-z) - \ln p - 1 \right)'_z = \frac{z}{(1-z)^2} > 0.$$

Also

$$\frac{1}{1-0} + \ln(1-0) - \ln p - 1 = -\ln p < 0, \quad \text{and} \quad \lim_{z \rightarrow 1-} \left(\frac{1}{1-z} + \ln(1-z) - \ln p - 1 \right) = +\infty.$$

Denote by x_G and y_G corresponding solutions to the system (25). Then we can rewrite them in the form

$$y_G = R \left(G^{\frac{1}{2(\omega-1)}} \right), \quad x_G = R \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y_G}} \right).$$

Now we check that these solutions are defined correctly. It is true, since $G > 1$, then $R(\sqrt{G})$ exists. Also we get that

$\frac{8}{5} G^{1+\frac{1}{y(\omega-1)}} > 1$. Obviously $(1-y)^{\frac{2}{y}} < 1$. Thus, $\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y_G}} > 1$ and x_G is also defined correctly.

Next we check whether the point (x_G, y_G) is a point of minimum of the function $S^G(x, y)$, which is defined in (22). For this, it is necessary to check the positive definiteness of the matrix of second derivatives.

Let's find the second partial derivative with respect to x using (23):

$$\begin{aligned}
 \frac{\partial^2 S^G}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{S^G(x, y)}{x^2} \left(\frac{1}{1-x} + \ln(1-x) - \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y}} \right) - 1 \right) \right) = \\
 &= \frac{\partial}{\partial x} \left[\frac{S^G(x, y)}{x^2} \right] \left(\frac{1}{1-x} + \ln(1-x) - \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y}} \right) - 1 \right) +
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{S^G(x, y)}{x^2} \frac{\partial}{\partial x} \left(\frac{1}{1-x} + \ln(1-x) - \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y}} \right) - 1 \right) = \\
 & = \frac{\partial}{\partial x} \left[S^G(x, y) \right] \left(\frac{1}{1-x} + \ln(1-x) - \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y)^2} \right)^{\frac{1}{y}} \right) - 1 \right) + \frac{S^G(x, y)}{x(1-x)^2}.
 \end{aligned} \tag{26}$$

The second partial derivative with respect to y is equal to

$$\begin{aligned}
 \frac{\partial^2 S^G}{\partial y^2} &= \frac{\partial}{\partial y} \left[\frac{2}{xy^2} S^G(x, y) \left(\frac{1}{1-y} + \ln(1-y) - \frac{1}{2(\omega-1)} \ln G - 1 \right) \right] = \\
 &= \frac{\partial}{\partial y} \left(\frac{2}{xy^2} S^G(x, y) \right) \left(\frac{1}{1-y} + \ln(1-y) - \frac{1}{2(\omega-1)} \ln G - 1 \right) + \\
 &+ \frac{2}{xy^2} S^G(x, y) \frac{\partial}{\partial y} \left(\frac{1}{1-y} + \ln(1-y) - \frac{1}{2(\omega-1)} \ln G - 1 \right) = \\
 &= \frac{\partial}{\partial y} \left[\frac{2}{xy^2} S^G(x, y) \right] \left(\frac{1}{1-y} + \ln(1-y) - \frac{1}{2(\omega-1)} \ln G - 1 \right) + \frac{2S^G(x, y)}{xy(1-y)^2}.
 \end{aligned} \tag{27}$$

The mixed derivative is equal to

$$\begin{aligned}
 \frac{\partial^2 S^G}{\partial x \partial y} &= \frac{\partial^2 S^G}{\partial y \partial x} = \frac{\partial}{\partial x} \left(\frac{2}{xy^2} S^G(x, y) \left(\frac{1}{1-y} + \ln(1-y) - \frac{1}{2(\omega-1)} \ln G - 1 \right) \right) = \\
 &= \frac{\partial}{\partial x} \left[\frac{2}{xy^2} S^G(x, y) \right] \left(\frac{1}{1-y} + \ln(1-y) - \frac{1}{2(\omega-1)} \ln G - 1 \right).
 \end{aligned} \tag{28}$$

From (26), (27), (28) and (25) we get

$$\frac{\partial^2 S^G}{\partial x^2}(x_G, y_G) = \frac{S^G(x_G, y_G)}{x_G(1-x_G)^2} > 0; \quad \frac{\partial^2 S^G}{\partial y^2}(x_G, y_G) = \frac{2S^G(x_G, y_G)}{x_G y_G(1-y_G)^2} > 0; \quad \frac{\partial^2 S^G}{\partial x \partial y}(x_G, y_G) = 0.$$

The positive definiteness of the matrix of second derivatives follows from Sylvester's criterion:

$$\begin{aligned}
 \Delta_1 &= \frac{\partial^2 S^G}{\partial x^2} > 0; \\
 \Delta_2 &= \left| \begin{pmatrix} \frac{\partial^2 S^G}{\partial x^2} & \frac{\partial^2 S^G}{\partial x \partial y} \\ \frac{\partial^2 S^G}{\partial x \partial y} & \frac{\partial^2 S^G}{\partial y^2} \end{pmatrix} \right|_{(x_G, y_G)} = \left| \begin{pmatrix} \frac{\partial^2 S^G}{\partial x^2}(x_G, y_G) & 0 \\ 0 & \frac{\partial^2 S^G}{\partial y^2}(x_G, y_G) \end{pmatrix} \right| = \\
 &= \frac{\partial^2 S^G}{\partial x^2}(x_G, y_G) \frac{\partial^2 S^G}{\partial y^2}(x_G, y_G) > 0.
 \end{aligned}$$

Note the monotonicity of the function $S^G(x_G, y_G)$ of the parameter G .

$$\begin{aligned}
 \frac{d}{dG} S^G(x_G, y_G) &= \frac{\partial S^G}{\partial x}(x_G, y_G) \frac{\partial x_G}{\partial G} + \frac{\partial S^G}{\partial y}(x_G, y_G) \frac{\partial y_G}{\partial G} + \frac{\partial S^G}{\partial G}(x_G, y_G) = \\
 &= \frac{\partial S^G}{\partial G}(x_G, y_G) = \frac{\partial}{\partial G} \left(\left(\frac{8}{5} \frac{G^{\left(\frac{1}{y(\omega-1)}+1\right)} (1-y)^{-\frac{2}{y}}}}{1-x} \right)^{\frac{1}{x}} \right) = \\
 &= \frac{1}{x} \left(\frac{8}{5} \frac{G^{\left(\frac{1}{y(\omega-1)}+1\right)} (1-y)^{-\frac{2}{y}}}}{1-x} \right) \frac{8}{5} \left(\frac{1}{y(\omega-1)} + 1 \right) \frac{G^{\frac{1}{y(\omega-1)}} (1-y)^{-\frac{2}{y}}}{1-x} = \\
 &= \frac{1}{x} \left(\frac{1}{y(\omega-1)} + 1 \right) \left(\frac{8}{5} \right)^{\frac{1}{x}} \left((1-y)^{-\frac{2}{y}} \right)^{\frac{1}{x}} \frac{1}{(1-x)^{\frac{1}{x}}} \left(G^{\left(\frac{1}{y(\omega-1)}+1\right)} \right)^{\frac{1}{x}-1} = \\
 &= \frac{1}{x} \left(\frac{1}{y(\omega-1)} + 1 \right) S^G(x_G, y_G) \frac{1}{G} > 0.
 \end{aligned}$$

Let us transform the first equation from (25):

$$\begin{aligned} \frac{x_G}{1-x_G} + \ln(1-x_G) &= \ln \left(\frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y_G)^2} \right)^{\frac{1}{y_G}} \right) \Leftrightarrow \\ \Leftrightarrow \frac{8G}{5} \left(\frac{G^{\frac{1}{\omega-1}}}{(1-y_G)^2} \right)^{\frac{1}{y_G}} &= \exp \left\{ \frac{x_G}{1-x_G} \right\} (1-x_G). \end{aligned}$$

Then we can write the function $S^G(x_G, y_G)$ in the form:

$$S^G(x_G, y_G) = \left(\frac{8G^{\left(\frac{1}{y_G(\omega-1)}+1\right)} (1-y_G)^{-\frac{2}{y_G}}}{5(1-x_G)} \right)^{\frac{1}{x_G}} = \left(\frac{e^{\frac{x_G}{1-x_G}} (1-x_G)}{(1-x_G)} \right) = e^{\frac{1}{1-x_G}}. \quad (29)$$

Thus, the condition (17) is satisfied if

$$\begin{aligned} \left(S^G(x_G, y_G) \right)^{\frac{1}{H}} \frac{\pi^2 N^2}{8} \exp \left[-\frac{1}{q} (\sqrt{G}-1)^q \right] &\leq \nu \Leftrightarrow \\ \Leftrightarrow \frac{\pi^2 N^2}{8} \exp \left[\frac{1}{H(1-x_G)} - \frac{1}{q} (\sqrt{G}-1)^q \right] &\leq \nu. \end{aligned}$$

This completes the proof of the theorem. $\square$

4. Simulation results for the optimized models

Using Theorem 3, we construct models $\tilde{Z}$ of φ -sub-Gaussian generalized fractional Brownian motion Z with the function $\varphi(x) = \frac{x^2}{\omega} \mathbb{I}_{\{|x|<1\}} + \frac{|x|^\omega}{\omega} \mathbb{I}_{\{|x|\geq 1\}}$, $\omega \geq 2$, such that $\tilde{Z}$ approximates Z with given reliability $1-\nu = 0.99$ and accuracy $\delta = 0.01$ in the space $C([0, 1])$.

The examples of such models for several values of the Hurst index H and two cases of ω are presented below.

1. $H = \frac{7}{8}$, $\omega = 3$. In this case we have

- $c = 0.0264278$, $\mu = 3.47172 \times 10^{-10}$, $N \geq 1042.53$, $G = 253.316$, $x_G = 0.941959$, $y_G = 0.728925$,
- $\alpha = 0.824214$, $b = 0.600791$.

Thus, it is enough to set $N = 1043$, and as a result we get the following model approximating generalized fractional Brownian motion with parameter $H = \frac{7}{8}$ with reliability 0.99 and accuracy 0.01 in $C([0, 1])$ (Fig. 1).

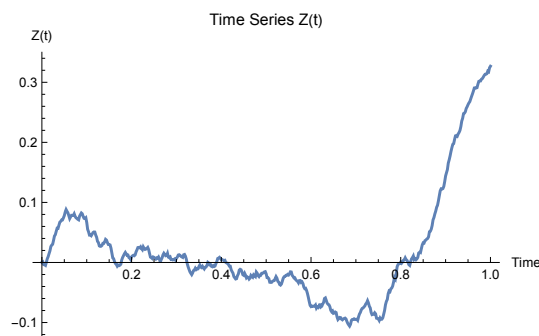


Fig. 1. The model of φ -GFBM with parameter $H = \frac{7}{8}$

2. $H = \frac{8}{9}$, $\omega = 3$. In this case we have:

- $c = 0.0234107$, $\mu = 8.9692 \times 10^{-10}$, $N \geq 853.88$, $G = 247.095$, $x_G = 0.941809$, $y_G = 0.728297$,
- $\alpha = 0.837164$, $b = 0.609704$.

Thus, it is enough to set $N = 854$, and as a result we get the following model, which approximates the process of generalized fractional Brownian motion with parameter $H = \frac{8}{9}$ with reliability 0.99 and accuracy 0.01 in $C([0, 1])$ (Fig. 2).

3. $H = \frac{9}{10}$, $\omega = 3$. In this case we have:

- $c = 0.021007$, $\mu = 1.9722 \times 10^{-9}$, $N \geq 726.468$, $G = 242.174$, $x_G = 0.941687$, $y_G = 0.727786$,
- $\alpha = 0.847518$, $b = 0.616812$.

Thus, it is enough to set $N = 727$, and as a result we get the following model, which approximates the process of generalized fractional Brownian motion with parameter $H = \frac{9}{10}$ with reliability 0.99 and accuracy 0.01 in $C([0, 1])$ (Fig. 3).

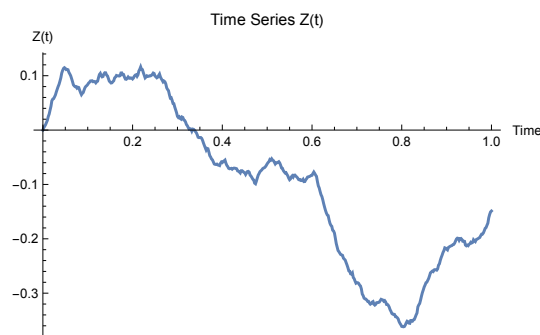


Fig. 2. The model of φ -GFBM with parameter $H = \frac{8}{9}$

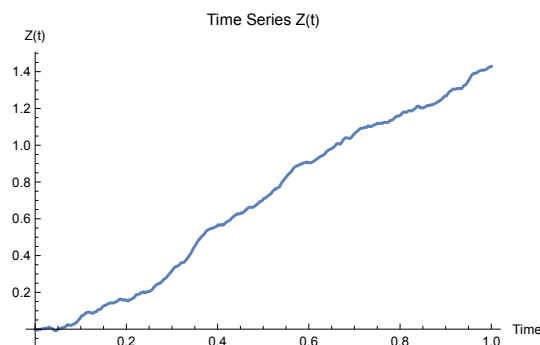


Fig. 3. The model of φ -GFBM with parameter $H = \frac{9}{10}$

4. $H = \frac{15}{16}$, $\omega = 3$.

In this case we have:

- $c = 0.0129786$, $\mu = 4.00608 \times 10^{-8}$, $N \geq 406.697$, $G = 225.521$, $x_G = 0.941249$, $y_G = 0.725958$,
- $\alpha = 0.882421$, $b = 0.640601$.

Thus, it is enough to set $N = 407$, and as a result we get the following model, which approximates the process of generalized fractional Brownian motion with parameter $H = \frac{15}{16}$ with reliability 0.99 and accuracy 0.01 in $C([0, 1])$ (Fig. 4).

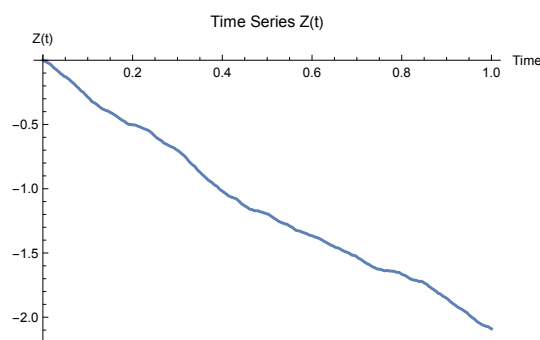


Fig. 4. The model of φ -GFBM with parameter $H = \frac{15}{16}$

5. $H = \frac{3}{4}$, $\omega = 3$. Now, let us take reliability $1 - \nu = 0.95$ and accuracy $\delta = 0.05$. This should reduce the number of terms in the model.

In this case we have:

- $c = 0.0537337$, $\mu = 2.84388 \times 10^{-6}$, $N \geq 732.124$ (for comparison, for reliability $1 - \nu = 0.99$ and accuracy $\delta = 0.01$ we have got $N \geq 6865.58$),
- $G = 276.498$, $x_G = 0.942482$, $y_G = 0.731111$, $\alpha = 0.706862$, $b = 0.516794$.

Thus, it is enough to set $N = 733$, and as a result we get the following model, which approximates the process of generalized fractional Brownian motion with parameter $H = \frac{3}{4}$ with reliability 0.95 and accuracy 0.05 in $C([0, 1])$ (Fig. 5).

6. Let us consider how the behaviour of the simulated trajectory changes in the case when $\omega = 5$, that is, when

$$\varphi(x) = \begin{cases} \frac{|x|^5}{5}, & |x| \geq 1; \\ \frac{x^2}{5}, & |x| < 1. \end{cases}$$

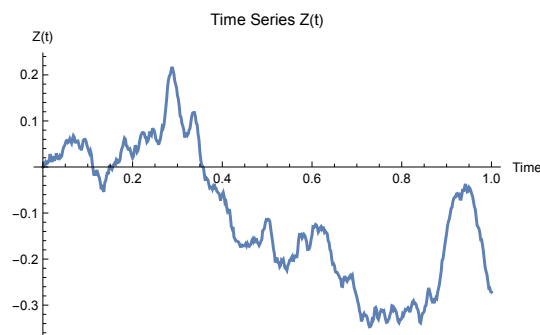


Fig. 5. The model of φ -GFBM with parameter $H = \frac{3}{4}$

The following example is given for $H = \frac{8}{9}$, reliability $1 - \nu = 0.99$ and accuracy $\delta = 0.01$. In this case we have:

- $c = 0.0234107$, $\mu = 1.32915 \times 10^{-6}$, $N \geq 1284.88$, $G = 510.929$, $x_G = 0.937944$, $y_G = 0.644661$,
- $\alpha = 0.833728$, $b = 0.537471$.

Thus, it is enough to set $N = 1285$, and as a result we get the following model, which approximates the process of generalized fractional Brownian motion with parameter $H = \frac{8}{9}$, $\omega = 5$ with reliability 0.99 and accuracy 0.01 in $C([0, 1])$ (Fig. 6).

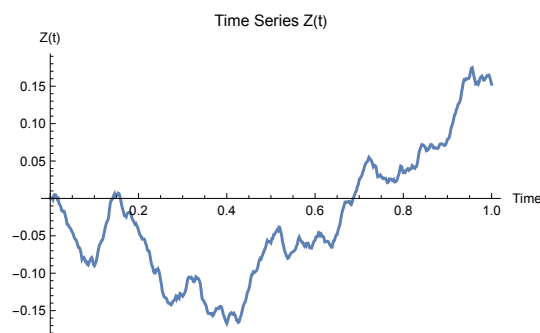


Fig. 6. The model of φ -GFBM with parameter $H = \frac{8}{9}$, $\omega = 5$

Here we see that with a larger value of the parameter ω , larger fluctuations of the simulated process are observed.

Remark 2. All calculations and simulations were performed in Wolfram Mathematica. The authors can share the script of the code with an interested reader.

Discussion and conclusions

In Tab. 1, we compare the results obtained in our work with the results obtained in the article (Vasylyk & Lovytska, 2021). We can conclude that we have found the minimum necessary number of terms N for the models in case of the optimal values of the parameters. As expected, in our case the obtained values of N are significantly smaller than the corresponding numbers for the models with the values of parameters that are not optimal.

Table 1

Comparison of the results					
H	ω	$1 - \nu$	δ	N in the article (Vasylyk & Lovytska, 2021)	N in the optimized model
$\frac{7}{8}$	3	0.99	0.01	1362	1043
$\frac{8}{9}$	3	0.99	0.01	1111	854
$\frac{9}{10}$	3	0.99	0.01	942	727
$\frac{15}{16}$	3	0.99	0.01	523	407
$\frac{3}{4}$	3	0.95	0.05	1006	733
$\frac{8}{9}$	5	0.99	0.01	1693	1285

Authors' contribution: Olga Vasylyk — conceptualization and methodology; Kostiantyn Shurubura — methodology and research, software, optimization of a model, simulation and analysis of the results, preparation of literature review.

Acknowledgments, funding sources. The authors would like to thank the anonymous referee(s) who provided useful and detailed comments on an earlier version of the manuscript. Also we are grateful to Prof. Yuliya Mishura for the very fruitful discussion, comments and recommendations which helped us to improve this paper.

References

- Banna, O., Mishura, Y., Ralchenko, K., & Shklyar, S. (2019). *Fractional brownian motion. approximations and projections*. Wiley-ISTE, 288 p.
- Buldygin, V., & Kozachenko, Y. (2000). *Metric characterization of random variables and random processes*. Translations of Mathematical Monographs. 188. Providence, RI: AMS, American Mathematical Society. 257 p.
- Cai, G. Q., Huan, R. H., & Zhu, W. Q. (2022). *Modeling and simulation of stochastic processes. book chapter in stochastic processes: Fundamentals and emerging applications*. Nova science publishers. <https://doi.org/10.52305/KMDP8644>
- Dzhaparidze, K. O., & Zanten, J. H. (2002). A series expansion of fractional brownian motion. *CWI. Probability, Networks and Algorithms [PNA]*, R 0216, .
- Giuliano, R., Kozachenko, Y., & Nikitina, T. (2003). Spaces of φ -subgaussian random variables. *Rendiconti Accademia Nazionale delle Scienze XL. Memorie di Matematica e Applicazioni* 121., XXVII, 95-124.
- Kozachenko, Y., & Koval'chuk, Y. (1998). Boundary value problems with random initial conditions and series of functions of $Sub_{\varphi}(\Omega)$. *Ukrainian Math. J.*, 50(4), 572-585.
- Kozachenko, Y., & Rozora, I. (2019). Construction of the karhunen-loeve model for an input Gaussian process in a linear system by using the output process. *Theory of Probability and Mathematical Statistics*, 99, 101-112. 10.1090/tpms/1084
- Kozachenko, Y., Sottinen, T., & Vasylyk, O. (2005). Simulation of Weakly Self-Similar Stationary Increment $Sub_{\varphi}(\Omega)$ -Processes: A Series Expansion Approach. *Methodology and Computing in Applied Probability*, 7, 379-400. 10.1007/s11009-005-4523-y
- Kozachenko, Y., & Vasylyk, O. (2006). Simulation of fractional Brownian motion with given reliability and accuracy in $C([0, 1])$. *Theory of Stochastic Processes*, 12 (28)(3-4), 59-66.
- Linda, J., & Allen, S. (2010). *An introduction to stochastic processes with applications to biology*. Routledge Taylor & Francis Group.
- Mode, C. J. (2011). *Applications of monte carlo methods in biology, medicine and other fields of sciences*. Intech Open.
- Molz, F. J., & Liu, H. H. (1997). Fractional Brownian motion and fractional Gaussian noise in subsurface hydrology: A review, presentation of fundamental properties, and extensions. *Water Resources Research*, 33(10), 2273-2286. 10.1029/97WR01982
- Pashko, A., Lukovych, O., Rozora, I., Oleshko, T., & Vasylyk, O. (2019). Analysis of simulation methods for fractional Brownian motion in the problems of intelligent systems design. In *Ieee international conference on advanced trends in information theory* (p. 373-378). ATIT 2019 - Proceedings.
- Pashko, A., & Rozora, I. (2018). Accuracy of simulation for the network traffic in the form of fractional Brownian motion. In *14th international conference on advanced trends in radioelectronics, telecommunications and computer engineering, tcset 2018 - proceedings* (p. 840-845).
- Pashko, A., Rozora, I., & Synyavskiy, O. (2021). Estimation of Hurst index and traffic simulation. In *Advances in computer science for engineering and education iv. lecture notes on data engineering and communications technologies. vol.83* (p. 37-46). Springer. 10.1007/978-3-030-80472-5_4
- Vasylyk, O., & Lovytska, I. (2021). Simulation of a strictly φ -sub-Gaussian generalized fractional Brownian motion. *Bulletin of TSNUK. Series: Physics and Mathematics*, 1, 11-19. <https://doi.org/10.17721/1812-5409.2021/1.1>
- Vasylyk, O., Rozora, I., Ianevych, T., & Lovytska, I. (2021). On some method on model construction for strictly φ -sub-Gaussian generalized fractional Brownian motion. *Bulletin of TSNUK. Series: Physics and Mathematics*, 2, 18-25. <https://doi.org/10.17721/1812-5409.2021/2.3>
- Watson, G. (1944). *A treatise of the theory of bessel functions*. Cambridge University Press, Cambridge, England.

Отримано редакцією журналу / Received: 01.02.24
Прорецензовано / Revised: 10.03.25
Схвалено до друку / Accepted: 14.03.25

Ольга ВАСИЛИК, д-р фіз.-мат. наук, доц.

ORCID ID: 0000-0002-0880-3751

e-mail: vasylyk@matan.kpi.ua

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

Костянтин ШУРУБУРА, асп.

ORCID ID: 0009-0003-5535-2644

e-mail: shurubura@hotmail.com

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

ОПТИМІЗАЦІЯ МОДЕЛІ φ -СУБГАУССОВОГО УЗАГАЛЬНЕНОГО ДРОБОВОГО БРОУНІВСЬКОГО РУХУ

Методи моделювання випадкових процесів використовують у багатьох наукових сферах, й особливу увагу приділяють дробовому броунівському руху. Ученими отримано багато результатів щодо моделювання гауссового дробового броунівського руху. У пропонованій статті розглянуто моделювання процесів із більш загального класу, а саме моделювання строго φ -субгауссового узагальненого дробового броунівського руху, визначеного на відріжку $[0, 1]$. Спочатку в роботі наведено певну попередню інформацію, необхідну для розуміння результатів нашого дослідження. В основній частині роботи для строго φ -субгауссового узагальненого дробового броунівського руху з деякою заданою функцією φ знайдено оптимальні значення параметрів, необхідних для моделювання із заданою надійністю та точністю в просторі $C([0, 1])$. Це зумовлює значне зменшення кількості доданків, необхідних для забезпечення заданої надійності та точності нашої моделі. Також у роботі наведено деякі результати розрахунків та приклади моделювання. Як й очікувалося, отримані значення кількості доданків у моделі є значно меншими за відповідні числа для моделей зі значеннями параметрів, що не є оптимальними.

Ключові слова: φ -субгауссовий узагальнений дробовий броунівський рух, оптимізація моделі, надійність і точність моделі, моделювання, випадковий процес.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.2

DOI: <https://doi.org/10.17721/1812-5409.2025/1.4>

Volodymyr ZUBCHENKO, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0002-1513-9326
e-mail: volodymyr.zubchenko@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Petro ZELENKO, Master's Student
ORCID ID: 0009-0005-9671-0465
e-mail: petro.zelenko@uni-ulm.de
Ulm University, Ulm, Germany

ENTROPIC RISK MEASURE EVAR REAL DATA TESTING

Investigating risk measures is crucial for understanding and managing the uncertainties inherent in various financial and economic activities. Risk measures provide quantitative assessments of potential losses, enabling businesses and financial institutions to make informed decisions and mitigate adverse outcomes. The background of this research lies in the increasing complexity and interconnectedness of financial markets, which have increased the impact of risks and the need for robust risk management frameworks. Traditional measures, such as Value at Risk (VaR), have been widely used, but their limitations, particularly during financial crises, have highlighted the necessity for more comprehensive approaches like Expected Shortfall (ES) and Entropic Value at Risk (EVAR).

The relevance of investigating risk measures is underscored by the financial turmoil experienced during global crises, such as the 2008 financial meltdown and the COVID-19 pandemic. These events exposed the vulnerabilities of existing risk management practices and underscored the need for more resilient and accurate risk assessment tools. The purpose of this research is to evaluate and compare different risk measures, providing insight into their strengths and weaknesses, and to develop more effective strategies for risk management.

In the paper, we investigate the performance of the Entropic Value-at-Risk (EVAR) risk measure on real financial data, comparing it with traditional risk measures such as Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). We consider historical return data from the S&P 500 index and Bank of America Corp. shares (2009–2019) and apply various parametric models, including normal and Laplace distributions. We also employ standard backtesting methods (coverage and independence tests) to assess the empirical adequacy of these measures. Our analysis highlights the mathematical properties of EVAR and its ability to detect heavy-tailed behavior more effectively than VaR, while providing coherent and potentially more robust estimates in volatile market conditions.

Future research perspectives include the development of new risk measures that can better capture systemic risks in financial markets. In addition, integrating machine learning and artificial intelligence with traditional risk measures presents an avenue to improve predictive accuracy and robustness. This ongoing research is important for building more resilient financial systems and for protecting economic stability.

Key words: *Entropic Value-at-Risk, EVAR, coverage tests, heavy tails, real data analysis, normal distribution, Laplace distribution.*

AMS 2020 classification: 62P05, 91-10.

Introduction

Using risk measures is crucial for quantifying potential losses in investments or business activities. This allows companies and investors to better understand the risks they are taking on and how these risks can impact their financial results. For example, statistical metrics such as the standard deviation or value-at-risk (VaR) help assess asset volatility and predict possible fluctuations in their value.

In addition, risk measures are essential for effective risk management. They enable organizations to identify which actions are necessary to minimize potential losses. This could include diversifying an investment portfolio, hedging risks, or implementing internal controls and procedures. By taking these steps, organizations can reduce the impact of adverse events and increase their resilience to financial shocks.

Finally, risk measures are necessary for regulatory compliance. Many financial institutions must adhere to specific standards and regulations that require regular risk assessment and reporting. For example, banks must maintain certain capital levels according to the Basel Agreements, which require the use of various risk measures to evaluate credit, market, and operational risks. This ensures the stability of the financial system and protects the interests of depositors and investors.

The simplest and best-known measure is VaR (Value-at-Risk). Provided that a stock's fluctuation over time has no peaks, VaR is useful for predicting the risk associated with a portfolio. But when we face financial crises of different types, the chance of VaR to give a complete overview of possible risks decreases, as it is not sensitive to anything beyond its loss threshold. In addition, VaR is not coherent. This problem is partially solved by CVaR (Conditional Value-at-Risk), but not well enough. An important disadvantage of CVaR is that it cannot be computed efficiently, even for a sum of arbitrary independent random variables.

Furthermore, financial markets may face high volatility and instability. In such circumstances, traders and managers use more stringent and conservative measures to manage risk rather than VaR and CVaR. In this case a more complicated measure can be used, namely a coherent measure of risk called Entropic Value-at-Risk (EVAR), introduced in (Ahmadi-Javid, 2012). The basic properties of this measure are described in the papers (Ahmadi-Javid, 2012; Ahmadi-Javid & Fallah-Tafti, 2019; Ahmadi-Javid & Pichler, 2017). A quantitative analysis of various risk measures, including EVAR, is performed in (Pichler, 2017). Delbaen provides a relation between EVAR and other commonotone risk measures in (Delbaen, 2019).

The properties of the entropic risk measure EVaR in relation to selected distributions are considered in the paper (Mishura et al., 2024). In continuation of the specified work, in this paper we consider EVaR real data testing.

Considering Entropic Value at Risk in real data testing is important for several reasons:

1. Enhanced Risk Measurement: EVaR provides a more comprehensive risk assessment by capturing tail risk more effectively than traditional measures like VaR or CVaR. This is particularly useful in scenarios with extreme events or heavy-tailed distributions, where conventional methods might underestimate the risk.
2. Robustness to Model Assumptions: EVaR is less sensitive to specific assumptions about the distribution of returns. This makes it a more robust risk measure when dealing with real-world data, which often deviates from theoretical models.
3. Improved Decision-Making: By incorporating EVaR into risk management practices, organizations can make better informed decisions. This can lead to more effective risk mitigation strategies, optimized capital allocation, and improved financial stability, especially in volatile or uncertain market conditions.

EVaR testing with real data ensures that the risk measure is practical and reliable, providing valuable information on the actual risk profile of investments or business activities.

1. Key definitions

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and let $\mathcal{L}_0 = \mathcal{L}_0(\Omega, \mathcal{F}, \mathbf{P})$ be the space of random variables $X: \Omega \rightarrow \mathbb{R}$.

Among various risk measures, the following risk measure is often used in modern risk management practice (Föllmer & Schied, 2016; Holton, 2003).

Definition 1. The *Value-at-Risk* (VaR) with confidence level $\alpha \in (0, 1)$ of a random variable X is the smallest number y such that X does not exceed it with a minimum probability of $1 - \alpha$, i.e.

$$\text{VaR}_\alpha(X) := \inf\{x \in \mathbb{R} : F_X(x) > 1 - \alpha\}.$$

Definition 2. The *Conditional Value-at-Risk* (CVaR) at level α is defined as:

$$\text{CVaR}_\alpha(X) := \frac{1}{1 - \alpha} \mathbb{E}[X | X \geq \text{VaR}_\alpha(X)].$$

However, in this paper, we focus on another risk measure, namely EVaR, introduced in (Ahmadi-Javid, 2012) according to the following definition.

Definition 3. Let $X \in \mathcal{L}_0$. Assume that its moment-generating function $m_X(t) = \mathbf{E}e^{tX}$ is well defined for all $t \geq 0$. An *entropic risk measure* EVaR with a confidence level $\alpha \in [0, 1)$ (or a risk level $1 - \alpha$) is defined as follows:

$$\text{EVaR}_\alpha(X) := \inf_{t \geq 0} t^{-1} \log \left(\frac{1}{1 - \alpha} m_X(t) \right), \quad \alpha \in [0, 1). \quad (1)$$

There is also a related risk measure Entropic Drawdown at Risk (EDaR) described in detail in (Cajas, 2021).

Definition 4. Assume that the moment-generating function of its drawdown series $DD(X, j)$, defined as

$$DD(X, j) = \max_{0 \leq t \leq j} \left(\sum_{i=0}^t X_i \right) - \sum_{i=0}^j X_i,$$

exists. The *Entropic Drawdown at Risk* (EDaR) at confidence level $\alpha \in (0, 1)$ is defined as:

$$\text{EDaR}_\alpha(X) = \text{EVaR}_\alpha(DD(X)).$$

Normal distribution is one of the most frequently used in various models. The formula derived in this theorem is given in (Ahmadi-Javid, 2012).

Theorem 1. Let $X \sim N(\mu, \sigma^2)$ be a normal distribution. Then $\forall \alpha \in (0, 1)$:

$$\text{EVaR}_\alpha(X) = \mu + \sigma \sqrt{2 \log \frac{1}{1 - \alpha}}.$$

We will also need the following result. The proof can be found in (Mishura et al., 2024). Let us now consider losses which have a Laplace distribution.

Theorem 2. Let $X \sim L(\mu, b)$ have a Laplace distribution. Then for all $\alpha \in (0, 1)$,

$$\text{EVaR}_\alpha(X) = \mu - bW_{-1}(\gamma) \left(1 + \frac{2}{W_{-1}(\gamma)} \right)^{\frac{1}{2}}, \quad (2)$$

where $\gamma = -2e^{-2}(1 - \alpha)$, W - Lambert function. The expression for $\text{EVaR}_\alpha(X)$ is continuous in α .

2. EVaR real data testing

The main task of risk measures is to estimate possible losses for real processes, whose distribution is not known in advance. There are many problems with the analysis of real data. First of all, we need to build a mathematical probability model, and only then apply risk measures. In this part, we will look at certain models for working with real data and compare the most well-known probabilistic measures. After each model, we will perform a back-testing procedure, which will also show the advantages and disadvantages of risk measures. We will consider three risk measures: VaR, ES (CVaR), and EVaR. All subsequent actions will be performed only for two thresholds or confidence levels: 95 % and 99 %. All computer modeling will be done using the R programming language.

Consider historical data on the value of the S&P500 index and the value of Bank of America Corp. shares from 2009 to 2019. An important feature of these data is their significant correlation. Their behavior is very similar, although the spread of returns is different. The first model we will consider is based on the assumption of a normal distribution of returns $r(t)$, where $r(t) = \frac{p_t}{p_{t-1}} - 1$ for price p_t in time t .

$$r(t) \sim N(\mu, \sigma^2).$$

Our goal is to determine the daily risk estimate for all of these measures. To calculate the value in time t , we will use the last 250 observations. Based on this sample, we estimate the distribution parameters with the method of moments. That is

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^n r(t),$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{t=1}^n (r(t) - \hat{\mu})^2.$$

To calculate EVaR, we apply the following formula:

$$EVaR_{\alpha}(X) = \mu + \sigma \sqrt{2 \log \frac{1}{1-\alpha}}.$$

For the obtained estimates, we will calculate the risk value using the following formulas:

$$VaR_{\alpha}(X) = \hat{\mu} + \sqrt{\hat{\sigma}^2} F_{N(0,1)}^{-1}(1-\alpha),$$

$$CVaR_{\alpha}(X) = \hat{\mu} + \sqrt{\hat{\sigma}^2} \frac{\phi(F_{N(0,1)}^{-1}(\alpha))}{1-\alpha}.$$

The previous history of 250 days is used to calculate the risk value. The construction process is as follows: we start from the 250th day, calculate a one-day measure using the last 250 observations, and so on until the last day of the period under consideration. The results are shown in Fig. 1.

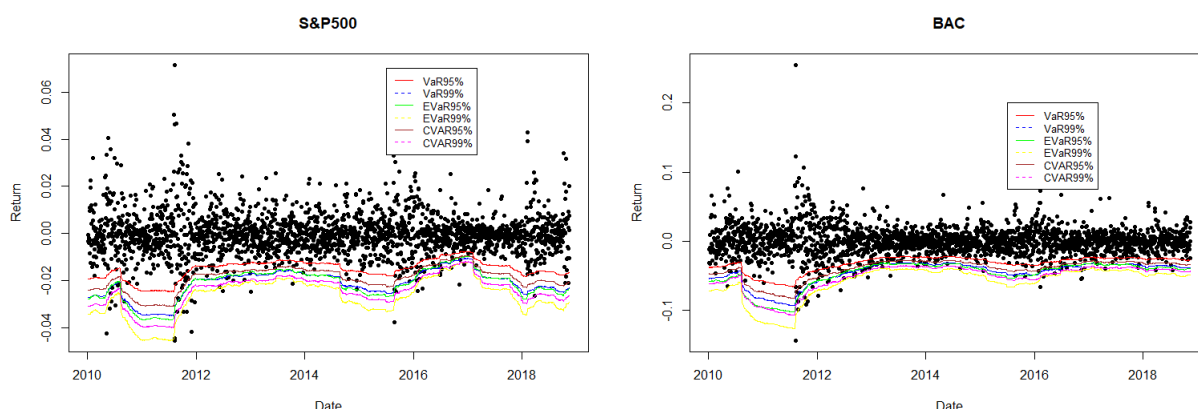


Fig. 1. Normal model

As can be seen from the return charts, the BAC has higher volatility, although both charts look quite similar. Coverage tests are often used to assess the quality of risk measures models. They assess whether the frequency of outliers corresponds to the loss quantile that the risk measure should reflect. Let us apply the standard test, which considers the hypothesis of a binomial distribution of exceedances, i.e.

$$H_0 : \sum_{t=0}^T 1(VaR_{\alpha}(X) < X) \sim Bin(n, \alpha),$$

where n is the number of observations, α is the theoretical probability of excess, which is equal to the confidence level for

VaR. We obtained the following confidence intervals for any VaR model in Tab. 1 since they depend only on the number of observations.

Table 1

Confidence intervals for VaR for binomial coverage test, S&P500 and BAC

Window size / Accuracy	5%	1%
100	[99,140]	[12,37]
250	[92,132]	[11,35]
500	[81,119]	[9,32]
1000	[58,91]	[6,26]

This test cannot be applied directly to CVaR and EVaR. For them, we will estimate the theoretical probability of $\hat{\alpha}$ of outperformance using the simulation method. That is, we simulate 100,000 price trajectories, each step of which has a theoretical distribution with parameters that we determined using the method of moments to calculate CVaR/EVaR at that moment. For each model, the confidence intervals for these risk measures depend on the distribution. The results are shown in Tab. 2.

Table 2

Confidence intervals for CVaR and EVaR for binomial coverage test, S&P500 and BAC

Window size/Measure	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	[34,60]	[2,18]	[16,36]	[0,11]
250	[38,66]	[4,21]	[17,37]	[1,12]
500	[56,89]	[12,37]	[31,56]	[6,26]
1000	[78,116]	[21,52]	[45,74]	[11,35]

The important point of this test is that, on the one hand, a failure to fall within the confidence interval should reject the model for the specified risk measure. However, the measure may be too conservative and the data may not quite fit the model. In this case, it is worth paying attention to whether the number of outputs for the risk measure exceeds the right-hand side of the confidence interval. The confidence intervals show that the risk models will perform differently depending on the amount of data used for parametric estimates of the distribution of the daily risk measure. The second test that we will use is the independence test in Christophersen format (Holton, 2003). This is a likelihood ratio test that looks for unusually frequent consecutive outliers.

$$q_0^* = P(I_t = 0 | I_{t-1} = 0),$$

$$q_1^* = P(I_t = 0 | I_{t-1} = 1).$$

We test the independence hypothesis, i.e. $H_0 : q_0^* = q_1^*$. To do this, we calculate the likelihood ratio

$$\Lambda = \left(\frac{\hat{q}}{\hat{q}_0}\right)^{\alpha_{00}} \left(\frac{1-\hat{q}}{1-\hat{q}_0}\right)^{\alpha_{01}} \left(\frac{\hat{q}}{\hat{q}_1}\right)^{\alpha_{10}} \left(\frac{1-\hat{q}}{1-\hat{q}_1}\right)^{\alpha_{11}}.$$

Where α_{ij} is the number of pairs such that $(I_t, I_{t-1}) = (i, j)$,

$$\hat{q} = \frac{\alpha_{00} + \alpha_{01}}{\alpha_{00} + \alpha_{01} + \alpha_{10} + \alpha_{11}}, \hat{q}_0 = \frac{\alpha_{00}}{\alpha_{00} + \alpha_{01}}, \hat{q}_1 = \frac{\alpha_{10}}{\alpha_{10} + \alpha_{11}}.$$

Then $H_0 : -2 \log(\Lambda) \sim \chi^2(1, 0)$. That is, for our confidence levels, the resulting statistics should not exceed 3.841 for 95 % and 6.635 for 99 %. An important point to note when applying this test is that it is meaningless if there are no outliers. Now, let us calculate the number of excesses of the estimates obtained using the risk measures for this model. Here and in the following, we will mark in green the hits in the confidence interval, in red the outliers, and in yellow the boundary values. The results are shown in Tab. 3 and 4.

Table 3

Coverage and independence tests, S&P500

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	140	43	69	25	51	22
250	105	44	57	25	45	22
500	105	41	61	25	42	20
1000	80	31	48	22	27	15

Table 4

Coverage and independence tests, BAC

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%	Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	136	46	74	27	48	12	100	1.36460684	0.01414058	2.57964038	0.61885905	0.19177997	0.12146825
250	127	46	73	24	44	11	250	1.9476527	5.8437422	0.9803172	1.2730132	0.8568735	0.1089113
500	119	50	72	31	49	22	500	4.5024847	4.0761754	3.5401206	6.3371277	0.3974163	5.1948436
1000	98	39	52	27	36	22	1000	3.0779306	2.7028451	4.1526883	2.7729393	0.0614779	0.9348360

As can be seen from the results for the first dataset, the VaR at the low confidence level only satisfies the tests at the medium 250 and short 100 data horizons used for parameter estimation. CVaR and EVaR perform well for the high confidence level in a medium window of size 500 and a large window of size 1000. EVaR also performed well for a low confidence level for a medium window of 500. For the second dataset, the situation is almost identical, except that for the large window of size 1000, only EVaR performed well. Moreover, EVaR again proves the result that it is the upper bound of the two previous measures by Chernow's inequality. As can be seen from the results, the independence test failed CVaR and VaR for a low confidence level for a less volatile asset. Combining the results of the tests, it can be determined that the normal model can be applied to these historical data only for a window of 500, using EVaR for both data sets and CVaR for the second set.

When building a model, it is worth looking at how the distribution of the model matches the actual distribution of the data. This is called a goodness-of-fit test. The null hypothesis here is obviously that the predicted distribution is identical to the true distribution. Conventional statistical analysis cannot be applied here because the distribution changes every day. This can be circumvented as follows. Let $U_t = \Phi_{t-1|t}(L_t)$ be the probability of loss of L_t at time t according to the distribution that we assume for time t at time $t-1$. Then our hypothesis becomes $H_0 : U_t \sim Unif(0, 1)$. There is a more extended test by Berkowitz (Holton, 2003), which consists of testing compliance with the standard normal distribution of $N'_j = \Phi(U_t)$. The results are shown in Fig. 2.

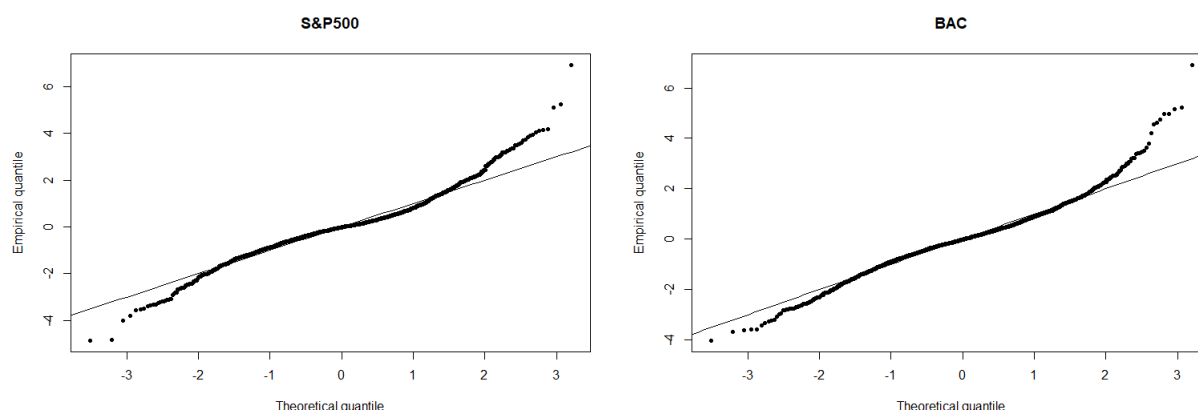


Fig. 2. Distribution test for Normal model

Consider the following model with the basic assumption: $r(t) \sim L(\mu, b)$ – that is, the return has a Laplace distribution. This distribution has heavier tails than the normal distribution. Then, for this model, the corresponding risk measures will be calculated as follows:

$$VaR_{\alpha}(X) = \mu - b \log(2 - 2\alpha),$$

$$CVaR_{\alpha}(X) = \mu + \lambda(1 - \log(2 - 2\alpha)) = \lambda + VaR_{\alpha}(X).$$

Let us use the result from the paper (Mishura et al., 2024):

$$EVaR_{\alpha}(X) = \mu - bW_{-1}(-2e^{-2}(1 - \alpha))\sqrt{1 + \frac{2}{W_{-1}(-2e^{-2}(1 - \alpha))}}.$$

The estimates of the method of moments for the Laplace distribution are as follows:

$$\hat{\mu} = med(\{X_i\}),$$

$$\hat{b} = \frac{1}{n} \sum_{i=1}^T |X_i - \hat{\mu}|.$$

We get the following result in Fig. 3.

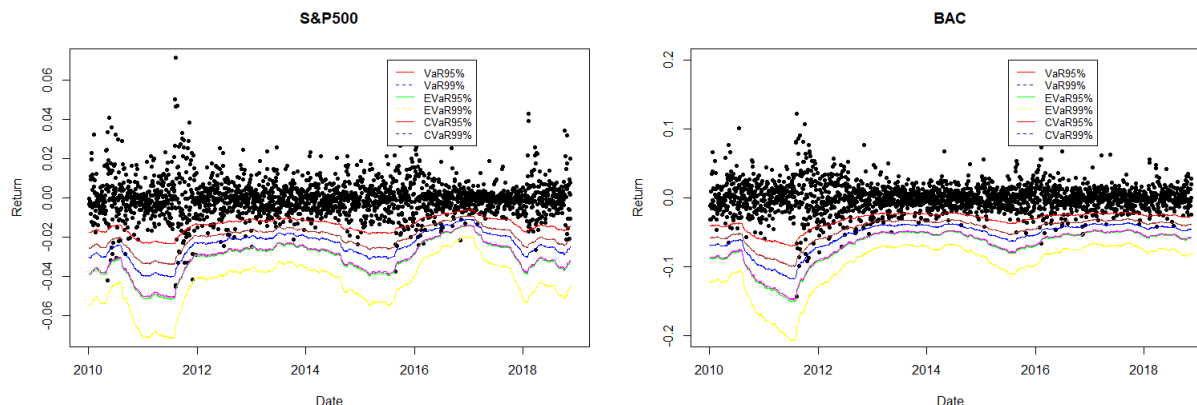


Fig. 3. Laplace model

All risk measures have become smoother and there are fewer jumps. The distribution test has become much better than before. We can see this in Fig. 4.

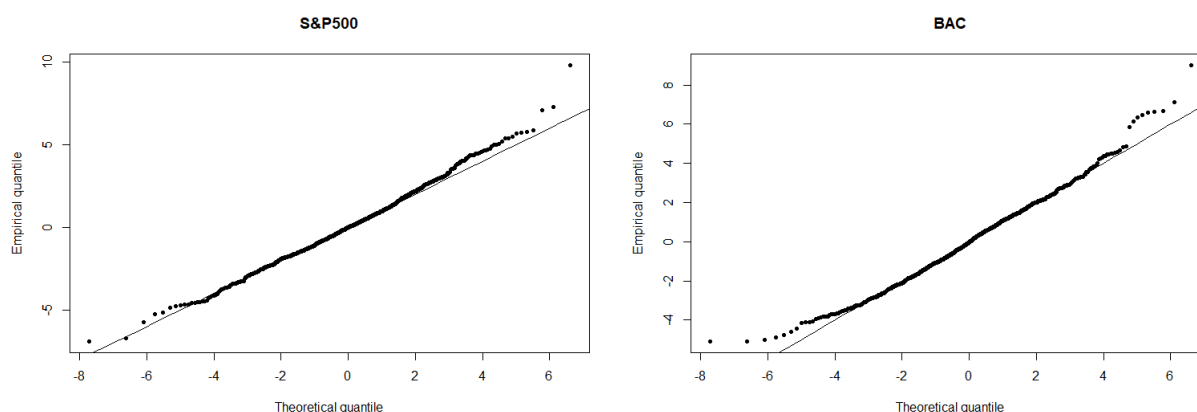


Fig. 4. Distribution test for Laplace model

Since our theoretical distribution has changed, the confidence intervals for the coverage test have also changed. The new confidence intervals are shown in the Tab. 5 for *S&P500* and *BAC*, respectively.

Table 5

Confidence intervals of the coverage test for the Laplace model CVaR/EVaR, *S&P500* and *BAC*, respectively

Window size/Measure	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	[42,70]	[4,21]	[4,17]	[0,5]
250	[37,64]	[3,19]	[4,15]	[0,5]
500	[31,57]	[2,17]	[3,14]	[0,5]
1000	[22,43]	[1,14]	[2,11]	[0,4]

Having performed the coverage and independence tests for this model, we obtained the following results, which are summarized in the following Tab. 6 and 7.

Table 6

Coverage and independence tests for the Laplace model, *S&P500*

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	6.76286394 58	0.43301158 87	4.25351324 80	0.16547162 59	1.82392322 35	0.00083963 06
250	6.09850707 3	1.76344426 9	5.99758341 4	3.16423247 5	1.62628760 2	0.00807175 1
500	7.31379254 7	13.3244410 00	15.7601586 15	2.45033873 0	0.01825730 8	0.00909091 3
1000	0.93885240 8	8.16754568 1	6.70819398 0	3.33485255 0	0.04613822 0	0.03382956 5

Table 7

Coverage and independence tests for the Laplace model, BAC

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	15.39163481 2	5.52352225 3	1.80041703 6	3.88575107 4	0.37385948 3	0.00335993 3
250	6.803391148 0	1.14908924 5	0.65721851 5	0.00358583 6	1.76344426 9	4.92135323 9
500	5.242324821 6	3.46675071 9	4.07617544 5	3.88165991 5	0.38611208 6	0.00909091 3
1000	8.50802462 1	5.56215684 1	4.78312724 1	15.2586275 1	1.66610288 1	0.0121621 1

For the first dataset, EVaR and CVaR for the high confidence level passed the tests for all windows. For the second, more volatile dataset, EVaR and CVaR for all levels worked well for all windows except the large one, while VaR worked well only for the high level. In the large window, as in the previous model, almost no one did well. This is most likely due to too much smoothing of the parameters of the theoretical distribution.

As we have already mentioned, the data distribution changes every day, so it is also worth considering models when the variance also changes over time. The generalized autoregressive conditional heteroscedasticity model (GARCH) will help us with this. Recently, models of this type have been used very often for EVaR. Such models are well suited for volatile markets(Ahmed, Soleymani, Ullah, & Hasan, 2021). The first model of this class is GARCH(1,1) with a mean. It has the following form:

$$r(t) = \mu + z_t = \mu + \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \omega + \alpha_1 z_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

where ϵ_t is the error, which can have different distributions. For this model, we first assume that $\epsilon_t \sim L(0, 1)$. In other words, we actually have the same Laplace model, but with a completely different approach to estimating the variance. The formulas remain the same as in the previous model, but we calculate the variance as a forecast of the previous values. The result is shown in the Fig. 5.

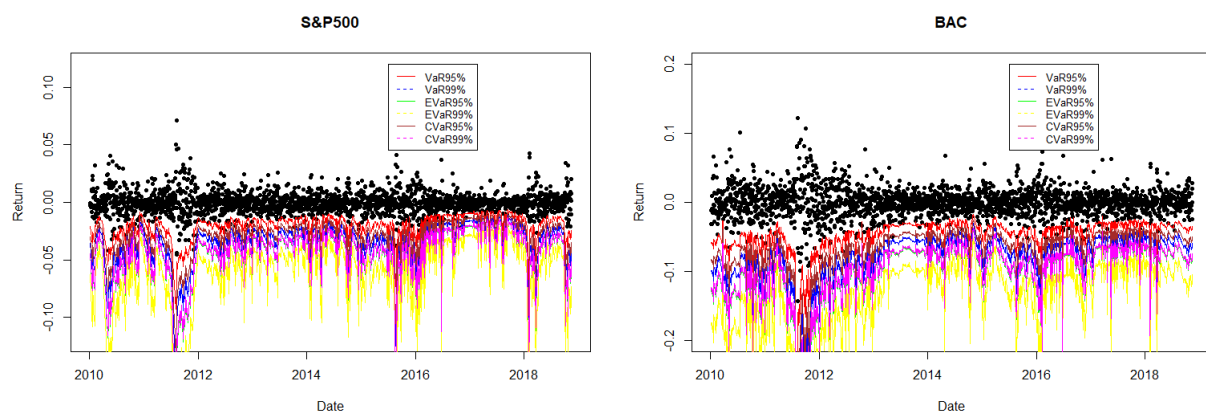


Fig. 5. GARCH(1,1) - Laplace model

The distribution test has become worse than before. This is shown in Fig. 6.

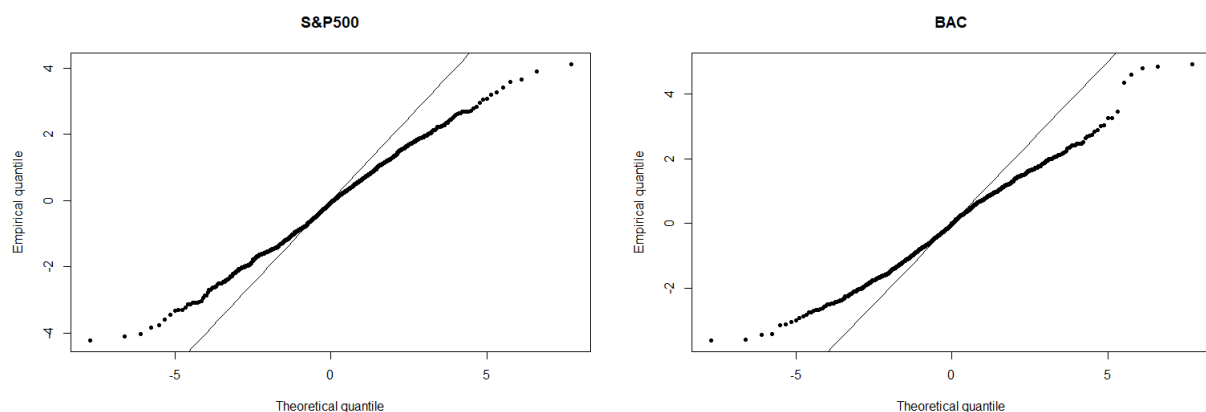


Fig. 6. Distribution test for GARCH(1,1)-Laplace model

As we can see, all the measures start to behave more conservatively, they are more sensitive to fluctuations. Since our theoretical distribution has changed, the confidence intervals for the coverage test have also changed. The new confidence intervals are shown in Tab. 8 for *S&P500* and *BAC*, respectively.

Table 8

Confidence intervals of the coverage test for the GARCH model with Laplace distribution for errors, CVaR/EVaR, S&P500 and BAC, respectively

Window size/Measure	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	[39,67]	[3,20]	[4,16]	[0,5]
250	[36,63]	[3,19]	[4,15]	[0,5]
500	[31,56]	[2,17]	[3,14]	[0,5]
1000	[21,43]	[1,14]	[2,11]	[0,4]

The coverage and independence tests yielded the following results, which are summarized in the Tab. 9 and 10.

Table 9

Coverage and independence tests for GARCH model with mean and Laplace distribution for errors, S&P500

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	43	3	9	0	0	0
250	39	2	6	0	0	0
500	30	1	3	0	0	0
1000	20	0	1	0	0	0

Table 10

Coverage and independence tests for GARCH model with mean and Laplace distribution for errors, BAC

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	38	4	5	0	0	0
250	30	5	6	1	0	0
500	26	5	7	0	0	0
1000	18	3	4	0	0	0

As can be seen from the tables, CVaR and EVaR have almost no excess. All measures showed fewer outliers than the lower bound of the confidence interval. We can say that the models do not work very well, although they protect against possible unlikely shocks. The results of the independence test can be disregarded because the number of outliers is too small. Formally, only the EVaR for high levels passes the coverage test, but since it yields 0 outliers, we should not rely entirely on these tests. It is obvious from the data that the model greatly overestimates the tails of the distribution. Therefore, the only argument for accepting the EVaR at a high confidence level is the need to have a very conservative measure, or the expectation of certain possible significant shocks. In general, it is better not to use this model for available data. Now let us consider the GARCH(1,1) model with mean and normal error distributions.

$$r(t) = \mu + z_t = \mu + \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \omega + \alpha_1 z_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

The result is shown in Fig. 7.

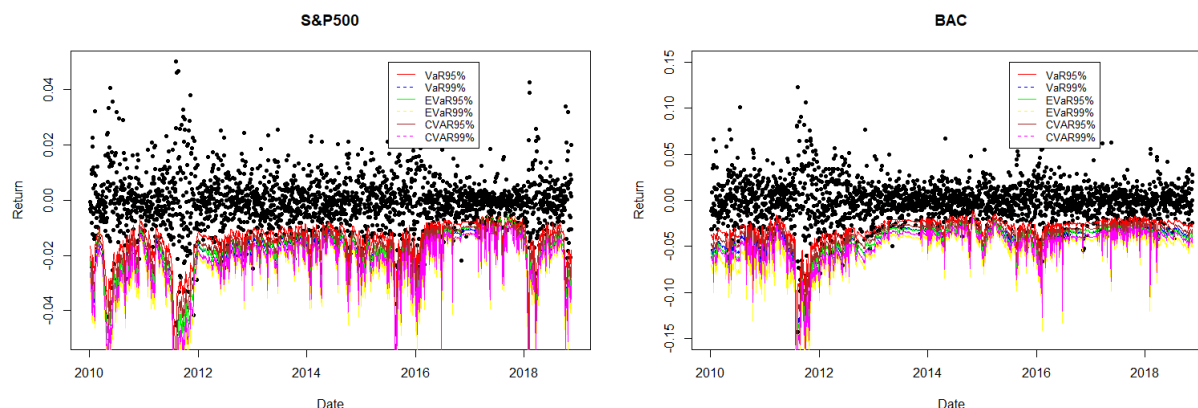


Fig. 7. GARCH(1,1) - Normal model

All risk measures have become smoother and there are fewer jumps. The distribution test has become much better than before. This we can see in Fig. 8.

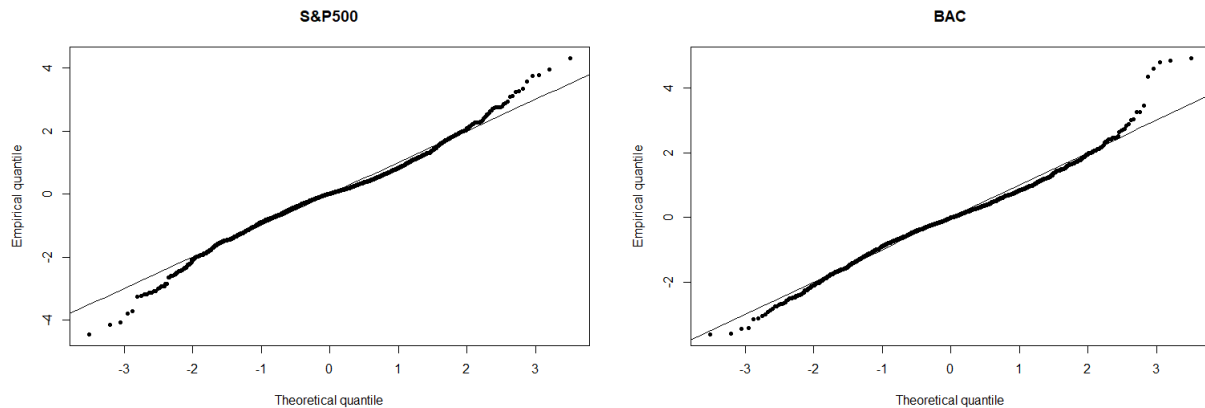


Fig. 8. Distribution test for GARCH(1,1) - Normal model

As we can see from the graphs, this model works quite well on our data, although it is more sensitive to disturbances, it fits the real data better. For this model we construct same Tab. 11, 12 and 13 with methodology as before.

Table 11

Confidence intervals of the coverage test for the GARCH model with mean and normal error distributions, CVaR/EVaR, S&P500 and BAC, respectively

Window size/Measure	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	[34,60]	[2,18]	[22,44]	[1,14]
250	[31,57]	[2,17]	[19,39]	[1,12]
500	[27,51]	[2,16]	[16,35]	[0,11]
1000	[19,40]	[1,13]	[10,26]	[0,9]

Table 12

Coverage and independence tests for the GARCH model with mean and normal error distributions, S&P500

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	98	39	61	22	49	20
250	96	32	46	20	33	14
500	83	27	43	15	29	12
1000	62	20	36	13	22	8

Table 13

Coverage and independence tests for the GARCH model with mean and normal error distributions, BAC

Window size/Measure	VaR95%	VaR99%	CVaR95%	CVaR99%	EVaR95%	EVaR99%
100	113	39	60	28	42	15
250	110	33	52	18	30	10
500	92	26	41	18	24	11
1000	69	23	34	12	20	6

The result is quite different from the normal distribution model. VaR works well for medium- and long-observation horizons for both data sets and for both confidence levels. CVaR and EVaR work well only for a large window size. Almost all the measures on this model pass the independence test. Although the correlation in the distribution test is only 0.992 for 250 values for prediction and the overall test fails, this model is quite good for this data set.

In addition, the EDaR risk measure was considered. This measure of risk is quite often compared to EVaR. However, because for the process $DD(X)$ does not have an elementary distribution for which it is possible to obtain the exact EVaR formula, we managed to obtain estimates only numerically for the sample. The calculation time is almost identical to that of EVaR if we use the nonparametric approach, which is much longer for the parametric approach we use. Actually in the Tab. 14 of time spent the advantage of parametric approach over non-parametric approach is clearly visible. In addition, this measure proved to be so conservative and distribution-insensitive that none of the tests presented can characterize this measure and provide results for backtesting. As can be seen in Fig. 9, the EDaR estimate strongly exceeds all values, making it impossible to perform an outlier analysis.

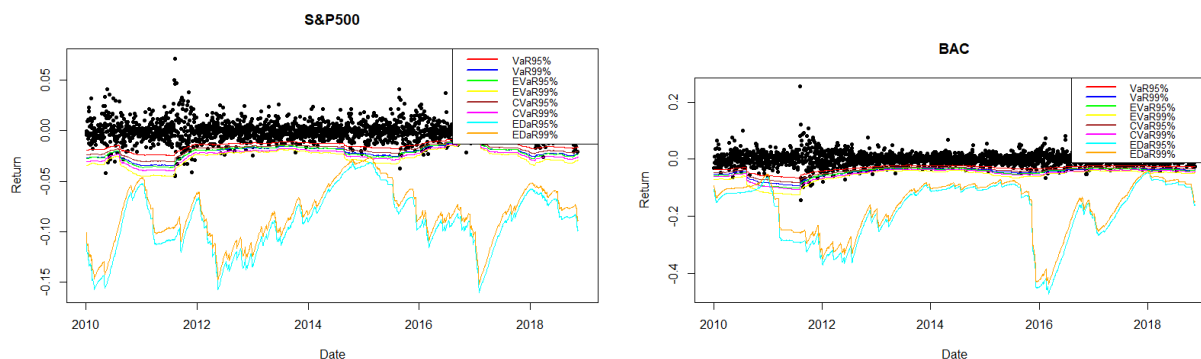


Fig. 9. Normal model in comparison to EDaR

Table 14

Time consumption in seconds to calculate a one-day risk value over the entire data interval

Window size	Parametric approach			Nonparametric approach			
	VaR	CVaR	EVaR	VaR	CVaR	EVaR	EDaR
100	0.0254	0.0271	0.0248	0.1712	0.4212	15.718	15.52
250	0.0358	0.0298	0.0295	0.1637	0.7374	21.174	22.461
500	0.0292	0.0352	0.0351	0.1607	1.0237	32.288	34.382
1000	0.0267	0.0258	0.0238	0.1159	1.4567	34.909	40.201

In summarizing the results of the models considered, the best fit to the data was shown by the Laplace model for EVaR at any confidence level and CVaR at a high confidence level. The GARCH model with mean and normal error distributions for VARs of all levels and medium and large windows, and the CVaR for the low confidence level, proved to be the second best. If a significant increase in volatility and unexpected shocks is expected, the EVaR model for GARCH with a Laplace error distribution for high confidence levels can be used.

To summarize the models reviewed, we can say that EVaR performs quite well on data with high volatility and heavy tails, when VaR does not work. EVaR performed on a par with CVaR, which it outperformed on data with heavy tails that were not taken into account by the theoretical model. In general, we confirmed that EVaR is an upper bound for VaR and CVaR. CVaR and EVaR were much more likely to pass the loss independence test, even if the model was not fully fit to the real data. However, if the model even slightly predicted heavier tails than the real data, the EVaR significantly exceeded the possible risk. All risk measures significantly and correctly increased their risk estimates when data volatility increased. This can be seen in the similarity of the results for the two data sets. To summarize the results, we can say that EVaR models perform well with these tests if the model underestimates the severity of the tails. But in case of overestimation, it is better to use other tests.

Discussion and conclusions

In this paper, we consider the Entropic Value-at-Risk (EVaR) risk measure on real data testing. Recently, this measure has been increasingly considered in scientific research and applied to solve practical problems. There are not many models for EVaR described in the current scientific literature. Most often, only the standard normal distribution model is considered. The lack of great results in this area is due to the complexity of the calculations for different models.

The result of this work gives the possibility of further studying the behavior of EVaR for the already defined models, as well as expanding the models for which EVaR can be analytically solved. In the future, one can either try to analytically approximate the EVaR value or develop a solution for working with mixtures of distributions.

The testing of some models on real data was performed using coverage tests, normality tests, and loss independence tests. These tests are standard in risk assessment practice and are widely used in model back-testing. As real data, we used publicly available data on the value of the *S&P500* index and the value of Bank of America Corp. shares for the period from 2009 to 2019. Each model was tested on a standard VaR task: to build a daily risk estimate for different available data volumes, for building estimates and risk models. The models for VaR, CVaR, and EVaR were compared for different data, different models, and different model parameters. The most commonly used GARCH models for financial forecasting were built and tested with different error distributions and the presence of a mean. The optimal models were selected from those proposed for the historical data. EVaR proved to be very sensitive to the heavy tails included in the model. It turned out that the GARCH models with Laplace errors, which are widely used today in practice, overestimated the real risk. The basic relationship between VaR, CVaR, and EVaR was confirmed in practice. For almost all models, the EVaR behavior was similar to CVaR and significantly outperformed the VaR behaviour. The comparison of EVaR and EDaR was also examined, and the difficulty of calculating the values was compared by comparing the time consumed to calculate the values.

From all of the above, we can conclude that the EVaR risk measure considered is relevant and has great promise. In our opinion, the study of EVaR is very promising, as this risk measure provides a better assessment of losses with heavy tails. Although, of course, this measure has disadvantages related to the complexity of numerical calculations, analysis of

the impact of model parameters, and the use of complex models. In this view, there are not many methods for practical application of this measure. Some of them have been reviewed and elaborated in this paper. Given the superiority of this measure over others, we can conclude that further study of this measure and the development of methods for its effective application to solve various financial problems are needed.

Authors' contribution: Volodymyr Zubchenko – conceptualization, methodology, analysis of sources, preparation of literature review and theoretical foundations of research; Petro Zelenko – empirical data collection and validation, software, empirical research.

References

- Ahmadi-Javid, A. (2012). Entropic value-at-risk: A new coherent risk measure. *Journal of Optimization Theory and Applications*, 155, 1105-1123. <https://doi.org/10.1007/s10957-011-9968-2>
- Ahmadi-Javid, A., & Fallah-Tafti, M. (2019). Portfolio optimization with entropic value-at-risk. *European Journal of Operational Research*, 279(1), 225-241. <https://doi.org/10.1016/j.ejor.2019.02.007>
- Ahmadi-Javid, A., & Pichler, A. (2017). An analytical study of norms and banach spaces induced by the entropic value-at-risk. *Mathematics and Financial Economics*, 11, 1-24. <https://doi.org/10.1007/s11579-017-0197-9>
- Ahmed, D., Soleymani, F., Ullah, M. Z., & Hasan, H. (2021). Managing the risk based on entropic value-at-risk under a normal-rayleigh distribution. *Applied Mathematics and Computation*, 402, 126-129. <https://doi.org/10.1016/j.amc.2021.126129>
- Cajas, D. (2021). Entropic portfolio optimization: A disciplined convex programming framework. *SSRN Electronic Journal*, 1-22. <https://doi.org/10.2139/ssrn.3792520>
- Delbaen, F. (2019). Remark on the paper "Entropic value-at-risk: A new coherent risk measure" by Amir Ahmadi-Javid, J. Optim. Theory Appl., 155(3) (2001) 1105-1123". In P. Barrieu (Ed.), *Risk and stochastic* (p. 151-158). World Scientific. https://doi.org/10.1142/9781786341952_0009
- Föllmer, H., & Schied, A. (2016). *Stochastic finance, an introduction in discrete time*. De Gruyter. <https://doi.org/10.1515/9783110463453>
- Holton, G. (2003). *Value-at-risk: Theory and practice*. Academic Press.
- Mishura, Y., Raichenko, K., Zelenko, P., & Zubchenko, V. (2024). Properties of the entropic risk measure evar in relation to selected distributions. *Modern Stochastics: Theory and Applications*, 11(4), 373-394. <https://doi.org/10.15559/24-VMSTA255>
- Pichler, A. (2017). A quantitative comparison of risk measures. *Annals of Operations Research*, 254, 251-275. <https://doi.org/10.1007/s10479-017-2397-3>

Отримано редакцією журналу / Received: 03.01.25

Прорецензовано / Revised: 04.03.25

Схвалено до друку / Accepted: 11.03.25

Володимир ЗУБЧЕНКО, канд. фіз.-мат. наук, доц.

ORCID ID: 0000-0002-1513-9326

e-mail: volodymyr.zubchenko@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Петро ЗЕЛЕНЬКО, магістр

ORCID ID: 0009-0005-9671-0465

e-mail: petro.zelenko@uni-ulm.de

Ульмський університет, Ульм, Німеччина

ТЕСТУВАННЯ ЕНТРОПІЙНОЇ МІРИ РИЗИКУ EVAR НА РЕАЛЬНИХ ДАНИХ

Дослідження мір ризику є надзвичайно важливим для розуміння та управління невизначеностями, що властиві різним фінансовим та економічним активностям. Міри ризику надають кількісну оцінку потенційних втрат, що дає змогу бізнесу та фінансовим установам ухвалювати обґрунтовані рішення та зменшувати негативні наслідки. Основою пропонованого дослідження є дедалі більша складність і взаємопов'язаність фінансових ринків, що посилює вплив ризиків і потребує їхнього управління. Традиційні міри, такі як Value at Risk (VaR), широко використовуються, але їхні обмеження, особливо під час фінансових криз, підкреслюють необхідність більш комплексних підходів, таких як Expected Shortfall (ES) та Entropic Value at Risk (EVAR). Актуальність дослідження мір ризику підтверджується фінансовими потрясіннями, що спостерігалися під час глобальних криз, таких як фінансова криза 2008 р. та пандемія COVID-19. Ці події виявили вразливість наявних практик управління ризиками та підкреслили потребу в стійкіших і точніших інструментах оцінювання ризиків. Метою нашого дослідження є оцінка та порівняння різних мір ризику, надання уявлення про їхні переваги та недоліки, а також розроблення ефективніших стратегій управління ризиками. Проаналізовано ефективність ентропійної міри ризику Value-at-Risk (EVAR) на реальних фінансових даних, та виконано її порівняння з традиційними ризиковими мірами, такими як Value-at-Risk (VaR) та Conditional Value-at-Risk (CVAR). Розглянуто історичні дані щодо дохідності індексу S&P 500 та акцій Bank of America Corp. (2009–2019) та застосовано різні параметричні моделі, включно з моделями на основі нормального розподілу і розподілу Лапласа. Також використано стандартні методи бек-тестування (тести покриття та незалежності) для оцінювання емпіричної адекватності цих мір. Проведений аналіз підкреслює математичні властивості EVAR і здатність цієї міри ефективніше виявляти важкі хвости розподілу порівняно з VaR, забезпечуючи при цьому когерентні та потенційно надійніші оцінки в умовах волатильних ринків. Перспективи майбутніх досліджень передбачають розроблення нових мір ризику, які можуть краще враховувати системні ризики на фінансових ринках. Крім того, інтеграція машинного навчання та штучного інтелекту з традиційними мірами ризику відкриває можливості для підвищення точності та надійності прогнозів. Ці дослідження є важливими для побудови стійкіших фінансових систем і забезпечення економічної стабільності.

Ключові слова: ентропійна міра ризику, EVAR, тестування на реальних даних, нормальний розподіл, розподіл Лапласа.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.21, 51-75

DOI: <https://doi.org/10.17721/1812-5409.2025/1.5>

Volodymyr LEVCHENKO, PhD Student

ORCID ID: 0009-0007-9681-8753

e-mail: vlad.levchenko@protonmail.com

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Rostyslav YAMNENKO, DSc (Phys. & Math.)

ORCID ID: 0000-0002-9612-7959

e-mail: rostyslav.yamnenko@knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

ESTIMATION OF RUIN PROBABILITY FOR RANDOM SUM OF NEGATIVE BINOMIALLY DISTRIBUTED NUMBER OF SUB-GAUSSIAN CLAIMS

This paper investigates the properties of a classical risk process in which the number of insurance claims follows a negative binomial distribution, and the claim sizes belong to the class of strictly sub-Gaussian random variables of a generalized type. This class is broad and includes both standard sub-Gaussian and Gaussian random variables. The core assumption is that each random variable in this class satisfies a specific inequality that constrains its exponential growth, allowing for control over the behavior of its probabilistic tail. It is also assumed that the insurance premium collected over time is described by a continuous, monotone increasing function whose increments are bounded by another continuous, increasing function. This condition enables flexible modeling of premium growth over time.

The aim of the paper is to obtain estimates for the probability of ruin in the described risk process model. To achieve this, the approach is based on the method of metric entropy, which allows for the consideration of both the complex nature of the claim count distribution and the properties of strictly sub-Gaussian claim sizes. This approach makes it possible to derive estimates even in cases where the exact distribution of the claims is unknown, provided that they exhibit sub-Gaussian behavior.

Key words: negative binomial distribution, sub-Gaussian random variable, ruin probability, claim distribution, metric entropy.

AMS 2020 classification: 60G50, 91G05.

Introduction

Negative binomial and Poisson distributions are popular models for describing the number of claims, see e.g. (Klugman, Panjer, & Willmot, 2012; Boucher, Denuit, & Guillén, 2008; Ismail & Zamani, 2013; Shi & Valdez, 2014; David & Jemna, 2015). The Poisson distribution assumes that the variance and mean of the claims should be equal, which is often not the case for real data. The negative binomial distribution doesn't have this limitation. So, let's develop a risk model using the negative binomial number of claims on the basis of the classic Poisson model.

$$X(t) = u + ct - \sum_{i=1}^{N_t} Y_i, \quad t \geq 0, \quad (1)$$

where u is the initial capital of an insurance company, $c > 0$ is the intensity of receiving insurance premiums and

- (I) Y_i are independent and identically distributed random variables representing the values of insurance claims, $E Y_i = m_1$, $E Y_i^2 = m_2$.
- (II) $Y_i \in \text{Sub}_\varphi(\Omega)$, i.e. Y_i belong to the space of φ -sub-Gaussian random variables with norm $\tau = \tau_\varphi(Y_i) < \infty$.
- (III) N_t is a process that describes the number of insurance claims that happened between 0 and t according to negative binomial distribution, i.e. $P(N_t = k) = C_{k+r-1}^k (1-\theta)^k \theta^{r-t}$, where $k \in \{0, 1, 2, \dots\}$ and r acts as an analog of the claims intensity, such that the expected number of claims during the interval (t, s) is proportional to its length, $E(N(s) - N(t)) = \frac{rs(1-\theta)}{\theta} - \frac{rt(1-\theta)}{\theta} = \frac{r(s-t)(1-\theta)}{\theta}$. In other words, $N(s) - N(t) \sim \text{NB}(r(s-t), \theta)$.
- (IV) N_t and the sequence $(Y_1, Y_2, Y_3, \dots)$ are mutually independent. And let's also assume that the number of claims on independent time intervals are independent.

The risk process described by (1) can also be expressed as:

$$X(t) = u + Ct - \sum_{i=1}^{N_t} (Y_i + \alpha) = u + (C - \alpha\mu)t - \sum_{i=1}^{N_t} Y_i, \quad t \geq 0, \quad (2)$$

where α is the mean value of an insurance payment.

This work focuses on the cumulative total of payments:

$$S(t) = \sum_{i=1}^{N_t} Y_i, \quad (3)$$

under the conditions outlined above.

Properties of similar sums with φ -sub-Gaussian claim sizes with Poisson or binomial number of claims are studied in papers (Yamnenko & Vasylyk, 2007; Saienko & Yamnenko, 2013) and (Yamnenko & Lamin, 2022).

The characteristics of sub-Gaussian, φ -sub-Gaussian and strictly φ -sub-Gaussian random variables and processes can be found in the books (Buldygin & Kozachenko, 2000; Kozachenko, Vasylyk, & Yamnenko, 2008; Vershynin, 2018) and other studies, e.g. (Kozachenko & Kovalchuk, 1998; Giuliano Antonini, Kozachenko, & Nikitina, 2003; Kozachenko, Yamnenko, & Vasylyk, 2005; Hopkalo, Kozachenko, Vasylyk, & Sakhno, 2020).

1. φ -sub-Gaussian random values and processes

This section contains the necessary definitions for φ -sub-Gaussian random variables and processes.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a standard probability space and let $(\mathbf{T}, \rho)$ be a pseudometric (metric) compact space equipped by pseudometric (metric) ρ .

Definition 1. (Buldygin & Kozachenko, 2000) A continuous even convex function φ is said to be an Orlicz N -function if it is strictly increasing for $x > 0$, $\varphi(0) = 0$ and

$$\frac{\varphi(x)}{x} \rightarrow 0 \text{ as } x \rightarrow 0 \quad \text{and} \quad \frac{\varphi(x)}{x} \rightarrow \infty \text{ as } x \rightarrow \infty.$$

Condition Q. (Guiliano, Kozachenko, & Nikitina, 2023)

We say that an N -function φ satisfies condition Q if

$$\liminf_{x \rightarrow 0} \frac{\varphi(x)}{x^2} = \alpha > 0. \quad (4)$$

It is permitted α to be infinite.

Definition 2. (Buldygin & Kozachenko, 2000) Let φ be an Orlicz N -function satisfying condition Q. The random variable ξ belongs to the space $\text{Sub}_\varphi(\Omega)$ (a space of φ -sub-Gaussian random variables), if it is centered, i.e. $E\xi = 0$, the moment generating function $E \exp\{\lambda\xi\}$ exists for all $\lambda \in \mathbb{R}$ and there exists a positive constant a such that the following inequality

$$E \exp(\lambda\xi) \leq \exp(\varphi(a\lambda)) \quad (5)$$

holds for all $\lambda \in \mathbb{R}$.

Theorem 1. (Buldygin & Kozachenko, 2000) The space $\text{Sub}_\varphi(\Omega)$ is a Banach space with respect to the norm

$$\tau_\varphi(\xi) = \inf\{a \geq 0 : E \exp(\lambda\xi) \leq \exp(\varphi(a\lambda)), \lambda \in \mathbb{R}\}$$

and the inequality

$$E \exp(\lambda\xi) \leq \exp(\varphi(\lambda\tau_\varphi(\xi))), \quad (6)$$

holds for all $\lambda \in \mathbb{R}$. Moreover, for all $r > 0$ there exists constant $c_r > 0$ such that $(E \xi^r)^{1/r} \leq c_r \tau_\varphi(\xi)$.

Definition 3. (Buldygin & Kozachenko, 2000) Metric entropy with regard to pseudometric (metric) ρ , or just metric entropy is a function

$$H_{(T,\rho)}(u) = H_T(u) = H(u) = \begin{cases} \log N_{(T,\rho)}(u), & \text{if } N_{(T,\rho)}(u) < +\infty \\ +\infty, & \text{if } N_{(T,\rho)}(u) = +\infty \end{cases},$$

where $N_{(T,\rho)}(u) = N_T(u) = N(u)$ denotes the least number of closed ρ balls with radius u covering the space (T, ρ) .

Example. If $\mathbf{T}$ is an interval $[a, b]$ and ρ is the Euclidean distance then

$$\ln \left(\max \left\{ \frac{b-a}{2u}, 1 \right\} \right) \leq H(u) \leq \ln \left(\frac{b-a}{2u} + 1 \right).$$

Definition 4. The random process $X = \{X(t), t \in T\}$ is called φ -sub-Gaussian process if for all $t \in T$ $X(t) \in \text{Sub}_\varphi(\Omega)$.

Definition 5. (Kozachenko & Kovalchuk, 1998) A family of random variables Δ from the space $\text{Sub}_\varphi(\Omega)$ is called *strictly* $\text{Sub}_\varphi(\Omega)$, if there exists a constant $C_\Delta > 0$ such that for arbitrary finite set $I : \xi_i \in \Delta, i \in I$, and for any $\lambda_i \in \mathbb{R}$ the following inequality takes place

$$\tau_\varphi \left(\sum_{i \in I} \lambda_i \xi_i \right) \leq C_\Delta \left(E \left(\sum_{i \in I} \lambda_i \xi_i \right)^2 \right)^{1/2}. \quad (7)$$

If Δ is a family of strictly $\text{Sub}_\varphi(\Omega)$ random variables, then the linear closure $\overline{\Delta}$ of the family Δ in the space $L_2(\Omega)$ is also strictly $\text{Sub}_\varphi(\Omega)$ family of random variables. Linearly closed families of strictly $\text{Sub}_\varphi(\Omega)$ random variables form a space of *strictly* φ -sub-Gaussian random variables. This space is denoted by $\text{SSub}_\varphi(\Omega)$.

When $\varphi(x) = x^2/2$ the space $\text{SSub}_\varphi(\Omega)$ is called the space of *strictly sub-Gaussian* random variables and is denoted as $\text{SSub}(\Omega)$. The space of jointly Gaussian random variables belongs to the space $\text{SSub}(\Omega)$ and $\tau^2(\xi) = E \xi^2$, i.e. $C_\Delta = 1$.

Definition 6. (Buldygin & Kozachenko, 2000) A random process $X = \{X(t), t \in T\}$ is a *strictly* φ -sub-Gaussian process if the corresponding family of random variables belongs to the space $\text{SSub}_\varphi(\Omega)$. The defining constant of this family is called the defining constant of the process X and is denoted by C_X .

Condition Σ . Suppose there exists such a continuous monotonically increasing function $\sigma = \{\sigma(h), h > 0\}$, that $\sigma(h) \rightarrow 0$, as $h \rightarrow 0$, and the following inequality for increments of the process is true

$$\sup_{\rho(t,s) \leq h} \tau_{\varphi}(Y(t) - Y(s)) \leq \sigma(h). \quad (8)$$

If a process $X(t)$ is continuous in norm $\tau_{\varphi}(\cdot)$ then the function $\sigma(h) = \sup_{\rho(t,s) \leq h} \tau_{\varphi}(X(t) - X(s))$ satisfies condition Σ .

2. Negative binomial random sums

In this section, we study the properties of sums whose number of terms follows a non-negative binomial distribution, as they appear in model (2). The following lemmas provide results concerning the moments of such sums and estimates for their moment-generating functions.

Lemma 1. Let $\{S(t) = \sum_{i=1}^{N_t} Y_i(t), t \geq 0\}$ be a random process satisfying the conditions (I)–(IV). Then for all $t \geq 0$, $0 \leq t_1 \leq t_2$, and $\lambda \in \mathbb{R}$ the following expressions hold true

$$\begin{aligned} \mathbb{E} S(t) &= \frac{(1-\theta)rt}{\theta} m_1, \quad \text{Var } S(t) = \frac{(1-\theta)rt}{\theta} (m_2 - m_1^2) + m_2 \frac{(1-\theta)rt}{\theta^2}, \\ \mathbb{E}(S(t_2) - S(t_1)) &= \frac{(1-\theta)r(t_2 - t_1)}{\theta} m_1, \\ \text{Var}(S(t_2) - S(t_1)) &= \frac{(1-\theta)r(t_2 - t_1)}{\theta} (m_2 - m_1^2) + m_2 \frac{(1-\theta)r(t_2 - t_1)}{\theta^2} \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \exp\{\lambda S(t)\} &\leq \left(\frac{\theta}{1 - (1-\theta) \exp\{\varphi(\lambda\tau)\}} \right)^{rt}, \\ \mathbb{E} \exp\{\lambda(S(t_2) - S(t_1))\} &\leq \left(\frac{\theta}{1 - (1-\theta) \exp\{\varphi(\lambda\tau)\}} \right)^{r(t_2 - t_1)}. \end{aligned}$$

Proof. Recall that for a random variable with a negative binomial distribution $\xi \sim \text{NB}(r, p)$, we have its moment-generating function is given by $M_{\xi}(\lambda) = \mathbb{E} \exp\{\lambda \xi\} = \left(\frac{p}{1 - (1-p)e^{\lambda}} \right)^r$. Considering that $S(t)$ has the compound negative binomial distribution, from (6) we have the following:

$$\begin{aligned} \mathbb{E} \exp\{\lambda S(t)\} &= \mathbb{E} \exp\left\{ \lambda \sum_{i=1}^{N_t} Y_i \right\} = \mathbb{E} \left[\mathbb{E} \left\{ \exp\left\{ \sum_{i=1}^{N_t} \lambda Y_i \right\} \middle| N_t \right\} \right] = \\ &= \sum_{n=0}^{\infty} \mathbb{E} \exp\left\{ \sum_{i=1}^n \lambda Y_i \right\} \mathbf{P}(N_t = n) = \sum_{n=0}^{\infty} \prod_{i=1}^n \exp\{\lambda Y_i\} \mathbf{P}(N_t = n) = \sum_{n=1}^{\infty} (\mathbb{E} \exp\{\lambda Y_1\})^n \mathbf{P}(N_t = n) \leq \\ &\leq \sum_{n=1}^{\infty} (\exp\{\varphi(\lambda \tau_{\varphi}(Y_1))\})^n \mathbf{P}(N_t = n) \leq \left(\frac{\theta}{1 - (1-\theta) \exp\{\varphi(\lambda\tau)\}} \right)^{rt}. \end{aligned}$$

Since $N_{t_1} \leq N_{t_2}$ for $t_1 \leq t_2$, similarly we get

$$\begin{aligned} \mathbb{E} \exp\{\lambda(S(t_2) - S(t_1))\} &= \mathbb{E} \exp\left\{ \lambda \sum_{i=N_{t_1}+1}^{N_{t_2}} Y_i \right\} = \mathbb{E} \left(\mathbb{E} \exp\left\{ \lambda \sum_{i=N_{t_1}+1}^{N_{t_2}} Y_i \right\} \middle| N_{t_1}, N_{t_2} \right) = \\ &= \sum_{k=0}^{\infty} \sum_{m=k}^{\infty} (\mathbb{E} \exp(\lambda Y_1))^{m-k} \mathbf{P}(N_{t_1} = k) \mathbf{P}(N_{t_2} = m | N_{t_1} = k) = \\ &= \sum_{k=0}^{\infty} \sum_{m=k}^{\infty} (\mathbb{E} \exp(\lambda Y_1))^{m-k} \mathbf{P}(N_{t_1} = k) \mathbf{P}(N_{t_2} - N_{t_1} = m - k) = \\ &= \sum_{k=0}^{\infty} \mathbf{P}(N_{t_1} = k) \sum_{j=0}^{\infty} (\mathbb{E} \exp(\lambda Y_1))^j \mathbf{P}(N_{t_2} - N_{t_1} = j) = \sum_{j=0}^{\infty} (\mathbb{E} \exp(\lambda Y_1))^j \mathbf{P}(N_{t_2} - N_{t_1} = j) \leq \\ &\leq \sum_{j=0}^{\infty} \exp\{\varphi(\lambda \tau_{\varphi}(Y_1))\}^j \mathbf{P}(N_{t_2} - N_{t_1} = j) = \left(\frac{\theta}{1 - (1-\theta) \exp\{\varphi(\lambda\tau)\}} \right)^{r(t_2 - t_1)}. \end{aligned}$$

The first two moments for a compound distribution (see, e.g. (Klugman et al., 2012)) are

$$\mathbb{E} S(t) = \mathbb{E} N_t \mathbb{E} Y_1, \quad \mathbb{E} S^2(t) = \mathbb{E} N_t \text{Var } Y_1 + \mathbb{E} Y_1^2 \mathbb{E} N_t^2, \quad \text{Var } S(t) = \mathbb{E} N_t \text{Var } Y_1 + \mathbb{E} Y_1^2 \text{Var } N_t,$$

therefore

$$\mathbb{E} S(t) = \frac{(1-\theta)rt}{\theta} m_1, \quad \text{Var } S(t) = \frac{(1-\theta)rt}{\theta} (m_2 - m_1^2) + m_2 \frac{(1-\theta)rt}{\theta^2}.$$

and

$$\begin{aligned} E(S(t_2) - S(t_1)) &= E(N_{t_2} - N_{t_1}) E Y_1 = \frac{(1 - \theta)r(t_2 - t_1)}{\theta} m_1, \\ \text{Var}(S(t_2) - S(t_1)) &= \text{Var} S(t_2 - t_1) = \frac{(1 - \theta)r(t_2 - t_1)}{\theta} (m_2 - m_1^2) + m_2 \frac{(1 - \theta)r(t_2 - t_1)}{\theta^2}. \end{aligned}$$

□

We now examine the estimates for the probability that the sums $S(t)$ in (3) exceed a given level $x \geq 0$, where $f(t)$ is a continuous function representing the total insurance premiums collected over time t .

First, we consider the conditions necessary to bound the supremum of the difference $S(t) - f(t)$. Additionally, we derive an estimate for the exponential moment of this supremum.

From Lemma 1 it follows that

$$\begin{aligned} E(S(t_2) - S(t_1))^2 &= \text{Var}(S(t_2) - S(t_1)) + [E(S(t_2) - S(t_1))]^2 = \\ &= \frac{(1 - \theta)r(t_2 - t_1)}{\theta^2} [\theta^2(m_2 - m_1^2) + m_2 + (1 - \theta)r(t_2 - t_1)m_1^2] \leq \\ &\leq \frac{(1 - \theta)r(t_2 - t_1) \max\{t_2 - t_1, 1\}}{\theta^2} [\theta^2(m_2 - m_1^2) + m_2 + (1 - \theta)rm_1^2]. \end{aligned}$$

Then let's choose

$$\sigma(h) = \kappa h^{\frac{1}{2}} \mathbf{1}_{h < 1} + \kappa h \mathbf{1}_{h \geq 1}, \quad (9)$$

as a function $\sigma(h)$ satisfying assumption (8) for the process $S(t)$, where $\kappa = \frac{(1 - \theta)r}{\theta^2} [\theta^2(m_2 - m_1^2) + m_2 + (1 - \theta)rm_1^2]$ and $\mathbf{1}$ is the indicator function.

Let $B \subset \mathbf{T}$ be a compact set. Put $0 < \beta \leq 1$ be such a number that $\beta \leq \sigma\left(\inf_{s \in B} \sup_{t \in B} \rho(t, s)\right)$.

Lemma 2. Let $X(t) = S(t) - f(t)$ with $\{S(t) = \sum_{i=1}^{N_t} Y_i, t \in \mathbf{T}\}$ be a random process satisfying assumptions (I)–(IV). Let's assume also that increments of continuous function $f = \{f(t), t \in \mathbf{T}\}$ are bounded by a monotone increasing function $\delta = \{\delta(s), s > 0\}$: $|f(u) - f(v)| \leq \delta(\rho(u, v))$, and let $\{q_k, k \in \mathbb{N}\}$ be such a sequence that $q_k > 1$, and $\sum_{k=1}^{\infty} \frac{1}{q_k} \leq 1$. Then for all $\lambda > 0$ and $p \in (0, 1)$ the following estimate holds true

$$\begin{aligned} E \exp \left\{ \lambda \sup_{t \in B} X(t) \right\} &\leq \sup_{u \in B} \left[\left(\frac{\theta}{1 - (1 - \theta) \exp\{\varphi(q_1 \lambda \tau)\}} \right)^{ru/q_1} \exp\{-\lambda f(u)\} \right] \times \\ &\times \left[\prod_{k=1}^{\infty} \left(N_B \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right)^{\frac{1}{q_k}} \right] \left[\prod_{k=2}^{\infty} \left(\frac{\theta}{1 - (1 - \theta) \exp\{\varphi(q_k \lambda \tau)\}} \right)^{\frac{r \beta^2 p^{2k}}{\kappa^2 q_k}} \exp \left\{ \lambda \delta \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right\} \right]. \end{aligned} \quad (10)$$

Proof. Put $\varepsilon_k = \sigma^{(-1)}(\beta p^k) = \frac{1}{\kappa^2} \beta^2 p^{2k}$ and denote by V_{ε_k} set of closed balls of radii ε_k , which forms the minimal coverage of the compact B . The number of points in the set V_{ε_k} equals to $N_B(\varepsilon_k)$. From Lemma 1 and Chebyshev's inequality for any $\varepsilon > 0$ we get

$$\mathbf{P} \{ |S(t) - S(s)| > \varepsilon \} \leq \frac{1}{\varepsilon^2} E \left(\sum_{n=1}^{N_t} Y_n - \sum_{k=1}^{N_s} Y_k \right)^2 = \varepsilon^{-2} [\sigma(t - s)]^2.$$

So, the processes $S(t)$ and $X(t) = S(t) - f(t)$ are continuous in probability. Therefore the set $V = \bigcup_{k=1}^{\infty} V_{\varepsilon_k}$ is a separant of X , and with probability one the following equality holds true

$$\sup_{t \in B} X(t) = \sup_{t \in V} X(t). \quad (11)$$

Consider α -mapping of the V into V_{ε_n} (see (Buldygin & Kozachenko, 2000)). Let's apply Chebyshev's inequality. It follows from (6), (9) and lemma 1 that

$$\mathbf{P} \left\{ |S(t) - S(\alpha_n(t))| > p^{\frac{n}{2}} \right\} \leq \frac{E(S(t) - S(\alpha_n(t)))^2}{p^n} = \frac{\sigma^2(|t - \alpha_n(t)|)}{p^n} < \frac{\sigma^2(\varepsilon_n)}{p^n} = \frac{[\sigma(\sigma^{(-1)}(\beta p^n))]^2}{p^n} = \beta^2 p^n.$$

This last inequality means that

$$\sum_{n=1}^{\infty} P \left\{ |S(t) - S(\alpha_n(t))| > p^{\frac{n}{2}} \right\} < \infty.$$

Therefore, it follows from the Borel-Cantelli lemma that $S(t) - S(\alpha_n(t)) \rightarrow 0$ with probability one. The same statement is valid for the process X .

The set V is countable, therefore $X(t) - X(\alpha_n(t)) \rightarrow 0$ as $n \rightarrow \infty$ for all t simultaneously. Let t be an arbitrary point from the set V . Put $t_m = \alpha_m(t)$, $t_{m-1} = \alpha_{m-1}(t_m)$, ..., $t_1 = \alpha_1(t_2)$ for any integer $m \geq 1$. Then for all $m \geq 2$ the following

inequality takes place

$$X(t) = X(t_1) + \sum_{k=2}^m (X(t_k) - X(t_{k-1})) + X(t) - X(\alpha_m(t)) \leq \max_{u \in V_{\varepsilon_1}} X(u) + \sum_{k=2}^m \max_{u \in V_{\varepsilon_k}} (X(u) - X(\alpha_{k-1}(u))) + X(t) - X(\alpha_m(t)).$$

Therefore it follows from (11) that with probability one

$$\sup_{t \in B} X(t) = \sup_{t \in V} X(t) \leq \lim_{m \rightarrow \infty} \inf \left(\max_{u \in V_{\varepsilon_1}} X(u) + \sum_{k=2}^m \max_{u \in V_{\varepsilon_k}} (X(u) - X(\alpha_{k-1}(u))) \right).$$

Next, applying Hölder's inequality and Fatou's lemma, we obtain for all $\lambda > 0$ that

$$\begin{aligned} \mathbb{E} \exp \left\{ \lambda \sup_{t \in B} X(t) \right\} &\leq \mathbb{E} \liminf_{m \rightarrow \infty} \exp \left\{ \lambda \left(\max_{u \in V_{\varepsilon_1}} X(u) + \sum_{k=2}^m \max_{u \in V_{\varepsilon_k}} (X(u) - X(\alpha_{k-1}(u))) \right) \right\} \leq \\ &\leq \liminf_{m \rightarrow \infty} \mathbb{E} \exp \left\{ \lambda \left(\max_{u \in V_{\varepsilon_1}} X(u) + \sum_{k=2}^m \max_{u \in V_{\varepsilon_k}} (X(u) - X(\alpha_{k-1}(u))) \right) \right\} \leq \\ &\leq \liminf_{m \rightarrow \infty} \left(\mathbb{E} \exp \{ q_1 \lambda \max_{u \in V_{\varepsilon_1}} X(u) \} \right)^{\frac{1}{q_1}} \prod_{k=2}^m \left(\mathbb{E} \exp \{ q_k \lambda \max_{u \in V_{\varepsilon_k}} (X(u) - X(\alpha_{k-1}(u))) \} \right)^{\frac{1}{q_k}} \leq \\ &\leq \left(\mathbb{E} \exp \left\{ q_1 \lambda \max_{u \in V_{\varepsilon_1}} X(u) \right\} \right)^{\frac{1}{q_1}} \prod_{k=2}^{\infty} \left(\mathbb{E} \exp \left\{ q_k \lambda \max_{u \in V_{\varepsilon_k}} (X(u) - X(\alpha_{k-1}(u))) \right\} \right)^{\frac{1}{q_k}} = I_1 \cdot \prod_{k=2}^{\infty} I_k. \end{aligned}$$

Let's look at each multiplier I_k from the above separately. It follows from the Lemma 1 that

$$\mathbb{E} \exp \{ q_1 \lambda S(u) \} \leq \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_1 \lambda \tau) \}} \right)^{r t}$$

and

$$\mathbb{E} \exp \{ q_k \lambda (S(u) - S(\alpha_{k-1}(u))) \} \leq \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_k \lambda \tau) \}} \right)^{r |u - \alpha_{k-1}(u)|} \leq \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_k \lambda \tau) \}} \right)^{r \varepsilon_k}.$$

Using these inequalities, we get

$$\begin{aligned} I_1 &\leq \left(\sum_{u \in V_{\varepsilon_1}} \mathbb{E} \exp \{ q_1 \lambda S(u) \} \exp \{ -q_1 \lambda f(u) \} \right)^{\frac{1}{q_1}} \leq \\ &\leq \left(\sum_{u \in V_{\varepsilon_1}} \max_{u \in V_{\varepsilon_1}} \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_1 \lambda \tau) \}} \right)^{r u} \exp \{ -q_1 \lambda f(u) \} \right)^{\frac{1}{q_1}} \leq \\ &\leq (N_B(\varepsilon_1))^{\frac{1}{q_1}} \sup_{u \in B} \left[\left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_1 \lambda \tau) \}} \right)^{r u / q_1} \exp \{ -\lambda f(u) \} \right]. \end{aligned}$$

Due to the assumption on boundedness of increments of the function f

$$f(u) - f(\alpha_{k-1}(u)) \leq \delta(\rho(u, \alpha_{k-1}(u))) \leq \delta(\varepsilon_k).$$

Then in the similar way as for I_1 we get the following estimate for $k \geq 2$

$$\begin{aligned} I_k &\leq \left(N_B(\varepsilon_k) \max_{u \in V_{\varepsilon_k}} \mathbb{E} \exp \{ q_k \lambda (S(u) - S(\alpha_{k-1}(u))) \} \exp \{ -q_k \lambda (f(u) - f(\alpha_{k-1}(u))) \} \right)^{\frac{1}{q_k}} \leq \\ &\leq (N_B(\varepsilon_k))^{\frac{1}{q_k}} \left(\max_{u \in V_{\varepsilon_k}} \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_k \lambda \tau) \}} \right)^{r \varepsilon_k} \exp \{ q_k \lambda \delta(\varepsilon_k) \} \right)^{\frac{1}{q_k}} \leq \\ &\leq (N_B(\varepsilon_k))^{\frac{1}{q_k}} \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_k \lambda \tau) \}} \right)^{r \varepsilon_k / q_k} \exp \{ \lambda \delta(\varepsilon_k) \} = \\ &= \left(N_B \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right)^{\frac{1}{q_k}} \left(\frac{\theta}{1 - (1 - \theta) \exp \{ \varphi(q_k \lambda \tau) \}} \right)^{\frac{r \beta^2 p^{2k}}{\kappa^2 q_k}} \exp \left\{ \lambda \delta \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right\}. \end{aligned}$$

This is where the assertion of the lemma follows. $\square$

Applying Chebyshev's inequality to Lemma 2, one can obtain an upper estimate of the ruin probability for our risk model.

Theorem 2. Let $X(t) = S(t) - f(t)$, where $\{S(t) = \sum_{i=1}^{N_t} Y_i, t \in B\}$ is a cumulative claim process satisfying assumptions (I)–(IV). Assume also that increments of continuous premium function $f = \{f(t), t \in \mathbf{T}\}$ are bounded by such a continuous monotone increasing function $\delta = \{\delta(s), s > 0\}$ that $|f(u) - f(v)| \leq \delta(\rho(u, v))$, and let $r = \{r(u), u \geq 1\}$ be such a non-decreasing monotone continuous function that $r(u) > 0$ as $u > 1$, and $s(t) = r(\exp\{t\}), t \geq 0$, is a convex function. Then for all $0 \leq \beta \leq \min\{\sigma(b-a)/2, 1\}$ the following estimates hold true

$$\mathbf{P} \left\{ \sup_{t \in B} (S(t) - f(t)) > x \right\} \leq Z_r(\beta, x), \quad \mathbf{P} \left\{ \inf_{t \in B} (S(t) - f(t)) < -x \right\} \leq Z_r(\beta, x),$$

$$\mathbf{P} \left\{ \sup_{t \in B} |S(t) - f(t)| > x \right\} \leq 2Z_r(\beta, x),$$

where

$$Z_r(\beta, x) = \inf_{p \in (0, 1)} r^{(-1)} \left(\frac{1}{\beta p} \int_0^{\beta p} r(N_B(u^2/\kappa^2)) du \right) \times$$

$$\times \inf_{\lambda > 0} \exp \left\{ \eta_\varphi(\lambda, p) + \sum_{k=2}^{\infty} \left[\frac{r\beta^2 p^{3k-1}(1-p)}{\kappa^2} \left(\ln \theta - \ln \left(1 - (1-\theta) \exp \left\{ \varphi \left(\frac{\lambda \tau}{(1-p)p^{k-1}} \right) \right\} \right) \right) + \lambda \delta \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right] - \lambda x \right\},$$

$$\eta_\varphi(\lambda, p) = \sup_{u \in B} \left[(\ln \theta - \ln(1 - (1-\theta) \exp\{\varphi(\lambda \tau/(1-p))\})) (1-p)ru - \lambda f(u) \right].$$

Proof. Let's apply Lemma 2 for the sequence $q_k = \frac{1}{(1-p)p^{k-1}}$, where $p \in (0, 1)$. Since

$$\ln \sup_{u \in B} \left[\left(\frac{\theta}{1 - (1-\theta) \exp\{\varphi(\lambda \tau/(1-p))\}} \right)^{(1-p)ru} \exp\{-\lambda f(u)\} \right] =$$

$$= \sup_{u \in B} \left[(\ln \theta - \ln(1 - (1-\theta) \exp\{\varphi(\lambda \tau/(1-p))\})) (1-p)ru - \lambda f(u) \right] = \eta_\varphi(\lambda, p),$$

then

$$\mathbf{E} \exp \left\{ \lambda \sup_{t \in B} X(t) \right\} \leq \exp \left\{ \eta_\varphi(\lambda, p) + \sum_{k=1}^{\infty} (1-p)p^{k-1} H_B \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) + \right.$$

$$\left. + \sum_{k=2}^{\infty} \left[\frac{r\beta^2 p^{3k-1}(1-p)}{\kappa^2} \left(\ln \theta - \ln \left(1 - (1-\theta) \exp \left\{ \varphi \left(\frac{\lambda \tau}{(1-p)p^{k-1}} \right) \right\} \right) \right) + \lambda \delta \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right] \right\}. \quad (12)$$

Properties of $r(t)$ imply

$$\exp \left\{ \sum_{k=1}^{\infty} (1-p)p^{k-1} H_B \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right\} = r^{(-1)} \left(r \left(\exp \left\{ \sum_{k=1}^{\infty} (1-p)p^{k-1} \ln N_B \left(\frac{\beta^2 p^{2k}}{\kappa^2} \right) \right\} \right) \right) \leq$$

$$\leq r^{(-1)} \left(\sum_{k=1}^{\infty} (1-p)p^{k-1} r(N_B(\sigma^{(-1)}(\beta p^k))) \right) \leq r^{(-1)} \left(\frac{1}{\beta p} \int_0^{\beta p} r(N_B(\sigma^{(-1)}(u))) du \right), \quad (13)$$

where the function $\sigma(u)$ is defined in (9). Then the statement of the theorem follows from the Lemma 2, inequalities (12), (13) and the Chebyshev's one. $\square$

Corollary 1. Let $f(t) = ct$ describes the linear incoming flow of premiums during time interval $[0, T]$. Then $\delta(u) = cu$ and

$$\eta_\varphi(\lambda, p) = \left[(\ln \theta - \ln(1 - (1-\theta) \exp\{\varphi(\lambda \tau/(1-p))\})) (1-p)r - \lambda c \right] T$$

if $(\ln \theta - \ln(1 - (1-\theta) \exp\{\varphi(\lambda \tau/(1-p))\})) (1-p)r - \lambda c > 0$ and $\eta_\varphi(\lambda, p) = 0$ elsewhere.

Corollary 2. Let $\varphi(u) = u^2/2$, i.e. claim sizes are sub-Gaussian random variables. Let also $f(t) = ct, t \in [0, T]$ as in Corollary 1. Then

$$\eta_\varphi(\lambda, p) = \left[\left(\ln \theta - \ln \left(1 - (1-\theta) \exp \left\{ \frac{\lambda^2 \tau^2}{2(1-p)^2} \right\} \right) \right) (1-p)r - \lambda c \right] T$$

if $(\ln \theta - \ln(1 - (1-\theta) \exp \left\{ \frac{\lambda^2 \tau^2}{2(1-p)^2} \right\})) (1-p)r - \lambda c > 0$ and $\eta_\varphi(\lambda, p) = 0$ elsewhere.

Discussion and conclusions

The results of Lemma 2 can be extended to other sequences q_k , allowing for alternative estimates of the ruin probability. By specifying the exact form of the Orlicz N-function φ deeper optimization in Theorem 2 can be achieved. Of particular interest is the space of sub-Gaussian random variables, where $\varphi(x) = x^2/2$.

Authors' contribution: Rostyslav Yamnenko – conceptualization and theoretical foundations of research; Volodymyr Levchenko – making the mathematical calculations.

References

- Boucher, J.-P., Denuit, M., & Guillén, M. (2008). Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions. *Variance*, 2(1), 135–162.
- Buldygin, V., & Kozachenko, Y. (2000). *Metric characterization of random variables and random processes* (2nd ed.). American Mathematical Society, Providence, RI. <https://doi.org/10.1090/mmono/188>
- David, M., & Jemna, D.-V. (2015). Modeling the frequency of auto insurance claims by means of Poisson and negative binomial models. *Analele stiintifice ale Universitatii "Al. I. Cuza" din Iasi. Stiinte economice/Scientific Annals of the "Al. I. Cuza", 62, 151–168.*
- Giuliano Antonini, R., Kozachenko, Y. V., & Nikitina, T. (2003). Space of φ -sub-gaussian random variables. *Rendiconti, Accademia Nazionale delle Scienze detta dei XL, Memorie di Matematica e Applicazioni*, XXVII(121°), 92–124.
- Hopkalo, O. M., Kozachenko, Y. V., Vasylyk, O. I., & Sakhno, L. M. (2020). Some properties and estimates for φ -sub-gaussian stochastic processes. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 4, 18–22. <https://doi.org/10.17721/1812-5409.2019/4.3>
- Ismail, N., & Zamani, H. (2013). Estimation of claim count data using negative binomial, generalized Poisson, zero-inflated negative binomial and zero-inflated generalized Poisson regression models. In *Casualty actuarial society e-forum*, Vol. 41 (pp. 1–28).
- Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). *Loss models: from data to decisions* (Vol. 715). John Wiley & Sons.
- Kozachenko, Y., & Kovalchuk, Y. (1998). Boundary value problems with random initial conditions, and functional series from $\text{Sub}_s(\Omega)$. *Ukrainian Mathematical Journal*, 50(4), 504–515. <https://doi.org/10.1007/BF02487389>
- Kozachenko, Y., Yamnenko, R., & Vasylyk, O. (2005). Upper estimate of overrunning by $\text{Sub}_s(\Omega)$ random process the level specified by continuous function. *Random Operators and Stochastic Equations*, 13(2), 111–128. <https://doi.org/10.1515/156939705323383832>
- Saienko, M., & Yamnenko, R. (2013). Sub-Gaussian risk processes with dependent moments of claims incoming and contracts signing. *Scientific Bulletin of Uzhhorod University. Series of Mathematics and Informatics*, 24(2), 176–184 [in Ukrainian]. [Саєнко, М., & Ямненко, Р. (2007). Субгауссові процеси ризику із залежними і незалежними моментами настання страхових випадків й укладання договорів. *Науковий вісник Ужгородського університету. Серія "Математика і інформатика"*, 24(2), 176–184].
- Shi, P., & Valdez, E. A. (2014). Multivariate negative binomial models for insurance claim counts. *Insurance: Mathematics and Economics*, 55, 18–29. <https://doi.org/10.1016/j.insmatheco.2013.11.011>
- Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science* (Vol. 47). Cambridge university press. <https://doi.org/10.1017/9781108231596>
- Yamnenko, R., & Lamin, A. (2022). Estimation of ruin probability for binomially distributed number of φ -sub-Gaussian claims. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 2, 20–27. <https://doi.org/10.17721/1812-5409.2022/2.2>
- Yamnenko, R., & Vasylyk, O. (2007). Some properties of random Poisson sums with φ -sub-Gaussian terms. *Applied Statistics. Actuarial and Financial Mathematics*, 1, 133–148 [in Ukrainian]. [Василик, О. & Ямненко, Р. (2007). Деякі властивості випадкових пуассонівських сум із φ -субгауссовими доданками. *Прикладна статистика, актуарна та фінансова математика*, 1, 133–148].

Отримано редакцією журналу / Received: 02.01.25

Прорецензовано / Revised: 12.02.25

Схвалено до друку / Accepted: 13.02.25

Володимир ЛЕВЧЕНКО, асп.

ORCID ID: 0009-0007-9681-8753

e-mail: vlad.levchenko@protonmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

Ростислав ЯМНЕНКО, д-р фіз.-мат. наук

ORCID ID: 0000-0002-9612-7959

e-mail: rostyslav.yamnenko@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

ОЦІНЮВАННЯ ЙМОВІРНОСТІ БАНКРУТСТВА ДЛЯ ВИПАДКОВОЇ СУМИ НЕГАТИВНО БІНОМІАЛЬНО РОЗПОДІЛЕНОГО ЧИСЛА СУБГАУССОВИХ ПОЗОВІВ

У пропонованій роботі досліджено властивості класичного процесу ризику, в якому кількість страхових подій має негативний біноміальний розподіл, а розміри позовних вимог належать до класу узагальнених строго субгауссових випадкових величин. Цей клас є досить широким й охоплює як стандартні субгауссові, так і гауссові випадкові величини. В основі такого підходу лежить вимога, щоб кожна випадкова величина в цьому класі задовольняла певну нерівність, яка обмежує її експоненціальне зростання, що дає змогу контролювати поведінку її ймовірнісного хвоста. При цьому також передбачається, що страховий дохід, який надходить упродовж часу, є неперервною монотонно зростаючою функцією, природи якої не перевищують значення іншої неперервної зростаючої функції. Така умова сприяє гнучкому регулюванню зміні доходу з плином часу.

Метою роботи є отримання оцінок ймовірності банкрутства для описаної моделі процесу ризику. Для досягнення цієї мети використовується підхід, заснований на методі метричної ентропії. Цей метод дає змогу враховувати як складну природу розподілу кількості страхових подій, так і властивості строго субгауссових розмірів позовних вимог. Застосування такого підходу робить можливим отримання оцінок і в ситуаціях, коли точний вигляд розподілу позовів невідомий, але відомо, що вони мають субгауссову поведінку.

Ключові слова: негативний біноміальний розподіл, субгауссова випадкова величина, ймовірність банкрутства, розподіл страхових позовів, метрична ентропія.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

УДК 519.21

DOI: <https://doi.org/10.17721/1812-5409.2025/1.6>

Олег МАКАРЧУК, канд. фіз.-мат. наук, доц., докторант
ORCID ID: 0000-0002-1001-8568
e-mail: makarchuk@imath.kiev.ua; makolpet@gmail.com
Інститут математики НАН України, Київ, Україна

Богдан ХАЛЕЦЬКИЙ, асп.
ORCID ID: 0009-0004-5162-825X
e-mail: aumykun@gmail.com

Центральноукраїнський державний університет імені Володимира Винниченка, Кропивницький, Україна

СТРУКТУРА РОЗПОДІЛУ ОДНОГО КЛАСУ УЗАГАЛЬНЕНИХ ЗГОРТОК БЕРНУЛЛІ

Відповідно до теореми Джессена – Вінтнера сума збіжного з ймовірністю одиниця випадкового ряду з незалежними дискретно розподіленими доданками має чистий лебегівський тип розподілу, тобто дискретний, абсолютно неперервний або сингулярний. Теорема Леві надає необхідні та достатні умови того, що відповідний розподіл (розподіл типу Джессена – Вінтнера) є дискретним. У загальному випадку проблема знаходження критеріїв абсолютної неперервності (сингулярності) розподілів типу Джессена – Вінтнера є складною, тому розглядається дослідниками в певних класах.

Важливим класом розподілів, що відповідають теоремі Джессена – Вінтнера, є класичні згортки Бернуллі та їхні відповідні узагальнення (узагальнені згортки Бернуллі), що були предметом дослідження багатьох науковців. Інтерес до відповідної проблематики є високим і нині.

Особливі складнощі в дослідженні лебегівської структури розподілів типу Джессена – Вінтнера, зокрема в класі узагальнених згорток Бернуллі, виникають для випадку узагальнених згорток Бернуллі із суттєвими перекриттями. Для узагальнених згорток Бернуллі із суттєвими перекриттями класичні підходи теорії міри, що ґрунтуються зокрема на теоремі Какутані, у багатьох випадках дають змогу отримати лише певні достатні умови абсолютної неперервності (сингулярності) відповідної функції розподілу.

Крім класичної проблеми поглиблення теореми Джессена – Вінтнера для узагальнених згорток Бернуллі, актуальними є: проблема дослідження тополого-метричної структури спектра розподілу, фрактальний аналіз носіїв розподілу, асимптотичні властивості характеристичної функції на нескінченності тощо.

У пропонованій роботі визначені необхідні та достатні умови абсолютної неперервності та сингулярності розподілу одного класу узагальнених симетричних згорток Бернуллі. Відповідні результати є узагальненням результатів щодо лебегівської структури розподілу випадкової величини з незалежними s -ковими цифрами. Акцентовано також на доведенні того, що спектр відповідного класу розподілів, розглянутих у роботі, має додатну міру Лебега.

Ключові слова: структура розподілу, теорема Леві, теорема Джессена – Вінтнера, згортка Бернуллі, абсолютно неперервний розподіл, сингулярний розподіл.

Класифікація відповідно до AMS 2020: 60G30, 11A67.

Вступ

Актуальність дослідження. Нехай $s \in \mathbb{N}, s > 1$, $\sum_{n=1}^{\infty} a_n$ – абсолютно збіжний ряд, $A = (d_{ij})$ – матриця дійсних чисел, де $i \in \{0; 1; \dots; s-1\}, j \in \mathbb{N}$, та для деякого $L > 0$ виконується умова

$$|d_{ij}| \leq L \quad \forall i \in \{0; 1; \dots; s-1\}, \quad j \in \mathbb{N}. \quad (1)$$

Розглянемо множини

$$M = \{\sum_{n=1}^{\infty} x_n a_n \mid x_n \in \{d_{0n}, \dots, d_{(s-1)n}\} \forall n \in \mathbb{N}\},$$

$$T(d_{j_1 1}; d_{j_2 2}; \dots; d_{j_k k}) = \{\sum_{n=1}^k d_{j_n n} a_n + \sum_{n=k+1}^{\infty} x_n a_n \mid x_n \in \{d_{0n}, \dots, d_{(s-1)n}\} \forall n \in \mathbb{N}\},$$

де $j_l \in \{0; 1; \dots; s-1\}$ для кожного $l \in \{1; 2; \dots; k\}$.

Спектром S_G функції розподілу $G(x)$ називають множину чисел t таких, що $G(t + \varepsilon) - G(t - \varepsilon) > 0$ для кожного $\varepsilon > 0$.

Нехай (ψ_k) – послідовність незалежних дискретно розподілених випадкових величин, що набувають значень $d_{0k}, \dots, d_{(s-1)k}$ з ймовірностями $p_{0k}, p_{1k}, \dots, p_{(s-1)k}$ відповідно, причому $\sum_{j=0}^{s-1} p_{jk} = 0$. Розглянемо випадкову величину

$$\psi = \sum_{k=1}^{\infty} \psi_k a_k.$$

Ймовірнісну міру $\psi(\cdot)$, що відповідає випадковій величині ψ , називають узагальненою згорткою Бернуллі.

За теоремою Джессена – Вінтнера (Jessen, & Wintner, 1935) розподіл ψ є чистим. За теоремою Леві (Levy, 1931) розподіл ψ є дискретним, лише коли

$$\prod_{n=1}^{+\infty} \max(p_{0n}; p_{1n}; \dots; p_{(s-1)n}) > 0.$$

Об'єкт дослідження – узагальнені симетричні згортки Бернуллі.

Предметом дослідження є лебегівська структура одного класу узагальнених симетричних згорток Бернуллі.

Мета та завдання роботи: визначити необхідні та достатні умови того, що відповідні розподіли класу узагальнених симетричних згорток Бернуллі, які розглянуто в роботі, мають дискретний, абсолютно неперервний (сингулярний) тип розподілу.

Якщо для деякого натурального M виконується така умова: для довільного натурального $n \geq M$ та довільних наборів $(j_1; j_2; \dots; j_n)$ і $(i_1; i_2; \dots; i_n)$ таких, що $j_l, i_l \in \{0; 1; \dots; s-1\}$ для кожного $l \in \{1; 2; \dots; k\}$, множини $T(d_{j_1}; d_{j_2}; \dots; d_{j_k})$ та $T(d_{i_1}; d_{i_2}; \dots; d_{i_k})$ мають не більш ніж одну спільну точку, то відповідні ймовірнісні міри $\psi(\cdot)$ являють собою узагальнені симетричні згортки Бернуллі без суттєвих перекриттів. Інакше $\psi(\cdot)$ є узагальненою симетричною згортокою Бернуллі із суттєвими перекриттями.

Проблемі вивчення лебегівської структури розподілу ψ для певних матриць A та послідовностей (a_n) присвячено велику кількість робіт (Гончаренко, Працьовитий, & Торбін, 2005; Albeverio et al., 2007b; Peres, Schlag, & Solomyak, 2000). Випадок $a_n = \frac{1}{s^n}, d_{jn} = j$ для кожного $j \in \{0; 1; \dots; s-1\}, n \in N$ розглянуто в роботі (Kakutani, 1948), де були знайдені необхідні та достатні умови сингулярності (абсолютної неперервності) розподілу ψ ; та в роботах (Chatterji, 1964; Marsaglia, 1971) для деяких часткових значень s . Топологічно подібним до класичного s -кового зображення є Q_s -зображення (Albeverio et al., 2002), що є частковим випадком загальнішого $\tilde{Q}$ -зображення (Albeverio et al., 2007a). У роботах (Albeverio et al., 2002; Albeverio et al., 2007a) були знайдені необхідні та достатні умови дискретності й абсолютної неперервності (сингулярності) розподілу випадкової величини з незалежними $\tilde{Q}$ -цифрами, що є узагальненням результатів роботи (Kakutani, 1948).

Добре відомо (Goncharenko, Pratsiovytyi, & Torbin, 2005), що за умови $p_{jn} > 0$ для кожного $j \in \{0; 1; \dots; s-1\}, n \in N$ множина S_{F_ψ} збігається з M . Топологічні та фрактальні властивості множини M для деяких послідовностей (a_n) і матриць A досліджували в роботах (Kenyon, 1997; Nitecki, 2015). Випадок $a_n = \frac{1}{s^n}, d_{jn} = j$ для кожного $j \in \{0; 1; \dots; s-1\}, n \in N$, де $(d_0, d_1, \dots, d_{s-1})$ – набір цілих чисел, що утворюють повну систему лишків за модулем s , розглянуто в роботі (Kenyon, 1997), де сформульовано гіпотезу про те, що $\lambda(M) > 0$, де $\lambda(\cdot)$ – класична міра Лебега. Відповідну гіпотезу доведено в роботі (Nitecki, 2015) з використанням топологічних методів.

Методи дослідження: ймовірнісні міри, що фігурують у роботі (Kakutani, 1948) та їхні аналоги (Albeverio et al., 2007a) відповідають випадку узагальнених згорток Бернуллі без суттєвих перекриттів. Поглиблені варіанти теореми Какутані, дають змогу за певних умов разом з теоремою Леві розв'язати повністю проблему поглиблення теореми Джессена – Вінтнера у відповідних класах. У загальному випадку теорема Какутані дає змогу будувати певні класи мір еквівалентних (ортогональних) до класичної міри Лебега. У роботі (Albeverio et al., 2007b) був представлений підхід зі зведення проблеми поглиблення теореми Джессена – Вінтнера для одного класу узагальнених симетричних згорток Бернуллі із суттєвими перекриттями до відповідних розподілів з незалежними $\tilde{Q}$ -цифрами. Потрібно зауважити, що реалізувати відповідний підхід роботи (Albeverio et al., 2007b) в загальному випадку достатньо важко.

У пропонованому дослідженні поглиблено результати робіт (Kakutani, 1948) та (Nitecki, 2015).

1. Лебегівська структура розподілу ψ .

Надалі будемо розглядати випадок, коли для кожного натурального $n: a_n = \frac{1}{s^n}$ та $(d_{0n}; d_{1n}; \dots; d_{(s-1)n})$ є набором цілих чисел, що утворюють повну систему лишків за модулем s та задовольняють умову (1). Головним результатом цієї роботи є теорема 1, що являє собою поглиблення результатів роботи (Kakutani, 1948).

Теорема 1. Випадкова величина ψ має чистий розподіл, причому

1) дискретний лише тоді, коли

$$\prod_{n=1}^{+\infty} \max(p_{0n}; p_{1n}; \dots; p_{(s-1)n}) > 0;$$

2) сингулярний лише тоді, коли

$$\prod_{n=1}^{+\infty} \max(p_{0n}; p_{1n}; \dots; p_{(s-1)n}) = 0$$

та

$$\prod_{n=1}^{+\infty} \sum_{j=0}^{s-1} \sqrt{\frac{p_{jn}}{s}} = 0;$$

3) абсолютно неперервний лише тоді, коли

$$\prod_{n=1}^{+\infty} \sum_{j=0}^{s-1} \sqrt{\frac{p_{jn}}{s}} > 0.$$

Для доведення теореми 1 необхідно довести наступні лема.

Лема 1. Якщо

$$\prod_{n=1}^{+\infty} (s \max_{0 \leq j \leq s-1} p_{jn}) = G < +\infty,$$

то $F_\psi(x)$ задовольняє умову Ліпшиця

$$|F_\psi(x_1) - F_\psi(x_2)| \leq G|x_1 - x_2| \quad \forall x_1, x_2 \in R.$$

Доведення. Ураховуючи теорему Леві, бачимо, що розподіл ψ є неперервним. Зрозуміло, що для довільної обмеженої послідовності цілих чисел (b_n) типи розподілів випадкових величин ψ та

$$\psi + \sum_{j=1}^{+\infty} \frac{b_j}{s^j} = \sum_{j=1}^{+\infty} \frac{\psi_j + b_j}{s^j}$$

є однаковими. Отже, не обмежуючи загальності, нехай для кожного натурального n

$$0 = d_{0n} < d_{1n} < \dots < d_{(s-1)n} \leq L. \quad (2)$$

Для кожного цілого r та натурального n розглянемо величину:

$$S(r; n) = \sum_{(j_1; j_2; \dots; j_n)} \prod_{k=1}^n p_{j_k k},$$

де сумування беремо за всіма наборами $(j_1; j_2; \dots; j_n)$ таким, що $j_k \in \{0; 1; \dots; s-1\}$ для кожного $k \in \{1; \dots; n\}$ і виконується рівність

$$\frac{r}{s^n} = \sum_{k=1}^n \frac{d_{j_k k}}{s^k}. \quad (3)$$

Якщо рівність (3) не справджується за всіма відповідними наборами $(j_1; j_2; \dots; j_n)$, то покладемо $S(r; n) = 0$. Розглянемо функції

$$f_n(t) = \prod_{k=1}^n \sum_{j=0}^{s-1} p_{jk} t^{d_{jk} s^{-k}}.$$

Бачимо, що

$$f_n(t) = \sum_r S(r; n) t^{rs^{-n}}.$$

Розглянемо многочлени

$$h_n(t) = f_n(t^{s^n}) = \sum_r S(r; n) t^r.$$

Індукцією за n покажемо, що

$$S(r; n) \leq \prod_{k=1}^n \max_{0 \leq j \leq s-1} p_{jk} \quad \forall r \in \mathbb{Z}, n \in \mathbb{N}. \quad (4)$$

За $n = 1$ нерівність (4) очевидно виконується. Нехай умова (4) справджується для деякого натурального n . Зрозуміло, що

$$h_{n+1}(t) = h_n(t^{s^n}) \sum_{j=0}^{s-1} p_{j(n+1)} t^{d_{j(n+1)}},$$

звідки для заданого r існує ціле j таке, що $d_{j(n+1)} \equiv r \pmod{s}$ і виконується умова

$$S(r; n+1) = S(r; n) p_{j(n+1)} \leq \max_{0 \leq j \leq s-1} p_{j(n+1)} \cdot \prod_{k=1}^n \max_{0 \leq j \leq s-1} p_{jk} = \prod_{k=1}^{n+1} \max_{0 \leq j \leq s-1} p_{jk}.$$

Позначимо

$$W_{n+1} = \sum_{k=n+1}^{+\infty} \frac{\psi_k}{s^k}$$

та обрахуємо величину

$$F_\psi\left(\frac{r+1}{s^n}\right) - F_\psi\left(\frac{r}{s^n}\right) = P\left(\psi \in \left[\frac{r}{s^n}; \frac{r+1}{s^n}\right]\right)$$

для натуральних n та $r+1$.

Ураховуючи умову (2), бачимо, що $\psi \in \left[\frac{r}{s^n}; \frac{r+1}{s^n}\right]$ лише тоді, коли

$$\sum_{k=1}^n \frac{\psi_k}{s^k} = \frac{r-l}{s^n}; \quad W_{n+1} \in \left[\frac{l}{s^n}; \frac{l+1}{s^n}\right],$$

де $l \in \{0; 1; \dots; \left[\frac{L}{s-1}\right]\}$. Ураховуючи нерівність (4), маємо

$$\begin{aligned} F_\psi\left(\frac{r+1}{s^n}\right) - F_\psi\left(\frac{r}{s^n}\right) &= \sum_l P\left(\sum_{k=1}^n \frac{\psi_k}{s^k} = \frac{r-l}{s^n}\right) P\left(W_{n+1} \in \left[\frac{l}{s^n}; \frac{l+1}{s^n}\right]\right) = \\ &= \sum_l S(r-l; n) P\left(W_{n+1} \in \left[\frac{l}{s^n}; \frac{l+1}{s^n}\right]\right) \leq \prod_{k=1}^n \max_{0 \leq j \leq s-1} p_{jk} \sum_l P\left(W_{n+1} \in \left[\frac{l}{s^n}; \frac{l+1}{s^n}\right]\right) \leq \frac{G}{s^n}. \end{aligned}$$

Для натуральних $m, k+1, n$ таких, що $m > k$ маємо

$$F_\psi\left(\frac{m}{s^n}\right) - F_\psi\left(\frac{k}{s^n}\right) = \sum_{j=1}^{m-k} F_\psi\left(\frac{k+j}{s^n}\right) - F_\psi\left(\frac{k+j-1}{s^n}\right) \leq \frac{Gm}{s^n} - \frac{Gk}{s^n}.$$

Ураховуючи останню нерівність і неперервність $F_\psi(x)$, легко отримати потрібне. □

Зауваження 1. Якщо $p_{ij} = \frac{1}{s} \forall i \in \{0; 1; \dots; s-1\}, j \in \mathbb{N}$, то з лема 1 розподіл ψ абсолютно неперервний і відповідно $\lambda(M) > 0$, що є поглибленням результату роботи (Nitecki, 2015).

Лема 2. Якщо виконується умова

$$\prod_{n=1}^{+\infty} \sum_{j=0}^{s-1} \sqrt{\frac{p_{jn}}{s}} > 0,$$

то розподіл ψ абсолютно неперервний.

Доведення. Не обмежуючи загальності, нехай виконується умова (2). Нехай (ξ_k) – послідовність незалежних випадкових величин, які набувають значень $(d_{0n}; d_{1n}; \dots; d_{(s-1)n})$ з рівними ймовірностями $\frac{1}{s}$. Розглянемо випадкову величину

$$\xi = \sum_{k=1}^{\infty} \frac{\xi_k}{s^k}.$$

За лемою 1 розподіл ξ абсолютно неперервний. Визначимо дві послідовності ймовірнісних просторів $\{(\Omega_k, A_k, \mu_k)\}$ і $\{(\Omega_k, A_k, \nu_k)\}$ так $\Omega_k = \{d_{0k}; d_{1k}; \dots; d_{(s-1)k}\}$, σ – алгебра A_k визначена на всіх підмножинах Ω_k :

$$\begin{aligned} \mu_k(d_{0k}) &= p_{0k}, \dots, \mu_k(d_{(s-1)k}) = p_{(s-1)k}; \\ \nu_k(d_{0k}) &= \frac{1}{s}, \nu_k(d_{1k}) = \frac{1}{s}, \dots, \nu_k(d_{(s-1)k}) = \frac{1}{s}. \end{aligned}$$

Зрозуміло, що $\mu_k \ll \nu_k$ для кожного $k \in \mathbb{N}$. Розглянемо нескінченні добутки ймовірнісних просторів

$$(\Omega, A, \mu) = \prod_{k=1}^{\infty} (\Omega_k, A_k, \mu_k); \quad (\Omega, A, \nu) = \prod_{k=1}^{\infty} (\Omega_k, A_k, \nu_k).$$

За теоремою Какутані (Kakutani, 1948), $\mu \ll \nu$ тоді і тільки тоді, коли

$$\prod_{k=1}^{\infty} \int_{\Omega_k} \sqrt{\frac{d\mu_k}{d\nu_k}} d\nu_k > 0 \Leftrightarrow \prod_{n=1}^{+\infty} \sum_{j=0}^{s-1} \sqrt{\frac{p_{jn}}{s}} > 0.$$

Остання умова виконується за умовою теореми. Розглянемо вимірне відображення $\Omega \xrightarrow{\varphi} M$, визначене рівністю

$$\forall \omega \in \Omega: \varphi(\omega) = \sum_{k=1}^{\infty} \frac{\omega_k}{s^k}.$$

Визначимо образи μ^* і ν^* мір μ і ν під дією φ так:

$$\mu^*(E) = \mu(\varphi^{-1}(E)); \quad \nu^*(E) = \nu(\varphi^{-1}(E))$$

для довільної борелівської підмножини E . Міра μ^* збігається з імовірнісною мірою P_{ψ} , а міра ν^* – з імовірнісною мірою P_{ξ} , еквівалентною мірі Лебега. З абсолютної неперервності міри μ відносно міри ν випливає абсолютна неперервність міри μ^* відносно міри ν^* .

Оскільки $\nu^* \sim \lambda$, то розподіл ψ абсолютно неперервний.

Лема 3. Якщо виконуються умови

$$\begin{aligned} \prod_{n=1}^{+\infty} \max(p_{0n}; p_{1n}; \dots; p_{(s-1)n}) &= 0; \\ \prod_{n=1}^{+\infty} \sum_{j=0}^{s-1} \sqrt{\frac{p_{jn}}{s}} &= 0, \end{aligned}$$

то розподіл ψ сингулярний.

Доведення. Не обмежуючи загальності, нехай виконується умова (2). Нехай

$$\tau_k = \sum_{j=1}^k \frac{\psi_j}{s^j}; \quad \eta_k = \sum_{n=1}^{+\infty} \frac{\psi_{k+n}}{s^n}.$$

Зрозуміло, що $\psi = \tau_k + \frac{\eta_k}{s^k}$ для кожного натурального k . Очевидно, що випадкова величина τ_k набуває значень $c_1, c_2, \dots, c_{s^k}$ з імовірностями $q_1, q_2, \dots, q_{s^k}$, причому для довільного $j \in \{1; 2; \dots; s^k\}$ існує набір чисел $(\alpha_1, \alpha_2, \dots, \alpha_k)$, кожне з яких належить множині $\{0; 1; \dots; s-1\}$ і виконується рівність $q_j = p_{\alpha_1 1} p_{\alpha_2 2} \dots p_{\alpha_k k}$.

Маємо

$$F_{\psi}(x) = P\left(\tau_k + \frac{\eta_k}{s^k} < x\right) = \sum_{j=1}^{s^k} q_j \cdot P\left(c_j + \frac{\eta_k}{s^k} < x\right) = \sum_{j=1}^{s^k} q_j \cdot P\left(\eta_k < s^k \cdot (x - c_j)\right) = \sum_{j=1}^{s^k} q_j \cdot F_{\eta_k}(s^k \cdot (x - c_j)).$$

Припустимо, що $F_{\psi}(x)$ абсолютно неперервна, тоді для кожного натурального k функція $F_{\eta_k}(x)$ також абсолютно неперервна і, враховуючи останню рівність, маємо

$$p_{\psi}(x) = \sum_{j=1}^{s^k} q_j s^k p_{\eta_k}(s^k \cdot (x - c_j))$$

майже скрізь.

Позначимо

$$H = \int_{-\infty}^{+\infty} \sqrt{p_{\psi}(x)} dx$$

та

$$B_k = \int_{-\infty}^{+\infty} \sqrt{p_{\eta_k}(x)} dx.$$

Оскільки $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ для довільних невід'ємних чисел a, b , то маємо

$$\begin{aligned} A &= \int_{-\infty}^{+\infty} \sqrt{\sum_{j=1}^{s^k} q_j s^k p_{\eta_k}(s^k(x - c_j))} dx \leq \sum_{j=1}^{s^k} \sqrt{q_j s^k} \int_{-\infty}^{+\infty} p_{\eta_k} \cdot (s^k \cdot (x - c_j)) dx = \\ &= B_k \sum_{j=1}^{s^k} \frac{\sqrt{q_j s^k}}{s^k} = B_k \prod_{n=1}^k \sum_{j=1}^{s-1} \sqrt{\frac{p_{jn}}{s}}. \end{aligned}$$

Оскільки для кожного натурального k

$$\sum_{n=1}^{+\infty} \frac{d_{(s-1)n}}{s^n} \leq \sum_{n=1}^{+\infty} \frac{L}{s^n} = \frac{L}{s-1},$$

то $P\left(\eta_k \notin \left[0; \frac{L}{s-1}\right]\right) = 0$, тому

$$\int_{-\infty}^{+\infty} \sqrt{p_{\eta_k}(x)} dx = \int_0^{\frac{L}{s-1}} \sqrt{p_{\eta_k}(x)} dx \leq \int_0^{\frac{L}{s-1}} \frac{1 + p_{\eta_k}(x)}{2} dx = \frac{\frac{L}{s-1} + 1}{2}.$$

Отже, отримали

$$H \leq \prod_{n=1}^k \sum_{j=0}^{s-1} \sqrt{\frac{p_{jn}}{s}} \cdot \frac{\frac{L}{s-1} + 1}{2} \rightarrow 0 \quad (k \rightarrow \infty),$$

звідки $H = 0$, тому майже скрізь $p_{\psi}(x) = 0$, прийшли до суперечності.

□

Дискусія і висновки

Робота має теоретичний характер. Щодо випадкової величини ψ , розглянутій у роботі, актуальними є такі задачі: вивчення топологічної структури множини M та $S_{F_{\psi}}$; знаходження необхідних і достатніх умов того, що $F_{\psi}(x)$ задовольняє умову Ліпшиця; знаходження точних значень показників Гельдера для $F_{\psi}(x)$. Зауважимо, що для випадку $s = 2, d_{0n} = 1, d_{1n} = 0$ для кожного $n \in N$, відомо (Nitecki, 2015), що M є скінченним об'єднанням відрізків, або множиною гомеоморфною множині Кантора або ж канторвалом. Навіть для попереднього випадку знайдені лише деякі достатні умови того, коли M є канторвалом. Топологічна структура $S_{F_{\psi}}$ у більшості випадків добре вивчена для узагальнених симетричних згорток Бернуллі без суттєвих перекриттів (Goncharenko, Pratsiovytyi, & Torbin, 2005). Проблема знаходження необхідних і достатніх умов того, що $F_{\psi}(x)$ задовольняє умову Ліпшиця та знаходження точних значень показників Гельдера для $F_{\psi}(x)$ є відкритою навіть для класичних симетричних згорток Бернуллі (Peres, Schlag, & Solomyak, 2000).

Предметом подальших досліджень є задача вивчення топологічної структури множини M та лебегівської структури розподілу випадкової величини ψ за умови, що $(d_{0n}; d_{1n}; \dots; d_{(s-1)n})$ не утворює повну систему лишків для нескінченної множини значень n .

Внесок авторів: Олег Макачук – концептуалізація та методика наукового дослідження; Богдан Халецький – аналіз джерел, підготовка огляду літератури, ідейна складова леми 1.

Список використаних джерел

- Гончаренко, Я. В., Працьовитий, М. В., & Торбін, Г. М. (2005). Тополого-метричні і фрактальні властивості множини неповних сум знакододатного ряду та розподілів на ній. *Науковий часопис НПУ імені М.П. Драгоманова. Серія 1. Фізико-математичні науки*, 6, 210–224.
- Albeverio, S., Koshmanenko, V. D., Pratsiovytyi, M. V., & Torbin, G. M. (2007a). On fine structure of singularly continuous probability measures and random variables with independent $\bar{Q}$ -symbols. *Methods of functional analysis and topology*, 17(2), 97–111.
- Albeverio, S., Goncharenko, Y. V., Pratsiovytyi, M. V., & Torbin, G. M. (2007b). Convolutions of distributions of random variables with independent binary digits. *Random Operators and Stochastic Equations*, 15(1), 89–97. <https://doi.org/10.1515/ROSE.2007.006>
- Albeverio, S., Koshmanenko, V. D., Pratsiovytyi, M. V., & Torbin, G. M. (2002). *Q-representation of real numbers and fractal probability distributions*. ArXiv. <https://doi.org/10.48550/arXiv.math/0308007>
- Chatterji, S. (1964). Certain induced measures and the fractional dimensions of their "supports". *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 3, 184–192. <https://doi.org/10.1007/BF00534907>
- Jessen, B., & Wintner, A. (1935). Distribution function and Riemann Zeta-function. *Transactions of the American Mathematical Society*, 38(1), 48–88.
- Kakutani, S. (1948). Equivalence of infinite product measures. *Annals of Mathematics*, 49, 214–224. <https://doi.org/10.2307/1969123>
- Kenyon, R. (1997). Projecting the one-dimensional Sierpinski gasket. *Israel Journal of Mathematics*, 97, 221–238. <https://doi.org/10.1007/BF02774038>
- Levy, P. (1931). Sur les sries don't les termes sont des variables independantes. *Studia Mathematica*, 3(1), 119–155. <https://doi.org/10.4064/SM-3-1-119-155>
- Marsaglia, G. (1971). Random variables with independent binary digits. *Annals of Mathematical Statistics*, 42(2), 1922–1929. <https://doi.org/10.1214/aoms/1177693058>
- Nitecki, Z. (2015). Cantorvals and subsum sets of null sequences. *The American mathematical monthly*, 122(9), 862–870. <https://doi.org/10.4169/amer.math.monthly.122.9.862>
- Peres, Y., Schlag, W., & Solomyak, B. (2000). Sixty years of Bernoulli convolutions. *Fractal Geometry and Stochastics II. Progress in Probability*, 46, 39–65. <https://doi.org/10.1007/978-3-0348-8380>

References

- Albeverio, S., Koshmanenko, V. D., Pratsiovytyi, M. V., & Torbin, G. M. (2007a). On fine structure of singularly continuous probability measures and random variables with independent $\bar{Q}$ -symbols. *Methods of functional analysis and topology*, 17(2), 97–111.
- Albeverio, S., Goncharenko, Ya. V., Pratsiovytyi, M. V., & Torbin, G. M. (2007b). Convolutions of distributions of random variables with independent binary digits. *Random Operators and Stochastic Equations*, 15(1), 89–97. <https://doi.org/10.1515/ROSE.2007.006>
- Albeverio, S., Koshmanenko, V. D., Pratsiovytyi, M. V., & Torbin, G. M. (2002). *Q-representation of real numbers and fractal probability distributions*. ArXiv. <https://doi.org/10.48550/arXiv.math/0308007>

- Chatterji, S. (1964). Certain induced measures and the fractional dimensions of their "supports". *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 3, 184–192. <https://doi.org/10.1007/BF00534907>
- Goncharenko, Ya. V., Pratsiovytyi, M. V., & Torbin, G. M. (2005). Topological-metric and fractal properties of the set of incomplete sums of a sign-positive series and distributions on it. *Scientific Journal of National Pedagogical Dragomanov University, Series 1. Physical and Mathematical Sciences*, 6, 210–224.
- Jessen, B., & Wintner, A. (1935). Distribution function and Riemann Zeta-function. *Transactions of the American Mathematical Society*, 38(1), 48–88.
- Kakutani, S. (1948). Equivalence of infinite product measures. *Annals of Mathematics*, 49, 214–224. <https://doi.org/10.2307/1969123>
- Kenyon, R. (1997). Projecting the one-dimensional Sierpinski gasket. *Israel Journal of Mathematics*, 97, 221–238. <https://doi.org/10.1007/BF02774038>
- Levy, P. (1931). On series whose terms are independent variables. *Mathematical Studies*, 3(1), 119–155. <https://doi.org/10.4064/SM-3-1-119-155>
- Marsaglia, G. (1971). Random variables with independent binary digits. *Annals of Mathematical Statistics*, 42(2), 1922–1929. <https://doi.org/10.1214/aoms/1177693058>
- Nitecki, Z. (2015). Cantorvals and subsum sets of null sequences. *The American mathematical monthly*, 122(9), 862–870. <https://doi.org/10.4169/amer.math.monthly.122.9.862>
- Peres, Y., Schlag, W., & Solomyak, B. (2000). Sixty years of Bernoulli convolutions. *Fractal Geometry and Stochastics II. Progress in Probability*, 46, 39–65. <https://doi.org/10.1007/978-3-0348-8380>

Отримано редакцією журналу / Received: 29.04.24

Прорецензовано / Revised: 07.03.25

Схвалено до друку / Accepted: 12.03.25

Oleg MAKARCHUK, PhD (Phys. & Math.), Assoc. Prof., Doctoral Student
ORCID ID: 0000-0002-1001-8568
e-mail: makarchuk@imath.kiev.ua; makolpet@gmail.com
Institute of Mathematics NAS of Ukraine, Kyiv, Ukraine

Bohdan KHALETSKY, PhD Student
ORCID ID: 0009-0004-5162-825X
e-mail: aumykun@gmail.com
Volodymyr Vynnychenko Central Ukrainian State University, Kropyvnytskyi, Ukraine

STRUCTURE OF THE DISTRIBUTION OF ONE CLASS OF GENERALIZED BERNOULLIE CONVOLUTIONS

According to the Jessen – Wintner theorem, the sum of a convergent with probability unity random series with independent discretely distributed terms has a pure Lebesgue type of distribution, i.e. discrete, absolutely continuous, or singular. Levi's theorem gives necessary and sufficient conditions for the corresponding distribution (Jessen-Wintner type distribution) to be discrete. In the general case, the problem of finding criteria for absolute continuity (singularity) of Jessen-Wintner type distributions is complex, therefore it is considered by researchers in certain classes.

An important class of distributions that satisfy the Jessen – Wintner theorem are classical Bernoulli convolutions and their corresponding generalizations (generalized Bernoulli convolutions), which have been the subject of research by many scientists. Interest in the corresponding problem is still high.

Special difficulties in the study of the Lebesgue structure of Jessen – Wintner type distributions, in particular in the class of generalized Bernoulli convolutions, arise for the case of generalized Bernoulli convolutions with substantial overlaps. For generalized Bernoulli convolutions with substantial overlaps, classical approaches of measure theory, based in particular on the Kakutani theorem, in many cases allow obtaining only certain sufficient conditions for the absolute continuity (singularity) of the corresponding distribution function.

In addition to the classical problem of deepening the Jessen-Wintner theorem for generalized Bernoulli convolutions, the following are relevant: the problem of studying the topological-metric structure of the distribution spectrum, fractal analysis of the distribution supports, asymptotic properties of the characteristic function at infinity, and others.

The paper specifies necessary and sufficient conditions for the absolute continuity and singularity of the distribution of one class of generalized symmetric Bernoulli convolutions. The corresponding results are a generalization of the results with respect to the Lebesgue structure of the distribution of a random variable with independent s -digits. The emphasis is also placed on proving that the spectrum of the corresponding class of distributions considered in the paper has a positive Lebesgue measure.

Keywords: distribution structure, Levy theorem, Jessen – Wintner theorem, Bernoulli convolution, absolutely continuous distribution, singular distribution.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.2

DOI: <https://doi.org/10.17721/1812-5409.2025/1.7>

Hening HUANG, PhD, Researcher
ORCID ID: 0000-0003-1882-0042
e-mail: heninghuang1@gmail.com
Teledyne RD Instruments, San Diego, USA

THE THEORY OF INFORMITY: A NOVEL PROBABILITY FRAMEWORK

This paper proposes a novel probability framework, called the theory of informity. We define a mathematical quantity called "informity" to quantitatively measure the degree of informativeness of a probability distribution (or a probability system). We also define two other quantities: cross-informity and joint informity. We propose an informity metric that can be used as an alternative to entropy metric. The informities for twelve continuous distributions are given. Three examples are presented to demonstrate the practicability of the proposed informity metric.

Keywords: *informity, information content, probability, probability distribution.*

AMS 2020 classification: 60E05, 94A17.

Introduction

Shannon (1948) introduced the concepts of "information content" and "information entropy". The Shannon information content (also called self-information, surprisal, or Shannon information) of an outcome x_i with probability $P(x_i)$ is defined as

$$I(x_i) := -\log_2 P(x_i). \quad (1)$$

For a discrete random variable X with the probability mass function (PMF) $P(x)$, the entropy (also called Shannon entropy or information entropy) of X is defined as

$$H(X) := -\sum_{i=1}^N P(x_i) \log_2 P(x_i) = E[-\log_2 P(X)], \quad (2)$$

where N is the number of outcomes of the probability system $\{X; P(x)\}$, and $E[\cdot]$ means taking the "mathematical expectation".

For a continuous random variable Y with the probability density function (PDF) $p(y)$, the entropy of Y is defined as

$$H(Y) := -\int p(y) \log_2 p(y) dy = E[-\log_2 p(Y)]. \quad (3)$$

When discussing Shannon information theory, "information" is actually shorthand for "information content," which refers to the amount of information conveyed by the outcome x_i based on its probability $P(x_i)$. Since most people may not explicitly think of "information" as shorthand for "information content," we use "information content" to avoid confusion. Furthermore, the notion of "information content" does not involve the meaning of information.

Needless to say, Shannon information content and entropy are key concepts in Shannon information theory, which provides a framework for transmission, storage, compression, and processing of information in communications engineering. However, the use of Shannon information content and entropy in fields other than communications engineering may not be consistent with our common sense about information or informativeness. For example, entropy metric is used in machine learning for splitting decision tree nodes. However, the so-called "information gain" is actually the "Shannon information content loss" corresponding to the reduction of entropy. Petty (2018) discussed some of the shortcomings of Shannon entropy as a measure of information content in indirect measurements of continuous variables. They highlighted a number of significant limitations of Shannon entropy as a basis for quantifying the information content of remote sensing observations. Based on a historical review of the problems with conceptualization of information content, Schroeder (2004) demonstrated the need for an alternative to entropy for measuring information content more generally. He stated, "...there is nothing wrong with entropy as long as we analyze communication systems and are interested in the 'engineering problem'... The question is whether we can use entropy equally successfully for measuring information [content] in different contexts. Here I believe the answer is in the negative."

Shannon information theory links "Shannon information content" with "probability" as shown in Eq. (1). However, "probability" is a (normalized) *relative* quantity ranging from 0 to 1 and has no unit, while "Shannon information content" is an *absolute* quantity ranging from 0 to infinity and has units such as "bits". By definition, a relative quantity is expressed as a ratio or percentage compared to (or normalized to) a reference value, whereas an absolute quantity is a numerical value that does not rely on comparison with another value. The Appendix explains how the *relative* quantity "probability" is converted into the *absolute* quantity "Shannon information content". Since Shannon information content and entropy are absolute quantities, they can be encoded, transmitted, compressed, and stored in communications engineering. However, in other fields or in daily life, people are more familiar with and use the *relative* quantity "probability". For example, when talking about the possibility of rain tomorrow, people would say "the probability of rain tomorrow is 80 %", rather than "the chance of rain tomorrow is 0.32 bits".

The aim of this study is to develop a novel probability framework, called the theory of informity, that is more applicable to fields other than communications engineering. Unlike Shannon information theory, which deals with *absolute* information content and uses log-transformation of probabilities, the proposed informity theory treats probability as *relative* information content and does not use log-transformation. The greater the probability of an event, the greater the relative information content of the event (i. e. the event is more informative). Conversely, the smaller the probability of an event, the smaller the relative information content of the event (i. e. the event is less informative). We believe this view of probability is intuitive and consistent with our common sense or perceptual understanding about information or informativeness. The relative information

© Huang Hening, 2025

content has the opposite meaning from Shannon information content, which is the absolute information content explained in the Appendix.

In the following sections, Section 2 defines a mathematical quantity called "infirmity" to quantitatively measure the degree of informativeness of a probability distribution (or a probability system). Section 3 defines cross-infirmity and joint infirmity. Section 4 presents infirmity metric. Section 5 provides discussion. Section 6 gives three application examples. Section 7 provides conclusions.

1. The definition of infirmity

We consider the word "infirmity" to be a technical term meaning "the degree of informativeness". This is similar to the technical term "uncertainty" in measurement science, which means "degree of uncertainty". The mathematical definition of infirmity is different for discrete random variables and continuous random variables.

1.1. Discrete random variables

Definition 1. Let $\beta(X)$ denote the infirmity of a discrete random variable X with the PMF $P(x)$. $\beta(X)$ is given by the mathematical expectation of $P(X)$

$$\beta(X) := E[P(X)] = \sum_{i=1}^N [P(x_i)]^2. \quad (4)$$

In the context of relative information content, $\beta(X)$ is the average amount of relative information contained in the probability system $\{X; P(x)\}$. Similar to how $H(X)$ is called discrete entropy, $\beta(X)$ is called discrete infirmity. $\beta(X)$ quantitatively measures the degree of informativeness (or certainty) of the probability system $\{X; P(x)\}$. It has the opposite meaning from the discrete entropy $H(X)$. $\beta(X)$ ranges between 0 and 1 and is dimensionless.

Consider tossing a coin, which is a two-state probability system: $\{X; P(x)\} = \{\text{head, tail}; P(\text{head}), P(\text{tail})\}$, where $P(\text{tail}) = 1 - P(\text{head})$. The infirmity $\beta(X)$ of this system can be calculated as

$$\beta(X) = [P(\text{head})]^2 + [P(\text{tail})]^2. \quad (5)$$

Figure 1 shows the infirmity $\beta(X)$ as a function of $P(\text{head})$, compared with the corresponding entropy (in bits): $H(X) = -P(\text{head})\log_2 P(\text{head}) - P(\text{tail})\log_2 P(\text{tail})$.

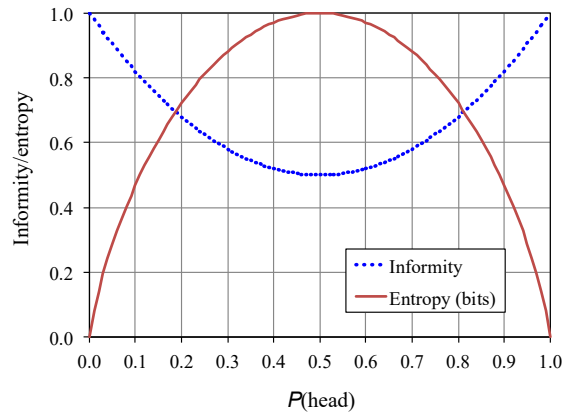


Fig. 1. Infirmity of the coin-tossing system as a function of the bias represented by the probability of heads, compared with the corresponding entropy

As can be seen in Fig. 1, when the coin is fair, i. e. $P(\text{head}) = P(\text{tail}) = 0.5$, the infirmity $\beta(X)$ is the lowest and the entropy $H(X)$ is the largest, indicating that the system is least informative or has the greatest uncertainty. When a coin is highly biased towards the tail or head, the infirmity increases and the entropy decreases. In the most extreme case of $P(\text{tail}) = 1$ or $P(\text{head}) = 1$, the infirmity is the highest: $\beta(X) = 1$ and the entropy is the lowest: $H(X) = 0$, indicating that the system is highest informative or has the greatest certainty.

1.2. Continuous random variables

Definition 2. Let $\beta(Y)$ denote the infirmity of a continuous random variable Y with the PDF $p(y)$. $\beta(Y)$ is given by the mathematical expectation of $p(Y)$

$$\beta(Y) := E[p(Y)] = \int [p(y)]^2 dy. \quad (6)$$

Similar to how $H(Y)$ is called continuous entropy, $\beta(Y)$ is called continuous infirmity. $\beta(Y)$ quantitatively measures the degree of informativeness (or certainty) of the probability system $\{Y; P(y)\}$. It has the opposite meaning from the continuous entropy $H(Y)$. Table 1 shows the continuous infirmities of twelve probability distributions.

The transition from discrete to continuous infirmity can be justified as follows. The continuous distribution is approximated by discrete quantization (divided into intervals of size Δy) to form a discrete probability system: $\{Y_\Delta; p(y)\Delta y\}$. The discrete infirmity of Y_Δ is given by,

$$\beta(Y_\Delta) = \sum_{i=1}^N [p(y_i)\Delta y]^2 = \Delta y \sum_{i=1}^N [p(y)]^2 \Delta y. \quad (7)$$

Normalizing $\beta(Y_\Delta)$ by Δy and then taking the limit of the normalized $\beta(Y_\Delta)$ yields

$$\lim_{\Delta y \rightarrow 0} \frac{\beta(Y_\Delta)}{\Delta y} = \lim_{\Delta y \rightarrow 0} \sum_{i=1}^N [p(y)]^2 \Delta y = \int [p(y)]^2 dy = \beta(Y). \quad (8)$$

This procedure shows that while continuous infirmity is analogous to discrete infirmity, it is not exactly a limit in the strict sense, but rather a renormalized version that accounts for the continuous nature of the underlying probability space. Note that this process is essentially similar to the transition from discrete to continuous (differential) entropy.

Table 1

Continuous informities of twelve probability distributions			
Distribution	Probability density function $p(y)$	Informity $\beta(Y)$	Support
Uniform	$\frac{1}{2a}$	$\frac{1}{2a}$	$[-a, a]$
Triangular	$\begin{cases} \frac{(y+a)}{a^2}, & -a \leq y \leq 0 \\ \frac{(a-y)}{a^2}, & 0 \leq y \leq a \end{cases}$	$\frac{2}{3a}$	$[-a, a]$
Quadratic	$\frac{3}{4a} \left[1 - \left(\frac{y}{a} \right)^2 \right]$	$\frac{3}{5a}$	$[-a, a]$
Raised cosine	$\frac{1}{2a} \left[1 + \cos \left(\frac{\pi}{a} y \right) \right]$	$\frac{3}{4a}$	$[-a, a]$
Half-cosine	$\frac{\pi}{4a} \cos \left(\frac{\pi}{2a} y \right)$	$\frac{\pi^2}{16a}$	$[-a, a]$
Normal	$\frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{y-\mu}{\sigma} \right)^2 \right]$	$\frac{1}{2\sigma\sqrt{\pi}}$	$(-\infty, \infty)$
Exponential	$\lambda \exp(-\lambda y)$	$\frac{\lambda}{2}$	$[0, \infty)$
Cauchy	$\frac{1}{\pi y \left[1 + \frac{1}{y^2} (y - y_0)^2 \right]}$	$\frac{1}{2\pi y}$	$(-\infty, \infty)$
Rayleigh	$\frac{y}{\sigma^2} \exp \left(-\frac{y^2}{2\sigma^2} \right)$	$\frac{\sqrt{\pi}}{4\sigma}$	$[0, \infty)$
Gamma	$\frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \exp(-\beta y)$	$\frac{\beta}{[\Gamma(\alpha)]^2} 2^{-(2\alpha-1)} \Gamma(2\alpha-1)$	$[0, \infty)$
Laplace	$\frac{1}{2b} \exp \left(-\frac{ y-\mu }{b} \right)$	$\frac{1}{4b}$	$(-\infty, \infty)$
Weibull	$\frac{k}{\lambda^k} y^{k-1} \exp \left(-\frac{y^k}{\lambda^k} \right)$	$\frac{k}{\lambda} 2^{-\frac{2k-1}{k}} \Gamma \left(\frac{2k-1}{k} \right)$	$[0, \infty)$

Just like the continuous entropy $H(Y)$ does not have all the properties of the discrete entropy $H(X)$, the continuous informity $\beta(Y)$ does not have all the properties of the discrete informity $\beta(X)$. Unlike $\beta(X)$, which is dimensionless, $\beta(Y)$ is not dimensionless. $\beta(Y)$ has dimensions of $[y]^{-1}$. If the quantity Y has physical units of length, the inverse of $\beta(Y)$ also has units of length. Sussmann (1997) introduced a quantity called "uncertainty length" in quantum mechanics.

Definition 3. The uncertainty length is given by the reciprocal of the continuous informity $\beta(Y)$

$$\delta(Y) := \frac{1}{\beta(Y)} = \frac{1}{\int [p(y)]^2 dy}. \quad (9)$$

For the normal distribution $N(\mu, \sigma^2)$, $\delta(Y) = 2\sigma\sqrt{\pi} = 3.545\sigma$. Therefore, the uncertainty length $\delta(Y)$ is proportional to the standard deviation σ .

2. The definitions of cross-informity and joint informity

Definition 4. The cross-informity of two discrete random variables X_1 and X_2 with PMFs $P_1(x)$ and $P_2(x)$ is given by

$$\beta(X_1 \cap X_2) := \sum_{i=1}^N P_1(x_i) P_2(x_i) = E_{P_1}[P_2(x)] = E_{P_2}[P_1(x)], \quad (10)$$

where $E_{P_1}[P_2(x)]$ is the expected value of $P_2(x)$ with respect to $P_1(x)$, and $E_{P_2}[P_1(x)]$ is the expected value of $P_1(x)$ with respect to $P_2(x)$. Since $\beta(X_1 \cap X_2) = \beta(X_2 \cap X_1)$, the cross-informity is symmetric.

Definition 5. The cross-informity of two continuous random variables Y_1 and Y_2 with PDFs $p_1(y)$ and $p_2(y)$ is given by

$$\beta(Y_1 \cap Y_2) := \int p_1(y) p_2(y) dy = E_{p_1}[p_2(y)] = E_{p_2}[p_1(y)], \quad (11)$$

where $E_{p_1}[p_2(y)]$ is the expected value of $p_2(y)$ with respect to $p_1(y)$, and $E_{p_2}[p_1(y)]$ is the expected value of $p_1(y)$ with respect to $p_2(y)$. The symmetry also holds, i.e. $\beta(Y_1 \cap Y_2) = \beta(Y_2 \cap Y_1)$.

Definition 6. Like $\beta(Y)$, $\beta(Y_1 \cap Y_2)$ has dimensions of $[y]^{-1}$. Thus, similar to the definition of uncertainty length, the discrepancy length is given by the reciprocal of the cross-informity $\beta(Y_1 \cap Y_2)$

$$\delta(Y_1 \cap Y_2) := \frac{1}{\beta(Y_1 \cap Y_2)} = \frac{1}{\int p_1(y) p_2(y) dy}. \quad (12)$$

If the distributions of Y_1 and Y_2 are identical, the discrepancy length $\delta(Y_1 \cap Y_2)$ reduces to the uncertainty length $\delta(Y)$.

Definition 7. The joint informity of two discrete random variables X_1 and X_2 is given by

$$\beta(X_1, X_2) := \sum_{i=1}^N \sum_{j=1}^N [P(x_{1,i}, x_{2,j})]^2, \quad (13)$$

where $P(x_{1,i}, x_{2,j})$ is the joint PMF of X_1 and X_2 . If X_1 and X_2 are independent,

$$\beta(X_1, X_2) = \sum_{i=1}^N [P_1(x_i)]^2 \sum_{j=1}^N [P_2(x_j)]^2 = \beta(X_1) \beta(X_2). \quad (14)$$

Definition 8. The joint informity of two continuous random variables Y_1 and Y_2 is given by

$$\beta(Y_1, Y_2) := \iint [p(y_1, y_2)]^2 dy_1 dy_2, \quad (15)$$

where $p(y_1, y_2)$ is the joint PDF of Y_1 and Y_2 . If Y_1 and Y_2 are independent,

$$\beta(Y_1, Y_2) = \iint [p_1(y_1)]^2 [p_2(y_2)]^2 dy_1 dy_2 = \beta(Y_1) \beta(Y_2). \quad (16)$$

The joint informity $\beta(Y_1, Y_2)$ has dimensions of $[y]^{-2}$. As a two-dimensional analogy to the one-dimensional "uncertainty length", Sussmann (1997) introduced a quantity called "uncertainty area" for the Wigner quasi-distribution in quantum mechanics.

The justification for the transition from discrete to continuous cross-informity or joint informity is similar to the procedure of the transition from discrete to continuous informity described in Section 2, so it is not repeated here.

Definition 9. The uncertainty area is given by the reciprocal of the joint informity $\beta(Y_1, Y_2)$

$$\delta(Y_1, Y_2) := \frac{1}{\beta(Y_1, Y_2)} = \frac{1}{\iint [p(y_1, y_2)]^2 dy_1 dy_2}. \quad (17)$$

If Y_1 and Y_2 are independent, the uncertainty area is the product of the two uncertainty lengths

$$\delta(Y_1, Y_2) = \frac{1}{\beta(Y_1) \beta(Y_2)} = \delta(Y_1) \delta(Y_2). \quad (18)$$

3. Informity metric

Informity can be used as a metric to compare statistical distributions (models) or splits (in machine learning). Informity metric provides an alternative to entropy metric or information criteria based on maximum likelihood, such as Akaike information criterion (AIC) and Bayesian information criterion (BIC).

Proposition 1. Suppose that among a set of candidate distributions, we want to select one that best fits the given data or information. Let $p_k(y)$ denote the PDF of the k -th candidate distribution and $\beta_k(Y)$ denote the corresponding informity. Then, the best distribution is the one with the maximum informity. That is,

$$\text{best distribution} = \arg \max_k [\beta_k(Y)], \quad k = 1, 2, 3 \dots M, \quad (19)$$

where M is the number of candidate distributions.

We designate Proposition 1 as "the maximum informity criterion". The maximum informity criterion can be seen as the counterpart of the minimum entropy criterion proposed by Huang (2024), and in fact it can be derived from the minimum entropy criterion because informity is inversely proportional to entropy. In principle, the informity and entropy metrics are composable. However, the informity metric does not necessarily give exactly the same results as the entropy metric. This is because the continuous informity is computed in the original probability space, while the continuous entropy is computed in the log-transformed probability space.

Definition 10. Suppose a set is split into several disjoint subsets. The sum of the elements in the subsets is equal to the elements in the original set. The informity gain $\Delta\beta$ is the difference in informity before and after the split. That is,

$$\Delta\beta := \beta_{\text{after}}(X) - \beta_{\text{before}}(X), \quad (20)$$

where $\beta_{\text{before}}(X)$ is the informity before the split and $\beta_{\text{after}}(X)$ is the informity after the split. $\beta_{\text{after}}(X)$ is calculated as the weighted average of the informities of the subsets.

Proposition 2. When training decision trees in machine learning, consider M candidate splits for a mixed class (the original set). Let $\Delta\beta_k$ denote the informity gain of the k -th candidate split. The best split among the M candidate splits is the one with the maximum informity gain. That is,

$$\text{best split} = \arg \max_k [\Delta\beta_k], \quad k = 1, 2, 3 \dots M. \quad (21)$$

We designate Proposition 2 as "the maximum informity gain criterion". The maximum informity gain criterion can be seen as the counterpart of the "information gain" criterion based on the maximum entropy reduction (e.g. Zhou, 2019a).

4. Discussion

The continuous informity has the opposite meaning from the continuous entropy. They provide insights into a probability system from different perspectives.

Figure 2 shows the continuous informity and continuous entropy of the normal distribution as a function of the standard deviation σ (scale).

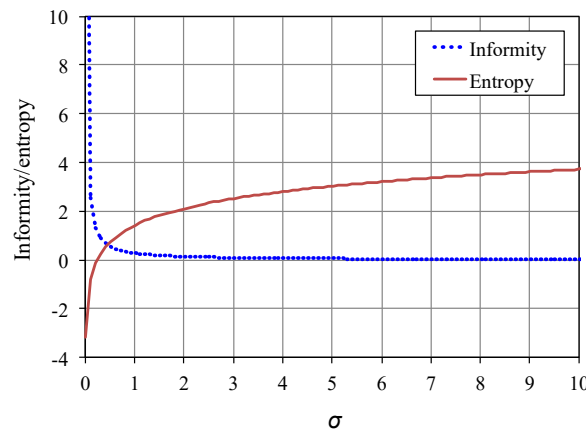


Fig. 2. Continuous informity and continuous entropy of normal distribution as a function of the standard deviation σ (scale)

It can be seen from Fig. 2 that, as the scale parameter σ increases, the continuous informity decreases and the continuous entropy increases. When the continuous informity is very large, the continuous entropy becomes negative.

Remark 1. When the population parameters (e.g. location and scale parameters) are known, the true value of the continuous informity can be obtained. When the population parameters are unknown and estimated from a sample, the continuous informity can be calculated based on the estimated parameter values and becomes a statistic. Therefore, using the proposed informity theory in practice may involve point estimates of population parameters.

Consider a sample (n observations) drawn from the normal distribution of Y with mean μ and standard deviation σ . When μ and σ are known, the true value of the informity $\beta(Y)$ is $1/(2\sigma\sqrt{\pi})$. When μ and σ are unknown, the sample mean $\bar{y}$ is an unbiased estimate of μ and $s/c_{4,n}$ is an unbiased estimate of σ , where s is the sample standard deviation and $c_{4,n}$ is the bias-correction factor; $c_{4,n} = \sqrt{\frac{2}{n-1} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}}$; $\Gamma(\cdot)$ stands for Gamma function (Wadsworth, 1989). Accordingly, the estimated informity is

$$\hat{\beta}(Y) = \frac{c_{4,n}}{2s\sqrt{\pi}}. \quad (22)$$

Remark 2. In this paper, the sum of the squares of PMF $\sum_{i=1}^N P(x_i)^2$ is defined as the discrete informity and the integral of the square of PDF $\int [p(y)]^2 dy$ is defined as the continuous informity. The author found through online search that these two quantities or their related quantities have different names and uses in some literature. The sum of the squares of PMF is called "repeat rate" or the Simpson index, which is a measure of concentration with a long history (Rousseau, 2018). The quantity (1- repeat rate) is called Gini impurity, which measures the disorder or impurity of a dataset. It is also known as logical entropy, which is interpreted as "the probability that a distinction or dit (elements in different blocks) is obtained in two independent draws from the underlying set" (Ellerman, 2013; 2022). On the other hand, Lad et al. (2015) defined the *negative half* of the integral of the square of PDF as differential *extropy*, i. e. $J(Y) = -\frac{1}{2} \int [p(y)]^2 dy$. However, this definition of differential *extropy* is inconsistent with their definition of discrete *extropy*: $J(X) = -\sum_{i=1}^N [1 - P(x_i)] \log_2 [1 - P(x_i)]$. Note that if $P(x_i)$ is replaced by the density value $p(y)$, the differential *extropy* would be undefined when $p(y) > 1$. Lad et al. (2015) acknowledged this problem and used an approximation: $J(X) \approx 1 - \frac{1}{2} \sum_{i=1}^N P(x_i)^2$. However, we argue that it is questionable to deduce $J(Y) = -\frac{1}{2} \int [p(y)]^2 dy$ from the approximation $J(X) \approx 1 - \frac{1}{2} \sum_{i=1}^N P(x_i)^2$. Moreover, puzzlingly, $J(Y)$ is always negative. This casts doubt about the true meaning of $J(Y)$.

Although these two quantities are not new, their specific perspective as the average amount of information contained in a probability system is novel and unique. Therefore, the idea of informity is new and the proposed informity theory is a novel probability framework.

5. Application examples

5.1. Selecting distribution of calibration error

Consider the calibration error of a measuring instrument. The calibration error is confined within the interval $[-a, +a]$, but has a central tendency. In measurement uncertainty analysis, we want to select the best distribution among a set of candidate distributions for the calibration error. Huang (2024) considered five candidate distributions: uniform, triangular, quadratic, raised cosine, and half-cosine. He determined that the raised cosine distribution is the best distribution according to the minimum entropy criterion.

We evaluate these five candidate distributions using the informity metric. Similar to defining the "unit entropy" by setting $a = 1$ in Huang (2024), we define "unity informity" by setting $a = 1$ in the informity expressions of these five distributions shown in Tab. 1. The results for unit informity are shown in Tab. 2, together with the unit entropy.

Table 2

Unit informities and unit entropies of five candidate distributions		
Distribution	Unit informity	Unit entropy (nats)
Uniform	0.5	0.693
Triangular	0.667	0.5
Quadratic	0.6	0.567
Raised cosine	0.75	0.386
Half-cosine	0.617	0.548

As can be seen from Tab. 2, the raised cosine distribution has the highest unit informity and the lowest unit entropy among the five candidate distributions. Therefore, it is the best distribution according to the maximum informity criterion and the minimum entropy criterion. This result is consistent with Castrup's (2004) argument based on physical reality considerations.

Note from Tab. 2 that, as expected, the uniform distribution has the lowest unit informity and the highest unit entropy. It is the least informative distribution among the five candidate distributions from both the informity and entropy perspectives.

5.2. Evaluating four distributions given a sample drawn from a Weibull distribution

Xie (2011) evaluated four candidate distributions: exponential, Weibull, Gamma, and normal to fit a sample ($n=200$) randomly drawn from a Weibull distribution with shape 1.8 and scale 1, using Akaike information criterion (AIC) and Pearson's Q-statistic. He estimated the parameters using the routine maximum likelihood estimation procedure, which gives: exponential distribution: rate = 1.167, Weibull distribution: shape = 1.662 and scale = 0.9588, Gamma distribution: shape = 2.255 and scale = 0.3081, and normal distribution: mean = 0.8570 and standard deviation = 0.5277. Huang (2024) considered this example using the entropy metric. We consider this example using the informity metric. Table 3 shows the results obtained using the four different metrics.

As can be seen from Tab. 3, the Weibull distribution is the best distribution according to the AIC (Xie, 2011) or the minimum entropy criterion (Huang, 2024). The Gamma distribution is the best distribution according to the Q-statistic (Xie, 2011) or

the maximum informity criterion (this study). Xie (2011) pointed out that the Gamma distribution was also significantly supported by the data, as evidenced by its low ΔAIC value 4.7. Xie (2011) also pointed out, "AIC selects a "good" model which is not necessarily the true model". His argument should also apply to the other three metrics. Therefore, we can say that the four metrics give consistent results for this example.

Table 3

Evaluation of four candidate distributions using four different metrics				
Distribution	AIC (Xie, 2011)	Q-statistic (Xie, 2011)	Entropy (nats) (Huang, 2024)	Informity (this study)
Exponential	340.3	48.1	0.846	0.5835
Weibull	276.6	7.07	0.680	0.5835
Gamma	281.3	6.34	0.694	0.6010
Normal	315.9	27.6	0.780	0.5346

5.3. Informity gain when splitting a mixed class

Zhou (2019a) considered an example of splitting a mixed class (a set) of 5 blues and 5 greens. He split the class into two branches (two subsets): the left branch had 4 blues and the right branch had 1 blue and 5 greens. He calculated that the entropy before the split was 1 and after the split was 0.39. The information gain (IG), defined as the difference in entropy before and after the split, is 0.61 bits.

We consider this example using the informity gain. Before the split, the informity of the mixed class $\beta_{\text{before}}(X) = 0.5$. After the split, $\beta_{\text{after}}(X) = 0.8333$. Then, the informity gain $\Delta\beta = 0.3333$.

Zhou (2019b) also considered this example using Gini impurity metric. Gini impurity (denoted by G) can be written as

$$G(X) := \sum_{i=1}^N P(x_i)[1 - P(x_i)] = 1 - \sum_{i=1}^N P(x_i)^2. \quad (23)$$

Gini gain ΔG is defined as

$$\Delta G := G_{\text{before}}(X) - G_{\text{after}}(X). \quad (24)$$

Note that Gini impurity G is related to the discrete informity as $G(X) = 1 - \beta(X)$. Therefore, the Gini gain ΔG is the same as the informity gain $\Delta\beta$. For this example, $\Delta G = \Delta\beta = 0.3333$.

Discussion and conclusions

The proposed informity theory (a novel probability framework) is established in the original probability space; it does not involve the log-transformation of probabilities like Shannon information theory and is more applicable to fields other than communications engineering. The mathematical quantity "informity" provides a quantitative measure of the degree of informativeness of a probability system. The idea of informity is intuitive and consistent with our common sense or perceptual understanding about information or informativeness. Informity is an important property of a probability distribution. It is the counterpart of entropy and provides insights into probability distributions from a different perspective than entropy. We have derived the informities for twelve continuous distributions. Informities for many more distributions need to be derived. Furthermore, just as the entropy for a distribution is often included in Wikipedia, the informity for the distribution may also be included in Wikipedia.

The proposed informity metric can be used as an alternative to entropy metric. For a given set of data or information, the proposed maximum informity criterion can be used to select the best distribution among a set of candidate distributions. When training decision trees in machine learning, the proposed maximum informity gain criterion can be used to select the best split among a set of candidate splits. The three examples presented demonstrate the practicability of the proposed informity metric.

The author envisions that the proposed informity theory will have broad application prospects. Currently, the author has used the proposed informity theory to develop a new index for measuring the tail-heaviness of a probability distribution (called the tail-heaviness index), two new indices for measuring the difference between two probability distributions (one called the "distribution similarity index (DSI)" and the other called the "distribution discrepancy index (DDI)"), and a new index for measuring the peakedness of a probability distribution (called the peakedness index). The manuscripts describing these developments have been posted as preprints on ResearchGate. Further research is needed to extend and enhance the proposed informity theory and its applications.

Appendix

Converting the *relative* quantity "probability" into the *absolute* quantity "Shannon information content"

Consider a discrete random variable X with the PMF $P(x)$ that has N possible outcomes. The total amount of information contained in the probability system $\{X; P(x)\}$ is normalized to 1, i. e. $\sum_{i=1}^N P(x_i) = 1$. We view probability $P(x_i)$ as the *relative* information content of the outcome x_i . Let n_i denote the reciprocal of $P(x_i)$, i. e. $n_i = \frac{1}{P(x_i)}$. This means that the outcome x_i has one chance of occurring out of n_i possible outcomes. In Shannon information theory, n_i is "the number of messages in the set" and "... this number or any monotonic function of this number can be regarded as a measure of the information [content] produced when one message is chosen from the set, all choices being equally likely (Shannon, 1948)." Therefore, either n_i or any monotonic function of n_i can be regarded as the *absolute* amount of information conveyed (or held) by the outcome x_i , i. e. the *absolute* information content of the outcome x_i .

The *absolute* amount of information contained in the probability system $\{X; P(x)\}$ can be calculated as

$$\sum_{i=1}^N P(x_i)n_i = \sum_{i=1}^N \frac{1}{n_i}n_i = N. \quad (25)$$

At first glance, this equation seems trivial. However, if we take the logarithm of n_i (note: logarithm is a monotonic function), Eq. (25) becomes the expression for entropy

$$H(X) = \sum_{i=1}^N P(x_i) \log_2(n_i) = - \sum_{i=1}^N P(x_i) \log_2 P(x_i). \quad (26)$$

It is important to note that the logarithm of a quantity *only* changes the scale of the quantity; it *does not* change the nature or innate character of the quantity. Therefore, both Shannon information content $\log_2(n_i) = -\log_2 P(x_i)$ and entropy $H(X)$ are the *absolute* quantities.

However, for a continuous random variable Y with the PDF $p(y)$, we cannot define the absolute information content of a y value. But we can define the absolute information content of a probability interval as the reciprocal of the coverage probability of the probability interval. For example, for a coverage probability of 0.9, the absolute amount of information contained in the probability interval is $1/0.9=11.11$, or $-\log_2 0.9 = 0.152$ bits.

References

- Castrup, H. (2004). Selecting and applying error distributions in uncertainty analysis. *Measurement Science Conference*. Anaheim. http://www.isgmax.com/Articles_Papers/Selecting%20and%20Applying%20Error%20Distributions.pdf
- Ellerman, D. (2013). An Introduction to Logical Entropy and its Relation to Shannon Entropy. *International Journal of Semantic Computing*, 7(2), 121–145. <https://doi.org/10.1142/S1793351X13400059>
- Ellerman, D. (2022). Introduction to logical entropy and its relationship to Shannon entropy. *4open*, 5, 1–33. <https://doi.org/10.1051/fopen/2021004>
- Huang, H. (2024). A minimum entropy criterion for distribution selection for measurement uncertainty analysis. *Measurement Science and Technology*, 35, 035014. <https://iopscience.iop.org/article/10.1088/1361-6501/ad1476>
- Lad, F., Sanfilippo, G. & Agrò, G. (2015). Extropy: Complementary Dual of Entropy. *Statistical Science*, 30(1), 40–58. <https://doi.org/10.1214/14-STS430>
- Petty, G. W. (2018). On some shortcomings of Shannon entropy as a measure of information content in indirect measurements of continuous variables. *Journal of Atmospheric and Oceanic Technology*, 35(5), 1011–1021. <https://doi.org/10.1175/JTECH-D-17-0056.1>
- Rousseau, R. (2018). The repeat rate: from Hirschman to Stirling. *Scientometrics*, 116, 645–653. <https://doi.org/10.1007/s11192-018-2724-8>
- Shannon, C. E. (1948). A mathematical theory of communications. *The Bell System Technical Journal*, 27(3), 379–423.
- Schroeder, M. J. (2004). An alternative to entropy in the measurement of information. *Entropy*, 6, 388–412.
- Sussmann, G. (1997). Uncertainty relation: from inequality to equality. *Zeitschrift für Naturforschung A*, 52, 1–2. <https://doi.org/10.1515/zna-1997-1-214>
- Wadsworth, Jr. H. M. (1989). Summarization and interpretation of data. In Harrison M Wadsworth Jr. (Ed.), *Handbook of Statistical Methods for Engineers and Scientists* (2.1–2.21). McGRAW-HILL Inc.
- Xie, G. (2011). *Further developments of two point process models for fine-scale time series*. Massey University, New Zealand.
- Zhou, V. (2019a). A simple explanation of information gain and entropy. <https://victorzhou.com/blog/information-gain>
- Zhou, V. (2019b). A simple explanation of Gini impurity. <https://victorzhou.com/blog/gini-impurity>

Отримано редакцією журналу / Received: 19.01.25

Прорецензовано / Revised: 13.03.25

Схвалено до друку / Accepted: 14.03.25

Хенінг ХУАНГ, д-р філософії, наук. співроб.

ORCID ID: 0000-0003-1882-0042

e-mail: heninghuang1@gmail.com

Теледін РД Інструментс, Сан-Дієго, США

ТЕОРІЯ ІНФОРМАТИВНОСТІ: НОВИЙ ПІДХІД ДО ТЕОРІЇ ЙМОВІРНОСТЕЙ

Запропоновано новий підхід до теорії ймовірностей, що отримав назву теорія інформативності. Для кількісного вимірювання рівня інформативності ймовірнісного розподілу (або системи ймовірностей) визначено математичну величину під назвою "інформативність". Також встановлено дві інші величини: перехресну інформативність і спільну інформативність. Запропоновано метрику інформативності, яку можна використовувати як альтернативу метриці ентропії. Наведено значення інформативності для дванадцяти неперервних розподілів. Подано три приклади, що демонструють практичність запропонованої метрики інформативності.

Ключові слова: інформативність, інформаційний контент, ймовірність, розподіл ймовірностей.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

ДИФЕРЕНЦІАЛЬНІ РІВНЯННЯ, МАТЕМАТИЧНА ФІЗИКА ТА МЕХАНІКА

УДК 539.375

DOI: <https://doi.org/10.17721/1812-5409.2025/1.8>

Олександр КІПНІС, канд. фіз.-мат. наук, ст. наук. співроб.

ORCID ID: 0000-0001-6747-8584

e-mail: a.l.kipnis@gmail.com

Інститут механіки ім. С. П. Тимошенка НАН України, Київ, Україна

СТИСНЕННЯ ШАРУВАТОГО КОМПОЗИТУ З КОМПОНЕНТАМИ, ЩО ПРОКОВЗУЮТЬ, УЗДОВЖ ДВОХ ПАРАЛЕЛЬНИХ МІЖФАЗНИХ СТРУКТУРНИХ ТРІЩИН

Досліджено задачу лінеаризованої теорії стійкості про стиснення кусково-однорідного тіла, що складається з нелінійно пружної смуги, яка без тертя проковзує між двома нелінійно пружними півплощинами, та послаблене двома паралельними тріщинами, розташованими на прямолінійних межах поділу середовищ. Застосовано апробований аналітико-числовий метод, що полягає у зведенні вихідної крайової задачі з використанням загальних представлень розв'язків лінеаризованих рівнянь рівноваги через потенціальні гармонічні функції до інтегрального рівняння Фредгольма першого роду. Зазначене розв'язальне інтегральне рівняння задачі одержано в загальній формі для широкого класу комбінацій двох стисливих або нестисливих матеріалів, для пружних потенціалів яких реалізується випадок рівних коренів характеристичних рівнянь, досліджених чисельно з використанням методу Бубнова – Гальоркіна у випадку, коли високоеластичні матеріали тіла описують потенціалом гармонічного типу. Визначено критичні значення параметрів навантаження, що відповідають старту руйнування такого тіла, та досліджено характер їхньої залежності від фізико-механічних і геометричних параметрів тіла.

Ключові слова: шаруватий композитний матеріал, міжфазна тріщина, високоеластичні матеріали, критерій руйнування, потенціал гармонічного типу.

Класифікація відповідно до AMS 2020: 74R99, 74G60.

Вступ

Наявність тріщин на межі поділу середовищ в елементах конструкцій, виконаних із шаруватих композитних матеріалів, може суттєво впливати на параметри їхньої міцності та надійності, обраховані за припущення відсутності міжфазних дефектів. Відповідні задачі механіки руйнування матеріалів найменш досліджені у випадку, коли навантаження до таких елементів конструкції прикладене паралельно до площин розташування тріщин, оскільки за вказаної геометрії навантаження незастосовними (Guz, 2014) є класичні критерії руйнування (Griffith, 1920; Irwin, 1957) та сучасні моделі для зон передруйнування (Камінський, Дудик, & Решітник, 2023).

Для дослідження означених неklasичних проблем руйнування в межах лінеаризованої теорії стійкості (Guz, Bogdanov, & Nazarenko, 2020) був запропонований критерій, який асоціює старт процесу руйнування тіл у процесі їхнього стиснення вздовж тріщин із втратою стійкості рівноваги локальної частини матеріалу, що оточує тріщину. Детальну бібліографію стосовно задач руйнування, досліджених із застосуванням цього критерію, подано, наприклад, в оглядовій роботі (Guz, 2014).

Нижче в межах тривимірної лінеаризованої теорії стійкості деформівних тіл досліджено задачу плоскої деформації про стиснення необмеженого шаруватого композиту, компоненти якого проковзують одне відносно одного без тертя, уздовж двох паралельних міжфазних структурних тріщин однакової довжини (рис. 1, а). Визначено критичні значення параметрів навантаження та досліджено їхню залежність від фізико-механічних і геометричних параметрів композиту.

1. Постановка задачі

Розглянемо необмежене кусково-однорідне тіло періодичної структури, утворене композицією шарів двох різних нелінійно пружних матеріалів "1" і "2", що перебуває в умовах стиснення вздовж прямолінійних меж поділу середовищ, що проковзують одне відносно одного без тертя, та послаблене двома паралельними міжфазними тріщинами однакової довжини (рис. 1а). Нехай товщина d_1 шару, виконаного з матеріалу "1", суттєво більша за товщину d_2 шару, виконаного з матеріалу "2" та за довжину $2a$ тріщин. Відповідно до термінології (Гузь, 2008), зазначений спосіб з'єднання шарів композиту відповідає "найслабшому" типу зв'язку між компонентами кусково-однорідного тіла, а міжфазні тріщини, за поданих співвідношень між геометричними параметрами можна характеризувати як "структурні".

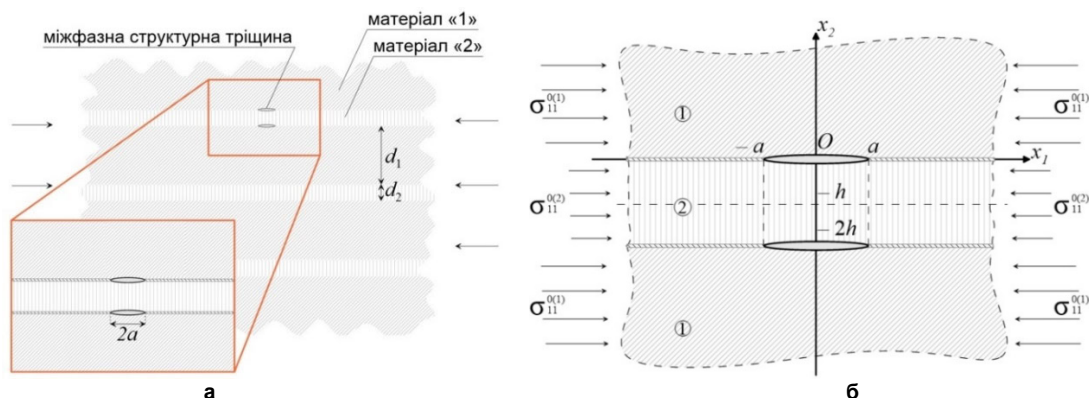


Рис. 1. Шаруватий композит в умовах стиснення вздовж двох міжфазних структурних тріщин

Ураховуючи малість тріщин приходимо до задачі плоскої деформації для нескінченної смуги $-2h \leq x_2 \leq 0$ ($d_2 = 2h$), виконаної з нелінійно пружного матеріалу "2", що проковзує без тертя між півплощинами $x_2 \geq 0$ і $x_2 \leq -2h$, виконаними з нелінійно пружного матеріалу "1". Прямолинійні межі поділу середовищ $x_2 = 0$ і $x_2 = -2h$ містять міжфазні тріщини завдовжки $2a$ (див. рис. 1б).

Нехай на нескінченності матеріали стискаються вздовж осі Ox_1 рівномірно розподіленими навантаженнями

$$\sigma_{11}^{(i)} = \text{const}, i = 1, 2; \sigma_{11}^{(1)} \neq \sigma_{11}^{(2)} \quad (1)$$

таким способом, щоб гарантувати однакові видовження вздовж осі Ox_1 для матеріалів півплощини та смуги:

$$\lambda_1^1 = \lambda_1^2 = \lambda_1 = \text{const}, \lambda_1 < 1, \quad (2)$$

де λ_1^1, λ_1^2 – коефіцієнти укорочення матеріалів "1" та "2", що обумовлені стискувальними зусиллями $\sigma_{11}^{(1)}$ та $\sigma_{11}^{(2)}$, відповідно.

У цьому випадку докритичний напружено-деформований стан (НДС) у кожній із компонент кусково-однорідного тіла буде статично визначеним, однорідним і визначається виразами для переміщень

$$u_i^{(i)} = (\lambda_i - 1)x_i, i = 1, 2. \quad (3)$$

Тут і далі верхнім індексом "1" або "2" позначені величини для відповідного матеріалу, верхнім індексом "0" – величини, що належать до початкового (докритичного, незбуреного) стану, а збурення цих величин не позначаються додатковим індексом.

Обмежуючись розглядом згинальної (Гузь, 2008; Guz, Nazarenko, & Starodubtsev, 1991) форми втрати стійкості та беручи до уваги симетрію конфігурації рис. 1б відносно лінії $x_2 = -h$, крайові умови сформульованої задачі запишемо так:

$$\begin{aligned} u_1^{(2)} &= 0, t_{22}^{(2)} = 0 \quad (x_2 = -h, 0 \leq |x_1| < \infty); \\ t_{22}^{(1)} &= t_{22}^{(2)}, t_{21}^{(1)} = 0, t_{21}^{(2)} = 0 \quad (x_2 = 0, 0 \leq |x_1| < \infty); \\ t_{22}^{(2)} &= 0 \quad (x_2 = 0, |x_1| \leq a); u_2^{(1)} = u_2^{(2)} \quad (x_2 = 0, |x_1| > a). \end{aligned} \quad (4)$$

У (4) $t_{kl}^{(i)}, i, k, l = 1, 2$ – збурення компонент несиметричного тензору напружень Піоли – Кірхгофа $\tilde{t}^{(i)}, \tilde{u}^{(i)}$ – вектор збурення переміщень.

Зазначимо, що задачу, аналогічну до тієї, що зображена на рис. 1б для випадку ідеального контакту між компонентами тіла, розглянуто в (Guz, 1994). З використанням чисельного методу скінченних різниць визначені критичні значення параметрів навантаження для багатьох конкретних комбінацій двох лінійно пружних матеріалів.

2. Розв'язальне інтегральне рівняння

Перевагою чисельно-аналітичного методу, що пропонуємо використовувати у розв'язанні поставленої задачі в цій роботі, є можливість одержання розв'язального інтегрального рівняння в загальній формі для комбінації двох різних нелінійно пружних матеріалів з довільною структурою пружного потенціалу, а конкретизацію моделей матеріалів проводити лише на фінальному етапі чисельного дослідження інтегрального рівняння. Цей метод базується на підході, який був запропонований для дослідження аналогічних задач про стиснення *однорідного* тіла вздовж приповерхневої (Nazarenko, 1986) та двох паралельних тріщин (Guz, Nazarenko, & Starodubtsev, 1991). Згодом зазначений підхід було поширено на випадок міжфазних тріщин (Bogdanov, Nazarenko, & Kipnis, 2024a; 2024b) і зон проковзування (Bogdanov, Nazarenko, & Kipnis, 2025) в напівобмежених *кусково-однорідних* тілах із шаром покриття.

Умови однорідності докритичного НДС (1) – (3) є необхідними умовами застосування загальних представлень розв'язків лінеаризованих рівнянь рівноваги через потенціальні гармонічні функції (Guz, Bogdanov, & Nazarenko, 2020). Використовуючи ці представлення та застосовуючи апарат інтегрального перетворення Фур'є за аналогією до роботи (Bogdanov, Nazarenko, & Kipnis, 2024a), вихідна крайова задача з умовами (4) зводиться до задачі на власні значення для наступного інтегрального рівняння Фредгольма першого роду відносно невідомої безрозмірної функції $f(\eta)$:

$$\begin{aligned} \int_0^1 K(\xi, \eta) f(\eta) d\eta &= 0, \beta = h/a, 0 \leq \xi < 1, 0 \leq \eta \leq 1; \\ K(\xi, \eta) &= \int_0^\infty \frac{\delta_1(n^{-1/2}\beta\lambda)}{\delta(n^{-1/2}\beta\lambda)} \frac{(\cos \lambda \eta - \cos \lambda) \cos \lambda \xi}{\lambda} d\lambda, \\ \delta_1(\mu) &= k + \mu(\coth \mu - \coth^{-1} \mu), \delta(\mu) = r_2 \coth \mu - r_1 \delta_1(\mu), \\ k &= \frac{k_2^2 k_4^2 - k_1^2 k_5^2}{k_2^2 k_5^2}, l = \frac{k_2^1 k_4^1 - k_1^1 k_5^1}{k_2^1 k_5^1}, r_1 = l^{-1} \left(p_2^1 \frac{k_1^1}{k_2^1} - p_1^1 \right), r_2 = p_2^2 \frac{k_1^2}{k_2^2} - p_1^2, \end{aligned} \quad (5)$$

де l, k_i^j, p_i^j – відомі величини, що визначені в (Bogdanov, Nazarenko, & Kipnis, 2024a) і залежать від типу пружного потенціалу складових компонентів тіла.

Наведене інтегральне рівняння (5) одержане в загальному вигляді для комбінації двох стисливих або нестисливих матеріалів, для пружних потенціалів яких справджується випадок рівних коренів характеристичних рівнянь (Guz, Bogdanov, & Nazarenko, 2020).

Отже, вихідна лінеаризована задача із граничними умовами (4) зводиться до задачі (5) на власні значення відносно параметра укорочення $\lambda_1 < 1$ (значення $\lambda_1 = 1$ відповідає недеформованому стану). Параметр λ_1 характеризує

докритичний стан і нелінійним чином входить до ядра інтегрального рівняння (5). Іншими словами, розшукується перше значення параметра $\lambda_1 < 1$ (за зменшення λ_1 , починаючи зі значення $\lambda_1 = 1$, що відповідає неперервному збільшенню зовнішнього стискувального навантаження), за якого інтегральне рівняння Фредгольма першого роду (5) має неєдиний розв'язок.

3. Аналіз числових результатів

Як приклад розглянемо випадок, коли матеріали "1" і "2" описують двоконстантним пружним потенціалом *гармонічного типу* для стисливих тіл.

Потенціал гармонічного типу характеризується двома сталими: μ_i і ν_i для кожного з матеріалів тіла, що визначають, відповідно, жорсткість матеріалу та його здатність до стисливості. Сталі μ_i є аналогом модуля жорсткості в лінійній теорії пружності; сталі ν_i – аналогом коефіцієнта Пуассона.

Числове дослідження реалізовано з використанням методу Бубнова – Гальоркіна.

На рис. 2а зображено залежність критичного значення відносного укорочення $\varepsilon_1 = 1 - \lambda_1$ в задачі, що розглядається, від значення відносної (нормованої на довжину тріщини) ширини смуги $\beta = h/a$ для різних значень $g = \mu_2/\mu_1$ відношення жорсткості матеріалу "2" до жорсткості матеріалу "1" за значень $\nu_1 = \nu_2 = 0.3$.

На рис. 2б зображено криві залежностей $\varepsilon_1 = \varepsilon_1(\beta)$ для деяких різних значень параметрів ν_1, ν_2 , що визначають здатність матеріалів до стисливості, за значення $g = 3$.

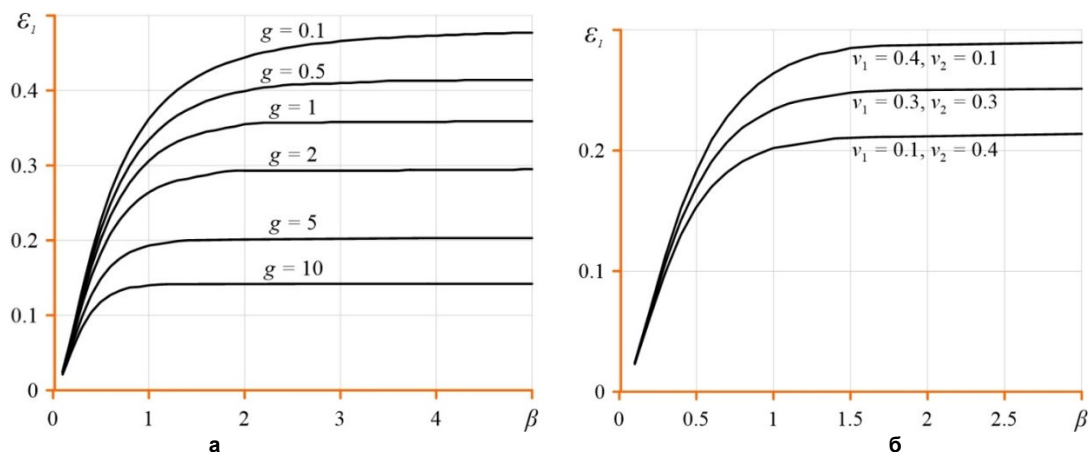


Рис. 2. Залежності критичного значення відносного укорочення ε_1 від значення геометричного параметру $\beta = h/a$ для: а – різних значень параметру $g = \mu_2/\mu_1$ за $\nu_1 = \nu_2 = 0.3$; б – різних значень параметрів ν_1, ν_2 за $g = 3$

Аналіз результатів показує, що за $\beta \rightarrow \infty$ критичні значення ε_1 асимптотично наближаються до значень, що відповідають внутрішній (структурній) нестійкості за стискання смуги, що проковзує без тертя між двома півплощинами (за відсутності тріщин), за тих самих значень фізико-механічних параметрів тіла (докл. див. Bogdanov, Nazarenko, & Kipnis, 2024b). Іншими словами, за зменшення довжини тріщини щодо ширини смуги розв'язок задачі, що розглядаємо, наближається до розв'язку аналогічної задачі для тіла без тріщини. Останнє цілком узгоджується з міркуваннями фізичного характеру.

Дискусія і висновки

У межах лінеаризованої теорії стійкості деформівних тіл аналітико-чисельний підхід, розвинений у (Bogdanov, Nazarenko, & Kipnis, 2024a, 2024b) для напівобмежених кусково-однорідних тіл із шаром покриття було успішно адаптовано до використання для задач про стиснення необмежених шаруватих композитів уздовж двох паралельних міжфазних структурних тріщин та визначені критичні значення параметрів навантаження задачі у випадку, коли матеріали тіла описуються пружним потенціалом гармонічного типу (стисливі матеріали, рівні корені характеристичних рівнянь).

Було показано, що коли матеріал "2" "тонкого" шару достатньо жорсткий щодо матеріалу "1" "товстого" шару шаруватого композиту (і, напр., виконує функцію армування) і довжина тріщин порівняно мала, то з достатньою для інженерних розрахунків точністю наявність дефектів у вигляді двох паралельних міжфазних структурних тріщин можна не враховувати в оцінюванні параметрів міцності такого структурно неоднорідного тіла, використовуючи значення цих параметрів, обраховані для аналогічного тіла без дефектів.

Список використаних джерел

- Гузь, А. Н. (2008). *Основи механіки руйнування композитів при сжатии*. Літера
 Камінський, А. О., Дудик, М. В., & Решітнік, Ю. В. (2023). Двопараметрична модель зони передруйнування у квазіпружному матеріалі біля вершини міжфазної тріщини. *Вісник Київського національного університету імені Тараса Шевченка. Фізико-математичні науки*, 2, 116–119. <https://doi.org/10.17721/1812-5409.2023/2.17>
 Bogdanov, V. L., Nazarenko, V. M., & Kipnis, A. L. (2025). Compression of a semi-bounded body with a coating layer along the interface sliding zone. *Z Angew Math Mech.*, 105, e202400799. <https://doi.org/10.1002/zamm.202400799>

- Bogdanov, V. L., Nazarenko, V. M., & Kipnis, O. L. (2024a). Compression of a semibounded body with thin coating layer along the interface near-surface crack. Part I. *International Applied Mechanics*, 60(5), 511–524. <https://doi.org/10.1007/s10778-025-01303-2>
- Bogdanov, V. L., Nazarenko, V. M., & Kipnis, O. L. (2024b). Compression of a semibounded body with thin coating layer along interface near surface crack. Part II. *International Applied Mechanics*, 60(6), 641–652. <https://doi.org/10.1007/s10778-025-01316-x>
- Griffith, A. A. (1920). The phenomenon of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London. Series A*, 221, 163–198. <https://doi.org/10.1098/rsta.1921.0006>
- Guz, A. N. (2014). Establishing the foundations of the mechanics of fracture of materials compressed along cracks (review). *International Applied Mechanics*, 50(1), 1–57. <https://doi.org/10.1007/s10778-014-0609-y>
- Guz, A. N., Bogdanov, V. L., & Nazarenko, V. M. (2020). Basic principles of fracture mechanics of materials loaded along cracks. In A. N. Guz, V. L. Bogdanov, & V. M. Nazarenko (Eds.), *Fracture of materials under compression along cracks. Advanced Structure Materials* (Vol. 138, pp. 61–148). Springer. https://doi.org/10.1007/978-3-030-51814-1_2
- Guz, A. N., Nazarenko, V. M., & Starodubtsev, I. P. (1991). A planar problem of failure of structural materials in compression along two parallel cracks. *Soviet Applied Mechanics*, 27, 352–360. <https://doi.org/10.1007/BF00896513>
- Guz, I. A. (1994). Investigation of the stability of a composite in compression along two parallel structural cracks at the layer interface. *International Applied Mechanics*, 30(11), 841–847.
- Irwin, G. R. (1957). Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics*, 24, 361–364. <https://doi.org/10.1115/1.4011547>
- Nazarenko, V. M. (1986). Two-dimensional problem of the fracture of materials in compression along surface cracks. *Soviet Applied Mechanics*, 22(10), 970–978. <https://doi.org/10.1007/BF01273678>

References

- Bogdanov, V. L., Nazarenko, V. M., & Kipnis, A. L. Compression of a semi-bounded body with a coating layer along the interface sliding zone. *Z Angew Math Mech.*, 105, e202400799 (2025). <https://doi.org/10.1002/zamm.202400799>
- Bogdanov, V. L., Nazarenko, V. M., & Kipnis, O. L. (2024a). Compression of a semibounded body with thin coating layer along the interface near-surface crack. Part I. *Int Appl Mech*, 60(5), 511–524. <https://doi.org/10.1007/s10778-025-01303-2>
- Bogdanov, V. L., Nazarenko, V. M., & Kipnis, O. L. (2024b). Compression of a semibounded body with thin coating layer along interface near surface crack. Part II. *Int Appl Mech*, 60(6), 641–652. <https://doi.org/10.1007/s10778-025-01316-x>
- Griffith, A. A. (1920). The phenomenon of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London. Series A*, 221, 163–198. <https://doi.org/10.1098/rsta.1921.0006>
- Guz, A. N. (2008). *Fundamentals of the fracture mechanics of compressed composites*. Litera [in Russian].
- Guz, A. N. (2014). Establishing the foundations of the mechanics of fracture of materials compressed along cracks (review). *International Applied Mechanics*, 50(1), 1–57. <https://doi.org/10.1007/s10778-014-0609-y>
- Guz, A. N., Bogdanov, V. L., & Nazarenko, V. M. (2020). Basic principles of fracture mechanics of materials loaded along cracks. In A. N. Guz, V. L. Bogdanov, & V. M. Nazarenko (Eds.), *Fracture of materials under compression along cracks. Advanced Structure Materials* (Vol. 138, pp. 61–148). Springer. https://doi.org/10.1007/978-3-030-51814-1_2
- Guz, A. N., Nazarenko, V. M., & Starodubtsev, I. P. (1991). A planar problem of failure of structural materials in compression along two parallel cracks. *Soviet Applied Mechanics*, 27, 352–360. <https://doi.org/10.1007/BF00896513>
- Guz, I. A. (1994). Investigation of the stability of a composite in compression along two parallel structural cracks at the layer interface. *International Applied Mechanics*, 30(11), 841–847.
- Irwin, G. R. (1957). Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics*, 24, 361–364. <https://doi.org/10.1115/1.4011547>
- Kaminsky, A. O., Dudyk, M. V., & Reshitnyk, Y. V. (2023). Two-parameter model of the prefraction zone in a quasi-elastic material near the tip of an interface crack. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 2, 116–119 [in Ukrainian]. <https://doi.org/10.17721/1812-5409.2023/2.17>
- Nazarenko, V. M. (1986). Two-dimensional problem of the fracture of materials in compression along surface cracks. *Soviet Applied Mechanics*, 22(10), 970–978. <https://doi.org/10.1007/BF01273678>

Отримано редакцією журналу / Received: 13.11.24

Прорецензовано / Revised: 11.03.25

Схвалено до друку / Accepted: 12.03.25

Alexander KIPNIS, PhD (Phys. & Math.), Senior Researcher

ORCID ID: 0000-0001-6747-8584

e-mail: a.i.kipnis@gmail.com

S. P. Timoshenko Institute of Mechanics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine

COMPRESSION OF A LAYERED COMPOSITE WITH SLIDING COMPONENTS ALONG TWO PARALLEL INTERFACE STRUCTURAL CRACKS

The problem of the linearized stability theory on compression of a piecewise-homogeneous body consisting of a nonlinearly elastic strip that slides frictionlessly between two nonlinearly elastic half-planes and weakened by two parallel cracks located on the rectilinear interface was studied. A proven analytical-numerical method is used, which consists in reducing the original boundary value problem using general representations of solutions of linearized equilibrium equations through potential harmonic functions to the Fredholm integral equation of the first kind. The specified solving integral equation of the problem was obtained in a general form for a wide class of combinations of two compressible or incompressible materials for which the case of equal roots of the characteristic equations is realized. The critical values of the load parameters corresponding to the start of the fracture of such a body were determined and the nature of their dependence on the physical-mechanical and geometrical parameters of the body was investigated.

Key words: layered composite material; interface crack; hyperelastic materials; fracture criterion; harmonic potential.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

УДК 539.3

DOI: <https://doi.org/10.17721/1812-5409.2025/1.9>

Веніамін КУБЕНКО, акад. НАН України

ORCID ID: 0000-0002-8863-8580

e-mail: vdk@inmech.kiev.ua

Інститут механіки ім. С. П. Тимошенка НАН України, Київ, Україна

Ігор ЯНЧЕВСЬКИЙ, д-р фіз.-мат. наук

ORCID ID: 0000-0002-7113-2276

e-mail: i.yanchevskiy@kpi.ua

Національний технічний університет України "КПІ ім. І. Сікорського", Київ, Україна

Олександр ОСТОС, д-р філософії

ORCID ID: 0009-0007-3546-2247

e-mail: ostos.alexander1994@gmail.com

Інститут механіки ім. С. П. Тимошенка НАН України, Київ, Україна

ДИНАМІЧНІ ХАРАКТЕРИСТИКИ ІДЕАЛЬНОЇ СТИСЛИВОЇ РІДИНИ, ЩО ЗБУДЖУЄТЬСЯ СФЕРИЧНИМ СЕГМЕНТОМ У ЦИЛІНДРИЧНІЙ ПОРОЖНИНІ

Представлено розв'язання осесиметричної задачі визначення потенціалу швидкостей ідеальної стисливої рідини у нескінченній круговій циліндричній порожнині зі сферичним сегментом, що вібрує за заданим законом. Розв'язок задачі побудовано на принципі суперпозиції потенціальних функцій, на перерозвиненні частинних розв'язків рівняння Гельмгольца, записаного в циліндричних (сферичних) координатах, по сферичних (циліндричних) хвильових функціях відповідно, що дало можливість, застосувавши також теореми додавання циліндричних функцій, представити повний потенціал швидкостей рідини як у сферичній, так і в циліндричній системах координат. Задовольняючи граничним умовам на поверхні сферичного сегмента та на циліндричній поверхні, задачу зведено до розв'язання нескінченної системи алгебраїчних рівнянь для коефіцієнтів розвинення сферичних і циліндричних складових шуканих потенціалів. Числові експерименти дають змогу виявити так звані резонансоподібні частоти, за яких гідродинамічні параметри суттєво збільшуються. Результати роботи можуть бути використані під час проєктування акустичних датчиків і решіток на основі сферичних ковпачків.

Ключові слова: циліндрична порожнина, стислива рідина, сферичний сегмент, гідродинамічні параметри.

Класифікація відповідно до AMS 2020: 33E15, 35Q35, 76N30.

Вступ

Дослідження взаємодії акустичних хвиль з різноманітними перешкодами викликають значний науковий і практичний інтерес. Вони передбачають розгляд так званих задач багатократного розсіювання. Актуальність розробки ефективних методів їхнього розв'язання пов'язана із широкими можливостями впровадження результатів у різноманітних галузях людської діяльності. Це, наприклад, керування рухом занурених у рідину частинок акустичним полем, акустична левітація, генерація бульбашок у рідині під дією швидко рухомих твердих тіл тощо. Сучасна медицина є ще однією галуззю застосування ефектів дифракції акустичних хвиль із бульбашками газу – газові мікробульбашки, які вводять у кровоносну систему живих організмів, дають змогу візуалізувати морфологію та різні процеси у серцево-судинній системі й інших органах.

У більшості випадків у задачах багатократного розсіювання розглянуто перешкоди сферичної форми. Найповніше сучасний стан досліджень за цією тематикою відображено в монографіях (Жук, О., Гузь, & Жук Я., 2023), (Кубенко, & Янчевський, 2024) та (Doinikov, 2005). Зокрема, монографія (Жук О., Гузь, & Жук Я., 2023) містить низку теоретичних задач динаміки акустичного збурення рідких крапель в іншій рідині. У монографії (Кубенко, & Янчевський, 2024) викладено методи і результати досліджень взаємодії акустичних хвиль із твердими та деформівними кулями. Вивченню дифракції акустичних хвиль на системах тіл як сферичної, так і циліндричної форм, присвячено монографію (Doinikov, 2005). Методика розв'язання відповідних задач ґрунтується на застосуванні принципу суперпозиції, методу розділення змінних і формул перерозвинення для скалярних гармонічних/хвильових функцій, за допомогою яких задачі зводяться до нескінченних систем алгебраїчних рівнянь щодо коефіцієнтів розвинення загальних розв'язків диференціальних рівнянь математичної моделі.

Менш дослідженими залишаються методи розв'язання задач багатократного розсіювання в механічних системах, що містять об'єкти неканонічної форми. У прикладному контексті певний інтерес представляють дослідження дифракції акустичних хвиль на об'єктах у вигляді сферичного сегменту, кулях із круговим отвором, сегментованих кулях із різними фізичними параметрами. У публікації (Xia et al., 2018) вивчено характеристики та закономірності акустичного випромінювання сферичним сегментом в безмежному середовищі. Показано, що отримані аналітично результати добре узгоджуються з експериментальними даними, зокрема, було досліджено властивості односпрямованих широкосмугових п'єзоелектричних акустичних датчиків, що використовують осесиметричні коливання як повних, так і неповних сферичних п'єзоелектричних оболонок методом електромеханічного модального аналізу. Також було проаналізовано спрямованість випромінювання складного масиву підводних акустичних решіток на напівсферичній перегородці на основі методу граничних елементів. Подібну задачу розсіювання плоскої хвилі деформівним сегментом замкненої кулі було розв'язано шляхом зведення до нескінченної системи лінійних алгебраїчних рівнянь типу Фредгольма (Liang et al., 2022) та (Резуненко, 2016), на основі якої досліджено спектр власних коливань кулі. У публікації (Резуненко, 2009) запропоновано методику розв'язання задачі дифракції плоскої акустичної хвилі на кулі з коловим отвором. У роботі (Кубенко, & Савін, 1995) викладено аналітичний метод розрахунку гідродинамічних параметрів ідеальної нестисливої рідини в нескінченній циліндричній порожнині, що збуджується сферичним сегментом на її осі. Проведені авторами (Кубенко, & Савін, 1995) дослідження свідчать про те, що за наявності

циліндричної порожнини амплітудні значення тиску та швидкості більші порівняно з випадком безмежного середовища. При цьому вони підвищуються зі збільшенням відносного радіуса сегмента.

Метою пропонованої роботи є дослідження особливостей хвильового поля рідини в циліндричній порожнині, обумовлених гармонічними коливаннями сферичного сегмента. Очікуємо, що певні співвідношення між радіусами порожнини

і сегмента та частоти збуджувального впливу можуть суттєво позначитися на характеристиках акустичного поля.

1. Постановка задачі

Розглянемо нескінченну кругову циліндричну порожнину радіусом ρ_0 , що заповнена ідеальною стисливою рідиною. Усередині порожнини симетрично відносно її осі перебуває сферичний сегмент, геометрію якого визначаємо радіусом R і кутом θ_0 (рис. 1.1).

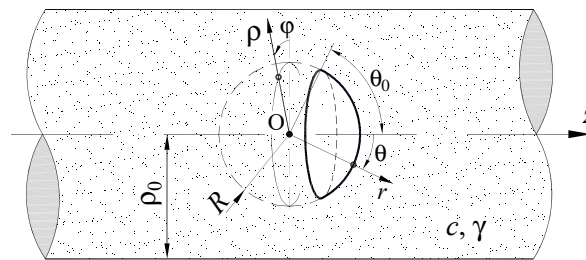


Рис. 1.1. Системи координат

Віднесемо порожнину до циліндричних координат (ρ, z, ϕ) так, що вісь порожнини збігається з віссю z . Із сегментом пов'яжемо сферичну систему координат $Or\theta\phi$ з початком у центрі сегменту, а зенітний кут θ виміряємо від осі z (рис. 1.1). Оскільки наявна осьова симетрія задачі, то координата ϕ випадає з розгляду.

Вважаємо, що сегмент здійснює гармонічні за часом осесиметричні радіальні коливання, амплітуда швидкості яких є заданою функцією частоти ω і кута θ – $V(\theta, t) = V(\theta)e^{-i\omega t}$.

Під час розв'язання задачі доцільно ввести до розгляду безрозмірнісні позначення, прийнявши за основні величини радіус циліндричної порожнини ρ_0 і характеристики рідини, що заповнює порожнину, а саме – швидкість поширення в ній звуку c та густину γ :

$$\bar{\rho} = \frac{\rho}{\rho_0}; \bar{r} = \frac{r}{\rho_0}; \bar{z} = \frac{z}{\rho_0}; \bar{R} = \frac{R}{\rho_0}; \bar{\omega} = \frac{\omega \rho_0}{c}; \bar{v} = \frac{v}{c}; \bar{p} = \frac{p}{\gamma c^2}; \bar{\kappa} = \kappa \rho_0; \bar{\Phi} = \frac{\Phi}{c \rho_0}. \quad (1.1)$$

Тут v – швидкість руху рідини, p – гідродинамічний тиск.

Надалі використовуємо тільки безрозмірнісні величини, тому риска над ними буде опущена.

Позначимо потенціали швидкостей розсіяних хвиль у рідині всередині сферичної області $0 \leq \rho < R$ через Φ_1 , а ззовні неї ($\rho \geq R$) – через Φ_0 . Ці потенціали є розв'язками рівняння Гельмгольца (Кубенко, & Янчевський, 2024)

$$\Delta \Phi_j + \kappa^2 \Phi_j = 0 \quad (j = 0, 1) \quad (1.2)$$

і мають задовольняти певним граничним умовам. Зокрема, на твердій поверхні циліндричної порожнини нормальна складова швидкості рідини має дорівнювати нулю –

$$\left. \frac{\partial \Phi_0}{\partial \rho} \right|_{\rho=\rho_0} = 0. \quad (1.3)$$

За умови, що на сегменті і поза його поверхнею маємо безвідривну взаємодію "зовнішньої" та "внутрішньої" до області $\rho = R$ рідин, то граничні умови полягають у рівності їхніх нормальних складових швидкостей

$$\left. \frac{\partial \Phi_0}{\partial r} \right|_{r=R} = \left. \frac{\partial \Phi_1}{\partial r} \right|_{r=R} \quad (0 \leq \theta < \pi) \quad (1.4)$$

та рівності тиску на умовній поверхні розділення областей:

$$\Phi_0|_{r=R} = \Phi_1|_{r=R} \quad (\theta_0 \leq \theta < \pi). \quad (1.5)$$

Варто зауважити, що в (1.2) і далі множник $e^{-i\omega t}$ маємо на увазі; Δ – диференціальний оператор Лапласа, а у записі (1.3)–(1.5) враховані співвідношення лінійної гідроаеропружності для швидкості і гідродинамічного тиску в рідині:

$$\mathbf{v}_j = \text{grad } \Phi_j; \quad p_j = i\omega \Phi_j \quad (j = 0, 1). \quad (1.6)$$

Умову для заданої швидкості на поверхні сегмента запишемо так:

$$\left. \frac{\partial \Phi_0}{\partial r} \right|_{r=R} = V(\theta) \quad (0 \leq \theta < \theta_0). \quad (1.7)$$

Потенціал поля розсіяних сегментом хвиль у порожнині має задовольняти також умовам випромінювання Зоммерфельда на нескінченності (Кубенко, & Янчевський, 2024)

$$\left(\frac{d\Phi_0}{d\rho} - i\kappa \Phi_0 \right) \rightarrow 0, \quad \Phi_0 \rightarrow 0 \quad \text{при} \quad \sqrt{\rho^2 + z^2} \rightarrow \infty. \quad (1.8)$$

Також обидва потенціали мають бути обмеженими на осі z і в початку координат (точці O):

$$\begin{aligned}\Phi_0 &\rightarrow \text{const} \quad \text{при} \quad \rho \rightarrow 0; \\ \Phi_1 &\rightarrow \text{const} \quad \text{при} \quad r \rightarrow 0.\end{aligned}\quad (1.9)$$

Визначення гідродинамічних параметрів рідини (1.6) полягає в розв'язанні рівнянь (1.2) із граничними умовами багатозв'язної задачі (1.4), (1.5), (1.7)–(1.9).

2. Побудова потенціалів

Використаємо метод розділення змінних у відповідних системах координат. Виходячи з виду граничних поверхонь задачі й користуючись принципом суперпозиції, хвильовий потенціал Φ_0 акустичного поля в циліндричній порожнині представимо у вигляді суми циліндричного $\Phi_{0\text{cyl}}(\rho, z)$ і сферичного $\Phi_{0\text{sph}}(r, \theta)$ потенціалів –

$$\Phi_0 = \Phi_{0\text{cyl}}(\rho, z) + \Phi_{0\text{sph}}(r, \theta). \quad (2.1)$$

Циліндричний і сферичний розв'язки рівняння Гельмгольца (1.2) мають вигляд (Кубенко, & Янчевський, 2024)

$$\Phi_{0\text{cyl}}(\rho, z) = \int_{-\infty}^{\infty} B(\xi) J_0(\sqrt{\kappa^2 - \xi^2} \rho) e^{i\xi z} d\xi; \quad (2.2)$$

$$\Phi_{0\text{sph}}(r, \theta) = \sum_{n=0}^{\infty} X_n h_n(\kappa r) P_n(\cos \theta). \quad (2.3)$$

Потенціал Φ_1 рідини всередині сферичної області $\rho < R$ обираємо у формі

$$\Phi_1(r, \theta) = \sum_{n=0}^{\infty} Z_n j_n(\kappa r) P_n(\cos \theta). \quad (2.4)$$

Тут $j_n(z)$, $h_n(z)$ – сферичні функції Бесселя і Ганкеля першого роду порядку n , відповідно; $J_0(z)$ – циліндрична функція Бесселя нульового порядку; $P_n(x)$ – функція Лежандра; X_n , Z_n – невідомі константи; $B(\xi)$ – невідома густина; ξ – константа поділу.

Вибір необхідного розв'язку з множини (2.2)–(2.4) реалізується шляхом задоволення граничних умов. Відзначимо, що умови (1.8) і (1.9) вже задоволені вибором розв'язків (2.2)–(2.4). Складові $\Phi_{0\text{cyl}}$ і $\Phi_{0\text{sph}}$ потенціалу Φ_0 у формулі (2.1), згідно з (2.2) і (2.3), записані в системах координат, координатна поверхня яких збігається з відповідною граничною поверхнею гідропружної системи. Для задоволення граничних умов на кожній такій поверхні необхідно потенціал Φ_0 представити в системі координат, пов'язаній із цією поверхнею. Для цього скористаємося співвідношеннями (Кубенко, & Янчевський, 2024), що дозволяють циліндричні хвильові функції виразити через сферичні хвильові функції, і навпаки.

У підсумку вирази для потенціалу Φ_0 сумарного акустичного поля рідини в циліндричній і сферичній системах координат можуть бути записані так:

$$\Phi_0(\rho, z) = \int_{-\infty}^{\infty} A(\xi) H_0(\sqrt{\kappa^2 - \xi^2} \rho) e^{i\xi z} d\xi + \int_{-\infty}^{\infty} B(\xi) J_0(\sqrt{\kappa^2 - \xi^2} \rho) e^{i\xi z} d\xi; \quad (2.5)$$

$$\Phi_0(r, \theta) = \sum_{n=0}^{\infty} [X_n h_n(\kappa r) + Y_n j_n(\kappa r)] P_n(\cos \theta), \quad (2.6)$$

причому між невідомими функціями і коефіцієнтами виконуються рівності:

$$A(\xi) = \frac{1}{2\kappa} \sum_{n=0}^{\infty} X_n i^{-n} P_n\left(\frac{\xi}{\kappa}\right); \quad Y_n = i^n (2n+1) \int_{-\infty}^{\infty} B(\xi) P_n\left(\frac{\xi}{\kappa}\right) d\xi. \quad (2.7)$$

Тут $H_0(z)$ – циліндрична функція Ганкеля нульового порядку 1-го роду.

3. Виконання граничних умов

Співвідношення (2.1)–(2.7) дають можливість перейти до задоволення граничних умов. Умова (2.3) на поверхні циліндричної порожнини з урахуванням виразів (2.7) і єдиності перетворення Фур'є дає змогу записати потенціал Φ_0 у сферичній системі координат через коефіцієнти X_n (Кубенко, & Янчевський, 2024)

$$\Phi_0(r, \theta) = \sum_{n=0}^{\infty} [X_n h_n(\kappa r) - j_n(\kappa r) \frac{2n+1}{2\kappa} \sum_{m=0}^{\infty} X_m i^{n-m} q_{nm}] P_n(\cos \theta), \quad (3.1)$$

$$\text{де } q_{nm} = \int_{-\infty}^{\infty} \frac{H_1(\sqrt{\kappa^2 - \xi^2} \rho_0)}{J_1(\sqrt{\kappa^2 - \xi^2} \rho_0)} P_m\left(\frac{\xi}{\kappa}\right) P_n\left(\frac{\xi}{\kappa}\right) d\xi.$$

На наступному етапі подання потенціалів (3.1) і (3.4) використовують для задоволення умов (1.4), (1.5). Унаслідок ортогональності поліномів Лежандра з умови (1.4) одержимо

$$Z_n = X_n \frac{h'_n(\kappa R)}{j'_n(\kappa R)} - \frac{2n+1}{2\kappa} \sum_{m=0}^{\infty} X_m i^{n-m} q_{nm}. \quad (3.2)$$

Тут штрих позначає похідну спеціальної функції за радіальною координатою, наприклад $j'_n(\kappa R) = \frac{d}{dr} [j_n(\kappa r)] \Big|_{r=R}$.

Невідомі коефіцієнти X_n визначаємо з граничних умов (1.5) і (1.7), які можуть бути об'єднані так:

$$H(\theta_0 - \theta) \cdot [v_0(R, \theta) - V(\theta)] + H(\theta - \theta_0) \cdot [p_0(R, \theta) - p_1(R, \theta)] = 0, \quad (3.3)$$

де $H(\theta)$ – функція Гевісайда.

Підставляючи у (3.3) вирази (2.4) і (2.6) з урахуванням (3.2), отримаємо рівняння:

$$0 = H(\theta_0 - \theta) \left\{ V(\theta) - \sum_{n=0}^{\infty} \left[X_n h'_n(\kappa R) - \frac{2n+1}{2\kappa} j'_n(\kappa R) \sum_{m=0}^{\infty} X_m i^{n-m} q_{nm} \right] P_n(\cos \theta) \right\} + \\ + H(\theta - \theta_0) i \kappa \sum_{n=0}^{\infty} X_n \left[h_n(\kappa R) - \frac{h'_n(\kappa R)}{j'_n(\kappa R)} j_n(\kappa R) \right] P_n(\cos \theta),$$

яке, помноживши на $P_f(\cos \theta) \sin \theta$ та проінтегрувавши за кутовою координатою, дає змогу одержати нескінченну систему алгебраїчних рівнянь щодо коефіцієнтів X_n :

$$\sum_{n=0}^{\infty} X_n \left[h'_n(\kappa R) \alpha_{fn} - i \omega \left[h_n(\kappa R) - \frac{h'_n(\kappa R)}{j'_n(\kappa R)} j_n(\kappa R) \right] \beta_{fn} - \sum_{m=0}^{\infty} i^{m-n} q_{mn} \alpha_{fm} \frac{2m+1}{2\kappa} j'_m(\kappa R) \right] = W_f \quad (n = 0, 1, \dots). \quad (3.4)$$

Тут позначено:

$$W_f = \int_0^{\theta_0} V(\theta) P_f(\cos \theta) \sin \theta d\theta; \quad \alpha_{fn} = \int_0^{\theta_0} P_f(\cos \theta) P_n(\cos \theta) \sin \theta d\theta; \quad \beta_{fn} = \delta_{fn} \frac{2}{2n+1} - \alpha_{fn}.$$

Можна припустити, що система (3.4) належить до класу регулярних і її розв'язок можна побудувати методом усичення з параметром N , тобто X_n варто розглядати як розв'язок скінченної системи з N -рівнянь. Параметр усичення при цьому визначається пробним шляхом за критерієм досягнення достатньої точності задоволення граничних умов задачі.

Аналіз системи лінійних алгебраїчних рівнянь (3.4) показує, що вона структурно подібна до розрахункової системи у випадку нестисливої рідини (Кубенко, & Савін, 1995), водночас замість степеневих функцій (див. Кубенко, & Савін, 1995) тут використано сферичні функції Бесселя. Це обумовлено залученням принципово інших трансляційних формул і теорем додавання спеціальних функцій, що і визначає певну складність методу розв'язання розглянутої задачі порівняно з представленою у роботі (Кубенко, & Савін, 1995) для нестисливої рідини.

Визначивши коефіцієнти X_n ($n = 0, N$), на наступному етапі обчислюємо гідродинамічний тиск і швидкість точок акустичного середовища за співвідношеннями (1.6), у яких потенціали Φ_j ($j = 0, 1$) визначаємо виразами (2.4), (3.1) і (3.2). Зокрема, для радіальних швидкостей точок рідини у сферичній системі координат будуть справедливими такі рівності:

$$v_0(r, \theta, t) = e^{-i\omega t} \sum_{n=0}^N \left[X_n \frac{h'_n(\kappa r)}{j'_n(\kappa r)} - \frac{2n+1}{2\kappa} \sum_{m=0}^{\infty} X_m i^{n-m} q_{nm} \right] j'_n(\kappa r) P_n(\cos \theta); \quad (3.5) \\ v_1(r, \theta, t) = e^{-i\omega t} \sum_{n=0}^N \left[X_n \frac{h'_n(\kappa r)}{j'_n(\kappa r)} - \frac{2n+1}{2\kappa} \sum_{m=0}^{\infty} X_m i^{n-m} q_{nm} \right] j'_n(\kappa r) P_n(\cos \theta).$$

Тут ураховано множник $e^{-i\omega t}$, який до цього мали на увазі.

Дійсна частина гідродинамічного параметру визначає його значення на початку періоду, а уявна – через чверть періоду. Амплітудне значення обчислюємо за теоремою Піфагора від дійсної та уявної частин.

4. Числові результати

Під час проведення числових експериментів ми вважали, що циліндрична порожнина заповнена водою, густина і швидкість поширення в ній звуку дорівнюють $\gamma = 1000 \text{ кг/м}^3$ і $c = 1500 \text{ м/с}$, відповідно.

У розв'язанні нескінченної системи алгебраїчних рівнянь (3.4) параметр усичення прийнято таким, що дорівнює $N = 40$, а безрозмірна частота коливань сегмента ω варіювалася в діапазоні від 0.2 до 10 з кроком 0.05 ($\kappa = \omega$).

Рисунок 4.1 ілюструє обчислений на основі співвідношень (1.6), (2.4), (3.1) і (3.2) розподіл амплітудного значення гідродинамічного тиску в околі сферичного сегмента з кутом розкриття $2\theta_0 = \pi/2$ за різних значень його радіуса на тій частоті збудження ω , що є однією з резонансоподібних для цієї механічної системи (Жук О., Гузь, & Жук Я., 2023; Кубенко, & Янчевський, 2024).

Рисунок 4.1а відповідає параметрам $R = 0.25$ і $\omega = 7.0$, рис. 4.1б – $R = 0.5$ і $\omega = 8.4$, рис. 4.1в – $R = 0.75$ і $\omega = 5.65$. Функцію $V(\theta)$ обрано у вигляді $V(\theta) = \cos(\pi\theta/2\theta_0)$.

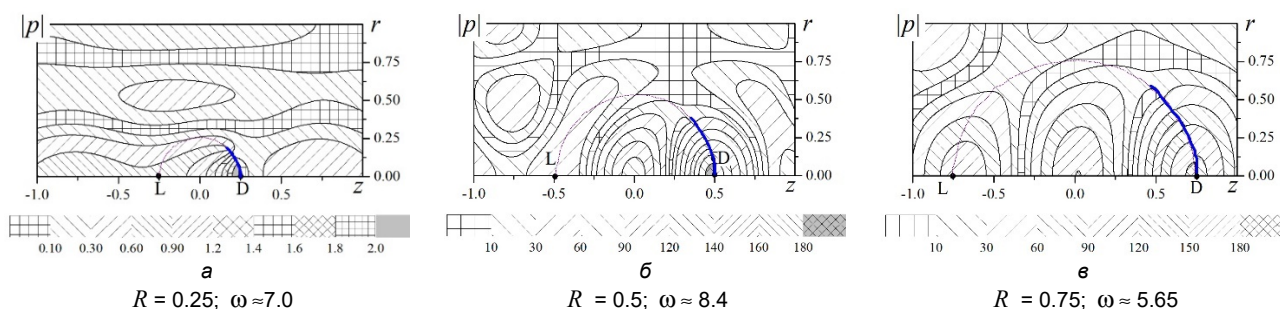


Рис. 4.1. Поле тиску в околі сегмента

З рисунку видно, що тиск набуває найбільших значень в околі сегмента, а максимальний тиск спостерігаємо вздовж осьової координати z в околі точки D з координатами $(+R, 0)$.

На рис. 4.1 також показано виконання граничної умови (1.7) про рівність гідродинамічного тиску зовні і зсередини умовної поверхні розділення областей ($r = R$). Видно, що збільшення відносного радіуса сегмента за відповідних значень частоти збуджувального впливу суттєво позначається на характеристиках акустичного поля, а саме розподіл тиску стає більш нерівномірним і вираженим в околі початку координат і точці D. За відносно великих радіусів сегмента ($R \geq 0.5$) виявлено утворення кількох ексцентричних областей зміни значень тиску, на відміну від випадку $R = 0.25$.

Варто також зауважити, що "умовно резонансні частоти" механічної системи можуть бути визначені виключно через аналіз усіченої розрахункової системи рівнянь (3.4) – на цих частотах детермінант матриці набуває великих значень.

На рис. 4.2 для сегмента з $R = 0.5$ проведено сканування гідродинамічного тиску у точці D від різних значень частоти коливань сегмента з $\theta_0 = \pi/4$ (рис. 4.2a) і з $\theta_0 = \pi/2$ (рис. 4.2б). На цьому рисунку $\text{Re}(p)$ та $\text{Im}(p)$ відображають дійсну й уявну частину тиску, відповідно.

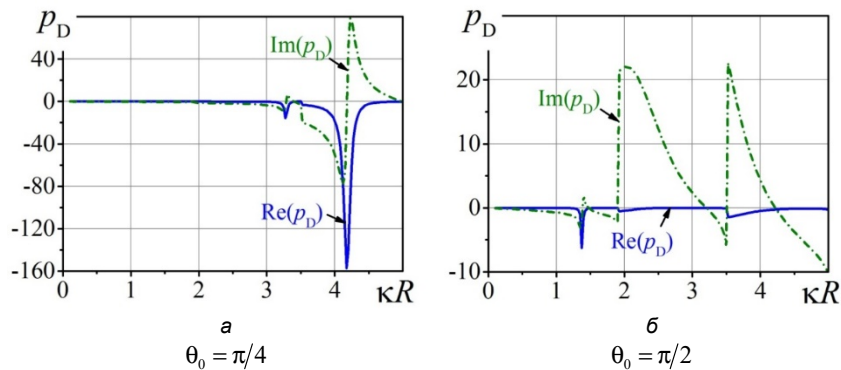


Рис. 4.2. Тиск у точці D за $R = 0.5$

З аналізу представлених на рис. 4.2a результатів випливає, що сплески значень тиску спостерігаємо за великих значень частоти, а саме за $\omega > 3$. Резонансоподібну поведінку тиску виявлено за $\kappa R \approx 4.2$ ($\omega \approx 8.4$). На рис. 4.2б помітно більше резонансоподібних частот. Також варто зауважити, що уявна частина тиску має більші значення, ніж дійсна.

На рис. 4.3 представлені розподіли дійсних та уявних частин радіальної та кутової швидкостей уздовж осьової координати r за частоти $\omega = 8.4$ та фіксованого значення кутової координати $\theta_1 = \pi/2$. Рисунок 4.3a побудовано для сегмента розміром $\theta_0 = \pi/4$, рис. 4.3б – для $\theta_0 = \pi/2$.

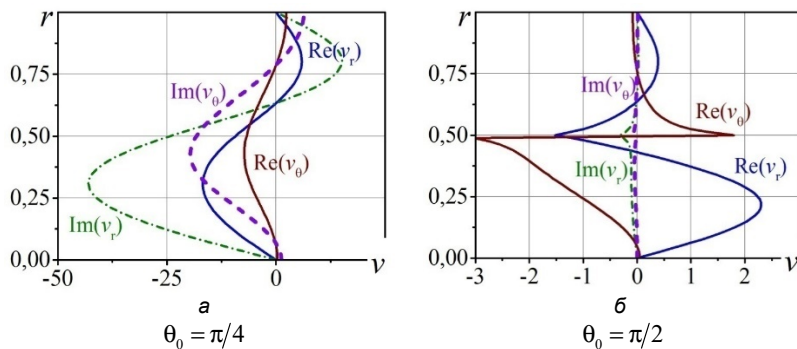


Рис. 4.3. Розподіл швидкості вздовж осі r за $\theta_1 = \pi/2$

Із цих рисунків видно, що виконується гранична умова як на твердій поверхні циліндричної порожнини (1.5), оскільки за $r = 1$ маємо $|v_r| \approx 0$, так і на уявній границі поділу рідини за $r = R$. З рисунку також випливає, що у випадку $\theta_0 = \pi/4$ (рис. 4.3a) уявні частини як радіальної, так і кутової швидкості набувають вищих амплітудних значень, ніж відповідні їхні дійсні частини. За $\theta_0 = \pi/2$ (рис. 4.3б) – навпаки. Визначення причини цього явища потребує окремого дослідження.

Аналіз представлених на рис. 4.3 кривих свідчить, що амплітуда швидкості для сегмента з удвічі меншим кутом розкриття ($\theta_0 = \pi/4$), приблизно в десять разів більша. Це може бути обумовлено локалізацією області збурення рідини.

Для аналізу впливу циліндричної границі на дифракційні процеси в околі сферичного сегмента були обчислені гідродинамічні параметри рідини за коливань сегмента в безмежному середовищі. Усунення граничної умови (1.3) у цьому випадку приводить до аналогічних до (3.4) і (3.5) розрахункових виразів, у яких необхідно покласти $q_{nm} = 0$.

На рис. 4.4 представлено залежність амплітуди швидкості рідини у точці D від частоти коливань сегментів з різними кутами θ_0 у циліндричній порожнині (криві cyl) і у безмежному середовищі (криві unl).

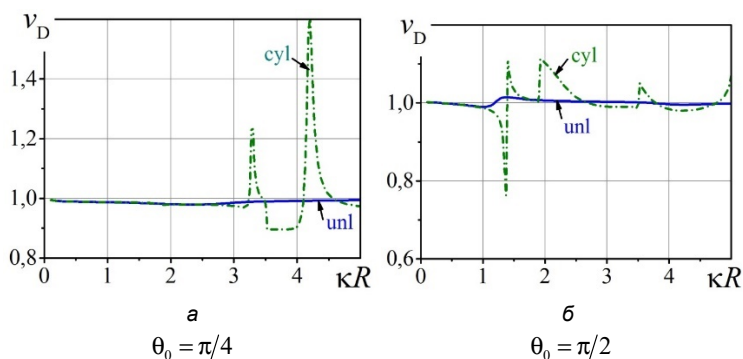


Рис. 4.4. Швидкість у точці D за $R = 0.5$

З рис. 4.4 можна зробити висновок, що в безмежному середовищі амплітуда швидкості рідини в точці D майже не залежить від частоти коливання сегмента. За наявності циліндричної порожнини (криві cyl) чітко виявляються «умовно резонансні» частоти. Рисунок 4.4 підтверджує той факт, що для сегмента з меншим кутом θ_0 ($\theta_0 = \pi/4$) максимальні значення швидкості у лобовій точці D значно більші, ніж для випадку $\theta_0 = \pi/2$, і це також відбувається за більших частот збудження сегмента.

На рис. 4.5 криві compr відображають розподіл радіальної швидкості стисливої рідини уздовж зенітного кута θ за фіксованого значення частоти коливання сегмента ($\omega = 8.4$).

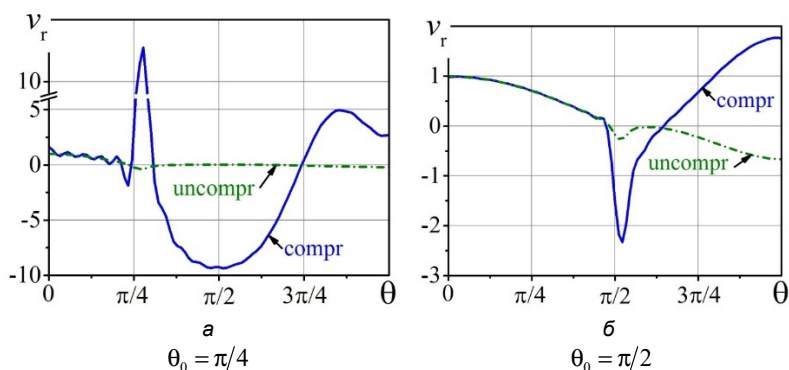


Рис. 4.5. Розподіл швидкості за перерізом сегмента за $R = 0.5$; $\omega = 8.4$

Криві uncompr на цьому рисунку відображають результати розв'язання задачі про коливання також сегмента у нестисливій рідині, що побудовані на підставі представлених у роботі (Kubenko, & Savin, 1995) виразів. З порівняння кривих видно, що на поверхні сегмента результати приблизно збігаються. Суттєву відмінність спостерігається на уявній межі розділення розрахункових областей ($\theta > \theta_0$). Амплітуда швидкості коливань сегмента у постановці ідеально стисливої рідини (крива compr) значно вища, ніж для випадку нестисливої рідини. Це пояснюємо врахуванням стисливості рідини в розглянутій постановці.

Дискусія і висновки

У роботі запропоновано аналітичний метод дослідження гідродинамічних параметрів у заповненій ідеальною стислою рідиною нескінченній циліндричній порожнині, що містить вібруючий осесиметрично сферичний сегмент. Розв'язок граничної задачі багатократного розсіяння побудовано у вигляді розвинення за власними функціями в циліндричних і сферичних координатах, пов'язаних із кожною граничною поверхнею. Для задоволення граничних умов використано співвідношення, що виражають циліндричні хвильові функції через сферичні, та навпаки. Це дало змогу одержати нескінченну систему алгебраїчних рівнянь щодо коефіцієнтів шуканого розв'язку, під час розв'язання якої використано метод усічення.

Проведені числові експерименти показали, що в околі досліджуваного сферичного сегмента тиск набуває найбільших значень, причому зі зменшенням кутової довжини сегмента чи зі збільшенням його відносного радіуса абсолютні значення тиску та швидкості зростають. При цьому так звані умовно резонансні частоти, за яких відбувається це зростання, можуть бути визначені виключно через аналіз детермінанту розрахункової системи рівнянь. Наявність циліндричної порожнини обумовлює появу нових умовно резонансних частот, яких не було у відповідній постановці задачі з безмежним середовищем.

На відміну від задачі для нестисливої рідини, де усталені процеси моделюються рівнянням Лапласа, застосування рівняння Гельмгольца для моделювання динамічних процесів у стисливій рідині приводить до необхідності застосування принципово інших базових співвідношень, зокрема і трансляційних теорем та формул додавання спеціальних функцій. Хоча як у випадку стисливої, так і у випадку нестисливої рідин для отримання прийнятної точності результатів знадобилося приймати практично однакові параметри усічення для розв'язання розрахункових нескінченних систем лінійних алгебраїчних рівнянь. Для стисливої рідини у циліндричній порожнині характерне

існування умовно резонансних частот, за яких амплітудні значення гідродинамічних параметрів у разі перевищують амплітуду зовнішнього навантаження. Проте модель нестисливої рідини не дає змоги виявляти такі ефекти.

Внесок авторів: Веніамін Кубенко – постановка задачі, концептуалізація, методологія, написання чорнового варіанту рукопису; Ігор Янчевський – методологія, побудова графіків, числові експерименти, аналіз результатів; Олександр Остос – аналіз джерел, побудова аналітичного розв'язку, аналіз результатів, оформлення рукопису.

Список використаних джерел

- Жук, О. П., Гузь, О. М., & Жук Я. О. (2023). *Радіаційні сили акустичного поля в рідині з включеннями*. Альянт.
- Кубенко, В. Д., & Янчевський, І. В. (2024). *Дифракція акустичних хвиль на системах неоднотипних тіл*. Політехніка.
- Резуценко, В. А. (2009). Дифракция плоской акустической волны на сфере с круговым отверстием. *Вісник Харківського національного університету ім. В. Н. Каразіна. Серія: Математика, прикладна математика і механіка*, 850, 71–77.
- Резуценко, В. А. (2016). Сфера, составленная из мягкого и жесткого круговых сегментов, в поле плоской акустической волны. *Вісник Харківського національного університету ім. В. Н. Каразіна. Серія: Радіофізика та електроніка*, 25, 58–65.
- Doinikov, A. A. (2005). *Bubble and Particle Dynamics in Acoustic Fields: Modern Trends and Applications*. Research Signpost.
- Kubenko, V. D., & Savin, V. A. (1995). Determination of the dynamic characteristics of an ideal incompressible liquid excited by a spherical segment in a cylindrical cavity. *International Applied Mechanics*, 31, 567–574.
- Liang, X.-X., Linz, N., Freidank, S., Paltauf, G., & Vogel, A. (2022). Comprehensive analysis of spherical bubble oscillations and shock wave emission in laser-induced cavitation. *Journal of Fluid Mechanics*, 940, A5.
- Xia, J., Huang, H., Zhang, C., & Li, Q. (2018). Analysis on acoustic directivity of spherical cap transducers. *Acta Acustica*, 43, 592–599.

References

- Zhuk, O. P., Guz, O. M., & Zhuk, Ya. O. (2023). *Radiation forces of acoustic field in liquid with inclusions*. Aliant [in Ukrainian].
- Kubenko, V. D., & Yanchevskyi, I. V. (2024). *Acoustic waves diffraction on systems of heterogeneous bodies*. Polytechnic [in Ukrainian].
- Doinikov, A. A. (2005) *Bubble and Particle Dynamics in Acoustic Fields: Modern Trends and Applications*. Research Signpost.
- Xia J., Huang H., Zhang C., & Li Q. (2018). Analysis on acoustic directivity of spherical cap transducers. *Acta Acustica*, 43, 592–599.
- Liang, X.-X., Linz, N., Freidank, S., Paltauf, G., & Vogel, A. (2022). Comprehensive analysis of spherical bubble oscillations and shock wave emission in laser-induced cavitation. *Journal of Fluid Mechanics*, 940, A5.
- Rezunencko, V. A. (2009). Diffraction of a plane acoustic wave on a sphere with a circular hole. *Bulletin of V. N. Karazin Kharkiv National University. Series: Mathematics, Applied Mathematics and Mechanics*, 850, 71–77 [In Russian].
- Rezunencko, V. A. (2016). A sphere composed of soft and rigid circular segments in the field of a plane acoustic wave. *Visnyk of V. N. Karazin Kharkiv National University. Series: Radiophysics and Electronics*, 25, 58–65 [In Russian].
- Kubenko, V. D., & Savin, V. A. (1995). Determination of the dynamic characteristics of an ideal incompressible liquid excited by a spherical segment in a cylindrical cavity. *International Applied Mechanics*, 31, 567–574.

Отримано редакцією журналу / Received: 26.06.24

Прорецензовано / Revised: 07.02.25

Схвалено до друку / Accepted: 12.03.25

Veniamin KUBENKO, Active Member of the National Academy of Sciences of Ukraine

ORCID ID: 0000-0002-8863-8580

e-mail: vdk@inmech.kiev.ua

S. P. Tymoshenko Institute of Mechanics of NAS of Ukraine, Kyiv, Ukraine

Ihor YANCHEVSKYI, DSc, Prof.

ORCID ID: 0000-0002-7113-2276

e-mail: i.yanchevskiy@kpi.ua

National Technical University of Ukraine "I. Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Oleksandr OSTOS, PhD (Phys. & Math.)

ORCID ID: 0009-0007-3546-2247

e-mail: ostos.alexander1994@gmail.com

S. P. Tymoshenko Institute of Mechanics of NAS of Ukraine, Kyiv, Ukraine

DYNAMIC CHARACTERISTICS OF AN IDEAL COMPRESSIBLE FLUID EXCITED BY A SPHERICAL SEGMENT IN A CYLINDRICAL CAVITY

The manuscript presents solving of axisymmetric problem of determining the velocity potential of an ideal compressible fluid in an infinite circular cylindrical cavity with a spherical segment, vibrating according to a given law. The solution of the problem is based on the principle of superposition of potential functions, on the redevelopment of partial solutions of the Helmholtz equation, written in cylindrical (spherical) coordinates, in spherical (cylindrical) wave functions, respectively, which made it possible, also using the theorems of addition of cylindrical functions, to represent the full potential of fluid velocities in both spherical and cylindrical coordinate systems. Satisfying the boundary conditions on the surface of the spherical segment and on the cylindrical surface, the problem is reduced to solving an infinite system of algebraic equations for the coefficients of the expansion of the spherical and cylindrical components of the desired potentials. Numerical experiments allow us to identify the so-called "resonance-like frequencies", at which the hydrodynamic parameters increase significantly. The results of the work can be used in the design of acoustic sensors and arrays based on spherical caps.

Key words: cylindrical cavity, compressible fluid, spherical segment, hydrodynamic parameters

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.6

DOI: <https://doi.org/10.17721/1812-5409.2025/1.10>

Gennadiy SANDRAKOV, DSc (Phys. & Math.), Leading Researcher
ORCID ID: 0000-0001-9133-605X
e-mail: gennadiy.sandrakov@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Mykola KYRYCHEK, Student
e-mail: mykolakyrychek@gmail.com
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

MODELING AND HOMOGENIZATION OF PROCESSES IN CYLINDRICAL LAYERED COMPOSITES LIKE CAPACITORS AND TRANSFORMERS

Capacitors and transformers are widely used devices. There is a trend towards minaturization of such devices. Thus, a natural problem of modeling thermal and electromagnetic processes in such devices arises. Understanding the models of such processes allows, for example, to optimize or make cheaper of these devices. But, the geometry of such devices is very specific. For example, a regular capacitor is a long strip with an insulator rolled into a cylinder. Direct modeling of processes in such a device is complicated, since there are very many layers, which leads to equations with highly oscillating coefficients. On the other hand, methods of homogenization theory can be applied to such equations and equations with constant coefficients can be obtained, the solutions of which approximate the solutions of the equations under consideration. The solution of equations with constant coefficients is much simpler both by numerical and analytical methods. As a first step in this direction, one dynamic equation will be considered here, which corresponds to the modeling of wave and thermal processes. Thus, the problem of modeling wave and thermal processes in a cylinder consisting of a large number of alternating layers with different characteristics is considered. Using the asymptotic methods of the homogenization theory, this problem is approximately reduced to modeling processes in a homogeneous (homogenized) cylinder. Asymptotic expansions and homogenized problems are presented, the solutions of which determine the approximate asymptotic of solutions to the problem under consideration. Formulas for calculating the characteristics of the homogeneous cylinder and accuracy estimates of the approximations are given. The presented accuracy estimates of such approximations are essential for numerical calculations and computer modeling of the problems with guaranteed accuracy. In addition, the given approximations are undoubtedly useful for understanding wave and thermal processes in devices similar to capacitors and transformers. This approach can be considered as a basis for considering more general electromagnetic and thermal processes, which are modeled by Maxwell's equations coupled with the heat equation.

Keywords: initial boundary value problems, homogenized problems, wave equations, heat equations, asymptotic expansions, heat conductivity coefficients, wave propagation speed coefficients.

AMS 2020 classification: 80M40, 35B27, 80M35.

Introduction

One of the homogenization theory founders, Nikolai S. Bakhvalov, spoke at lectures about the creation of a real rolling mill. This mill rolled out an iron billet into a long thin strip. Then this strip is folded into a roll and moved to cool. The cooling time of the roll was calculated as a metal cylinder. The place of cooling was allocated in accordance with this time. But, the real cooling time turned out to be much longer than the calculated one, which led to overflow of the cooling place. In addition, the increase in cooling time led to the loss of some relevant properties of the roll metal, which had a catastrophic effect on the profitability of the rolling mill. As a result, this real rolling mill had to be closed.

The story, of course, is about the usefulness of calculation formulas for cylindrical layered composite media and materials, since the metal in the roll is not solid metal and can be separated by air layers. For ordinary periodically layered composite media, for example, such formulas for heat conductivity equation are known (Bensoussan, Lions, & Papanicolaou, 1978; Bakhvalov, & Panasenko, 1989). The formulas were obtained by methods of asymptotic expansions for solutions of the wave and heat equation at the beginning of the development of the homogenization theory. But, we are not aware of works where the formulas were obtained for cylindrical layered composites and Bakhvalov's problem was solved. Here the problem will be considered in various layer geometric configurations, in particular for cylindrical layered composites, which are similar to some models for capacitors and transformers.

The usefulness of layered composites for influencing thermal processes is known actually in electrodynamics. Indeed, transformer cores are periodically assembled from thin plates separated by insulating layers. Such a periodic structure of the plates dampens the Foucault currents (Landau, & Lifshitz, 1984), which prevents energy losses for heating the cores. Transformer cores can also be made from long strips with an insulating layer rolled into a toric roll, which simulates a layered structure in cylindrical coordinates and resembles the metal roll. It is known (Bertotti, 1998) that Foucault currents can be modeled by diffusion equations for the magnetic field strengths. Here we consider only one such equation for cylindrical layered composites, which is also used to model thermal processes and is called the heat conduction equation. Note also that diffusion equations model filtration processes. Some filters actually used have a cylindrical layered structure. Therefore, the methods discussed here may be relevant for modeling filtration processes in porous cylindrical layered media with various layer configurations.

Respectively, the transformer cores are the basis for the winding, consisting of turns of conductors, which wrap around part of the magnetic system of the transformer. As a simple model of such a system, one can consider a cylindrical layered composite, the layers of which are formed by a large number of rings or a helix with many turns separated by some material. Often in practice, these can be turns of insulated wires surrounded by air. Such composites will be discussed as well. Layers of such composites are called solenoids in electrodynamics (Vagner, Lembrikov, & Wyder, 2004).

At last, some real capacitors are made from a very thin strip of foil, lined with a thin insulator, which is rolled into a cylinder starting from some core. It is for the modeling of thermal processes in such composites resembling the metal roll that this

© Sandrakov Gennadiy, Kyrychek Mykola, 2025

article was conceived. The problem of describing thermal fields in capacitors was posed by Vladimir V. Mokhor as significant for understanding the distribution of heat, since sometimes capacitors overheat, which leads to irreversible consequences. The use of homogenization methods in solving this problem has led to some applications of these methods to the modeling of layered composites like transformers. Naturally, to fully simulate such processes in composites, it is necessary to consider the complete system of heat conductivity and Maxwell equations (Landau, & Lifshitz, 1984). The approach used here can be considered as a basis for modeling more general processes.

To model processes in cylindrical layered composites, we will use the homogenization theory. The theory arose about fifty years ago with the advent of artificial composite materials. Such materials usually have a periodic structure, which is technologically convenient. Indeed, by fixing the cell of periodicity, it is possible to copy this cell to fill a certain volume, just as city squares are covered with tiles. There are usually a lot of such cells in artificial composites. Therefore, when modeling processes in the composites, it is necessary to choose a very fine mesh to take into account the structure of each cell, which requires very significant computer resources.

Initially, the homogenization theory was developed by N. S. Bakhvalov, I. Babushka, A. Bensoussan, J. L. Lions, and G. Papanicolaou to simplify computer simulations processes in composite materials with a periodic structure; relevant references can be found, for example, in (Bensoussan, Lions, & Papanicolaou, 1978; Bakhvalov, & Panasenko, 1989). According to the theory, for example, modeling of heat processes in composites can be replaced with guaranteed accuracy by modeling the heat process in homogeneous (homogenized) media. In a mathematical sense, this means that the solution of the heat conduction problem for the equation with periodic rapidly oscillating coefficients is close to the solution of the problem for the equation with constant coefficients. The corresponding theorems and accuracy estimates were proved, for example, in (Bensoussan, Lions, & Papanicolaou, 1978; Bakhvalov, & Panasenko, 1989) for sufficiently regular data.

Later, it turned out that it is much easier to prove the convergence or multiscale convergence of solutions to solutions of some homogenized multiscale problems, starting essentially from (Allaire, 1992). This approach, usually, does not require additional data regularity that is essential for nonlinear problems as, for example, in (Piat, & Sandrakov, 2004; Sandrakov, 2004a; 2005) for variational inequalities whose solutions are not regular as a rule. However, the relevant convergence theorems are based entirely on a priori solution estimates, and for problems with additional parameters such estimates may not be exact, which can lead to incorrect homogenized multiscale problems. Thus, inferring homogenized problems from solution estimates requires high mathematical skills.

In contrast to the method of asymptotic expansions, since the construction of asymptotics leads to a small residual by definition and, therefore, a priori estimates provide guaranteed accuracy of the asymptotic approximation, which implies convergence of solutions. This approach allows to separate fast and slow variables, similar methods are used, for example, in (Diaz et al., 2019a; 2019b). But, for homogenized multiscale problems such variables are not separated and the type of the corresponding problems is not obvious. In addition, the accuracy of the approximations was not clarified and proved in the multiscale case, further details and relevant references can be found, for example, in (Mielke, Reichelt, & Thomas, 2014; Jager, & Woukeng, 2021; Garttnerat, et al., 2022; Ray, & Schulz, 2023; Garttnerat, Knabner, & Ray, 2024). An approach leading also to homogenized multiscale problems based on the periodic unfolding method is presented, for example, in (Cioranescu, Damlamian, & Griso, 2018; 2023). It is obvious that such multiscale convergence methods are convenient justification methods for results of the homogenization theory.

The homogenization method was originally based on asymptotic solution expansions for second order divergent equations according to (Bensoussan, Lions, & Papanicolaou, 1978; Bakhvalov, & Panasenko, 1989). Here we will use a method of asymptotic solution expansions for elliptic equations of arbitrary order, the use of which does not require high skills as detailed in (Sandrakov, 1991). This approach makes it possible to consider various problems for modeling of processes in cylindrical layered composites like capacitors and transformers. This will allow us to consider in sufficient detail the actual problems for such modeling under various geometric assumptions. The exact formulation of such problems is presented in the next section. Various variants of geometric structures for cylindrical layered composites are discussed in sections 3–5. Composites consisting of periodic layers, rolled strips, and layers of periodic rings will be considered sequentially in the sections. Finally, section 6 deals with composites with helix layers that model some transformers.

1. Models for processes in cylindrical layered composites

Models for wave and heat processes in composite materials will be considered. Precise definitions will be needed for the exact formulation of the models. Let

$$\Omega = B \times (0, l) = \{(x, y) \in \mathbb{R}^2 : \kappa < x^2 + y^2 < R^2\} \times (0, l)$$

be a cylinder of radius R and length l with a removed core of radius κ . Here B denotes an open circle of radius R without origin with some neighborhood of the positive radius κ . The origin is removed here to correctly use cylindrical coordinates, and some neighborhood is removed to correctly define the boundary conditions, otherwise such conditions would be quite specific.

The coordinates $(r, \varphi, z) \in (0, R) \times [0, 2\pi) \times (0, l)$ will be applied on the domain Ω . We will also use functional spaces, which are defined, for example, in (Bensoussan, Lions, & Papanicolaou, 1978; Bakhvalov, & Panasenko, 1989; Duvaut, & Lions, 1972). Such spaces are necessary for the correct setting of the initial boundary value problems considered here, as is customary in mathematical and numerical physics.

Let the functions $u_0, u_1 \in C_0^\infty(\Omega)$ and $f \in C_0^\infty((0, T) \times \Omega)$ be given, where the positive T is fixed. We define the function $u = u(t, x, y, z) \in C^0([0, T]; H_0^1(\Omega))$ as a solution to the initial boundary value problem for the wave equation

$$\begin{aligned} m^\varepsilon u''_{tt} + s^\varepsilon u'_t - \operatorname{div}(A^\varepsilon \nabla u) &= g^\varepsilon f \quad \text{in } \Omega \times (0, T), \\ u|_{t=0} &= u_0, \quad u'_t|_{t=0} = u_1 \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega \times (0, T), \end{aligned} \quad (1)$$

where the equation coefficients of the problem characterize the component properties of the composite materials and ε^{-1} is a large parameter specified below as the number of layers. The equation of this problem is understood in the distribution sense (Duvaut, & Lions, 1972).

In the case when $m^\varepsilon = 0$ and the initial condition $u'|_{t=0} = u_1$ is removed, the coefficient s^ε characterizes the density (normalized to the specific heat capacity) of the components and similarly for $g^\varepsilon = s^\varepsilon$, the matrix A^ε determines the heat conductivity of the components in the directions of coordinate axes, and this problem models the heat process in the domain Ω as it is generally accepted. However, the case of non-zero m^ε may be used to model heat processes at extreme temperatures, as suggested, for example, in (Saedodin, & Barforoush, 2017) with reference to the work of J. C. Maxwell 1867. This case is also called the telegraph equation or the relativistic heat equation.

In general, the coefficient m^ε characterizes the density of the components and similarly for $g^\varepsilon = m^\varepsilon$, the matrix A^ε determines the permeability of the components in the directions of coordinate axes, and the problem models the wave process with damping in the domain Ω , where s^ε is the characteristic of such damping. Sometimes, Maxwell's equations can be reduced to wave equations with damping. In this case, the matrix A^ε is usually diagonal and can characterize the electrical or magnetic permeability of the components, and the coefficient s^ε determines the specific conductivity of the components when Ohm's law is fulfilled. But, for models of electromagnetic processes, it is necessary to consider a system of equations for electric and magnetic fields. Here we are only studying one equation as a model for wave or heat processes, which can be considered as basic for the general case.

For thermal processes, for example, the condition $u = 0$ on the boundary is called the homogeneous Dirichlet condition and means that the temperature on the boundary of the cylinder is zero, which is convenient in the mathematical sense. If for some process the ambient temperature is equal to $\eta \neq 0$, then the boundary condition should be replaced by the condition $u = \eta$. But, one can replace the unknown function $u = u(t, x, y, z)$ in problem (1) with a new unknown function $v = u - \eta$ for which the homogeneous boundary condition is fulfilled. It is possible to consider other types of boundary conditions in (1), for example, the Neumann or mixed boundary conditions. In thermal physics and electrodynamics, it is also customary to reduce the equations of the problem to a dimensionless form; for simplicity, this is not done here.

Let us now define the coefficients in (1), which are specific to layered composites. Using cylindrical coordinates, we rewrite the equation of the problem in the form

$$m^\varepsilon u''_{tt} + s^\varepsilon u'_t - \partial_r (a_1^\varepsilon \partial_r u) - r^{-1} a_1^\varepsilon \partial_r u - r^{-2} \partial_\varphi (a_2^\varepsilon \partial_\varphi u) - \partial_z (a_3^\varepsilon \partial_z u) = g^\varepsilon f \quad (2)$$

in $\Omega \times (0, T)$, where $\partial_r, \partial_\varphi, \partial_z$ denote partial derivatives with respect to cylindrical coordinates. It is assumed that the matrix $A^\varepsilon = A^\varepsilon(r)$ is diagonal and depends only on the radial variable r . The last assumption will also be made for these other coefficients $m^\varepsilon = m^\varepsilon(r)$, $s^\varepsilon = s^\varepsilon(r)$ and $g^\varepsilon = g^\varepsilon(r)$.

Fixing a sufficiently large integer N , we will consider $\varepsilon = R/N$ as a small parameter and represent the cylinder as the union of tubes

$$\Omega = \bigcup_{k=1}^N B_k \times (0, l) = \bigcup_{k=1}^N \{ (x, y) \in \mathbb{R}^2 : \kappa + \varepsilon(k-1) \leq \sqrt{x^2 + y^2} < \kappa + \varepsilon k \} \times (0, l),$$

where B_k is a ring of thickness ε and suitable radii for $k = 1, \dots, N$.

Let us define the coefficients of equation (2) by the equalities

$$m^\varepsilon = m(r/\varepsilon), \quad s^\varepsilon = s(r/\varepsilon), \quad a_i^\varepsilon = a_i(r/\varepsilon), \quad g^\varepsilon = g(r/\varepsilon), \quad (3)$$

where the functions $m = m(\tau)$, $s = s(\tau)$, $a_i = a_i(\tau)$ and $g = g(\tau)$ for $i = 1, 2, 3$ are 1-periodic (periodic with period 1 relative to the variable $\tau \in \mathbb{R}$) and are defined on the period by the following equalities

$$\begin{aligned} m(\tau) &= m_0 \text{ if } \tau \in [0, \sigma] \cup [1 - \sigma, 1] \text{ and } m(\tau) = m_1 \text{ if } \tau \in (\sigma, 1 - \sigma), \\ a_i(\tau) &= a_0^i \text{ if } \tau \in [0, \sigma] \cup [1 - \sigma, 1] \text{ and } a_i(\tau) = a_1^i \text{ if } \tau \in (\sigma, 1 - \sigma) \end{aligned} \quad (4)$$

for $i = 1, 2, 3$. Similarly, $s(\tau)$ is equal to s_0 or s_1 and $g(\tau)$ is equal to g_0 or g_1 , where all constants $m_0, m_1, a_0^i, a_1^i, g_0, g_1$ and σ are positive and fixed. In addition $\sigma \in (0, 0.5)$. Further, all 1-periodic functions defined on the period are continued through the periodicity. Schematically, the coefficient model is presented in Fig. 1.

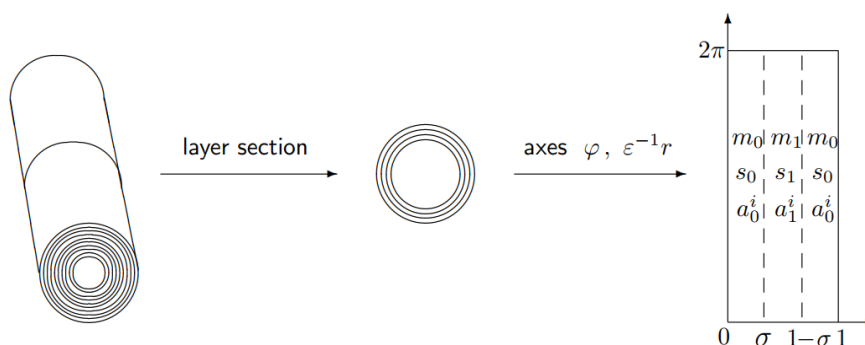


Fig. 1. Model of coefficients of cylindrical layered composites

2. Asymptotic expansions and homogenization

Under the assumptions made, problem (1) has a unique solution, for example, in accordance with (Duvaut, & Lions, 1972). Thus, one can try to construct an approximation of this solution for sufficiently small ε . Following (Sandakov, 1991), one may choose an asymptotic approximation for the solution of problem (1) in the form

$$u_\varepsilon = v(t, r, \varphi, z) + \varepsilon v_1(t, r, \varphi, z, r/\varepsilon) + \varepsilon^2 v_2(t, r, \varphi, z, r/\varepsilon), \quad (5)$$

where by definition the functions $v_1(t, r, \varphi, z, \tau)$ and $v_2(t, r, \varphi, z, \tau)$ are 1-periodic with respect to the variable $\tau \in \mathbb{R}$.

To determine the terms of asymptotic expansion (5), the following differentiation formula for functions of a special form $w = w(t, r, \varphi, z, r/\varepsilon)$ will be used

$$\partial_r w(t, r, \varphi, z, r/\varepsilon) = (\varepsilon^{-1} \partial_\tau w(t, r, \varphi, z, \tau) + \partial_r w(t, r, \varphi, z, \tau))|_{\tau=r/\varepsilon}. \quad (6)$$

The main idea of constructing asymptotic expansions is to observe that the variables t, r, φ, z, τ are independent inside the brackets of the formula. In the sense that if $\partial_\tau w(t, r, \varphi, z, \tau) = 0$, for example, in (6) for all variables, then this is also true for $\tau = r/\varepsilon$. Note that the first term in expansion (5) does not depend on r/ε , which is determined by the kernel of the relevant operator on 1-periodic functions according to (Sandrakov, 1991). For divergent operators of the second order, the kernel is generated by 1, which is well known in homogenization theory (Bensoussan, Lions, & Papanicolaou, 1978).

Let's substitute asymptotic approximation (5) into equation (2) and take formula (6) into account. Then, retaining only the essential terms, we obtain

$$\begin{aligned} m v''_{tt} + s v'_t - \varepsilon^{-1} \partial_\tau (a_1 \partial_\tau v_1) - \varepsilon^{-1} \partial_\tau (a_1 \partial_r v) - \\ - \partial_\tau (a_1 \partial_\tau v_2) - \partial_r (a_1 \partial_r v) - \partial_\tau (a_1 \partial_r v_1) - \partial_r (a_1 \partial_\tau v_1) - \\ - r^{-1} a_1 \partial_\tau v_1 - r^{-1} a_1 \partial_r v - r^{-2} a_2 \partial_\varphi^2 v - a_3 \partial_z^2 v - \varepsilon r^{-1} a_1 \partial_r v_1 - \\ - \varepsilon \partial_\tau (a_1 \partial_r v_2) - \varepsilon \partial_r (a_1 \partial_r v_1) - \varepsilon \partial_r (a_1 \partial_\tau v_2) = g f \end{aligned} \quad (7)$$

for $\tau = r/\varepsilon$, where $m = m(\tau)$, $s = s(\tau)$, $a_i = a_i(\tau)$ and $g = g(\tau)$ for $i = 1, 2, 3$ in accordance with definition (3).

Equating the coefficients at ε^{-1} in asymptotic equality (7), we have

$$-\partial_\tau (a_1 \partial_\tau v_1) = \partial_\tau (a_1 \partial_r v). \quad (8)$$

This equality will be fulfilled if we choose $v_1 = \Phi(\tau) \partial_r v(t, r, \varphi, z)$ and define the function $\Phi = \Phi(\tau)$ as a solution of the following 1-periodic problem

$$-\partial_\tau (a_1 \partial_\tau \Phi) = \partial_\tau (a_1) \text{ in } (0, 1). \quad (9)$$

It is known (Bensoussan, Lions, & Papanicolaou, 1978) that such a solution to problem (9) exists and is defined up to a constant, which will be fixed by the condition $\langle \Phi \rangle = \int_0^1 \Phi(\tau) d\tau = 0$. Indeed, we have $a_1 \partial_\tau (\Phi + \tau) = C$ and $\partial_\tau \Phi = a_1^{-1} C - 1$. Therefore, we obtain $C = \langle a_1^{-1} \rangle^{-1}$ and

$$\Phi = \int_0^\tau (a_1^{-1}(\eta) \langle a_1^{-1} \rangle^{-1} - 1) d\eta + \langle \Phi \rangle.$$

Equating the coefficients at ε^0 in asymptotic equality (7) and using the definition already made that $v_1 = \Phi \partial_r v$, we get

$$\begin{aligned} -\partial_\tau (a_1 \partial_\tau v_2) = \partial_\tau (a_1 \Phi) \partial_r^2 v + (a_1 + a_1 \partial_\tau \Phi) \partial_r^2 v + \\ + r^{-1} (a_1 + a_1 \partial_\tau \Phi) \partial_r v + r^{-2} a_2 \partial_\varphi^2 v + a_3 \partial_z^2 v + g f - m v''_{tt} - s v'_t. \end{aligned} \quad (10)$$

By analogy with (8), consider the equality as a 1-periodic problem for $v_2 = v_2(\tau)$ with fixed variables t, r, φ, z . It is known (Bensoussan et al., 1978) that the problem is correct if and only if the right-hand side is orthogonal to 1 in the space $L^2(0, 1)$ and therefore we obtain

$$\hat{m} v''_{tt} + \hat{s} v'_t - \hat{a}_1 \partial_r^2 v - r^{-1} \hat{a}_1 \partial_r v - r^{-2} \hat{a}_2 \partial_\varphi^2 v - \hat{a}_3 \partial_z^2 v = \hat{g} f \text{ in } \Omega \times (0, T), \quad (11)$$

where, using that $a_1 + a_1 \partial_\tau \Phi = \langle a_1^{-1} \rangle^{-1}$, the constants are defined by the equalities

$$\begin{aligned} \hat{m} = \langle m \rangle = \int_0^1 m(\tau) d\tau, \quad \hat{s} = \langle s \rangle, \quad \hat{g} = \langle g \rangle, \\ \hat{a}_1 = \langle a_1^{-1} \rangle^{-1}, \quad \hat{a}_2 = \langle a_2 \rangle, \quad \hat{a}_3 = \langle a_3 \rangle. \end{aligned} \quad (12)$$

Thus, starting from equation (2) with complex rapidly oscillating coefficients, we have obtained an equation with constant coefficients, which is called *homogenized*. Adding initial and boundary conditions to this equation, as in problem (1), we get a homogenized problem of the same type, but with constant coefficients.

It is known (Duvaut, & Lions, 1972) that the initial boundary value problem for the homogenized equation has a unique regular solution. Similarly, the unique regular solvability holds for the heat case (when formally $m^\varepsilon = 0$ and the condition $u'_t|_{t=0} = u_1$ is removed). By the method of separation of variables, the solution can be constructed as a convergent series, which cannot be done for problem (1). The series can be used for a computer representation. Moreover, each term of the series is a particular solution of the homogenized problem with suitable initial conditions. Also, the homogenized problem can be solved numerically by standard methods.

Consequently, two terms in the asymptotic approximation (5) are defined. In order to determine the third term $v_2 = v_2(t, r, \varphi, z, r/\varepsilon)$, we subtract equation (11) from equality (10). Then, we obtain

$$-\partial_\tau (a_1 \partial_\tau v_2) = \tilde{a}_1 \partial_r^2 v + r^{-2} \tilde{a}_2 \partial_\varphi^2 v + \tilde{a}_3 \partial_z^2 v + \tilde{g} f - \tilde{m} v''_{tt} - \tilde{s} v'_t, \quad (13)$$

where the determined coefficients $\tilde{a}_1 = \partial_\tau (a_1 \Phi)$, $\tilde{a}_2 = a_2 - \hat{a}_2$, $\tilde{a}_3 = a_3 - \hat{a}_3$, $\tilde{g} = g - \hat{g}$, $\tilde{m} = m - \hat{m}$ and $\tilde{s} = s - \hat{s}$ are 1-periodic functions and such that they are orthogonal to 1 in the space $L^2(0, 1)$. Note, for example, that the equalities $\langle \partial_\tau (a_1 \Phi) \rangle = -\langle (a_1 \Phi) \partial_\tau 1 \rangle = 0$ are valid, which means that $\partial_\tau (a_1 \Phi)$ and 1 are orthogonal. Here we have used the integration by parts formula and the 1-periodicity of the function $a_1 \Phi$.

Hence, equality (13) and therefore equality (10) are fulfilled if we choose

$$v_2 = \Phi_1 \partial_r^2 v + r^{-2} \Phi_2 \partial_\varphi^2 v + \Phi_3 \partial_z^2 v + \Phi_g f - \Phi_m v''_{tt} - \Phi_s v'_t \quad (14)$$

and define the functions Φ_m and Φ_s as 1-periodic solutions of the problems

$$\partial_\tau (a_1 \partial_\tau \Phi_m) = \tilde{m}, \quad \partial_\tau (a_1 \partial_\tau \Phi_s) = \tilde{s} \text{ in } (0, 1), \quad (15)$$

for example. It is known (Bensoussan, Lions, & Papanicolaou, 1978) that the solutions to problems (15) exist and are determined up to constants, which will be fixed by the conditions $\langle \Phi_m \rangle = \langle \Phi_g \rangle = 0$. So, one can define the 1-periodic functions Φ_1, Φ_2, Φ_3 and Φ_g in (14). With this choice of the function $v_2 = v_2(t, r, \varphi, z, z/\varepsilon)$, all coefficients at ε^0 in asymptotic equality (10) are equal to zero. Likewise, one can construct additional terms $v_3 = v_3(t, r, \varphi, z, z/\varepsilon)$, $v_4 = v_4(t, r, \varphi, z, z/\varepsilon)$ in expansion (5) according to (Sandrakov, 1991). But, the available terms will be enough for approximating the solution of problem (1).

Thus, the asymptotic equality (7) is fulfilled with accuracy up to $O(\varepsilon)$, since the solution to the homogenized problem is regular due to the assumptions on the initial data and so the terms in (5) are regular. So, the asymptotic approximation $u_a = u_a(t, r, \varphi, z, z/\varepsilon)$ also approaches the equation (2) with accuracy up to $O(\varepsilon)$. Hence, following (Sandrakov, 1998), this can be verified by energetic methods that the last statement implies, for example, the following accuracy estimate

$$\begin{aligned} & \|r^{1/2}(\partial_r u - \theta^\varepsilon \partial_r v + r^{-1} \partial_\varphi u - r^{-1} \partial_\varphi v + \partial_z u - \partial_z v)\|_{C^0([0,T];L^2(\Omega))}^2 \\ & + \|u'_t - v'_t\|_{C^0([0,T];L^2(\Omega))}^2 + \|u - v\|_{C^0([0,T];L^2(\Omega))}^2 \leq C \varepsilon, \end{aligned} \quad (16)$$

with $\theta^\varepsilon = \theta(r/\varepsilon)$, where $\theta(\tau) = 1 + \partial_\tau \Phi(\tau)$, $\Phi = \Phi(\tau)$ is the solution of problem (9), u is the solution of problem (1) with coefficients from (3), v is the solution of homogenized problem with coefficients from (12) and the constant C does not depend on ε when $0 < \varepsilon \leq \varepsilon_0$ for the corresponding ε_0 .

The specified accuracy estimate may be proved as follows. It is known (Duvaut, & Lions, 1972) and can be directly verified that the solution of problem (1) satisfies the following energy estimate

$$\begin{aligned} & \|u'_t\|_{C^0([0,T];L^2(\Omega))}^2 + \|u\|_{C^0([0,T];L^2(\Omega))}^2 + \|\nabla u\|_{C^0([0,T];L^2(\Omega))}^2 \leq \\ & \leq C \left(\|f\|_{L^1(0,T;L^2(\Omega))}^2 + \|u_0\|_{H_0^1(\Omega)}^2 + \|u_1\|_{L^2(\Omega)}^2 \right), \end{aligned}$$

where C does not depend on ε . This energy estimate is also valid for a solution of the homogenized problem for equation (11). Indeed, this homogenized problem can be written in the form (1), but for an equation with constant coefficients, for which the energy estimate is well known. Further, from the equation of this problem, we can express $v''_{tt}|_{t=0}$ as an element of the space $C_0^\infty(\Omega)$. Hence, v'_t satisfies the similar problem, but with the right side $f'_t \in C_0^\infty((0,T) \times \Omega)$. Thus, the energy estimate is carried out for v'_t with a constant C independent of ε . Continuing by induction, we conclude that for the n -th derivative the inclusion $v_t^{(n)} \in C^0([0,T];H_0^1(\Omega))$ holds for the required n .

Now, using elliptic regularity (according to (Kinderlehrer, & Stampacchia, 1980), for example) and writing the homogenized equation in the following form

$$\operatorname{div}(\hat{a} \nabla v) = \hat{m} v''_{tt} + \hat{s} v'_t - \hat{g} f$$

we get that $v_t^{(n-2)} \in C^0([0,T];H^3(\Omega) \cap H_0^1(\Omega))$, where $\hat{a}$ is a diagonal matrix with constant elements from (12). Accordingly, we also have that $v_t^{(n-4)} \in C^0([0,T];H^5(\Omega) \cap H_0^1(\Omega))$. Continuing by induction and using embedding theorems, we conclude that $v \in C^n([0,T] \times \bar{\Omega})$ with the required n and norm estimates that do not depend on ε . (Similar inclusions are fulfilled for thermal processes when $m^e = 0$.) Such inclusions imply the smoothness of the solution of the homogenized problem in cylindrical coordinates, which completely describes the structure of the terms of the asymptotic expansion (5). Namely, each such term is a smooth function multiplied by a bounded one, which follows from the properties of the solutions of problem (9) and similar problems, for example, (15).

According to the construction of u_a , the right side of the equation for $w = u - u_a$ is equal to εF , where the function F is the sum of terms, each of which is a smooth (and so bounded) function multiplied by a bounded one, which follows from definitions and the asymptotic equality (7), based on formula (6). This yields the estimate (16), since the gradient in cylindrical coordinates is $(\partial_r w, r^{-1} \partial_\varphi w, \partial_z w)$ and the relevant Jacobian is r . It is also used here that $\partial_r u_a = \theta^\varepsilon \partial_r v + O(\varepsilon)$ according to (5) and (6), since $v_1 = \Phi(r/\varepsilon) \partial_r v$ by definition. More precisely and correctly, the terms v_1 and v_2 in the asymptotic expansion (5) should be multiplied by a smooth cutoff function (equal to zero at the boundary and to one for $(x, y, z) \in \Omega$ distant from the boundary by more than ε). But, such a multiplication does not change the estimate (16), as is well known (Bensoussan, Lions, & Papanicolaou, 1978; Bakhvalov, & Panasenko, 1989).

Note that the accuracy estimate (16), proved for smooth and compactly supported data, easily implies convergence for irregular energy data by density. Indeed, suppose that $f \in L^1(0,T;L^2(\Omega))$ in (1), and the other conditions remain unchanged. It is known (Duvaut, & Lions, 1972) that there exists a sequence of $\{f_i\}_{i=1}^\infty \subset C_0^\infty((0,T) \times \Omega)$ such that

$$\|f - f_i\|_{L^1(0,T;L^2(\Omega))} \leq \delta_i \rightarrow 0 \quad \text{for} \quad i \rightarrow \infty.$$

For the solution of problem (1) with the right-hand side f_i , estimate (16) holds with a constant C_i that does not depend on ε , but may depend on i .

Hence, using the energy estimate for these problems (with f , f_i and $f - f_i$ as right-hand sides of the equation) and the triangle inequality, we obtain the following estimate

$$\begin{aligned} & \|r^{1/2}(\partial_r u - \theta^\varepsilon \partial_r v + r^{-1} \partial_\varphi u - r^{-1} \partial_\varphi v + \partial_z u - \partial_z v)\|_{C^0([0,T];L^2(\Omega))}^2 \\ & + \|u'_t - v'_t\|_{C^0([0,T];L^2(\Omega))}^2 + \|u - v\|_{C^0([0,T];L^2(\Omega))}^2 \leq C_i \varepsilon + \delta_i. \end{aligned}$$

(We use here, for example, that $\|u - v\| \leq \|u - u_i\| + \|v - v_i\| + \|u_i - v_i\|$ in natural notations, for example, v_i is the solution to problem (11) with the right-hand side f_i .)

For every positive δ one can take and fix i such that $\delta_i < \delta/2$. Now, there is ε_0 such that $C_i \varepsilon < \delta/2$ for $0 < \varepsilon < \varepsilon_0$. By definition, this means that the right hand side in the last highlighted inequality tends to zero as $\varepsilon \rightarrow 0$. This completes the

proof of the convergence of solutions of problem (1) to the solution of the homogenized problem as $\varepsilon \rightarrow 0$ in the case of irregular energy data.

Thus, we have a fairly good accuracy estimate for solutions and gradients for sufficiently small ε . In this case, for modeling wave processes (respectively, heat processes when $m^\varepsilon = 0$) in cylindrical layered composites, instead of solving problem (1) (for example, numerically through computer simulations), it is possible to solve the homogenized problem with guaranteed accuracy. Naturally, the numerical solution of problem (1) is much more difficult than the solution of the homogenized problem, since a very fine calculation grid is required to take into account the geometry of a very large number of layers in Ω , since the coefficients in (3) are rapidly oscillating.

Remark that the estimate (16) is fulfilled when all constants $m_0, m_1, a_0^i, a_1^i, g_0$ and g_1 in (4) are positive and fixed. If one of these constants is small compared to ε , for example, the case $a_0^1 = C \varepsilon^2$ is interesting. In this case, composites with highly variable properties should be considered. The homogenization problem in Cartesian coordinates for this case was considered, for example, in (Sandakov, 1998; 2004b). Interestingly, the composites discussed in (Sandakov, 2004b) have the properties of semiconductors for waves, which transmit or not waves depending on their frequencies. This is a consequence of the fact that the homogenization problem in Cartesian coordinates is hyperbolic (parabolic) problem for integro-differential equations. Optimization issues for such equations were considered, for example, in (Anikushyn, Lyashko, & Samosonok, 2024; Anikushyn, & Hranishak, 2025) and for the parabolic case in (Kapustyan et al., 2025). Using the approaches of these works, for example, it is possible to optimize the properties of the homogenized equation coefficients.

All coefficients of the homogenized equation in (12) are calculated as averages, except for $\hat{a}_1$. The coefficient is calculated using the so-called harmonic mean. As an example, for air the heat conductivity is $a_0^1 = a_0^2 = a_0^3 = 0.02$, and for iron we have $a_1^1 = a_1^2 = a_1^3 = 100$ according to reference books in suitable dimensional values.

Hence, $\langle a_1^{-1} \rangle = 50 \cdot 2 \sigma + 0.01 \cdot (1 - 2 \sigma)$ and, for example, when $\sigma = 0.25$, we obtain

$$\hat{a}_1 = \langle a_1^{-1} \rangle^{-1} = (25.005)^{-1} \approx 0.04, \quad \hat{a}_2 = \hat{a}_3 \approx 50.01.$$

In addition, we get

$$\begin{aligned} \hat{a}_1 &\approx 0.42, \quad \hat{a}_2 = \hat{a}_3 \approx 99.95 \quad \text{for } \sigma = 0.025, \\ \hat{a}_1 &\approx 0.02, \quad \hat{a}_2 = \hat{a}_3 \approx 5.01 \quad \text{for } \sigma = 0.475. \end{aligned}$$

For comparison, for glass $a_1^1 = a_1^2 = a_1^3 = 1$. Let $a_0^1 = a_0^2 = a_0^3 = 0.02$ also. Then, we have

$$\begin{aligned} \hat{a}_1 &\approx 0.04, \quad \hat{a}_2 = \hat{a}_3 \approx 0.51 \quad \text{for } \sigma = 0.25, \\ \hat{a}_1 &\approx 0.29, \quad \hat{a}_2 = \hat{a}_3 \approx 0.95 \quad \text{for } \sigma = 0.025, \\ \hat{a}_1 &\approx 0.02, \quad \hat{a}_2 = \hat{a}_3 \approx 0.07 \quad \text{for } \sigma = 0.475. \end{aligned}$$

This explains why air layer materials are good heat insulators in a direction perpendicular to the layers. In reality, this fact for multiple layers has long been used in thermoses. There was a report by a specialist in composite mechanics, at which homogenization methods for conventional composites with periodic structure were stated to be in good agreement with experimental data even at $\varepsilon = 0.1$. But, we cannot give an exact reference. As a naive example of thermoses, it can be seen that thermal insulation is well achieved even for two or three layers of thermal insulation. Naturally, this will be even better for composites with many layers considered here. In conventional layered geometries, this fact is actually used in multiple glazing for good heat insulation. Of course, all these conclusions are applicable to wave processes, since $\hat{a}_1$ determines the speed of wave propagation in the direction perpendicular to the layers. It is known, for example, that multiple glazing is a good sound insulator.

Comparing these two cases with the predominance of air (for $\sigma = 0.475$), it can be seen that the heat insulation properties of composites with iron and glass are not very different, at least for the direction perpendicular to the layers with a large number of such alternating layers. But, this is not entirely true for other directions. Nevertheless, glass flask thermoses have long been used for a long time, and iron flask thermoses are more recently available. Accordingly, in the case of wave processes, a system of repeating iron sheets can be used for sound insulation.

3. Models of processes in cylindrical rolled layered composites

We have considered composites which are the union of thin cylindrical tubes. Some capacitors are made from a very thin strip of foil, lined with a thin insulator, which is rolled into a cylinder. This case can be considered by the same homogenization method. Indeed, consider a new coordinate

$$\rho = r + \varepsilon (\varphi/2\pi).$$

It would be possible to use this coordinate as new instead of r in equation (2). But, we will use the coordinate only in coefficients. Note that fixing $r = r_0 > \kappa$, we have $\rho = r_0 + \varepsilon k$ for $\varphi = 2\pi k$, where $k = 1, \dots, N$. Thus, the curve $\rho = r_0 + \varepsilon (\varphi/2\pi)$ describes a spiral with step ε for positive φ , which is called *Archimedean*, and models the cylindrical rolled layered composite if we add the coordinate $z \in (0, l)$ and constrain the value of the variable φ .

Now we define the coefficients of equation (2) by the equalities

$$m^\varepsilon = m(\rho/\varepsilon), \quad s^\varepsilon = s(\rho/\varepsilon), \quad a_i^\varepsilon = a_i(\rho/\varepsilon), \quad g^\varepsilon = g(\rho/\varepsilon), \quad (17)$$

where the functions $m = m(\tau)$, $s = s(\tau)$, $a_i = a_i(\tau)$ and $g = g(\tau)$ for $i = 1, 2, 3$ are 1-periodic and are defined on the period by equalities (4). We remark, for example, that if $m^\varepsilon = m(\rho/\varepsilon) = m_0$ for $r = r_0$ and $\varphi = 0$, then $m^\varepsilon = m(\rho/\varepsilon) = m(1 + r_0/\varepsilon) = m_0$ for $r = r_0$ and $\varphi = 2\pi$ by 1-periodicity, and we can assume that $\varphi \in [0, 2\pi)$ in (17). Schematically, such a coefficient model is presented in Fig. 2.

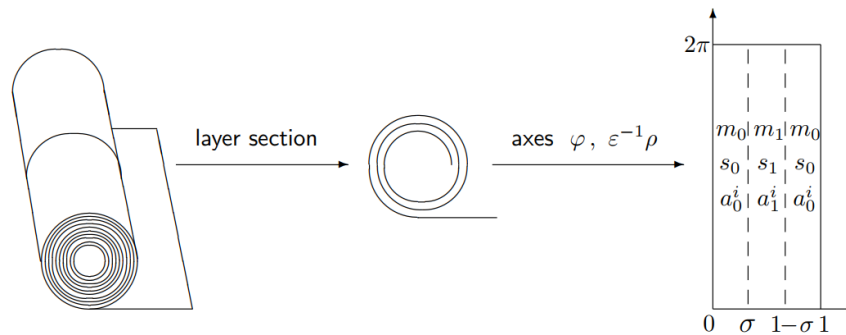


Fig. 2. Model of coefficients of cylindrical rolled layered composites

Further, one may choose an asymptotic approximation for the solution of problem (1) with coefficients (17) in the form

$$u_a = v(t, r, \varphi, z) + \varepsilon v_1(t, r, \varphi, z, \rho/\varepsilon) + \varepsilon^2 v_2(t, r, \varphi, z, \rho/\varepsilon), \quad (18)$$

where, as before, the functions $v_1(t, r, \varphi, z, \tau)$ and $v_2(t, r, \varphi, z, \tau)$ are 1-periodic relative to the variable $\tau \in \mathbb{R}$ by definition.

Hence, the differentiation formula (6) should now be replaced by the formulas

$$\begin{aligned} \partial_r w(t, r, \varphi, z, \rho/\varepsilon) &= (\varepsilon^{-1} \partial_\tau w(t, r, \varphi, z, \tau) + \partial_r w(t, r, \varphi, z, \tau))|_{\tau=\rho/\varepsilon}, \\ \partial_\varphi w(t, r, \varphi, z, \rho/\varepsilon) &= \left((\partial_\varphi + (2\pi)^{-1} \partial_\tau) w(t, r, \varphi, z, \tau) \right)|_{\tau=\rho/\varepsilon} = \\ &= (\partial_\varphi w(t, r, \varphi, z, \tau) + (2\pi)^{-1} \partial_\tau w(t, r, \varphi, z, \tau))|_{\tau=\rho/\varepsilon}, \end{aligned}$$

because $\tau = \rho/\varepsilon = (r/\varepsilon) + (2\pi)^{-1}\varphi$. Taking the last formula into account, we obtain

$$\partial_\varphi (a_2^\varepsilon \partial_\varphi v) = (\partial_\varphi + (2\pi)^{-1} \partial_\tau) (a_2(\tau) \partial_\varphi v) = a_2 \partial_\varphi^2 v + (2\pi)^{-1} (\partial_\tau a_2) \partial_\varphi v, \quad (19)$$

since v does not depend on τ . Therefore, retaining the essential terms, instead of (7) we have

$$\begin{aligned} m v''_{tt} + s v'_t - \varepsilon^{-1} \partial_\tau (a_1 \partial_\tau v_1) - \varepsilon^{-1} \partial_\tau (a_1 \partial_r v) - \\ - \partial_\tau (a_1 \partial_\tau v_2) - \partial_r (a_1 \partial_r v) - \partial_\tau (a_1 \partial_r v_1) - \partial_r (a_1 \partial_\tau v_1) - \\ - r^{-1} a_1 \partial_\tau v_1 - r^{-1} a_1 \partial_r v - r^{-2} a_2 \partial_\varphi^2 v - r^{-2} (2\pi)^{-1} (\partial_\tau a_2) \partial_\varphi v - \\ - a_3 \partial_z^2 v - \varepsilon r^{-1} a_1 \partial_r v_1 - \varepsilon \partial_\tau (a_1 \partial_r v_2) - \varepsilon \partial_r (a_1 \partial_\tau v_1) - \varepsilon \partial_r (a_1 \partial_\tau v_2) = g f \end{aligned} \quad (20)$$

for $\tau = r/\varepsilon$, where $m = m(\tau)$, $s = s(\tau)$, $a_i = a_i(\tau)$ and $g = g(\tau)$ for $i = 1, 2, 3$ in accordance with definition (17).

Further, we get the 1-periodic problem (9) and equality (10), in which the term $r^{-2} a_2 \partial_\varphi^2 v$ should be replaced by

$$r^{-2} a_2 \partial_\varphi^2 v + r^{-2} (2\pi)^{-1} (\partial_\tau a_2) \partial_\varphi v$$

according to (20). Applying the averaging to the new equality (10), we have exactly the homogenized equation (11), since $\langle \partial_\tau (a_2) \rangle = -\langle (a_2) \partial_\tau 1 \rangle = 0$ through using the integration by parts formula and the 1-periodicity of the function $a_2 = a_2(\tau)$. Therefore, one more term $r^{-2} \Phi_2^1 \partial_\varphi v$ should be added to definition (14), where (in addition to problems (15) and similar problems for $\Phi_1, \Phi_2, \Phi_3, \Phi_g$) the function Φ_2^1 is a 1-periodic solution of the problem

$$-\partial_\tau (a_1 \partial_\tau \Phi_2^1) = (2\pi)^{-1} \partial_\tau a_2 \quad \text{in } (0, 1).$$

As before, the regular solution to this problem exists and is uniquely determined up to a constant. Consequently, this allows us to completely determine by definition (14) the function $v_2 = v_2(t, r, \varphi, z, \rho/\varepsilon)$ and achieve the fulfillment of equation (2) up to $O(\varepsilon)$. Therefore, following (Sandakov, 2004b), this can be proven by energetic methods that the last statement implies the accuracy estimate (16). Similar results are valid for the case of heat processes (when formally $m^\varepsilon = 0$ and the condition $u'_t|_{t=0} = u_1$ is removed). The results are natural, since at the micro level cylindrical layered composites and cylindrical rolled layered composites are indistinguishable, at least for wave and heat flows. This may not be true for electric and magnetic flows, since Maxwell's equations form a system of equations.

4. Models of processes in cylindrical layered composites, the layers of which are formed by rings

We have considered cylindrical layered composites which are the union of thin cylindrical tubes. Now we will assume that each tube is a union of rings surrounded by some material. Let, as before, the cylindrical domain without the core is such union. In addition, we will assume that $L = l/\varepsilon$ is an integer for simplicity.

Let us denote by $Y = (0, 1)^2$ the cell, which will be the periodicity cell in what follows. Further, we assume that $Y = Y_0 \cup Y_1 \cup \partial Y_1$ is the union of two open sets with the regular boundary $\partial Y_1 = \partial Y_0$, where Y_0 and Y_1 have non-zero Lebesgue measures. Let's also define the following functions

$$\begin{aligned} m(\tau, \xi) &= m_0 \text{ if } (\tau, \xi) \in Y_0 \text{ and } m(\tau, \xi) = m_1 \text{ if } (\tau, \xi) \in Y_1, \\ a_i(\tau, \xi) &= a_0^i \text{ if } (\tau, \xi) \in Y_0 \text{ and } a_i(\tau, \xi) = a_1^i \text{ if } (\tau, \xi) \in Y_1 \end{aligned} \quad (21)$$

for $i = 1, 2, 3$. Similarly, $s(\tau, \xi)$ is equal to s_0 or s_1 and $g(\tau, \xi)$ is equal to g_0 or g_1 , where all constants $m_0, m_1, a_0^i, a_1^i, g_0$ and g_1 are positive and fixed. Further, the functions are continued by 1-periodicity. But, now 1-periodicity means periodicity with period 1 with respect to the two variables $\tau \in \mathbb{R}$ and $\xi \in \mathbb{R}$.

Next, we define the coefficients of equation (2) by the equalities

$$m^\varepsilon = m(r/\varepsilon, z/\varepsilon), \quad s^\varepsilon = s(r/\varepsilon, z/\varepsilon), \quad a_i^\varepsilon = a_i(r/\varepsilon, z/\varepsilon), \quad g^\varepsilon = g(r/\varepsilon, z/\varepsilon) \quad (22)$$

for $i = 1, 2, 3$. Schematically, such a coefficient model is presented in Fig. 3.

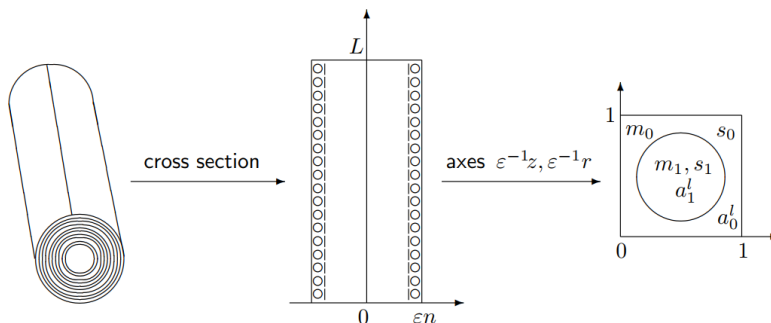


Fig. 3. Model of coefficients of cylindrical composites with layers from rings

Note that the further constructions of asymptotic solution expansions and results do not change for the more general case $Y = Y_0 \cup Y_1 \cup Y_2 \cup \partial Y_1 \cup \partial Y_2$, for example, it can be an insulated wire surrounded by air. For simplicity, we restrict ourselves to the case. It would also be possible to choose a relevant non-degenerate parallelogram as the periodicity cell. But, using a non-degenerate change of variables, the case can always be reduced to the one under consideration.

In this case, to construct an asymptotic expansion, following (Sandakov, 1991), one may choose an asymptotic approximation for the solution of (1) in the form

$$u_a = v(t, r, \varphi, z) + \varepsilon v_1(t, r, \varphi, z, r/\varepsilon, z/\varepsilon) + \varepsilon^2 v_2(t, r, \varphi, z, r/\varepsilon, z/\varepsilon), \quad (23)$$

where the functions $v_1(t, r, \varphi, z, \tau, \xi)$ and $v_2(t, r, \varphi, z, \tau, \xi)$ are 1-periodic relative to $(\tau, \xi) \in \mathbb{R}^2$ and the first term does not depend on $\tau = r/\varepsilon$ and $\xi = z/\varepsilon$, since the kernel of the relevant operator on 1-periodic functions is generated by 1.

Now, to define the terms of asymptotic expansion (23), the following differentiation formulas for functions of a special form $w = w(t, r, \varphi, z, r/\varepsilon, z/\varepsilon)$ will be used

$$\begin{aligned} \partial_r w(t, r, \varphi, z, r/\varepsilon, z/\varepsilon) &= (\varepsilon^{-1} \partial_\tau w(t, r, \varphi, z, \tau, \xi) + \partial_r w(t, r, \varphi, z, \tau, \xi))|_{\tau=r/\varepsilon, \xi=z/\varepsilon}, \\ \partial_z w(t, r, \varphi, z, r/\varepsilon, z/\varepsilon) &= (\varepsilon^{-1} \partial_\xi w(t, r, \varphi, z, \tau, \xi) + \partial_z w(t, r, \varphi, z, \tau, \xi))|_{\tau=r/\varepsilon, \xi=z/\varepsilon}. \end{aligned} \quad (24)$$

Let's substitute asymptotic approximation (23) into equation (2) and take formulas (24) into account. Then, retaining only the essential terms, we obtain

$$\begin{aligned} & -\varepsilon^{-1} \partial_\tau (a_1 \partial_\tau v_1) - \varepsilon^{-1} \partial_\xi (a_3 \partial_\xi v_1) - \varepsilon^{-1} \partial_\tau (a_1 \partial_r v) - \varepsilon^{-1} \partial_\xi (a_3 \partial_z v) - \\ & - \partial_\tau (a_1 \partial_\tau v_2) - \partial_\xi (a_3 \partial_\xi v_2) - \partial_r (a_1 \partial_r v) - \partial_z (a_3 \partial_z v) + m v''_{tt} - \\ & - \partial_\tau (a_1 \partial_r v_1) - \partial_r (a_1 \partial_\tau v_1) - \partial_\xi (a_3 \partial_z v_1) - \partial_z (a_3 \partial_\xi v_1) + s v'_t - \\ & - r^{-1} a_1 \partial_\tau v_1 - r^{-1} a_1 \partial_r v - r^{-2} a_2 \partial_\phi^2 v - \varepsilon \partial_r (a_1 \partial_r v_1) = g f \end{aligned} \quad (25)$$

for $\tau = r/\varepsilon$ and $\xi = z/\varepsilon$, where $m = m(\tau, \xi)$, $s = s(\tau, \xi)$, $a_i = a_i(\tau, \xi)$, $g = g(\tau, \xi)$ according to definition (22).

Equating the coefficients at ε^{-1} in asymptotic equality (25), we have

$$-\partial_\tau (a_1 \partial_\tau v_1) - \partial_\xi (a_3 \partial_\xi v_1) = \partial_\tau (a_1 \partial_r v) + \partial_\xi (a_3 \partial_z v). \quad (26)$$

This equality will always be satisfied if we choose

$$v_1 = \Phi^1(\tau, \xi) \partial_r v(t, r, \varphi, z) + \Phi^3(\tau, \xi) \partial_z v(t, r, \varphi, z)$$

and define the functions $\Phi^1 = \Phi^1(\tau, \xi)$ and $\Phi^3 = \Phi^3(\tau, \xi)$ as solutions of the following 1-periodic problems choose

$$\begin{aligned} & -\partial_\tau (a_1 \partial_\tau \Phi^1) - \partial_\xi (a_3 \partial_\xi \Phi^1) = \partial_\tau (a_1) \quad \text{in } Y = (0, 1)^2, \\ & -\partial_\tau (a_1 \partial_\tau \Phi^3) - \partial_\xi (a_3 \partial_\xi \Phi^3) = \partial_\xi (a_3) \quad \text{in } Y = (0, 1)^2. \end{aligned} \quad (27)$$

It is known (Bensoussan, Lions, & Papanicolaou, 1978) that the solutions to problems (27) exist and are determined up to constants, which will be fixed by the condition

$$\langle \langle \Phi^i \rangle \rangle = \int_0^1 \int_0^1 \Phi^i(\tau, \xi) d\tau d\xi = 0 \quad \text{for } i = 1, 3.$$

Note that the solvability follows from the equalities $\langle \langle \partial_\tau (a_1) \rangle \rangle = -\langle \langle a_1 \partial_\tau 1 \rangle \rangle = 0$. Problems (27) can be considered as problems for an elliptic operator on the 2-torus as a manifold without boundary. It is known (Kinderlehrer, & Stampacchia, 1980) that the solutions are Holder functions, since $a_i \in L^\infty(Y)$ for $i = 1, 3$. Moreover, if the boundary ∂Y_1 is quite regular, then $\Phi^i \in W_{per}^{1, \infty}(Y)$, where the index *per* means 1-periodicity with respect to $(\tau, \xi) \in \mathbb{R}^2$, since the coefficients of these operator are constants and Φ^i are smooth outside ∂Y_1 . So the solutions to the problems are quite regular. In the previous cases, the regularity follows from explicit formulas, for example, for solutions of problem (9).

Equating the coefficients at ε^0 in asymptotic equality (25) and using the choice already made, that $v_1 = \Phi^1 \partial_r v + \Phi^3 \partial_z v$, we obtain

$$\begin{aligned} & -\partial_\tau (a_1 \partial_\tau v_2) - \partial_\xi (a_3 \partial_\xi v_2) = g f - m v''_{tt} - s v'_t + \\ & + (a_1 + a_1 \partial_\tau \Phi^1) \partial_r^2 v + (a_1 \partial_\tau \Phi^3) \partial_r \partial_z v + (a_3 + a_3 \partial_\xi \Phi^3) \partial_z^2 v + (a_3 \partial_\xi \Phi^1) \partial_z \partial_r v + \\ & + r^{-1} (a_1 + a_1 \partial_\tau \Phi^1) \partial_r v + r^{-1} (a_1 \partial_\tau \Phi^3) \partial_z v + r^{-2} a_2 \partial_\phi^2 v \\ & + \partial_\tau (a_1 \Phi^1) \partial_r^2 v + \partial_\tau (a_1 \Phi^3) \partial_r \partial_z v + \partial_\xi (a_3 \Phi^3) \partial_z^2 v + \partial_\xi (a_3 \Phi^1) \partial_z \partial_r v. \end{aligned} \quad (28)$$

By analogy with (26), consider the equality as a 1-periodic problem for the function $v_2 = v_2(\tau, \xi)$ with fixed t, r, φ, z . It is known (Bensoussan, Lions, & Papanicolaou, 1978) that the problem is correct if and only if the right-hand side is orthogonal to 1 in the space $L^2(Y)$. Consequently, we get

$$\begin{aligned} \hat{m} v''_{tt} + \hat{s} v'_t - \hat{a}_1^1 \partial_\tau^2 v - r^{-1} \hat{a}_1^1 \partial_r v - \hat{a}_1^3 \partial_r \partial_z v - r^{-1} \hat{a}_1^3 \partial_r v - \\ - \hat{a}_3^3 \partial_z^2 v - \hat{a}_3^1 \partial_z \partial_r v - r^{-2} \hat{a}_2 \partial_\varphi^2 v = \hat{g} f \quad \text{in } \Omega \times (0, T), \end{aligned} \quad (29)$$

where the following notation was introduced for the constants

$$\begin{aligned} \hat{m} = \langle\langle m \rangle\rangle = \int_0^1 \int_0^1 m(\tau, \xi) d\tau d\xi, \quad \hat{s} = \langle\langle s \rangle\rangle, \quad \hat{g} = \langle\langle g \rangle\rangle, \\ \hat{a}_2 = \langle\langle a_2 \rangle\rangle, \quad \hat{a}_1^1 = \langle\langle a_1 + a_1 \partial_\tau \Phi^1 \rangle\rangle, \quad \hat{a}_1^3 = \langle\langle a_1 \partial_\tau \Phi^3 \rangle\rangle, \\ \hat{a}_3^3 = \langle\langle a_3 + a_3 \partial_\xi \Phi^3 \rangle\rangle, \quad \hat{a}_3^1 = \langle\langle a_3 \partial_\xi \Phi^1 \rangle\rangle. \end{aligned} \quad (30)$$

Thus, starting from equation (2) with complex oscillating coefficients, we have obtained an equation with constant coefficients, which is homogenized. Adding boundary and initial conditions to this equation, as in problem (1), we obtain a homogenized problem. Note that, in the general case, the matrix $\hat{a} = \{\hat{a}_i^j\}_{i,j=1,3}$ is not diagonal and, therefore, homogenized media may be anisotropic. However, it is known (Bensoussan et al., 1978) that the matrix with constant components is both symmetric and elliptic. It is also known that the following inequalities hold in the matrix sense

$$\langle\langle a^{-1} \rangle\rangle^{-1} \leq \hat{a} \leq \langle\langle a \rangle\rangle,$$

where a is the diagonal matrix with diagonal entries $a_1 = a_1(\tau, \xi)$ and $a_3 = a_3(\tau, \xi)$.

To explicitly calculate the elements of the homogenized matrix $\hat{a}$, it is necessary to solve problems (27). In the general case, this is done numerically through computer simulations. Note also that if the coefficients in (21) do not depend on $\xi \in [0, 1]$, that is $m = m(\tau)$, $s = s(\tau)$, $a_i = a_i(\tau)$, $g = g(\tau)$. Then the matrix $\hat{a}$ is diagonal and the homogenized problem (29) coincides with problem (11), where the coefficients are defined in (12), which coincide with (30). Indeed, in the case, the second equation in (27) is trivial and $\Phi^3 = 0$. In addition, the first equation coincides with equation (9).

In turn, if the coefficients in (21) do not depend on $\tau \in [0, 1]$, that is, the coefficients have the form $m = m(\xi)$, $s = s(\xi)$, $a_i = a_i(\xi)$, $g = g(\xi)$. Then the matrix $\hat{a}$ is also diagonal and the homogenized problem (29) coincides with problem (11), but the coefficients of this problem should be determined by the following equalities

$$\begin{aligned} \hat{m} = \langle m \rangle = \int_0^1 m(\xi) d\xi, \quad \hat{s} = \langle s \rangle, \quad \hat{g} = \langle g \rangle, \\ \hat{a}_1 = \langle a_1 \rangle, \quad \hat{a}_2 = \langle a_2 \rangle, \quad \hat{a}_3 = \langle a_3^{-1} \rangle^{-1}. \end{aligned}$$

Therefore, in comparison with (12), good heat (wave) insulation will be in the direction of the axis of the cylinder, which is layered composed of pancakes.

In the general case, the matrix $\hat{a}$ is diagonal if the coefficients $a_1 = a_1(\tau, \xi)$ and $a_3 = a_3(\tau, \xi)$. in definition (21) are symmetric with respect to the coordinate axes passing through the center $(0.5, 0.5)$ of the periodicity cell Y according to (Bakhvalov, & Panasenko, 1989). For example, if Y_1 is a square or circle centered at this point and of suitable size, as in Fig. 2, it is known (Duvaut, & Lions, 1972) that the homogenized problem (29) (with initial and boundary conditions from (1)) has a unique regular solution. Similar results are valid for the case of heat processes (when formally $m^\varepsilon = 0$ and the condition $u'_t|_{t=0} = u_1$ is removed). The solution can be found numerically, which is a standard problem in computational mathematics. In the case of the diagonal matrix $\hat{a}$, the solution can be constructed in the form of a convergent series by the method of separation of variables. Moreover, each term of the series is a particular solution of the homogenized problem with suitable initial conditions

Thus, two terms in the asymptotic approximation (23) for the solution of problem (1) are defined. In order to define the third term $v_2 = v_2(r, \varphi, z, r/\varepsilon, z/\varepsilon)$, we subtract equation (29) from equality (28). Then, we obtain

$$\begin{aligned} -\partial_\tau(a_1 \partial_\tau v_2) - \partial_\xi(a_3 \partial_\xi v_2) = \tilde{g} f - \tilde{m} v''_{tt} - \tilde{s} v'_t + r^{-2} \tilde{a}_2 \partial_\varphi^2 v + \\ + \tilde{a}_1^1 \partial_r^2 v + r^{-1} \tilde{a}_1^r \partial_r v + \tilde{a}_1^3 \partial_r \partial_z v + r^{-1} \tilde{a}_1^z \partial_z v + \tilde{a}_3^3 \partial_z^2 v + \tilde{a}_3^1 \partial_z \partial_r v, \end{aligned} \quad (31)$$

where the coefficients $\tilde{g}$, $\tilde{m}$, $\tilde{s}$, $\tilde{a}_2$, $\tilde{a}_1^1$, $\tilde{a}_1^r$, $\tilde{a}_1^z$, $\tilde{a}_3^3$, $\tilde{a}_3^1$ are 1-periodic functions and are orthogonal to 1 in $L^2(Y)$.

Therefore, equality (31) and therefore equality (28) are fulfilled if we choose

$$\begin{aligned} v_2 = \Phi_g f - \Phi_m v''_{tt} - \Phi_s v'_t + r^{-2} \Phi_2 \partial_\varphi^2 v + \Phi_1^1 \partial_r^2 v + \\ + r^{-1} \Phi_1^r \partial_r v + \Phi_1^3 \partial_r \partial_z v + r^{-1} \Phi_1^z \partial_z v + \Phi_3^3 \partial_z^2 v + \Phi_3^1 \partial_z \partial_r v, \end{aligned} \quad (32)$$

where, for example, the functions $\Phi_3^1 = \Phi_3^1(\tau, \xi)$ is 1-periodic solution of the following problem

$$-\partial_\tau(a_1 \partial_\tau \Phi_3^1) - \partial_\xi(a_3 \partial_\xi \Phi_3^1) = \tilde{a}_3^1 \quad \text{in } Y. \quad (33)$$

The problem is quite similar to problems (27) and is solvable up to a constant fixed by the condition $\langle\langle \Phi_3^1 \rangle\rangle = 0$. Hence, 1-periodic problems for Φ_g , Φ_m , Φ_s , Φ_2 , Φ_1^1 , Φ_1^r , Φ_1^3 , Φ_1^z , Φ_3^3 can be defined in the same way as before.

This approach allows us to completely determine $v_2 = v_2(t, r, \varphi, z, r/\varepsilon, z/\varepsilon)$ and achieve the fulfillment of equation (2) up to $O(\varepsilon)$. Therefore, following (Sandakov, 2004b), this can be verified by energetic methods that the last statement implies, for example, the following accuracy estimate

$$\begin{aligned} \|r^{1/2}(\partial_r u - \theta_1^\varepsilon \partial_r v + r^{-1} \partial_\varphi u - r^{-1} \partial_\varphi v + \partial_z u - \theta_3^\varepsilon \partial_z v)\|_{C^0([0, T]; L^2(\Omega))}^2 \\ + \|u'_t - v'_t\|_{C^0([0, T]; L^2(\Omega))}^2 + \|u - v\|_{C^0([0, T]; L^2(\Omega))}^2 \leq C \varepsilon, \end{aligned} \quad (34)$$

with $\theta_1^\varepsilon = \theta_1(r/\varepsilon, z/\varepsilon)$ and $\theta_3^\varepsilon = \theta_3(r/\varepsilon, z/\varepsilon)$, where $\theta_1(\tau, \xi) = 1 + \partial_\tau \Phi^1 + \partial_\xi \Phi^1$, $\theta_3(\tau, \xi) = 1 + \partial_\tau \Phi^3 + \partial_\xi \Phi^3$, $\Phi^1 = \Phi^1(\tau, \xi)$ and $\Phi^3 = \Phi^3(\tau, \xi)$ are the solutions of problem (27), u is the solution of problem (1) with coefficients from (22), v is the solution of

homogenized problem with coefficients from (30) and the constant C does not depend on ε when $0 < \varepsilon \leq \varepsilon_0$ for the corresponding ε_0 .

Thus, just like before, we have a fairly good accuracy estimate for solutions and flows for sufficiently small ε . In this case, for modeling wave processes (correspondingly, heat processes when $m^\varepsilon = 0$) in such composites, instead of solving problem (1) (for example, numerically through computer simulations), it is possible to solve the homogenized problem for equation (29) with guaranteed accuracy.

It is useful to note that representation (32) as well as (14) contain quite a lot of terms. Therefore, when proving the accuracy estimate (34), it will be necessary to estimate the contribution of each such term. In fact, in order to prove the estimate, it suffices to use only two terms

$$u_a^1 = v(t, r, \varphi, z) + \varepsilon v_1(t, r, \varphi, z, r/\varepsilon, z/\varepsilon)$$

in the asymptotic approximation (23) for the solution of problem (1). Further details can be found in (Sandrakov, 2004b), where the main idea is that the right side of the equation for $w = u - u_a^1$ can contain a term $\operatorname{div}(\varepsilon F)$, which is estimable using integration by parts. But, without understanding the structure of the function $v_2 = v_2(t, r, \varphi, z, r/\varepsilon, z/\varepsilon)$ it would be impossible to find out the exact form of homogenized equation (29), which is the basis of the method here.

5. Models for processes in cylindrical layered composites, the layers of which are formed by helices

We have considered composites which are the union of thin cylindrical tubes, which are formed by a large number of rings separated by a material. Some transformers have a cylindrical layer structure around the core, the layers of which consist of conductor turns in the form of helices. Thus, one can naturally consider a model for a cylindrical layered composite, the layers of which are formed by a helix with many turns separated by some material (or several materials, for example, an insulator and air). This case can be considered by the same homogenization method.

Indeed, let $z \in (0, l)$ and $\varphi \in [0, 2\pi]$ just like before, consider a new coordinate

$$\zeta = z + \varepsilon(\varphi/2\pi).$$

It would be possible to use this coordinate as new instead of z in equation (2). But, we will use the coordinate only in the coefficients defined by (21). Let, as before, we have

$$\Omega = \bigcup_{k=1}^N B_k \times (0, l) = \bigcup_{k=1}^N \{ (x, y) \in \mathbb{R}^2 : \kappa + \varepsilon(k-1) \leq \sqrt{x^2 + y^2} < \kappa + \varepsilon k \} \times (0, l),$$

where B_k is a ring of thickness ε and suitable radii for $k = 1, \dots, N$. Remark that by fixing a positive $\kappa < r_0 < R$, we can consider the following curve in Cartesian coordinates

$$x = r_0 \cos(t), \quad y = r_0 \sin(t), \quad z = \varepsilon(t/2\pi) \quad \text{for } 0 \leq t < \infty, \quad (35)$$

which is called a cylindrical helix and passes through one of the tubes $B_k \times (0, l)$, since $0 < r_0 < R$. Consequently, the curve $r = \sqrt{x^2 + y^2} = r_0$, $z = \varepsilon(\varphi/2\pi)$ is the cylindrical helix also for the parameter $\varphi = t$.

Now, we define the coefficients of equation (2) by the equalities

$$m^\varepsilon = m(r/\varepsilon, \zeta/\varepsilon), \quad s^\varepsilon = s(r/\varepsilon, \zeta/\varepsilon), \quad a_i^\varepsilon = a_i(r/\varepsilon, \zeta/\varepsilon), \quad g^\varepsilon = g(r/\varepsilon, \zeta/\varepsilon) \quad (36)$$

for $i = 1, 2, 3$. Here the functions $m = m(\tau, \xi)$, $s = s(\tau, \xi)$, $a_i = a_i(\tau, \xi)$ and $g = g(\tau, \xi)$ for $i = 1, 2, 3$ are 1-periodic with respect to $(\tau, \xi) \in \mathbb{R}^2$ and are defined by equalities (21) on the periodicity cell $Y = (0, 1)^2$.

Note, for example, that if $m^\varepsilon = m(r/\varepsilon, \zeta/\varepsilon) = m_0$ for $r = r_0$, $\varphi = 0$ and $z = z_0$, then $m^\varepsilon = m(r/\varepsilon, 1 + z_0/\varepsilon) = m_0$ for $r = r_0$, $\varphi = 2\pi$ and $z = z_0$ by 1-periodicity, and we can assume that $\varphi \in [0, 2\pi]$ in (36). In fact, the latter possibility follows from the periodicity of the trigonometric functions in (35) with the period 2π .

Further, to construct an asymptotic expansion, one may choose an asymptotic approximation for the solution of problem (1) with coefficients (36) in the form

$$u_a = v(t, r, \varphi, z) + \varepsilon v_1(t, r, \varphi, z, r/\varepsilon, \zeta/\varepsilon) + \varepsilon^2 v_2(t, r, \varphi, z, r/\varepsilon, \zeta/\varepsilon),$$

where the functions $v_1(t, r, \varphi, z, \tau, \xi)$ and $v_2(t, r, \varphi, z, \tau, \varphi)$ are 1-periodic relative to $(\tau, \xi) \in \mathbb{R}^2$ and the first term does not depend on $\tau = r/\varepsilon$ and $\xi = z/\varepsilon$.

Hence, the differentiation formula (24) should now be replaced by the formulas

$$\begin{aligned} \partial_r w(t, r, \varphi, z, r/\varepsilon, \zeta/\varepsilon) &= (\varepsilon^{-1} \partial_\tau w(t, r, \varphi, z, \tau, \xi) + \partial_r w(t, r, \varphi, z, \tau, \xi))|_{\tau=r/\varepsilon, \xi=\zeta/\varepsilon}, \\ \partial_z w(t, r, \varphi, z, r/\varepsilon, \zeta/\varepsilon) &= (\varepsilon^{-1} \partial_\xi w(t, r, \varphi, z, \tau, \xi) + \partial_z w(t, r, \varphi, z, \tau, \xi))|_{\tau=r/\varepsilon, \xi=\zeta/\varepsilon}, \\ \partial_\varphi w\left(t, r, \varphi, z, \frac{r}{\varepsilon}, \frac{\zeta}{\varepsilon}\right) &= \left((\partial_\varphi + (2\pi)^{-1} \partial_\xi) w(t, r, \varphi, z, \tau, \xi) \right)|_{\tau=\frac{r}{\varepsilon}, \xi=\frac{\zeta}{\varepsilon}} = \\ &= (\partial_\varphi w(t, r, \varphi, z, \tau, \xi) + (2\pi)^{-1} \partial_\xi w(t, r, \varphi, z, \tau, \xi))|_{\tau=r/\varepsilon, \xi=\zeta/\varepsilon} \end{aligned} \quad (37)$$

because $\xi = \zeta/\varepsilon = (z/\varepsilon) + (2\pi)^{-1}\varphi$. Taking the last formula into account, we obtain

$$\partial_\varphi (a_2^\varepsilon \partial_\varphi v) = a_2(\tau, \xi) \partial_\varphi^2 v + (2\pi)^{-1} (\partial_\xi a_2(\tau, \xi)) \partial_\varphi v,$$

since v does not depend on τ and ξ . Consequently, retaining only the essential terms, we get (28), where $r^{-2} a_2 \partial_\varphi^2 v$ is replaced by $r^{-2} a_2 \partial_\varphi^2 v + r^{-2} (2\pi)^{-1} (\partial_\xi a_2) \partial_\varphi v$ for $\tau = r/\varepsilon$, $\xi = \zeta/\varepsilon$, in accordance with (37).

Further, we get the 1-periodic problems (27) and equality (28), in which the term $r^{-2} a_2 \partial_\varphi^2 v$ is replaced by the term

$$r^{-2} a_2 \partial_\varphi^2 v + r^{-2} (2\pi)^{-1} (\partial_\xi a_2) \partial_\varphi v$$

due to (37). Applying the mean to the new equality (28), we get exactly the homogenized equation (29), since $\langle \partial_\xi (a_2) \rangle = -\langle (a_2) \partial_\xi 1 \rangle = 0$. Here we have used the integration by parts and the 1-periodicity of the function $a_2(\tau, \xi)$.

Thus, one more term $r^{-2} \Phi_2^1 \partial_\varphi v$ should be added to definition (32), where, in addition to problem (33) and similar problems for $\Phi_g, \Phi_m, \Phi_s, \Phi_2, \Phi_1^r, \Phi_1^z, \Phi_1^3, \Phi_2^3, \Phi_3^3$, the function Φ_2^1 is a 1-periodic solution of the following problem

$$-\partial_{\tau}(a_1 \partial_{\tau} \Phi_2^1) - \partial_{\xi}(a_3 \partial_{\xi} \Phi_2^1) = (2\pi)^{-1}(\partial_{\xi} a_2) \quad \text{in } Y.$$

Again, the solution to this problem exists and is unique up to a constant, which allows us to completely determine $v_2 = v_2(t, r, \varphi, z, r/\varepsilon, \zeta/\varepsilon)$ and achieve the fulfillment of equation (2) up to $O(\varepsilon)$. Therefore, following (Sandrakov, 2004b), this can be verified by energetic methods that the accuracy estimate (34) is valid. Similar results are valid for the case of heat processes (when formally $m^{\varepsilon} = 0$ and the condition $u'|_{t=0} = u_1$ is removed). The results are natural, since at the micro level cylindrical layered composites, the layers of which are formed by a large number of rings or rings of helices, are indistinguishable, at least for wave and heat flows.

The interested reader may consider as a useful exercise the case in which (instead of equalities (22) or (36)) the coefficients in equation (2) are given by the following relations

$$m^{\varepsilon} = m(\rho/\varepsilon, \zeta/\varepsilon), \quad s^{\varepsilon} = s(\rho/\varepsilon, \zeta/\varepsilon), \quad a_i^{\varepsilon} = a_i(\rho/\varepsilon, \zeta/\varepsilon), \quad g^{\varepsilon} = g(\rho/\varepsilon, \zeta/\varepsilon)$$

where the functions $m = m(\tau, \xi)$, $s = s(\tau, \xi)$, $a_i = a_i(\tau, \xi)$ and $g = g(\tau, \xi)$ for $i = 1, 2, 3$ are defined by equalities (21) and the following coordinates $\rho = r + \varepsilon(\varphi/2\pi)$, $\zeta = z + \varepsilon(\varphi/2\pi)$ are used. Here, as usual, one can take $(r, \varphi, z) \in (0, R) \times [0, 2\pi) \times (0, l)$. We only note that the corresponding curves will be conical helical lines. An inquisitive reader can also try to consider this problem using the multiscale convergence method.

Discussion and conclusions

Initial boundary value problems for dynamic parabolic heat conduction equations and hyperbolic wave equations in periodic media having a cylindrical structure are considered. Analytical methods and an algorithm for determining asymptotic expansions and homogenized problems, the solutions of which provide approximate multi-scale asymptotic solutions of such problems, have been established. Asymptotic accuracy estimates are presented, which are actual in numerical calculations and computer simulation for models with guaranteed accuracy. Thus, it is shown in sufficient detail that homogenization methods can be useful for modeling wave and thermal processes in a cylinder consisting of alternating layers with different characteristics, which are similar to capacitors and transformers. This approach can also be implemented for modeling filtration processes in cylindrical filters. Undoubtedly, the extension of the methods to Maxwell's equations is very relevant and promising for modeling processes in various microdevices. The approach developed here can really be generalized to such a system of equations and is asymptotically accurate, since it takes into account the dependence on the parameters and geometry of the layers quite essentially.

Authors' contribution: Gennadiy Sandrakov – conceptualization, methodology and theoretical foundations of research; Mykola Kyrychek – analysis of sources, preparation of literature review and research.

Acknowledgments, funding sources. This work has been supported partially by Grant of the Ministry of Education and Science of Ukraine for perspective development of a scientific direction "Mathematical sciences and natural sciences" at Taras Shevchenko National University of Kyiv.

References

- Allaire, G. (1992). Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6), 1482–1518. <https://doi.org/10.1137/0523084>.
- Anikushyn, A., & Hranishak, K. (2025). Point control of linear hyperbolic integro-differential systems. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 79(2), 20–28. <https://doi.org/10.17721/1812-5409.2024/2.4>
- Anikushyn, A., Lyashko, V., & Samosonok, O. (2024). Optimal control of the systems governed by linear hyperbolic integro-differential equations. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 78(1), 111–118. <https://doi.org/10.17721/1812-5409.2024/1.22>
- Bakhvalov, N. S., & Panasenko, G. P. (1989). *Homogenization: averaging processes in periodic media*. Kluwer, Dordrecht.
- Bensoussan, A., Lions, J. L., & Papanicolaou, G. (1978). *Asymptotic analysis for periodic structures*. Amsterdam, North-Holland.
- Bertotti, G. (1998). *Hysteresis in magnetism*. Academic Press.
- Cioranescu, D., Damilamian, A., & Griso, G. (2018). *The periodic unfolding method*. Springer Nature, Singapore. <https://doi.org/10.1007/978-981-13-3032-2>
- Cioranescu, D., Damilamian, A., & Griso, G. (2023). The periodic unfolding method for countably many scales. *Advances in Mathematical Sciences and Applications*, 32(1), 17–58.
- Diaz, J. I., Gomez-Castro, D., Shaposhnikova, T. A., & Zubova, M. N. (2019a). A nonlocal memory strange term arising in the critical scale homogenization of a diffusion equation with dynamic boundary conditions. *Electronic J. Differential Equations*, 77, 1–13.
- Diaz, J. I., Gomez-Castro, D., Shaposhnikova, T. A., & Zubova, M. N. (2019b). Classification of homogenized limits of diffusion problems with spatially dependent reaction over critical-size particles. *Applicable Analysis*, 98, 232–255. <https://doi.org/10.1080/00036811.2018.1441997>
- Duvaut, G., & Lions, J. L. (1972). *Les inéquations en mécanique et en physique*. Dunod.
- Garttner, S., Frolkovic, P., Knabner, P., & Ray, N. (2022). Efficiency of micro-macro models for reactive two-mineral systems. *Multiscale Modeling & Simulation*, 20(1), 433–461. <https://doi.org/10.1137/20M1380648>
- Garttner, S., Knabner, P., & Ray, N. (2024). Local existence of strong solutions to micro-macro models for reactive transport in evolving porous media. *European Journal of Applied Mathematics*, 35(1), 127–154. <https://doi.org/10.1017/S095679252300013X>
- Jager, W., & Woukeng, L. (2021). Homogenization of Richards' equations in multiscale porous media with soft inclusions. *Journal of Differential Equations*, 281, 503–549. <https://doi.org/10.1016/j.jde.2021.02.012>
- Kapustyan, O., Kasimova, N., Sobchuk, V., & Stanzhytskyi, O. (2025). The averaging method for the optimal control problem of a parabolic inclusion with fast-oscillating coefficients on a finite time interval. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 79(2), 33–40. <https://doi.org/10.17721/1812-5409.2024/2.6>
- Kinderlehrer, D., & Stampacchia, G. (1980). *An introduction to variational inequalities and their applications*. Academic Press.
- Landau, L. D., & Lifshitz, E. M. (1984). *Electrodynamics of continuous media*. Pergamon Press.
- Mielke, A., Reichelt, S., & Thomas, M. (2014). Two-scale homogenization of nonlinear reaction-diffusion systems with slow diffusion. *Networks and Heterogeneous Media*, 9(2), 353–382. <https://doi.org/10.3934/nhm.2014.9.353>
- Piat, V. C., & Sandrakov, G. V. (2004). Homogenization of some variational inequalities for elasto-plastic torsion problems. *Asymptotic Analysis*, 40(1), 1–23. <https://doi.org/10.3233/asy-2004-637>
- Ray, N., & Schulz, R. (2023). Existence and uniqueness of solutions to a flow and transport problem with degenerating coefficients. *European Journal of Applied Mathematics*, 34(1), 55–76. <https://doi.org/10.1017/S0956792522000018>
- Saedodin, S., & Barforoush, M. S. M. (2017). An exact solution for thermal analysis of a cylindrical object using hyperbolic thermal conduction model. *Thermophysics and Aeromechanics*, 24(6), 909–920. <https://doi.org/10.1134/S0869864317060099>
- Sandrakov, G. V. (1991). Averaging principles for equations with rapidly oscillating coefficients. *Mathematics of the USSR – Sbornik*, 68(2), 503–553. <https://doi.org/10.1070/SM1991v068n02ABEH002111>
- Sandrakov, G. V. (1998). Homogenization of nonstationary problems of the theory of strong nonhomogeneous elastic media. *Doklady Akademii Nauk*, 358(3), 308–311.
- Sandrakov, G. V. (2004a). Homogenization of variational inequalities for problems with regular obstacles. *Doklady Akademii Nauk*, 397(2), 170–173.

- Sandrakov, G. V. (2004b). Multiphase models of nonstationary diffusion arising from homogenization. *Computational Mathematics and Mathematical Physics*, 44(10), 1741–1756.
- Sandrakov, G. V. (2005). Homogenization of variational inequalities for obstacle problems. *Sbornik: Mathematics*, 196(4), 541–560. <https://doi.org/10.1070/SM2005v196n04ABEH000891>
- Vagner, I. D., Lembrikov, B. I., & Wyder, P. (2004). *Electrodynamics of magnetoactive media*. Springer-Verlag.

Отримано редакцією журналу / Received: 13.01.25

Прорецензовано / Revised: 16.02.25

Схвалено до друку / Accepted: 13.03.25

Геннадій САНДРАКОВ, д-р фіз.-мат. наук, провід. наук. співроб.

ORCID ID: 0000-0001-9133-605X

e-mail: genadiy.sandrakov@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Микола КИРИЧЕК, студ.

e-mail: mykolakyruchek@gmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

МОДЕЛЮВАННЯ І ГОМОГЕНІЗАЦІЯ ПРОЦЕСІВ У ЦИЛІНДРИЧНИХ ШАРУВАТИХ КОМПОЗИТАХ, ТАКИХ ЯК КОНДЕНСАТОРИ ТА ТРАНСФОРМАТОРИ

Конденсатори та трансформатори є допоміжними пристроями, які широко використовуються в сучасному житті. Нині спостерігається тенденція до мініатюризації таких пристроїв. Отже, виникає природна проблема моделювання теплових і електромагнітних процесів у них. Розуміння моделей зазначених процесів дає змогу, наприклад, оптимізувати або здешевити такі пристрої. Проте геометрія таких пристроїв дуже специфічна. Звичайний конденсатор, наприклад, являє собою згорнуту в циліндр довгу стрічку, укриту ізолятором. Безпосереднє моделювання процесів у такому пристрої складне, оскільки існує дуже багато шарів, що призводить до рівнянь зі швидко осцилюючими коефіцієнтами. З іншого боку, до таких рівнянь можна застосувати методи теорії гомогенізації й отримати рівняння зі сталими коефіцієнтами, розв'язки яких апроксимують розв'язки рівнянь, які розглядаємо. Зрозуміло, що значно простіше розв'язувати рівняння зі сталими коефіцієнтами як чисельними, так й аналітичними методами. Як перший крок у цьому напрямі, у пропонованому дослідженні розглянуто задачу моделювання хвильових і теплових процесів у циліндрі, що складається з великої кількості шарів з різними характеристиками, які чергуються. За допомогою асимптотичних методів теорії гомогенізації цю задачу наближено зведено до моделювання процесів в однорідному (гомогенізованому) циліндрі. Наведено асимптотичні розклади та гомогенізовані задачі, розв'язки яких визначають наближену асимптотику розв'язків задачі, яку ми розглядаємо. Наведено формули для розрахунку характеристик однорідного циліндра й оцінки точності наближень. Подано також оцінки точності цих апроксимацій, що необхідні для чисельних розрахунків і комп'ютерного моделювання таких проблем з гарантованою точністю. Крім того, наведені наближення, безсумнівно, корисні для розуміння хвильових і теплових процесів у пристроях, подібних до конденсаторів і трансформаторів. Зазначений підхід можна розглядати як основу для моделювання загальніших електромагнітних і теплових процесів, що моделюються рівняннями Максвелла в поєднанні з рівняннями теплопровідності.

Ключові слова: початково-крайові задачі, гомогенізовані задачі, хвильові рівняння, рівняння теплопровідності, асимптотичні розв'язки, коефіцієнти теплопровідності, коефіцієнти швидкості поширення хвиль.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 531.8

DOI: <https://doi.org/10.17721/1812-5409.2025/1.11>

Kostiantyn KHOROSHEV, PhD (Phys. & Math.), Assoc. Prof.

ORCID ID: 0000-0002-3900-7838

e-mail: khoroshev.kostiantyn@iit.kpi.ua

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Kyryl DUCHENKO, Student

ORCID ID 0009-0002-7987-6891

e-mail: ge85fix@mytum.de

Technical University of Munich, Munich, Germany

Serhii KYKOT, PhD (Phys. & Math.), Assoc. Prof.

ORCID ID: 0000-0003-1035-8711

e-mail kykot.serhii@iit.kpi.ua

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

KINETOSTATIC ANALYSIS OF ASSUR DYADS

The core of computer-aided design systems for design and calculation of plane mechanisms is the modular mathematical methods of studying the dynamics of the mechanisms. The most common approach is Multibody Dynamics, where a mechanism is divided into links that are assumed to be rigid bodies. The number of links is equal to the number of analysis modules. Another approach takes into consideration the structure of a mechanism, which can allow a reduction in the number of modules that need analysis. This approach necessitates the development of a plane mechanism classification and a comprehensive set of mathematical methods for analyzing the kinematics, kinetostatics, and dynamics of all possible structural elements. Of all the classification systems, Artobolevsky's one is the most aligned with the stated needs. Class II plane mechanisms according to the pointed classification have been fully covered by modular kinematic analysis algorithms, but kinetostatic analysis remains limited to a graph-analytical approach. The paper presents a unified analytical method of kinetostatic analysis of Assur dyads, which, along with class I mechanisms, constitute the structural elements of the class II mechanisms. The method is based on the vector equations of kinetostatics, solved using vector algebra tools. A full kinematic analysis must be performed beforehand. Active forces, inertial forces and reactions in kinematic pairs are represented as vectors, decomposed in the local bases rigidly attached to the dyad's links. The bases must be determined during the position analysis. The presented method is appropriate to use in combination with kinematic analysis, which also relies on vector algebra methods. It can be algorithmized and integrated into computer-aided design systems for class II plane mechanisms.

Key words: modular approach, kinetostatic analysis, class II plane mechanisms, Assur dyads.

AMS 2020 classification: 70-10, 70B15, 70E99.

Introduction

Modern computer-aided design systems for plane mechanisms are based on mathematical methods of mechanism analysis, using a modular approach. The most common approach, multibody dynamics, involves decomposition of mechanisms into links that adhere to the principles of rigid body dynamics or analytical mechanics (Nikravesh, 1988; Roberson & Schwertassek, 1988). A classic example of this method is the library "Modelica. Mechanics. MultiBody" in the Modelica programming language (Otter, Elmqvist, & Mattsson, 2003). Another approach considers the structural properties of the mechanism (Brát, & Lederer, 1973; Hansen, 1996; Simionescu, 2016). The advantage of the last one is the reduction in the number of modules, which requires a well-developed classification of mechanisms and a complete set of mathematical methods of kinematic, kinetostatic, and dynamic analyses of any possible structural element.

The objects of the research are Assur dyads that form the II class of structural groups according to Artobolevsky's classification (Artobolevski, 1977). Assur dyads, together with class I mechanisms, form a set of structural elements of class II mechanisms in accordance with Artobolevsky's classification. This is the only class of mechanisms with completely algorithmized kinematic analysis. Projection approaches solving vector contour equations that describe the position of structural groups are given in (Artobolevski, 1977; Kinytskyi, 2002). The method of coordinate planes can be found in (Honchar, 2011). The vector solution of the Chace equation (Chace, 1963) enables the algorithmization of Assur dyads kinematics using vector algebra tools (Khoroshev, Duchenko, & Kykot, 2023; 2024).

There is still no general analytical approach to kinetostatic analysis of class II structural groups; grapho-analytical is the only known one (Kinytskyi, 2002; Honchar, 2011). There are some separate investigations devoted to analytical methods of kinetostatic analysis of specific plane mechanisms or kinematic chains. For instance, the paper (Durango, Calle, & Ruiz, 2010) presents an analytical method of kinetostatic analysis of the RRR-type dyad with consideration of frictional forces. Additionally, see (Uicker, Pennock, & Shigley, 2016; Jomartov, & Tuleshov, 2018; Abhery, 2022).

The aim of the research is to develop a unified approach of kinetostatic analysis of Assur dyads, which can be implemented in computer-aided design systems of class II mechanisms according to Artobolevsky.

1. General statements

Consider a plane kinematic chain with zero degree of freedom consisting of two links $D_m E$, $D_n E$ and three kinematic (rotational or translational) pairs D_m , E and D_n (Fig. 1). The kinematic pairs allow relative movements of the connected links in the plane of motion of the kinematic chain. Such a kinematic chain is called an Assur dyad and it is a class II structural group according to Artobolevsky's classification. Five types of dyads are distinguished (Artobolevski, 1977): RRR dyad – all pairs are rotational; RRT dyad – two pairs are rotational, one outer pair is translational; RTR dyad – two pairs are rotational, an inner one is translational; TRT dyad – two pairs are translational, one inner pair is rotational; RTT dyad – two pairs are translational, an outer one is rotational.

The numbers m and n are assigned to the links. Suppose that the considered kinematic chain is a part of a plane mechanism, kinematic pairs D_m and D_n are formed by the links of the group and external to the group links numbered k and p . Distances between kinematic pairs are $|D_mE|=l_m$ and $|D_nE|=l_n$. Points D_k and D_p are taken on links k and p . If pairs D_m and D_n are translational, points D_k and D_p should be chosen on the guidelines of the relative movement of the links connected in a kinematic pair. Assume $|D_kD_m|=l_k > 0$ and $|D_pD_n|=l_p > 0$.

The structural group is referred to a fixed rectangular coordinate system $D_0x_0y_0z_0$ with a right standard basis $\{i_0; j_0; k_0\}$ in $\mathbb{R}^3$. Unit vectors i_0 and j_0 lie in the plane of the group. Consider local coordinate systems $D_mx_my_mz_m$, $D_nx_ny_nz_n$, $D_kx_ky_kz_k$ and $D_px_py_pz_p$ with bases $\{i_m; j_m=k_m \times i_m; k_m=k_0\}$, $\{i_n; j_n=k_n \times i_n; k_n=k_0\}$, $\{i_k; j_k=k_k \times i_k; k_k=k_0\}$, $\{i_p; j_p=k_p \times i_p; k_p=k_0\}$ connected with links m , n , k and p so, that unit vectors i_m , i_n , i_k and i_p lie on lines D_mE , D_nE , D_kD_m and D_pD_n correspondently. In Figure 1 unit vectors k_0 , k_m , k_n , k_k and k_p are indicated implicitly, the force diagrams of structural groups are presented in the planes of their motion.

The dyad is a statically determined mechanical system (Artobolevski, 1977), that is, it undergoes a kinetostatic analysis, the purpose of which is to determine the unknown forces, taking into account the inertia forces of the links. We will focus on the determination of unknown reactions in kinematic pairs.

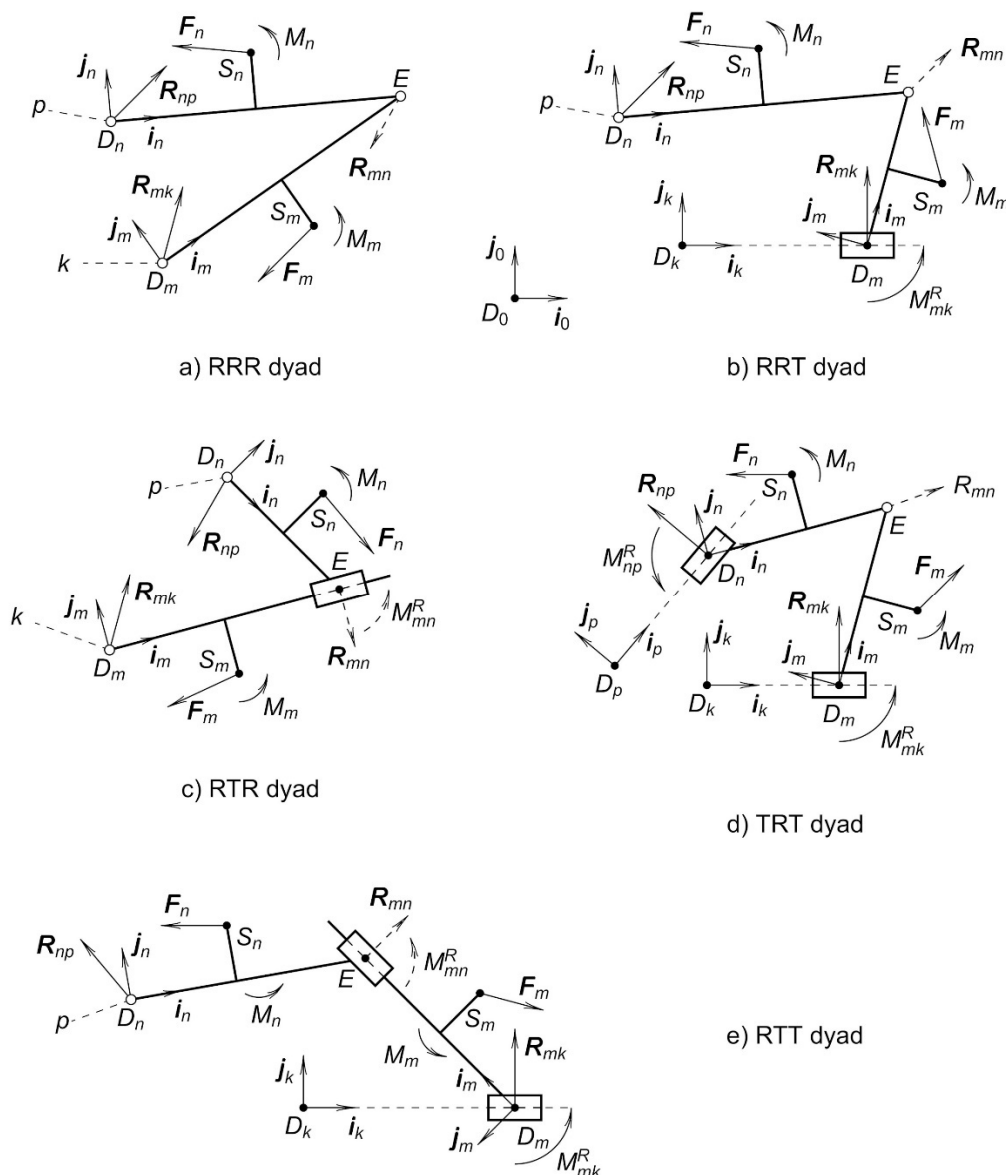


Fig. 1. Force diagrams of Assur dyads

Before compiling the equations of kinetostatics, it is assumed that a full kinematic analysis of the structural group was conducted in any way, so the positions, absolute velocities and accelerations of all points of the dyad, unit vectors $\mathbf{i}_m$, $\mathbf{i}_n$, absolute angular velocities ω_m , ω_n and angular accelerations ε_m , ε_n of all links of the group are known.

Assume γ is number of a link from the set $\{m, n\}$. The radius vector of the center of mass S_γ of the link γ , and its absolute acceleration $\mathbf{w}_{S_\gamma}$ can be presented in a form

$$\mathbf{r}_{S_\gamma} = \mathbf{r}_{D_\gamma} + x_{S_\gamma} \mathbf{i}_\gamma + y_{S_\gamma} \mathbf{j}_\gamma, \quad \mathbf{w}_{S_\gamma} = \mathbf{w}_{D_\gamma} - (\omega_\gamma^2 x_{S_\gamma} + \varepsilon_\gamma y_{S_\gamma}) \mathbf{i}_\gamma - (\omega_\gamma^2 y_{S_\gamma} - \varepsilon_\gamma x_{S_\gamma}) \mathbf{j}_\gamma, \quad (1)$$

where x_{S_γ} , y_{S_γ} are given coordinates of S_γ in coordinate systems $D_\gamma x_\gamma y_\gamma z_\gamma$; $\mathbf{r}_{D_\gamma}$ and $\mathbf{w}_{D_\gamma}$ are known radius vectors and absolute accelerations of points D_γ .

Using equation (1), for resultant vector Φ_γ and resultant moment $\mathbf{M}_\gamma^\Phi$ of inertia forces of link γ in basis $\{\mathbf{i}_\gamma; \mathbf{j}_\gamma; \mathbf{k}_\gamma\}$ we write

$$\Phi_\gamma = -m_\gamma \left((\mathbf{w}_{D_\gamma} \cdot \mathbf{i}_\gamma - \omega_\gamma^2 x_{S_\gamma} - \varepsilon_\gamma y_{S_\gamma}) \mathbf{i}_\gamma + (\mathbf{w}_{D_\gamma} \cdot \mathbf{j}_\gamma - \omega_\gamma^2 y_{S_\gamma} + \varepsilon_\gamma x_{S_\gamma}) \mathbf{j}_\gamma \right), \quad \mathbf{M}_\gamma^\Phi = -J_\gamma \varepsilon_\gamma \mathbf{k}_\gamma. \quad (2)$$

Here m_γ and J_γ are the mass of link γ and its moment of inertia relative to the axis, passing through the center of mass of the link perpendicular to the plane of the structural group.

Assume that link γ undergoes a known active load in the plane of the group, which is given by resultant force $\mathbf{F}_\gamma$ at the center of mass S_γ of the link, and resultant moment $\mathbf{M}_\gamma$. Since the resultant vectors of active forces and inertial forces are known and applied to the same point, we will simplify the presentation by using symbols $\mathbf{F}_\gamma$ and $\mathbf{M}_\gamma$ to represent $\mathbf{F}_\gamma + \Phi_\gamma$ and $\mathbf{M}_\gamma + \mathbf{M}_\gamma^\Phi$. It is obvious that $\mathbf{F}_\gamma \perp \mathbf{k}_\gamma$ and $\mathbf{M}_\gamma \parallel \mathbf{k}_\gamma$. The force diagrams presented in Figure 1 are made taking into account the proposed simplification. Moments are marked with their algebraic values.

All kinematic pairs are assumed to be free of any kind of friction. If a kinematic pairs in point D_γ is rotational, then the reactive forces arising from the external link can be written in the form of a uniform force $\mathbf{R}_{\gamma\delta} = (\mathbf{R}_{\gamma\delta} \cdot \mathbf{i}_\alpha) \mathbf{i}_\alpha + (\mathbf{R}_{\gamma\delta} \cdot \mathbf{j}_\alpha) \mathbf{j}_\alpha$, applied to point D_γ . Here $\delta = k$ for $\gamma = m$ and $\delta = p$ for $\gamma = n$; $\{\mathbf{i}_\alpha; \mathbf{j}_\alpha; \mathbf{k}_\alpha\}$ is an arbitrary from the described above bases; $(\mathbf{R}_{\gamma\delta} \cdot \mathbf{i}_\alpha)$, $(\mathbf{R}_{\gamma\delta} \cdot \mathbf{j}_\alpha)$ are the components of reaction $\mathbf{R}_{\gamma\delta}$ in this basis. In case of translational pair in point D_γ , reactive forces from the external link can be presented in the form of a resultant vector $\mathbf{R}_{\gamma\delta} = (\mathbf{R}_{\gamma\delta} \cdot \mathbf{j}_\beta) \mathbf{j}_\beta$ and resultant moment $\mathbf{M}_{\gamma\delta}^R = (\mathbf{M}_{\gamma\delta}^R \cdot \mathbf{k}_\beta) \mathbf{k}_\beta$, where $\{\mathbf{i}_\beta; \mathbf{j}_\beta; \mathbf{k}_\beta\}$ is the basis, which is related to the guideline of the relative movement of the links in the kinematic pair. Similar considerations and notations are performed for the representation of reactive forces in the internal pair in point E . Assume $\mathbf{R}_{mn}$, and in case it is needed $\mathbf{M}_{mn}^R$, are the vectors describing reactive forces in point E , acting on link m from the side of link n . Unknown reactions will be determined if the relations for calculation of the vectors or their components in any known basis are obtained.

2. The kinetostatic analysis of RRR dyad

Given: D_γ and E are rotational pairs, $l_\gamma > 0$, $\{\mathbf{i}_\gamma; \mathbf{j}_\gamma; \mathbf{k}_\gamma\}$, $\mathbf{F}_\gamma$, $\mathbf{M}_\gamma$ ($\gamma \in \{m, n\}$). Find: $\mathbf{R}_{mk}$, $\mathbf{R}_{np}$ and $\mathbf{R}_{mn}$.

The force diagram is presented in Figure 1(a). Consider resolution of vectors $\mathbf{R}_{mk}$ and $\mathbf{R}_{np}$ in given bases $\{\mathbf{i}_m; \mathbf{j}_m; \mathbf{k}_m\}$ and $\{\mathbf{i}_n; \mathbf{j}_n; \mathbf{k}_n\}$. The moment equation for link γ with respect to point E can be expressed in the form

$$\sum \mathbf{M}_{E\gamma} = (-l_\gamma \mathbf{i}_\gamma) \times \mathbf{R}_{\gamma\delta} + ((x_{S_\gamma} - l_\gamma) \mathbf{i}_\gamma + y_{S_\gamma} \mathbf{j}_\gamma) \times \mathbf{F}_\gamma + \mathbf{M}_\gamma = 0. \quad (3)$$

After elementary transformations, we obtain

$$\mathbf{R}_{\gamma\delta} \cdot \mathbf{j}_\gamma = l_\gamma^{-1} \left\{ (x_{S_\gamma} - l_\gamma) (\mathbf{F}_\gamma \cdot \mathbf{j}_\gamma) - y_{S_\gamma} (\mathbf{F}_\gamma \cdot \mathbf{i}_\gamma) + (\mathbf{M}_\gamma \cdot \mathbf{k}_\gamma) \right\}. \quad (4)$$

Equilibrium equation

$$\sum \mathbf{F}_{(m,n)} = \mathbf{R}_{mk} + \mathbf{F}_m + \mathbf{F}_n + \mathbf{R}_{np} = 0 \quad (5)$$

for the whole structural group can be rewritten as

$$(\mathbf{R}_{mk} \cdot \mathbf{i}_m) \mathbf{i}_m + \{(\mathbf{R}_{mk} \cdot \mathbf{j}_m) \mathbf{j}_m + \mathbf{F}_m + \mathbf{F}_n + (\mathbf{R}_{np} \cdot \mathbf{j}_n) \mathbf{j}_n\} + (\mathbf{R}_{np} \cdot \mathbf{i}_n) \mathbf{i}_n = 0. \quad (6)$$

In the case $\mathbf{i}_m \cdot \mathbf{j}_n \neq 0$, scalar multiplication of the last equation by either vector $\mathbf{j}_m$ or vector $\mathbf{j}_n$ leads to

$$\begin{aligned} \mathbf{R}_{mk} \cdot \mathbf{i}_m &= -(\mathbf{i}_m \cdot \mathbf{j}_n)^{-1} \{(\mathbf{R}_{mk} \cdot \mathbf{j}_m) (\mathbf{j}_m \cdot \mathbf{j}_n) + \mathbf{F}_m \cdot \mathbf{j}_n + \mathbf{F}_n \cdot \mathbf{j}_n + \mathbf{R}_{np} \cdot \mathbf{j}_n\}, \\ \mathbf{R}_{np} \cdot \mathbf{i}_n &= -(\mathbf{i}_n \cdot \mathbf{j}_m)^{-1} \{(\mathbf{R}_{np} \cdot \mathbf{j}_n) (\mathbf{j}_n \cdot \mathbf{j}_m) + \mathbf{F}_n \cdot \mathbf{j}_m + \mathbf{F}_m \cdot \mathbf{j}_m + \mathbf{R}_{mk} \cdot \mathbf{j}_m\}. \end{aligned} \quad (7)$$

Expressions (4) and (7) allow the calculation of all non-zero components of reactions $\mathbf{R}_{mk}$ and $\mathbf{R}_{np}$ in bases $\{\mathbf{i}_m; \mathbf{j}_m; \mathbf{k}_m\}$ and $\{\mathbf{i}_n; \mathbf{j}_n; \mathbf{k}_n\}$ respectively. From equilibrium equation for link m , we obtain the expression for determination of reaction $\mathbf{R}_{mn}$ in the inner pair in point E

$$\mathbf{R}_{mn} = -(\mathbf{R}_{mk} + \mathbf{F}_m) . \quad (8)$$

3. The kinetostatic analysis of TRT dyad

Given: D_m and D_n are translational pairs, E is a rotational pair, $I_\gamma \geq 0$, $\{\mathbf{i}_\gamma; \mathbf{j}_\gamma; \mathbf{k}_\gamma\}$, $\mathbf{F}_\gamma$, $\mathbf{M}_\gamma$ ($\gamma \in \{m, n\}$, if $I_\gamma = 0$, then $\mathbf{i}_\gamma = \mathbf{i}_\delta$), $\{\mathbf{i}_k; \mathbf{j}_k; \mathbf{k}_k\}$, $\{\mathbf{i}_p; \mathbf{j}_p; \mathbf{k}_p\}$. Find: $\mathbf{R}_{mk}$, $\mathbf{M}_{mk}^R$, $\mathbf{R}_{np}$, $\mathbf{M}_{np}^R$ and $\mathbf{R}_{mn}$.

The force diagram is presented in Figure 1(d). Since kinematic pairs in D_γ are translational, then $\mathbf{R}_{\gamma\delta} \parallel \mathbf{j}_\delta$. Equilibrium equation (5) for the whole structural group can be expressed as

$$(\mathbf{R}_{mk} \cdot \mathbf{j}_k) \mathbf{j}_k + \{\mathbf{F}_m + \mathbf{F}_n\} + (\mathbf{R}_{np} \cdot \mathbf{j}_p) \mathbf{j}_p = 0 . \quad (9)$$

If $\mathbf{i}_k \cdot \mathbf{j}_p \neq 0$, then, after scalar multiplication of equation (9) by either vector $\mathbf{i}_k$ or vector $\mathbf{i}_p$, we obtain

$$\mathbf{R}_{mk} \cdot \mathbf{j}_k = -(\mathbf{j}_k \cdot \mathbf{i}_p)^{-1} \{\mathbf{F}_m \cdot \mathbf{i}_p + \mathbf{F}_n \cdot \mathbf{i}_p\}, \quad \mathbf{R}_{np} \cdot \mathbf{j}_p = -(\mathbf{j}_p \cdot \mathbf{i}_k)^{-1} \{\mathbf{F}_m \cdot \mathbf{i}_k + \mathbf{F}_n \cdot \mathbf{i}_k\} . \quad (10)$$

Reactions $\mathbf{R}_{\gamma\delta}$ are therefore determined. To obtain relations for reactive moments $\mathbf{M}_{\gamma\delta}^R$ moment equation for link γ with respect to point E can be written. Finally, we get

$$\mathbf{M}_{\gamma\delta}^R = I_\gamma (\mathbf{i}_\gamma \times \mathbf{R}_{\gamma\delta}) + ((I_\gamma - x_{S\gamma}) \mathbf{i}_\gamma - y_{S\gamma} \mathbf{j}_\gamma) \times \mathbf{F}_\gamma - \mathbf{M}_\gamma . \quad (11)$$

Reaction $\mathbf{R}_{mn}$, as in the previous case, can be calculated using formula (8).

4. The kinetostatic analysis of RRT dyad

Given: D_n and E are rotational pairs, D_m is a translational pair, $I_n > 0$, $I_m \geq 0$, $\{\mathbf{i}_\gamma; \mathbf{j}_\gamma; \mathbf{k}_\gamma\}$, $\mathbf{F}_\gamma$, $\mathbf{M}_\gamma$ ($\gamma \in \{m, n\}$), $\{\mathbf{i}_k; \mathbf{j}_k; \mathbf{k}_k\}$. If $I_m = 0$, then $\mathbf{i}_m = \mathbf{i}_k$. Find: $\mathbf{R}_{mk}$, $\mathbf{M}_{mk}^R$, $\mathbf{R}_{np}$ and $\mathbf{R}_{mn}$.

The force diagram is presented in Figure 1(b). The component $\mathbf{R}_{np} \cdot \mathbf{j}_n$ can be calculated by formula (4). Equilibrium equation (5) for the whole structural group is presented as

$$(\mathbf{R}_{mk} \cdot \mathbf{j}_k) \mathbf{j}_k + \{\mathbf{F}_m + \mathbf{F}_n + (\mathbf{R}_{np} \cdot \mathbf{j}_n) \mathbf{j}_n\} + (\mathbf{R}_{np} \cdot \mathbf{i}_n) \mathbf{i}_n = 0 \quad (12)$$

and, after scalar multiplication by either vector $\mathbf{i}_k$ or $\mathbf{j}_n$, for $\mathbf{i}_n \cdot \mathbf{i}_k \neq 0$, it transforms into

$$\begin{aligned} \mathbf{R}_{mk} \cdot \mathbf{j}_k &= -(\mathbf{j}_k \cdot \mathbf{j}_n)^{-1} \{\mathbf{F}_m \cdot \mathbf{j}_n + \mathbf{F}_n \cdot \mathbf{j}_n + \mathbf{R}_{np} \cdot \mathbf{j}_n\}, \\ \mathbf{R}_{np} \cdot \mathbf{i}_n &= -(\mathbf{i}_n \cdot \mathbf{i}_k)^{-1} \{(\mathbf{R}_{np} \cdot \mathbf{j}_n)(\mathbf{j}_n \cdot \mathbf{i}_k) + \mathbf{F}_n \cdot \mathbf{i}_k + \mathbf{F}_m \cdot \mathbf{i}_k\} . \end{aligned} \quad (13)$$

Thus the relations for components of reactions $\mathbf{R}_{mk}$ and $\mathbf{R}_{np}$ in bases $\{\mathbf{i}_k; \mathbf{j}_k; \mathbf{k}_k\}$ and $\{\mathbf{i}_n; \mathbf{j}_n; \mathbf{k}_n\}$ are founded. Vectors $\mathbf{R}_{mn}$ and $\mathbf{M}_{mk}^R$ are calculated using formulas (8) and (11) respectively.

5. The kinetostatic analysis of RTR dyad

Given: D_m and D_n are rotational pairs, E is a translational pair, $I_n \geq 0$, $I_m > 0$, $\{\mathbf{i}_\gamma; \mathbf{j}_\gamma; \mathbf{k}_\gamma\}$, $\mathbf{F}_\gamma$, $\mathbf{M}_\gamma$ ($\gamma \in \{m, n\}$). If $I_n = 0$, then $\mathbf{i}_n$ equals to $\mathbf{i}_m$. Find: $\mathbf{R}_{mk}$, $\mathbf{R}_{np}$, $\mathbf{R}_{mn}$ and $\mathbf{M}_{mn}^R$.

The force diagram is presented in Figure 1(c). Assume that points D_m and D_n do not coincide and introduce new basis $\{\mathbf{i}_\beta; \mathbf{j}_\beta = \mathbf{k}_\beta \times \mathbf{i}_\beta; \mathbf{k}_\beta = \mathbf{k}_0\}$ with unit vector $\mathbf{i}_\beta = |D_m D_n|^{-1} (I_m \mathbf{j}_m - I_n \mathbf{j}_n)$ lying on the guideline $D_m D_n$. Here $|D_m D_n| = \sqrt{I_m^2 + I_n^2 - 2I_m I_n (\mathbf{i}_m \cdot \mathbf{i}_n)}$.

The moment equation for the whole dyad with respect to point D_m is

$$\begin{aligned} \sum \mathbf{M}_{Dm(m,n)} &= |D_m D_n| \mathbf{i}_\beta \times \mathbf{R}_{np} + (|D_m D_n| \mathbf{i}_\beta + x_{Sn} \mathbf{i}_n + y_{Sn} \mathbf{j}_n) \times \mathbf{F}_n + \mathbf{M}_n + \\ &+ (x_{Sm} \mathbf{i}_m + y_{Sm} \mathbf{j}_m) \times \mathbf{F}_m + \mathbf{M}_m = 0 . \end{aligned} \quad (14)$$

After elementary transformations, for component $\mathbf{R}_{np} \cdot \mathbf{j}_\beta$ it is obtained

$$\mathbf{R}_{np} \cdot \mathbf{j}_\beta = -(\mathbf{F}_n \cdot \mathbf{j}_\beta) - |D_m D_n|^{-1} \sum_{\gamma=m,n} \{x_{S\gamma} (\mathbf{F}_\gamma \cdot \mathbf{j}_\beta) - y_{S\gamma} (\mathbf{F}_\gamma \cdot \mathbf{i}_\gamma) + (\mathbf{M}_\gamma \cdot \mathbf{k}_\beta)\} . \quad (15)$$

Equilibrium equation for link n can be presented in the form

$$(\mathbf{R}_{np} \cdot \mathbf{i}_\beta) \mathbf{i}_\beta + \{(\mathbf{R}_{np} \cdot \mathbf{j}_\beta) \mathbf{j}_\beta + \mathbf{F}_n\} - (\mathbf{R}_{mn} \cdot \mathbf{j}_m) \mathbf{j}_m = 0 , \quad (16)$$

which for $\mathbf{i}_\beta \cdot \mathbf{i}_m \neq 0$ implies

$$\begin{aligned} \mathbf{R}_{np} \cdot \mathbf{i}_\beta &= -(\mathbf{i}_\beta \cdot \mathbf{i}_m)^{-1} \{(\mathbf{R}_{np} \cdot \mathbf{j}_\beta)(\mathbf{j}_\beta \cdot \mathbf{i}_m) + \mathbf{F}_n \cdot \mathbf{i}_m\} , \\ \mathbf{R}_{mn} \cdot \mathbf{j}_m &= (\mathbf{j}_m \cdot \mathbf{j}_\beta)^{-1} \{(\mathbf{R}_{np} \cdot \mathbf{j}_\beta) + \mathbf{F}_n \cdot \mathbf{j}_\beta\} . \end{aligned} \quad (17)$$

Relations for components of vectors $\mathbf{R}_{mn}$ and $\mathbf{R}_{np}$ in basis $\{\mathbf{i}_m; \mathbf{j}_m; \mathbf{k}_m\}$ and $\{\mathbf{i}_\beta; \mathbf{j}_\beta; \mathbf{k}_\beta\}$ are therefore found. Applying equilibrium and moment equation with respect to D_m for link m , force $\mathbf{R}_{mk}$ and moment $\mathbf{M}_{mn}^R$ are determined as

$$\mathbf{R}_{mk} = -(\mathbf{R}_{mn} + \mathbf{F}_m) , \quad (18)$$

$$\mathbf{M}_{mn}^R = -I_m (\mathbf{R}_{mn} \cdot \mathbf{j}_m) \mathbf{k}_m - (x_{Sm} \mathbf{i}_m + y_{Sm} \mathbf{j}_m) \times \mathbf{F}_m - \mathbf{M}_m.$$

6. The kinetostatic analysis of RTT dyad

Given: D_m and E are translational pairs, D_n is rotational pair, $I_n \geq 0$, $I_m > 0$, $\{\mathbf{i}_\gamma; \mathbf{j}_\gamma; \mathbf{k}_\gamma\}$, $\mathbf{F}_\gamma$, $\mathbf{M}_\gamma$ ($\gamma \in \{m, n\}$). If $I_n = 0$, then $\mathbf{i}_n$ equals to $\mathbf{i}_m$. Find: $\mathbf{R}_{mk}$, $\mathbf{M}_{mk}^R$, $\mathbf{R}_{np}$, $\mathbf{R}_{mn}$ and $\mathbf{M}_{mn}^R$.

The force diagram is presented in Figure 1(e). Equilibrium equation for link m can be written as

$$(\mathbf{R}_{mk} \cdot \mathbf{j}_k) \mathbf{j}_k + \mathbf{F}_m + (\mathbf{R}_{mn} \cdot \mathbf{j}_m) \mathbf{j}_m = 0, \quad (19)$$

and, after scalar multiplication by either vector $\mathbf{i}_k$ or $\mathbf{i}_m$, for $\mathbf{j}_k \cdot \mathbf{i}_m \neq 0$ leads to

$$(\mathbf{R}_{mk} \cdot \mathbf{j}_k) = -(\mathbf{j}_k \cdot \mathbf{i}_m)^{-1} (\mathbf{F}_m \cdot \mathbf{i}_m), \quad (\mathbf{R}_{mn} \cdot \mathbf{j}_m) = -(\mathbf{j}_m \cdot \mathbf{i}_k)^{-1} (\mathbf{F}_m \cdot \mathbf{i}_k). \quad (20)$$

Forces $\mathbf{R}_{mn}$ and $\mathbf{R}_{mk}$ can therefore be calculated. From equilibrium equation (5) for the whole structural group we find

$$\mathbf{R}_{np} = -(\mathbf{R}_{mk} + \mathbf{F}_m + \mathbf{F}_n). \quad (21)$$

Using the moment equations with respect to point E for link n and then for link m , for moments $\mathbf{M}_{mn}^R$ and $\mathbf{M}_{mk}^R$ we obtain

$$\begin{aligned} \mathbf{M}_{mn}^R &= -\mathbf{M}_{nm}^R = -I_n \mathbf{j}_n \times \mathbf{R}_{np} + ((x_{Sn} - I_n) \mathbf{i}_n + y_{Sn} \mathbf{j}_n) \times \mathbf{F}_n + \mathbf{M}_n, \\ \mathbf{M}_{mk}^R &= I_m \mathbf{i}_m \times \mathbf{R}_{mk} - ((x_{Sm} - I_m) \mathbf{i}_m + y_{Sm} \mathbf{j}_m) \times \mathbf{F}_m - \mathbf{M}_m - \mathbf{M}_{mn}^R. \end{aligned} \quad (22)$$

Discussion and conclusions

A unified approach to the kinetostatic analysis of Assur dyads is presented, which allows to conduct the kinetostatic analysis of any class II plane mechanism according to Artobolevsky's classification. The method is based on the solution of the vector equations of kinetostatics by vector algebra tools. Representations of active forces, inertial forces, and reactions in kinematic pairs are introduced in the form of decompositions in the local bases. These are rigidly attached to the dyad's links determined within the position analysis. The suggested method is appropriate to apply in combination with positional and kinematic analysis of Assur dyads, which are also based on vector algebra methods (Khoroshev, Duchenko, & Kykot, 2023; 2024). The obtained calculation schemes are convenient for algorithmization and can be an integral part of computer-aided design systems for class II plane mechanisms according to Artobolevsky's classification.

Authors' contribution: Kostiantyn Khoroshev – conceptualization, formal analysis, methodology, writing – original draft; Kyryl Duchenko – formal analysis, writing – original draft; Serhii Kykot – formal analysis.

References

- Abhery, K. (2022). An analytical parametric method for force analysis of planar mechanisms with sliding frictions. *International Journal of Mechanical Engineering Education*, 50(3), 667–691. <https://doi.org/10.1177/03064190211035186>
- Artobolevski, I. (1977). *Théorie Des Mécanismes Et Des Machines*. Éditions Mir.
- Brát, V., & Lederer, P. (1973). KIDYAN: Computer-aided kinematic and dynamic analysis of planar mechanisms. *Mechanism and Machine Theory*, 8(4), 457–467. [https://doi.org/10.1016/0094-114X\(73\)90020-7](https://doi.org/10.1016/0094-114X(73)90020-7)
- Chace, M. A. (1963). Vector analysis of linkages. *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, 85(3), 289–297. <https://doi.org/10.1115/1.3669867>
- Durango, S., Calle, G., & Ruiz, O. (2010). Analytical method for the kinetostatic analysis of the second-class RRR assur group allowing for friction in the kinematic pairs. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 32(3), 200–207. <https://doi.org/10.1590/S1678-58782010000300002>
- Hansen, M. R. (1996). A general method for analysis of planar mechanisms using a modular approach. *Mechanism and Machine Theory*, 31(8), 1155–1166. [https://doi.org/10.1016/0094-114X\(96\)84606-4](https://doi.org/10.1016/0094-114X(96)84606-4)
- Honchar, M. O. (2011). *Theory of mechanisms and machines*. Vydavnychiy dim Vinnychenko [in Ukrainian]. [Гончар, М. О. (2011). Теорія механізмів і машин. Видавничий дім Винниченка].
- Jomartov, A., & Tuleshov, A. (2018). Vector method for kinetostatic analysis of planar linkages. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 40(2), 56. <https://doi.org/10.1007/s40430-018-1022-y>
- Khoroshev, K., Duchenko, K., & Kykot, S. (2024). The direct vector approach to the position analysis of Assur dyads. *Mechanism and Machine Theory*, 195, 105608. <https://doi.org/10.1016/j.mechmachtheory.2024.105608>
- Khoroshev, K. G., Duchenko, K. O., & Kykot, S. V. (2023). The vector algebra approach to the kinematic analysis of the structural groups of the 2nd class by Artobolevsky. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 2, 160–163. <https://doi.org/10.17721/1812-5409.2023/2.28>
- Kinytskyi, Ya. T. (2002). *Theory of mechanisms and machines*. Naukova dumka [in Ukrainian]. [Кіницький, Я. Т. (2002). Теорія механізмів і машин. Наукова думка].
- Nikraves, P. E. (1988). *Computer-Aided Analysis of Mechanical Systems*. Prentice Hall.
- Otter, M., Elmquist, H., & Mattsson, S. E. (2003, November 3–4). The New Modelica MultiBody Library. In P. Fritzson (Ed.), *Proceedings of the 3rd International Modelica Conference* (pp. 311–330). Linköpings universitet. https://modelica.org/events/Conference2003/papers/h37_Otter_multibody.pdf
- Roberson, R. E., & Schwertassek, R. (1988). *Dynamics of Multibody Systems*. Springer Verlag.
- Simionescu, P. A. (2016, August 21–24). MEKIN2D: Suite for planar mechanism kinematics. *Proceedings of the ASME Design Engineering Technical Conference*, 5B-2016. Charlotte. <https://doi.org/10.1115/DETC2016-59086>
- Uicker, J. J., Pennock, G. R., & Shigley, J. E. (2016). *Theory of Machines and Mechanisms* (5th ed.). Oxford University Press/Covey.

Отримано редакцією журналу / Received: 30.10.24
Прорецензовано / Revised: 17.01.25
Схвалено до друку / Accepted: 28.01.25

Костянтин ХОРОШЕВ, канд. фіз.-мат. наук, доц.

ORCID ID: 0000-0002-3900-7838

e-mail: khoroshev.kostiantyn@iit.kpi.ua

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

Кирил ДУЧЕНКО, студ.

ORCID ID 0009-0002-7987-6891

e-mail: ge85fix@mytum.de

Мюнхенський технічний університет, Мюнхен, Німеччина

Сергій КИКОТЬ, канд. фіз.-мат. наук, доц.

ORCID ID: 0000-0003-1035-8711

e-mail: kykot.serhii@iit.kpi.ua

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

КІНЕТОСТАТИЧНИЙ АНАЛІЗ ДІАД АССУРА

Ядром комп'ютерних систем автоматизованого проєктування та розрахунку плоских механізмів є математичні методи модульного характеру дослідження динаміки механізмів. Найпоширеніший підхід – динаміка багатокомпонентних механічних систем, у якому механізм розбивають на ланки, що являють собою абсолютно тверді тіла. Кількість ланок дорівнює кількості модулів аналізу. Інший підхід ураховує структуру механізму, що може сприяти зменшенню кількості модулів, які потребують аналізу. Однак для такого підходу необхідний розвиток системи класифікації механізмів і повного набору математичних методів кінематичного, кінетостатичного та динамічного аналізу всіх можливих структурних елементів. Найрозвиненішою в цьому сенсі є класифікація за Артоблевським. Сукупність плоских механізмів, що формують найпростіший II клас за Артоблевським, повністю забезпечена алгоритмізованими підходами модульного кінематичного аналізу. Ці методи аналітичні та підлягають алгоритмізації. Кінетостатичний аналіз для цього класу досі представлений тільки графо-аналітичним підходом. У пропонованій роботі подано уніфікований аналітичний метод кінетостатичного аналізу діад Ассура, що разом з механізмами I класу формують множину структурних елементів механізмів II класу за Артоблевським. В основу методу покладено векторні рівняння кінетостатики, що розв'язуються за допомогою математичного апарату векторної алгебри. Уведено представлення активних сил, сил інерції та реакцій у кінематичних парах у вигляді векторів, що розкладені за відомими з аналізу положень локальними базисами, жорстко зв'язаними з ланками структурних груп. Запропонований підхід рекомендуємо використовувати в поєднанні з кінематичним аналізом, що також базується на інструментарії векторної алгебри, а також він може бути алгоритмізований й утворювати невід'ємну частину систем автоматизованого проєктування плоских механізмів II класу за Артоблевським.

Ключові слова: модульний підхід, кінетостатика, плоскі механізми, діади Ассура.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

UDC 519.7

DOI: <https://doi.org/10.17721/1812-5409.2025/1.12>

Igor AIZENBERG, PhD (Comp. Science), Prof.

ORCID ID: 0000-0002-5994-6568

e-mail: igor.aizenberg@manhattan.edu

Manhattan University, Riverdale, New York, USA

Alexander VASKO, PhD Student

ORCID ID: 0009-0006-1527-505X

e-mail: oleksandrvasco@uzhnu.edu.ua

Uzhhorod National University, Uzhhorod, Ukraine

COMPARATIVE ANALYSIS OF CNNMVN AND MLMVN
AS FREQUENCY DOMAIN CNN CONVOLUTIONS

Each convolutional layer in any convolutional neural network produces a feature map containing the most important information, which a network needs to recognize respective images. To further improve these neural networks and better understand their capabilities, it is essential to discover, which features are actually extracted and how the images to be recognized are transformed by convolutions resulted from the learning process. This paper presents a comparative analysis of convolutions obtained via two complex-valued neural networks based on multi-valued neurons. The first network is a convolutional neural network based on multi-valued neurons (CNNMVN) which has a traditional convolutional neural network topology except of that it employs complex-valued convolutional kernels in its convolutional part and multi-valued neurons in its fully connected part. The second one is the multi-valued neural network based on multi-valued neurons (MLMVN) which is a fully connected multilayer neural network employed as a convolutional network in the frequency domain.

Considering that both neural networks are complex-valued and the obtained filters operate in the complex domain, the conducted research indicates that the kernels of both networks produce filters similar to existing digital image processing filters. The analysis of CNNMVN kernels revealed that they implement unsharp masking filters and edge detection filters for identifying shapes in images, while the MLMVN kernels enhance specific frequency sub-bands. The latter means that the respective filters are mostly not similar to the ones known as unsharp masking or sharpening filters. Thus, the kernels of both convolutional networks contribute to improving image recognition performance in their own ways.

Key words: convolutional neural networks; CNNMVN; complex-valued neural networks; multi-valued neuron; MLMVN; image recognition; frequency domain.

AMS 2020 classification: 68T07.

Introduction

A convolutional neural network (CNN) has a special place in the field of neural networks and deep learning since its introduction in (Krizhevsky, Sutskever, & Hinton, 2012). Originally inspired by the human visual system (Jarrett et al., 2009; LeCun et al., 2004), CNNs become a primary tool for visual information processing. In general, all real-world images may contain many details that may not be always useful for classification or recognition. The main distinction of CNN from the other neural networks is that CNN has special convolutional layers that can first extract useful information from visual data and then perform classification and recognition based only on it. This approach showed excellent results compared to regular feedforward neural networks (Beysolow li, 2017; Caldeira et al., 2019; Gad, 2018; Kaur, & Garg, 2020; Lin et al., 2019; Lin et al., 2020; Sarabu, & Santra, 2021; Venkatesan, & Li, 2017; Wu, Zhang, & Zhao, 2020; Xiao et al., 2018; Yar et al., 2021).

Among all known neural networks, we would like to focus our attention in this work on complex-valued networks. They use complex numbers as weights, employ complex-valued activation functions, and their inputs and outputs are also complex-valued, accordingly. Various neural networks based on complex-valued neurons with a complex hyperbolic tangent activation function were considered during the last decades (Hirose, 2011; Kobayashi, 2017; Valle, 2014). These networks demonstrated a number of advantages over their real-valued counterparts (Boonsatit et al., 2022; Nitta, 2004; Nitta, 2008; Suresh, Sundararajan, & Savitha, 2013; Zhang et al., 2022). The first complex-valued convolutional neural network (CV-CNN) was introduced in (Bruna et al., 2015) and the first quaternion-based convolutional neural network (QV-CNN) was introduced in (Zhu et al., 2018). Over the past years, significant progress has been made in the development and application of CV-CNNs and QV-CNN, leading to their successful implementation across various domains. Subsequent studies (Chatterjee et al., 2023; Fuchs et al., 2021; Guberman, 2016; Hongo et al., 2020; Popa, 2017; Rawat, Rana, & Kumar, 2021; Sunaga, Natsuaki, & Hirose, 2020; Yadav, & Jeripothula, 2023; Zhang et al., 2017) introduce both CV-CNNs and QV-CNNs as powerful tools for complex image processing tasks. Notably, these models have demonstrated their effectiveness in synthetic-aperture radar (SAR) image analysis, where they have been employed for both image recognition and noise filtering. In addition to SAR applications, CV-CNNs and QV-CNNs have proven to be valuable in broader image recognition tasks, offering advantages in handling complex-valued data and capturing richer spatial and spectral features compared to traditional real-valued CNNs.

Another class of complex-valued neural networks is based on multi-valued neurons (MVN). MVN and MVN-based networks are comprehensively described in (Aizenberg, 2011). The multilayer feedforward neural network with multi-valued neurons (MLMVN) was first introduced in (Aizenberg, & Moraga, 2007). The following features should be mentioned among advantages: derivative-free learning algorithm not suffering from the local minima problem; it often learns faster and develop better generalization capability compared to regular MLP (Aizenberg, 2011; Aizenberg, & Moraga, 2007). A convolutional neural network based on MVN (CNNMVN) was introduced in (Aizenberg, & Vasko, 2020), and its modification was presented in (Aizenberg, Herman, & Vasko, 2022). CNNMVN is a network with a standard convolutional neural network topology. It consists of the convolutional layers that can be followed by pooling layers, other convolutional layers or the dense layers.

© Aizenberg Igor, Vasko Alexander, 2025

It was shown in (Aizenberg, & Vasko, 2024) that CNNMVN is the further generalization of MLMVN. The use of MLMVN as a frequency domain CNN (MLMVN-FCNN) was suggested in (Aizenberg, & Vasko, 2023). According to the Convolution Theorem, the convolution operation can be calculated in both spatial and frequency domains. Thus, it was suggested to interpret and use the first hidden layer of MLMVN as a convolutional layer in the frequency domain.

In this paper, we present a deep analysis of feature maps for both CNNMVN and MLMVN-FCNN. This should help to better understand what kind of information is extracted and which convolutional kernels are constructed during the learning process.

1. CNNMVN

CNNMVN combines the principles of convolutional neural networks and multi-valued neurons (Aizenberg, Herman, & Vasko, 2022; Aizenberg, & Vasko, 2020; 2024). It has a topology traditional for CNNs, employs complex-valued weights and continuous MVN activation function for its convolutional kernels, and has an MVN-based fully connected part.

As any convolutional neural network, CNNMVN consists of convolutional, pooling and the dense layers. For proper processing, intensities of a real-valued input image should be converted to complex-valued ones according to the following rule

$$z_j = e^{i\varphi_j}, \quad \varphi_j = \frac{x_j - \min\{X\}}{\max\{X\} - \min\{X\}} \cdot \lambda, \quad (1)$$

where x_j and z_j stand for a real-valued datapoint and a datapoint projected onto the unit circle, respectively; φ_j is the argument (phase) of the projected datapoint z_j ; $\min\{X\}$ and $\max\{X\}$ are respectively the minimum and maximum values over a real-valued dataset; parameter λ defines a pre-determined part of the unit circle to which real-valued data are projected (a standard value is $\lambda = 3\pi/2$).

Each convolutional layer consists of kernels whose purpose is to extract some useful features from an input image. All kernels slide over an input image (or a feature map) performing a convolution operation. It computes a weighted sum of the input values within its receptive field. This process helps to extract specific features such as edges, shapes and textures from images to be analyzed. Convolutional kernels also capture spatial relationships between pixels. It is widely known that adjacent pixels in an image are often highly correlated, and kernels can identify local patterns because of this property. Also, in a case when multiple convolutional layers are used, kernels can identify larger features by combining smaller local features that were extracted in preceding layer(s).

A layer following a convolutional one can be either another convolutional layer or a pooling layer. If the latter is the case, its main purpose is downsampling and reduction of the information to be processed. The convolutional and the pooling layers are followed by fully connected layers. This part of CNNMVN is nothing else but MLMVN.

A CNNMVN learning algorithm, that is the process of error backpropagation and error correction in CNNMVN was presented in (Aizenberg & Vasko, 2020), further develop in (Aizenberg, Herman, & Vasko, 2022) and then presented in detail in (Aizenberg, & Vasko, 2024).

2. MLMVN as a Frequency domain CNN (MLMVN-FCNN)

The idea to use MLMVN as a frequency-domain convolutional neural network was introduced in (Aizenberg, & Vasko, 2023) and then considered in detail in (Aizenberg, & Vasko, 2024). It was inspired by the fact that convolution operation can be performed in either of the spatial or frequency domains. The convolution of two signals $f(t)$ and $g(t)$ can be obtained as

$$(f * g)(t) = \mathcal{F}^{-1}(F \cdot G), \quad (2)$$

where $\mathcal{F}^{-1}$ stands for the inverse Fourier transform, F and G are the Fourier transforms of signals $f(t)$ and $g(t)$. According to (2) the Convolutions Theorem states that spatial domain convolution of two signals is equivalent to the frequency-domain product of the respective Fourier transforms of these signals.

Let us now consider the first hidden layer in MLMVN. Let vector $(x_1, \dots, x_n)$ and vector $(w_0^i, w_1^i, \dots, w_n^i)$ be the i^{th} neuron's inputs and weights, respectively. Then a weighted sum z_i is calculated as

$$z_i = w_0^i + w_1^i x_1 + w_2^i x_2 + \dots + w_n^i x_n \quad (3)$$

According to (3), if the input vector $(x_1, \dots, x_n)$ represents a vectorized 2D Fourier transform of an image, we can assume that each neuron in the first layer of MLMVN performs a convolution in the frequency domain. Consequently, the first layer can be interpreted as a convolutional layer operating in the frequency domain and the neurons' weights interpreted as convolutional kernels in the frequency domain.

This approach has several key distinctions when compared to traditional CNNs. The first one is that in the first hidden layer, the result of each neuron's convolution is not a feature map as it is in regular CNNs. Instead, it is a sum of the frequency-domain feature map coefficients. And as a result, there is no possibility to apply another convolutional layer. The second distinction is that the kernel size in the spatial domain cannot be manually set, and therefore, we cannot directly control the size of the features being extracted.

To improve performance and network's generalization capability, some modifications including frequency domain pooling were suggested. The main purpose of the frequency domain pooling is to reduce the number of calculations and remove frequencies, which are not important for recognition of objects of interest. Frequency domain pooling is reduced to extraction of Fourier coefficients corresponding only to the 2D frequencies, which are essential to recognition of certain objects – from the 1st frequency to the k -th frequency, which is a cut-off frequency in this case.

3. Comparative analysis of convolutions

As it was said before, the main purpose of any convolutional layer is to extract important features from an image to make it easier to classify/recognize it. Considering that both CNNMVN and MLMVN-FCNN are complex valued, it is interesting to visualize and compare feature maps that are generated in convolutional layers. Analyzing these feature maps, we may better understand and hopefully improve the learning process of both networks.

We used in our experiments two popular benchmarks: MNIST – the dataset of handwritten digits (LeCun et al., n. d.) and Fashion MNIST – the dataset of fashion images (Xiao, Rasul, & Vollgraf, 2017). Both datasets contain 60 000 images for training and 10 000 images for testing. All images are gray-scale and therefore their intensity range is $[0, 255]$. The size of each image is 28×28 pixels.

To visualize and analyze the feature maps, we used trained networks, which were presented in (Aizenberg, & Vasko, 2024). For CNNMVN we analyzed models with same topology for both MNIST and Fashion MNIST datasets. It was 32C5-h256-o10 topology, where 32C5 stands for the convolutional layer with 32 kernels of size 5×5 , h-256 is the number of neurons in the hidden layer and

σ_{10} is the number of output neurons. The accuracy was 97,71% and 88,06% for MNIST and Fashion MNIST datasets, respectively. To analyze and visualize feature maps of MLMVN-FCNN we used a model with the h2048- σ_{10} topology. For this model, we got 98.11% accuracy on the MNIST dataset with 5 extracted 2D frequencies. To analyze Fashion MNIST feature maps, we used a model with the h3072- σ_{10} topology. For this model, we got 90% accuracy with 7 extracted 2D frequencies.

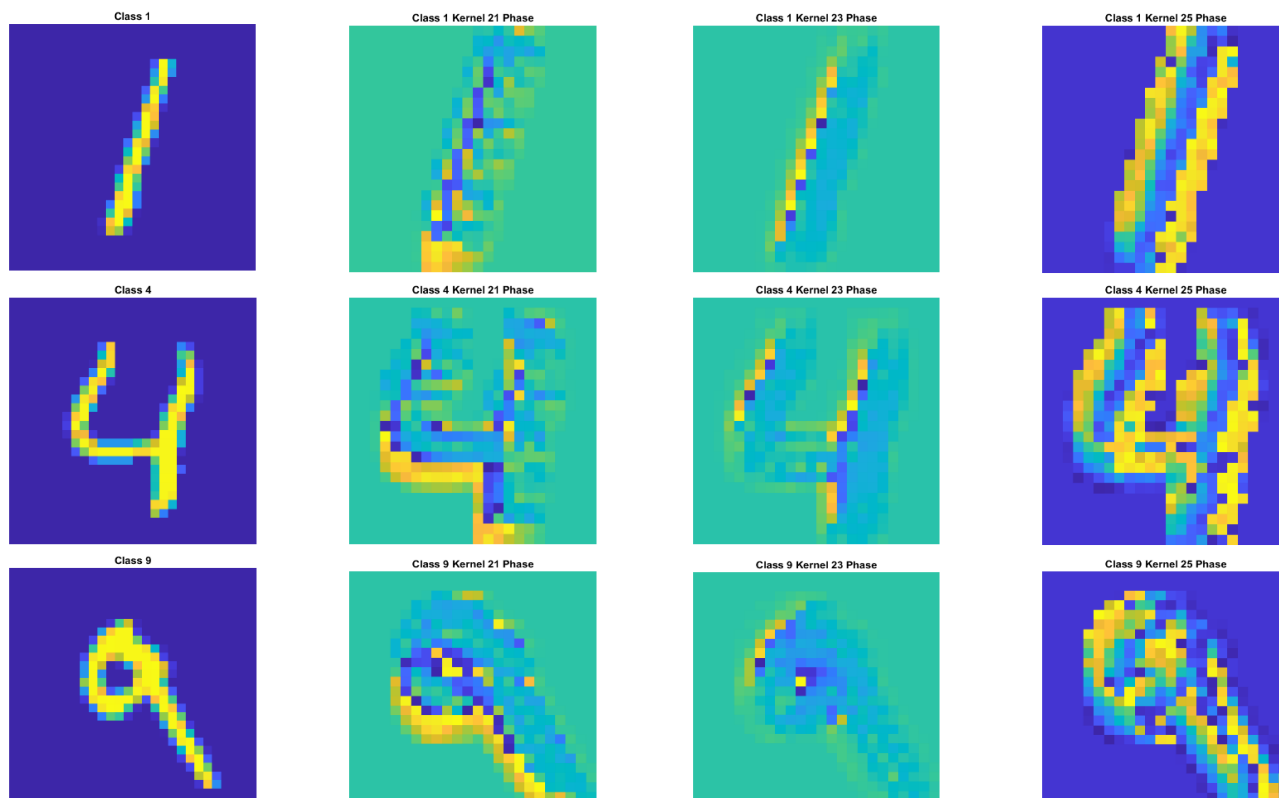


Fig. 1 (a). Original images of 1, 4 and 9 classes

Fig. 1 (b). Phases of the feature maps for 21, 23 and 25 kernels of CNNMVN in MNIST dataset

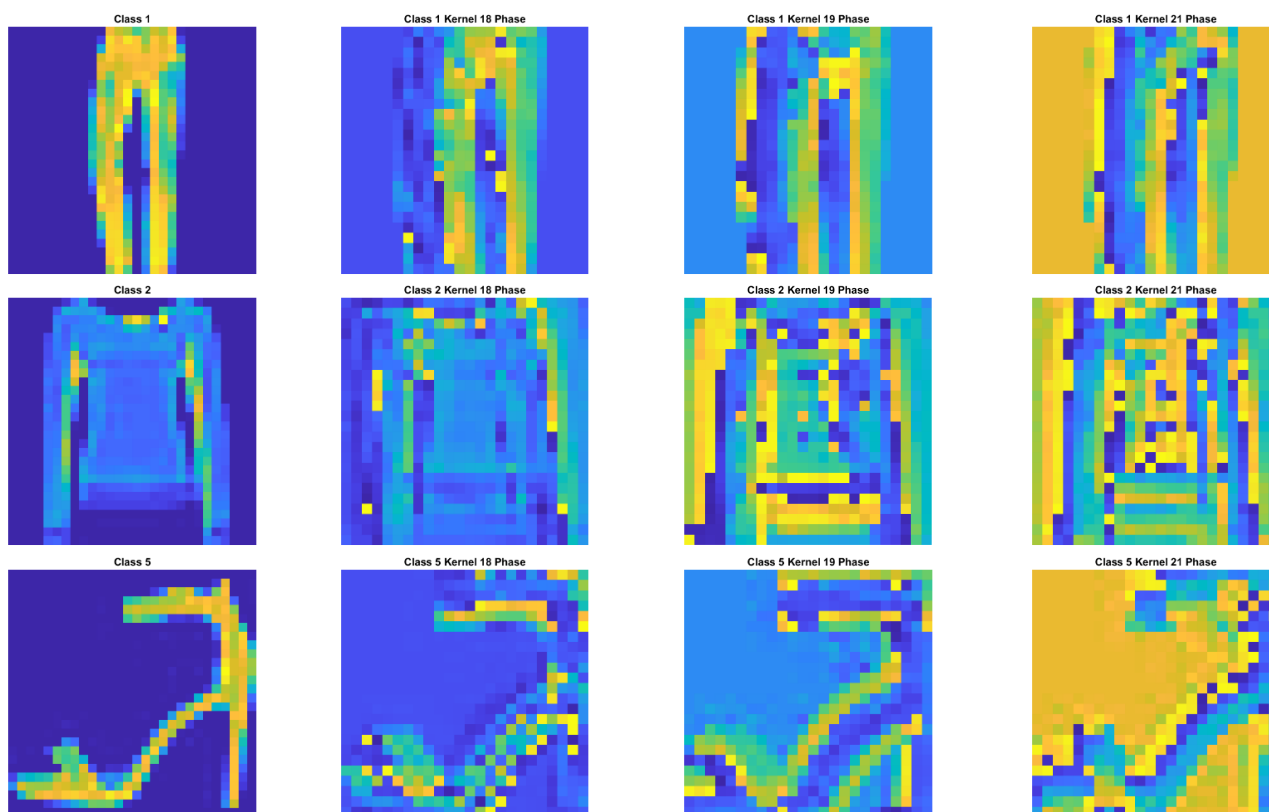


Fig. 1 (c). Original images of 1, 2 and 5 classes

Fig. 1 (d). Phases of the feature maps for 18, 19 and 21 kernels of CNNMVN in Fashion MNIST dataset

The visualization of the feature maps for MNIST and Fashion MNIST datasets of CNNMVN are presented in Fig. 1. In Fig. 1 (a) we observe the phases of the original images of MNIST dataset for the classes 1,4 and 9, transformed according to (1). Figure 1 (b) presents the phases of the feature maps generated by the kernels 21, 23 and 25 of CNNMVN after the learning process was completed. One can observe that kernel 21 is performing horizontal and diagonal lines detection while kernel 23 better detects vertical lines. The kernel 25 performs high frequency emphasizing, that is, it performs filtering, which can be compared to unsharp masking. Similar results are observed for the Fashion MNIST dataset. Fig. 1 (c) illustrates the phases of the original images for the classes 1, 2 and 6, transformed by (1). Figure 2 (d) depicts phases of the respective feature maps for kernels 18, 19, and 21 after the CNNMVN training. We can observe that kernels 18 and 19 are performing high-pass filtering, which leads to shape highlighting of the image. Kernel 21 performs some kind of high frequency emphasizing which can be compared to unsharp masking. Summarizing the above, we can conclude that CNNMVN kernels produce feature maps similar to those of traditional CNNs. Like their real-valued counterparts, CNNMVN kernels aims to detect shapes and edges in an input image through the development of various types of filters.

While CNNMVN performs convolution in the spatial domain, MLMVN-FCNN performs it in the frequency domain. As mentioned earlier, this approach has several limitations. We cannot manually set the kernel size in the spatial domain and therefore we cannot have control over the size of details extracted by a respective convolution. Additionally, as we are dealing in this case with the frequency domain kernels, the learning process can be resulted in kernels that often are difficult to compare to their spatial domain counterparts.

The visualization of the MLMVN feature maps for both datasets is presented in Fig. 2. In Fig. 2 (a) samples from the MNIST dataset after frequency domain pooling are presented with focusing on classes 1, 4, and 9, and with the frequencies (1–5) extracted. Fig. 2 (c) depicts the Fashion MNIST classes after the frequency domain pooling, specifically classes 1, 2, and 6, with the frequencies (1–7) were extracted. Figures 2 (b) and 2 (d) visualize the feature maps resulted from the learning process of MLMVN. Since the convolution process in MLMVN is performed in the frequency domain, we applied the inverse Fourier transform to the feature maps, and then their magnitudes were extracted to visualize them. It should also be mentioned that according to the frequency domain pooling algorithm, the 2D frequencies are extracted in a "diamond" shape (Fig. 3), have been flattened and then processed by the first hidden layer neurons in MLMVN. Thus, a network learns employing only Fourier coefficients extracted from the kernel according to the "diamond" rule. It should also be taken into consideration that kernels and feature maps have ripple effects because of this when visualized. It is significant when analyzing kernels and feature maps in the frequency domain.

It is clearly seen in Fig. 2 (b) that all filters resulted from the learning process in the frequency domain distinguish certain features in the MNIST images. These filters emphasize medium and high frequencies and can be compared to unsharp masking filters. At the same time, while unsharp masking filter usually emphasize all high frequencies and possibly some medium frequencies, MLMVN neurons as a result of learning emphasize very specific narrow frequency sub-bands. Thus, they develop filters distinguishing specific details. For example, as it is seen from Fig. 2 (b), neuron 351 distinguishes details of certain specific sizes by emphasizing respective medium frequencies. It can also be seen from Fig. 2 (b) that neuron 461 distinguishes vertically oriented details in images by emphasizing vertical frequencies, while neuron 466 distinguishes diagonally oriented details (from bottom-left to top-right) by emphasizing respective frequencies.

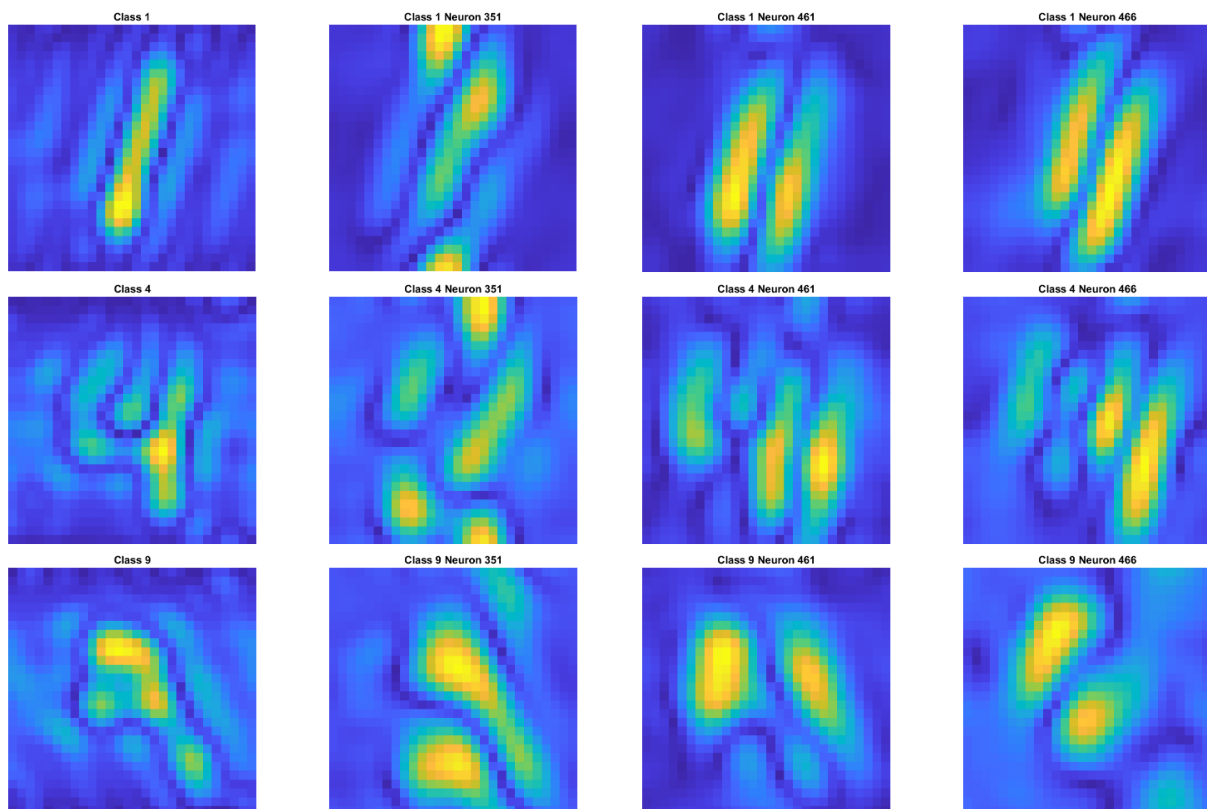


Fig. 2 (a). Classes 1,4, and 9 after frequency domain pooling

Fig. 2 (b). Magnitudes of the inverse Fourier transform for neurons 351, 461, and 466 neurons of MLMVN-FCNN in MNIST dataset

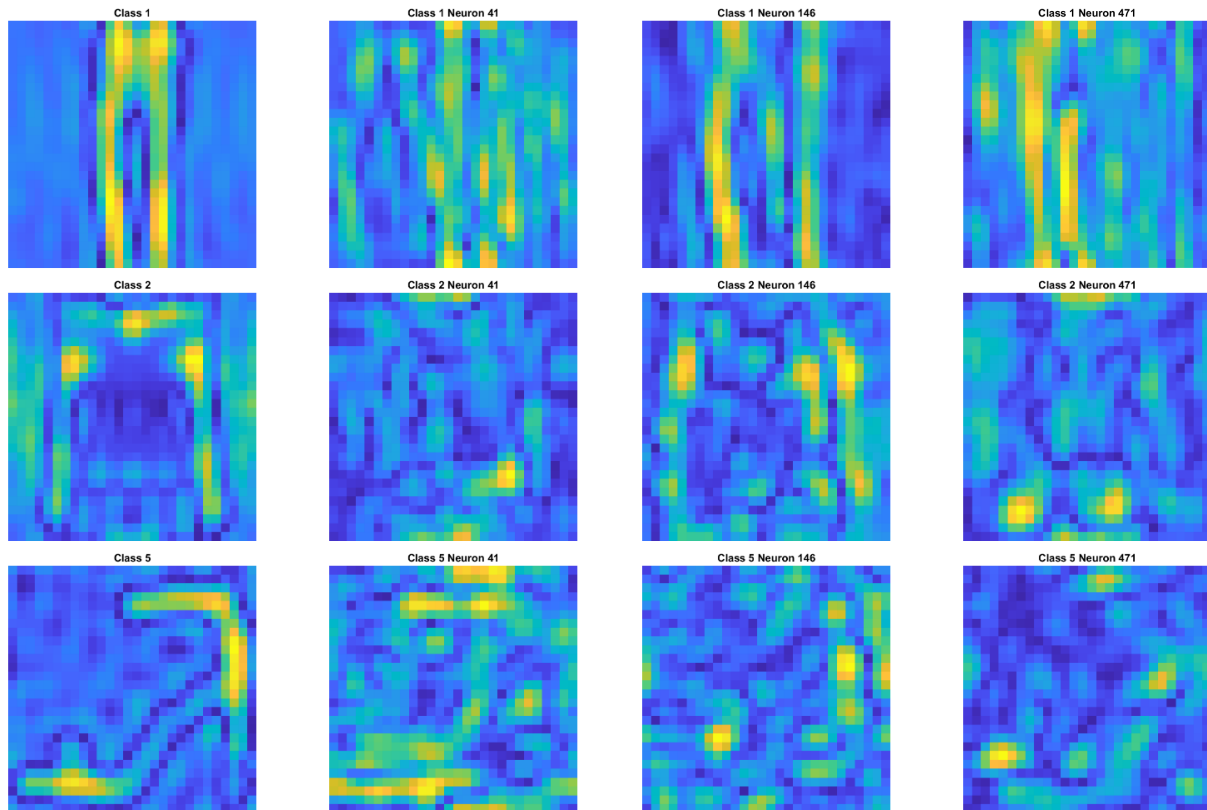


Fig. 2 (c). Classes 1,2 and 6 after frequency domain pooling

Fig. 2 (d). Magnitudes of the inverse Fourier transform for neurons 41, 146, and 471 of MLMVN-FCNN in Fashion MNIST dataset

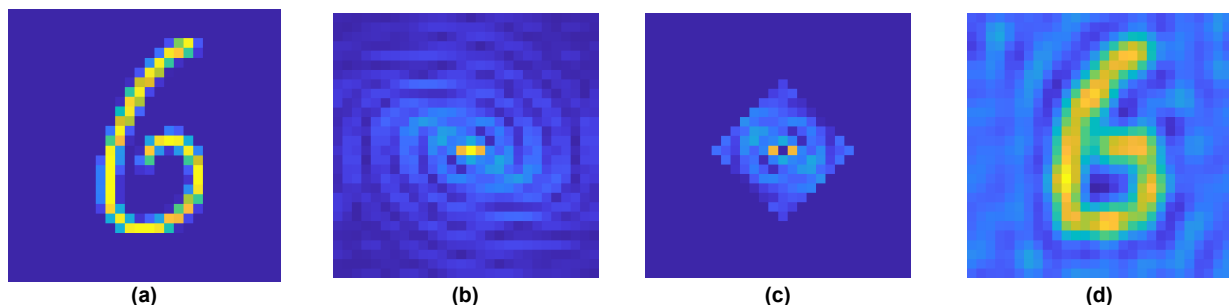


Fig. 3. Frequency pooling process: a) original image; b) Fourier frequencies of the image with circular shift; c) extracted Fourier coefficients corresponding to the frequencies 1-5 around the DC frequency; d) inverse Fourier transform of the extracted Fourier coefficients corresponding to the frequencies 1-5

A similar result can be observed in Fig. 2 (d) for the Fashion MNIST images, but with some distinctions. Since these images are more complicated and lost some details due to the frequency domain pooling (the frequencies (1–7) were extracted), they retain the general shape. For example, neuron 41 detects horizontally oriented details in the images from classes 2 and 6. At the same time, while there are no horizontally oriented details in the image belonging to class 1, it highlights some isolated elements belonging to vertically oriented details in this image. Neuron 146 highlights vertically oriented details for all classes and the kernel 471 highlights various shapes for different classes by emphasizing medium frequencies.

It is worth looking at the illustrations showing exactly which frequencies are amplified by the filters resulting from training both neural networks. To discover this, we need to look at Fourier power spectra of the respective filter kernels. To obtain respective power spectra of CNNMVN kernels, we took Fourier transform of the spatial domain kernels resulting from the learning process. Figure 4 (a) represents respective power spectra of kernels 21, 23, 25 for the MNIST dataset, and Fig. 4 (b) represents respective powers spectra of kernels 18, 19, 21 for the Fashion MNIST dataset. Thus, power spectra in Fig. 4 (a) represent the filters used to create feature maps shown in Fig. 1 (b), while power spectra in Fig. 4 (b) represent the filters used to create feature maps shown in Fig. 1 (d).

Power spectra of MLMVN-FCNN kernels can easily be obtained by taking absolute values of the respective weights and reshaping them into 2D structures. Figure 5 (a) represents respective power spectra of kernels 351, 451, 466 for the MNIST dataset. We may observe from Fig. 5 (a) that medium frequencies are emphasized by neuron 351, and this leads to the extraction of details of certain specific sizes as it can be seen in Fig. 2 (b). This example also shows that vertical frequencies are emphasized by neuron 461, and this is resulted in distinguishing vertical details as can be seen in Fig. 2 (b). As we see from Fig. 5 (a), neuron 466 emphasizes frequencies corresponding to the diagonally oriented image elements, and this can be seen in Fig. 2 (b). Figure 5 (b) represents respective powers spectra of kernels 41, 146, 471 for the Fashion MNIST dataset, that is they represent the filters used to create feature maps shown in Fig. 2 (d).

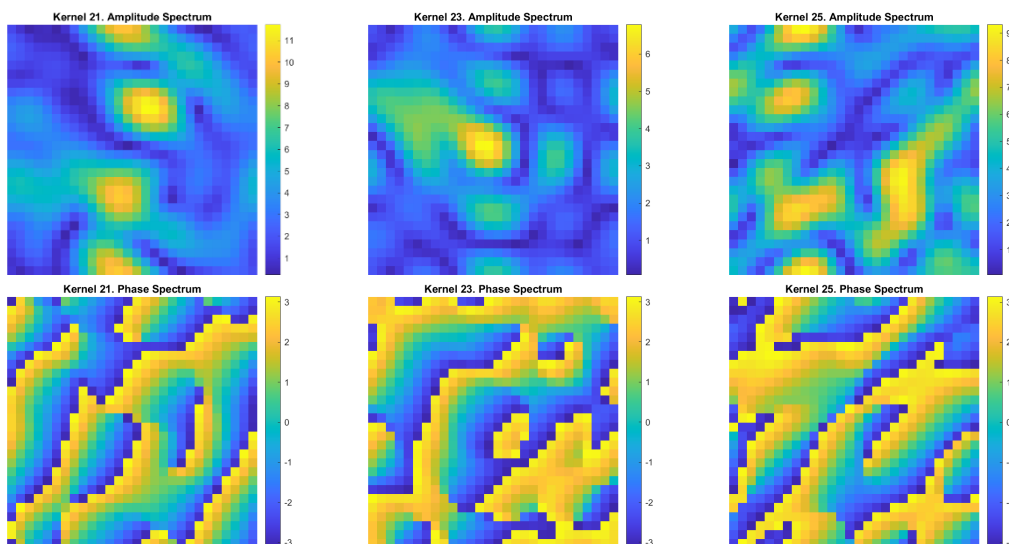


Fig 4 (a). Amplitude and Phase spectrum of kernels 21, 23, 25 of CNNMVN in MNIST dataset

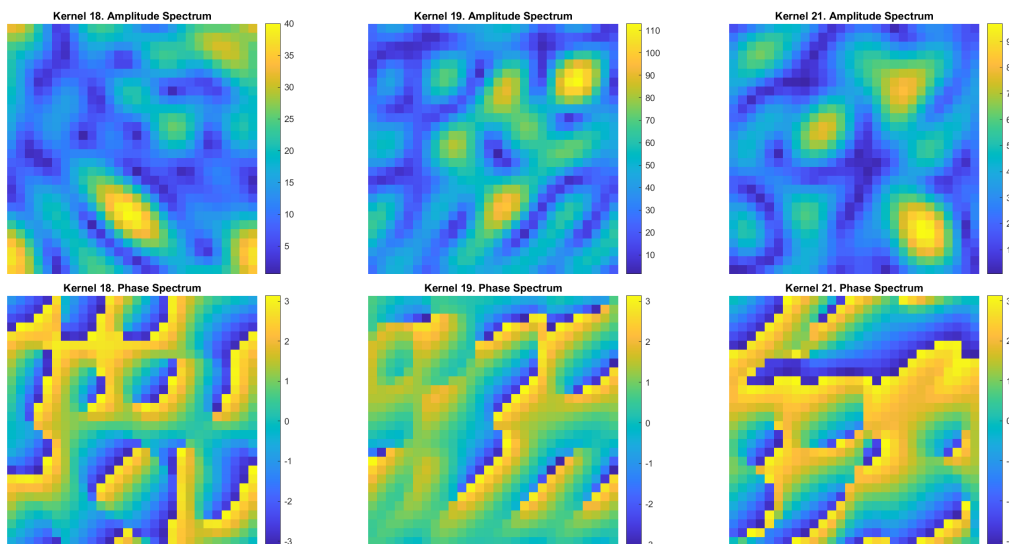


Fig. 4 (b). Amplitude and Phase spectrum of kernels 18, 19, 21 of CNNMVN in Fashion MNIST dataset

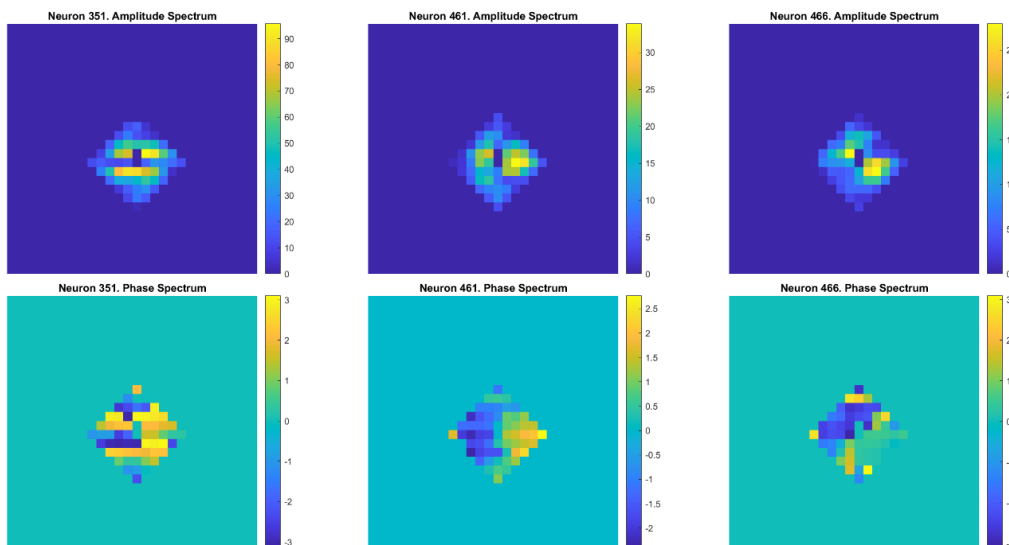


Fig. 5 (a). Amplitude and Phase spectrum of kernels 351, 451, 466 of MLMVN-FCNN in MNIST dataset

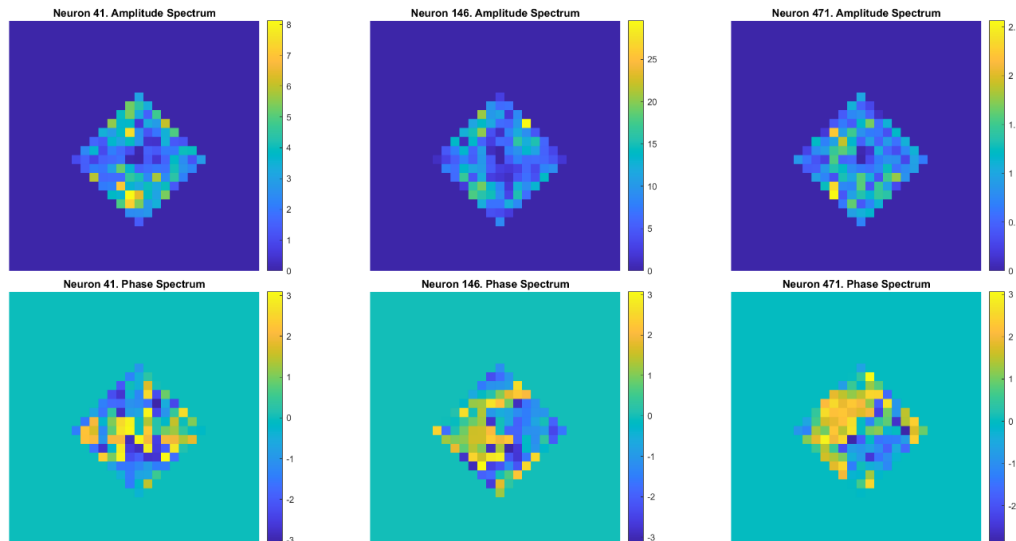


Fig. 5 (b). Amplitude and Phase spectrum of kernels 41, 146, 471 of MLMVN-FCNN in Fashion MNIST dataset

Considering that both networks reached high accuracy rate we can conclude that the extracted features, which are visualized in Fig. 2 (b) and 2 (d) are crucial for image classification. Phase spectra of the respective kernels are also shown in Fig. 4 (a), 4 (b), 5 (a) and 5 (b), but it can be seen that it does not carry some visually significant information. However, it is worth noting that visualization of the magnitude of complex-valued feature maps may not provide a complete understanding of their nature.

Discussion and conclusions

Summarizing the above, we can conclude that both CNNMVN and MLMVN-FCNN networks successfully extract meaningful features from input images despite using different approaches. Both models demonstrate their ability to generate meaningful feature maps that emphasize important image details such as edges, shapes, and specific frequency bands. CNNMVN operates in the spatial domain and produces feature maps similar to those in traditional CNNs. Its kernels are effective at detecting specific patterns such as horizontal, vertical, and diagonal lines, as well as performing high-pass filtering to highlight image details. In contrast, MLMVN-FCNN operates in the frequency domain, emphasizing specific frequency subbands that are difficult to directly compare to spatial domain kernels.

Future work will focus on investigating which regions, shapes and features of an image most frequently activate the kernels, allowing for a deeper understanding of their formation and functionality.

Authors' contribution: Igor Aizenberg – conceptualization; methodology, analysis of sources, preparation of literature review or theoretical foundations of research; Alexander Vasko – conceptualization; methodology, software, empirical data collection and validation, empirical research, analysis of sources, preparation of literature review or theoretical foundations of research.

References

- Aizenberg, I. (2011). *Complex-Valued Neural Networks with Multi-Valued Neurons* (Vol. 353). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-20353-4>
- Aizenberg, I., Herman, J., & Vasko, A. (2022). A Convolutional Neural Network with Multi-Valued Neurons: A Modified Learning Algorithm and Analysis of Performance. *2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference*, 0585–0591. <https://doi.org/10.1109/UEMCON54665.2022.9965659>
- Aizenberg, I., & Moraga, C. (2007). Multilayer Feedforward Neural Network Based on Multi-valued Neurons (MLMVN) and a Backpropagation Learning Algorithm. *Soft Computing*, 11(2), 169–183. <https://doi.org/10.1007/s00500-006-0075-5>
- Aizenberg, I., & Vasko, A. (2020). Convolutional Neural Network with Multi-Valued Neurons. *2020 IEEE Third International Conference on Data Stream Mining & Processing*, 72–77. <https://doi.org/10.1109/DSMP47368.2020.9204076>
- Aizenberg, I., & Vasko, A. (2023). MLMVN as a Frequency Domain Convolutional Neural Network. *2023 International Conference on Computational Science and Computational Intelligence*, 341–347. <https://doi.org/10.1109/CSCI62032.2023.00061>
- Aizenberg, I., & Vasko, A. (2024). Frequency-Domain and Spatial-Domain MLMVN-Based Convolutional Neural Networks. *Algorithms*, 17(8), Article 8. <https://doi.org/10.3390/a17080361>
- Beysolow II, T. (2017). Convolutional Neural Networks (CNNs). In T. Beysolow II, *Introduction to Deep Learning Using R* (pp. 101–112). Apress. https://doi.org/10.1007/978-1-4842-2734-3_5
- Boonsatit, N., Rajendran, S., Lim, C. P., Jirawattanapanit, A., & Mohandas, P. (2022). New Adaptive Finite-Time Cluster Synchronization of Neutral-Type Complex-Valued Coupled Neural Networks with Mixed Time Delays. *Fractal and Fractional*, 6(9), 515. <https://doi.org/10.3390/fractalfract6090515>
- Bruna, J., Chintala, S., LeCun, Y., Piantino, S., Szlam, A., & Tygert, M. (2015). A mathematical motivation for complex-valued convolutional networks. <https://doi.org/10.48550/ARXIV.1503.03438>
- Caldeira, M., Martins, P., Cecilio, J., & Furtado, P. (2019). Comparison Study on Convolution Neural Networks (CNNs) vs. Human Visual System (HVS). In S. Kozielski, D. Mrozek, P. Kasprowski, B. Malysiak-Mrozek, & D. Kostrzewa (Eds.), *Beyond Databases, Architectures and Structures. Paving the Road to Smart Data Processing and Analysis* (Vol. 1018, pp. 111–125). Springer International Publishing. https://doi.org/10.1007/978-3-030-19093-4_9
- Chatterjee, S., Tummala, P., Speck, O., & Nürnberger, A. (2023). Complex Network for Complex Problems: A comparative study of CNN and Complex-valued CNN. <https://doi.org/10.48550/ARXIV.2302.04584>
- Fuchs, A., Rock, J., Toth, M., Meissner, P., & Pernkopf, F. (2021). Complex-valued Convolutional Neural Networks for Enhanced Radar Signal Denoising and Interference Mitigation. *2021 IEEE Radar Conference* (pp. 1–6). <https://doi.org/10.1109/RadarConf2147009.2021.9455296>
- Gad, A. F. (2018). Convolutional Neural Networks. In A. F. Gad, *Practical Computer Vision Applications Using Deep Learning with CNNs* (pp. 183–227). Apress. https://doi.org/10.1007/978-1-4842-4167-7_5
- Guberman, N. (2016). *On Complex Valued Convolutional Neural Networks* (arXiv:1602.09046). arXiv. <http://arxiv.org/abs/1602.09046>
- Hirose, A. (2011). Complex-valued Neural Networks. *IEEE Transactions on Electronics, Information and Systems*, 131(1), 2–8. <https://doi.org/10.1541/ieejieiss.131.2>
- Hongo, S., Isokawa, T., Matsui, N., Nishimura, H., & Kamiura, N. (2020). Constructing Convolutional Neural Networks Based on Quaternion. *2020 International Joint Conference on Neural Networks*, 1–6. <https://doi.org/10.1109/IJCNN48605.2020.9207325>
- Jarrett, K., Kavukcuoglu, K., Ranzato, M. A., & LeCun, Y. (2009). What is the best multi-stage architecture for object recognition? *2009 IEEE 12th International Conference on Computer Vision*, 2146–2153. <https://doi.org/10.1109/ICCV.2009.5459469>
- Kaur, P., & Garg, R. (2020). Towards Convolution Neural Networks (CNNs): A Brief Overview of AI and Deep Learning. In G. Ranganathan, J. Chen, & Á. Rocha (Eds.), *Inventive Communication and Computational Technologies* (Vol. 89, pp. 399–407). Springer Singapore. https://doi.org/10.1007/978-981-15-0146-3_38

- Kobayashi, M. (2017). Symmetric Complex-Valued Hopfield Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 28(4), 1011–1015. <https://doi.org/10.1109/TNNLS.2016.2518672>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25. <https://doi.org/10.1145/3065386>
- LeCun, Y., Cortes, C., & Burges, C. J. C. (n. d.). *The MNIST Database of handwritten digits*. [Dataset]. Retrieved August 9, 2024, from <https://www.kaggle.com/datasets/zalando-research/fashionmnist>
- LeCun, Y., Fu Jie, Huang, & Bottou, L. (2004). Learning methods for generic object recognition with invariance to pose and lighting. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004*, 2, 97–104. <https://doi.org/10.1109/CVPR.2004.1315150>
- Lin, L., Liang, L., Jin, L., & Chen, W. (2019). Attribute-Aware Convolutional Neural Networks for Facial Beauty Prediction. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 847–853. <https://doi.org/10.24963/ijcai.2019/119>
- Lin, W., Ding, Y., Wei, H.-L., Pan, X., & Zhang, Y. (2020). LdsConv: Learned Depthwise Separable Convolutions by Group Pruning. *Sensors*, 20(15), 4349. <https://doi.org/10.3390/s20154349>
- Nitta, T. (2004). Orthogonality of Decision Boundaries in Complex-Valued Neural Networks. *Neural Computation*, 16(1), 73–97. <https://doi.org/10.1162/08997660460734001>
- Nitta, T. (2008). The uniqueness theorem for complex-valued neural networks with threshold parameters and the redundancy of the parameters. *International Journal of Neural Systems*, 18(02), 123–134. <https://doi.org/10.1142/S0129065708001439>
- Popa, C.-A. (2017). Complex-valued convolutional neural networks for real-valued image classification. *2017 International Joint Conference on Neural Networks* (pp. 816–822). <https://doi.org/10.1109/IJCNN.2017.7965936>
- Rawat, S., Rana, K. P. S., & Kumar, V. (2021). A novel complex-valued convolutional neural network for medical image denoising. *Biomedical Signal Processing and Control*, 69, 102859. <https://doi.org/10.1016/j.bspc.2021.102859>
- Sarabu, A., & Santra, A. K. (2021). Human Action Recognition in Videos using Convolution Long Short-Term Memory Network with Spatio-Temporal Networks. *Emerging Science Journal*, 5(1), 25–33. <https://doi.org/10.28991/esj-2021-01254>
- Sunaga, Y., Natsuaki, R., & Hirose, A. (2020). Similar land-form discovery: Complex absolute-value max pooling in complex-valued convolutional neural networks in interferometric synthetic aperture radar. *2020 International Joint Conference on Neural Networks* (pp. 1–7). <https://doi.org/10.1109/IJCNN48605.2020.9207122>
- Suresh, S., Sundararajan, N., & Savitha, R. (2013). *Supervised Learning with Complex-valued Neural Networks* (Vol. 421). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-29491-4>
- Valle, M. E. (2014). Complex-Valued Recurrent Correlation Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 25(9), 1600–1612. <https://doi.org/10.1109/TNNLS.2014.2341013>
- Venkatesan, R., & Li, B. (2017). Modern and Novel Usages of CNNs. In R. Venkatesan & B. Li, *Convolutional Neural Networks in Visual Computing* (1st ed., pp. 117–146). CRC Press. <https://doi.org/10.4324/9781315154282-5>
- Wu, D., Zhang, J., & Zhao, Q. (2020). A Text Emotion Analysis Method Using the Dual-Channel Convolution Neural Network in Social Networks. *Mathematical Problems in Engineering*, 2020(1), 6182876. <https://doi.org/10.1155/2020/6182876>
- Xiao, H., Rasul, K., & Vollgraf, R. (2017). *Fashion-MNIST: A Novel Image Dataset for Benchmarking Machine Learning Algorithms* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.1708.07747>
- Xiao, L., Zhang, H., Chen, W., Wang, Y., & Jin, Y. (2018). Transformable Convolutional Neural Network for Text Classification. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 4496–4502. <https://doi.org/10.24963/ijcai.2018/625>
- Yadav, S., & Jerripothula, K. R. (2023). FCCNs: Fully Complex-valued Convolutional Networks using Complex-valued Color Model and Loss Function. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 10655–10664. <https://doi.org/10.1109/ICCV51070.2023.00981>
- Yar, H., Abbas, N., Sadad, T., & Iqbal, S. (2021). Lung Nodule Detection and Classification using 2D and 3D Convolution Neural Networks (CNNs). In L. M. Goyal, T. Saba, A. Rehman, & S. Larabi-Marie-Sainte, *Artificial Intelligence and Internet of Things* (1st ed., pp. 365–386). CRC Press. <https://doi.org/10.1201/9781003097204-17>
- Zhang, Z., Wang, H., Xu, F., & Jin, Y.-Q. (2017). Complex-Valued Convolutional Neural Network and Its Application in Polarimetric SAR Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 7177–7188. <https://doi.org/10.1109/TGRS.2017.2743222>
- Zhang, Z., Wang, Z., Chen, J., & Lin, C. (2022). *Complex-Valued Neural Networks Systems with Time Delay: Stability Analysis and (Anti-)Synchronization Control* (Vol. 4). Springer Nature Singapore. <https://doi.org/10.1007/978-981-19-5450-4>
- Zhu, X., Xu, Y., Xu, H., & Chen, C. (2018). *Quaternion Convolutional Neural Networks*, 631–647. <https://doi.org/10.48550/arXiv.1903.00658>

Отримано редакцією журналу / Received: 21.01.25

Прорецензовано / Revised: 24.02.25

Схвалено до друку / Accepted: 06.03.25

Ігор АЙЗЕНБЕРГ, д-р філософії, проф.

ORCID ID: 0000-0002-5994-6568

e-mail: igor.aizenberg@manhattan.edu

Манхеттенський університет, Рівердейл, Нью-Йорк, США

Олександр ВАСЬКО, асп.

ORCID ID: 0009-0006-1527-505X

e-mail: oleksandrvasco@uzhnu.edu.ua

ДВНЗ "Ужгородський національний університет", Ужгород, Україна

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЗГОРТОК, ОТРИМАНИХ МЕРЕЖАМИ CNNMVN ТА MLMVN, ЯК ЗГОРТКОВОЇ МЕРЕЖІ В ЧАСТОТНІЙ ОБЛАСТІ

Кожен згортковий шар у будь-якій згортковій нейронній мережі створює карту ознак, що містить найважливішу інформацію, необхідну мережі для розпізнавання відповідних зображень. Для подальшого вдосконалення цих нейронних мереж і кращого розуміння їхніх можливостей важливо виявити, які саме ознаки витягаються і як змінюються зображення, що підлягають розпізнаванню, унаслідок згортки, отриманих у процесі навчання.

У пропонованій статті представлено порівняльний аналіз згортки, отриманих за допомогою двох комплекснозначних нейронних мереж, які базуються на багатозначних нейронах. Перша мережа – це згорткова нейронна мережа на основі багатозначних нейронів (CNNMVN), що має традиційну топологію згорткової нейронної мережі, за винятком того, що вона використовує комплекснозначні згорткові ядра у своїй згортковій частині, та багатозначні нейрони у повноз'єзній частині. Друга мережа – це багатозначна нейронна мережа на основі багатозначних нейронів (MLMVN), що є повноз'єзною багатошаровою нейронною мережею, яку використовують як згорткову мережу в частотній області.

Ураховуючи, що обидві нейронні мережі є комплекснозначними, а отримані фільтри працюють у комплексній площині, проведене дослідження показує, що ядра обох мереж формують фільтри, подібні до наявних цифрових фільтрів для оброблення зображень. Аналіз ядер CNNMVN виявив, що вони реалізують фільтри нерізкого маскуванню та фільтри виявлення країв для ідентифікації форм на зображеннях, тоді як ядра MLMVN підсилюють окремі частотні піддіапази, що ускладнює їхнє пряме порівняння з просторовими ядрами. Отже, ядра обох згорткових мереж сприяють поліпшенню точності розпізнавання зображень. Візуалізація цих ядер демонструє, що комплекснозначні згорткові нейронні мережі функціонують аналогічно до їхніх дійснозначних аналогів.

Ключові слова: згорткові нейромережі, комплекснозначні мережі, багатозначні нейрони, CNNMVN, MLMVN, розпізнавання зображень, частотна область.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 517.9

DOI: <https://doi.org/10.17721/1812-5409.2025/1.13>

Konstantin ATOYEV, PhD (Biol.), Senior Researcher

ORCID ID: 0000-0001-9892-054X

e-mail: konstantin_atoyev@yahoo.com

V. M. Glushkov Institute of Cybernetics, NAS of Ukraine, Kyiv, Ukraine

Pavlo KNOPOV, DSc (Phys. & Math.), Prof., Corresponding Member of the National Academy of Sciences of Ukraine

ORCID ID: 0000-0001-6550-2237

e-mail: knopov1@yahoo.com

V. M. Glushkov Institute of Cybernetics, NAS of Ukraine, Kyiv, Ukraine

SIX-SECTOR LORENZ MODEL FOR CRITICAL INFRASTRUCTURE RISK MANAGEMENT

The work is dedicated to the problem of developing mathematical methods for assessing the vulnerability of critical infrastructure (CI).

The main goal of the study is to create a mathematical model for forecasting crisis phenomena, ranking threat levels, and identifying weak links in CI that significantly affect the deformation of the security space. The model includes an economic block and a risk assessment block. The economic block is described by a six-sector Lorenz model with variable coefficients, which integrates similarly described economic sectors into a unified structure. Each sector is considered in terms of productivity, employment, and structural disruptions in sectors directly related to CI activities: energy, transport infrastructure, water supply networks and hydraulic structures, chemical industry, production, storage, and transportation of hazardous materials, as well as food and medical supplies, and information technology. The risk assessment block uses catastrophe theory methods to evaluate vulnerability of CI. Risk is assessed based on the proximity of the model parameters to their bifurcation values, upon reaching which the system transitions from a stationary state, characterized by acceptable vulnerability, to a state of increased vulnerability.

The main results of the work are related to the numerical study of the developed model. The conditions for the transition of interconnected economic sectors to a regime of deterministic chaos, due to changes in specific demand and supply for production system outputs and the number of jobs in them, as well as changes in the rate of elimination of structural disruptions, have been identified. It has been shown that there are parameter value ranges in the model where minor changes lead to significant transformations in the CI system and its security.

Further development of this work will focus on determining the dependence of the control parameters of the developed model on technological, environmental, economic, and social factors that characterize the functioning of CI.

Keywords: *Lorentz model, mathematical modeling, critical infrastructure, deterministic chaos, risk assessment.*

AMS 2020 classification: 62H30, 34C60.

Introduction

The increase in the number of technogenic disasters, along with the scale of the damage they cause, has focused attention on the problem of a crisis in management systems. This crisis is associated with a narrowing of the forecasting horizon, manifesting in the fact that low-probability events occur much more frequently than predicted by traditional safety assessment methods. Due to globalization, the technosphere has transformed into a system characterized by the following features: 1) the presence of multiple development trajectories and "windows of vulnerability"; 2) a limited timeframe for decision-making; 3) the rapid pace of changes in order parameters that define human safety; 4) a strong dependence on initial conditions, making it so that small influences can lead to significant divergences in development trajectories (Wilkinson, Elani, & Eidinow, 2002).

The increasing role of minor impacts expands the scope of traditional risks and gives rise to fundamentally new ones. These are associated with events of enormous destructive power, prolonged in time, unfolding on a transnational scale, where their effects are felt in geographical zones far from the incident site (Renn, 2003). This has led to a sharp increase in the vulnerability of critical infrastructure (CI). Until recently, security policy in the field of CI protection was based on addressing issues related only to the crisis in management systems (Ermoliev, & von Winterfeldt, 2012; Norkin et al., 2018; Pepelyaev, Golodnikov, & Golodnikova, 2022; Zgurovsky, 2024). However, in the context of military conflicts, it is necessary to approach the problem of CI protection from a new perspective. There is a need to significantly revise the old security paradigms, considering the role that CI has come to play in modern warfare as a target for attacks. It is essential to consider the likelihood of deliberate destruction of energy facilities, hydraulic structures, and hazardous enterprises in the nuclear and chemical industries. Therefore, the task of developing and implementing automated systems that enable forecasting and prevention of crisis phenomena, ranking threat levels, and identifying weak points that significantly influence the deformation of the security space has become highly relevant. The development of mathematical models and methods to address these tasks is the goal of this work.

In works (Sergienko, Yanenko, & Atoev, 1997; Atoev et al., 2017; Atoev, & Knopov, 2022), mathematical models for analyzing the vulnerability of CI have been developed based on the theory of smooth mappings, using examples from housing and communal services, potentially hazardous facilities, and environmentally polluting enterprises. The advantages of the proposed approach include the ability to: 1) forecast and manage complex risks; 2) rank different types of threats, identify weak links in CI, and quantitatively assess the proximity to critical parameter values indicating increase in risk; 3) calculating, using optimal control theory, the trajectories of changes in CI parameters to reduce the risks of man-made disasters and minimize the costs of their elimination.

In works (Atoev, 2005; Atoev, & Knopov, 2023), multi-sector Lorenz models were developed to study the interconnections between different sectors of the economy and to determine the conditions for instability in such complex systems. These models integrate in a holistic structure various economic sectors, described in the same way. Each sector is considered in terms of productivity levels, the number of jobs, and structural disruptions. They allow for the calculation of the losses from instability in the functioning of such complex systems as CI. The measure of these losses is the level of structural disruptions.

To assess risks for CI, we will combine these two approaches. The model should include: 1) a block for evaluating the proximity of parameters characterizing CI to their bifurcation values, the attainment of which sharply increases the probability

© Atoev Konstantin, Knopov Pavlo, 2025

of its destruction; 2) an economic block for assessing the impact of CI destruction on the economy. Furthermore, the system of interconnections between these blocks should be considered.

2. Mathematical model of vulnerability assessment

The mathematical formalism of catastrophe theory (Poston, & Stewart, 1978) allows for the calculation of bifurcation values, curves, and surfaces of control parameters. The risk of the system transitioning from one steady state (normal) to another (catastrophe) will be assessed by the degree of proximity of the control parameter values to their bifurcation values.

The function $U(x)$, called the potential, will be introduced as follows:

$$U(x) = - \int F(x) dx$$

where $F(x)$ is a function describing the motion in a potential field $U(x)$.

In mechanics, quantum mechanics, and field theory, Hamiltonian systems are primarily considered. However, in "real life", we often deal with dissipative systems. To study the system behavior near local extrema function $U(x)$, the function $F(x)$ is expanded into a series in the neighborhood of stationary states, and the expansion is limited to a few small terms. To assess risk, we will use the "cusp" catastrophe, whose potential has the following form:

$$U(x) = -\frac{x^4}{4} - A_1 \frac{x^2}{2} - A_2 x.$$

The manifold of the catastrophe is defined by the equation

$$-\frac{\partial U(x)}{\partial x} = \frac{dx}{dt} = x^3 + A_1 x + A_2. \quad (1)$$

The polynomial (1), according to Sturm's theorem regarding the number of real roots of a polynomial (Andronov, Witt, & Khaikin, 1959), has at most three and at least one real root. The number of roots depends on the discriminant $\Delta = 4A_1^3 + 27A_2^2$. If the condition $\Delta < 0$ is satisfied, there are three real roots. The system has three steady states, two of which are stable. The first stable steady state represents the norm (an acceptable level of vulnerability of the system), while the second represents a crisis (a dangerous level of vulnerability). When $\Delta > 0$ there is one real root and two imaginary ones. The bifurcation curve has the following form:

$$\Delta(A_1, A_2) = 0. \quad (2)$$

The measure of the risk for the transition of the system from one steady state, corresponding to the norm, to another, corresponding to a crisis, is the degree of proximity of parameters A_1 and A_2 to their bifurcation values A_1^* and A_2^* . Since the parameter A_2 characterizes rapid impacts on the system, it was used to assess the risk in (Atoyev, Tomin, & Aksionova, 2007; Atoyev et al., 2017). The bifurcation value of this parameter is determined from equation (2)

$$A_2^* = \pm 2 \frac{A_1}{3} (A_1/3)^{1/2}.$$

The risk value R can be assessed as follows:

$$R = 1 - (A_2^* - A_2)/A_2^*. \quad (3)$$

If $A_2 = 0$, then $R = 0$, and if $A_2 = A_2^*$, then $R = 1$. If $A_2 > A_2^*$, we will also consider $R = 1$. For the chosen acceptable level of risk R_{ac} , the corresponding value A_{2ac} is determined from (3), which will be used in further research to determine the strategy for ensuring the specified level of CI security.

Risk assessment for critical infrastructure requires reliable information about the objects being studied. The absence of such information, due to incomplete sampling or the stochastic nature of the research object, leads to uncertainty (Pepelyaev, Golodnikov, & Golodnikova, 2024). In this situation, the effectiveness of deterministic models decreases, making the issue of accounting for stochasticity particularly relevant (Wagenmakers et al., 2005).

There are several approaches to addressing this issue. In (Ebeling, & Engel-Herbert, 1982), the case of non-equilibrium stochastic systems approaching a stationary state was examined. It was shown that for gradient systems, the set of transitions (catastrophes) coincides under both deterministic and stochastic descriptions and consists of all points in the control space where the number and type of extremal points change. When the modeled process is stochastic and governed by the Langevin equation, equation (1) transforms into the following form:

$$-\frac{\partial U(x)}{\partial x} = \frac{dx}{dt} = x^3 + A_1 x + A_2 + D^{1/2} \xi(t)$$

where D is a constant, $\xi(t)$ is an independent Gaussian δ -correlated fluctuating quantity. According to (Ebeling, & Engel-Herbert, 1982), the system is described using a single, globally stable stationary probability distribution $P^0(x, A)$, which is related to $U(x, A)$, as follows:

$$P^0(x, A) = (1/n) \exp[2U(x, A)/D].$$

3. Model for assessing the economic consequences of CI disruptions

For the economic block of the model, we will use a multi-sector Lorenz model. In this context, as in (Atoyev, & Knopov, 2023), we will consider the following assumptions: 1) there is competition for labor among different sectors of the economy, so the growth of the production function in one sector hinders job creation in others; 2) fluctuations in energy prices, changes in the balance of natural and material resources, changes due to the disruption of CI, and the introduction of new technologies that significantly alter productivity – all of these provoke random disturbances $w_{ij}(t)$ that become additional factors for increasing the level of structural disruptions; 3) processes in different sectors of the economy may proceed at different speeds, with time scaling across sectors carried out using parameters ε_i ; 4) An increase in the vulnerability level of CI x leads to a decrease in the system's capacity to recover from structural disruptions; 5) control parameters A_{1i} and A_{2i} change with an increase in structural disruptions. The mathematical model that integrates the economic and security blocks has the following form:

$$\begin{aligned} \varepsilon_i \frac{dX_i}{dt} &= \sigma_i(Y_i - X_i) + \delta_i \dot{w}_{ij}, \\ \varepsilon_i \frac{dY_i}{dt} &= [F_i(X_1, X_2, \dots, X_6) - Z_i]X_i - Y_i + \delta_i \dot{w}_{ij}, \end{aligned} \quad (4)$$

$$\begin{aligned}\varepsilon_i \frac{dZ_i}{dt} &= X_i Y_i - b_i (1 - k_i x) Z_i + \delta_i \dot{w}_{ij}, \\ \varepsilon_x \frac{dx}{dt} &= x^3 + \left[A_{1i} + \sum_{i=1}^6 l_i Z_i \right] x + A_{2i} + \sum_{i=1}^6 m_i Z_i + \delta_i \dot{w}_{ij},\end{aligned}$$

where X_i are normalized levels of productivity, Y_i are normalized jobs, Z_i are normalized level of structural disturbances in the following sectors: energy facilities and networks ($i=1$); 2) transport service($i=2$); hydraulic structures, water supply networks ($i=3$); chemical industry, production, transportation and storage of hazardous materials ($i=4$); food and medical supply facilities ($i = 5$); communication and information technology ($i = 6$); x is normalized level of vulnerability of CI; σ_i, r_i, b_i ($i = \overline{1,6}$) are parameters of the Lorenz model, which are functions of specific demands and supplies for the products of the respective production systems and the number of jobs occupied in them; $F_i = r_i(1 - \sum_{i=1}^6 a_{ij} X_i)$, a_{ij} are parameters characterizing competition between different sectors of the economy in labor markets ($i \neq j$); k_i, l_i, m_i are parameters characterizing the impact of critical infrastructure vulnerability on economic sectors, and the impact of structural disruptions on the dynamics of critical infrastructure vulnerability, respectively; w_{ij} are independent standard Wiener processes with parameters $E(w_{ij}(t) - w_{ij}(s)) = 0$; $E(w_{ij}(t) - w_{ij}(s))^2 = |t - s|$; δ_i are intensities of disturbance.

According to the results obtained in (Atojev, 2005), the parameters σ_i, r_i, b_i are related to the characteristics that determine the functioning of economic sectors as follows

$$\sigma_i = (\alpha_{1i} \beta_{2i}) / (\alpha_{2i} \gamma_{2i}), \quad r_i = (\beta_{1i} \gamma_{1i}) / (\beta_{2i} \gamma_{2i}), \quad b_i = \zeta_i / (\alpha_{2i} \gamma_{2i}), \quad (5)$$

where α_{1i} and α_{2i} are parameters that characterize adaptive capabilities; β_{1i} are demands for the activity of the i -th manufacturing system, normalized to the unit of measurement of the material production system, which is considered to be a job in the corresponding industry Y_i ; β_{2i} are supplies normalized to the unit of function of the i -th manufacturing system X_i ; γ_{1i} are demands for increasing the number of jobs, normalized to the unit of X_i ; γ_{2i} are supplies for increasing the number of jobs, normalized to the unit of Y_i ; ζ_i are the specific rates of disruption elimination.

The investigation of the classic Lorenz model (Lorenz, 1963) shows that increasing the parameter r is crucial for the onset of turbulence. The operating modes change at the following bifurcation values of r : $r = 13.926$, $r = 24.06$, $r = 24.74$. When $r > 24.74$, the system transitions to a regime of metastable chaos – a strange attractor emerges (Kaplan, & Yorke, 1979). In model (4) the transition of parameters r_i to the range of values corresponding to the regime of metastable chaos is achieved by increasing the specific demand for the products of the manufacturing system and the number of jobs employed (parameters β_{1i} and γ_{1i}) or by reducing their specific supply (parameters β_{2i} and γ_{2i}).

Let's consider how the rate of disruption elimination affects the level of security of CI. As follows from (5), a decrease in the rates of disruption elimination ζ_i , leads to decline in parameters b_i . Rates ζ_i change significantly faster than the adaptive capabilities α_{2i} and the supply of increasing the number of jobs γ_{2i} . Within short time intervals α_{2i} and γ_{2i} can be consider constants. In these conditions, we will associate the decrease in b_i only with the reduction in ζ_i .

Figure 1 shows how the dynamics of vulnerability x , the discriminant Δ , and the proximity of the current values of the parameters A_2 to their bifurcation values A_2^* change with the decrease of the parameters b_i of model (2). For simplicity in modeling, it was assumed that the coefficients for different sectors are the same.

Figure. 2 shows how the level of the total structural disruption of the CI ($\sum \int Z_i(t) dt$) changes with the variation of the parameters r_i . The results of the research show that there are regions of parameters r_i . in which the level of structured disruption increases sharply. For the energy sector ($i=1$), when the demand-to-supply ratio reaches a level corresponding to the values of parameter r_1 in the range (22,24), the level of total disruption increases from 3856.4 to 8077.6. A further increase in r_1 initially reduces this level to 3925.9 at $r_1 = 26$, but then the level of structural disturbances begins rise again.

In Figure 3, it is shown how the dynamics of disturbances Z_i change with a decrease in parameters b_i , as well as the projections of the phase space of system (2) onto the planes $Z_i - X_i$ and $Z_i - Y_i$, using one of the sectors of the economy related to the energy system as an example ($i = 1 = E$). The simulation results presented in Fig. 1 and Fig. 2 were obtained with the following values for the model parameters: $A_{1i} = 2$, $A_{2i} = 1$, $k_i = 0.5$, $l_i = m_i = 0.015$, $r_i = 24$, $\varepsilon_i = \varepsilon_x = 0.2$.

4. Dependence on small changes in parameters

The formation of a complex and interconnected network of modern society significantly deforms the fields of natural, technological, and economic risks. The nodes of this network serve as unique acupuncture points of vulnerability for CI, where small influences can lead to unforeseen consequences on a large scale with irreversible damage (a kind of domino effect). This occurs due to the density of connections between nodes, unprecedented lengthening of causal relationships, and the speed at which material values and information circulate through the links of this network. The behavior of such systems often becomes unstable (Laszlo, 1991).

Therefore, the network society can be seen as a system near an unstable stationary state within a bifurcation zone, where the likelihood of sudden and unforeseen events increases. Very small fluctuations can lead to abrupt changes in the socio-economic and natural-technological characteristics of society, with the amplitude of these jumps potentially being quite significant (Zgurovsky, & Pankratova, 2007; Ermoliev, & von Vinterfeldt, 2012).

Figure 4 illustrates the results of modeling, where a very slight change in the parameter l_i leads to a significant shift in the operating mode of system (4). When $l_i = 1$ the level of structural disruptions reaches maximum $\sum \int Z_i(t) dt = 8077.6$. With a small increase or decrease in l_i by a value of 10^{-9} , the level of structural disruptions significantly decreases (as shown in Figures 4b and 4d). Moreover, in this case the phase portraits of the system differ significantly.

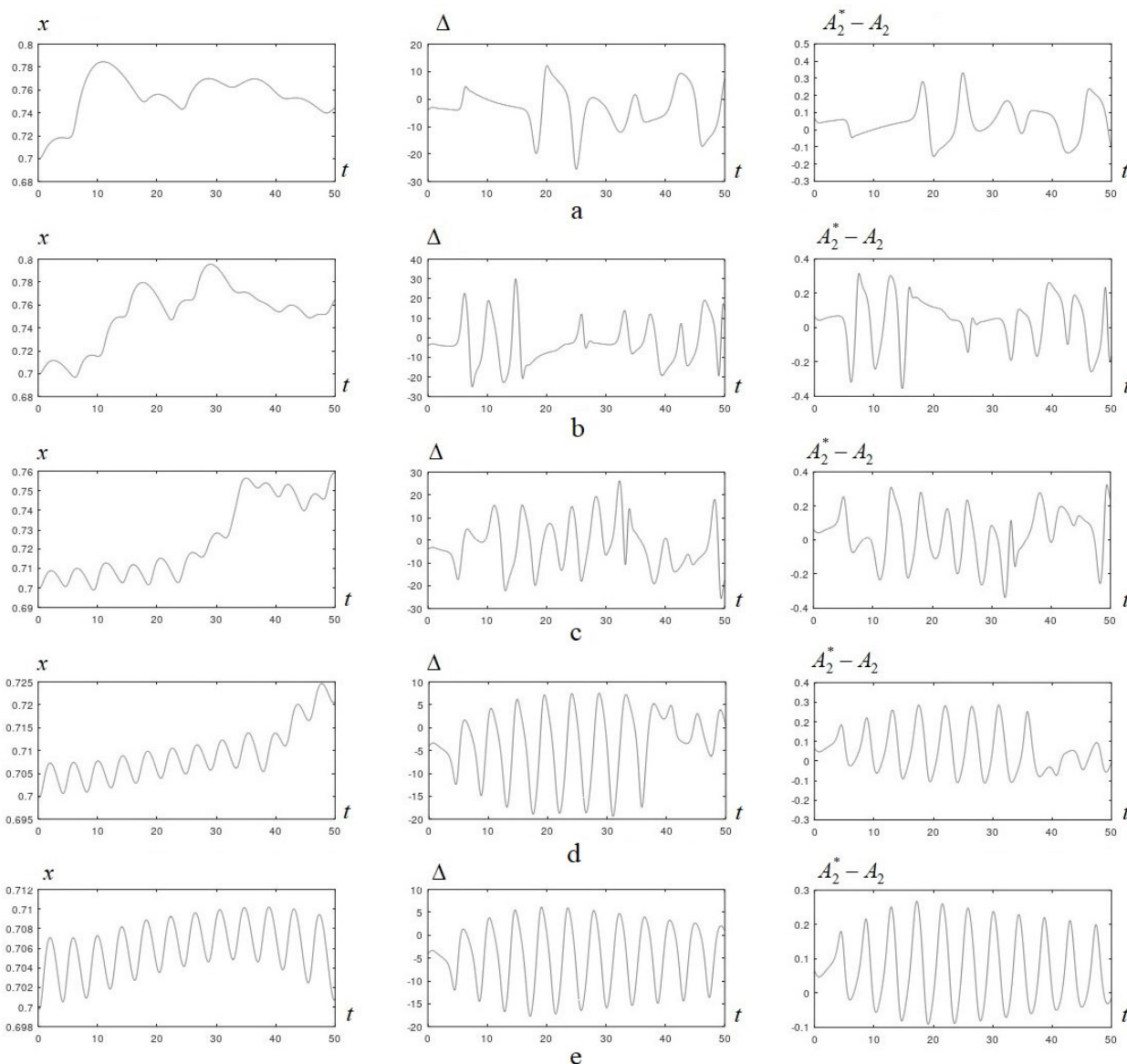


Fig. 1. The impact of changes in the rate of disruption elimination on the dynamics of changes in the vulnerability levels of CI:
 $b_i = 5/3$ (a); $b_i = 8/3$ (b); $b_i = 12/3$ (c); $b_i = 15.4825/3$ (d); $b_i = 16/3$ (e)

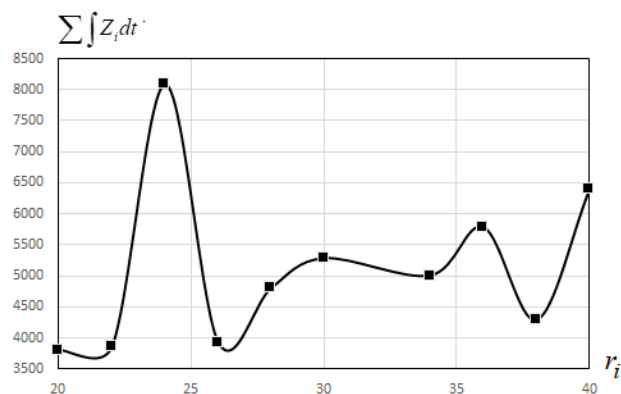


Fig. 2. Dependence of total structural disruptions of CI on parameter r_i .

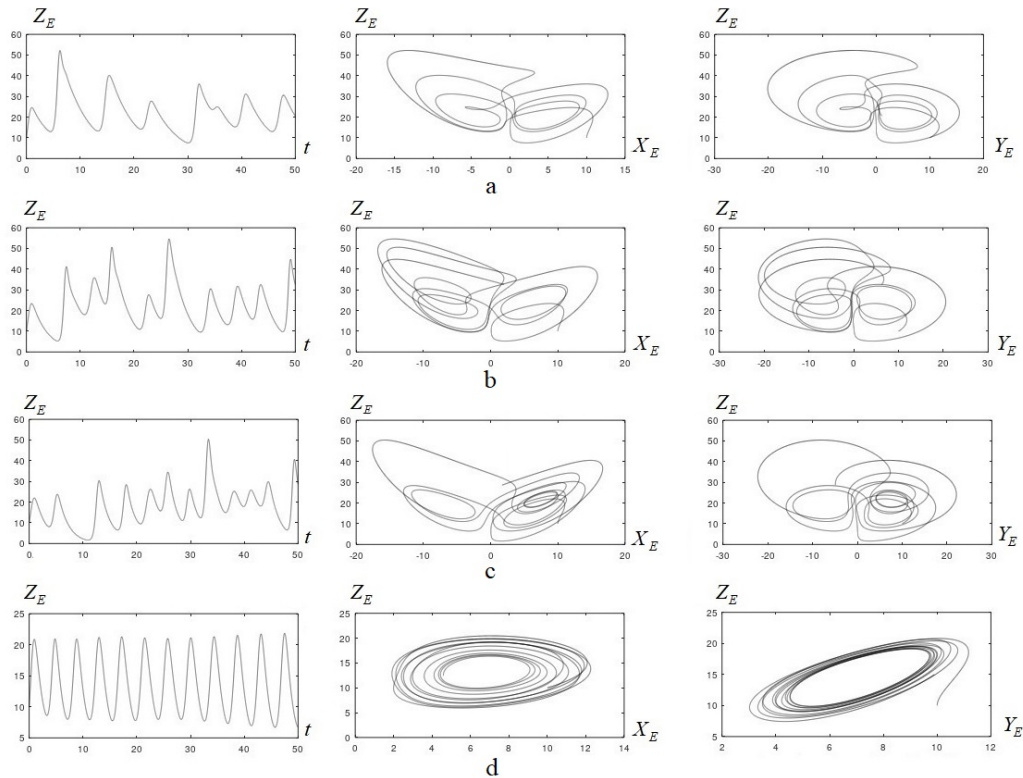


Fig. 3. The impact of changes in the rate of disruption elimination on the dynamics of changes in disturbances in energetic sector: $b_i = 5/3$ (a); $b_i = 8/3$ (b); $b_i = 12/3$ (c); $b_i = 16/3$ (d)

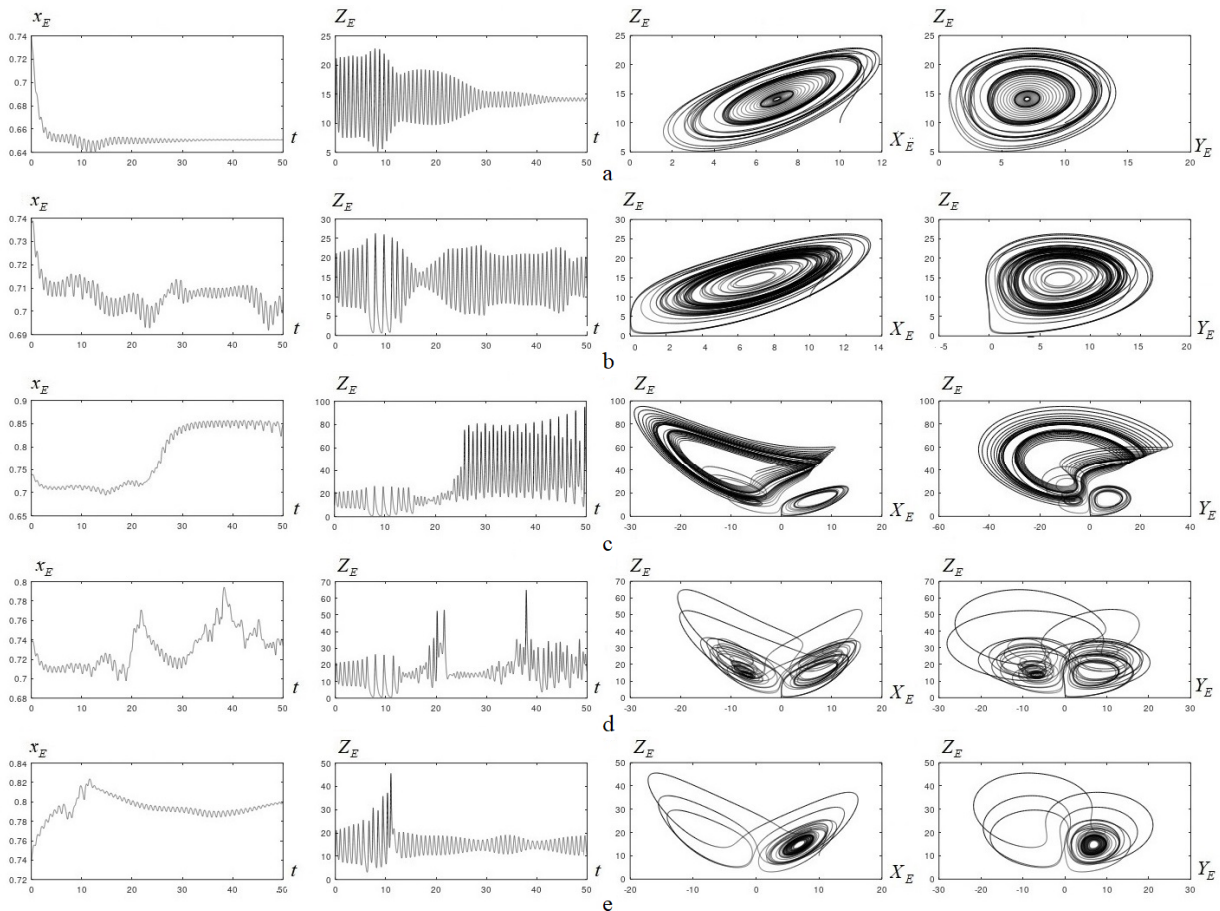


Fig. 4. Changes in the vulnerability of CI and in the level of structural disruptions in the energy sector with an increase in parameter l_i : $l_i = 0.954$, $\sum \int Z_i(t)dt = 3853.6$ (a); $l_i = 0.999999999$, $\sum \int Z_i(t)dt = 3870.9$ (b); $l_i = 1.0$, $\sum \int Z_i(t)dt = 8077.6$ (c); $l_i = 1.000000001$, $\sum \int Z_i(t)dt = 5192.8$ (d); $l_i = 1.046$, $\sum \int Z_i(t)dt = 4159.6$ (e)

Thus, a slight change in the balance of supply and demand can significantly alter the development trajectory of the system, change the vulnerability of CI, and affect the scale of negative consequences in case of its disruption. At the same time, due to this high sensitivity to small parameter changes under chaotic conditions, it is possible to shift the system from one unstable limit cycle to another, thereby qualitatively altering the system's behavior with minimal costs for changes in the interconnected sectors of the economy.

The further development of this work will be aimed at determining the dependence of the control parameters A_i and A_2 of model (1) on the technological, environmental, economic, and social factors that characterize the functioning of CI. These parameters may be determined using approaches, elaborated in (Atoyev et al., 2017; Ermoliev et al., 2020; Atoyev, & Knopov, 2023; Pepelyaev, Golodnikov, & Golodnikova, 2023; Shulzhenko, Nechaieva, & Leshchenko, 2024).

Discussion and conclusions

A mathematical approach has been proposed to assess the security level of critical infrastructure (CI) objects and the consequences of their destruction for interrelated sectors of the economy. This approach allows for the evaluation of threat levels, defined using the formalism of smooth function theory, based on the proximity of CI parameters to their bifurcation values, the achievement of which sharply increases the likelihood of their destruction. This way, the "safety margin" of objects is calculated. It is assumed that the probability of an attack on a CI object is higher the greater its vulnerability level. The assessment of the impact of CI destruction consequences on the economy is carried out using a multi-sectoral Lorenz model, which determines the dynamics of productivity levels, employment, and structural disruptions in sectors directly related to CI activities: energy, transportation, hydraulic structures and water supply networks, chemical industry, production, storage and transportation of hazardous materials, food and medical supply facilities, and information technology.

The study investigated the impact of changes in the speed of eliminating structural disruptions on the dynamics of vulnerability levels in CI. Conditions for the transition of interrelated sectors of the economy into a state of deterministic chaos were identified as a result of changes in the specific demands and supplies of products in production systems and the number of occupied jobs, as well as changes in the speed of eliminating structural disruptions. It was demonstrated that there are regions of parameter values in the model where minor changes can lead to significant transformations in the CI system and its security space.

Authors' contribution: Konstantin Atoyev – conceptualization, methodology, empirical data collection and validation, empirical research, analysis of sources, preparation of literature review or theoretical foundations of research; Pavel Knopov – methodology, analysis of sources, preparation of literature review or theoretical foundations of research.

Acknowledgments, funding sources. This work was supported by a joint project between the International Institute for Applied Systems Analysis (IIASA) and the National Academy of Sciences of Ukraine (NASU) on "Integrated robust modeling and management of food-energy-water-land use nexus for sustainable development", and by the project "Comprehensive analysis of robust preventive and adaptive measures of food, energy, water and social management in the context of systemic risks and consequences of COVID-19" (0122U000552) of National Academy of Sciences of Ukraine).

References

- Andronov, A., Witt, A., & Khaikin, S. (1959). *Theory of oscillations*. Physics and Mathematics Literature [In Russian]. [Андронов, А., Витт, А., & Хайкин, С. (1959). *Теория колебаний*. Физико-математическая литература].
- Atoyev, K. (2005). The Challenges to Safety in the East Mediterranean: Mathematical Modeling and Risk Management of Marine Ecosystems. In E. Levner, I. Linkov, & J.-M. Proth (Eds.), *Strategic Management of Marine Ecosystems*, (pp. 179–197). Dordrecht: Springer. http://dx.doi.org/10.1007/1-4020-3198-X_12
- Atoyev, K., Ermoliev, T., & Knopov P. (2024). Mathematical Modeling of Interconnections Between Ecological, Food, and Economic Dimensions of Security. In *Nexus of Sustainability, Studies in Systems, Decision and Control* (A. Zagorodny, V. Bogdanov, & A. Zaporozhts, Eds.), 559, 33–60. https://doi.org/10.1007/978-3-031-66764-0_2
- Atoyev, K., & Knopov P. (2022). Application of robust methods for estimation of distribution parameters with a priori constraints on parameters in economics and engineering. *Cybernetics and Systems Analysis*, 58(5), 713–20. <https://doi.org/10.1007/s10559-022-00504-8>
- Atoyev, K., & Knopov, P. (2023). Mathematical modeling of climate change impact on relationships of economic sectors. *Cybernetics and Systems Analysis*, 59(4), 535–545. <https://doi.org/10.1007/s10559-023-00589-9>
- Atoyev, K., Knopov, P., Pepelyaev, V., Kisala, P., Romaniuk, R., & Kalimoldayev M. (2017). The mathematical problems of complex systems investigation under uncertainties. In W. Wojcik, & J. Sikora (Eds.), *Recent Advances in Information Technology*, 6, 135–171. <https://doi.org/10.1201/9781351243179-6>
- Atoyev, K., Tomin, A., & Aksionova, T. (2007). Global changes, new risks, and novel methods and tools of their assessment. Modeling and management of environmental security in Ukraine. In I. Linkov, R. J. Wenning, & G. A. Kiker (Eds.), *Managing Critical Infrastructure Risks*, Springer Netherlands (pp. 339–351). https://doi.org/10.1007/978-1-4020-6385-5_18
- Ebeling, W., & Engel-Herbert, H. (1982). Extremal Principles and Catastrophe Theory for Stochastic Models of Nonlinear Irreversible Processes. *Thermodynamics and Kinetics of Biological Processes*, 197–215.
- Ermoliev, Yu., & von Winterfeldt, D. (2012). Systemic risk and security management. *Managing Safety of Heterogeneous Systems*, 19–49.
- Ermoliev, Yu., Zagorodny, A., Bogdanov, V., Ermolieva, T., Havlik, P., Obersteiner, M., & Rovenskaya, E. (2020). Linking distributed sectorial optimization models under asymmetric information: towards robust food-water-environmental nexus. In Yu. Ermoliev, A. Zagorodny, V. Bogdanov, T. Ermolieva, & Yu. Kostyuchenko (Eds.), *FEW Nexus for sustainable Development: Integrated Modeling & Robust Management* (pp. 303–322).
- Kaplan, J., & Yorke, J. (1979). Preturbulence: A regime observed in a fluid flow model of Lorenz. *Communications in Mathematical Physics*, 67(2), 93–108. <https://doi.org/10.1007/BF01221359>
- Laszlo E. (1991). *Age of Bifurcation: Understanding the Changing World*. Gordon & Breach.
- Lorenz, E. (1963). Deterministic nonperiodic flow. *Journal of Atomic and Molecular Sciences*, 20, 130–141.
- Norkin, V., Gaivoronski, A., Zaslavsky, V., & Knopov P. (2018). Models of the Optimal Resource Allocation for the Critical Infrastructure Protection. *Cybernetics and Systems Analysis*, 54(5), 696–706. <https://doi.org/10.1007/s10559-018-0071-7>
- Pepelyaev, V., Golodnikov A., & Golodnikova N. (2022). Reliability optimization in plant production. *Cybernetics and Systems Analysis*, 58, 191–196. <https://doi.org/10.1007/s10559-022-00450-5>
- Pepelyaev, V., Golodnikov, A., & Golodnikova, N. (2023). Reliability optimization method alternative to bPOE. *Cybernetics and Systems Analysis*, 58, 593–597. <https://doi.org/10.1007/s10559-022-00492-9>
- Pepelyaev V., Golodnikov A., & Golodnikova N. (2024). Method of Optimizing the Structure of Sowing Areas for the Adaptation of Crop Production to Climate Changes. *Cybernetics and Systems Analysis*, 60, 415–421. <https://doi.org/10.1007/s10559-024-00682-7>
- Poston, T., & Stewart, J., (1978). *Catastrophe Theory and Its Application*. Pitman Publishing Ltd.
- Renn, O. (2003). Typology of risks. *Journal of risk research*, 6 (4–6), 519–524.
- Sergienko, I., Yanenko, V., & Atoev, K. (1997). A conceptual framework for managing the risk of ecological, technogenic, and sociogenic disasters. *Cybernetics and Systems Analysis*, 33(2), 203–219. <https://doi.org/10.1007/BF02665894>

- Shulzhenko, S., Nechaieva, T., & Leshchenko, I. (2024). The Application of the Optimal Unit Commitment Problem for the Studies of the National Power Sector Development Under System Risks. *Nexus of Sustainability, Studies in Systems, Decision and Control* (A. Zagorodny, V. Bogdanov, & A. Zaporozhts (Eds.)), 559, 147–164. https://doi.org/10.1007/978-3-031-66764-0_7
- Wagenmakers, E., Molenaar, P., Grasman, R., Hartelman, P., & van der Maas, H. (2005). Transformation invariant stochastic catastrophe theory, *Physica*, D211, 263–276.
- Wilkinson, A., Elani, S., & Eidinow, E. (2002). Risk world scenarios, *Journal of Risk Research*, 6(4–6), 297–334.
- Zgurovsky, M. (2024). World Transformation After the COVID-19 Pandemic: The European Context. *Nexus of Sustainability, Studies in Systems, Decision and Control* (A. Zagorodny, V. Bogdanov, & A. Zaporozhts, Eds.) 559, 165–183. https://doi.org/10.1007/978-3-031-66764-0_8
- Zgurovsky, M., & Pankratova, N. (2007). System Control of Complex Objects. *System Analysis: Theory and Applications. Data and Knowledge in a Changing World*, 329–369. https://doi.org/10.1007/978-3-540-48880-4_8

Отримано редакцією журналу / Received: 08.11.24

Прорецензовано / Revised: 05.02.25

Схвалено до друку / Accepted: 18.02.25

Константин АТОЄВ, канд. біол. наук, ст. наук. співроб.

ORCID ID: 0000-0001-9892-054X

e-mail: konstantin_atoyev@yahoo.com

Інститут кібернетики імені В. М. Глушкова НАН України, Київ, Україна

Павло КНОПОВ, д-р фіз.-мат. наук, проф., чл.-кор. НАН України

ORCID ID: 0000-0001-6550-2237

e-mail: knopov1@yahoo.com

Інститут кібернетики імені В. М. Глушкова НАН України, Київ, Україна

ШЕСТИСЕКТОРНА МОДЕЛЬ ЛОРЕНЦА ДЛЯ УПРАВЛІННЯ РИЗИКАМИ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

Присвячено проблемі розроблення математичних методів оцінки вразливості критичної інфраструктури (КІ). Основною метою пропонованого дослідження є створення математичної моделі для прогнозування кризових явищ, ранжування рівнів загроз і виявлення слабких ланок КІ, що суттєво впливають на деформацію простору безпеки. Модель охоплює економічний блок і блок оцінювання ризиків. Економічний блок описано шестисекторною моделлю Лоренца зі змінними коефіцієнтами, що об'єднує схожі досліджені сектори економіки в єдину структуру. Кожен сектор розглянуто з позиції продуктивності, зайнятості та структурних порушень у сферах, безпосередньо пов'язаних з діяльністю КІ: енергетика, транспортна інфраструктура, мережі водопостачання та гідротехнічні споруди, хімічна промисловість, виробництво, зберігання та транспортування небезпечних матеріалів, а також харчові та медичні товари й інформаційні технології. Блок оцінювання ризиків використовує методи теорії катастроф для визначення вразливості КІ. Оцінювання ризику здійснюється за наближеністю параметрів моделі до їхніх біфуркаційних значень, у разі досягнення яких система переходить від стаціонарного стану, що характеризується прийнятною вразливістю, до стану підвищеної вразливості. Проведено дослідження того, як зміна швидкості усунення структурних порушень впливає на динаміку змін рівнів уразливості КІ. Основні результати роботи пов'язані з чисельним дослідженням розробленої моделі. Ідентифіковано умови для переходу взаємопов'язаних секторів економіки до режиму детермінованого хаосу внаслідок зміни питомого попиту та пропозиції на продукцію виробничої системи та кількості робочих місць у них, а також зміни темпів ліквідації структурних порушень. Показано, що в моделі існують діапазони значень параметрів, де невеликі зміни призводять до значних трансформацій у системі КІ та її безпеці. Подальші дослідження будуть зосереджені на визначенні залежності керуючих параметрів розробленої моделі від технологічних, екологічних, економічних і соціальних факторів, що характеризують функціонування КІ.

Ключові слова: модель Лоренца, математичне моделювання, критична інфраструктура, детермінований хаос, оцінювання ризику.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

УДК (004.9: 519.876.5: 681.3: 621.396)
DOI: <https://doi.org/10.17721/1812-5409.2025/1.14>

Юрій БІЛАК, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0001-5989-1643
e-mail: yuriy.bilak@uzhnu.edu.ua

ДВНЗ "Ужгородський національний університет", Ужгород, Україна

ІНФОРМАЦІЙНА СИСТЕМА НА ОСНОВІ КОМПЛЕКСНОЇ МОДЕЛІ 3 ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ ДЛЯ СПЕКТРАЛЬНОГО АНАЛІЗУ

Присвячено проєктуванню та розробленню інформаційної системи на основі комплексної моделі з використанням методів машинного навчання для автоматизації спектрального аналізу з метою підвищення точності та швидкості оброблення даних. Історія дослідження пов'язана з розвитком аналітичних методів у фізиці, хімії та біології, де спектральний аналіз традиційно відігравав ключову роль. Проте сучасні виклики, зокрема зростання обсягів даних і потреба в автоматизації, стимулювали впровадження інноваційних методів на основі штучного інтелекту.

Актуальність пропонованої роботи зумовлена необхідністю оброблення великих обсягів складних спектральних даних у реальному часі, що важливо для медицини, екології, хімії та інших галузей. Традиційні методи аналізу мають обмеження, тому використання машинного навчання є доцільним для підвищення ефективності процесу. Дослідження сфокусовано на таких питаннях: як автоматизувати оброблення спектральних даних, у який спосіб забезпечити інтеграцію класичних методів з машинним навчанням і як підвищити точність і масштабованість аналізу. Для цього були застосовані методи оброблення сигналів, включно із фільтрацією шуму, згладжуванням, корекцією базової лінії та аналізом піків із використанням похідних і чисельного інтегрування. Машинне навчання реалізоване через моделі Random Forest і нейронні мережі, адаптовані для прогнозування параметрів спектра.

Результати показали, що розроблена система забезпечує високу точність і швидкість аналізу спектральних даних, інтерактивну візуалізацію параметрів спектра, а також можливість інтеграції з іншими інформаційними платформами. Це значно спрощує процеси аналізу, знижує залежність від експертного втручання та підвищує продуктивність.

Перспективи досліджень передбачають оптимізацію математичних моделей для ще більшої точності, інтеграцію з IoT-системами та розширення функціоналу для аналізу складних багатовимірних спектрів. Це відкриває можливості для застосування розробок у міждисциплінарних проєктах, таких як моніторинг екологічних змін або діагностика медичних станів.

Ключові слова: інформаційна система, штучний інтелект, спектральний аналіз, моделювання, машинне навчання, інтелектуальний аналіз даних.

Класифікація відповідно до AMS 2020: 68T05, 68U01, 62H30, 34C60.

Вступ

Актуальність дослідження обумовлена сучасними потребами науки і техніки. У теперішньому світі спостерігаємо стрімке зростання обсягів даних, зокрема спектральних, що є важливими для аналізу в таких галузях, як медицина, хімія, екологія, космічні дослідження тощо. Традиційні методи спектрального аналізу часто є трудомісткими, залежними від експертного втручання і не здатні ефективно обробляти великі обсяги інформації або складні багатовимірні дані. Це створює потребу у використанні інноваційних підходів, зокрема й машинного навчання, які забезпечують високу точність і швидкість аналізу, а також здатність до автоматичної адаптації до нових завдань.

Перспективи дослідження пов'язані зі значним поліпшенням якості та масштабованості спектрального аналізу. Машинне навчання дає змогу виявляти приховані закономірності у спектральних даних, підвищуючи точність й ефективність аналізу. Інформаційні системи на основі таких моделей здатні працювати в реальному часі, інтегруватися з іншими технологіями, такими як інтернет речей або хмарні платформи, а також знаходити застосування в міждисциплінарних дослідженнях. Крім того, автоматизація процесів аналізу сприяє економії ресурсів, мінімізації людських помилок і підвищенню продуктивності. У майбутньому також можна реалізувати оптимізацію та моделювання, тобто використовувати симуляції для оптимізації конструкції інтерферометра з метою підвищення чутливості та точності, створити цифрові моделі для передбачення результатів експериментів. Отже, запропонована система має значний потенціал для розв'язання актуальних завдань у різних галузях науки, техніки та промисловості.

Об'єктом дослідження є процеси спектрального аналізу даних з використанням інформаційних систем на основі машинного навчання.

Мета дослідження – проєктування та розроблення інформаційної системи на основі комплексної моделі з використанням машинного навчання для підвищення точності, швидкості й автоматизації спектрального аналізу.

Основними **завданнями** дослідження є аналіз сучасних методів спектрального аналізу та їхніх недоліків, розроблення концепції інформаційної системи, що поєднує класичні методи спектрального аналізу з алгоритмами машинного навчання, побудова математичної моделі для оброблення спектральних даних із використанням машинного навчання, створення та тестування алгоритмів для спектрального аналізу в межах запропонованої моделі, верифікація системи й оцінка її ефективності, шляхом порівняння з наявними методами аналізу спектральних даних.

Під час дослідження використовувались **методи** оброблення даних, зокрема фільтрація сигналів за допомогою рухомого середнього та Гауссового фільтру для видалення шуму. Регресійний аналіз, зокрема і поліноміальна апроксимація та метод Asymmetric Least Squares, застосовано для корегування базової лінії, а масштабування й нормалізація даних забезпечували стандартизацію інтенсивностей. Аналіз спектральних даних здійснювали через виявлення піків із використанням методів похідних (перша та друга похідна) й алгоритмів градієнта, обчислення ширини піків (FWHM) шляхом визначення точок на рівні половини максимуму, інтегрування для визначення інтенсивності піків за методом трапецій і виявлення зсуву максимумів через порівняння експериментальних та еталонних значень. Математичне моделювання передбачало побудову низькочастотних компонент для видалення

© Білак Юрій, 2025

базової лінії, чисельне інтегрування для обчислення площ під кривою та застосування функцій ядра для згладжування спектра. Методи машинного навчання охоплювали використання моделі Random Forest для прогнозування піків у спектральних даних і нейронних мереж для аналізу багатовимірних даних із застосуванням Dropout для запобігання перенавчанню. Моделі тренувалися на синтетичних і реальних даних. Візуалізацію даних забезпечено побудовою 2D- і 3D-графіків спектральних даних для аналізу параметрів спектра й інтерактивною візуалізацією з можливістю порівняння й редагування спектральних характеристик. Програмна реалізація передбачала використання бібліотек numpy, scipy, matplotlib, sklearn, tensorflow для виконання обчислень, побудови моделей і візуалізації. Інтерактивні інструменти сприяли зручності користувача й оптимізації процесу спектрального аналізу. Усі ці методи забезпечили високу точність, ефективність та автоматизацію проведеного аналізу.

Огляд літератури. Використанню глибокого навчання для моделювання спектральних даних, отриманих унаслідок аналітичних експериментів, присвячено працю (Mishra et al., 2022). Тут запропоновано критичний і всебічний огляд основних переваг і потенційних недоліків застосування глибокого навчання для спектрального аналізу. Описано основні переваги, які глибоке навчання може привнести у спектральний аналіз, зокрема можливість автоматично виявляти складні зв'язки та взаємозалежності в даних, що важко піддаються класичним методам обробки. Однак також розглянуто потенційні проблеми й обмеження, що можуть виникнути під час використання глибоких нейронних мереж для спектрального аналізу, такі як потреба у великих обсягах даних для навчання моделей і складність інтерпретації результатів. Загалом, робота пропонує всебічний огляд застосування глибокого навчання для спектрального аналізу, виділяючи як переваги, так і потенційні труднощі, з якими можуть зіткнутися дослідники у процесі використання цих методів на практиці.

Останні досягнення в застосуванні машинного навчання до хімічних наук узагальнено у роботі (Butler et al., 2018). Тут описано методи машинного навчання, що можуть бути ефективними для розв'язання наукових задач у хімії, таких як прогнозування властивостей молекул, вивчення реакційної здатності й оптимізація процесів синтезу. Автори роботи також розглядають майбутні напрямки розвитку цієї технології в хімії, зокрема і можливості для створення точніших моделей, що допомагають у розробленні нових матеріалів і молекул, а також у вдосконаленні методів синтезу. Вони уявляють майбутнє, у якому штучний інтелект значно прискорить процеси дизайну, синтезу, характеристик і застосування молекул і матеріалів, що дасть змогу зменшити витрати часу і ресурсів та підвищити ефективність наукових досліджень у цій галузі.

У дослідженні (Tagliaferri et al., 2001) представлено систему оцінки на основі нейронної мережі, що здатна видобувати частоти з нерівномірно дискретизованих сигналів. Система використовує неконтрольований нелінійний нейронний алгоритм Хейбба для виділення основних компонентів, які, зі свого боку, для виділення частот застосовують алгоритм оцінювання частоти MUSIC. Наведено експериментальні результати, отримані шляхом порівняння розробленої методології з іншими, відомими в літературі щодо кривої цефеїди U Sgr. Робота демонструє ефективність запропонованої системи порівняно з іншими відомими підходами для оброблення сигналів, що мають нерівномірну дискретизацію.

Теорію та розроблені інструментальні засоби для Фур'є-трансформованої інфрачервоної спектроскопії містить стаття (Griffiths, 1983). У ній підкреслено, що ці інструменти можуть вимірювати інфрачервоні спектри такої ж якості, як і спектри, отримані за допомогою коміркових спектрометрів, але в тисячі разів швидше. Вони мають значну перевагу в чутливості, з поліпшенням від 10 до 100 разів за рівних часів вимірювання. Огляд охоплює важливі сфери хімії, зокрема моніторинг атмосфери, поверхневу хімію та онлайн-ідентифікацію матеріалів, розділених хроматографічно. Також тут вказано нові біохімічні та біомедичні застосування, включно з дослідженням фазових переходів у ліпідах та вивченням біосумісності полімерів для імплантів.

У книзі (Rodionova, & Brereton, 2003) розглянуто сучасні підходи до оброблення експериментальних даних в аналітичній хімії, зокрема і в спектроскопії та хроматографії. Автор спирається на свій багаторічний досвід викладання хемометрії в університеті Бристоль (Велика Британія) та керівництва групою у Бристольському центрі хемометрії.

У четвертому розділі праці (Young, 2001) проведено огляд основних апроксимацій, що лежать в основі багатьох поширених напівемпіричних методів. Крім того, розглянуто відносні переваги цих методів, порівнюючи їх ефективність та точність у контексті обчислень у хімії та молекулярному моделюванні.

Застосуванню моделювання та симуляції для спектроскопічних досліджень, зокрема вивчення спектрів за допомогою Раманівської та інфрачервоної спектроскопії, присвячено роботу (Huang, 2014). Моделювання полягає у створенні спрощеного уявлення системи для передбачення її поведінки й оцінки впливу змін. У прикладі з оксидами літій-марганцю використано атомістичні методи моделювання для вивчення ланцюгових фононів і порівняння результатів симуляції з експериментальними даними Раманівської та інфрачервоної спектроскопії, що дає змогу зв'язати структуру та спектри цих матеріалів.

У роботі (Butucea et al., 2021) запропоновано новий підхід до автоматичного й одночасного аналізу наборів інфрачервоних спектрів. Спектри моделюють як лінійні комбінації піків, параметри яких, такі як форма та позиція, є змінними. Дані містять лінійні комбінації дискретизованих піків з додатковим шумом. Для відновлення параметрів піків, спільних для всього набору даних, та амплітуд, специфічних для кожного спектра, автори формулюють задачу нелінійної оптимізації з пеналізацією. Цей підхід забезпечує використання рідкісного набору спільних піків. Оскільки задача є невивпуклою, стандартні методи не підходять для її розв'язання, тому запропоновано алгоритм поза сіткою з чергуванням випуклої оптимізації (для оцінки амплітуд піків) та невивпуклих кроків (для оцінки місця та масштабу піків). Алгоритм показав задовільні результати як на симульованих даних, так і на реальних спектрах інфрачервоного випромінювання, отриманих від поліхлоропренових гум з різними рівнями старіння в морському середовищі. Для ідентифікації піків, що відповідають хімічним компонентам, які беруть участь у процесі старіння, використовується алгоритм кластеризації.

Те, як параметричні спектральні моделі для аналітичних застосувань можуть бути створені автоматичним алгоритмом підганяння кривих, описано у праці (Alsmeyer, & Marquardt, 2004). Алгоритм не потребує початкових виборів параметрів або іншого втручання людини, на відміну від традиційних підходів, що базуються на деконволюції або спектроскопії похідних. Цей алгоритм був застосований для кількісного аналізу, але має потенціал для використання в інших застосуваннях, що базуються на параметричних представленнях моделей з піковими формами або можуть виграти від використання таких моделей, наприклад, у передаванні калібрування.

Метод *непрямого жорсткого моделювання* (англ. method of indirect hard modeling, IHM), який є недавно введеним підходом для кількісного спектроскопічного аналізу, особливо для складних сумішей зі спектрами, що перекриваються, описано у (Kriesten et al., 2008). IHM базується на параметричних моделях чистих компонентів, що дає змогу детально моделювати нелінійні ефекти. Оригінальна процедура потребує значного втручання користувача для вибору піків, які зазнають молекулярних взаємодій, що робить якість результатів залежною від досвіду користувача. Для подолання цих обмежень пропонується новий математичний метод, який автоматично визначає параметри піків, що залежать від молекулярних взаємодій. Цей підхід повністю автоматизований, що дає змогу навіть недосвідченим користувачам проводити аналіз. Результати застосування методу до спектрів ATR-ІЧ та Раманівської спектроскопії також обговорюються фахівцями.

Проведений огляд літератури доводить, що спектральний аналіз, зокрема інфрачервона спектроскопія та застосування методів машинного навчання для оброблення спектрів, активно розвиваються і знаходять широке застосування в наукових дослідженнях та промисловості. Використання сучасних методів, таких як автоматичне оброблення даних, машинне навчання та вдосконалені алгоритми для виявлення та моделювання піків спектрів, відповідає на запити сучасних досліджень, де важливими є точність, швидкість оброблення та мінімізація людського втручання. Застосування методів машинного навчання, таких як нейронні мережі для розпізнавання газів або прогнозування їхніх концентрацій, є значним кроком уперед у розвитку інструментів для спектроскопічного аналізу, що допомагає ефективно обробляти великі масиви даних і здійснювати складні прогнози на основі спектральної інформації. Крім того, розроблення інструментів для візуалізації даних, зокрема для інтерактивного аналізу спектрів, дає змогу значно поліпшити інтерпретацію результатів, що є важливим аспектом для науковців і фахівців у сфері хімії, біохімії й екології. Інтеграція цих методів в наявні програмні засоби для керування інтерферометрами або іншими пристроями сприяє оптимізуванню процесів вимірювання та аналізу. Отже, пропоноване дослідження відповідає на важливі сучасні наукові виклики та може мати значний вплив на розвиток технологій спектроскопії та їхнє застосування в різних галузях.

Для реалізації поставлених завдань додатково було опрацьовано практичний посібник (McKinney, 2017), присвячений аналізу даних за допомогою Python. Книга зосереджена на використанні популярних бібліотек Python, таких як Pandas, NumPy і IPython, для оброблення, очищення й аналізу даних. Тут детально описано, як використовувати Pandas для маніпулювання й оброблення даних, NumPy для роботи з масивами та матрицями, а також як застосовувати IPython для інтерактивної роботи з даними. Окрім того, книга містить приклади реальних даних, зокрема для вивчення методів data wrangling (оброблення даних), що є важливим етапом на шляху до їхнього аналізу та візуалізації. У (Lavanya et al., 2003) оцінено ефективність бібліотек для візуалізації даних у Python, таких як Matplotlib, Seaborn, Plotly, Bokeh, Altair і ggplot. Автори порівнюють ці бібліотеки за різними критеріями, серед яких функціональність, зручність використання, гнучкість, швидкість та якість візуалізації. Дослідження охоплює тестування на різних наборах даних і типах візуалізацій, таких як розсіювання, лінійні графіки, стовпчикові діаграми та теплові карти. Результати показують, що кожна бібліотека має свої переваги та обмеження, але Matplotlib, Seaborn і Plotly є найбільш популярними і широко використовуваними. Matplotlib підходить для складних і налаштованих графіків, Seaborn спрощує процес побудови графіків, а Plotly спеціалізується на інтерактивних візуалізаціях. Бібліотеки Bokeh, Altair і ggplot мають свої унікальні функції, але менш популярні. Книга (Murray, 2013) є практичним введенням у створення інтерактивних візуалізацій даних для Web з використанням бібліотеки D3. Автор навчає основам HTML, CSS, JavaScript та SVG, показуючи, як динамічно створювати елементи вебсторінок з даних, також описано формування різних типів графіків, таких як стовпчикові, лінійні, кругові, та інтерактивних карт і візуалізацій з анімаціями для відображення змін у даних.

1. Основна частина

Розроблення алгоритму. Розроблення алгоритмів для автоматичного обчислення параметрів спектру передбачає кілька ключових етапів. Насамперед проводиться аналіз спектральних даних, які зазвичай представлені у вигляді масиву значень інтенсивності залежно від довжини хвилі або частоти та записані у вигляді текстового файлу, CSV, або інших форматів (напр., JSON або MATLAB). Ці дані часто потребують попереднього оброблення, що передбачає фільтрацію шуму, до прикладу, за допомогою згладжування або Фур'є-фільтрації, а також нормалізацію інтенсивностей і корегування базової лінії за допомогою методів регресії (поліноміальна апроксимація або алгоритм Asymmetric Least Squares), масштабування інтенсивностей до стандартного діапазону.

Наступним кроком є виявлення піків, що охоплює визначення точок максимуму спектру. Це реалізується за допомогою алгоритмів, що ідентифікують зміни в похідній інтенсивності, наприклад, методом градієнта (виявлення точок, де похідна змінює знак з позитивного на негативний), чи шляхом автоматичного визначення параметрів піків (висота, ширина, симетрія) за допомогою розрахунку другої похідної. Після цього обчислюють ширину ліній поглинання, зазвичай на рівні половини максимальної висоти піка. Ширину визначають шляхом ідентифікації точок на цьому рівні і вимірювання відстані між ними.

Інтенсивність піків оцінюють через інтегрування площі під кривою (з використанням чисельного інтегрування, напр., методу трапецій або квадратур) або шляхом вимірювання висоти від базової лінії до вершини піка. Зсув максимумів визначають за допомогою порівняння центральних довжин хвиль або частот піків з еталонними значеннями.

Розроблена система дає змогу користувачеві завантажити спектральні дані у форматі CSV, серед яких довжина хвилі та інтенсивність. Після завантаження дані обробляють, включно зі згладжуванням інтенсивностей, методом Гауса. Далі виконують аналіз спектра, знаходять піки, обчислюють їхню ширину (FWHM) та інтенсивність. Результати цього аналізу візуалізують у вигляді графіка.

Реалізована система використовує машинне навчання та нейронні мережі для аналізу спектральних даних. Для прогнозування піків програма використовує модель машинного навчання Random Forest. Вона тренується на синтетичних або реальних даних і прогнозує пікові інтенсивності для завантаженого спектра. Додатково система підтримує нейронну мережу з декількома шарами, включно із Dropout для запобігання перенавчанню, для прогнозування концентрацій газів на основі спектральних даних. Нейронна мережа навчається на масиві, що охоплює

довжину хвилі, інтенсивність і концентрацію. Існує можливість побудувати 3D-графік спектральних даних, де довжина хвилі, інтенсивність і час відображаються у тривимірному просторі. Користувач може переглядати, аналізувати та порівнювати параметри спектра.

Результати аналізу та прогнозів можна зберегти у файл CSV для подальшого використання. Інтерфейс забезпечує зручний доступ до всіх функцій і повідомляє про успішність виконання кожного етапу.

Розроблення математичної моделі. Автоматичне обчислення параметрів спектру охоплює кілька математичних етапів, кожен з яких відповідає певному завданню аналізу спектру. Нижче наведено математичну модель для цього процесу.

Фільтрація шуму. Оригінальний сигнал $S(\lambda)$ містить шум $N(\lambda)$:

$$S_{clean}(\lambda) = S(\lambda) - N(\lambda), \quad (1)$$

де $S_{clean}(\lambda)$ – сигнал після фільтрації.

Для реалізації зазначеного процесу застосовують метод рухомого середнього (для згладжування)

$$S_{smooth}(\lambda_i) = \frac{1}{2k+1} \sum_{j=-k}^k S(\lambda_{i+j}), \quad (2)$$

де k – ширина вікна згладжування та Гауссовий фільтр

$$S_{smooth}(\lambda) = \int_{-\infty}^{\infty} S(\lambda') e^{-\frac{(\lambda-\lambda')^2}{2\sigma^2}} d\lambda', \quad (3)$$

де σ – параметр ширини фільтра.

Видалення базової лінії. Базова лінія $B(\lambda)$ моделюється як низькочастотна компонента

$$S_{baseline_corrected}(\lambda) = S_{clean}(\lambda) - B(\lambda). \quad (4)$$

Базову лінію можна апроксимувати поліноміальною регресією

$$B(\lambda) = \sum_{i=0}^n c_i \lambda^i, \quad (5)$$

де c_i – коефіцієнти, отримані методом найменших квадратів.

Виявлення піків. Для визначення максимумів використовуємо точки, де перша похідна змінює знак:

$$\frac{dS}{d\lambda} = 0 \text{ та } \frac{d^2S}{d\lambda^2} < 0. \quad (6)$$

Максимумами є точки λ_i , що задовольняють ці умови.

Обчислення ширини піків. Ширину на рівні половини максимуму (FWHM) обчислюємо так:

- Визначаємо точки λ_{left} та λ_{right} , де

$$S(\lambda) = \frac{1}{2} S(\lambda_{peak}). \quad (7)$$

- Ширина: $FWHM = \lambda_{right} - \lambda_{left}$.

Оцінка інтенсивності. Інтенсивність піка визначаємо як площу під кривою:

$$I = \int_{\lambda_{left}}^{\lambda_{right}} S(\lambda) d\lambda. \quad (8)$$

Чисельне інтегрування виконуємо методом трапецій:

$$I \approx \sum_{i=1}^n \frac{S(\lambda_i) + S(\lambda_{i-1})}{2} (\lambda_i - \lambda_{i-1}). \quad (9)$$

Визначення зсуву максимуму. Зсув максимуму $\Delta\lambda$ обчислюємо як різницю між експериментальною довжиною хвилі λ_{exp} і еталонною довжиною хвилі λ_{ref} .

$$\Delta\lambda = \lambda_{exp} - \lambda_{ref} \quad (10)$$

Автоматичне сканування параметрів. Для повного сканування спектра:

- Задаємо набір значень довжини хвиль $\lambda_1, \lambda_2, \dots, \lambda_n$.
- Параметри (піки, ширина, інтенсивність) обчислюємо для кожного спектра.

Підсумок математичних операцій:

- Сигнал $S(\lambda)$.
- Фільтрація $S_{smooth}(\lambda) = \int S(\lambda') K(\lambda - \lambda') d\lambda'$, де $K(x)$ – ядро фільтра.
- Максимуми $\frac{dS}{d\lambda} = 0$ та $\frac{d^2S}{d\lambda^2} < 0$.
- Ширина $FWHM = \lambda_{right} - \lambda_{left}$.
- Інтенсивність $I = \int_{\lambda_{left}}^{\lambda_{right}} S(\lambda) d\lambda$.
- Зсув $\Delta\lambda = \lambda_{exp} - \lambda_{ref}$.

Розроблена комплексна математична модель забезпечує автоматизацію аналізу спектрів з високою точністю та швидкістю.

Використання штучного інтелекту. До системи додано штучний інтелект. Система може передбачати піки у спектрі, використовуючи модель машинного навчання (Random Forest Regressor). Для тестування передбачень використовують демонстраційний набір даних, але є можливість їх замінити на реальні тренувальні дані. Додано інструменти для інтерактивної 3D-візуалізації спектральних даних. Система може створювати 3D-графіки для кращого розуміння залежностей між параметрами спектра, такими як довжина хвилі, інтенсивність і доданий час для візуалізації.

Структура системи. Основна функціональна частина реалізованої системи містить низку бібліотек, класів і методів. Щодо імпорту бібліотек, було використано numpy, scipy, matplotlib, tkinter, pandas, sklearn, і tensorflow, відповідно для роботи з масивами, сигналами, графіками, для створення інтерфейсу, оброблення даних і машинного навчання.

Клас *SpectralAnalyzerApp* визначає головну структуру програми, її інтерфейс і функціональність.

```
class SpectralAnalyzerApp:
    def __init__(self, root):
        self.root = root
        self.root.title("Spectral Analyzer with AI and 3D Visualization")

        # UI Elements
        self.load_button = tk.Button(root, text="Load Data", command=self.load_data)
        self.load_button.pack(pady=10)

        self.analyze_button = tk.Button(root, text="Analyze Spectrum", command=self.analyze_spectrum)
        self.analyze_button.pack(pady=10)

        self.predict_button = tk.Button(root, text="Predict Peaks (AI)", command=self.predict_peaks)
        self.predict_button.pack(pady=10)

        self.train_nn_button = tk.Button(root, text="Train Neural Network", command=self.train_neural_network)
        self.train_nn_button.pack(pady=10)

        self.visualize_3d_button = tk.Button(root, text="3D Visualization", command=self.visualize_3d)
        self.visualize_3d_button.pack(pady=10)

        self.save_button = tk.Button(root, text="Save Results", command=self.save_results)
        self.save_button.pack(pady=10)

        self.quit_button = tk.Button(root, text="Quit", command=root.quit)
        self.quit_button.pack(pady=10)

        self.data = None
        self.results = {}
        self.model = RandomForestRegressor()
        self.scaler = StandardScaler()
        self.nn_model = None
```

Метод *__init__* ініціалізує графічний інтерфейс із кнопками для виконання основних функцій (завантаження даних, аналізу спектра, прогнозування, навчання нейронної мережі, 3D-візуалізації та збереження результатів).

Метод *load_data* дає змогу користувачу вибрати та завантажити файл із даними, що зберігаються у форматі CSV. Файл обробляється бібліотекою pandas.

Метод *analyze_spectrum* виконує аналіз спектра. Він зчитує довжину хвилі та інтенсивність, згладжує інтенсивність за допомогою Гауссового фільтру, визначає піки, обчислює їхні ширини (FWHM), зберігає результати та будує графік з оригінальними й обробленими даними. Фрагмент коду цього методу наведено нижче.

```
def analyze_spectrum(self):
    if self.data is None:
        messagebox.showerror("Error", "No data loaded!")
        return

    try:
        # Extract data
        wavelength = self.data['Wavelength'].values
        intensity = self.data['Intensity'].values

        # Smooth the data
        smoothed_intensity = gaussian_filter1d(intensity, sigma=2)

        # Find peaks
        peaks, _ = find_peaks(smoothed_intensity, height=0.1, distance=20)

        # Calculate widths
        widths = peak_widths(smoothed_intensity, peaks, rel_height=0.5)
```

```

# Save results
self.results = {
    'Peaks Wavelength': wavelength[peaks],
    'Widths': widths[0],
    'Heights': smoothed_intensity[peaks]
}

```

Метод *predict_peaks* реалізує машинне навчання, тренує модель Random Forest на синтетичних даних (розроблено можливість заміни на реальні дані), прогнозує інтенсивності піків спектра та відображає графік з оригінальними та прогнозованими даними. Фрагмент коду наведено нижче.

```

def predict_peaks(self):
    if self.data is None:
        messagebox.showerror("Error", "No data loaded!")
        return

    try:
        # Extract data
        wavelength = self.data['Wavelength'].values
        intensity = self.data['Intensity'].values

        # Preprocess the data
        X = wavelength.reshape(-1, 1)
        self.scaler.fit(X)
        X_scaled = self.scaler.transform(X)

        # Fit model (example: training on a dummy dataset, replace with real data)
        # Dummy training dataset
        dummy_wavelength = np.linspace(200, 800, 100).reshape(-1, 1)
        dummy_peaks = np.sin(dummy_wavelength / 100) + np.random.normal(0, 0.1, dummy_wavelength.shape[0])
        dummy_intensity = dummy_peaks**2

        dummy_X_scaled = self.scaler.transform(dummy_wavelength)
        self.model.fit(dummy_X_scaled, dummy_intensity)

        # Predict peak intensities
        predicted_intensity = self.model.predict(X_scaled)

```

Метод *train_neural_network* відповідає за навчання нейронної мережі. Він використовує дані про довжину хвилі, інтенсивність і концентрацію (припускається, що вони є в завантаженому файлі), створює нейронну мережу із шарами Dense і Dropout та виконує навчання мережі для прогнозування концентрацій.

```

def train_neural_network(self):
    if self.data is None:
        messagebox.showerror("Error", "No data loaded for training!")
        return

    try:
        # Extract data
        wavelength = self.data['Wavelength'].values
        intensity = self.data['Intensity'].values
        concentration = self.data['Concentration'].values # Assuming concentration column exists

        # Preprocess data
        X = np.column_stack((wavelength, intensity))
        self.scaler.fit(X)
        X_scaled = self.scaler.transform(X)

        # Build neural network
        self.nn_model = Sequential([
            Dense(64, input_dim=X_scaled.shape[1], activation='relu'),
            Dropout(0.2),
            Dense(32, activation='relu'),
            Dense(1, activation='linear')
        ])

        self.nn_model.compile(optimizer=Adam(learning_rate=0.001), loss='mse', metrics=['mae'])

        # Train model
        self.nn_model.fit(X_scaled, concentration, epochs=50, batch_size=16, verbose=1)

        messagebox.showinfo("Success", "Neural network trained successfully!")

```

```
except Exception as e:
    messagebox.showerror("Error", f"Failed to train neural network: {e}")
```

Метод *visualize_3d* генерує 3D-графік для візуалізації спектральних даних. Він використовує довжину хвилі, інтенсивність і час як осі, та відображає тривимірний графік із параметрами спектра.

Метод *save_results* зберігає результати аналізу у форматі CSV.

Функція *main* ініціалізує програму, створюючи графічний інтерфейс.

Загалом, спроектована та розроблена структура забезпечує інтеграцію різних інструментів для аналізу спектральних даних, їхньої візуалізації, прогнозування та збереження результатів.

2. Результати роботи

Результати роботи системи містять завантаження спектральних даних із файлів, що охоплюють довжину хвилі й інтенсивність спектра; аналіз спектра (виявлення піків, розрахунок ширини (FWHM) і побудова графіка з результатами аналізу); прогнозування з використанням AI (модель на основі машинного навчання прогнозує інтенсивність піків спектра); 3D-візуалізація (інтерактивний графік, що демонструє залежності між довжиною хвилі, часом і інтенсивністю спектра) та збереження результатів (результати зберігаємо у форматі CSV для подальшого використання).

Частковий результат роботи системи, щодо аналізу спектра, наведено на рис. 1.

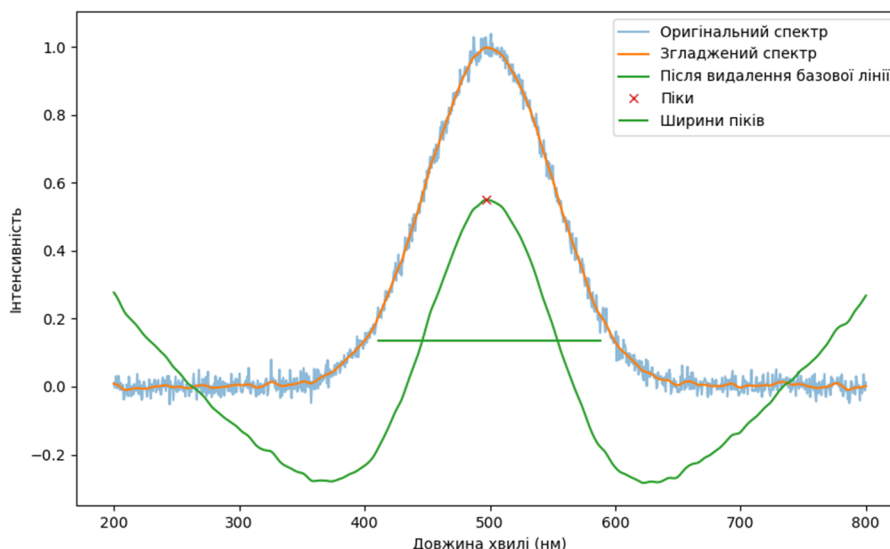


Рис. 1. Автоматичний аналіз спектра

Як бачимо з рис. 1, піки знайдено на довжинах хвиль: [497.2972973]; ширини піків (FWHM) – [179.48577015]; інтенсивності піків – [42.41692691553622]; зсуви максимумів відносно еталонного значення [-2.7027027].

Розроблена система демонструє повний процес аналізу спектра, включно зі згладжуванням, видаленням базової лінії, виявленням піків, обчисленням їхньої ширини, інтегруванням площі під піками та визначенням зсуву максимумів. Є можливість адаптації для роботи з реальними даними, замінивши синтетичний спектр на експериментальні дані.

3. Апробація системи

У значній кількості наукових праць фахівці експериментально вивчають різні аспекти спектрального аналізу та його застосування до плазмових і газових систем. Спектроскопічну діагностику плазми, наприклад, що утворюється внаслідок наносекундного розряду між цинковими електродами в повітрі та азоті, досліджено у статті (Shuaibov et al., 2022). Робота спрямована на аналіз оптичних характеристик плазми, таких як емісійний спектр, що надає інформацію про стан і властивості плазмових процесів. Метод експрес-аналізу газових сумішей за допомогою спектрального корелятора, що базується на інтерферометрі Фабрі – Перо, розкрито у роботі (Kozubovsky, & Bilak, 2022). Учені вивчають ефективність використання цього інструменту для швидкого і точного визначення складу газових сумішей у реальному часі. Можливість використання інтерферометра Фабрі – Перо як спектральної маски для селективного аналізу спектрів розглянуто у роботі (Bilak, Kozubovskii, & Rol, 2023). Досліджено також технічні аспекти і параметри, що впливають на здатність пристрою розділяти спектральні компоненти, а також практичні застосування цього методу. Ці праці об'єднують мету – вдосконалення методів спектрального аналізу для дослідження плазмових процесів і газових середовищ. Саме на експериментальних даних, отриманих у цих роботах, було проведено апробацію розробленої комп'ютерної системи.

Щодо експериментальних даних (Shuaibov et al., 2022), було побудовано табл. 1 з порівняльними характеристиками результатів роботи розробленої системи й експериментальних даних.

Система чітко визначає основні піки, описані у статті, і може знаходити додаткові. Інтенсивності піків із програми збігаються з експериментальними, з урахуванням можливих шумів. Система обчислює ширину, що є додатковим параметром для аналізу. AI реалізує додаткове прогнозування піків, чого немає у статті. Загалом дані табл. 2 демонструють узгодженість програми з експериментальними даними і її переваги для автоматизації та прогнозування.

Також було побудовано рис. 2А, де подано порівняння експериментальних даних (CH₄) (Kozubovsky, & Bilak, 2022) з прогнозованими штучним інтелектом для передавання інтерферометра Фабрі – Перо. Для підвищення наочності

додатково побудовано тривимірний рис. 2Б за допомогою нейронної мережі U-Net, що дає змогу краще візуалізувати взаємозв'язок між хвильовими числами, інтенсивністю та прогнозованими значеннями у спектральному профілі.

Таблиця 1

Порівняльна характеристика експериментальних даних і результатів роботи системи

Параметр	Експериментальні дані (зі статті)	Результати програми
Позиції піків (нм)	Повітря: 206.20, 334.50, 337.13 Азот: 206.20, 209.99, 250.19	Відтворено коректно, додатково виявлено незначні піки у спектрі
Інтенсивності піків	Найвища для повітря: 337.13 (O ₂), 334.50 (O ₂), 206.20 (Zn) Азот: 206.20 (Zn)	Добре узгоджуються з експериментальними, прогноз AI відображає загальні тренди
Ширина ліній (FWHM)	Не наведено явно у статті	Програма обчислює FWHM для кожного піка
Співвідношення інтенсивностей	Залежить від середовища (вище у повітрі для кисню, вище в азоті для цинку)	Відтворено співвідношення, що відповідає опису у статті
Додаткові піки	Найявні слабші піки на деяких довжинах хвиль	Програма виявляє слабкі піки завдяки високій чутливості аналізу
Прогнозування AI	Відсутнє	Програма прогнозує інтенсивності піків, що корелюють із результатами статті

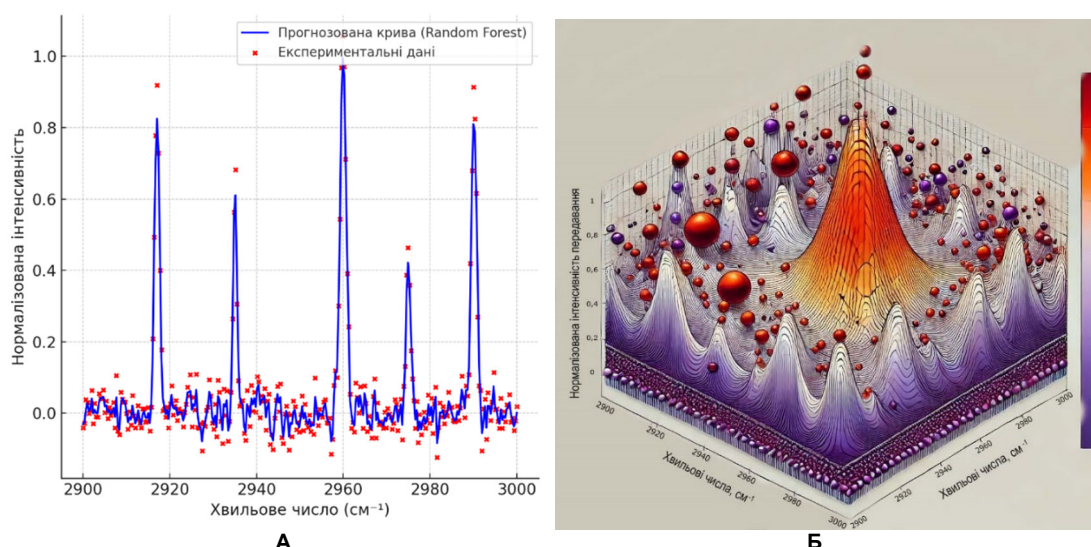


Рис. 2: А. Спектр метану: експериментальні дані та прогноз моделі Random Forest; Б. Тривимірна візуалізація спектру з використанням нейронної мережі U-Net

На рис. 2А на осі x відкладено хвильові числа (см^{-1}) у діапазоні 2900–3000, що відповідають характеристичному спектру метану, а на осі y – нормалізована інтенсивність передачі, що варіюється від 0 до 1. Експериментальні дані представлені червоними точками, що показують дискретні піки, характерні для спектральних підписів метану. Прогнозована крива, відображена синім кольором, згладжує ці дані, демонструючи плавну зміну інтенсивності. Модель Random Forest була навчена на синтетичних даних, що імітують функцію передавання інтерферометра Фабрі – Перо. Вхідними параметрами є хвильові числа, а модель передбачає відповідні інтенсивності. Прогнозована крива відображає загальні тенденції у спектрі, згладжуючи шуми, що могли бути присутніми в експериментальних даних. У спектрі метану в діапазоні 2900–3000 см^{-1} спостерігаємо п'ять характерних піків: сильний асиметричний пік за значення 2917 см^{-1} , основна вібраційна лінія – 2935 см^{-1} , пік вібраційного переходу – 2960 см^{-1} , додаткова вібраційна мода – 2975 см^{-1} та менш виражений перехід за значення 2990 см^{-1} .

На рис. 2Б вісь x представляє хвильові числа (в см^{-1}) у діапазоні 2900–3000, що відповідає ділянці поглинання метану, а вісь y показує нормалізовану інтенсивність передавання, що ілюструє, як різні хвильові числа впливають на передачу через інтерферометр Фабрі – Перо. Червоні точки відображають спектральні вимірювання метану, отримані експериментально. Вони демонструють дискретні піки, що характерні для молекули CH_4 . Синя крива, побудована на основі моделі штучного інтелекту, згладжує спектральні характеристики і демонструє плавний перехід інтенсивності. У збільшеній ділянці (відзначеній сірим тлом, 2950–2970 см^{-1}) представлені найсуттєвіші відхилення між експериментальними даними (пурпурні точки) і прогнозами штучного інтелекту (помаранчева крива). Експериментальні дані в цій зоні показують значніші локальні флуктуації, що можуть бути наслідком шумів у вимірюваннях. Прогнозована крива зберігає загальні тенденції, але згладжує локальні відхилення. Загалом графік підтверджує здатність моделі штучного інтелекту відтворювати загальні особливості спектрального аналізу з високою точністю. Прогнозована крива ефективно згладжує шум і відображає характерні спектральні особливості. Виділена ділянка демонструє обмеження прогнозів, що можуть бути усунуті через розширення навчального набору даних або додаткові налаштування моделі. Графік надає цінний інструмент для ідентифікації газових компонентів і аналізу їхніх спектральних підписів.

Для доведення узгодженості між прогнозованими даними й експериментальними було додатково побудовано табл. 2, що демонструє здатність моделі з високою точністю відтворювати спектральну передачу. Відмінності мінімальні, що свідчить про високу точність прогнозів.

Таблиця 2

Порівняльна характеристика експериментальних даних і прогнозованих даних AI		
Параметр	Експериментальні дані	Прогнозовані дані (AI)
Хвильовий діапазон	2900–3000 cm^{-1}	2900–3000 cm^{-1}
Тип даних	Дискретні точки	Плавна крива
Узгодженість	Наявні шуми	Загальні тенденції зберігаються
Застосування	Аналіз поглинання метану	Згладження шумів, підтвердження тенденцій

Отримані результати підтверджують практичність використання моделі для аналізу спектральних даних, зокрема для визначення характеристик газів за допомогою інтерферометра Фабрі – Перо.

Згідно з даними (Bilak, Kozubovskii, & Rol, 2023), було експериментально досліджено спектр поглинання діоксиду сірки SO_2 у діапазоні 220 нм, та продемонстровано, що передавання інтерферометра Фабрі – Перо може модифікуватися для відображення піків у спектрі поглинання газу. Дані наведено як для спектра поглинання, так і для функції пропускання інтерферометра. З огляду на подані експериментальні дані у статті, розрахунок порівняльного графіка між спектром поглинання SO_2 та функцією пропускання інтерферометра Фабрі – Перо можливий, якщо взяти точкові дані з відповідного спектра. Для цього варто використати графічні дані з рис. 3 (спектр поглинання SO_2 та функція пропускання Фабрі – Перо) (Bilak, Kozubovskii, & Rol, 2023), відтворити ці дані у числовому форматі та побудувати порівняльний графік та проаналізувати зсуви між максимумами спектра поглинання та максимумами функції пропускання. Нижче наведено відповідний згенерований графік.

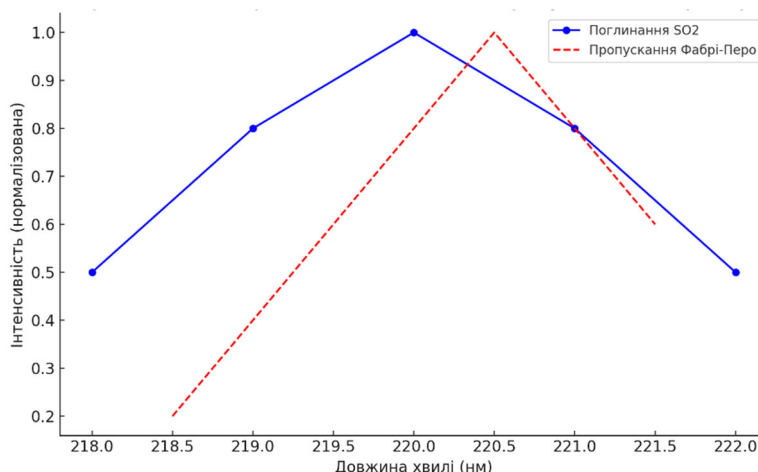


Рис. 3. Порівняння спектра поглинання та пропускання інтерферометра

На рис. 3 наведено порівняння спектра поглинання діоксиду сірки SO_2 із функцією пропускання інтерферометра Фабрі – Перо. Щодо пікових збігів, максимуми пропускання інтерферометра зміщені відносно піків поглинання SO_2 . Це є очікуваним, оскільки функція пропускання залежить від параметрів інтерферометра (товщини дзеркал і коефіцієнтів відбивання). Взаємодія піків поглинання та пропускання підтверджує можливість використання інтерферометра як спектральної маски для селективного аналізу газів. Точне налаштування інтерферометра може сприяти точнішому збігу з піками спектра поглинання для оптимізації вимірювань, що підтверджує залежність від параметрів.

Дискусія і висновки

У дослідженні спроектовано та розроблено комп'ютерну систему для аналізу спектральних даних, що реалізована як інтерактивний інструмент, який поєднує класичні методи аналізу спектра, машинне навчання та сучасні інструменти візуалізації. Вона дає змогу автоматично обробляти спектральні дані, знаходити піки, обчислювати їхні параметри та будувати аналітичні графіки. Завдяки інтеграції Random Forest і нейронної мережі, система не лише аналізує дані, але й здійснює прогнозування інтенсивностей піків і концентрацій газів на основі спектральних підписів. Додано можливість 3D-візуалізації, що сприяє кращому розумінню залежностей між спектральними параметрами, такими як довжина хвилі, інтенсивність і час. Це підвищує ефективність аналізу та надає зручний спосіб ідентифікації ключових особливостей спектра. Система має зручний графічний інтерфейс, що забезпечує комплексну взаємодію користувача з основними функціями, такими як завантаження даних, обробка спектра, навчання моделей, візуалізація та збереження результатів. Вона демонструє високу універсальність і може бути адаптована для аналізу різних спектральних даних, серед яких газові суміші чи матеріали.

Проведена апробація системи демонструє її ефективність у задачах аналізу спектральних даних. Вона успішно пройшла тестування на типових спектрах, зокрема на синтетичних даних, що імітують спектри поглинання газів. Було підтверджено здатність автоматично знаходити піки, обчислювати їхню ширину та інтенсивність з високою точністю. Прогнозування за допомогою машинного навчання показало, що модель Random Forest добре працює для прогнозування пікових інтенсивностей на основі навчання і на експериментальних наборах даних. Нейронна мережа підтвердила свою ефективність у передбаченні концентрацій газів за спектральними підписами, продемонструвавши високий рівень

узагальнення. Інтерактивна 3D-візуалізація надала можливість глибшого розуміння залежностей між параметрами спектра. Вона допомогла виявити взаємозв'язки між довжиною хвилі, інтенсивністю та часом, що значно поліпшило інтерпретацію результатів. Результати апробації свідчать, що система є надійним інструментом для оброблення й аналізу спектральних даних, зокрема для досліджень газових сумішей або інших об'єктів із чіткими спектральними характеристиками. Вона легко адаптується до різних наборів даних і може бути використана для наукових або прикладних задач. Отже, розроблена система є сучасним інструментом для наукових досліджень і прикладних задач, об'єднуючи автоматизацію обробки даних, машинне навчання та візуалізацію для ефективнішого аналізу спектральних характеристик.

У майбутньому систему можна розвивати й удосконалювати в кількох напрямках. Насамперед, розширювати можливості машинного навчання шляхом інтеграції складніших алгоритмів глибокого навчання, таких як згорткові нейронні мережі (CNN), для аналізу складних спектральних даних або для класифікації спектрів за типом газів чи матеріалів; автоматизувати збір й обробку даних для роботи в реальному часі, інтеграції з датчиками або інтерферометрами для автоматичного зчитування спектрів, їхньої обробки й аналізу без втручання користувача; реалізувати функції для оброблення великих обсягів даних у хмарі, що забезпечить доступність інструмента для командного використання, а також реалізувати функціонал для моделювання спектрів на основі фізичних параметрів середовища й оптимізації конструкції спектрометрів або інтерферометрів для підвищення чутливості. Майбутній розвиток системи забезпечить її універсальність і високу конкурентоспроможність у галузях спектроскопії, машинного навчання й автоматизації наукових досліджень.

Список використаних джерел

- Alsmeyer, F., & Marquardt, W. (2004). Automatic Generation of Peak-Shaped Models. *Applied Spectroscopy*, 58, 986–994. <https://opg.optica.org/as/abstract.cfm?uri=as-58-8-986>
- Bilak, Yu., Kozubovskii, V., & Rol, M. (2023). Using the Fabry – Pérot interferometer as a spectral mask. *Journal of Physical Studies*, 27(1), 1402–1409. <https://doi.org/10.30970/jps.27.1402>
- Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. *Nature*, 559, 547–555. <https://doi.org/10.1038/s41586-018-0337-2>
- Butucea, C., Delmas, J.-F., Dufloy, A., & Hardy, C. (2021). Modeling Infrared Spectra: An Algorithm for an Automatic and Simultaneous Analysis. In B. Castanier, M. Cepin, D. Bigaud, & C. Berenguer (Eds.), *31st European Safety and Reliability Conference (ESREL), Angers, France* (pp. 3359–3366). https://doi.org/10.3850/978-981-18-2016-8_732-cd
- Griffiths, P. (1983). Fourier Transform Infrared Spectrometry. *Science*, 222(4621), 297–302. <https://doi.org/10.1126/science.6623077>
- Huang, L. (2014). Modeling and Simulation in Spectroscopic Study. *Spectral Analysis Review*, 2(3), 7–9. <https://doi.org/10.4236/sar.2014.23023>
- Kozubovsky, V. R., & Bilak, Y. Y. (2022). Express Analysis of Gas Mixtures Using a Spectral Correlator Based on the Fabry – Perot Interferometer. *Journal of Applied Spectroscopy*, 89, 495–499. <https://doi.org/10.1007/s10812-022-01385-7>
- Kriesten, E., Alsmeyer, F., Bardow, A., & Marquardt, W. (2008). Fully automated indirect hard modeling of mixture spectra. *Chemometrics and Intelligent Laboratory Systems*, 91(2), 181–193. <https://doi.org/10.1016/j.chemolab.2007.11.004>
- Lavanya, A., Sindhuja, S., Gaurav, L., & Ali, W. (2023). Assessing the Performance of Python Data Visualization Libraries. *International Journal of Computer Engineering in Research Trends*, 10, 28–39. <https://doi.org/10.22362/ijcert/2023/v10/i01/v10i0104>
- Mishra, P., Passos, D., Marini, F., Xu, J., Amigo, J. M., Gowen, A. A., Jansen, J. J., Biancolillo, A., Roger, J. M., Rutledge, D. N., & Nordon, A. (2022). Deep learning for near-infrared spectral data modelling: Hypes and benefits. *TRAC Trends in Analytical Chemistry*, 157(2), 116804. <https://doi.org/10.1016/j.trac.2022.116804>
- Murray, S. (2013). *Interactive Data Visualization for the Web*. O'Reilly Media, Inc. <https://www.oreilly.com/library/view/interactive-data-visualization/9781449340223/>
- Rodionova, O. Y., & Brereton, R. G. (2005). Chemometrics: Data Analysis for the Laboratory and Chemical Plant. *Journal of Analytical Chemistry*, 60, 994–996. <https://doi.org/10.1007/s10809-005-0223-6>
- Shuaibov, O. K., Hrytsak, R. V., Minya, O. I., Malinina, A. A., Bilak, Yu. Yu., & Gomoki, Z. T. (2022). Spectroscopic diagnostics of overstressed nanosecond discharge plasma between zinc electrodes in air and nitrogen. *Journal of physical studies*, 26(2), 2501–2509. <https://doi.org/10.30970/jps.26.2501>
- Tagliaferri, R., Ciaramella, A., Barone, F., Longo, G., & Milano, L. (2001). Neural Networks for Spectral Analysis of Unevenly Sampled Data. In A. J. Banday, S. Zaroubi, & M. Bartelmann (Eds.), *Mining the Sky* (pp. 386–390). Springer Nature. https://doi.org/10.1007/10849171_47
- McKinney, W. (2017). *Python for Data Analysis* (2nd ed.). O'Reilly Media. <https://www.oreilly.com/library/view/python-for-data/9781491957653>
- Young, C. D. (2001). *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/0471220655>

References

- Alsmeyer, F., & Marquardt, W. (2004). Automatic Generation of Peak-Shaped Models. *Applied Spectroscopy*, 58, 986–994. <https://opg.optica.org/as/abstract.cfm?uri=as-58-8-986>
- Bilak, Yu., Kozubovskii, V., & Rol, M. (2023). Using the Fabry – Pérot interferometer as a spectral mask. *Journal of Physical Studies*, 27(1), 1402–1409. <https://doi.org/10.30970/jps.27.1402>
- Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. *Nature*, 559, 547–555. <https://doi.org/10.1038/s41586-018-0337-2>
- Butucea, C., Delmas, J.-F., Dufloy, A., & Hardy, C. (2021). Modeling Infrared Spectra: An Algorithm for an Automatic and Simultaneous Analysis. In B. Castanier, M. Cepin, D. Bigaud, & C. Berenguer (Eds.), *31st European Safety and Reliability Conference (ESREL), Angers, France* (pp. 3359–3366). https://doi.org/10.3850/978-981-18-2016-8_732-cd
- Griffiths, P. (1983). Fourier Transform Infrared Spectrometry. *Science*, 222(4621), 297–302. <https://doi.org/10.1126/science.6623077>
- Huang, L. (2014). Modeling and Simulation in Spectroscopic Study. *Spectral Analysis Review*, 2(3), 7–9. <https://doi.org/10.4236/sar.2014.23023>
- Kozubovsky, V. R., & Bilak, Y. Y. (2022). Express Analysis of Gas Mixtures Using a Spectral Correlator Based on the Fabry – Perot Interferometer. *Journal of Applied Spectroscopy*, 89, 495–499. <https://doi.org/10.1007/s10812-022-01385-7>
- Kriesten, E., Alsmeyer, F., Bardow, A., & Marquardt, W. (2008). Fully automated indirect hard modeling of mixture spectra. *Chemometrics and Intelligent Laboratory Systems*, 91(2), 181–193. <https://doi.org/10.1016/j.chemolab.2007.11.004>
- Lavanya, A., Sindhuja, S., Gaurav, L., & Ali, W. (2023). Assessing the Performance of Python Data Visualization Libraries. *International Journal of Computer Engineering in Research Trends*, 10, 28–39. <https://doi.org/10.22362/ijcert/2023/v10/i01/v10i0104>
- Mishra, P., Passos, D., Marini, F., Xu, J., Amigo, J. M., Gowen, A. A., Jansen, J. J., Biancolillo, A., Roger, J. M., Rutledge, D. N., & Nordon, A. (2022). Deep learning for near-infrared spectral data modelling: Hypes and benefits. *TRAC Trends in Analytical Chemistry*, 157(2), 116804. <https://doi.org/10.1016/j.trac.2022.116804>
- Murray, S. (2013). *Interactive Data Visualization for the Web*. O'Reilly Media, Inc. <https://www.oreilly.com/library/view/interactive-data-visualization/9781449340223/>
- Rodionova, O. Y., & Brereton, R. G. (2005). Chemometrics: Data Analysis for the Laboratory and Chemical Plant. *Journal of Analytical Chemistry*, 60, 994–996. <https://doi.org/10.1007/s10809-005-0223-6>
- Shuaibov, O. K., Hrytsak, R. V., Minya, O. I., Malinina, A. A., Bilak, Yu. Yu., & Gomoki, Z. T. (2022). Spectroscopic diagnostics of overstressed nanosecond discharge plasma between zinc electrodes in air and nitrogen. *Journal of physical studies*, 26(2), 2501–2509. <https://doi.org/10.30970/jps.26.2501>
- Tagliaferri, R., Ciaramella, A., Barone, F., Longo, G., & Milano, L. (2001). Neural Networks for Spectral Analysis of Unevenly Sampled Data. In A. J. Banday, S. Zaroubi, & M. Bartelmann (Eds.), *Mining the Sky* (pp. 386–390). Springer Nature. https://doi.org/10.1007/10849171_47
- McKinney, W. (2017). *Python for Data Analysis* (2nd ed.). O'Reilly Media. <https://www.oreilly.com/library/view/python-for-data/9781491957653/>

Young, C. D. (2001). *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*. John Wiley & Sons, Ltd.
<https://doi.org/10.1002/0471220655>

Отримано редакцією журналу / Received: 26.12.24
Прорецензовано / Revised: 22.02.25
Схвалено до друку / Accepted: 06.03.25

Yurii BILAK, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0001-5989-1643
e-mail: yuriy.bilak@uzhnu.edu.ua
Uzhgorod National University, Uzhgorod, Ukraine

INFORMATION SYSTEM BASED ON A COMPLEX MODEL USING MACHINE LEARNING FOR SPECTRAL ANALYSIS

The research is devoted to the design and development of an information system based on a complex model using machine learning methods to automate spectral analysis to increase the accuracy and speed of data processing. The history of the research is connected with the development of analytical methods in physics, chemistry and biology, where spectral analysis has traditionally played a key role. However, modern challenges, in particular the growth of data volumes and the need for automation, have stimulated the introduction of innovative methods based on artificial intelligence.

The relevance is due to the need to process large volumes of complex spectral data in real time, which is important for medicine, ecology, chemistry and other industries. Traditional analysis methods have limitations, so the use of machine learning is appropriate to increase the efficiency of the process.

The research focuses on the following issues: how to automate spectral data processing, how to ensure the integration of classical methods with machine learning, and how to increase the accuracy and scalability of the analysis. For this purpose, signal processing methods were applied, including noise filtering, smoothing, baseline correction, and peak analysis using derivatives and numerical integration. Machine learning was implemented through Random Forest models and neural networks adapted for predicting spectrum parameters.

The results showed that the developed system provides high accuracy and speed of spectral data analysis, interactive visualization of spectrum parameters, as well as the ability to integrate with other information platforms. This significantly simplifies analysis processes, reduces dependence on expert intervention, and increases productivity.

Research prospects include optimizing mathematical models for even greater accuracy, integration with IoT systems, and expanding the functionality for analyzing complex multidimensional spectra. This opens up opportunities for application in interdisciplinary projects, such as monitoring environmental changes or diagnosing medical conditions.

Keywords: *information system, artificial intelligence, spectral analysis, modeling, machine learning, intelligent data analysis.*

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

УДК 004.4+004.89+004.738.52

DOI: <https://doi.org/10.17721/1812-5409.2025/1.15>

Галина ДОЛЕНКО¹, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0001-5708-4880
e-mail: galyna.dolenko@knu.ua

Дарія МАНОВИЦЬКА², канд. техн. наук, провід. інж.-програм.
ORCID ID: 0000-0003-4567-8901
e-mail: manovytska@nas.gov.ua

Олена СУББОТИНА², наук. співроб.
ORCID ID: 0009-0008-3681-0263
e-mail: subotina@nas.gov.ua

Юрій ЧАПЛІНСЬКИЙ², канд. техн. наук, ст. наук. співроб.
ORCID ID: 0000-0001-7697-3958
e-mail: chaplinskiy@nas.gov.ua

Ігор БУДЯНСЬКИЙ³, інж.-програміст III
ORCID ID: 0009-0002-5107-9778
e-mail: igor.budyansky@yale.edu

¹Київський національний університет імені Тараса Шевченка, Київ, Україна

²Інститут кібернетики імені В. М. Глушкова НАН України, Київ, Україна

³Єльський університет, Нью-Гейвен, США

РОЗРОБЛЕННЯ ЧАТ-БОТУ СЛУЖБИ ПІДТРИМКИ КЛІЄНТІВ МЕРЕЖІ ЗАРЯДНИХ СТАНЦІЙ ЕЛЕКТРОМОБІЛІВ

Запропоновано сценарій автоматизованої цілодобової взаємодії з користувачами зарядних станцій для електромобілів на основі використання штучного інтелекту. Описано актуальність інтелектуального віртуального помічника для підтримки клієнтів службами власників мереж зарядних станцій електромобілів. Обґрунтовано необхідність застосування інтелектуальних технологій, таких як генеративний штучний інтелект компанії OpenAI, для створення чат-боту. Описано необхідність та актуальність інтеграції створеного чат-боту в найбільший за охопленням аудиторії месенджер WhatsApp. Для реалізації розробки проаналізовано і продемонстровано можливість застосування найпрогресивніших технологій сервісу OpenAI, а саме: моделлю для генерування тексту відповідей було обрано модель GPT-4o-mini, яку кастомізовано за допомогою системного контексту. Модель було конфігуровано у класі OpenAIAssistant, де сформульовано підхід, що вказує, як модель має діяти за визначеним алгоритмом / сценарієм. Розроблено і представлено алгоритм дії моделі. Вона повинна вміти вести комунікацію з користувачем різними мовами. Комунікація моделі з користувачем проходить за конкретно заданим промптами сценарієм, а також запрограмована на відповіді, які вона має видавати на запитання не зі сценарію / алгоритму, на невідому їй мову, та запрограмована промптами на захист від помилкових відповідей, які іноді допускає генеративний штучний інтелект. Реалізація чат-боту здійснена із застосуванням технологій на основі сервісів Whapi.Cloud (платформа для автоматизації роботи з API месенджера WhatsApp), Vultr Cloud Provider (провайдер для хостингу серверної частини додатку). Бізнес-логіка чат-бота реалізована на JavaScript, NodeJS, Express.

Практична реалізація розробки засвідчила конструктивність запропонованого підходу, що було підтверджено під час проведення тестування в месенджері WhatsApp. Запропонована стаття становить інтерес для фахівців у сфері генеративного штучного інтелекту.

Ключові слова: штучний інтелект, чат-бот, WhatsApp, API, OpenAI, Whapi.Cloud.

Класифікація відповідно до AMS 2020: 68P20, 68P10, 68T20, 68U35, 68W27.

Вступ

Нині підтримка клієнтів компаніями проходить період трансформації завдяки впровадженню генеративного штучного інтелекту (ШІ), що забезпечує більш персоналізований та ефективніший підхід до підтримки клієнтів (Rogers, 2016). Чат-боти на основі оброблення природної мови (англ. *Natural Language Processing*, NLP) та машинного навчання (англ. *machine learning*, ML) стали ключовим інструментом в такій трансформації, імітуючи розмову людини та працюючи в режимі реального часу (Castro, & New, 2016; Adam, Wessel, & Benlian, 2021). Також підтримується функція управління діалогом: керування потоком розмови, підтримання контексту й узгодженості.

NLP допомагає чат-боту розуміти й інтерпретувати користувацьке введення природною мовою, дає змогу надавати контекстно-залежні відповіді та поліпшує розмовні можливості. ML допомагає чат-ботам навчатися на взаємодії з користувачами й адаптувати свої відповіді з часом (Adamoroulou, & Moussiades, 2020). Завдяки ML підвищується здатність чат-бота обробляти складні запити та підвищувати точність розпізнавання намірів користувача, які стоять за запитом, щоб надати релевантні відповіді. У цих умовах використання методів підтримки ухвалення інфраструктурних рішень, що пов'язані з електромобілями, гарантує підвищення ефективності задоволення потреб користувачів чат-ботом (Dolenko, Manovytska, & Semenova, 2021).

Основні переваги чат-ботів:

- **Миттєва підтримка 24/7:** Віртуальні помічники доступні на вебсайтах, у месенджерах і соціальних мережах, обробляючи численні запити одночасно (Krishnan et al., 2022).
- **Персоналізація:** Завдяки інтеграції з CRM й аналізу клієнтських даних чат-боти забезпечують індивідуальні рекомендації та пропонують релевантні продукти або послуги (Paliwal, Bharti, & Mishra, 2020).

© Доленко Галина, Мановицька Дарія, Субботіна Олена, Чаплінський Юрій, Будянський Ігор, 2025

ISSN 1812-5409 (Print), ISSN 2218-2055 (Online)

• **Масштабованість:** Здатність обслуговувати великий обсяг запитів під час пікових навантажень без зниження якості (Zierau et al., 2021).

• **Багатомовність:** Підтримка клієнтів різними мовами для глобального охоплення (Aleedy, Shaiba, & Bezbradica, 2019).

• **Автоматизація рутинних завдань:** Швидке вирішення простих запитів і допомога операторам у складних випадках через збирання й аналіз інформації (Adam, Wessel, & Benlian, 2021).

Технологічні можливості:

• **Обробка природної мови (NLP):** Інтерпретація запитів клієнтів і формування контекстно-залежних відповідей.

• **Машинне навчання (ML):** Підвищення ефективності взаємодії через аналіз історичних даних і адаптацію до потреб клієнтів (Paliwal, Bharti, & Mishra, 2020; Mykhailov, 2023).

• **Інтеграція з платформами:** Узгоджена підтримка на різних каналах зв'язку, включно із соціальними мережами, месенджерами та вебсайтами (Krishnan et al., 2022; Nakonechnyi et al., 2022).

Чат-боти оптимізують операції, знижуючи навантаження на персонал, поліпшують якість обслуговування та збільшують задоволеність клієнтів. Завдяки технологіям ШІ компанії отримують ефективне, інноваційне рішення для взаємодії з клієнтами (Zierau et al., 2021), що забезпечує масштабування роботи чат-ботів.

З розвитком використання електромобілей розвиваються і технології, що підтримують їхнє використання. Чат-боти, що надають підтримку у використанні електромобілей – це сучасні інструменти на основі ШІ, які стали невід'ємною частиною оптимізації взаємодії з клієнтами в автомобільному секторі. Ці інтерактивні боти доступні на різних платформах, таких як платформи обміну повідомленнями, вебсайти та додатки, гарантуючи миттєву підтримку в разі потреби.

Підтримка клієнтів, що пов'язана з питаннями зарядження електромобілів, є дуже важливою, оскільки, наприклад, зарядне обладнання для електромобілів визначається певною множиною параметрів, а саме режими/рівні зарядження, потужність зарядного пристрою, стандарти роз'ємів, принцип використання тощо. Уже тепер існують мобільні додатки (PlugShare, To-U, TOKA, YASNO E-mobility тощо), що надають можливість пошуку найближчої зарядної станції, статус зарядної станції, планування маршруту до зарядної станції. Проте власник електромобіля бажав би знати більше, наприклад, чи підходить ця зарядна станція до його електромобіля, чи підходить роз'єм зарядної станції, який режим зарядження буде оптимальним, яка буде вартість зарядження тощо. Чат-боти на базі моделі GPT (Generative Pretrained Transformer) можуть надати підтримку в отриманні відповіді на ці важливі питання. Основою для використання таких чат-ботів є формування певних запитань (запитів, промптів). Це базується на реалізації певної промпт-компетентності. Під промпт-компетентністю будемо розуміти уміння розпізнати, яка інформація необхідна моделі штучного інтелекту для створення відповідних відповідей, сформулювати цю інформацію чітко й точно, і вносити послідовні коригування інформації відповідно до можливостей моделі, наприклад, конкретні умови чи обмеження. Промпт-компетентність формується на основі промпт-інжинірингу (prompt engineering), що є процесом створення та формулювання оптимальних промптів (запитань або запитів) для генеративних моделей ШІ для досягнення найкращого результату згенерованого контенту. У процесі створення запитань (промптів) дотримуються базових принципів у їхньому формуванні, наприклад, визначають, для кого формується завдання, формулюють конкретне завдання, яка генеративна модель GPT має виконати, надають необхідний контекст пошуку, що допоможе моделі GPT сфокусуватися на генерації більш релевантних відповідей, обирають бажаний формат виведення відповіді або стиль представлення інформації, визначають необхідні деталі, обмеження або вимоги.

Отже, чат-боти з моделлю штучного інтелекту GPT дають змогу, використовуючи відповідне серверне та хмарне програмне забезпечення, реалізувати сценарії від формування запитань (промптів) у межах моделей GPT до отримання відповідей на сформоване завдання.

Метою статті є розроблення багатомовного чат-боту, що інтегрований у сервіс WhatsApp та забезпечує допомогу клієнтів компанії, яка займається менеджментом зарядних станцій для електромобілів.

1. Основні результати

Підтримка клієнтів, що будується на спілкуванні з працівником служби підтримки за допомоги телефонних дзвінків, не дає змоги реалізувати оперативну та якісну допомогу з питань процесу зарядження електромобілей (стандартні відповіді на певної мірою унікальні запитання, фахівець служби підтримки не завжди має повну інформацію щодо конкретного електромобіля). Питання щодо процесу зарядження завжди вимагають швидкої, конкретної та оперативної відповіді. У пропонуванні роботі досліджено електромобілі, що використовують акумуляторні батареї як джерело руху та мають зарядний порт для зовнішньої зарядки. Як характеристики електромобілей розглянуто модель, виробника, тип або категорію акумуляторної батареї, тип порту зарядки тощо. Як елемент екосистеми електромобілів проаналізовано зарядні станції, що характеризуються власником, виробником, типами роз'ємів, доступними режимами зарядження, вартістю зарядження тощо.

Одним із сучасних засобів розв'язання проблеми підтримки клієнтів з питань зарядження електромобілей є використання моделі GPT-4o mini, що розроблена компанією OpenAI. Головні переваги моделі – компактність і швидкість. Ця модель генерує текст швидше за інші моделі штучного інтелекту, що є ключовою перевагою, коли важлива оперативність надання відповіді.

Отже, необхідно створити чат-бот, який має забезпечити надання можливості подання запитів й отримання відповідей різними мовами та підтримати виконання сценаріїв взаємодії, що охоплюють основні запити клієнтів. Такий чат-бот буде інтегровано в сервіс WhatsApp через API, використовуючи платформу Whapi.Cloud. Вибір WhatsApp визначався тим, що за даними інформаційного ресурсу Similarweb за 2023 рік сервіс WhatsApp є найпопулярнішим месенджером у 63 країнах, зокрема Німеччині, Іспанії. Whapi.Cloud являє собою API-шлюз, що дає змогу надсилати, отримувати та відстежувати повідомлення у WhatsApp.

Для безперебійної роботи чат-боту використовують платформу Vultr Cloud, що забезпечує розгортання інфраструктури чат-бота за допомогою хмарної платформи.

Робота чат-бота базується на певному сценарії. Будемо розуміти під сценарієм чат-бота деталізовану послідовність взаємопов'язаних кроків, що виконуються під час взаємодії з користувачами. Цей деталізований план

визначає послідовність від формулювання запитів (промптів) до отримання відповіді з урахуванням їхнього оброблення та можливості виникнення непередбачуваних ситуацій.

Під час створення сценарію проведено попередній аналіз основних запитів клієнтів, а також урахувано можливі невизначені запити, відповіді іншими мовами, що не підтримуються чат-ботом, або запити поза сценарієм.

Створення сценарію починається з визначення мети чат-бота. У нашому випадку метою є забезпечення цілодобової підтримки клієнтів електрозарядних станцій і збір інформації щодо зворотного зв'язку. Чітке розуміння мети допомагає спрямувати наступний (другий) крок – визначення цільової аудиторії. Цільовою аудиторією чат-боту є власники електромобілів із різними рівнями технічної підготовки. Розуміння цієї аудиторії допомагає створити найбільш ефективний і корисний сценарій. Третій крок – аналіз запитів: були враховані найпоширеніші ситуації, з якими стикаються користувачі, використовуючи дані попередніх взаємодій. Визначення основних запитів зробило можливим побудову сценарію, що покриває найтипівші ситуації. Після аналізу запитів переходять до визначення та реалізації основних гілок сценарію. Отже четвертий крок – розроблення основних гілок сценарію, а саме: відповіді на основні запити та передбачення роботи з некоректними запитами. Кожна з гілок має відповідати на конкретний запит. Передбачені різні варіанти розвитку подій і забезпечені логічні переходи між гілками. Робочий процес, що описує сценарій чат-боту, представлено на рис. 1.

Для того, щоб виконати описаний сценарій необхідно зробити таке:

1. Вибрати модель штучного інтелекту для генерації тексту.

- Використано модель GPT-4o-mini, адаптовану через інженерію підказок. Системний контекст спрямовує модель для роботи в межах визначеного сценарію.

- Модель отримує вхідний запит користувача разом із конфігураційними параметрами, генерує відповідь відповідно до заданого сценарію.

2. Зробити інтеграцію з платформою WhatsApp через Whapi.Cloud API.

- Створено тестовий канал для обміну повідомленнями.

- API-токен використовуємо для аутентифікації та передавання даних між сервером і WhatsApp.

3. Створити інфраструктуру серверної частини.

- Використано хмарний провайдер Vultr для хостингу серверної частини додатку.

- Налаштовано вебсервер за допомогою Express для оброблення запитів і взаємодії з вебхуками платформи Whapi.Cloud.

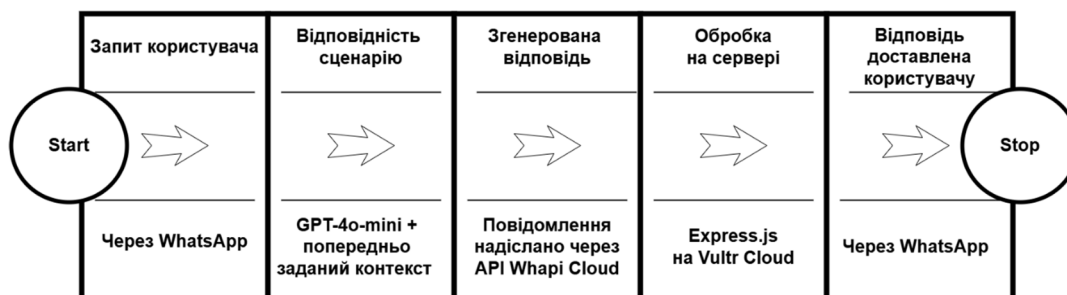


Рис. 1. Схема робочого процесу чат-бота

2. Приклади роботи системи

Як приклади роботи продемонструємо найважливіші частини коду з особливостями використання технологій OpenAI у поєднанні з API-токеном отриманим через Whapi.Cloud для процесу створення повністю функціонального чат-бота, що дає змогу здійснювати автоматизовану взаємодію з користувачами через чат за заданими сценаріями.

Після того, як модель gpt-4o-mini було кастомізовано за допомогою системного контексту <https://platform.openai.com/docs/guides/prompt-engineering#tactic-instruct-the-model-to-answer-using-a-reference-text>, визначаємо клас ScriptSupplier, що керує сценарієм запитань і відповідей.

```
export class ScriptSupplier {
```

```
  getScript() {
    return
```

```
    Question: I can't charge my car.
```

```
    Answer: For Elinta charging station, please first take the charging cable out of your car (or borrow it at reception), then connect the cable from the charging station to the car, then pay with a credit card at the charging station.
```

```
    If you are using xCharge fast charging station, first pay with credit card, then plug the cable from the charging station into the car.
```

```
    If it doesn't work the first time, please remove the cable from the car and try again.
```

```
    Question: Charging cable cannot be removed after charging
```

```
    Answer: Please read the car instructions how to unlock the cable from your particular car, often it is a button next to the cable socket in the car.
```

```
    Question: Charging cable is stuck and cannot be removed after charging.
```

```
    Answer: Please read your car instructions how to unlock the cable from your particular car. If you don't have the instructions, please write here the brand and model of your car and we'll try to help you to disconnect the cable.
```

Question: How much does charging cost?

Answer: It costs 0.55 euros per kWh.

```
`;
}
}
```

Новизною є навчання й обмеження цієї моделі спеціальними запитами (промптами) так, щоб уникнути випадків надання ШІ неправильно інформації користувачу, для чого конфігуруємо модель у класі OpenAIAssistant, де вказуємо, що модель має діяти за наведеним сценарієм, а також зазначаємо, які відповіді мають бути на запитання не зі сценарію та на невідому мову.

```
import OpenAI from "openai";
import * as os from "os";
const openai = new OpenAI();

export class OpenAIAssistant {

  _openai = null;
  _scriptSupplier = null;

  static _MODEL = "gpt-4o-mini";

  constructor({ scriptSupplier }) {
    this._openai = new OpenAI();
    this._scriptSupplier = scriptSupplier;
  }

  _getConfigPrompt() {
    const script = this._scriptSupplier.getScript();
    return `
```

You are a customer support assistant assistant. You support languages: Spanish, English, Ukrainian, German.

You must answer the users' questions according to this script:

\${script}

If a customer asks something that is not in the script, reply "I cannot answer this question. Here are the questions I can answer without an operator" and then list a short summary of the questions from the script.

If a customer asks a question in a language other than Spanish, English, Ukrainian or German, reply "I cannot understand this language. I only support Spanish, English, Ukrainian and German".

```
`;
}
```

```
async generateResponse({ query }) {
  const configPrompt = this._getConfigPrompt();
  const completion = await openai.chat.completions.create({
    model: OpenAIAssistant._MODEL,
    messages: [
      { role: "system", content: configPrompt },
      { role: "user", content: query },
    ],
  });
  const response = completion.choices[0].message;
  return {
    role: response.role,
    content: response.content,
    refusal: response.refusal,
  }
}
```

Чат-бот підтримує багатомовність. Для кожної мови було налаштовано окремий сценарій для взаємодії.

Це забезпечує глобальний доступ для користувачів, незалежно від їхніх мовних переваг.

Особливість цього чат-боту досягається і завдяки використанню сучасних хмарних технологій для забезпечення стабільності роботи під навантаженням. Для цього обрано платформу Vultr Cloud, як засіб реалізації хостингу серверної частини. Це надає можливість для чат-боту створення віртуальної машини з публічною IP-адресою для налаштування сервера, який зможе приймати запити й отримувати відповіді від чат-боту з будь-якої точки світу.

Прикладом високоефективної інтеграції є використання функції Express для налаштування вебсервера, що приймає запити.

```

import 'dotenv/config';
import express from 'express';
import { OpenAIAssistant } from './openaiAssistant.js';
import { WhatsAppBot } from './whatsappBot.js';
import { ScriptSupplier } from './scriptSupplier.js'
const scriptSupplier = new ScriptSupplier();
const openAiAssistant = new OpenAIAssistant({ scriptSupplier });
const whatsappBot = new WhatsAppBot({ aiAssistant: openAiAssistant });
const app = express();
app.use(express.json());
app.post('/webhook', async (req, res) => {
  try {
    const messages = req.body.messages;
    for (const message of messages) {
      await whatsappBot.handleMessage(message);
    }
  } catch (e) {
    console.log(e);
  }
  res.send('Ok');
});
app.listen(80, () => {
  console.log('Server started');
});

```

Після цього необхідно дозволити Whapi.Cloud надсилати повідомлення на вебсервер через вебхуки. Вебсервер встановлює так звані вебхуки, що базуються на URL-з'єднанні між чат-ботом та чат-платформою. Вебхуки дають можливість безпечно відправляти й отримувати повідомлення через прості HTTP-запити. Щоб отримувати вхідні дані з месенджера, такі як вхідне повідомлення, вхідний дзвінок або сповіщення, використовують вебхук. Ця технологія дає змогу миттєво отримувати сповіщення про будь-яку подію у WhatsApp і реагувати на неї.

Отже, дві системи (API-шлюз, під'єднаний до WhatsApp, і чат-бот) обмінюються інформацією без необхідності будь-яких зовнішніх дій.

Технічно, отримавши повідомлення, система виконає POST-запит з JSON-тілом на налаштовану URL-адресу скрипта. Тіло JSON містить усю інформацію про вхідне повідомлення, таку як текст, номер відправника, час відправлення тощо.

Для тестування роботи чат-боту для WhatsApp використовували платформу Vultr Cloud.

Чат-бот швидко й точно відповідає на запитання, пов'язані із зарядженням електромобілів, що підтверджує його практичну придатність. Чат-бот надає відповіді на запитання про вартість зарядки, під'єднання кабелю й інші питання.

Рис. 2 демонструє приклад взаємодії з користувачем у месенджері WhatsApp.

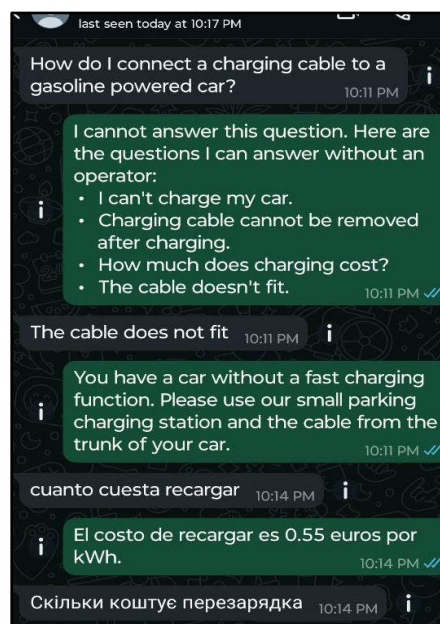


Рис. 2. Приклад взаємодії із чат-ботом у WhatsApp

Дискусія і висновки

У публікації представлено результати розроблення чат-бота, що працює на моделі штучного інтелекту GPT-4o mini. Ця модель надає чат-боту належної швидкості й оперативності, що є дуже важливими у процесі реалізації підтримки

клієнтів службами власників мереж зарядних станцій електромобілів. Інтеграція створеного чат-боту в сервіс WhatsApp надає можливість великому колу власників електромобілів використати переваги ШІ у процесі підготовки до заряджання або в самому процесі заряджання. Запропоновано сценарій чат-боту. Процес взаємодії з користувачами надає можливість формувати певну множину запитів (промтів) та подавати відповіді різними мовами. У процес реалізації чат-боту залучено мережеве кросплатформове середовище Node.js разом із вебфреймворком Express.js. Використання Whapi.Cloud дає змогу чат-боту взаємодіяти із сервісом WhatsApp. Для організації серверної частини чат-боту використано Vultr Cloud.

Завдяки практичній реалізації чат-боту у компаній з управління зарядними станціями електромобілів розширюються можливості взаємодії з клієнтами щодо кола питань підтримки та функціональності, підвищення швидкості й оперативності надання відповідей. Під час тестування чат-бота було досягнуто таких результатів:

- Чат-бот зміг обробляти 95 % запитів без втручання оператора.
- Середній час відповіді становив 1 секунду, що значно швидше порівняно з традиційними методами підтримки (напр., телефонні дзвінки або електронна пошта).
- Точність відповідей на основні запити (про вартість зарядки, підключення кабелю) досягала 98 %.
- Чат-бот підтримував три мови (англійську, іспанську, українську) з точністю перекладу 95 %.

Крім того, чат-бот знизив навантаження на операторів на 40 %, а кількість скарг від клієнтів зменшилася на 25 % завдяки швидкій і точній підтримці. У майбутньому планується впровадити SSL-шифрування для забезпечення безпеки даних, додати функцію зв'язку з оператором для складних запитів і розширити підтримку мов (напр., французькою, китайською та італійською).

Рекомендовано в майбутньому перейти до використання підходу CoT (Chain-of-Thought) (Wei et al., 2022). CoT базується на ланцюгу запитань, що дає змогу моделям штучного інтелекту розкласти складні багатетапні проблеми на простіші підзадачі. Це може бути особливо корисним для чат-бота, який працює з електромобілями, оскільки користувачі часто ставлять складні запитання, що вимагають поетапного розв'язання. Крім того, CoT може бути інтегрований із семантичними засобами, такими як онтології, що стосуються електромобілів. Це допоможе чат-боту краще розуміти контекст запитів і надавати точніші відповіді.

Внесок авторів: Галина Доленко – концептуалізація, методика; Дарія Мановицька – методика, програмне забезпечення; Олена Субботіна – редагування, аналіз джерел; Юрій Чаплінський – концептуалізація, підготовка огляду літератури; Ігор Будянський – консультування.

Список використаних джерел

- Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31(2), 427–445. <https://doi.org/10.1007/s12525-020-00414-7>.
- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Aleedy, M., Shaiba, H., & Bezbradica, M. (2019). Generating and analyzing chatbot responses using natural language processing. *International Journal of Advanced Computer Science and Applications*, 10(9), 60–68. <https://dx.doi.org/10.14569/IJACSA.2019.0100910>
- Castro, D., & New, J. (2016). *The promise of artificial intelligence*. Center for Data Innovation. <https://www2.datainnovation.org/2016-promise-of-ai.pdf>
- Dolenko G., Manovytska, D., & Semenova, N. (2021). Problems of decision making in subject domain "electric vehicle". *CEUR Workshop Proceedings*, 3018, 175–183. https://ceur-ws.org/Vol-3018/Paper_17.pdf
- Krishnan, C., Gupta, A., Gupta, A., & Singh, G. (2022). Impact of artificial intelligence-based chatbots on customer engagement and business growth. In T. P. Hong, L. Serrano-Estrada, A. Saxena, & A. Biswas (Eds.), *Deep Learning for Social Media Data Analytics. Studies in Big Data*, vol 113 (pp. 195–210). Springer International Publishing. http://dx.doi.org/10.1007/978-3-031-10869-3_11
- Mykhailov, N. O. (2023). Classifications of users on online platforms using machine learning techniques. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 4, 66–71. <https://doi.org/10.17721/1812-5409.2022/4.8>
- Nakonechnyi, O. G., Kapustian, O. A., Shevchuk, I. M., Loseva, M. V., & Kosukha, O. Y. (2022). A intellectual system of analysis of reactions to news based on data from Telegram channels. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 3, 55–61. <https://doi.org/10.17721/1812-5409.2022/3.7>
- Paliwal, S., Bharti, V., & Mishra, A.K. (2020). Ai chatbots: Transforming the digital world. In V. Balas, R. Kumar, & R. Srivastava (Eds.), *Recent Trends and Advances in Artificial Intelligence and Internet of Things, Intelligent Systems Reference Library*, vol. 172 (pp. 455–482). Springer International Publishing. https://doi.org/10.1007/978-3-030-32644-9_34
- Rogers, D. L. (2016). *The digital transformation playbook: Rethink your business for the digital age*. Columbia University Press. <https://doi.org/10.7312/roge17544>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E.H., Xia, F., Le, Q., & Zhou, D. (2022). *Chain of Thought Prompting Elicits Reasoning in Large Language Models*. <https://doi.org/10.48550/arXiv.2201.11903>
- Zierau, N., Flock, K., Janson, A., Söllner, M., & Leimeister, J. M. (2021). The Influence of AI-Based Chatbots and Their Design on Users Trust and Information Sharing in Online Loan Applications. In *Proceedings of the 54th Hawaii International Conference on System Sciences 2021 (HICSS 54)* (pp. 5483–5492). <https://dx.doi.org/10.24251/HICSS.2021.666>

References

- Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31(2), 427–445. <https://doi.org/10.1007/s12525-020-00414-7>.
- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Aleedy, M., Shaiba, H., & Bezbradica, M. (2019). Generating and analyzing chatbot responses using natural language processing. *International Journal of Advanced Computer Science and Applications*, 10(9), 60–68. <https://dx.doi.org/10.14569/IJACSA.2019.0100910>
- Castro, D., & New, J. (2016). *The promise of artificial intelligence*. Center for Data Innovation. <https://www2.datainnovation.org/2016-promise-of-ai.pdf>
- Dolenko G., Manovytska, D., & Semenova, N. (2021). Problems of decision making in subject domain "electric vehicle". *CEUR Workshop Proceedings*, 3018, 175–183. https://ceur-ws.org/Vol-3018/Paper_17.pdf
- Krishnan, C., Gupta, A., Gupta, A., & Singh, G. (2022). Impact of artificial intelligence-based chatbots on customer engagement and business growth. In T. P. Hong, L. Serrano-Estrada, A. Saxena, & A. Biswas (Eds.), *Deep Learning for Social Media Data Analytics. Studies in Big Data*, vol 113 (pp. 195–210). Springer International Publishing. http://dx.doi.org/10.1007/978-3-031-10869-3_11
- Mykhailov, N. O. (2023). Classifications of users on online platforms using machine learning techniques. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 4, 66–71. <https://doi.org/10.17721/1812-5409.2022/4.8>
- Nakonechnyi, O. G., Kapustian, O. A., Shevchuk, I. M., Loseva, M. V., & Kosukha, O. Y. (2022). A intellectual system of analysis of reactions to news based on data from Telegram channels. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 3, 55–61. <https://doi.org/10.17721/1812-5409.2022/3.7>

- Paliwal, S., Bharti, V., & Mishra, A.K. (2020). Ai chatbots: Transforming the digital world. In V. Balas, R. Kumar, & R. Srivastava (Eds.), *Recent Trends and Advances in Artificial Intelligence and Internet of Things, Intelligent Systems Reference Library*, vol. 172 (pp. 455–482). Springer International Publishing. https://doi.org/10.1007/978-3-030-32644-9_34
- Rogers, D. L. (2016). *The digital transformation playbook: Rethink your business for the digital age*. Columbia University Press. <https://doi.org/10.7312/roge17544>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E.H., Xia, F., Le, Q., & Zhou, D. (2022). *Chain of Thought Prompting Elicits Reasoning in Large Language Models*. <https://doi.org/10.48550/arXiv.2201.11903>.
- Zierau, N., Flock, K., Janson, A., Söllner, M., & Leimeister, J. M. (2021). The Influence of AI-Based Chatbots and Their Design on Users Trust and Information Sharing in Online Loan Applications. In *Proceedings of the 54th Hawaii International Conference on System Sciences 2021 (HICSS 54)* (pp. 5483–5492). <https://dx.doi.org/10.24251/HICSS.2021.666>

Отримано редакцією журналу / Received: 17.01.25

Прорецензовано / Revised: 08.03.25

Схвалено до друку / Accepted: 12.03.25

Galyna DOLENKO¹, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0001-5708-4880
e-mail: galyna.dolenko@knu.ua

Dariia MANOVYTSKA², PhD (Engin.), Leading Programmer
ORCID ID: 0000-0003-4567-8901
e-mail: manovytska@nas.gov.ua

Olena SUBBOTINA², Researcher
ORCID ID: 0009-0008-3681-0263
e-mail: subbotina@nas.gov.ua

Yuriy CHAPLINSKY², PhD (Engin.), Senior Researcher
ORCID ID: 0000-0001-7697-3958
e-mail: chaplinskiy@nas.gov.ua

Igor BUDYANSKY³, Software Engineer III
ORCID ID: 0009-0002-5107-9778
e-mail: igor.budyansky@yale.edu

¹Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

²V. M. Glushkov Institute of Cybernetics of the NAS of Ukraine, Kyiv, Ukraine

³Yale University, New Haven, USA

CHATBOT FOR CUSTOMER SUPPORT OF A COMPANY WHICH OPERATES CHARGING STATIONS FOR ELECTRIC CARS

A scenario for automated, around-the-clock interaction with users of electric vehicle charging stations, based on artificial intelligence, has been proposed. The relevance of an intelligent virtual assistant for supporting customers by the service departments of companies managing charging station networks is discussed. The necessity of employing market-leading technologies, such as OpenAI's generative artificial intelligence models, for the creation of a chatbot is substantiated. The importance and timeliness of integrating the developed chatbot into WhatsApp, the messenger with the largest audience reach, are described. To implement the solution, the potential for applying the most advanced OpenAI service technologies was analyzed and demonstrated. Specifically, the GPT-4o-mini model was selected for generating text-based responses and customized using system context. The model was subsequently configured in the OpenAIAssistant class, where a method specifying how the model should function according to the provided algorithm or scenario was developed. An algorithm for the model's operation was designed and reviewed. The model must be capable of communicating with users in various languages. Communication is guided by a predefined scenario based on prompt templates. Additionally, the model is programmed to respond to questions that fall outside the provided scenario or algorithm, to questions posed in unknown languages, and includes prompt-based safeguards against incorrect answers that AI occasionally generates. For the chatbot's implementation, technologies were employed, including Whapi.Cloud services (a platform for automating interactions with the WhatsApp messenger API) and Vultr Cloud Provider (a provider for hosting the server-side application). The chatbot's business logic was implemented using programming languages and frameworks such as JavaScript, NodeJS, and Express.

The practical implementation of this development confirmed the feasibility and effectiveness of the proposed approach, as evidenced during testing within the WhatsApp messenger. The proposed article is of interest to software developers working in the field of generative artificial intelligence.

Keywords: Artificial Intelligence, Chatbot, WhatsApp, API, OpenAI, Whapi.Cloud.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 004.032.26

DOI: <https://doi.org/10.17721/1812-5409.2025/1.16>

Dmitriy KLYUSHIN, DSc (Phys. & Math.), Prof.

ORCID ID: 0000-0003-4554-1049

e-mail: dmytroklyushin@knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Oleksandr MAISTRENKO, PhD Student

ORCID ID: 0009-0003-3464-7777

e-mail: o.maistrenko@knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

USE OF IZHIKEVICH NEURONS IN HOPFIELD MODELS

The Izhikevich chaotic neuron model represents a considerable advancement in computational neuroscience by offering a mathematical framework that closely mirrors the behavior of biological neurons, especially in generating chaotic or complex spiking patterns seen in the brain. This model has garnered attention due to its ability to replicate a wide range of real-life neural dynamics, including bursting and tonic spiking, which are fundamental to understanding the complexities of neural communication and processing. On the other hand, Hopfield networks, a type of recurrent neural network, have long been recognized for their ability to serve as content-addressable memory systems, storing and recalling information based on associative dynamics. Often described as spin glass systems, Hopfield networks operate by finding stable states or patterns within the neural network, emulating certain memory functions of the human brain. Recently, research into innovative activation functions has opened new possibilities for enhancing the capabilities of recurrent networks. The chaotic activation functions, in particular, present an intriguing area of exploration within Hopfield networks. This article investigates the effects of embedding these chaotic activation functions in Hopfield networks, examining how they influence the network's stability, adaptability, and efficiency. Through this exploration, we aim to reveal the impact of chaos on the network's dynamics, providing insights that could potentially lead to improved performance in applications requiring complex memory and associative processing. The study contributes to the growing field of neuromorphic engineering, with implications for both artificial intelligence and neuroscience.

Key words: self-concept of the individual, dichotomous orientation of the transformation of the self-concept of the individual, constructive transformation of the self-concept of the individual, destructive transformation of the self-concept of the individual.

AMS 2020 classification: 92B20.

Introduction

In the field of computational neuroscience, the ability to accurately model and simulate the complex behavior of neurons is crucial for understanding brain function. One of the most promising advances in this area is the development of Izhikevich chaotic neurons. These models, named after neuroscientist Eugene Izhikevich, offer a powerful tool for replicating the highly dynamic and unpredictable behavior observed in real biological neurons.

Chaotic neurons differ from traditional neuron models by incorporating chaotic dynamics—nonlinear patterns that are sensitive to initial conditions, much like real brain activity. In biological systems, such chaotic patterns are essential for various cognitive processes, from memory encoding to decision-making, and even in maintaining the brain's flexibility to adapt to new information. Traditional neuron models, which often rely on simpler, more linear systems, struggle to capture this complexity.

The Izhikevich model strikes a balance between biological realism and computational efficiency, allowing for the simulation of large neural networks without requiring the immense computational resources that more complex models demand. This balance is critical in making large-scale simulations of brain function more accessible, aiding researchers in studying everything from learning and memory to neurological disorders.

The introduction of chaotic dynamics into neuron models represents a paradigm shift in computational neuroscience. By better capturing the intricacies of real-life neural processes, Izhikevich chaotic neurons allow scientists to explore new hypotheses about brain behavior, including how the brain processes information, adapts to change, and possibly even how it malfunctions in the case of diseases like epilepsy or schizophrenia.

As this research continues to evolve, the implications extend far beyond academic inquiry. By understanding how chaotic processes contribute to brain function, future developments could influence the creation of more advanced artificial intelligence systems, as well as new treatments for neurological conditions. In this sense, Izhikevich chaotic neurons represent not just a step forward for neuroscience but a key to unlocking deeper understanding of the brain's underlying mechanisms.

Activation functions play a vital role in transmitting signals between neuron layers in deep learning models. While sigmoid and hyperbolic tangent functions were commonly utilized in the past, recent research has revealed the benefits of using chaos-based activation functions (Dubey, Singh, & Chaudhuri, 2022). These functions offer features that enable faster and more efficient convergence during neural network training. By broadening the search space for optimal parameters, they reduce the risk of the model becoming stuck in local minima and support a more stable learning process (Wang, Ren, & Wang, 2022; Kilicarslan, Adem, & Celik, 2021).

Improved prediction accuracy and reduced training time for deep models can be achieved as a result. Furthermore, introducing chaos into activation functions helps resolve the issue of training reproducibility, as these functions are generally less sensitive to initial conditions (Chen, & Wang, 2023). However, it should be recognized that chaos-based activation functions might require more sophisticated optimization techniques and could also increase computational demands.

The Hopfield model plays a pivotal role in the realm of neural network theory, offering a profound mathematical foundation for understanding complex systems of artificial neurons. At its core, the Hopfield neural network is a recurrent, fully connected network, meaning that every node (or neuron) in the network is connected to every other node, forming a dense network topology. This full connectivity allows the network to engage in more intricate and dynamic interactions, making it capable of solving a variety of problems.

One of the distinctive features of the Hopfield model is its symmetric connection matrix. The weights between the neurons in this network are not only interconnected but also symmetrical—meaning that the weight from neuron i to neuron j is the

© Klyushin Dmitriy, Maistrenko Oleksandr, 2025

same as the weight from neuron j to neuron i . This symmetry is a crucial aspect of the Hopfield model, as it ensures that the network's dynamics can evolve in a manner that guarantees the convergence to stable equilibrium states. These equilibrium states, where the system eventually settles, are of great interest because they correspond to local minima of a function known as the network energy.

The network energy is typically represented as a negatively defined quadratic form, often visualized geometrically as a surface over an n -dimensional cube. The network evolves by updating the states of its neurons according to a set of rules that ultimately drive the system towards these minima. This energy function is a measure of how "far" the system is from its most stable configuration, and as the network processes its inputs, it reduces this energy until it finds a point where no further reduction is possible—this point is where the system reaches equilibrium.

The activation functions in Hopfield models are designed to emulate the behavior of neurons in the brain. Although linear and sigmoidal activation functions are commonly used, studies have demonstrated that chaotic activation functions can yield interesting results (Deng, Wong, & Lin, 2024a). Chaotic functions, for example, introduce an element of randomness into the problem-solving process, which can be particularly useful for tasks like optimization or pattern recognition, where variability and dynamic responses are essential (Deng, Wong, & Lin, 2024b).

Incorporating chaotic activation functions into Hopfield networks can enhance performance, especially when dealing with problems characterized by numerous local minima. Their complex, unpredictable nature helps the network navigate toward global optima more effectively (Magallon et al., 2024). However, this approach also presents challenges, such as increased complexity in training and higher computational costs, as these functions require more advanced optimization strategies and processing techniques.

Another advantage of using chaotic activation functions in Hopfield networks is their ability to create more intricate dynamic relationships between neurons. This can improve the network's ability to adapt to changes in input data or evolving environmental conditions. Such chaotic interactions also enable the model to better mimic complex dynamic processes, such as memory retention and bias learning (Cursino, & Dias, 2024; Krotov, 2023).

The use of chaotic activation functions in neural networks introduces a level of complexity that demands meticulous attention to both the fine-tuning of model parameters and the selection of the most appropriate training methods. Unlike traditional activation functions, chaotic functions exhibit non-linear and unpredictable behavior, which can create challenges during the training process. As such, any adjustments made to the model's parameters—such as learning rates, weight initialization, and network architecture – must be approached with great care. If these adjustments are inadequate or improperly balanced, the network may exhibit erratic or unstable behavior, potentially causing the learning process to stagnate or fail entirely. Furthermore, the chaotic nature of these functions means that the training dynamics may not converge in a straightforward manner, often requiring more advanced techniques like adaptive learning rates or specialized optimization algorithms to manage the inherent instability.

One of the key challenges with chaotic activation functions lies in their unpredictability, which can sometimes manifest as highly variable outputs that deviate significantly for minor changes in input. This unpredictability may not be conducive to tasks where high precision and consistent performance are paramount. For example, applications such as financial forecasting, medical diagnostics, or autonomous driving require stable and reliable decision-making processes, as even small fluctuations in the network's behavior could result in critical errors. In such contexts, the inherent variability of chaotic functions might compromise the model's ability to make consistently accurate predictions, which is crucial for applications that rely on a high degree of reliability and precision.

Looking ahead, the future of chaotic activation functions in neural networks holds promise. Advancements in computational power and optimization algorithms may address current limitations, paving the way for broader adoption. Hybrid approaches that combine chaos with traditional methods could offer a balance between precision and adaptability. Furthermore, the exploration of chaos theory in quantum computing and neuromorphic hardware could unlock new frontiers for neural network design.

There is a growing demand for more in-depth research on the application of chaotic activation functions within Hopfield networks, spurred by the increasing interest in neural networks and their potential roles in complex information processing tasks. Although some studies have already highlighted the promise of this approach, the subject is still not well represented in the scientific literature. Therefore, it is crucial to conduct more extensive experimental and theoretical investigations to explore the potential advantages and disadvantages of integrating chaotic activation functions into Hopfield networks (Liang et al., 2024; Kashyap et al., 2022).

1. Activation functions' comparison

Differentiable activation functions in neural networks provide several significant advantages over their nondifferentiable counterparts. Most importantly, they allow for the effective use of optimization algorithms like backpropagation during network training. The smooth nature of differentiable functions aids in finding optimal parameters through gradient search techniques, which streamlines the training process and improves the convergence rate.

It is also important to recognize that differentiable activation functions enable the application of regularization methods that help reduce overfitting and improve generalization capabilities. For example, techniques such as L1 or L2 regularization can be applied to model weights, helping to control their complexity and mitigate the risk of overfitting.

Additionally, differentiable activation functions create a smooth and continuous relationship between the network's inputs and outputs, allowing for gradual transitions between values and simplifying the optimization process. They also facilitate the use of more sophisticated optimization algorithms, such as adaptive gradient descent methods, which improve both the speed and efficiency of network convergence.

In (Lieberman-Aiden, 2009), it was proposed that a neuron's behavior can be described by a dynamical system, represented as a one-dimensional function that is both piecewise linear and piecewise continuous:

$$x(t + 1) = F(x(t), h(t)). \quad (1)$$

This category of functions is defined by the function $K(h)$, which is influenced by the slope angle. This relationship must satisfy the following conditions: 1) symmetry, such that $K(-h) = K(h)$; 2) a monotonic decrease for $h > 0$ (and consequently, a monotonic increase for $h < 0$); and 3) $K(0) = 2$. The following function meets these criteria

$$K(h) = \frac{2}{1+|h|/\mu}. \quad (2)$$

The behavior of the formal neuron is defined by a ratio:

$$x(t+1) = \begin{cases} 1, & h(t) > 0 \\ -1, & h(t) < 0 \end{cases} \quad (3)$$

It is well-known that a real neuron operates as a highly complex system with inertia (Freeman, 1987; Hirsch, 1989; Babloyantz, Nicolis, & Salazar, 1985; Hopfield, 1982). The characteristic time scales of its processes range from just a few milliseconds to several hundred milliseconds, or even longer (Skarda, & Freeman, 1987; Tsuda, 1992; Shepherd, 1990; Yao, & Freeman, 1990). In (Lieberman-Aiden, 2009), it was suggested that when the external input h is large, the neuron quickly reaches an excited state. However, when h is smaller, the transition to this state slows down. This behavior is captured by the chaotic Izhikevich activation function (1), where a decrease in h leads to an increase in $K(h)$, resulting in a longer transition time in (2). Despite this, the special point at $x = 1$ remains stable. Once $K(h)$ surpasses 1, $x = 1$ loses its stability, giving rise to chaotic behavior in the neuron.

The transition from order to chaos, as described in (3), is a phenomenon often observed in functions with discontinuities or sharp changes, referred to as a homoclinic explosion (Sparrow, 1982). Previous research has shown that the specific form of the function $K(h)$ does not significantly affect the outcomes, as long as the system follows this chaotic behavior. The critical parameter to focus on is the value $h = p$, where $K(h)$ equals 1, marking the point at which the shift to chaos occurs.

One drawback of the chaotic Izhikevich activation function mentioned earlier is its lack of differentiability for small values of $h(t)$. Differentiability is a critical property in many machine learning algorithms, particularly those relying on gradient-based optimization methods, such as backpropagation. The absence of differentiability in an activation function can lead to challenges in effectively training neural networks, as gradients are essential for adjusting weights to minimize error during the learning process.

This problem can be addressed by replacing the linear function within the Izhikevich activation model with a parametric Bézier function (see Fig. 1). Bézier functions, widely used in computer graphics for curve modeling, provide a flexible and smooth representation of data. In the context of neural networks, this replacement ensures that the activation functions remain differentiable, even for small values of $h(t)$. The smoothness and control offered by Bézier functions can facilitate more stable gradient computations, enhancing the training efficiency and convergence of neural networks.

Moreover, the parametric nature of Bézier functions introduces additional parameters that can be fine-tuned during the training process. These parameters offer a higher degree of flexibility, allowing the activation function to adapt to specific learning requirements of different layers or architectures. Importantly, this substitution preserves the bijectivity condition of the equation $x(p)$, where x represents the input vector, and p denotes the parameters of the Bézier function. Bijectivity is crucial in ensuring that every input maps uniquely to an output, maintaining the mathematical integrity and predictability of the activation function.

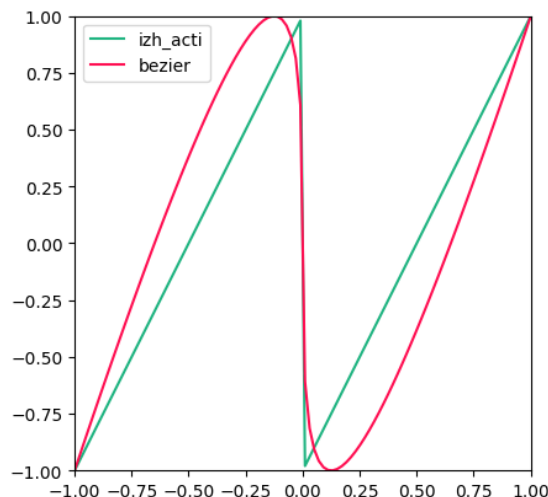


Fig. 1. The Izhikevich chaotic activation function and its altered form utilizing a parametric Bézier curve

The adaptation of parametric Bézier functions in place of traditional activation mechanisms exemplifies the innovative approaches being developed to address the inherent challenges in neural network training. This adjustment not only highlights the interplay between mathematical rigor and computational efficiency but also underscores the importance of interdisciplinary methods in advancing machine learning technologies. An experiment was conducted to evaluate the activation functions used in the Hopfield network, specifically the signum function, the chaotic Izhikevich activation function, and its differentiable variant. The input data comprised samples from normal distributions $N(\mu, \sigma^2)$, representing two distinct classes. The parameters μ and σ^2 were chosen to cover a range of distribution overlap scenarios (see Fig. 2). Training images were created based on the training data for each binary classification category, and the models' performance was measured on the test data using Leave-One-Out validation, with the F1-score as the primary metric.

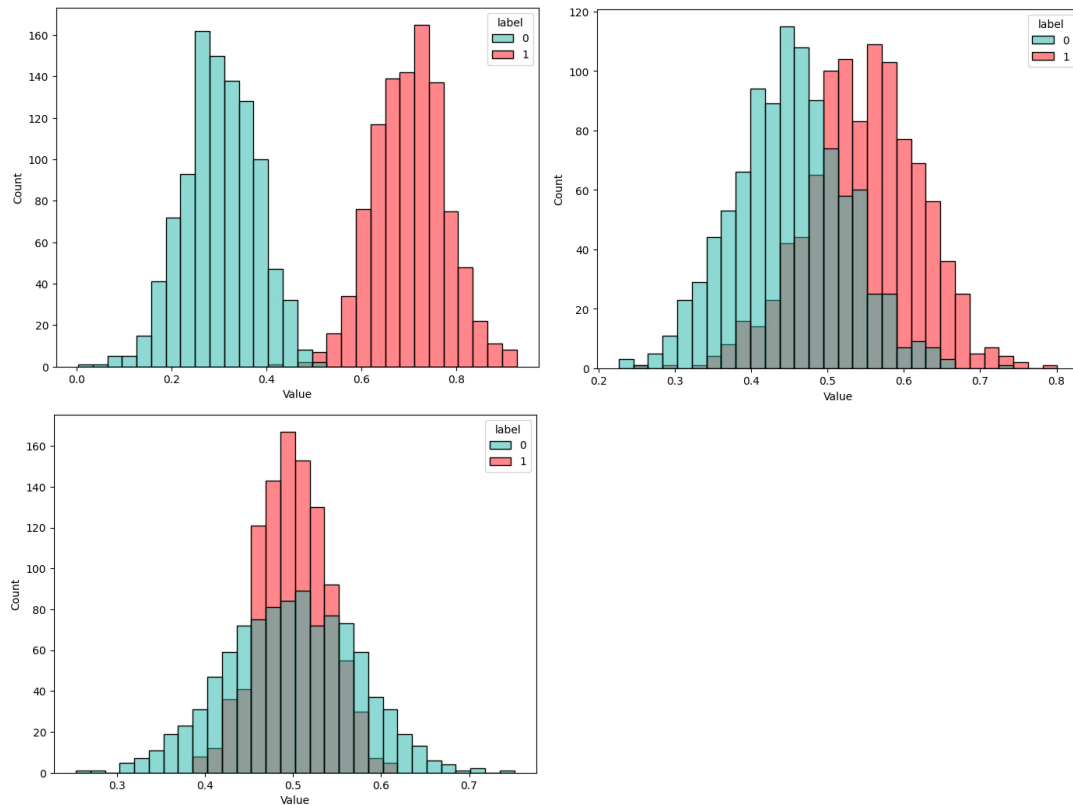


Fig. 2. Samples for binary classification derived from normal distributions with distinct variances and mean values

The null hypothesis posits that the selection of an activation function does not influence the outcomes of the model, whereas the alternative hypothesis asserts that chaotic activation functions provide enhanced results.

When analyzing cases where the distributions intersect, with parameters set as $\mu_1=0.34$, $\mu_2=0.66$, and $\sigma_1^2=\sigma_2^2=0.05$, an average F1-score of 0.79 was recorded after conducting 100 trials with the signum activation function. Through statistical power analysis, it is possible to ascertain the minimum number of experiments necessary to achieve a statistically significant increase in the mean F1-score. To reach statistical significance using a z-test for an improvement of at least 1 %, a minimum of 25,580 experiments per group would be required.

A total of 26,000 experiments were carried out for each of the activation functions under comparison (including signum, chaotic Izhikevich, and its differentiable variant), and the F1-scores were computed for each trial. The average F1-scores obtained were 0.86 for the signum function, 0.87 for the chaotic Izhikevich function, and 0.872 for the differentiable version. Following this, z-tests were conducted to analyze the results for each experimental pair (refer to Fig. 3).

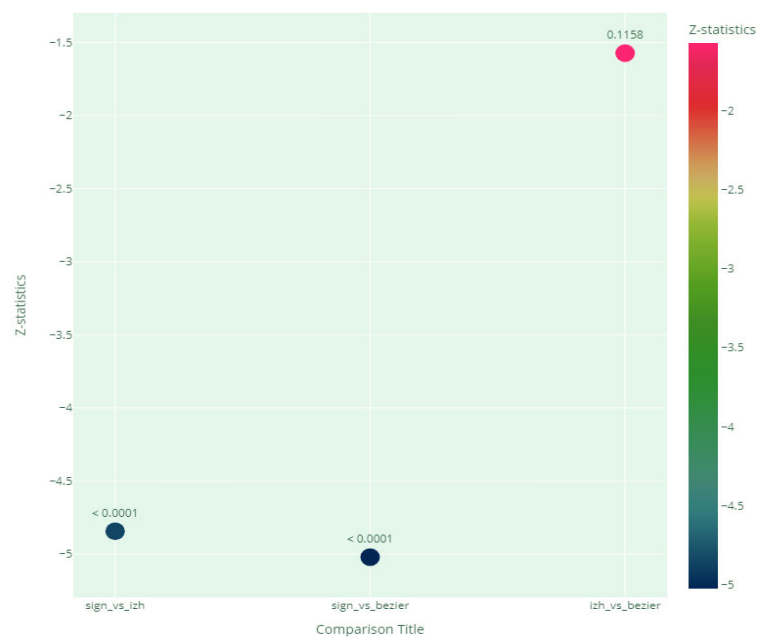


Fig. 3. Results of z-tests for experiments with activation functions

The findings indicated that the samples utilizing chaotic activation functions were statistically significantly different from those using the signum function. However, no statistically significant difference was found between the two chaotic functions, even though the modified function yielded slightly better outcomes, as demonstrated by the z-statistics.

Additional experiments with different parameters of normal distributions were conducted, and F1-scores were determined for each sample, with paired z-tests performed (see Fig. 4).

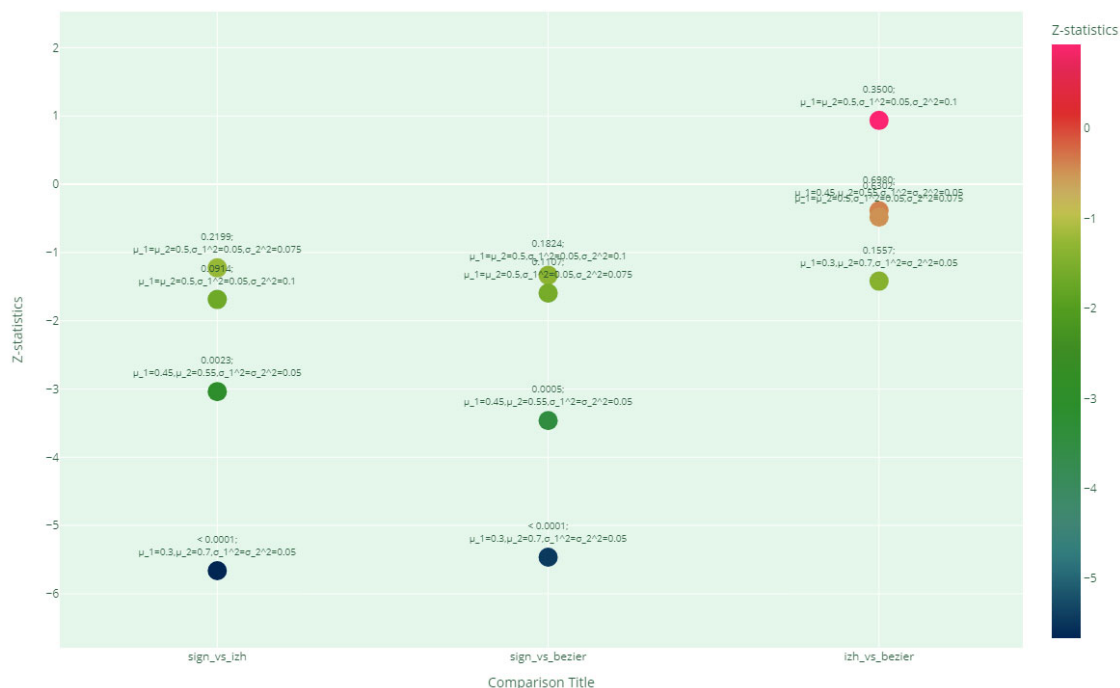


Fig. 4. Results of z-tests for experiments with activation functions for different distributions

In trials where the mean values varied, the previous trend persisted: chaotic functions consistently demonstrated statistically significant improvements over the signum function, with the differentiable modification exhibiting even greater performance. Conversely, in experiments where the means were identical, all functions delivered comparable results, revealing no statistically significant differences.

2. Experiment on real data

We will now investigate the efficacy of chaotic activation functions in an experiment utilizing real-world data.

In the 1960s, some of the earliest studies on malignant transformations emerged, focusing on X-chromatin content in somatic cells. These investigations underscored the instability of X-chromatin and its relationship with various functional changes in the body, as well as the general pathology of somatic cells. Significant alterations in the X-chromatin content of buccal epithelium and peripheral blood neutrophils were noted in individuals with tumors. Furthermore, it was shown that variations in the number of cells harboring X chromosomes are connected to abnormalities in the functional status of the heterozygous X chromosome (Radak, Lafta, & Fallahi, 2023).

Research emphasizing changes in buccal epithelial cells among cancer patients has garnered considerable interest. In the 1960s, H. Nieburgs (1968) and colleagues reported a notable redistribution of chromatin mass in somatic cells of 77 % of cancer patients, labeling these alterations as tumorigenic. These changes included the enlargement of epithelial cell nuclei and an increase in the size of "restricted" chromatin areas, which were surrounded by lighter regions (Boroday et al., 2016). Similar transformations were observed in cells from other organs, such as the liver and kidneys. In breast cancer patients, there was a reported increase in DNA content and the size of interphase nuclei in buccal epithelium. However, some researchers did not observe a significant difference in DNA content between patients and healthy individuals, as demonstrated in a study measuring the DNA content of buccal epithelial cells in men with bronchial epithelioma through cellular spectrophotometry (Ogden, Cowpe, & Green, 1990).

Another study analyzed a control group of 29 patients, 68 patients diagnosed with stage II breast cancer, and 33 patients with fibroadenomatosis, all confirmed through histological examination. The morphological dataset consisted of 20,256 images of interphase nuclei from buccal epithelium, with 6,752 nuclei scanned using three different formats: unfiltered, with a yellow filter, and with a purple filter.

The morphological samples were derived from smears of epithelial cells taken from the middle depth of the spinous layer of the oral mucosa, averaging 52 cells per preparation. The concentration of DNA-fuchsin in the epithelial cell nuclei was calculated by multiplying the optical density by the area. During the initial phase of the analysis, an image of chromatin distribution was generated as a 128 x 128-pixel matrix (Klyushin et al., 2021).

The model's input comprised three-channel (RGB) samples of fractal kernel sizes for each patient, exhibiting considerable variation in the number of elements across the samples. As a result, quantiles were calculated for each sample before training the neural network, where n signifies the number of elements in the smallest sample (refer to Fig. 5).

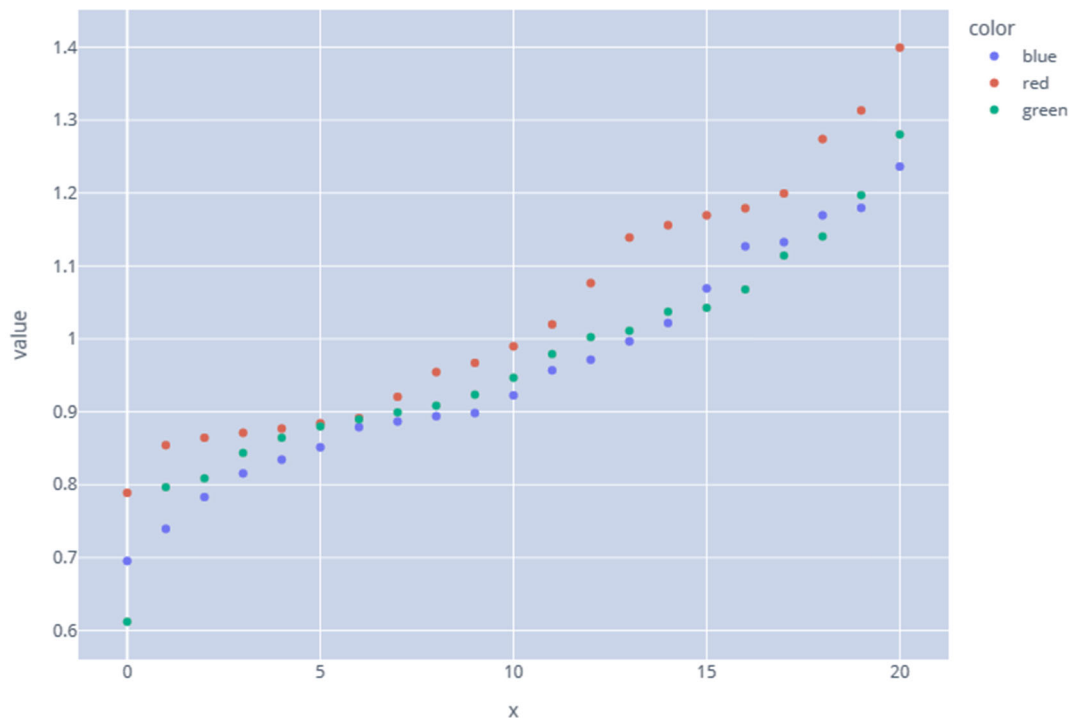


Fig. 5. An example of consolidation of training samples using quantiles

A total of 144 experiments were meticulously conducted, carefully divided into four distinct categories to facilitate a comprehensive analysis: cancer patients compared to healthy controls, cancer patients versus both healthy controls and fibroadenomatosis patients, cancer patients against fibroadenomatosis patients, and fibroadenomatosis patients compared to healthy controls. Each category represented a unique challenge in distinguishing between the various groups, and the experiments were designed to test multiple aspects of the data and model behavior under different conditions.

Training images were generated from the initial training data for each binary classification category, ensuring that the models were exposed to a balanced and varied set of input data. The performance of each model was rigorously evaluated using Leave-One-Out Cross-Validation (LOO-CV) on the test data. This method enabled the analysis of each individual data point's influence on the model's predictions. Key performance metrics, including precision, recall, and F1-score, were utilized to assess how well the models differentiated between the various groups, with a specific focus on achieving high accuracy and minimizing false positives and false negatives.

In the analysis comparing 'cancer patients and fibroadenomatosis patients,' it was observed that the differentiable modification introduced in the models yielded significantly better outcomes. Specifically, the use of the median as an aggregation function outperformed the arithmetic mean, consistently providing a more reliable aggregation of data, especially in the most successful experimental trials. The highest accuracy was recorded when utilizing input data in RGB format, offering the most comprehensive color representation for distinguishing between the patient groups. Additionally, input data formatted with single-channel inputs for red (R), green (G), and blue (B) channels also demonstrated promising results, further emphasizing the importance of selecting the right input format for model performance.

In the comparison between 'cancer patients and healthy controls,' the differentiable modification displayed comparable results to those achieved by the chaotic Izhikevich function, known for its non-linear and complex behavior. In most experiments, both chaotic functions outperformed the traditional signum function, showcasing the potential of chaotic dynamics in improving model predictions. Interestingly, the median and arithmetic mean showed nearly identical performance in related tests, with the best accuracy achieved when the input data was in the B (blue) and RGB formats, suggesting that these color channels were particularly effective in discriminating between cancer patients and healthy controls.

When examining the more complex comparison of 'cancer patients versus healthy controls and fibroadenomatosis patients,' all three activation functions—differentiable, chaotic, and signum—exhibited similar levels of effectiveness. This lack of a clear winner among the activation functions suggests that, in this case, the task of differentiating between the three categories posed a greater challenge for the models, regardless of the function used. Similarly, the aggregation techniques and input data formats did not show any clear advantage, with no single combination yielding significantly better results across all experiments.

In the comparison of 'fibroadenomatosis patients compared to healthy controls,' chaotic functions consistently outperformed the signum function across the majority of experiments. This trend aligns with previous findings and further supports the hypothesis that chaotic systems can enhance model performance in certain types of medical data classification. Moreover, in line with previous analyses, the median was found to be more effective than the arithmetic mean in aggregating training images, leading to better F1-scores in most experiments. The highest accuracy in this category was achieved with input data in both RGB and B formats, reinforcing the consistent success of these input channels across different classification tasks.

The most impressive results were seen in the experiments comparing 'cancer patients to fibroadenomatosis patients,' where the top five trials achieved an average F1-score of 0.84, reflecting a high level of model effectiveness in this binary

comparison. This stands in contrast to the lowest performance, observed in the category comparing 'cancer patients to healthy controls and fibroadenomatosis patients,' where the best five experiments yielded an average F1-score of only 0.57. This significant variation in performance can largely be attributed to the complexity introduced by the second class – healthy controls and fibroadenomatosis patients. The combined dataset for these two groups proved to be especially difficult for the model to differentiate from cancer patients, likely due to shared features or overlapping characteristics between healthy and fibroadenomatosis patients that were not easily separable.

Furthermore, the aggregation method once again played a crucial role in model performance. The median, as an aggregation technique, demonstrated greater effectiveness in most experiments, resulting in higher F1-scores compared to the arithmetic mean. This finding underscores the importance of selecting the right aggregation function to ensure that the training data is represented in a way that maximizes the model's ability to distinguish between patient groups. The input formats that consistently emerged as the most successful across the various experimental groups were RGB and B, both of which regularly appeared among the top five trials in each category, suggesting that these formats offer the most robust and informative data for the classification task at hand.

In summary, the results from these 144 experiments highlight the intricate dynamics of using machine learning models to distinguish between different patient categories based on complex medical data. The findings suggest that differentiable modifications, chaotic functions, and thoughtful data aggregation methods can significantly enhance model performance, especially when carefully paired with the most appropriate input data formats. These insights provide valuable guidance for future research and model optimization in the field of medical data classification.

Discussion and conclusions

Incorporating chaos into activation functions is becoming an increasingly important strategy for improving the efficiency, adaptability, and overall performance of deep learning models, particularly within the context of artificial neural networks (ANNs). Chaotic systems, by their very nature, exhibit complex, sensitive dependence on initial conditions, and by embedding such behavior into activation functions, models can simulate a more dynamic, unpredictable response to varying inputs. This adaptability is particularly beneficial in deep learning, where the capacity to handle diverse and fluctuating data patterns can significantly improve model performance. Traditional activation functions like ReLU or sigmoid, while effective in many scenarios, often struggle to account for the nonlinearities and complexities inherent in real-world datasets, leading to limited model flexibility and suboptimal generalization.

By introducing chaos into activation functions, the feature distribution space within the model is greatly expanded. This means that the neural network has access to a broader spectrum of representational possibilities, allowing it to adjust more effectively to changing or uncertain conditions in the data. Such a mechanism allows the network to better explore and map the relationships within the data, as chaotic functions inherently offer a rich diversity of behaviors, ranging from periodic to highly irregular patterns. This increased diversity enables models to capture more nuanced features and detect intricate patterns that might otherwise be missed by more conventional activation functions. Additionally, chaotic activation functions contribute to making the model more resilient to noise and fluctuations in the training data, as the chaotic behavior introduces a natural robustness to small, random perturbations. This becomes especially important in real-world applications, where data can often be incomplete, noisy, or otherwise unreliable.

Another significant advantage of chaotic activation functions is their ability to improve the model's generalization capabilities, especially when exposed to new, unseen data. When deep learning models are trained, they often face the challenge of overfitting to the training set, which can result in poor performance on out-of-sample data. The unpredictability introduced by chaotic functions helps to mitigate this risk by promoting more flexible learning, allowing the model to explore a wider range of potential solutions during training. This ability to generalize better means that chaotic activation functions enhance the model's ability to adapt to new scenarios, making it more versatile and better equipped to handle a broad variety of machine learning problems, from time series forecasting to pattern recognition.

Moreover, differentiable activation functions in neural networks provide several significant advantages over undifferentiable ones. Primarily, they enable the effective use of optimization algorithms, such as error backpropagation, for network training. The smooth properties of differentiable functions allow for the identification of optimal parameters via gradient descent, thereby streamlining the training process and improving the convergence rate.

This article presents evidence and highlights the advantages of chaotic activation functions within the Hopfield model, demonstrated through experiments with artificial samples generated from normal distributions, as well as real data related to breast cancer prediction. This area still requires further research and experimentation, and the development and implementation of new chaotic activation functions could be a promising direction for advancing deep learning. Such enhancements could benefit not only Hopfield models but also other algorithms and neural network architectures, potentially resulting in improved performance across a range of tasks.

Authors' contribution: Dmitriy Klyushin – conceptualization, methodology; Oleksandr Maistrenko – software, empirical data collection and validation, empirical research, analysis of sources, preparation of literature review or theoretical foundations of research.

Acknowledgments, funding sources. Authors would like to express their sincere gratitude to K. Golubeva and N. Borodai for generously providing the data used for training the models in this study.

References

- Babloyantz, A., Nicolis, S., & Salazar, M. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111, 152–156. [https://doi.org/10.1016/0375-9601\(85\)90444-X](https://doi.org/10.1016/0375-9601(85)90444-X)
- Boroday, N., Chekhun, V., Golubeva, E., & Klyushin, D. (2016). In vitro and in vivo densitometric analysis of DNA content and chromatin texture in tumor cell nuclei under the influence of a nano composite and magnetic field. *Advances in Cancer Research & Treatment*, 1–11. <https://doi.org/10.5171/2016.706183>
- Chen, X., & Wang, Y. (2023). A Chaotic Neuron and its Ability to Prevent Overfitting. *Frontiers in Computing and Intelligent Systems*, 5, 53–61. <https://doi.org/10.54097/fcis.v5i1.11673>
- Cursino, C., & Dias, L. (2024). Hybrid Hopfield Neural Network. *SN Computer Science*, 5, 25–74. <https://doi.org/10.1007/s42979-023-02575-6>
- Deng, Q., Wong, C., & Lin, H. (2024a). Chaotic dynamical system of Hopfield neural network influenced by neuron activation threshold and its image encryption. *Nonlinear Dynamics*, 112, 1–18. <https://doi.org/10.1007/s11071-024-09384-3>

- Deng, Q., Wong, C., & Lin, H. (2024b). Memristive Hopfield neural network dynamics with heterogeneous activation functions and its application. *Chaos Solitons & Fractals*, 178, 101–124. <https://doi.org/10.1016/j.chaos.2023.114387>
- Dubey, S. R., Singh, S. K., & Chaudhuri B. B. (2022). Activation Functions in Deep Learning: A comprehensive Survey and Benchmark. *Neurocomputing*, 503, 24–27. <https://doi.org/10.1016/j.neucom.2022.06.111>
- Freeman, W. (1987). Simulation of Chaotic EEG Patterns with a Dynamic Model of the Olfactory System. *Biological Cybernetics*, 56, 139–150. <https://doi.org/10.1007/BF00317988>
- Hirsch, M. (1989). Convergent Activation Dynamics in Continuous Time Neural Networks. *Neural Networks*, 2, 331–351. [https://doi.org/10.1016/0893-6080\(89\)90018-X](https://doi.org/10.1016/0893-6080(89)90018-X)
- Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A.*, 79 (Apr), 2554–2558. <https://doi.org/10.1073/pnas.79.8.2554>
- Kashyap, D., Dsouza, N., Shi, L., & Beymer, D. (2022). Hopfield Encoding Networks. <https://doi.org/10.13140/RG.2.2.17209.65123>
- Kilicarslan, S., Adem, K., & Celik M. (2021). An overview of the activation functions used in deep learning algorithms. *Journal of New Results in Science*, 10, 75–88. <https://doi.org/10.54187/jnrs.1011739>
- Klyushin, D., Golubeva, K., Boroday, N., & Shervarly, D. (2021). Breast Cancer Diagnosis Using Machine Learning and Fractal Analysis of Malignancy-Associated Changes in Buccal Epithelium. *Artificial Intelligence, Machine Learning, and Data Science Technologies Future Impact and Well-Being for Society*, 5, 1–18. <https://doi.org/10.1201/9781003153405-1>
- Krotov, D. (2023). A new frontier for Hopfield networks. *Nature Reviews Physics*, 5, 52–70. <https://doi.org/10.1038/s42254-023-00595-y>
- Liang, X., Yang, Y., Wang, R., & Chen, J. (2024). Synchronization of delayed stochastic reaction-diffusion Hopfield neural networks via sliding mode control. *Nonlinear Analysis Modelling and Control*, 29, 1–19. <https://doi.org/10.15388/namc.2024.29.34884>
- Lieberman-Aiden, E. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human Genome. *Science*, 326, 289–193. <https://doi.org/10.1126/science.1181369>
- Magallon, D., Garcia, J., Huerta, G., & Jaimes, R. (2024). Real-time neural identification using a recurrent wavelet first-order neural network of a chaotic system implemented in an FPAA. *Integration*, 96, 384–393. <https://doi.org/10.1016/j.vlsi.2023.102134>
- Nieburgs, H. (1968). Recent progress in the interpretation of malignancy associated changes (MAC). *Acta Cytologica*, 12, 445–453. <https://doi.org/10.1155/238921>
- Ogden, G., Cowpe, J., & Green, M. (1990). The effect of distant malignancy upon quantitative cytologic assessment of normal oral mucosa. *Cancer*, 65, 477–480. [https://doi.org/10.1002/1097-0142\(19900201\)65:3<477::AID-CNCR2820650317>3.0.CO;2-G](https://doi.org/10.1002/1097-0142(19900201)65:3<477::AID-CNCR2820650317>3.0.CO;2-G)
- Radak, M., Lafta, H., & Fallah, H. (2023). Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies. *Journal of Cancer Research and Clinical Oncology*, 49, 10473–10491. <https://doi.org/10.1007/s00432-023-04956-z>
- Shepherd, G. (1990). *The synaptic organization of the brain*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195159561.001.1>
- Skarda, C., & Freeman, W. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10, 161–195. <https://doi.org/10.1017/S0140525X00047336>
- Sparrow, C. (1982). The Lorenz Equations. Bifurcations, Chaos, and Strange Attractors. *Applied Mathematical Sciences*, 41, 26–50. <https://doi.org/10.1007/978-1-4612-5767-7>
- Tsuda, I. (1992). Dynamic link of memory – Chaotic memory map in nonequilibrium neural networks. *Neural Networks*, 5, 857–893. [https://doi.org/10.1016/S0893-6080\(05\)80029-2](https://doi.org/10.1016/S0893-6080(05)80029-2)
- Wang, X., Ren H., & Wang A. (2022). Smish: A Novel Activation Function for Deep Learning Methods. *Electronics*, 11, 540–545. <https://doi.org/10.3390/electronics11040540>
- Yao, Y., & Freeman, W. (1990). Model of Biological Pattern Recognition with Spatially Chaotic Dynamics. *Neural Networks*, 3, 153–170. [https://doi.org/10.1016/0893-6080\(90\)90086-Z](https://doi.org/10.1016/0893-6080(90)90086-Z)

Отримано редакцією журналу / Received: 02.11.24

Прорецензовано / Revised: 06.03.25

Схвалено до друку / Accepted: 12.03.25

Дмитро КЛЮШИН, д-р фіз.-мат. наук, проф.

ORCID ID: 0000-0003-4554-1049

e-mail: dmytroklyushin@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

Олександр МАЙСТРЕНКО, асп.

ORCID ID: 0009-0003-3464-7777

e-mail: o.maistrenko@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

ВИКОРИСТАННЯ НЕЙРОНІВ ІЖКЕВИЧА В МОДЕЛЯХ ГОПФІЛДА

Побудова моделі хаотичного нейрона Іжкєвича є значним проривом у галузі обчислювальної нейронауки, оскільки надає цьому явищу математичну основу, що наближено відтворює поведінку біологічних нейронів, зокрема складні хаотичні спайкові патерни, які спостерігаються в мозку. Ця модель привернула увагу вчених завдяки своїй здатності відтворювати широкий спектр реальних нейронних динамік, включно зі спалахами та тонічними спайками, що є фундаментальним для розуміння складності нейронної комунікації й оброблення інформації. З іншого боку, мережі Хопфілда, один із типів рекурентних нейронних мереж, давно визнані за їхню здатність функціонувати як системи пам'яті з доступом за змістом, зберігаючи та викликаючи інформацію на основі асоціативної динаміки. Часто описані як спінові скляні системи, мережі Хопфілда функціонують завдяки знаходженню стабільних станів або патернів у межах нейронної мережі, імітуючи певні функції пам'яті людського мозку. Нещодавні дослідження в галузі інноваційних активаційних функцій відкрили нові можливості для вдосконалення характеристик рекурентних мереж. Зокрема, хаотичні активаційні функції є цікавим напрямком для досліджень у мережах Хопфілда. У пропонуваній статті проаналізовано вплив упровадження цих хаотичних активаційних функцій у мережі Хопфілда, зокрема те, як їхня дія відображається на стабільності, адаптивності й ефективності роботи мережі. Мета роботи – розкрити вплив хаосу на динаміку мережі, надаючи нові уявлення, які потенційно можуть підвищити продуктивність у застосуваннях, що потребують складної пам'яті й асоціативної обробки. Це дослідження робить вагомий внесок у розвиток нейроморфної інженерії та має важливе значення для штучного інтелекту й нейронауки.

Ключові слова: нейрон Іжкєвича, мережа Гопфілда, хаотичні активаційні функції, нейронні мережі.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 004.94:519.6

DOI: <https://doi.org/10.17721/1812-5409.2025/1.17>

Igor MELNYK¹, DSc (Engin.), Prof.
ORCID ID: 0000-0003-0220-0615
e-mail: imelnik@phbme.kpi.ua

Alina POCHYNOK¹, PhD (Engin.), Senior Researcher
ORCID ID: 0000-0001-9531-7593
e-mail: alina_pochynok@yahoo.com

Mykhailo SKRYPKA¹, PhD Student
ORCID ID: 0009-0006-7142-5569
e-mail: scientetik@gmail.com

Olga DEMYANCHENKO², PhD (Engin.), Senior Lecturer
ORCID ID: 0000-0002-4693-0364
e-mail: olgademyanchenko@gmail.com

¹National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

²National University "Lviv Polytechnic", Lviv, Ukraine

METHOD OF INTERPOLATION USING ROOT-FRACTIONAL-RATIONAL FUNCTIONS OF DIFFERENT ORDERS

The possibility of interpolating different kinds of mathematical functions using different order root-fractional-rational functions, namely, second, third, and fourth, is considered in the article. Generally, it is shown that root-fractional-rational functions are given a precision interpolation with a small value of error even for stiff analytical dependences. Root-fractional-rational functions from second to fourth order are considered, and corresponding analytical relations for defining polynomial coefficients in the nominator and the denominator are given. It is also proven that the number of necessary points for interpolation corresponds to the value $2n + 1$, where n is the order of the root-fractional-rational function. Examples of interpolation of different functions for electrodynamics problems, simulation of magnetic lenses, probability tasks, and fuzzy-logic tasks are given; the error of interpolation for all considered examples is also defined. All presumptions of theoretical analysis are tested and verified using the original elaborated computer software, created using the Python programming language means. In the most considered examples, the resulting error of interpolation is smaller than a few percent. The graphic results of testing the proposed method of interpolation are also given.

Keywords: interpolation; root-fractional-rational function; electrodynamic tasks; functions of probability theory; fuzzy logic; membership functions.

AMS 2020 classification: 26A06, 26A09.

Introduction

Today, interpolation methods are widely used to solve complex problems in computational mathematics, applied physics, and mathematical statistics, to simplify computational algorithms, and to effectively process large amounts of numerical information. But the general problem with using this approach is that the accuracy of describing the behavior of rigid functions using polynomial dependencies is usually not adequate, since polynomial functions often produce large outliers with extrema in smooth sections of the interpolated functional dependency. In this case, the interpolation accuracy can be very low, and the difference between the original data and the interpolation result exceeds hundreds of percent (Chapra, & Canale, 2014; Jain, M., Iengar, & Jain, R., 2010; Mathews, & Fink, 1998; Bronshtein et al., 2007; Phillips, 2023; Epperson, 2013). Moreover, under this condition, even a qualitative description of a set of numerical data using polynomial interpolation methods is incorrect. In the classical theory of interpolation, this problem is solved using splines through interpolation on small segments of the function argument using known derived values, but the general disadvantage of this approach is that different functions are used in different sections of the argument change, and the number of such sections can be several thousand and even be more (Chapra, & Canale, 2014; Jain, M., Iengar, & Jain, R., 2010; Mathews, & Fink, 1998; Bronshtein et al., 2007; Phillips, 2023; Epperson, 2013).

In the years 2019–2020 in paper (Melnyk, & Pochynok, 2020), a slightly different approach to the interpolation of data sets of stiff functions that are described as ravine dependencies associated with the use of root-polynomial functions has been proposed. Initially, the proposed approach has been applied to interpolating the boundary trajectories of electron beams during their transportation in an ionized gas. In paper (Melnyk, & Pochynok, 2020) has been proposed the mathematical relation for root-polynomial functions, which is generally written as:

$$r(z) = \sqrt[n]{C_n z^n + C_{n-1} z^{n-1} + \dots + C_1 z + C_0}, \quad (1)$$

where z is the longitudinal coordinate, r is the radius of the boundary trajectory of the beam, n is the degree of the polynomial and the order of the root function, $C_0 - C_n$ is the value of the polynomial coefficients.

Research carried out and its results, reflected in paper (Melnyk, & Pochynok, 2020), showed that usually root polynomial functions (1) give an extremely small error in the interpolation of the boundary trajectories of the electron beams, no more than a fraction of a percent. Also, in paper (Melnyk, & Pochynok, 2020), simple analytical relations have been given for calculating the coefficients of root polynomial functions C_n for values of the order of the function n from 2 to 6. Subsequently, in paper (Melnyk, & Pochynok, 2023), algorithms for approximating ravine dependencies using root-polynomial functions also have been considered. The paper (Melnyk, & Pochynok, 2023) considered methods of approximation by the reference points, and by the tangents to ravine functions corresponded to a given set of reference points.

However, studies carried out in (Melnyk, & Pochynok, 2023) showed that not even for all ravine and vertex functions, the numerical interpolation algorithms converge, considered in (Melnyk, & Pochynok, 2020). It was justified that for ravine

© Melnyk Igor, Pochynok Alina, Skrypka Mykhailo, Demyanchenko Olga, 2025

functions, such algorithms converge in the case of dependencies with a narrow minimum region, and for functions with a plateau, especially in the case when the values of the numerical data are close to zero, the interpolation algorithms diverge. On the contrary, for vertex functions, the convergence of the developed algorithms is guaranteed in the case of a wide maximum area, that is, for functions with a plateau. It was also shown in (Melnyk, & Pochynok, 2023) that this problem of algorithms divergence can be avoided by using the method of mirroring the function values relative to the line corresponding to their average value along the ordinate axis. Then the corresponding values of $r(z)_{\text{mir}}$ are calculated through the following relation (Melnyk, & Pochynok, 2023):

$$r(z)_{\text{mir}} = \frac{r(z)_{\text{max}} + r(z)_{\text{min}}}{2} - \sqrt[n]{C_n z^n + C_{n-1} z^{n-1} + \dots + C_1 z + C_0}, \quad (2)$$

where $r(z)_{\text{max}}$ – maximal, and $r(z)_{\text{min}}$ – minimal values of function $r(z)$ for considered digital data set.

However, in papers (Melnyk, & Pochynok, 2023), it was shown that using the root polynomial function (1) from the second to the sixth order, it is possible to solve the problem of interpolation of the boundary trajectory of a short-focus electron beam propagated in an ionized gas with a very small error, less than a fraction of a percent. Solving this problem is extremely important from a practical point of view for analyzing the parameters of electron beams that are formed by high-voltage glow discharge electron guns.

Electron guns of this type have a number of advantages over traditional guns with hot cathodes (Melnyk, 2005; Melnyk, 2006) and today they are widely used in modern welding technologies (Melnyk, 2006; Melnyk et al., 2022; Druzhinin et al., 2015), the deposition of ceramic coatings for various purposes (Prikhna et al., 2022a; Grechanyuk et al., 2022; Zakharov et al., 2020; Zakharov, Rozenko, & Ilchenko, 2019), and the melting of refractory metals and non-metallic materials in order to purify them and improve their quality (Akhonin et al., 2018; Paton, Akhonin, & Berezos, 2018; Kemmotsu, Nagai, & Maeda, 2011). The main industries where high-voltage glow discharge electron guns are effectively used today are the electronics industry (Druzhinin et al., 2014; Druzhinin et al., 2015; Zakharov et al., 2020; Zakharov, Rozenko, & Ilchenko, 2019; Kemmotsu, Nagai, & Maeda, 2011), instrument making (Melnyk et al., 2022; Zakharov et al., 2020; Zakharov, Rozenko, & Ilchenko, 2019), mechanical engineering, production of high-energy equipment and electric vehicles, automotive, aviation, and space industries (Prikhna et al., 2022b). In general, methods for simulating electron beams formed in gas-discharge devices with gas ionization and beam interaction with ionized plasma are based on the basic principles of the theory of gas discharge (Melnyk, 2005; Smirnov, 2015; Lieberman, & Lichtenberg, 1994; Raizer, 1991).

There are also other approaches to the simulation of gas-discharge devices that form intensive electron beams for technological purposes. In particular, the works (Tselyuk et al., 2008; Rudychev, V., Lazurik, & Rudychev, Y., 2021) describe algorithms for the analysis of high-energy relativistic electron beams that are formed in ionized plasma. These methods are based on direct measurement of beam parameters, analysis of experimental data, and their subsequent statistical processing using modern computer technologies and suitable algorithms.

Thus, previous studies have shown that, in any case, the class of functions that are subject to interpolation and approximation using relations (1) and (2) is extremely limited; usually, these can only be ravine and vertex functions with a small, stiff degree. And even for the class of such functions, the simple numerical algorithms proposed and described in (Melnyk, & Pochynok, 2020; Melnyk, & Pochynok, 2023) often diverge. Therefore, the aim of this work is the development and software implementation of an interpolation algorithm for another, more wide class of interpolation functions, in particular root-fractional-rational ones.

1. Analysis of the results of previous research and general formulation of the problem

As substantiated in papers (Melnyk, & Pochynok, 2023), the convergence of interpolation methods based on the use of relations (1, 2) significantly depends on the maximum value of the derivative for the set of interpolated data and is not always guaranteed. A significant limitation is also that such interpolation methods are only suitable for ravine and vertex numerical dependencies with one global extremum. For ravine functions with a narrow minimum region, if their values in this region tend to zero, the proposed algorithms for interpolation by root polynomial functions also diverge.

Therefore, the aim of this paper is to analyze the possibilities of using other root-fractional-rational functions to solve interpolation problems.

In the general case, one can write the root-fraction-rational function as follows:

$$r(z) = \sqrt[n]{\frac{z^n + C_1 z^{n-1} + \dots + C_{n-1} z + C_n}{C_{n+1} z^n + \dots + C_{2n} z + C_{2n+1}}}, \quad (3)$$

where $C_1 \dots C_{2n+1}$ – the corresponded values of coefficients of root-fraction-rational function.

The provided research shown, that the root-fractional-rational functions written in the form (3) can be used to solve a wide range of practical engineering problems, in particular.

1. To interpolate the magnetic field of focusing lenses in tasks of electron optics (Schiller, Heisig, & Panzer, 1982; Szilagy, 2012; Humphries, 2013; Kasper, & Hawkes, 2018; Grivet, 1972).

2. To interpolate the boundary trajectory of electron beams in the case of their transportation in equipotential channels with a change in pressure along the transportation channel. From the point of view of the physical formulation and the numerical algorithms used, such a problem is considered as a problem of nonlinear electron optics (Schiller, Heisig, & Panzer, 1982; Szilagy, 2012; Humphries, 2013; Kasper, & Hawkes, 2018; Grivet, 1972; El-Kareh, 2012).

3. Interpolation of the values of probability distribution and probability density functions in problems of probability theory and mathematical statistics. The main feature here is the high degree of stiffness of such functions and the impossibility of obtaining an analytical expression for the inverse function in a simple algebraic form (Wentzel, & Ovcharov, 2002; Gubner, 2006).

4. Interpolation of membership function values in fuzzy logic problems. Typically, such functions are specified as piecewise continuous on separate intervals of argument values since they do not have an analytical description over the entire interval of numerical values (Zadeh, & Aliev, 2018). Under this condition, to describe the membership functions over the entire numerical interval of values of the argument, one usually uses either a piecewise continuous description of functions on finite numerical intervals of the argument according to given conditions or corresponding arithmetic-logical expressions (Melnyk, & Luntovskyy, 2022).

2. Main obtained results and analytical relations for determining the coefficients of root-fractional-rational functions

Let us assume that a vector of reference points along the coordinate z $\mathbf{V}_z = [z_1, z_2, \dots, z_{2n}, z_{2n+1}]$ and a vector of function values $r(z)$ $\mathbf{V}_r = [r_1, r_2, \dots, r_{2n}, r_{2n+1}]$, which are correspond to these points, are given. Under this condition, the determination of the coefficients $C_1 \dots C_{2n+1}$ of the root-fractional-rational function (3) is carried out by analytically solving the corresponding set of linear equations:

$$\begin{aligned} C_1 z_j^{n-1} + \dots + C_{n-1} z_j + C_n - \left(r_j(z_j) \right)^n \times (C_{n+1} z_j^n + \dots + C_{2n} z_j^n + C_{2n+1}) = \\ = \sum_{i=1}^n C_i z_j^i - \left(r_j(z_j) \right)^n \sum_{i=n+1}^{2n+1} C_i z_j^i = -z_j^n; j \in N, j \leq 2n+1. \end{aligned} \quad (4)$$

A generalized technique for the analytical solution of sets of linear equations of this type was described in papers (Melnik, & Pochynok, 2020). The set of equations is solved using the classical Gauss–Seidel method with the sequential elimination of the C_i variables, but to order at each iteration the coefficients k in front of the C_i variables and the coefficients b on the right side of the equations, such a simple ordered indexing system is used. Let us write $k_{m,j,l}$, and $b_{k,j}$, where m is the number of the current iteration, j is the number of the equation according to formula (4), and l is the number of coefficient C according to formula (3). Thus, the first values of the coefficients $k_{1,j,l}$ and $b_{1,j}$ are specified by the coordinates of points from the vectors $\mathbf{V}_z$ and $\mathbf{V}_r$, and the values of these coefficients at the next iterations are calculated iteratively as follows:

$$k_{m,j,l} = \frac{k_{m-1,j,l} - k_{m-1,j_e,l_e}}{k_{m-1,j_e,l} - k_{m-1,j_e,l_e}}, b_{m,j} = \frac{b_{m-1,j} - b_{m-1,j_e}}{k_{m-1,j_e,l} - k_{m-1,j_e,l_e}}, \quad (5)$$

where j_e – number of the equation that is subtracted from other equations of the set and, thus, excluded from consideration in subsequent iterations, and l_e – the number of the coefficient C , which is excluded from the set of equations and which, in connection with this, is not considered at subsequent iterations.

After calculating all the values of the intermediate coefficients $k_{1,j,l}$ and $b_{1,j}$, the coefficients C that are being sought are found through iterative relations (5), using them sequentially in reverse order.

From a computational point of view, the main feature of iterative relations (5) is that they contain only simple arithmetic operations of subtraction and division, which the computer processor can easily perform at the hardware level (Melnik, & Luntovskyy, 2022). An important advantage of the proposed algorithm is also that calculations using formula (5) for different values of j and l are carried out using the same ratios, so they can be performed by parallel computing. These features of the proposed algorithm greatly simplify the writing of program code, especially with the use of modern declarative, functional, structural, and matrix programming tools (Melnik, & Luntovskyy, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013). When using parallel computing, the speed of solving the interpolation problem can be increased in proportion to the number of generated parallel threads (Melnik, & Luntovskyy, 2022).

During the provided research and studies, the following iterative analytical relations were obtained to search for the coefficients of root-fractional rational functions of the second, third, and fourth orders. For example, for second order function:

$$\begin{aligned} r(z) = \sqrt{\frac{z^2 + C_1 z + C_2}{C_3 z^2 + C_4 z + C_5}}, \text{ the following iterative relations have been obtained:} \\ \begin{aligned} k_{1,1,1} &= z_1 - z_5; k_{1,1,3} = (r_5 z_5)^2 - (r_1 z_1)^2; k_{1,1,4} = r_5^2 z_5 - r_1^2 z_1; k_{1,1,5} = r_5^2 - r_1^2; b_{1,1} = z_5^2 - z_1^2; \\ k_{1,2,1} &= z_2 - z_5; k_{1,2,3} = (r_5 z_5)^2 - (r_2 z_2)^2; k_{1,2,4} = r_5^2 z_5 - r_2^2 z_2; k_{1,2,5} = r_5^2 - r_2^2; b_{1,2} = z_5^2 - z_2^2; \\ k_{1,3,1} &= z_3 - z_5; k_{1,3,3} = (r_5 z_5)^2 - (r_3 z_3)^2; k_{1,3,4} = r_5^2 z_5 - r_3^2 z_3; k_{1,3,5} = r_5^2 - r_3^2; b_{1,3} = z_5^2 - z_3^2; \\ k_{1,4,1} &= z_4 - z_5; k_{1,4,3} = (r_5 z_5)^2 - (r_4 z_4)^2; k_{1,4,4} = r_5^2 z_5 - r_4^2 z_4; k_{1,4,5} = r_5^2 - r_4^2; b_{1,4} = z_5^2 - z_4^2; \\ k_{2,1,1} &= \frac{k_{1,1,1}}{k_{1,1,5}}; k_{2,1,3} = \frac{k_{1,1,3}}{k_{1,1,5}}; k_{2,1,4} = \frac{k_{1,1,4}}{k_{1,1,5}}; b_{2,1} = \frac{b_{1,1}}{k_{1,1,5}}; k_{2,2,1} = \frac{k_{1,2,1}}{k_{1,2,5}}; k_{2,2,3} = \frac{k_{1,2,3}}{k_{1,2,5}}; k_{2,2,4} = \frac{k_{1,2,4}}{k_{1,2,5}}; b_{2,2} = \frac{b_{1,2}}{k_{1,2,5}}; \\ k_{2,3,1} &= \frac{k_{1,3,1}}{k_{1,3,5}}; k_{2,3,3} = \frac{k_{1,3,3}}{k_{1,3,5}}; k_{2,3,4} = \frac{k_{1,3,4}}{k_{1,3,5}}; b_{2,3} = \frac{b_{1,3}}{k_{1,3,5}}; k_{2,4,1} = \frac{k_{1,4,1}}{k_{1,4,5}}; k_{2,4,3} = \frac{k_{1,4,3}}{k_{1,4,5}}; k_{2,4,4} = \frac{k_{1,4,4}}{k_{1,4,5}}; b_{2,4} = \frac{b_{1,4}}{k_{1,4,5}}; \\ k_{3,1,1} &= \frac{k_{2,1,1} - k_{2,4,1}}{k_{2,1,4} - k_{2,4,4}}; k_{3,1,3} = \frac{k_{2,1,3} - k_{2,4,3}}{k_{2,1,4} - k_{2,4,4}}; b_{3,1} = \frac{b_{2,1} - b_{2,4}}{k_{2,1,4} - k_{2,4,4}}; k_{3,2,1} = \frac{k_{2,2,1} - k_{2,4,1}}{k_{2,2,4} - k_{2,4,4}}; k_{3,2,3} = \frac{k_{2,2,3} - k_{2,4,3}}{k_{2,2,4} - k_{2,4,4}}; \\ b_{3,2} &= \frac{b_{2,2} - b_{2,4}}{k_{2,2,4} - k_{2,4,4}}; k_{3,3,1} = \frac{k_{2,3,1} - k_{2,4,1}}{k_{2,3,4} - k_{2,4,4}}; k_{3,3,3} = \frac{k_{2,3,3} - k_{2,4,3}}{k_{2,3,4} - k_{2,4,4}}; b_{3,3} = \frac{b_{2,3} - b_{2,4}}{k_{2,3,4} - k_{2,4,4}}; k_{4,1,1} = \frac{k_{3,1,1} - k_{3,3,1}}{k_{3,1,3} - k_{3,3,3}}; \\ b_{4,1} &= \frac{b_{3,1} - b_{3,3}}{k_{3,1,3} - k_{3,3,3}}; k_{4,2,1} = \frac{k_{3,2,1} - k_{3,3,1}}{k_{3,2,3} - k_{3,3,3}}; b_{4,2} = \frac{b_{3,2} - b_{3,3}}{k_{3,2,3} - k_{3,3,3}}; C_1 = \frac{b_{4,1} - b_{4,2}}{k_{4,1,1} - k_{4,2,1}}; C_3 = b_{4,1} - k_{4,1,1} \frac{b_{4,1} - b_{4,2}}{k_{4,1,1} - k_{4,2,1}}; \\ C_4 &= b_{3,1} - k_{3,1,1} \frac{b_{4,1} - b_{4,2}}{k_{4,1,1} - k_{4,2,1}} - k_{3,1,3} \left(b_{4,1} - k_{4,1,1} \frac{b_{4,1} - b_{4,2}}{k_{4,1,1} - k_{4,2,1}} \right); C_5 = b_{2,1} - k_{2,1,1} C_1 - k_{2,1,3} C_3 - k_{2,1,4} C_4; \\ C_2 &= r_1^2 (z_1^2 C_3 + z_1 C_4 + C_5) - z_1 C_2 - z_1^2. \end{aligned} \end{aligned} \quad (6)$$

The relations, obtained for third and fourth-order root-fractional rational functions (3), are generally similar, but sophisticated. Let's consider now typical examples of using such relations to solve practical engineering problems related to the interpolation of numerical data.

It should be also pointed out, that convergence of proposed algorithm and error of calculation is generally depend on position of base points. The divergence is really possible, when the values of interpolated dataset in the region of minimum is closed to 0 because of possibility of negative values under the root. The condition $r_i(z) > 0$ is necessary to convergence the proposed algorithm, but in the case $r_i(z) \approx 0$ the divergence is possible, because the value $\frac{z^n + C_1 z^{n-1} + \dots + C_{n-1} z + C_n}{C_{n+1} z^n + \dots + C_{2n} z + C_{2n+1}}$ under the root can be negative. Corresponded example for homomorphic function will be considered in the next part of the paper. In this case the best solution is using of positive deviation for interpolated dataset. As for coefficients $C_1 \dots C_{2n+1}$, with the positive values of r they are always calculated correct and additional limitation doesn't need. The approach, described here is similar, as for root-polynomial functions, which has been considered carefully on the mathematical point of view in papers (Melnyk, & Pochynok, 2020; Melnyk, & Pochynok, 2023).

3. Test examples for solving interpolation tasks using root-fractional-rational functions

3.1. First example

It is well known that the axial distribution of the magnetic field of a short focusing lens can be calculated with fairly high accuracy using the Galejs model (Schiller, Heisig, & Panzer, 1982; Szilagyi, 2012; Humphries, 2013; Kasper, & Hawkes, 2018; Grivet, 1972):

$$B_{z0} = \frac{1,26 \cdot 10^{-6} I_l N_l}{2 S_l \pi} \left(\frac{(z_l - P_l) + \frac{S_l}{2}}{\sqrt{\left(\frac{D_l}{3}\right)^2 + \left((z_l - P_l) + \frac{S_l}{2}\right)^2}} - \frac{(z_l - P_l) - \frac{S_l}{2}}{\sqrt{\left(\frac{D_l}{3}\right)^2 + \left((z_l - P_l) - \frac{S_l}{2}\right)^2}} \right), \quad (7)$$

where B_{z0} – magnetic field induction, I_l – lens current, N_l – number of turns, z_l – lens thickness, R_l – winding radius relative to the axis of symmetry of the lens, D_l – winding diameter, $D_l = 2R_l$, S_l – non-magnetic gap size, P_l – position of the lens symmetry plane along the longitudinal coordinate z relative to the point $z = 0$. The corresponding geometric parameters of the short focusing magnetic lens are shown in Fig. 1.

Interpolate the Galejs function (9) with a fourth-order root-fractional rational function and determine the coefficients of this function $C_0 - C_9$ for the following parameters of the lens model: $I_l = 3$ A, $N_l = 1500$, $z_l = 40$ mm, $R_l = 60$ mm, $D_l = 40$ mm, $S_l = 20$ mm, $P_l = 0,1$ m. Calculate the interpolation error, determine the maximum and average value of the interpolation error in the interval $[0,1$ m; $0,13$ m].

The result of the interpolation of dependence (7) for the given parameters of the focusing lens by the fourth-order root-fraction-rational function of is shown in Fig. 2. The position of the basic calculation points $z_1 - z_9$ has been chosen as follows: $z_1 = 0,1$ m, $z_2 = 0,1015$ m, $z_3 = 0,103$ m, $z_4 = 0,1075$ m, $z_5 = 0,112$ m, $z_6 = 0,115$ m, $z_7 = 0,121$ m, $z_8 = 0,127$ m, $z_9 = 0,129$ m. The corresponding interpolation results are shown in Fig. 2. From the presented results it is clear that the interpolation error relative to the basic calculated values of the Galejs function (9) does not exceed 0.125%. Under this condition, the relative average error in the interval $[0,1$ m; $0,13$ m] is 0.0143%.

The coefficients of the fourth-order root-fractional-rational function, calculated from relations (8), for this example have the following values:

$$C_1 = -0.47038 \text{ m}; C_2 = 0.08212 \text{ m}^2; C_3 = -6.2884 \cdot 10^{-3} \text{ m}^3; C_4 = -1.7746 \cdot 10^{-4} \text{ m}^4; C_5 = -4.6373 \cdot 10^4 \text{ m}^{-4}; \\ C_6 = 1.9268 \cdot 10^4 \text{ m}^{-3}; C_7 = -3.04108 \cdot 10^3 \text{ m}^{-2}; C_8 = 216.733 \text{ m}^{-1}; C_9 = -5.91715.$$

3.2. Second example

Describe the error integral $\text{erf}(z)$ (Wentzel, & Ovcharov, 2002; Gubner, 2006) by a second-order root-fractional-rational function. Interpolate using the following reference points: $z_1 = 0,01$; $z_2 = 0,5$; $z_3 = 1,1$; $z_4 = 1,9$; $z_5 = 3$.

The interpolation result, has been obtained for the selected parameters of error function is shown in Fig. 3.

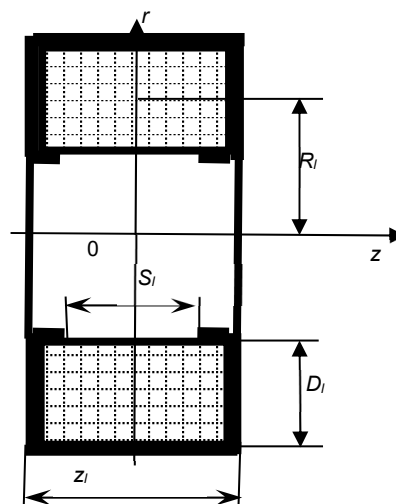


Fig. 1. Geometric parameters of the short focusing magnetic lens

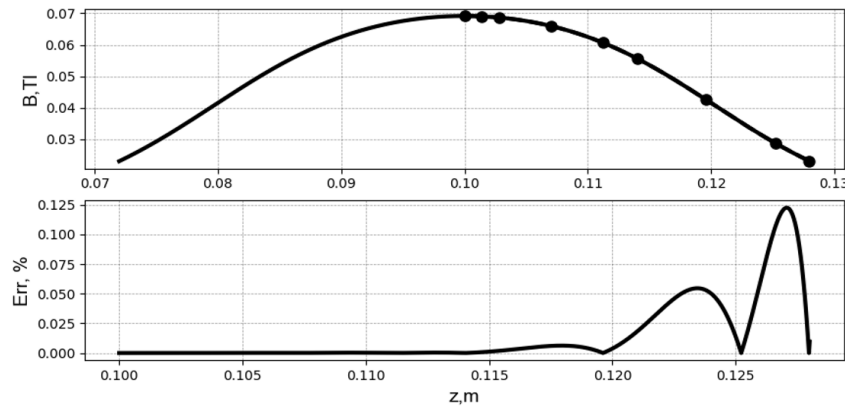


Fig. 2. The result of interpolation of the axial magnetic field distribution of a short focusing lens with a fourth-order root-fractional-rational function for first example

From the given graphical dependencies, it is clear that the maximum value of the interpolation error is observed near the value of the function argument $z = 0$ and is about 0.4 %. The average interpolation error is 0.1 %. The coefficients of the second-order root-fractional-rational function, calculated from relations (6), have the following values: $C_1 = 0.07482$; $C_2 = -6.9629 \cdot 10^{-4}$; $C_3 = 1.1851$; $C_4 = -0.88147$; $C_5 = 1.20276$.

From the point of view of further research of the described interpolation methods, it is interesting that for a root-fractional-rational function of the second order an inverse function $z(r) = r^{-1}(z)$ is existed. The corresponding analytical relationship for the functional dependence $z(r)$ can be written in the following form:

$$z(r) = \frac{C_1 - r^2 C_4}{2 \cdot (1 - r^2 C_3)} \pm \sqrt{\frac{(C_1 - r^2 C_4)^2}{4 \cdot (1 - r^2 C_3)^2} - \frac{C_2 - r^2 C_5}{1 - r^2 C_3}}. \quad (8)$$

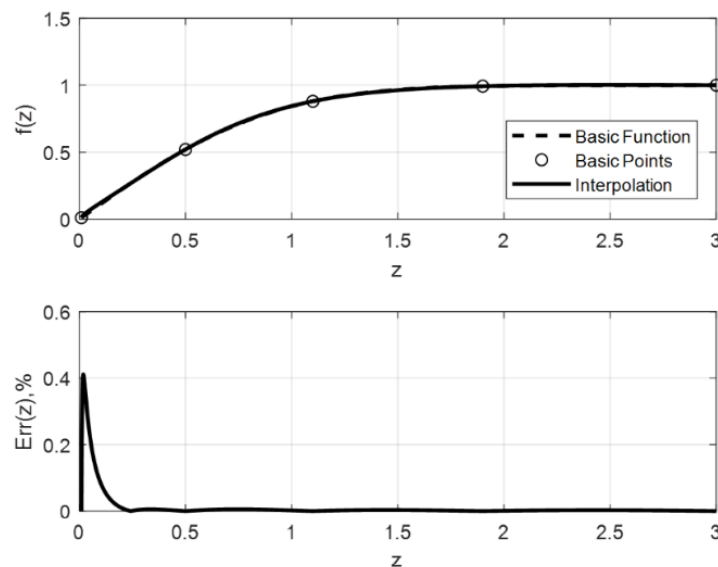


Fig. 3. Interpolation of the function $\text{erf}(z)$ by a second-order root-fractional-rational function for second example

For example, calculating the values of the inverse function in the error function $z(r) = \text{erf}^{-1}(z)$ using formula (8) makes it possible to construct an effective generator of random numbers distributed according to a normal law. It should also be noted that relation (8) is universal and can be used for any set of numerical data. Only the coefficients of the root-fractional-rational function $C_1 - C_5$ change.

It should be noted here that for a homomorphic single-valued functions $r(z)$, the sign – in front of the square root in relation (8) is an incorrect solution. It yields negative values of the inverse function $z(r)$, and by this reason is discarded. The corresponding dependencies for this example are shown in Fig. 4. On the other hand, for multi-valued polymorphic functions, both solutions should be considered, each of which is valid for the corresponding values of the argument of the original function $r(z)$.

3.3. Third example

Describe a Π - shaped fuzzy logic membership function defined by the following arithmetic-logical expressions (Zadeh, & Aliev, 2018):

$$F_s(z; a, b) = 1 \cdot (z \leq a) + \frac{1 + \cos\left(\frac{\pi \cdot (z - b)}{b - a}\right)}{2} \cdot (a < z \leq b) + 0 \cdot (z > b);$$

$$F_Z(z; c, d) = 0 \cdot (z \leq c) + \frac{1 + \cos\left(\frac{\pi \cdot (z - c)}{d - c}\right)}{2} \cdot (c < z \leq d) + 1 \cdot (z > d);$$

$$F_{\Pi}(z; a, b, c, d) = F_Z(z; c, d) F_S(z; a, b), a < b < c < d.$$
(9)

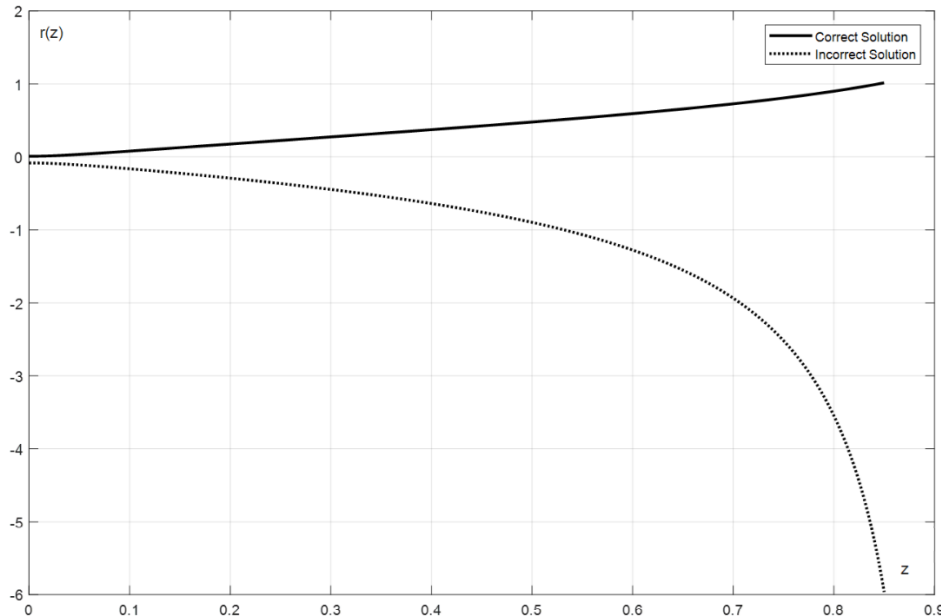


Fig. 4. Correct and incorrect solutions for inverse function $z(r)$ corresponding to relation (8) for second example

Calculations have to be carried out for the following values of the parameters of function: $a = 1$, $b = 2$, $c = 3$, and $d = 4$. Interpolation has to be provided using a fourth-order root-fractional-rational function on the interval of values of the argument z [1,5; 3,5].

To solve this task, the following reference interpolation points have been selected: $z_1 = 1.5$; $z_2 = 1.6$; $z_3 = 1.9$; $z_4 = 2.0$; $z_5 = 2.5$; $z_6 = 2.9$; $z_7 = 3.15$; $z_8 = 3.4$; $z_9 = 3.5$.

The result of interpolation for a given set of numerical data is presented in Fig. 5. The interpolation error in this case does not exceed 1.8 %, and the average error over the entire interpolation interval is 0.5326%. The coefficients of the fourth-order root-fractional rational function for this example have the following values: $C_1 = -9.987$; $C_2 = 35.1493$; $C_3 = -51.006$; $C_4 = 26.148$; $C_5 = 2.5061$; $C_6 = -25.134466$; $C_7 = 91.95123$; $C_8 = -145.085$; $C_9 = 84.18057$.

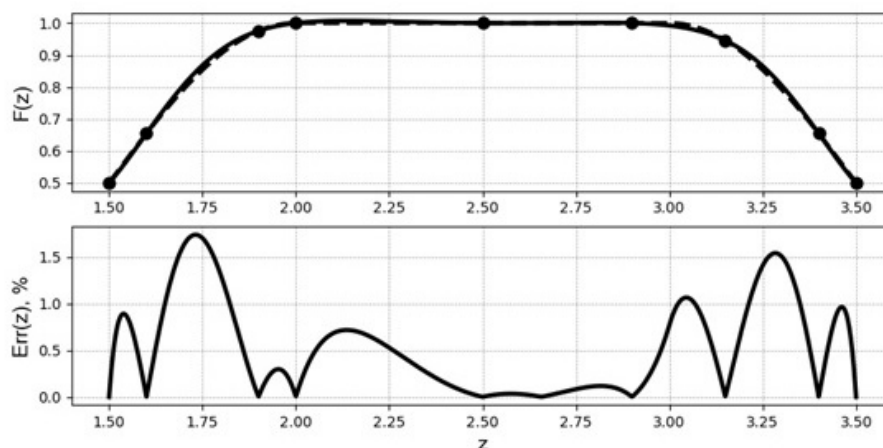


Fig. 5. Interpolation of the Π -shaped fuzzy logic membership function specified by arithmetic-logical relations (9) for third example

4. Features of the software implementation of the developed interpolation algorithm

The program code intended for calculating the coefficients of root-fractional-rational functions using obtained corresponded relations, similar to (6), has been written in the Python programming language using advanced developed libraries for performing mathematical calculations and displaying graphical data from the pip module. Therefore, additional system tools have been used to connect the corresponding advanced programming possibilities. In particular, a virtual environment has been created using the system commands of the Windows 10 operating system to define corresponding

virtual variables. After that, in this virtual environment, virtual memory was allocated to create a virtual disk. To perform these system functions, the venv software module has been used (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013).

The main features of writing program code were as follows.

1. In general, the program is written using well-known means of structural, declarative and functional programming in one program module.

2. Mathematical operations with data arrays have been performed using advanced developed mathematical libraries of the numpy module (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013). To operations with matrices, effective matrix programming tools have been used (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013). Using of such approach is greatly simplifies writing, reading and improving program code.

3. To display function graphs in the created graphical user interface window, developed graphic libraries of the matplotlib module have been used (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013).

4. Since the created software package is primarily intended to solve a wide range of practical problems related to the simulation of electron beams, to systematize the functions of the program and to simplify the work of users with it, each of these realized functions is implemented on separate tabs.

5. To solve complex problems of simulation electron beams, data exchange between separate, local tasks is provided. At the software level, data exchange between internal functions is implemented through the directive of data description global (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013). For example, the results of calculating the magnetic field of a focusing lens and their interpolation can be used to calculate the boundary trajectory of an electron beam. The algorithm for calculating the trajectory of electron beams as they propagate in an ionized gas, have been realized in this software, was considered and described in detail in (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013).

6. The obtained graphic dependencies can be saved in a file on the local computer disk in the tiff graphic format. For this, the standard functions of the matplotlib module graphic library are used (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013). To create and record a graphic file, the corresponding elements of the graphic interface are used on each tab, in particular, the event processing button and the text window for entering the file name (Robbins, 2022; McKinney, 2023; Chollet, 2022; Lutz, 2013).

The Study Properties of Root-Fractional-Rational Function tab, where the properties of mathematical functions describing sets of data for interpolation are studied, is shown in Fig. 6 a, and the Simulation of Magnetic Lens tab, on which the axial distribution of the magnetic field of short focusing lenses is interpolated, presented, correspondently, in Fig. 6 b.

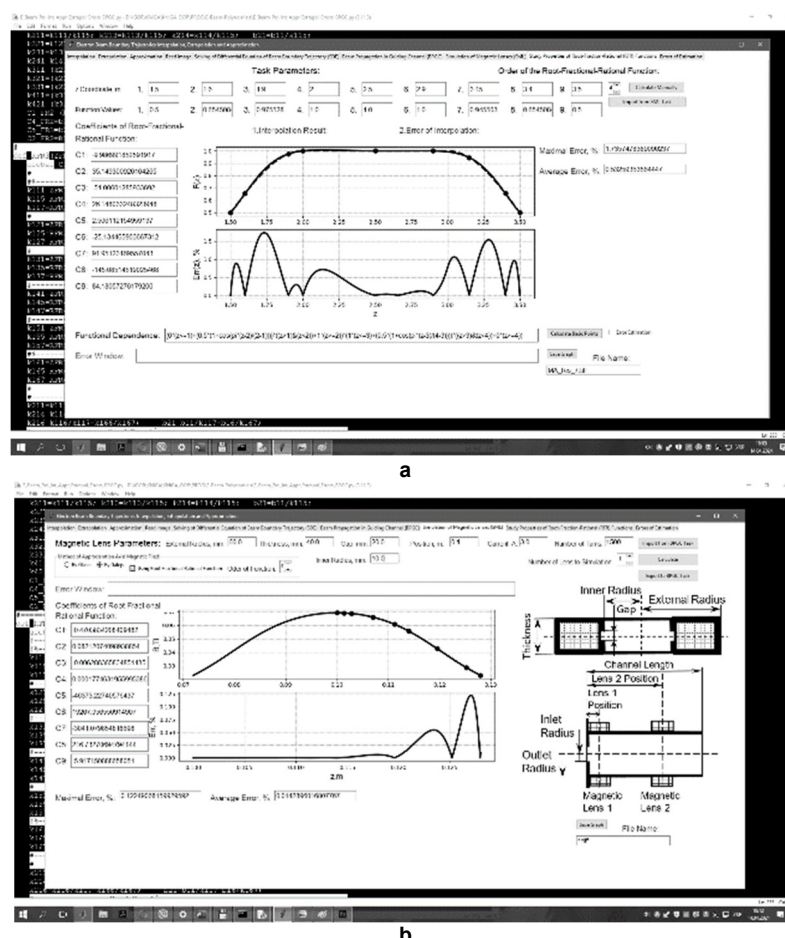


Fig. 6. Tabs Study Properties of Root-Fractional-Rational Function (a) and Simulation of Magnetic Lens (b) of the graphical user interface of the developed software package, has been elaborated for working with root-polynomial and root-fractional-rational functions (Melnik et al., 2023). Screen copies

Discussion and conclusions

Theoretical studies and test computer experiments with the developed software have shown that when using root-fractional rational functions of the second, third and fourth orders, it is possible, with high accuracy, to interpolate the values of various functions, in particular, those that have a high stiff degree. Often for this class of functions there are no simple analytical expressions, and they can only be described by numerical methods associated with solving complex systems of integral or differential equations. Testing of the developed interpolation algorithm have been carried out for a wide class of mathematical functions of various types used in various fields of engineering modeling, including problems in electrical engineering and electronic optics, in probability theory and mathematical statistics, as well as in applied problems of fuzzy logic theory. These test studies have shown that even for functions with a high stiff degree, the maximum interpolation error does not exceed several percent, and in most cases it can be range of a fraction of a percent. The theoretical results obtained in the performance of this work may be of interest to a wide range of mathematicians who are involved in functional analysis and mathematical processing of large data sets, in particular, in methods of numerical and statistical analysis.

Authors' contribution: Melnyk Igor – conceptualization and main ideas; Pochynok Alina – mathematical calculations and finding analytical dependences for coefficients of root-fractional-rational functions; Skrypka Mykhailo – realization the conception of simulation in virtual environment and writing Python programming code; Olga Demyanchenko – study the mathematical properties of root-fractional-rational functions and particularities of its convergence.

Acknowledgments, funding sources. All scientific research described in this work has been carried out in the Laboratory of Electron Beam Technological Devices of the National Technical University of Ukraine "Igor Sikorsky Kiev Polytechnic Institute" as part of scientific and technical research works commissioned by the Ministry of Education and Science of Ukraine, state registration number is 0124U001525.

References

- Akhonin, S. V., Berezos, B. A., Pikulin, A. N., Severin, A. Yu., Shvab, S. L., & Erokhin, A. G. (2018). Production of high-strength titanium alloy VT22 by electron beam melting method. *Advances in Electrometallurgy*, 3(131), 8–14 [in Russian]. [Ахонин, С. В., Березос, В. А. Пикулин, А. Н., Северин, А. Ю., Шваб, С. Л., & Ерохин, А. Г. (2018). Получение высокопрочного титанового сплава ВТ22 способом электронно-лучевой плавки, *Современная электрометаллургия*, 3(131), 8–14]. <https://patonpublishinghouse.com/sem/pdf/2018/sem201803all.pdf>
- Bronstein, I. N., Semendyayev, K. A., Musiol, G., & Mühlig, H. (2007). *Handbook of Mathematics* (5th ed.). Springer. <https://doi.org/10.1007/978-3-662-46221-8>
- Chapra, S. C., & Canale, R. P. (2014). *Numerical Methods for Engineers* (7th ed.). McGraw Hill.
- Chollet, F. (2022). *Deep Learning with Python* (2nd ed.). Manning. <https://doi.org/10.31211/interacoes.n42.2022.r1>
- Druzhinin, A. A., Ostrovskii, I. P., Khoverko, Y. N., Liakh-Kaguy, N. S., & Vuytsyk, A. M. (2014). Low temperature characteristics of germanium whiskers. *Functional materials*, 21(2), 130–136. <http://functmaterials.org.ua/contents/21-2/2; http://dx.doi.org/10.15407/fm21.02.130>
- Druzhinin, A. A., Bolshakova, I. A., Ostrovskii, I. P., Khoverko, Y. N., & Liakh-Kaguy, N. S. (2015). Low temperature magnetoresistance of InSb whiskers. *Materials Science in Semiconductor Processing*, 40, 550–555. <https://doi.org/10.1016/j.mssp.2015.07.030>
- Epperson, J. F. (2013). *An Introduction to Numerical Methods and Analysis* (2nd ed.). Wiley-Interscience. <https://doi.org/10.1002/9781119604570>
- Grechanyuk, I. M., Grechanyuk, M. I., Bagliuk, G. A., Grechanyuk, V. G., Khomenko, O. V., Dudnik, O. V., & Gots, V. I. (2022). Electron-Beam and Plasma Oxidation-Resistant and Thermal-Barrier Coatings Deposited on Turbine Blades Using Cast and Powder Ni(Co)CrAlY(Si) Alloys Produced by Electron Beam Melting IV. Chemical and Phase Composition and Structure of Cocrystal Powder Alloys and Their Use. *Powder Metallurgy and Metal Ceramics*, 61(7–8), 459–464. <https://doi.org/10.1007/s11106-023-00333-0>
- Grivet, P. (1972). *Electron Optics*. Elsevier. <https://doi.org/10.1016/C2013-0-02400-0>
- Gubner, J. A. (2006). *Probability and random processes for electrical and computer engineers*. Cambridge University Press. <http://www.cambridge.org/gb/academic/subjects/engineering/communications-and-signal-processing/probability-and-random-processes-electrical-and-computer-engineers>
- Jain, M. K., Jengar, S. R. K., & Jain, R. K. (2010). *Numerical Methods for Scientific and Engineering Computation*. New Age International Pvt. Ltd. https://www.academia.edu/8565134/Numerical_Methods_For_Scientific_And_Engineering_Computation_M_K_Jain_S_R_K_Iyengar_And_R_K_Jain
- Kasper, E., & Hawkes, P. (2018). *Principles of Electron Optics. Applied Geometrical Optics*. Elsevier Science. <https://doi.org/10.1016/C2015-0-06653-9>
- Kemmotsu, T., Nagai, T., & Maeda, M. (2011). Removal Rate of Phosphorous form Melting Silicon. *High Temperature Materials and Processes*, 30(1–2), 17–22. <https://doi.org/10.1515/htmp.2011.002>
- Lutz, M. (2013). *Learning Python* (5th ed.). O'Reilly. <https://www.oreilly.com/library/view/learning-python-5th/9781449355722/>
- Mathews, J. H., & Fink, K. D. (1998). *Numerical Methods. Using MATLAB* (3rd ed.). Prentice Hall. <https://convexoptimization.com/TOOLS/Fink.pdf>
- McKinney, W. (2023). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter* (3rd ed.). O'Reilly Media. <https://www.oreilly.com/library/view/python-for-data/9781098104023/>
- Melnyk, I. (2005). Numerical simulation of distribution of electric field and particle trajectories in electron sources based on high-voltage glow discharge. *Radioelectronic and Communication Systems*, 48(6), 61–71. <http://radioelektronika.org/article/view/S0735272705060087>
- Melnyk, I. (2006). Simulation of geometry of high voltage glow discharge electrodes' systems, formed profile electron beams. In A. M. Filachev (Ed.), *Proceedings of SPIE: Vol. 6278. Seventh Seminar on Problems of Theoretical and applied Electron and Ion Optics*. 627809-1–627809-13. <https://doi.org/10.1117/12.693202>
- Melnyk, I., & Luntovskyy, A. (2022). Estimation of Energy Efficiency and Quality of Service in Cloud Realizations of Parallel Computing Algorithms for IBN. In M. Klymash, M. Beshley, & A. Luntovskyy (Eds.), *Lecture Notes in Electrical Engineering: Vol. 831. Future Intent-Based Networking* (pp. 339–379). Springer, Cham. https://doi.org/10.1007/978-3-030-92435-5_20
- Melnyk, I., & Pochynok, A. (2020). Study of a Class of Algebraic Functions for Interpolation of Boundary Trajectories of Short-Focus Electron Beams. *System Researches and Information Technologies*, 3, 23–39 [in Ukrainian]. [Мельник, І. В., Починок, А. В. (2020). Дослідження класу алгебричних функцій для інтерполяції межових траєкторій короткофокусних електронних пучків, *Системні дослідження та інформаційні технології*, 3, 8–14]. <https://doi.org/10.20535/SRIT.2308-8893.2020.3.02>
- Melnyk, I., & Pochynok, A. (2023). Theoretical justification of application possibility of different order root-polynomial functions for interpolation and approximation of boundary trajectory of electron beam. *Radioelectronics and Communications Systems*, 66(2), 53–73. <https://doi.org/10.3103/S0735272723020024>
- Melnyk, I., Tuhai, S., Surzhykov, M., Shved, I., Melnyk, V., & Kovalchuk, D. (2022). Analytical Estimation of the Deep of Seam Penetration for the Electron-Beam Welding Technologies with Application of Glow Discharge Electron Guns. In *2022 IEEE 41st International Conference on Electronics and Nanotechnology* (pp. 1–5). IEEE. <https://doi.org/10.1109/ELNANO54667.2022.9927071>
- Paton, B. E., Akhonin, S. V., & Berezos, V. A. (2018). Development of technologies for electron-beam melting of metals at the E.O. Paton Electric Welding Institute NAS of Ukraine. *Advances Electrometallurgy*, 4(132), 19–35 [in Russian]. [Патон, Б. Е., Ахонин, С. В., & Березос, В. А. (2018). Развитие технологий электронно-лучевой плавки металлов в ИЭС им. Е. О. Патона НАН Украины. *Современная электрометаллургия*, 4(132), 19–35]. <http://dx.doi.org/10.15407/sem2018.04.01>
- Phillips, G. M. (2023). *Interpolation and Approximation by Polynomials*. Springer. <https://doi.org/10.1007/b97417>
- Prikhna, T. O., Grechanyuk, I. M., Karpets, M. V., Grechanyuk, M. I., Bagliuk, G. A., Grechanyuk, V. G., & Chornovol, V. A. (2022a). Electron-Beam and Plasma Oxidation-Resistant and Thermal-Barrier Coatings Deposited on Turbine Blades Using Cast and Powder Ni(Co)CrAlY(Si) Alloys I. Fundamentals of the Production Technology. Structure, and Phase Composition of Cast NiCrAlY Alloys. *Powder Metallurgy and Metal Ceramics*, 61(1–2), 70–76. <https://doi.org/10.1007/s11106-022-00296-8>
- Prikhna, T. O., Grechanyuk, I. M., Karpets, M. V., Grechanyuk, M. I., Bagliuk, G. A., Grechanyuk, V. G., & Khomenko, O. V. (2022b). Electron-Beam and Plasma Oxidation-Resistant and Thermal-Barrier Coatings Deposited on Turbine Blades Using Cast and Powder Ni(Co)CrAlY(Si) Alloys Produced by Electron-Beam Melting II. Structure and Chemical and Phase Composition of Cast CoCrAlY Alloys. *Powder Metallurgy and Metal Ceramics*, 61(3–4), 230–237. <https://doi.org/10.1007/s11106-022-00310-z>

- Raizer, Yu. P. (1991). *Gas Discharge Physics*. Springer. <https://link.springer.com/book/9783642647604>
- Rudychev, V. G., Lazurik, V. T., & Rudychev, Y. V. (2021). Influence of the electron beams incidence angles on the depth-dose distribution of the irradiated object. *Radiation Physics and Chemistry*, 186, 109527. <https://doi.org/10.1016/j.radphyschem.2021.109527>
- Schiller, S., Heisig, U., & Panzer, S. (1982). *Electron Beam Technology*. John Wiley & Sons.
- Smirnov, B. M. (2015). *Theory of Gas Discharge Plasma*. Springer Cham. <https://doi.org/10.1007/978-3-319-11065-3>
- Szilagy, M. (2012). *Electron and Ion Optics*. Springer. <https://doi.org/10.1007/978-1-4613-0923-9>
- Tseluyko, A. F., Lazurik, V. T., Ryabchikov, D. L., Maslov, V. I., & Sereda, I. N. (2008). Experimental study of radiation in the wavelength range 12.2-15.8 nm from a pulsed high-current plasma diode. *Plasma Physics Reports*, 34(11), 963–968. <https://doi.org/10.1134/S1063780X0811010X>
- Wentzel, E., & Ovcharov, L. (2002). *Applied Problems of Probability Theory*. Mir Publisher.
- Zadeh, L. A., & Aliev, R. A. (2018). *Fuzzy Logic Theory and Applications*. World Scientific.
- Zakharov, A., Rozenko, S., Litvintsev, S., & Ilchenko, M. (2020). Trisection Bandpass Filter with Mixed Cross-Coupling and Different Paths for Signal Propagation. *IEEE Microwave Wireless Component Letters*, 30(1), 12–15. <https://doi.org/10.1109/LMWC.2019.2957207>
- Zakharov, A., Rozenko, S., & Ilchenko, M. (2019). Varactor-Tuned Microstrip Bandpass Filter with Loop Hairpin and Combine Resonators. *IEEE Transactions on Circuits Systems. II. Experimental Briefs*, 66(6), 953–957. <https://doi.org/10.1109/LMWC.2019.2929650>

Отримано редакцією журналу / Received: 05.08.24

Прорецензовано / Revised: 24.01.25

Схвалено до друку / Accepted: 06.02.25

Ігор МЕЛЬНИК¹, д-р техн. наук, проф.

ORCID ID: 0000-0003-0220-0615

e-mail: imelnik@phbme.kpi.ua

Аліна ПОЧИНОК¹, канд. техн. наук, ст. наук. співроб.

ORCID ID: 0000-0001-9531-7593

e-mail: alina_pochynok@yahoo.com

Михайло СКРИПКА¹, асп.

ORCID ID: 0009-0006-7142-5569

e-mail: scientetik@gmail.com

Ольга ДЕМ'ЯНЧЕНКО², канд. техн. наук, ст. викл.

ORCID ID: 0000-0002-4693-0364

e-mail: olgademyanchenko@gmail.com

¹Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна²Національний університет "Львівська політехніка", Львів, Україна

МЕТОД ІНТЕРПОЛЯЦІЇ З ВИКОРИСТАННЯМ КОРЕНЕВО-ДРОБОВО-РАЦІОНАЛЬНИХ ФУНКЦІЙ РІЗНОГО ПОРЯДКУ

Розглянуто можливість інтерполяції різних видів математичних функцій з використанням коренево-дробово-раціональних функцій різного порядку, а саме другого, третього та четвертого. Показано, що коренево-дробово-раціональні функції надають точну інтерполяцію з малою похибкою навіть жорстких аналітичних залежностей. Розглянуто коренево-дробово-раціональні функції від другого до четвертого порядку та наведено відповідні аналітичні співвідношення для визначення поліноміальних коефіцієнтів у чисельнику та знаменнику. Описано аналітичний метод розв'язування системи лінійних алгебраїчних рівнянь для пошуку коефіцієнтів коренево-дробово-раціональних функцій, оснований на відомому методі виключення змінних Гаусса – Зейделя, який відрізняється тим, що аналітичні обчислення проведені ітераційно та, для систематизації поставленої математичної задачі, усі коефіцієнти пронумеровані. Нумерація змінних містить три коефіцієнти: номер ітерації, номер рівняння та номер змінної у цьому рівнянні. Такий систематизований підхід до розв'язування задачі у значній мірі спростив проведення аналітичних розрахунків. Також доведено, що кількість необхідних точок для інтерполяції відповідає значенню $2n+1$, де n – порядок коренево-дробово-раціональної функції. Наведено приклади інтерполяції різних функцій для задач електродинаміки, моделювання магнітних лінз, імовірнісних задач і задач нечіткої логіки, а також визначено похибку інтерполяції для всіх розглянутих прикладів. Зокрема, подано результати інтерполяції для розподілу магнітного поля короткої лінзи, для функції помилок $\text{erf}(x)$, для експоненціальної функції, для розподілу Бернуллі з різними коефіцієнтами та для П-подібної функції належності. Усі припущення теоретичного аналізу перевірені за допомогою оригінально розробленого програмного забезпечення для ЕОМ, створеного засобами мови програмування Python. Для більшості розглянутих прикладів отримана похибка інтерполяції становить менше кількох відсотків. Наведено також графічні результати тестування, отримані для запропонованого методу інтерполяції.

Ключові слова: інтерполяція, коренево-дробово-раціональна функція, задачі електродинаміки, функції теорії ймовірностей, нечітка логіка, функції належності.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.8

DOI: <https://doi.org/10.17721/1812-5409.2025/1.17>

Yevhen OSMAK, PhD Student

ORCID ID: 0009-0007-9081-5527

e-mail: yevhen.osmak@knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

MODELLING OF THE DECISION-MAKING PROCESS OF JOBS PRIORITIES DEFINITION UNDER CONDITIONS OF LIMITED RESOURCES: CASE OF A SERVICE COMPANY

This article focuses on the problem of decision-making regarding the prioritization, selection and scheduling of maintenance and repair jobs on complex industrial equipment under conditions of limited resources for the case of a service company. The author reviews modern approaches to predictive maintenance and jobs scheduling, namely job shop scheduling problem (JSSP), noting a gap in combining these two tasks for the case of service companies and absence of its strict scientific substantiation. As the first step, a mathematical model is proposed which formulates the problem as an integer linear programming problem (ILP) with the objective of maximizing revenue and minimizing the total duration of the work execution under the limited resources of the service company. The scheduling part is based on the flexible JSSP, one of the latest developments of JSSP, which allows to consider business-processes of the service company, which operates a fleet of almost identical mobile workshops. The peculiarity of the researched model is usage of time-indexed variables. Since the revenue is unknown at the moment of decision making, it is modelled with type-2 fuzzy logic approach, which mimics the real expert decision making. Application of type-2 fuzzy logic, in contrast to the type-1 fuzzy logic, allows to consider both uncertainty of experts in preliminary revenue evaluation, which could be modelled by type-1 fuzzy logic, and vagueness in determination of membership function. Directions for further research are outlined, among them numerical experiments on real and synthetically generated data, and further multi-criteria optimization considering equipment downtime. In further research it will be necessary to overcome the increased computing complexity of type-2 fuzzy logic. A few ideas are proposed to avoid this obstacle.

Keywords: predictive maintenance; data-driven methods; type-2 fuzzy logic; flexible job-shop scheduling; integer linear programming.

AMS 2020 classification: 90C11, 90C29, 90C70, 90C90.

Introduction

1. Modern predictive maintenance. Digital technologies, cheaper industrial sensors, and availability of broadband networks enabled real-time performance monitoring of complex industrial equipment. This speeds up the collecting and processing of digital data, and, most importantly, enables maintaining of high technical readiness of industrial equipment, which is a critical success indicator of industrial machinery.

In manufacturing, the connection between physical equipment, such as moving machines, stationary equipment, etc., in the form of primary data and IT systems is the basis for Industry 4.0. Rüßmann et al. (2015) analyze Industry 4.0 technological trends and explore its economic benefits. The data from various sources and business process management systems collection and evaluation are becoming the de facto standard for real-time decision-making under limited resources. These processes are widely known as the Internet of Things (IoT). IoT is a concept that involves connecting physical objects equipped with sensors to the Internet so that they can exchange data and interact with each other without direct human intervention.

Those approaches are especially relevant for industries where equipment requires millions of US dollars of investment and losses due to equipment downtime are estimated at thousands of dollars per hour (Siemens AG, 2023). Those industries are the construction, mining, and autonomous power systems, where equipment includes excavators, bulldozers, dump trucks, rollers, graders, industrial diesel generator sets, etc. The research in this article focuses on these industries.

Purpose of the paper. The agents in studied industries are equipment manufacturers and operators, as well as service companies. Due to necessity of high competent technicians, equipment operators often cannot retain required specialists and outsource maintenance to service companies (Baatartogtokh, Dunbar, & van Zyl, 2018).

Typically, a service company has the following goals: technical readiness increase, which benefits the manufacturer and operator; and own revenue increase under condition of adequate available resources utilization. The scheduling problem of a service company resembles Job-shop scheduling problem (JSSP). Finding the best schedule can be very easy or very difficult, depending on the shop environment, process constraints, etc. (Pinedo, 1995). In JSSP there is a set of n jobs to be executed on m machines, where each job consists of n_i operations. The goal of JSSP is to find a sequence of operations on the machines to optimize a productivity measure, typically the makespan, i.e. the time required to complete all jobs.

The JSSP is one of the combinatorial optimization problems that has been studied since the 1960s (Doh et al., 2013). The authors note the inherent complexity of JSSP, namely that it has been proven to be NP-complete in the strong sense (Lenstra, Kan, & Brucker, 1977). One of the development directions is the formulation with alternative machines, which is called the flexible job shop scheduling problem (FJSSP). In contrast to the classical problem, FJSSP has route flexibility, i.e. each job can be performed on different machines. FJSSP is exactly applicable in the studied industries for the service company case.

It remains unclear why the authors of previous studies have not considered the operating model of service companies. In fact, there is a gap between the predictive maintenance problem, which aims to determine the optimal time of maintenance intervention, and the problem of maintenance tasks scheduling. Although both problems are individually well developed, we couldn't find similar models for optimal management of the flow of potential failures signals from the predictive maintenance system in such a way that the selected for response signals are transformed into jobs in order to provide a high degree of technical readiness and maximize the service company revenue, considering limited resources of the service company. The purpose of this study is to conduct modeling for the described case. Since the potential revenue is not known, it is modeled using the type-2 fuzzy logic.

Literature review. Although JSSP has been thoroughly studied since the late 1960s, usually papers focus on the operator's point of view. The authors (Demir, İşleyen, 2013) argue that a complete review of mathematical formulations of

© Osmak Yevhen, 2025

FJSSP had not been considered before. The four most common formulations of FJSSP are proposed and compared. The recommendations for choosing models for specific cases are presented. Having considered those recommendations, a time-indexed model was selected as basic for this study.

An approach to prioritizing the incoming flow of jobs is proposed in (Doh et al., 2013). The main decision-making steps are selecting an operation/machine pair; and determining the sequence of tasks assigned to each machine. A practical approach to priority planning is proposed, in which these two decisions are made simultaneously.

A general overview of methods for solving the problem of dynamic FJSSP is given by the authors in (Mohan, Lanka, & Rao, 2019). The article discusses the concept and strategies of dynamic planning, methods of planning, and the planning system. The methods are divided into exact and approximate. Approximate methods include establishing prioritization rules; simulations; artificial neural networks; tabu search; genetic algorithms; ant colony optimization; particle swarm optimization; immune algorithms; fuzzy logic and hybrid approaches. The paper does not provide a comparative characterization of the above methods. This is another gap worth exploring.

More attention is paid to the comparison of constructing metaheuristics methods in (Thiruvady, Blum, & Ernst, 2020). The authors argue that metaheuristics are becoming popular for solving combinatorial optimization problems. Constructive heuristics and ant colony optimization are considered.

In recent years, machine learning methods have also become extremely popular for solving scheduling problems in a broad sense. Here are some of them.

In the paper (Boussadia, & Driss, 2021) an overview of machine learning methods for solving JSSP is provided. More extensive overview is provided in (Schlebusch, 2020) with focus on the reinforcement learning. In (Hameeda, & Schwung, 2023) graph neural networks are used. A hybrid evolutionary algorithm for solving the reactive FJSSP is proposed by the authors (Lina et al., 2012). Promising multilayer convolutional neural networks are used in (Shao, & Kim, 2022). The paper (Liu, Piplani, & Toro, 2022) describes an approach based on hierarchical and distributed architecture. The Double Deep Q-Network algorithm is used for agents teaching and for real-time decision making.

The (Li et al., 2021) is particularly interesting, while the type-2 fuzzy logic approaches are used in the duration of work modeling.

2. Main result: Modelling

Let's generally describe the modeled business process. A service company operates a machines and equipment fleet of its customers (referred to as machines). These machines automatically send real-time data in the form of so-called events. The events consist of information about deviations from the optimal parameters and are the basis for decision about preventive maintenance to avoid breakdowns. The service company, in turn, enriches the event data with data about inspections, lubricant analysis, etc. All of this is presented in a dashboard to a service company operator, who selects some of the events for further transformation into jobs to be performed. The operator has an expert understanding of the potential revenue, generated by converting a particular event into a job. The events selected by the operator are then included in the schedule according to FJSSP. The flexibility here means that a particular job can be performed by a mechanic with a skill level not lower than that required.

The aim of this study is to develop a model for the decision making regarding optimal events selection from the incoming events stream, provided by the predictive maintenance system, which are further transformed into jobs, for the case of a service company, which has 2 objectives: revenue maximization and optimal scheduling of the selected jobs, taking into account the limited available resources. The potential revenue can only be estimated by an expert at the stage of events analysis and modeled using type-2 fuzzy.

The time-indexed model is selected as the basic model (Demir, & İşleyen, 2013). Considering that all operations within a single job in the industries under study are performed within one mobile workshop and cannot be partially performed on another workshop, we will remove the concept of operations from the basic model, i.e. the number of operations in each job is always defined as one. This reduces the combinatorial complexity of the resulting model.

Mathematical formalization. The problem has m mobile workshops and n jobs. Each job can be performed by any of the available mobile workshops if the technician (operator of this mobile workshop) has qualification not lower than the required. At the zero moment of time all the mobile workshops are available. A workshop can only perform one job at a time. No overruns are allowed. The proposed model neglects the mobile workshops setup time, which includes receiving necessary materials and spare parts, and the time of transportation between machines locations. The objective is to minimize makespan C_{max} , which is the total duration of the selected works execution and to maximize expected revenue $\tilde{R}$.

Potential revenue $\tilde{r}_i$, which is generated by execution of job J_i is unknown but can be estimated by experts. To model this parameter, we will use type-2 fuzzy logic. This is justified by the real business process and allows us to consider the uncertainty of experts not only in estimating the potential revenue amount, but also in estimate reliability per se. Applying a fuzzy logic approach to some humanitarian problems shows much clearer results compared to traditional statistical models (Mashchenko, Kapustian, & Rubino, 2023). Interval type-2 fuzzy sets (IT2FS) are used in the form $\tilde{a}_i = (a_1, a_2, a_3, a_4, a_5)$, as it's suggested in (Li et al., 2021).

Definition. A type-2 fuzzy set, say $\tilde{A}$, with a membership function $\mu_{\tilde{A}}(x, u)$, where $x \in X$, and $u \in [0, 1]$ is defined as follows:

$$\tilde{A} = \{((x, u), \mu_{\tilde{A}}(x, u)) \mid \forall x \in X, \forall u \in [0, 1]\},$$

$$0 \leq \mu_{\tilde{A}}(x, u) \leq 1.$$

An IT2FS $\tilde{A}_I$, is a special case of T2FS $\tilde{A}_I = \{((x, u), 1) \mid \forall x \in X, \forall u \in [0, 1]\}$ as in (Mendel, & Liu, 2013)

Assume two IT2FS $\tilde{A} = (a_1, a_2, a_3, a_4, a_5)$ and $\tilde{B} = (b_1, b_2, b_3, b_4, b_5)$ are given.

Addition operator:

$$\tilde{A} + \tilde{B} = (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4, a_5 + b_5).$$

Ranking operator. Calculate the centroid values $C_{\tilde{A}}$ and $C_{\tilde{B}}$ for $\tilde{A}$ and $\tilde{B}$ respectively:

$$C_{\tilde{A}} = [C_{\tilde{A}}^U, C_{\tilde{A}}^L] \approx \begin{bmatrix} a_3 - \frac{(a_5 - a_4)(a_5 + 2a_4 - a_2 - 2a_3)}{6(a_4 - a_2)} \\ a_3 + \frac{(a_5 + a_2 - 2a_3)(a_5 + a_4 - 2a_2)}{6(a_4 - a_2)} \end{bmatrix}$$

$$C_{\tilde{B}} = [C_{\tilde{B}}^U, C_{\tilde{B}}^L] \approx \begin{bmatrix} b_3 - \frac{(b_5 - b_4)(b_5 + 2b_4 - b_2 - 2b_3)}{6(b_4 - b_2)} \\ b_3 + \frac{(b_5 + b_2 - 2b_3)(b_5 + b_4 - 2b_2)}{6(b_4 - b_2)} \end{bmatrix}$$

Next, compare $\tilde{A}$ and $\tilde{B}$ with the following logic:

If $C_{\tilde{A}} > (<) C_{\tilde{B}}$, then $\tilde{A} > (<) \tilde{B}$; otherwise

if $a_3 > (<) b_3$, then $\tilde{A} > (<) \tilde{B}$; otherwise

if $(a_5 - a_1) > (<) (b_5 - b_1)$, then $\tilde{A} > (<) \tilde{B}$.

Indices, parameters, and variables used in the model:

Indices

i, h : job index (1, ..., n); k : mobile workshop index; u : time period index (1, ..., T).

Parameters

n : total number of jobs, m : total number of mobile workshops; T : number of available for planning periods

p_{ki} : processing time of job J_i , if performed on machine k

$\tilde{r}_i$: expected revenue for job J_i , in a form of IT2FS $\tilde{r}_i = (r_1, r_2, r_3, r_4, r_5)$

a_{ki} : describes the capable mobile workshops set M_i is assigned to job J_i

$a_{kij} = \begin{cases} 1, & \text{if } J_i \text{ can be performed on machine } i \\ 0, & \text{otherwise} \end{cases}$

M : a large number (practical infinity, can be defined as $100 * T$)

Decision variables

C_{max} : makespan, i.e. total duration of all selected jobs execution, positive integer

$\tilde{R}$: total expected revenue as a result of all selected jobs execution

$$v_{ik} = \begin{cases} 1, & \text{if job } J_i \text{ is done on mobile workshop } k \\ 0, & \text{otherwise} \end{cases}$$

$$w_{iku} = \begin{cases} 1, & \text{if } J_i \text{ done on workshop } k \text{ during period } u \\ 0, & \text{otherwise} \end{cases}$$

$$z_{ihk} = \begin{cases} 1, & \text{if job } J_i \text{ precedes job } J_h \text{ on workshop } k \\ 0, & \text{otherwise} \end{cases}$$

$$g_i = \begin{cases} 1, & \text{if job } J_i \text{ is selected to be performed} \\ 0, & \text{otherwise} \end{cases}$$

w_{iku} is the binary time-indexed variable. It distributes jobs to available mobile workshops by time periods. Thomalla and Gomes formulated a model with time-indexed variables for FJSSP with parallel mobile workshops (Demir, & İşleyen, 2013, p. 983).

As a result, we get a mathematical formulation for the problem modeled in this paper.

$$\tilde{R} = \sum_{i=1, n} g_i * \tilde{r}_i \rightarrow \max. \quad (2.1)$$

$$C_{max} \rightarrow \min. \quad (2.2)$$

$$u * \sum_{k \in M_i} w_{iku} \leq C_{max} \quad \forall i: g_i = 1, u = \overline{1, T}. \quad (2.3)$$

$$\sum_u w_{iku} = p_{ki} * v_{ik} \quad \forall i: g_i = 1, \forall k \in M_i. \quad (2.4)$$

$$\sum_i w_{iku} * g_i \leq 1 \quad \forall k; u = 1, \dots, T. \quad (2.5)$$

$$p_{ki} * (w_{iku} - w_{iku+1}) + \sum_{r=u+2}^T w_{ikr} \leq p_{ki}, \forall i: g_i = 1, k, u = 1, \dots, T-2. \quad (2.6)$$

$$\sum_{k \in M_i} v_{ik} = 1 \quad \forall i: g_i = 1. \quad (2.7)$$

$$p_{ki} * w_{hku} \leq \sum_{r=1}^{u-1} w_{ikr} + M * (1 - z_{ihk}), \quad \forall i \leq h: g_i = 1, g_h = 1, \forall u = 2, \dots, T, \quad (2.8)$$

$$p_{kh} * w_{iku} \leq \sum_{r=1}^{u-1} w_{hkr} + M * (z_{ihk}), \quad \forall i \leq h: g_i = 1, g_h = 1, \forall u = 2, \dots, T \quad \forall k \in M_i \cap M_h. \quad (2.9)$$

$$w_{iku} \in \{0, 1\} \quad \forall i, k, u.$$

$$z_{ihk} \in \{0, 1\} \quad \forall i \leq h, \forall k \in M_i \cap M_h.$$

$$v_{ik} \in \{0, 1\} \quad \forall i, k.$$

$$g_{ik} \in \{0, 1\} \quad \forall i, k.$$

The objectives of the fuzzy multi-criteria optimization problem are maximizing the potential revenue (2.1), which is modeled by IT2FS, and minimizing the total duration of jobs execution (2.2).

Constraint (2.3) defines makespan C_{max} .

Constraint (2.4) defines total execution time of job J_i on mobile workshop k .

Constraint (2.5) forces each machine to process one job at a time.

Constraint (2.6) ensures that each job will not be interrupted until finished.

Constraint (2.7) ensures, that any job is performed on one and only one mobile workshop.

Constraint (2.8) and (2.9) ensures that jobs J_i and J_h cannot be performed in parallel on any of the mobile workshops $M_i \cap M_h$.

Potential methods of problem solving. Fuzzy artificial immune system algorithms are developed in (Li et al., 2021; Ma et al., 2016). Co-evolutionary genetic algorithm is proposed in (Deming, 2012). Multi objective problems are studied in (Tran et al., 2013); in particular, Pareto-optimality approach as hybridization of evolutionary algorithms and fuzzy logic (Kacemf, Hammadi, & Borne, 2002).

Discussion and conclusions

The results of this study show gap in mathematical models for the specific case of a service company performing jobs within the predictive maintenance process. This problem occurs in a wide range of industries. In fact, it is applicable to the entire class of specific equipment that requires deeply specialized technicians for maintenance. The proposed mathematical model formulates the problem, which occurs by service companies from these industries that need to prioritize and schedule maintenance and repair jobs under limited resources. A key feature of the model is the flexibility in assigning tasks to technicians based on their skill level. This is a common and important condition for service companies. The time-indexed formulation also ensures that jobs cannot be postponed once they have started at a particular mobile workshop.

However, there are some limitations and extensions to the model that should be discussed. The model does not consider aspects such as travel time between locations. This could have a significant impact on the optimal schedule, especially when working across large geographical areas.

The challenge for further research may also be the results comparison to the previous papers. One idea for overcoming this challenge is to compare with data of the manual business process execution. Another idea is to compare with the results for classical FJSSP, i.e., to ensure that the proposed approaches are not worse than those already developed and tested for solving FJSSP.

From the solution point of view, the formulated problem is a complex combinatorial optimization problem with objectives, containing IT2FS as parameters, which can be solved by heuristic methods, metaheuristics, and machine learning. Machine learning might be used for fuzzy parameters identification to reduce the manual work, required from the experts. In this case, experts are involved into the ML model training phase and the on-going results verification. The following study will focus on the experimental comparison of the solution methods efficiency. It is not yet clear whether it will be possible to keep the problem in the form of a multi-criteria optimization, or whether the criterion of makespan (minimizing the duration of jobs execution) will be replaced by a restriction on the downtime of mobile workshops.

Finally, while the model provides a framework for decision making, its actual application will require integration with an equipment monitoring system, where decisions are implemented.

As a conclusion, the proposed model provides a formal framework for solving the important problem of prioritizing and scheduling maintenance activities. The discussion highlights promising areas of research to extend the applicability of the model, paving the way for the development of effective decision support tools in the area. A mathematical model has been developed to solve the problem of selecting and scheduling maintenance and repair jobs for service companies under conditions of limited resources. The problem is reduced to a fuzzy two-criteria integer linear programming problem. The literature review revealed the absence of similar approaches that combine predictive maintenance and production schedule planning methods. Areas for further research could include the use of machine learning methods to estimate the components of fuzzy sets, conducting numerical experiments on real and synthetically generated data sets, and extending the model for multi-criteria optimization to consider additional constraints, such as equipment downtime.

References

- Baatarogtokh, B., Dunbar, W. S., & van Zyl, D. (2018). The state of outsourcing in the Canadian mining industry. *Resources Policy*, 59, 184–191. <https://doi.org/10.1016/j.resourpol.2018.06.014>
- Boussadia, N., & Driss, O. B. (2021). Machine Learning for Dynamic Job Shop Scheduling Problem: Literature Review. *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning*, 444–450. <https://doi.org/10.5220/0010736200003101>
- Deming, L. (2012). Co-evolutionary genetic algorithm for fuzzy flexible job shop scheduling. *Applied Soft Computing*, 12(8), 2237–2245. <https://doi.org/10.1016/j.asoc.2012.03.025>
- Demir, Y., & İşleyen, S. K. (2013). Evaluation of mathematical models for flexible job-shop scheduling problems. *Applied Mathematical Modelling*, 37(3), 977–988. <https://doi.org/10.1016/j.apm.2012.03.020>
- Doh, H.-H., Yu, J.-M., Kim, J.-S., Lee, D.-H., & Nam, S.-H. (2013). A priority scheduling approach for flexible job shops with multiple process plans. *International Journal of Production Research* 51(12), 3748–3764. <https://doi.org/10.1080/00207543.2013.765074>
- Hameeda, M. S. A., & Schwung, A. (2023). Graph neural networks-based Scheduler for Production planning problems using Reinforcement Learning. *Journal of Manufacturing Systems*, 69, 91–102. <https://doi.org/10.1016/j.jmsy.2023.06.005>
- Kacem, I., Hammadi, S., & Borne, P. (2002). Pareto-optimality approach for flexible job-shop scheduling problems: hybridization of evolutionary algorithms and fuzzy logic. *Mathematics and Computers in Simulation*, 60(3–5), 245–276. [https://doi.org/10.1016/S0378-4754\(02\)00019-8](https://doi.org/10.1016/S0378-4754(02)00019-8)
- Lenstra, J. K., Kan, A. H. G. R., & Brucker, P. (1977). Complexity of Machine Scheduling Problems. *Annals of Discrete*, 1, 343–362.
- Li, J., Liu, Z.-M., Li, Ch., & Zheng, Zh. (2021). Improved artificial immune system algorithm for type-2 fuzzy flexible job shop scheduling problem. *IEEE Transactions on Fuzzy Systems*, 29(11), 3234–3248. <https://doi.org/10.1109/TFUZZ.2020.3016225>
- Lina, L., Gena, M., Liang, Y., & Ohno, K. (2012). A Hybrid EA for Reactive Flexible Job-shop Scheduling. *Procedia Computer Science*, 12, 110–115.
- Liu, R., Piplani, R., & Toro, C. (2022). Deep reinforcement learning for dynamic scheduling of a flexible job shop. *International Journal of Production Research*, 13(60), 4049–4069. <https://doi.org/10.1080/00207543.2022.2058432>
- Ma, D., Han, X., Liang, Y., & Shi, X. (2016). A Fuzzy Immune GA Scheduling Framework for JSSP. *International Journal of Grid and Distributed Computing*, 9(8), 357–370. <https://doi.org/10.14257/ijgcd.2016.9.8.31>
- Mashchenko, S. O., Kapustian, O. A., & Rubino, B. (2023). On Kemeny Optimization Scheme for Fuzzy Set of Relations. *Axioms*, 12(12), 1067. <https://doi.org/10.3390/axioms12121067>
- Mendel, J. M., & Liu, X. (2013). Simplified Interval Type-2 Fuzzy Logic Systems. *IEEE Transactions on Fuzzy Systems*, 21(6), 1056–1069. <https://doi.org/10.1109/TFUZZ.2013.2241771>
- Mohan, J., Lanka, K., & Rao, A. N. (2019). A Review of Dynamic Job Shop Scheduling Techniques. *Procedia Manufacturing*, 30, 34–39. <https://doi.org/10.1016/j.promfg.2019.02.006>
- Pinedo, M. (1995). *Scheduling: Theory, Algorithms and Systems* (3 ed.). New Jersey, Englewood Cliffs: Springer.
- Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries, 9, 54–89. <https://doi.org/10.5281/zenodo.12804554>
- Schlebusch, D. (2020). Solving the Job-Shop Scheduling Problem with Reinforcement Learning. *University of Applied Sciences and Arts FHNW School of Engineering*.
- Shao, X., & Kim, C. S. (2022). An Adaptive Job Shop Scheduler Using Multi-Level Convolutional Neural Network and Iterative Local Search. *IEEE Access*, 4. <https://doi.org/10.1109/ACCESS.2022.3188765>
- Siemens AG. (2023). The True Cost of Downtime 2022. *SENSEYE PREDICTIVE MAINTENANCE*, 11. Retrieved 05 12, 2024 from <https://assets.new.siemens.com/siemens/assets/api/uuid:3d606495-dbe0-43e4-80b1-d04e27ada920/dics-b10153-00-7600truecostofdowntime2022-144.pdf>
- Thiruvady, D., Blum, C., & Ernst, A. T. (2020). Solution Merging in Matheuristics for Resource Constrained Job Scheduling. *Algorithms*, 13, 256. <https://doi.org/10.3390/a13100256>
- Tran, T.-D., Varela, R., González-Rodríguez, I., & Talbi, E.-Gh. (2013). Solving Fuzzy Job-Shop Scheduling Problems with a Multiobjective Optimizer. *The Fifth International Conference on Knowledge and Systems Engineering*, 197–209. https://doi.org/10.1007/978-3-319-02821-7_19

Отримано редакцією журналу / Received: 05.08.24

Прорецензовано / Revised: 24.01.25

Схвалено до друку / Accepted: 06.02.25

Євген ОСЬМАК, студ.
ORCID ID: 0009-0007-9081-5527
e-mail: yevhen.osmak@knu.ua
Київський національний університет імені Тараса Шевченка, Київ, Україна

РОЗРОБЛЕННЯ МОДЕЛІ УХВАЛЕННЯ РІШЕНЬ ЩОДО ПРІОРИТЕТІВ ВИКОНАННЯ РОБІТ В УМОВАХ ОБМЕЖЕНИХ РЕСУРСІВ ДЛЯ ВИПАДКУ СЕРВІСНОЇ КОМПАНІЇ

Присвячено проблемі ухвалення рішень щодо визначення пріоритетності, відбору та планування робіт із технічного обслуговування і ремонту складного промислового обладнання в умовах обмежених ресурсів на прикладі сервісного підприємства. Розглянуто сучасні підходи до прогнозування технічного обслуговування та планування робіт, а саме задачу планування роботи майстерні (JSSP), відзначаючи прогалину в об'єднанні цих двох задач для випадку сервісних компаній і відсутність її строгого наукового обґрунтування. Як перший крок запропоновано математичну модель, що формулює задачу як задачу цілочисельного лінійного програмування з метою максимізації доходу та мінімізації загальної тривалості виконання робіт за умови обмежених ресурсів сервісної компанії. Частина планування побудовано на основі сучасного підходу, а саме гнучкого JSSP (FJSSP), що дає змогу враховувати бізнес-процеси сервісної компанії, яка оперує парком практично однакових мобільних майстерень. Особливістю моделі є використання змінних, що індексовані за часом. Оскільки на момент ухвалення рішення дохід невідомий, то він моделюється за допомогою нечіткої логіки 2-го типу, яка імітує реальний процес ухвалення рішень експертом. Нечітка логіка 2-го типу дає змогу, на відміну від нечіткої логіки 1-го типу, урахувати як невизначеність експертів у попередній оцінці доходу, так і непевність у визначенні функції належності. Нами окреслено також напрямки подальших досліджень, серед яких багатокритеріальна оптимізація з урахуванням простоїв обладнання, чисельні експерименти на реальних і синтезованих даних. У подальших дослідженнях виникне необхідність подолання підвищеної складності обчислення нечіткої логіки 2-го типу, щодо уникнення якої запропоновано декілька ідей.

Ключові слова: *предиктивне технічне обслуговування, методи на основі даних, нечітка логіка 2-го типу, гнучке планування роботи цеху, лінійне програмування із цілими числами.*

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

DEFERRED RENDERING METHODOLOGIES ANALYSIS AND COMPARISON

Real-time photorealistic rendering has always been one of the long-standing goals in computer graphics. In particular, this objective often goes down to low-latency photorealistic two-dimensional image generation based on analytic three-dimensional environment description and spectator location and view parameters using a typical personal computer with consumer graphics hardware with relatively limited computing capability. The fundamental problems one has to solve to perform rendering are main view visibility determination and shading, i. e. finding the environment surface element that is observed in the particular pixel of the resulting image and calculating the amount of light this surface element reflects in the observer's direction for all image's pixels. Due to the computation capability constraints of the target hardware, it is crucial to perform these operations as efficiently as possible.

This work is aimed at analysis and experimental comparison of the modern methodologies used to tackle main view visibility determination and shading tasks. Due to modern lighting calculation methods' complexity and high geometric detail of the virtual environments, forward rendering has become impractical as a general-purpose rendering approach and only remains in use when specific geometry or material types are involved. Deferred rendering with G-buffer generation has become a state-of-the-art solution for products demanding high visual quality and fidelity. G-buffer generation can be performed in different ways including rasterized G-buffer generation and visibility buffer method. A theoretical overview and comparison of these techniques are presented in this work. Also, we use a demonstration application we implemented to perform an experimental comparison of G-buffer generation techniques in various conditions. Our experimental results can be used as a guidance when designing production rendering solutions.

Key words: computer graphics, real-time rendering, forward rendering, deferred rendering, G-buffer, visibility buffer.

AMS 2020 classification: 68U05.

Introduction

Computer graphics is an important area of computer science nowadays with rendering being one of the key objectives. Rendering (Pharr, Jakob, & Humphreys, 2016) in computer graphics generates an image of the analytically described environment viewed by a certain observer with a given location, orientation, and view properties. It allows computer users to observe virtual visual worlds. It is thus an integral part of computer simulation and modeling software packages, animated films and cinematic effects production software, computer games, etc. Real-time photorealistic rendering is a particularly crucial and relevant case – it is intended to interactively produce a photorealistic image sequence for the environment and observer that can change over time based on behavior algorithms and user's input, which can provide a real-life-like visual experience to the user. On the other hand, a typical interactive graphics application has to produce frames with a high enough refresh rate (60 frames per second or more) using the consumer graphics adapter in a personal computer, smartphone, or other device the application runs on. Since the demand for higher visual fidelity keeps increasing and the computational capabilities and memory size of graphics processors are limited, real-time photorealistic rendering techniques and algorithms remain relevant topics for further research aimed at minimizing the rendering software packages latency (time to build one frame) and the amount of memory required for their operation.

Different rendering approaches have been used for interactive rendering. A forward rendering method was one of the first and simplest methods used. Later, deferred rendering (Deering et al., 1988) became an important rendering technique replacing forward rendering in most scenarios (Pharr, & Fernando, 2005, chap. 9; Karis, 2013). The idea behind it is to generate a G-buffer – a set of large screen images containing the visible surface elements data (position, normal, material properties, etc.), which are then used to calculate lighting by a separate deferred procedure. G-buffer generation is typically much faster than complete lighting calculation, so this scheme provides significant speedup to the rendering algorithm, especially when using complex lighting models and processing scenes with complex geometry. However, due to the increasing complexity of virtual environments, the G-buffer generation stage can still pose a performance issue so it is rather important to perform it most efficiently.

This article provides an overview of the forward and deferred rendering methodologies and compares their advantages and disadvantages. Also, we perform an experimental comparison of two common G-buffer generation algorithms which are rasterized G-buffer and visibility buffer generation, and analyze the applicability of both techniques in different scenarios. These testing results can be used as a guidance when designing production rendering solutions.

1. Forward and deferred rendering

In general, interactive rendering intends continuous image (frame) generation of a given dynamic environment from a certain dynamic observer's point of view. The frame generation process can be separated into two fundamental stages: visibility determination and shading. Visibility determination is finding the environment surface element observed in the particular pixel of the resulting frame. Shading means calculating the amount of light this surface element reflects in the observer's direction. Rendering a frame is essentially solving visibility determination and shading tasks for every pixel of the resulting frame. It should be noted that while both of these problems are important, visibility determination is a primary one since it is impossible to perform any shading and generate a resulting image if visible surface elements are unknown. On the other hand, shading itself can be simplified at the cost of the photorealism and visual quality of the rendering solution. Due to this, initial graphics hardware with three-dimensional graphics acceleration implemented robust visibility determination in the first place while shading solutions were gradually evolving with graphics cards' hardware capabilities.

Graphics processing units use rasterizers to tackle visibility determination problem. Rasterizers are fixed-function hardware blocks that take the two-dimensional triangle defined in image space as input and output a set of pixels of the image

the input triangle covers. To use these blocks, graphics processors implement a graphics pipeline, which consists of three common stages:

- Vertex processing stage: it takes geometry data as input (for example, three-dimensional environment geometry) and transforms it into two-dimensional geometry defined in the image space.
- Rasterization stage: it uses hardware rasterizers to determine a set of covered image pixels based on the vertex processing stage's output.
- Pixel processing stage: it performs a certain program for every resulting pixel of the rasterization stage which must output some value to the output image or discard the processing. To avoid occluded triangle color output, the graphics pipeline can maintain a minimum camera-surface distance image (depth buffer, Z-buffer) to discard or overwrite existing results based on distance comparison.

Forward rendering is a rendering method that uses a graphics pipeline to rasterize all the environment geometry and performs shading in the pixel processing stage in these pipelines. This rendering method was implemented in the first non-programmable three-dimensional graphics accelerators and remains relevant for specific rendering scenarios. Its most important advantages are simplicity, wide hardware support (it only requires a basic graphics pipeline to be performed), and low memory consumption, as only base output and depth buffers are necessary.

The rise of the graphics hardware computing power and the expansion of its technical capabilities opened possibilities to increase the geometry scale and complexity and switch to advanced lighting models, which drastically improved the visual quality of virtual environments (Burley, 2012; Walter et al., 2007). However, the forward rendering method turned out to be ineffective in these conditions because new lighting models required a significantly larger amount of calculations. One reason for this is triangle overlap which reduces pixel processing efficiency as it often leads to duplicate pixel processing in the overlap area when triangles are processed in the order from the farthest to the closest to the observer. This phenomenon becomes common when using the pixel processing stage to perform computation-intensive lighting models and rendering complex geometry with a severe amount of triangle overlap, resulting in a crucial performance penalty. Another reason for performance degradation is the unbalanced workload of the rasterizers and multiprocessors that perform computationally heavy lighting calculations, which reduces general graphics pipeline processing throughput. The deferred rendering method (Deering et al., 1988; Liktov, & Dachsbacher, 2012; Mara, McGuire, & Luebke, 2013) was developed to tackle these problems. Instead of immediately calculating the lighting during geometry rasterization, this method uses a graphics pipeline pixel processing stage to generate a G-buffer – a set of textures that match the main screen in size and contain visible surface element parameters corresponding to a particular pixel. After G-buffer generation, the lighting is calculated once for every pixel by a separate procedure that uses G-buffer data as input. Example G-buffer images are displayed in Figure 1; they preserve typical surface parameters including the color of the surface element, the surface normal, the type of light reflection model, etc. Determining these values is usually significantly cheaper than lighting calculation in terms of operations to be performed. Hence, the triangle overlap issue's influence on performance reduces drastically. Another key advantage of this deferred shading method is compatibility with advanced shading algorithms that are impossible to implement in a pixel processing stage. The main disadvantage of deferred shading is the increased memory consumption required for G-buffer maintenance and performance overhead produced by writing data to G-buffer textures and reading from it. These may become notable performance issues in case many different surface parameters are required for shading. G-buffer management also causes problems with anti-aliasing support for deferred rendering due to severe memory consumption and processing requirement increase. Various techniques were developed to decrease the negative effect of this issue (Kerzner, & Salvi, 2014; Crassin et al., 2016).

Recently, the method of generating a G-buffer using a visibility buffer (V-buffer) has become popular (Burns, & Hunt, 2013). This method addresses the shortcomings of the classical rasterized G-buffer deferred rendering, which poses a performance issue when using a G-buffer to store an extensive amount of data. It narrows the data written during the pixel processing stage when rasterizing environment geometry to the minimal identifier of the rasterized triangle – the resulting buffer containing these identifiers is called visibility buffer. This significantly speeds up the geometry rasterization stage since such identifiers need a small amount of memory for storage and usually do not require any calculations in the pixel processing stage to be obtained. The G-buffer generation is performed by a separate procedure that takes the visibility buffer as an input and uses environment data to reconstruct the visible surface element for a particular pixel and calculate necessary parameters. To implement full simultaneous access to all the environment data required, modern graphics processor programming model support is necessary which slightly limits the applicability of the visibility buffer method for the old hardware. Surface element reconstruction requires software repeat of the operations performed at the rasterization stage (for example, transforming the vertices of the triangle, determining the point of its intersection with the corresponding pixel's sight ray, and others), which are not necessary when using rasterized G-buffer method. Therefore, the visibility buffer approach efficiency may be lower on a small amount of geometry that does not cause significant triangle overlap during rasterization, or simple lighting models that do not require a large number of geometric parameters. Hence, choosing the optimal G-buffer generation technique is a subtle compromise that needs careful research and testing.

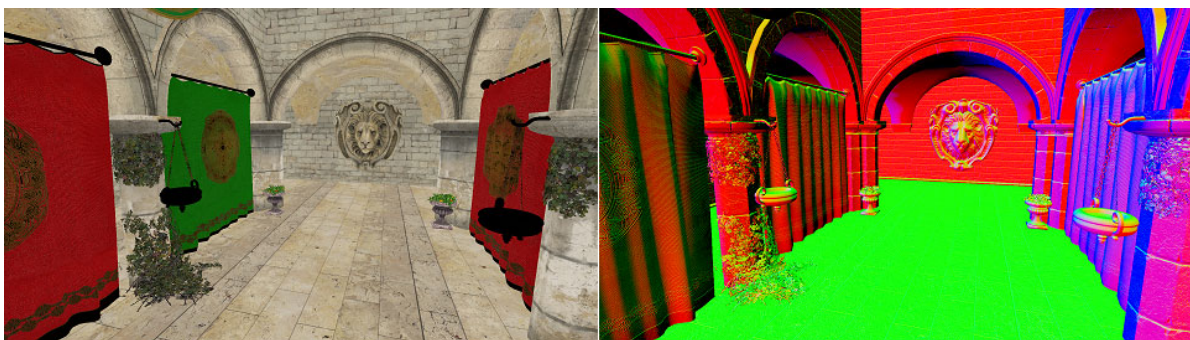


Fig. 1. G-buffer example. The left image contains diffuse color, right image preserves the surface normal (each of the 3D pixel normal's components is encoded in one of the image's color channels)

2. Testing and comparison

We implemented a demonstration application to perform an experimental comparison of two G-buffer generation methods described above. To achieve decent analysis variability, we performed testing with different options for the following parameters:

- Frame resolution: two common screen resolutions were used: 1920x1080 pixels (FullHD) and 2560x1440 (2K).
- Geometry amount: testing was performed using several common demonstration scenes containing different numbers of triangles: Cornell Box (36 triangles), Sponza (262267 triangles), and Bistro (2829801 triangles).
- G-buffer memory size: to simulate testing of different lighting models, three G-buffer layouts with common surface element parameters were implemented, each requiring 16, 32, and 64 bytes per pixel respectively.

The testing was performed on a personal computer with an NVidia RTX 2080 Super graphics card. Results are shown in Tab. 1 and 2. Each table corresponds to the screen resolution specified in the description; the table rows contain the time measurements for the specified scene, while the columns contain the measurements for the rendering using the specified G-buffer generation method with the G-buffer of a certain size. For example, the column labeled "Raster, 16 bpp" corresponds to rasterized G-buffer deferred rendering with a G-buffer size of 16 bytes per pixel, and the column "V-buffer, 64 bpp" corresponds to visibility buffer deferred rendering with a G-buffer size of 64 bytes per pixel. All time measurements are given in milliseconds.

Table 1

Time measurements for 1920x1080 resolution						
Time, ms	Raster, 16bpp	V-buffer, 16bpp	Raster, 32bpp	V-buffer, 32bpp	Raster, 64bpp	V-buffer, 64bpp
Cornell box	0.10	0.31	0.17	0.35	0.28	0.41
Sponza	0.35	0.37	0.58	0.48	0.89	0.54
Bistro	0.61	0.62	0.99	0.67	1.74	0.71

Table 2

Time measurements for 2560x1440 resolution						
Time, ms	Raster, 16bpp	V-buffer, 16bpp	Raster, 32bpp	V-buffer, 32bpp	Raster, 64bpp	V-buffer, 64bpp
Cornell box	0.21	0.54	0.35	0.62	0.52	0.75
Sponza	0.63	0.67	1.07	0.76	1.53	0.88
Bistro	1.10	1.03	1.78	1.16	3.42	1.24

As can be observed in the testing results, the visibility buffer turns out to be a powerful optimization when rendering complex environments with high geometry detail and using advanced lighting models that require large G-buffers. However, when dealing with smaller environment geometry amounts, classical deferred shading noticeably outperforms the visibility buffer approach due to surface element reconstruction calculations needed by the visibility buffer method which adds an unavoidable fixed performance price to it. Visibility buffer approach relative performance also tends to slightly improve with increasing screen resolution, because shortcomings of classic deferred rendering become more apparent when processing a larger number of pixels.

Discussion and conclusions

In this work, common rendering methodologies were reviewed and compared, including forward rendering and deferred rendering with its G-buffer generation algorithm variations – rasterized G-buffer and visibility buffer approaches. In addition to the theoretical discussion, a demonstration application was developed and an experimental comparison of G-buffer generation methods was performed in different data and working conditions. Experiments show that the rasterized G-buffer method is more efficient for simple graphics libraries designed to process simple visual environments, while the visibility buffer turns out to be superior when rendering complex scenes using advanced lighting models. Based on the testing data, the relative performance of different G-buffer generation approaches can be predicted for other rendering applications.

References

- Burley, B. (2012). *Physically-based shading at Disney*. Disney. https://media.disneyanimation.com/uploads/production/publication_asset/48/asset/s2012_pbs_disney_brdf_notes_v3.pdf
- Burns, C. A., & Hunt, W. A. (2013). The Visibility Buffer: A Cache-Friendly Approach to Deferred Shading. *Journal of Computer Graphics Techniques*, 2(2), 55–69. <http://jcgt.org/published/0002/02/04/>
- Crassin, C., McGuire, M., Fatahalian, K., & Lefohn, A. (2016). Aggregate G-Buffer Anti-Aliasing. *IEEE Transactions on Visualization and Computer Graphics*, 22(10), 2215–2228. <https://doi.org/10.1109/TVCG.2016.2586073>
- Deering, M., Winner, S., Schediwy, B., Duffy, C., & Hunt, N. (1988). The triangle processor and normal vector shader: a VLSI system for high performance graphics. *SIGGRAPH Computer Graphics*, 22(4), 21–30. <https://doi.org/10.1145/378456.378468>
- Karis, B. (2013). *Real Shading in Unreal Engine 4*. Epic Games. <https://cdn2.unrealengine.com/Resources/files/2013SiggraphPresentationsNotes-26915738.pdf>
- Kerzner, E., & Salvi, M. (2014). Streaming G-Buffer compression for multi-sample anti-aliasing. In *HPG '14: Proceedings of High Performance Graphics* (pp. 1–7). Eurographics Association.
- Liktor, G., & Dachsbacher, C. (2012). Decoupled deferred shading for hardware rasterization. In Spencer, S. N. (Ed.), *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (pp. 143–150). Association for Computing Machinery. <https://doi.org/10.1145/2159616.2159640>
- Mara, M., McGuire, M., & Luebke, D. (2013). *Lighting Deep G-Buffers: A single-pass, layered depth images with minimum separation applied to indirect illumination*. NVIDIA Corporation. https://research.nvidia.com/sites/default/files/pubs/2013-12_Lighting-Deep-G-Buffers/Mara2013DeepGBuffer.pdf
- Pharr, M., & Fernando, R. (2005). *GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation (Gpu Gems)*. Addison-Wesley Professional.
- Pharr, M., Jakob, W., & Humphreys, G. (2016). *Physically Based Rendering: From Theory to Implementation* (3rd ed.). Morgan Kaufmann Publishers Inc.
- Walter, B., Marschner, S. R., Li, H., & Torrance, K. E. (2007). Microfacet models for refraction through rough surfaces. In J. Kautz, & S. Pattanaik (Eds.), *Proceedings of the 18th Eurographics Conference on Rendering Techniques* (pp. 195–206). Eurographics Association.

Отримано редакцією журналу / Received: 06.12.24
Прорецензовано / Revised: 07.03.25
Схвалено до друку / Accepted: 13.03.25

Ростислав ПІКУЛЬСЬКИЙ, асп.
ORCID ID: 0000-0001-8973-8084
e-mail: rostyslav.pikulsky@gmail.com
Київський національний університет імені Тараса Шевченка, Київ, Україна

АНАЛІЗ І ПОРІВНЯННЯ МЕТОДОЛОГІЙ ВІДКЛАДЕНОГО РЕНДЕРИНГУ

Фотореалістичний рендеринг у реальному часі є однією з основних проблем у сфері комп'ютерної графіки. На практиці вона часто зводиться до генерації анімованої послідовності фотореалістичних зображень з низькою затримкою на основі аналітичного опису тривимірного віртуального середовища та параметрів спостерігача з використанням типового персонального комп'ютера або іншого пристрою з графічним обладнанням, що має відносно обмежені обчислювальні можливості. Фундаментальними задачами, які необхідно розв'язати для виконання рендерингу, є визначення видимості головного виду та затінення для кожного пікселя зображення, тобто пошук елемента поверхні середовища, що спостерігаємо в конкретному пікселі результуючого зображення та обчислення кількості світла, яке цей елемент поверхні віддзеркалює в напрямку спостерігача. Через обмежені обчислювальні можливості цільового обладнання екстремально важливо виконувати ці операції якомога ефективніше.

Пропоновану роботу присвячено аналізу й експериментальному порівнянню сучасних методологій, які використовують для розв'язання завдань визначення видимості та затінення головного виду. Через складність сучасних методів розрахунку освітлення та високу геометричну деталізацію віртуальних середовищ прямий рендеринг поступово втрапив практичність як загальноновживаний підхід до рендерингу та нині використовується лише в разі необхідності оброблення специфічних типів геометрії та матеріалів. Відкладений рендеринг з використанням G-буфера став найактуальнішим рішенням для графічних додатків, що потребують високої візуальної якості та точності. Генерація G-буфера може бути виконана різними способами, такими як метод растеризації G-буфера та метод буфера видимості. У дослідженні представлено також теоретичний огляд і порівняння цих методик. Крім того, розроблено демонстраційну програму та застосовано її для здійснення експериментального порівняння методів генерації G-буфера в різних умовах. Результати експериментів можна використовувати для прогнозування ефективності зазначених методів під час розроблення бібліотек рендерингу.

Ключові слова: комп'ютерна графіка, рендеринг у реальному часі, прямий рендеринг, відкладений рендеринг, G-буфер, буфер видимості.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 519.8

DOI: <https://doi.org/10.17721/1812-5409.2025/1.20>Volodymyr POLISHCHUK¹, DSc (Engin.), Prof.

ORCID ID: 0000-0003-4586-1333

e-mail: volodymyr.polishchuk@uzhnu.edu.ua

Yurii MLAVETS¹, PhD (Phys. & Math.), Assoc. Prof.

ORCID ID: 0000-0002-1480-9017

e-mail: yurii.mlavets@uzhnu.edu.ua

Lukáš DANKO², PhD.

ORCID ID: 0000-0001-7404-8857

e-mail: danko@utb.cz

Iwona WŁOCH³, PhD, Prof.

ORCID ID: 0000-0002-9969-0827

e-mail: iwloch@prz.edu.pl

¹Uzhhorod National University, Uzhhorod, Ukraine²Tomas Bata University in Zlin, Zlin, Czech Republic³Rzeszow University of Technology, Rzeszów, Poland

TECHNOLOGY FOR EVALUATING AND SELECTING COWORKING SPACES ACCORDING TO CLIENTS' TARGET NEEDS

The main goal of this study is to develop a technology for evaluating and selecting coworking spaces (CS) according to the target needs of clients. For this purpose, the following were developed: a fuzzy model for evaluating coworking; a model for aggregating the output data of the fuzzy model for evaluating coworking and the target needs of clients; a validation of the study based on 5 coworking spaces is presented. The research is based on the apparatus of fuzzy sets, which allows one to increase the degree of validity of decisions. The input data are given in estimation intervals that consider all possible scenarios. An intellectual analysis of knowledge is used to process input information, the subjectivity of experts is revealed, and the final assessment is presented in quantitative form. The value of the model is that it allows one to obtain a quantitative assessment of CS from input fuzzy expert data, and the assessment procedure for experts is simple and natural. The model considers the users' goals regarding an environment that increases work productivity, namely: environmental goals, social goals, and innovation goals. At the model's output, we have an initial quantitative assessment and a ranking series of CS. The research that was conducted contributes to the development of CS, increasing its efficiency and compliance with market needs. It is also useful for freelancers, startup teams, and small businesses, who receive a tool for choosing a space that meets their needs in terms of location, budget, and conditions.

Keywords: multi-criteria evaluation of alternatives, fuzzy sets, fuzzy logic, coworking spaces, expert evaluation, decision maker (DM).

AMS 2020 classification: 03B52, 03E72, 03H10.

Introduction

With the spread of remote work, startups, and freelancing, the demand for coworking spaces (CS) has increased significantly. This necessitates a systematic approach to their evaluation and selection. CS customers have different goals: access to technological equipment, convenient location, and a conducive environment for teamwork or individual productivity. Therefore, it is important to create technology that takes these needs into account. CS evaluation includes subjective factors (e.g., comfort or interior style) that are difficult to quantify. Fuzzy logic-based technologies allow for efficient work with such data (Rudnitskii et al., 2021). Moreover, the growing number of CSs forces owners to look for ways to improve their services. The development of evaluation technology will help both customers and operators better understand the key factors for success. The use of modern technologies, such as data analysis and fuzzy logic models (Arun et al., 2024), ensures objectivity of assessment and adaptability to changing customer needs. Choosing the optimal CS contributes to the efficient use of resources, increased employee productivity, and customer satisfaction, which is important for small and medium-sized businesses.

When choosing a coworking space, numerous criteria are taken into account, such as price, technical equipment, accessibility, aesthetics, etc. Fuzzy logic allows you to effectively integrate these heterogeneous parameters into a single model (Fast, & Jansson, 2024). The fuzzy set theory allows you to process fuzzy, blurred data, representing them in the form of a degree of belonging to certain sets (for example, "high level of comfort" or "moderate cost"). It also helps to avoid hard limits (for example, "price up to \$ 100"), which is a drawback of traditional evaluation models. In addition, fuzzy logic allows you to model the real decision-making process of a client, where certain criteria may partially satisfy needs. For example, a coworking space can be "quite convenient" and "almost affordable", which is taken into account in the overall assessment. Fuzzy logic also allows for more accurate and well-founded evaluation results, reducing the risk of erroneous decisions that may arise due to overly generalized or strict criteria. The use of fuzzy set theory is an advanced approach that contributes to the development of innovative analysis methods and provides effective, flexible, and realistic modeling of the process of evaluating and selecting coworking spaces, which meets modern challenges and needs.

Modern research places coworking spaces within a broader category of "new workspaces" (Mariotti et al., 2024), which have emerged as a result of digitalization, although they are not limited to it. Coworking spaces share similar features with spaces such as Fab labs, Makerspaces, Hackerspaces, Living labs, Innovation hubs, but at the same time have their unique characteristics (Micek, Baycan, & Lange, 2024; Tomaz, & Tabrizi, 2024). The beginning of the coworking movement is associated with the creation of The San Francisco Coworking Space in 2005, which became a model for shared workspaces

in the "dot-com zone" (De Peuter, Cohen, & Saraco, 2017). This concept gained considerable popularity worldwide, especially before the COVID-19 pandemic in 2020, which significantly affected the industry due to quarantines and safety measures. The long-term consequences of the pandemic are still being analyzed, including its impact on new workspaces (Mariotti et al., 2022). This demonstrates the dynamism of the market, its vulnerability to external factors, and the need for new approaches to evaluating and managing coworking spaces.

Our study is also seen as a tool that allows entrepreneurs, local government representatives, and other stakeholders to participate in decision-making on the opening of new coworking spaces to promote economic, business, and social transformation. In the context of business decision support systems, many important works can be highlighted. For example: (Piwowarski, & Nermend, 2022) examines the challenges of applying multi-criteria methods in complex business processes; (Yalcin, Kilic, & Delen, 2022) focuses on the use of business analytics and multi-criteria modeling for effective decision-making by enterprises; (Ceballos et al., 2021) presents a strategy that combines data and business rules based on multi-criteria analysis; (Wu et al., 2022) develops an integrated multi-criteria model for selecting a socially responsible investment portfolio; in the works (Buchnik, & Frenkel, 2023; Lin et al., 2022) a hybrid assessment model was created for the analysis of smart office buildings.

CS also plays an important role in supporting refugees affected by the war in Ukraine, providing them with workspace, social integration, and access to necessary resources. Coworking spaces often organize training, seminars, and workshops (Di Marino et al., 2023), helping refugees improve their professional skills, learn new technologies, or adapt to the local labor market. Coworking spaces often have resources for business development, mentoring, and legal advice, which helps refugees start their initiatives or businesses even in new conditions (Hendriks et al., 2024). The working environment provided by coworking spaces helps reduce stress and feelings of isolation, promoting psychological recovery (Figueiredo, Dierks, & Ferreira, 2024). However, there is no mathematically based research on the evaluation and selection of coworking spaces taking into account the target needs of clients.

The main goal of this study is to develop a technology for evaluating and selecting coworking spaces according to the target needs of clients. To do this, it is necessary to develop: a fuzzy model for evaluating coworking; a model for aggregating the output data of the fuzzy model for evaluating coworking and target needs of clients; a system of criteria for evaluating coworking spaces and a fuzzy knowledge base for evaluating coworking spaces according to the target needs of DM.

From the above, we can formulate the scientific hypothesis of this study as follows. If the selected coworking space satisfies all the needs of the employee, ensuring his development, then we can say about the high assessment of coworking, obtained on the basis of the constructed fuzzy model that takes into account the target needs of employees.

The research is based on the fuzzy set apparatus, which allows to increase the degree of validity of decisions. Input data is given in assessment intervals that take into account all possible scenarios. Intellectual knowledge analysis is used to process input information and reveal the subjectivity of experts, and the final assessment is presented in quantitative form. The value of the model is that it allows us to obtain a quantitative assessment of coworking spaces based on input fuzzy expert data, and the assessment procedure for experts is simple and natural. Input data is processed by appropriate fuzzy models, revealing the subjectivity of experts, and setting the parameters of models and target needs of clients makes it impossible for system analysts or experts to have a subjective influence on the assessment process and the final result of choosing a coworking space by clients. At the output of the model, we have an initial quantitative assessment and a ranking series of coworking spaces for selection by clients, taking into account their target needs.

The conducted research contributes to the development of coworking spaces, increasing their efficiency and compliance with market needs. It is also useful for freelancers, startup teams, and small businesses, who receive a tool for choosing a space that meets their needs in terms of location, budget, and conditions. It also helps coworking owners understand customer needs, optimize services, and increase competitiveness. Consulting companies can use the model to develop strategies for creating and managing coworking spaces. Therefore, the research contributes to the development of coworking spaces, increasing their efficiency and compliance with market needs.

The research is organized as follows: in section 2, a formal statement of the problem is formulated, a fuzzy model for evaluating coworking is described; a model for aggregating the output data of the fuzzy model for evaluating coworking spaces and target customer needs; a system of criteria for evaluating coworking spaces, as well as a fuzzy knowledge base for evaluating coworking spaces according to DM target needs. Section 3 presents the validation of the study based on 5 coworking spaces. Section 4 analyzes and discusses the results of the study, presents scientific and practical conclusions, and identifies areas for further research.

1. Main definitions and results

Formally, the technology of evaluation and selection of coworking spaces according to the target needs of clients can be presented as follows:

$$\{C, K_E, K_S, K_I, M_C, G_E, G_S, G_I, M_A | f\}. \quad (1)$$

Where: $C = \{c_1; c_2; \dots; c_n\}$ – a set of CS in a certain region for evaluation and selection by customers; K_E – a set of evaluation criteria for the coworking facility itself, this is its office infrastructure, location, and environmental policy; K_S – a set of evaluation criteria for coworking entities, social atmosphere, and knowledge exchange; K_I – a set of criteria for evaluating innovative potential, this is new funding opportunities, new opportunities for business, improving a higher level of productivity; M_C – a fuzzy coworking evaluation model; G_E – ecological goal (importance of the object); G_S – social goal (importance of the subject); G_I – innovative goal (interaction of subjects and business); M_A – a model for aggregating output data of a fuzzy M_C coworking evaluation model and goal user needs. As a result, we obtain the output score f based on which a ranking series of coworking spaces are built for customers to choose from, considering their needs.

Since our task is in the area of decision-making with the help of expert evaluation, we will introduce the following management subjects: a system analyst is a person who configures and manages the evaluation process; an expert forms a set of evaluation criteria for the relevant types of coworking spaces, selects a set of alternative CS and evaluates them; a decision-maker (DM) is a user (customer) who makes a decision regarding the choice of a coworking space, he sets his own target needs and, if necessary, enters additional parameters for the choice.

We will illustrate the technology of evaluation and selection of CS following the goal needs of users in the form of a structural diagram (Fig. 1).

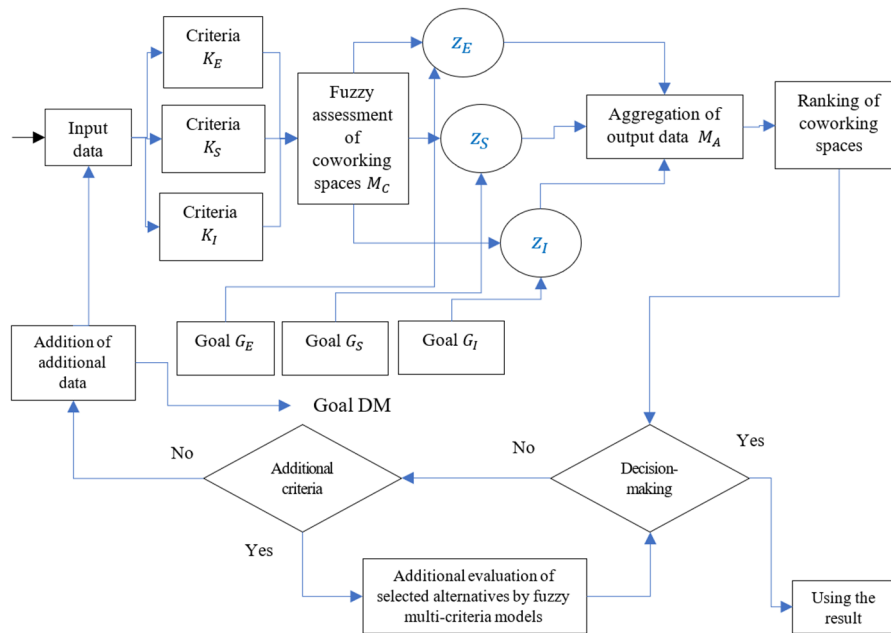


Fig. 1. Structural diagram of the technology evaluation and selection of CS

Figure 1 reflects the structural diagram of CS evaluation, based on the developed models, available input data, and the goals of DM. After evaluating and deriving a ranking series of CS, a decision is made about the choice. If the solution does not satisfy the DM, it is possible to involve additional criteria for the evaluation and use fuzzy models (Kelemen, Gavurova, & Polishchuk, 2022; Polishchuk, 2019). Another option is to review the DM's goal needs, change the parameters of the evaluation models, or involve new data by the system analyst. At the output for choosing a DM, there will be a series of coworking spaces and their aggregated normalized evaluations to increase the degree of validity of decisions.

The mathematical model for evaluating and selecting coworking spaces following the goal needs of customers will be presented in the form of two sub models: fuzzy evaluation of coworking spaces according to the proposed set of criteria and data aggregation using the goal needs of customers.

M_C is proposed – a fuzzy coworking evaluation model, which is based on the expert's opinions, in a certain range of coworking quality values according to certain indicators. At the first stage, the expert analyses a set of coworking spaces C , using his own experience and expresses his reasoning according to K_E, K_S, K_I sets of evaluation criteria.

In practice, people express their thoughts not in clear discrete values, but in intervals. For example: "the rating of coworking according to some criterion should be close to a number", or "the rating of coworking according to a certain criterion will most likely lie within certain limits", etc. Therefore, it is advisable to set the rating x according to each criterion in some interval of numbers $x \in [a_0; b_0]$. Let the expert express his opinion on some criterion according to one of the linguistic statements $CL = \{A; B; C; D\}$, where: $A = \{\text{coworking assessment according to a certain criterion is within the range from } a_1 \text{ to } b_1\}$; $B = \{\text{the evaluation of coworking according to a certain criterion is close to the number } b_1\}$; $C = \{\text{the evaluation of coworking according to a certain criterion is not greater than the number } b_1\}$; $D = \{\text{coworking assessment according to a certain criterion is not less than the number } b_1\}$, $\{a_1; b_1\} \in [a_0; b_0]$. The set of linguistic statements is open and, if necessary, the system analyst can supplement or change it.

Next, all fuzzy statements are described using fuzzy set theory and membership functions.

Statement $A = \{\text{coworking assessment according to a certain criterion is within the range from } a_1 \text{ to } b_1\}$ is treated as a fuzzy set and is formalized using a trapezoidal membership function (Kelemen, Gavurova, & Polishchuk, 2022):

$$\mu_A(x) = \begin{cases} 0, & x \leq a_0; \\ \frac{x-a_0}{a_1-a_0}, & a_0 \leq x < a_1; \\ 1, & a_1 \leq x < b_1; \\ \frac{b_0-x}{b_0-b_1}, & b_1 \leq x < b_0; \\ 0, & x \geq b_0. \end{cases} \quad (2)$$

The following linguistic statement $B = \{\text{the evaluation of coworking according to a certain criterion is close to the number } b_1\}$, is formalized using a triangular membership function (Kelemen, Gavurova, & Polishchuk, 2022):

$$\mu_B(x) = \begin{cases} 0, & x \leq a_0; \\ \frac{x-a_0}{b_1-a_0}, & a_0 < x \leq b_1; \\ \frac{b_0-x}{b_0-b_1}, & b_1 < x < b_0; \\ 0, & x \geq b_0. \end{cases} \quad (3)$$

Statement $C = \{\text{the evaluation of coworking according to a certain criterion is not greater than the number } b_1\}$, is characterized in fuzzy sets by a Z-like type of membership function (Kelemen, Gavurova, & Polishchuk, 2022):

$$\mu_C(x) = \begin{cases} 1, & 0 < x \leq b_1; \\ \frac{b_0 - x}{b_0 - b_1}, & b_1 < x \leq b_0; \\ 0, & x > b_0. \end{cases} \quad (4)$$

A final linguistic consideration $D = \{\text{coworking assessment according to a certain criterion is not less than the number } b_1\}$ represents the uncertainties of the S-shaped membership function type (Kelemen, Gavurova, & Polishchuk, 2022):

$$\mu_D(x) = \begin{cases} 0, & x \leq a_0; \\ \frac{x - a_0}{b_1 - a_0}, & a_0 < x \leq b_1; \\ 1, & x > b_1. \end{cases} \quad (5)$$

Therefore, the statements of the system analyst regarding coworking indicators will not represent discrete values, but intervals of values, according to the selected rule $CL = \{A; B; C; D\}$.

In the next stage, we will move from fuzzy evaluation intervals to one-point evaluation. For this, let DM set the weighting coefficients for each evaluation criterion $w_{kj}, k = \{E; S; I\}; j = \overline{1, k_k}$, for example from the interval $[1; 10]$. Next, their normalization is carried out within the limits of the appropriate set of criteria K_E, K_S, K_I :

$$\alpha_{kj} = \frac{w_{kj}}{\sum_{j=1}^{k_k} w_{kj}}, k = \{E; S; I\}; \alpha_{kj} \in [0; 1]. \quad (6)$$

For data fuzzification, we will find a weighted sum to determine the aggregated expert opinion within one set of evaluation criteria K_E, K_S, K_I :

$$\delta_{ki}(x) = \sum_{j=1}^{k_k} \mu_{CL_{kj}}^i(x) \cdot \alpha_{kj}, k = \{E; S; I\}, i = \overline{1, n}. \quad (7)$$

Thus, we will obtain for each set of criteria K_E, K_S, K_I a continuous set of values $\{\delta_{Ei}(x); \delta_{Si}(x); \delta_{Ii}(x)\} \in [0; 1]$ on the interval of evaluation points $x \in [a_0; b_0]$, according to the evaluated set of coworking spaces, $i = \overline{1, n}$. For the aggregated evaluation of the expert opinion $x_k \in [a_0; b_0]$ of the corresponding set of criteria $k = \{E; S; I\}$, the membership function $\theta(\varepsilon_{ki})$ of the maximum value of the weighted sum $\max_x \delta_{ki}(x) = \varepsilon_{ki}$ is selected. As a result, an aggregated conclusion of the expert's reasoning based on a set of criteria is obtained K_E, K_S, K_I .

At the third stage, the aggregated conclusions on sets of criteria are normalized according to the following formula:

$$m_k(c_i) = \frac{\varepsilon_{ki} \cdot \theta(\varepsilon_{ki})}{b_0}, k = \{E; S; I\}, i = \overline{1, n}. \quad (8)$$

The obtained scores $m_k(c_i) \in [0; 1]$ characterize the level of quality of coworking, respectively, evaluated according to the sets of criteria K_E, K_S, K_I . The higher the score, the better the space for work, which provides not only comfortable working conditions but also opportunities for innovative development in a society of like-minded people. Thus, there was a transition from fuzzy linguistic conclusions of the expert to a quantitative assessment. Adequate use of the fuzzy sets increases the degree of validity of future decisions.

Thus, there was a transition from the linguistic fuzzy conclusions of the expert to a quantitative assessment. Adequate use of the apparatus of fuzzy sets increases the degree of validity of future decisions.

During the expert evaluation, the expert formed a set of evaluation criteria and evaluated a set of coworking spaces based on them. Then, at the next – fourth stage, the DM must make a choice, considering its target needs G_E, G_S, G_I . Such goals are correlated according to the set of evaluation criteria K_E, K_S, K_I and are proposed by the authors, based on the experience in the subject area.

Since the task of evaluating alternative options of coworking spaces, consisting of three goals, then the vectors $m_E(C) = \{m_E(c_1), m_E(c_2), \dots, m_E(c_n)\}$, $m_S(C) = \{m_S(c_1), m_S(c_2), \dots, m_S(c_n)\}$ and $m_I(C) = \{m_I(c_1), m_I(c_2), \dots, m_I(c_n)\}$, are projected on the three-dimensional coordinate system x, y, z respectively. For each coworking space, the coordinates for the goals G_E, G_S, G_I are obtained in the form: $(m_E(c_1), m_S(c_1), m_I(c_1)), (m_E(c_2), m_S(c_2), m_I(c_2)), \dots, (m_E(c_n), m_S(c_n), m_I(c_n))$.

Next, a three-dimensional vector of DM goals $T^* = (A_1, A_2, A_3)$, is considered, which considers his wishes regarding the value of alternative coworking spaces, according to goals G_E, G_S, G_I . Let's model the vector of DM goals as follows (Kelemen, Gavurova, & Polishchuk, 2022; Polishchuk, 2019).

Let an object with 3 inputs and one output be analyzed:

$$U = (A_1, A_2, A_3). \quad (9)$$

Where U is the initial evaluation vector (u_1, u_2, u_3) , and its components can have values from the interval $[0; 1]$, and A_1, A_2, A_3 are input linguistic variables.

Qualitative terms from such term-sets are used to evaluate linguistic variables A_1, A_2, A_3 :

$$A_1 = (a_{11}, a_{12}, \dots, a_{1t_1}), A_2 = (a_{21}, a_{22}, \dots, a_{2t_2}), A_3 = (a_{31}, a_{32}, \dots, a_{3t_3}). \quad (10)$$

Knowledge of the vector of DM goals $T = (t_1, t_2, t_3)$ is obtained from the fuzzy knowledge base consisting of systems of logical statements - "If - Then, Else", which connect the values of the input variables A_1, A_2, A_3 with one of the possible values U .

$$\text{If } A_1 = a_{1t_1} \text{ and } A_2 = a_{2t_2} \text{ and } A_3 = a_{3t_3} \text{ Then } U = (u_1, u_2, u_3) \text{ Else ...} \quad (11)$$

Thus, the DM sets the linguistic desire for the vector of goals, which is translated into the vector of the initial quantitative and normalized assessment (u_1, u_2, u_3) . We denote accordingly $(u_1, u_2, u_3) = (t_1, t_2, t_3)$. A fuzzy knowledge base is proposed to evaluate coworking spaces according to the target needs of DM is as follows:

IF we have:

G_E – ecological goal (importance of the object): $\{a_{11}$ the need for the ratio of price and quality of coworking, then $u_1 = 0.4$; a_{12} priority need for a good location of the coworking space, then $u_1 = 0.6$; a_{13} the need for both high-quality office infrastructure and a good location, then $u_1 = 0.8$; a_{14} significant need for quality office infrastructure, good location and overview of environmental policy, then $u_1 = 1\}$.

AND G_S – social goal (importance of the subject): $\{a_{21}$ minimal need for a community of like-minded people, participation in joint projects, moral support, exchange of new ideas and knowledge, then $u_2 = 0.2$; a_{22} low need for a community of like-minded people, participation in joint projects, moral support, exchange of new ideas and knowledge, then $u_2 = 0.4$; a_{23} average need for a community of like-minded people, participation in joint projects, moral support, exchange of new ideas and knowledge, then $u_2 = 0.6$; a_{24} high need for a community of like-minded people, participation in joint projects, moral support, exchange of new ideas and knowledge, then $u_2 = 0.8$; a_{25} very high need for a community of like-minded people, participation in joint projects, moral support, exchange of new ideas and knowledge, then $u_2 = 1\}$.

AND G_I – innovative goal (interaction of subjects and business for the development of innovations): $\{a_{31}$ is not interested in the innovative component, then $u_3 = 0.2$; a_{32} may be some innovative possibilities, then $u_3 = 0.5$; a_{33} are interested in new opportunities for business and financing, then $u_3 = 0.8$; a_{34} the need for the best opportunities for the development of one's innovative product and its financing, then $u_3 = 1\}$;

THEN we can formulate the logical statement as follows $U = (u_1, u_2, u_3)$.

The system analyst can change quantitative levels or rules in goals. Therefore, the knowledge base is open, and the number of goals can be changed if necessary.

Next, M_A is proposed – a model for aggregating the raw data of the fuzzy coworking evaluation model M_C and goal customer needs.

For this, for all evaluated coworking spaces, there are values $Z_i = (z_{Ei}, z_{Si}, z_{Ii})$, $i = \overline{1, n}$, which characterize the relative estimates of the proximity of the evaluated coworking spaces to the vector of DM goals, for each individual goal G_E, G_S, G_I , while removing the issue of different evaluation scales (Polishchuk, 2019):

$$z_{Ei} = 1 - \frac{|u_1 - m_E(c_i)|}{\max\{u_1 - \min_i m_E(c_i); \max_i m_E(c_i) - u_1\}}, \quad (12)$$

$$z_{Si} = 1 - \frac{|u_2 - m_S(c_i)|}{\max\{u_2 - \min_i m_S(c_i); \max_i m_S(c_i) - u_2\}}, \quad (13)$$

$$z_{Ii} = 1 - \frac{|u_3 - m_I(c_i)|}{\max\{u_3 - \min_i m_I(c_i); \max_i m_I(c_i) - u_3\}}. \quad (14)$$

The evaluation of coworking spaces according to this approach will have meaning, with two or more alternative options.

Furthermore, for the aggregation of Z values, it is proposed to use the modelling of uncertainties of the "average value" form in a three-dimensional space, using a cone-shaped membership function in the space of estimates $[0; 1]$. The value of the center of the base of the cone will be the unit vector $(x_1^0; x_2^0; x_3^0) = (1; 1; 1)$, and scaling according to the coordinates of the vector Z_i was obtained experimentally – $(3; 3; 3)$. Then, the three-dimensional conical membership function will be given by the formula (Kelemen, Gavurova, & Polishchuk, 2022):

$$f_i = \begin{cases} 1 - \alpha_i, & \text{if } \alpha_i < 1 \\ 0, & \text{otherwise} \end{cases}, \quad (15)$$

$$\alpha_i = \frac{1}{3} \cdot \sqrt{(z_{Ei} - 1)^2 + (z_{Si} - 1)^2 + (z_{Ii} - 1)^2}, i = \overline{1, n}.$$

Thus, the initial estimates of f_i from the interval $[0; 1]$ respectively for n coworking spaces.

Based on the received evaluations, a ranking series of alternative options $C = \{c_1; c_2; \dots; c_n\}$ is built, and the DM chooses the best co-working according to its own target priorities:

$$C^* = \max_i f_i \quad i = \overline{1, n}. \quad (16)$$

The vector of DM goals ensures the construction of a ranking series of coworking space alternatives $C = \{c_1; c_2; \dots; c_n\}$, given by the evaluation vectors, while the safety of choosing alternative options according to the goal needs is increased.

Other important factors in choosing a coworking space can influence the decision-making process of the DM. For example, if an employee performs most of his work remotely, the focus should be on choosing a co-working space that is close to home or in a place that is easy to get to. Another example is if an employee regularly visits his clients, then their location is also important. In this case, from the obtained ranking series, the DM selects g ($g < n$) coworking spaces. Their DM is evaluated with the help of fuzzy multicriteria models previously developed by the authors (Kelemen, Gavurova, & Polishchuk, 2022; Polishchuk, 2019). Only additional important factors for the client are considered as a criterion set. At the same time, evaluations of alternatives should be measured f_i .

We present a system of input data for evaluating coworking spaces, which are the basis of the proposed fuzzy models of the researched problem. The criteria for evaluating coworking spaces are presented in the form of a question about substantiating the quality of a coworking space in terms of environmental friendliness, sociality, and innovativeness. Based on information and knowledge about coworking, the expert assigns a score from the interval $x \in [a_0; b_0]$ for each criterion according to the proposed fuzzy statements $CL = \{A; B; C; D\}$. The system of criteria is proposed as follows.

K_E – set of evaluation criteria for the coworking facility itself:

K_{E1} – the general design of the space includes all modern components of a co-working space, such as meeting rooms; halls for events (conference halls); individual boxes; telephone boxes; discussion booths (round tables); game areas; recreational/lounge areas; terraces; cafe/dining room and others.

K_{E2} – workplace quality: furniture made of ecological materials, availability of necessary equipment, availability of stationery, high-speed Internet.

K_{E3} – quality of additional services: room with office equipment, unlimited snacks/drinks or lunches, events, training, workshops, administrative support, virtual office, and technological platforms.

K_{E4} – the location of the co-working space represents a good transport network and infrastructure around the coworking space, such as restaurants, grocery stores.

K_{E5} – availability of support namely: administrative support, mail and package delivery, front desk and office management, event planning, catering and other services along with technical support to facilitate remote conferencing and presentation capabilities.

K_{E6} – ratio of price and quality of coworking space.

K_{E7} – coworking space provides an environmental policy, this is policy aimed at ensuring the effective use and reproduction of natural resources, environmental protection and ensuring the environmental safety of co-working space activities.

K_S – set of evaluation criteria for coworking entities:

K_{S1} – possibility of working in groups, exchange of new contacts.

K_{S2} – the possibility of working on joint or similar projects.

K_{S3} – availability of social contact, moral support, opportunities for communication, friendly atmosphere.

K_{S4} – brainstorming, exchange of new ideas, opportunities, knowledge, advice, and feedback between other customers of the coworking space.

K_I – set of criteria for evaluating innovative potential:

K_{I1} – availability of opportunities to create innovation (art object) in the form of a new product or service.

K_{I2} – availability of opportunities to increase turnover by promoting goods, services, and works of art, through the tools of industry events, social networks, quality marketing, etc.

K_{I3} – the possibilities of various financing tools for an innovative or artistic project, in the form of technical, legal, or economic support from the coworking space.

K_{I4} – availability of new opportunities for business from the coworking space side.

K_{I5} – a coworking space increases the productivity of work on an innovative or artistic project.

The given sets of criteria are open, and the model does not depend on the number of criteria. Experts can always add their metrics when customizing the rating model for different types of coworking spaces (Clifton et al., 2022).

2. Examples

The results of the research on the example of evaluating 5 coworking spaces $C = \{c_1; c_2; \dots; c_5\}$, the data of which are taken from open sources (Lange, & Bürkner, 2017). The calculation will be carried out based on a technology of evaluation and selection of co-working spaces according to the target needs of customers. For this purpose, an expert evaluation was conducted based on the given systems of criteria K_E, K_S, K_I . The evaluation was carried out by the authors within the field of innovative business, members of various commissions and competitions for evaluating grants, scientific, technical, and start-up projects. We set our statements on the interval of numbers $x \in [1; 10]$ using one of the linguistic statements CL. Also, let DM determine the weighting factors for each criterion. Input data and weighting factors are given in the Tab. 1.

Table 1

Input data from the expert assessment of coworking spaces							
System criteria	Name	Weight	Type of linguistic statement / estimation limits x				
			c ₁	c ₂	c ₃	c ₄	c ₅
K _E	K _{E1}	10	A / from 7 to 9	C / is not greater than 5	D / is not less than 3	A / from 5 to 7	B / close to 8
	K _{E2}	9	C / is not greater than 7	B / close to 9	B / close to 8	A / from 6 to 9	B / close to 8
	K _{E3}	7	D / is not less than 5	C / is not greater than 8	A / from 7 to 10	B / close to 9	B / close to 7
	K _{E4}	8	B / close to 7	B / close to 6	C / is not greater than 7	D / is not less than 7	D / is not less than 7
	K _{E5}	8	A / from 8 to 9	C / is not greater than 5	D / is not less than 5	A / from 8 to 9	A / from 7 to 9
	K _{E6}	9	C / is not greater than 6	A / from 7 to 9	C / is not greater than 6	D / is not less than 8	A / from 9 to 10
	K _{E7}	10	D / is not less than 6	B / close to 6	B / close to 8	A / from 5 to 8	A / from 6 to 9
K _S	K _{S1}	7	C / is not greater than 5	B / close to 6	C / is not greater than 8	A / from 5 to 9	B / close to 7
	K _{S2}	8	D / is not less than 7	D / is not less than 8	A / from 5 to 6	B / close to 9	C / is not greater than 6
	K _{S3}	7	A / from 7 to 9	A / from 6 to 8	C / is not greater than 8	A / from 7 to 9	D / is not less than 8
	K _{S4}	8	C / is not greater than 6	C / is not greater than 6	B / close to 5	C / is not greater than 6	A / from 3 to 5
K _I	K _{I1}	9	C / is not greater than 5	A / from 5 to 6	A / from 7 to 9	D / is not less than 9	C / is not greater than 5
	K _{I2}	10	A / from 7 to 9	B / close to 7	B / close to 5	A / from 7 to 9	B / close to 7
	K _{I3}	10	D / is not less than 6	C / is not greater than 8	B / close to 6	D / is not less than 8	B / close to 6
	K _{I4}	8	A / from 4 to 6	D / is not less than 7	A / from 4 to 6	A / from 6 to 9	D / is not less than 5
	K _{I5}	8	B / close to 7	C / is not greater than 7	B / close to 8	B / close to 9	A / from 5 to 7

First, weighting coefficients are normalized within the appropriate set of criteria K_E, K_S, K_I , according to formula (6). The result will be presented in the form of a vector: $\alpha_E = (0.16; 0.15; 0.12; 0.13; 0.13; 0.15; 0.16)$, $\alpha_S = (0.23; 0.27; 0.23; 0.27)$, $\alpha_I = (0.2; 0.22; 0.22; 0.18; 0.18)$. Next, we will move from the intervals of fuzzy evaluations of coworking spaces according to the criteria to a single point evaluation within the criteria system K_E, K_S, K_I . To do this, based on the input data, using the membership functions of criteria (2)–(5), we will construct graphs of experts' conclusions. And for data fuzzification, we will find the weighted sum within one set of evaluation criteria K_E, K_S, K_I according to formula (7). The results of the calculation, separately for the systems of criteria K_E, K_S, K_I , will be illustrated on the graphs of the functions built with the help of the specially designed software (Kelemen, Gavurova, & Polishchuk, 2022), Fig. 2.

For the aggregated assessment of the expert opinion of the corresponding system of criteria $k = \{E; S; I\}$, selected the membership function $\theta(\varepsilon_k)$ of the maximum value of the weighted sum $\max_x \delta_k(x) = \varepsilon_k$. Also, the aggregated conclusions on sets of criteria are normalized according to formula (8), the result will be presented in the form of Tab. 2.

Thus, we moved from the linguistic fuzzy conclusions of the expert to a quantitative assessment. Next, we move to the stage when DM sets its target needs G_E, G_S, G_I . If there is a significant need for high-quality office infrastructure, a good location and an overview of environmental policy ($u_1 = 1$) AND a high need for a community of like-minded people, participation in common projects, moral support, exchange of new ideas and knowledge ($u_2 = 0.8$) AND interested in new opportunities for business and financing ($u_3 = 0.8$) THEN $U = (1; 0.8; 0.8)$.

Further evaluation takes place according to the model of aggregation of output data M_A . First, for all evaluated co-working spaces, there are values $Z_i = (z_{Ei}, z_{Si}, z_{Ii}), i = \overline{1,5}$, according to formulas (12)–(14). We will present all the calculation results in Tab. 3.

Table 2

Results of aggregated findings of coworking spaces															
	c ₁			c ₂			c ₃			c ₄			c ₅		
	ε ₁	μ(ε ₁)	m(c ₁)	ε ₂	μ(ε ₂)	m(c ₂)	ε ₃	μ(ε ₃)	m(c ₃)	ε ₄	μ(ε ₄)	m(c ₄)	ε ₅	μ(ε ₅)	m(c ₅)
K _E	7	0.94	0.66	6	0.86	0.52	7	0.92	0.64	7	0.93	0.65	8	0.94	0.75
K _S	6	0.87	0.52	6	0.92	0.55	5	1	0.5	7	0.87	0.61	6	0.84	0.50
K _I	6	0.89	0.53	7	0.95	0.67	6	0.87	0.52	9	1	0.9	6	0.92	0.55

Table 3

Relative estimates of the proximity of the evaluated CS to the vector of DM goals

	c_1	c_2	c_3	c_4	c_5
Estimate z_E relative to the goal G_E	0.29	0	0.26	0.28	0.49
Estimate z_S relative to the goal G_S	0.07	0.17	0	0.36	0.01
Estimate z_I relative to the goal G_I	0.04	0.51	0	0.64	0.11

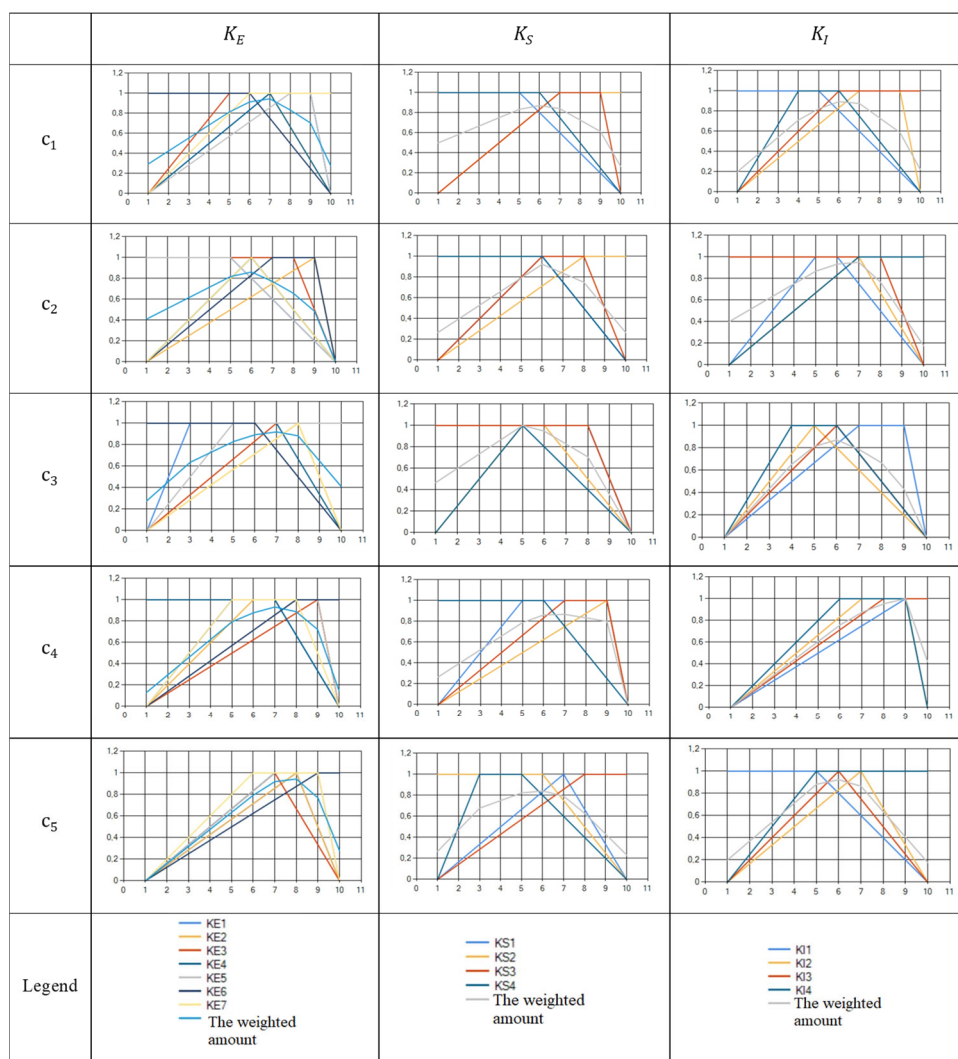


Fig. 2. A fragment of the calculation result

Next, a three-dimensional cone-shaped membership function according to formula (15) is used to aggregate Z values: $f_1 = 0.49$; $f_2 = 0.54$; $f_3 = 0.47$; $f_4 = 0.66$; $f_5 = 0.53$. Based on the received evaluations, a ranking series of coworking spaces is built, and the DM selects the best coworking space, (16): $(c_4; c_2; c_5; c_1; c_3)$. The results reveal that the best alternative solution for choosing a CS is – c_4 .

Discussion and conclusions

The research is based on the apparatus of fuzzy sets, which allows one to increase the degree of validity of decisions. The input data are given in estimation intervals that consider all possible scenarios. An intellectual analysis of knowledge is used to process input information, the subjectivity of experts is revealed, and the final assessment is presented in quantitative

form. The value of the model is that it allows one to obtain a quantitative assessment of CS from input fuzzy expert data, and the assessment procedure for experts is simple and natural. The input data are processed by appropriate fuzzy models, revealing that quality of additional services (workshops, virtual office, and tech platforms) and availability of support (administrative and technical support) are decisive factor in the case facility itself. The results highlight the importance of collaboration on joint projects and the availability of social contact (moral support and a friendly atmosphere) among the needs of users. We identified opportunities to create innovation, and spaces that increase the productivity in project-oriented production are among criteria for innovation potential in CS. At the output of the model, we have an initial quantitative assessment and a series of CS for users to choose from, considering their needs in a systematic manner.

The study presents a technology to evaluate and selecting CS following the target needs of customers to provide some information on the importance and definition of the needs and tensions related to these needs. The model considers the users' goals regarding an environment that increases work productivity, namely: environmental goal (importance of the object), social goal (importance of the subject), and innovation goal (interaction of subjects and business to develop innovation). A developed fuzzy knowledge base for evaluating coworking spaces for targeted DM needs can be easily adapted to various other DM goals, especially with increasing diversity among CS members. The obtained evaluation and the ranking series of alternative options for CS built on their basis reflects on the degree of validity for the decision. Furthermore, the model increases the degree of reasonableness of the choice of CS, and there is an opportunity to involve additional criteria for evaluation and use fuzzy multi-criteria models, revise the goal needs of DM, change the parameters of evaluation models for sustainability of ideas in communities.

A limitation of our study was the use of different types of membership functions for fuzzy CL statements and the aggregation of Z values. As is known, the choice of the type of membership functions and the use of different types of convolutions can lead to ambiguity in the results, but this does not affect the content of the evaluation, since the system's analyst can always adjust the levels of decision-making. Another limitation is related to the geographical context in which this study was conducted. Although the research was conducted on the territory of three countries: Ukraine, the Czech Republic, and Poland, it is also important to reproduce it in other countries. This allows us to compare the results, verify and unify the developed model.

The rationality of the obtained initial estimate f for building a ranking series of CS proves the advantages of the developed model. The reliability of the obtained results is confirmed by the results of research and is ensured by the justified use of the apparatus of fuzzy sets, intellectual analysis of knowledge and system approach towards creative communities. Thus, the hypothesis of this study was proved.

The main goal of the conducted research is to develop a technology decision support of evaluation and selection of coworking spaces according to the goal needs of customers. The technology is aimed at supporting the decision-making processes of users in CS regarding the evaluation and selection of them for the workplace, according to their own goal needs, impartial evaluation of coworking spaces, and increases the safety of their choice. At the same time, the following results were obtained:

- for the first time, a system of criteria for evaluating coworking spaces, a total of 16, was developed to justify their quality in terms of environmental friendliness, sociality, and innovativeness. The set of criteria is open, and the model does not depend on their number. Experts can always add their metrics when customizing the rating model for different types of coworking spaces;
- for the first time, a fuzzy model of evaluation and selection of coworking spaces was developed following the goal needs of customers. The model is based on the opinions of experts, in a certain interval of values of the quality of coworking according to certain indicators. At the output of the model, we receive an aggregated assessment based on an expert assessment of the infrastructure of the coworking facility, environmental policy, social atmosphere and knowledge exchange of coworking subjects, innovative potential, and also takes into account the goals of employees regarding their needs in the workspace in terms of environmental friendliness, sociality, and innovativeness. The fuzzy model consists of two sub-models: fuzzy evaluation of coworking according to the proposed set of criteria and data aggregation using goal customer needs;
- the research results were tested on the example of the evaluation of five coworking spaces, the real data of which were taken from open sources. At the same time, the adequacy of the fuzzy model, its sub-models, and system of criteria developed in the work was experimentally confirmed.

The obtained results demonstrate applied value and prove the scientific hypothesis of this study. The technology was presented to provide insight into decision making for entrepreneurial sustainability in CS by emphasizing the coworking facility, community, and the innovative potential.

Further research of the problem can be seen in the implementation of additional evaluation criteria, on their basis the development of a fuzzy multi-criteria model, and the adjustment of the system of criteria for clearly defined types of co-working spaces. The next stage, based on fuzzy models, will be the construction of software in the form of a web platform, which will contain an innovative tool for choosing the best CS, considering the goals and wishes of customers. Together, the developed fuzzy models, criteria information systems, and software support will constitute an innovative information technology for the evaluation and selection of co-working spaces, which will support the concepts of the European Green Deal and Industry 5.0.

Authors' contribution: Volodymyr Polishchuk – conceptualization, methodology, empirical research; Yurii Mlavets – conceptualization, methodology, empirical data collection and validation, analysis of sources, preparation of literature review or theoretical foundations of research; Lukáš Danko – conceptualization, methodology, empirical data collection and validation, analysis of sources, preparation of literature review or theoretical foundations of research; Iwona Włoch – conceptualization, methodology, empirical data collection and validation, empirical research.

References

- Arun, M., Efremov, C., Nguyen, V. N., Barik, D., Sharma, P., Bora, B. J., Kowalski, J., Le, H. C., Truong, T. H., & Cao, D. N. (2024). Fuzzy logic-supported building design for low-energy consumption in urban environments. *Case Studies in Thermal Engineering*, 64, 105384. <https://doi.org/10.1016/j.csite.2024.105384>
- Buchnik, T., & Frenkel, A. (2021). The lifestyles of millennial coworkers in urban spaces: the case of Tel-Aviv. *European Planning Studies*, 31(3), 528–553. <https://doi.org/10.1080/09654313.2021.1950641>

- Ceballos, R., Borrego, D., Gómez-López, M. T., & Gasca, R. M. (2021). Multi-criteria decision analysis for non-conformance diagnosis: A priority-based strategy combining data and business rules. *Expert Systems with Applications*, 183, 115212. <https://doi.org/10.1016/j.eswa.2021.115212>
- Clifton, J., Jacobs, G., Valeiras-Jurado, J., & Vandendaele, A. (2022). Governmentality-in-action. The pursuit of happiness and identity-work in graduate career coaching interaction. *Language and Dialogue*, 12(3), 335–359. <https://doi.org/10.1075/ld.00117.cli>
- De Peuter, G., Cohen, N. S., & Saraco, F. (2017). The ambivalence of coworking: On the politics of an emerging work practice. *European Journal of Cultural Studies*, 20(6), 687–706. <https://doi.org/10.1177/1367549417732997>
- Di Marino, M., Tabrizi, H. A., Chavoshi, S. H., & Sinitsyna, A. (2023). Hybrid cities and new working spaces – The case of Oslo. *Progress in Planning*, 170, 100712. <https://doi.org/10.1016/j.progress.2022.100712>
- Fast, K., & Jansson, A. (2024). Working in the comfort zone: Understanding coworking spaces as post-digital, post-work and post-tourist territory. *Digital Geography and Society*, 7, 100103. <https://doi.org/10.1016/j.diggeo.2024.100103>
- Figueiredo, S., Dierks, A., & Ferreira, R. (2024). Mental health screening in refugees communities: Ukrainian refugees and their post-traumatic stress disorder specificities. *European Journal of Trauma & Dissociation*, 8(1), 100382. <https://doi.org/10.1016/j.ejtd.2024.100382>
- Hendriks, T., Hassankhan, A., de Jong, J., & van Woerkom, M. (2024). Improving resilience and mental well-being among refugees residing at asylum centers in the Netherlands: A pre-post feasibility study. *Mental Health & Prevention*, 36, 200366. <https://doi.org/10.1016/j.mhp.2024.200366>
- Kelemen, M., Gavurova, B., & Polishchuk, V. (2022). A Complex Hybrid Model for Evaluating Projects to Improve the Sustainability and Health of Regions and Cities. *International Journal of Environmental Research and Public Health*, 19(13), 8217. <https://doi.org/10.3390/ijerph19138217>
- Lange, B., & Bürkner, H. J. (2017). Open workshops as sites of innovative socio-economic practices: approaching urban post-growth by assemblage theory. *Local Environment*, 23(7), 680–696. <https://doi.org/10.1080/13549839.2017.1418305>
- Lin, S.-H., Zhang, H., Li, J.-H., Ye, C.-Z., & Hsieh, J.-C. (2022). Evaluating smart office buildings from a sustainability perspective: A model of hybrid multi-attribute decision-making. *Technology in Society*, 68, 101824. <https://doi.org/10.1016/j.techsoc.2021.101824>
- Mariotti, I., Di Marino, M., & Bednář, P. (2022). *The COVID-19 pandemic and the future of working space*. Taylor & Francis. <https://doi.org/10.4324/9781003181163>
- Mariotti, I., Tomaz, E., Micek, G., & Mendez-Ortega, C. (2024). *Evolution of new working spaces: Changing Nature and Geographies*. Springer Nature. <https://doi.org/10.1007/978-3-031-50868-4>
- Micek, G., Baycan, T., & Lange, B. (2024). A Taxonomy of New Working Spaces. In I. Mariotti, E. Tomaz, G. Micek, & C. Méndez-Ortega (Eds.), *Evolution of New Working Spaces* (pp. 21–33). Springer, Cham. https://doi.org/10.1007/978-3-031-50868-4_3
- Piowowski, M., & Nermend, K. (2022). Issues of multi-criteria methods applicability supporting complex business process decision-making in management. *Procedia Computer Science*, 207, 4161–4170. <https://doi.org/10.1016/j.procs.2022.09.479>
- Polishchuk, V. V. (2019). Technology to improve the safety of choosing alternatives by groups of goals. *Journal of Automation and Information Sciences*, 51(9), 66–76. <https://doi.org/10.1615/JAutomatInfScien.v51.i9.60>
- Rudnitskii, A. G., Rudnitska, M. A., Tkachenko, L. V., & Pechuk, E. D. (2021). Application of fuzzy logic in finding the optimal filter in optoacoustics problems. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*, 1, 43–54. <https://doi.org/10.17721/1812-5409.2021/1.5>
- Tomaz, E., & Tabrizi, H. A. (2024). The Evolution of Non-traditional Workplaces: From Third Places to Hybrid Places. In I. Mariotti, E. Tomaz, G. Micek, & C. Méndez-Ortega (Eds.), *Evolution of New Working Spaces* (pp. 7–20). Springer, Cham. https://doi.org/10.1007/978-3-031-50868-4_2
- Wu, Q., Liu, X., Qin, J., Zhou, L., Mardani, A., & Deveci, M. (2022). An integrated multi-criteria decision-making and multi-objective optimization model for socially responsible portfolio selection. *Technological Forecasting and Social Change*, 184, 121977. <https://doi.org/10.1016/j.techfore.2022.121977>
- Yalcin, A. S., Kilic, H. S., & Delen, D. (2022). The use of multi-criteria decision-making methods in business analytics: A comprehensive literature review. *Technological Forecasting and Social Change*, 174, 121193. <https://doi.org/10.1016/j.techfore.2021.121193>

Отримано редакцією журналу / Received: 19.12.24

Прорецензовано / Revised: 10.03.25

Схвалено до друку / Accepted: 14.03.25

Володимир ПОЛІЩУК¹, д-р тех. наук, проф.
ORCID ID: 0000-0003-4586-1333
e-mail: volodymyr.polishchuk@uzhnu.edu.ua

Юрій МЛАВЕЦЬ¹, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0002-1480-9017
e-mail: yurii.mlavets@uzhnu.edu.ua

Лукаш ДАНКО², PhD,
ORCID ID: 0000-0001-7404-8857
e-mail: danko@utb.cz

Івона ВЛОХ³, д-р філософії, проф.
ORCID ID: 0000-0002-9969-0827
e-mail: iwloch@prz.edu.pl

¹ДВНЗ "Ужгородський національний університет", Ужгород, Україна

²Університет Томаша Бата в Зліні, Злін, Чеська Республіка

³Жешувський технологічний університет, Жешув, Польща

ТЕХНОЛОГІЯ ОЦІНЮВАННЯ ТА ВИБОРУ КОВОРКІНГ-ПРОСТОРІВ ВІДПОВІДНО ДО ЦІЛЬОВИХ ПОТРЕБ КЛІЄНТІВ

Основною метою пропонованого дослідження є розроблення технології оцінювання та вибору коворкінгів відповідно до цільових потреб клієнтів. Для цього були розроблені: нечітка модель оцінювання коворкінгу, модель агрегування вихідних даних нечіткої моделі для оцінювання коворкінгу та цільових потреб клієнтів, представлено валідацію дослідження на основі п'яти коворкінгів. Дослідження базується на апараті нечітких множин, що дає змогу підвищити рівень обґрунтованості рішень. Вхідні дані надано в інтервалах оцінювання, що враховують усі можливі сценарії. Для оброблення вхідної інформації використано інтелектуальний аналіз знань. У процесі роботи виявлено суб'єктивізм експертів, підсумкову оцінку виставлено в кількісному вигляді. Цінність моделі полягає в тому, що вона дає змогу отримати кількісну оцінку коворкінгів із вхідних нечітких експертних даних, а процедура оцінювання для експертів є простою та природною. Модель ураховує цілі користувачів щодо середовища, яке підвищує продуктивність праці, а саме: екологічні, соціальні й інноваційні цілі. На виході моделі ми маємо початкову кількісну оцінку та рейтинговий ряд коворкінгів. Проведені дослідження сприяють розвитку коворкінгів, підвищенню їхньої ефективності та відповідності потребам ринку. Це також корисно для фрилансерів, команд стартапів і малого бізнесу, які отримують інструмент для вибору простору, що відповідає їхнім потребам з позиції розташування, бюджету й умов.

Ключові слова: багатокритеріальна оцінка альтернатив, нечіткі множини, нечітка логіка, коворкінг, експертна оцінка, особа, що ухвалює рішення.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

UDC 004.8

DOI: <https://doi.org/10.17721/1812-5409.2025/1.21>Olena SKLIARENKO¹, PhD (Phys. & Math.), Assoc. Prof.

ORCID ID: 0000-0001-6555-1223

e-mail: olena.skliarenko@e-u.edu.uaAnatolii PASHKO², DSc (Phys. & Math.), Prof.

ORCID ID: 0000-0001-6944-8477

e-mail: anatoliipashko@knu.uaLeonid LYTVYENKO¹, PhD (Engin.)

ORCID ID: 0000-0002-0828-383X

e-mail: leonid.lytvynenko@e-u.edu.uaYanina KOLODINSKA¹, Senior Lecturer

ORCID ID: 0000-0002-3330-7565

e-mail: yanina.kolodinska@e-u.edu.ua¹Private Higher Educational Establishment "European University", Kyiv, Ukraine²Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

DEEP LEARNING MODELS FOR SOLVING THE PROBLEM OF COMPARING UNSTRUCTURED TEXTUAL INFORMATION

This study is devoted to the application of modern deep learning methods, such as universal multilingual sentence vector representations, for analyzing and comparing the similarity of unstructured text documents. The paper develops an algorithm for determining the similarity of multilingual texts using deep neural networks. The authors propose a supervision-free method for fast analysis and comparison of complex documents on similar topics written in different countries and in different languages. The presented methodology involves pre-processing documents, reformatting them for comparison, and applying deep learning methods to detect textual similarities. The algorithm is tested on the example of a comparative analysis of the similarity of the texts of the Tax Code of Ukraine and the German fiscal code. The Ukrainian fiscal code, presented in HTML format, was processed using the library for analyzing HTML documents and regular expressions, while the German fiscal code in PDF format required special analysis of unstructured content. In the course of the study, to improve the accuracy and overcome the identified limitations of universal multilingual vector representations, the Ukrainian document was translated into English using the Google Translate API. Solving the given task requires deep pre-processing of the data, which was carried out in this study. The article provides a step-by-step algorithm of the proposed methodology, which leads to the formation of a correlation table, which determines the similarity between similar multilingual text documents. Additionally, the authors support every step with their reasoning about how to interpret intermediate results and understand if the method should be adjusted to the need or the problems of the specific documents, like applying machine translation or advanced pre-processing.

The authors propose an algorithm for applying this method to the fiscal codes of European Union member states or other countries. In this research, we identified and described the limitations of this approach, demonstrated the verification of intermediate results, and supplemented the method with additional capabilities to increase its reliability. Recommendations are given on the possibilities of applying the proposed methodology, in particular, the development of algorithms for standardizing certain documents of the European Union.

Key words: *deep learning, universal vector representations of sentences, financial documents, fiscal code, regular expressions.*

AMS 2020 classification: 68T07, 68T50.

Introduction

We focus our research on the comparison of the Ukrainian fiscal code with the German fiscal code. These fiscal codes share common principles but have a very different structure and different sizes – and so making them difficult to compare manually.

As a part of the preparation for the research, we analyzed the frequency of the changes to the Ukrainian fiscal code, and we can see that it has received constant updates and amendments during the last decade shown on Fig. 1 which signifies the importance of our research not only from a scientific perspective, but also from the practical need.

Ukraine is actively moving to the European Union, and an important part of the movement is to harmonize Ukrainian legislative standards with the standards required in the European Union. At the same time every country, while following the general principles, adopts its own legislative documents which are not identical.

Natural language analysis has made many steps forward in recent years, and by applying modern Deep Learning techniques to document analysis and comparison we are offering a method to identify missing common parts. We do not aim to substitute human analysis, but aim to offer a robust, easily verifiable approach that will supplement it – so we're leaving generative models out of the scope of this research.

We're using Multilingual Universal Sentence Embeddings initially introduced by Google (see Basystiuk, & Melnykova, 2022; Shakhovska N., Basystiuk, & Shakhovska K., 2019) – which allows us to analyze texts in different languages.

In recent years, scientists have been using artificial intelligence technologies, in particular deep learning methods, to analyze unstructured data, recognize and compare natural language texts.

For example, works (Golomoziy et al., 2023; lanevych et al., 2023; Peng et al., 2023) investigated methods for classifying unstructured texts related to the processing of court documents. The transformation of unstructured text data into a statistical dataset allowed them to be used for analysis by scientists, journalists, government officials, politicians, and anyone else. This papers provide a literature review summarizing the methods and approaches used in the application of natural language processing (NLP) techniques specifically for the legal domain.

© Skliarenko Olena, Pashko Anatolii, Lytvynenko Leonid, Kolodinska Yanina, 2025

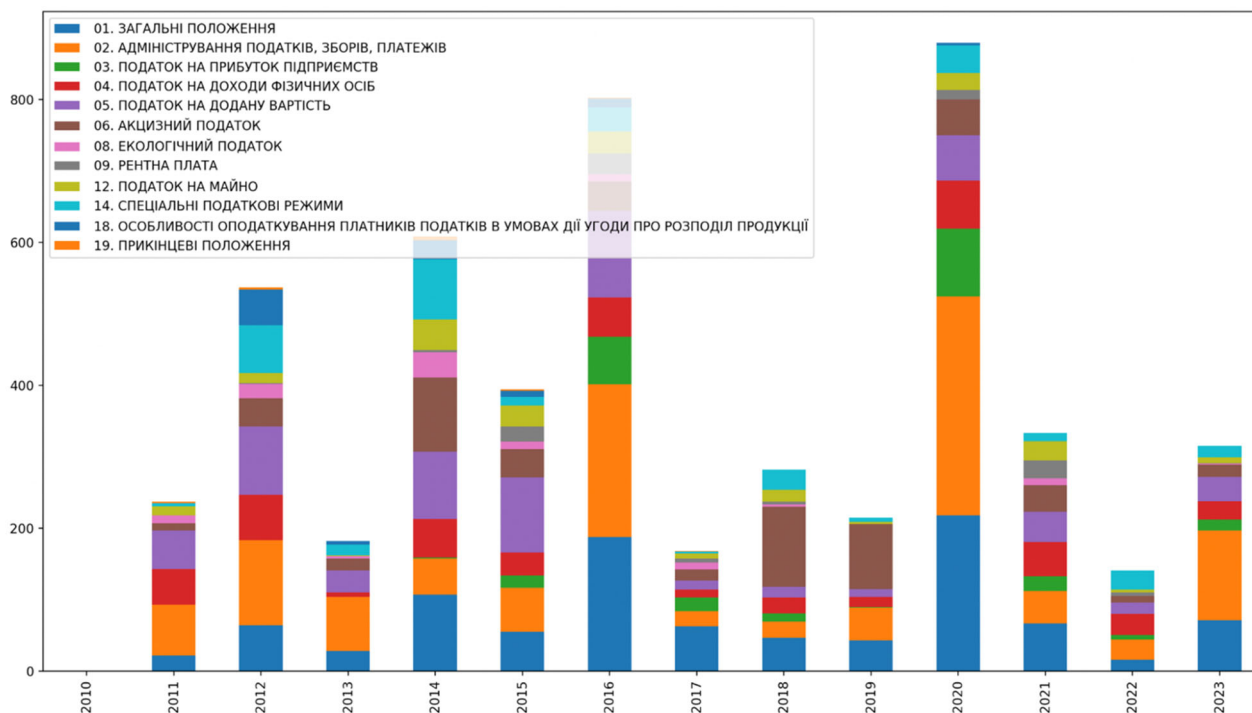


Fig. 1. Frequency of changes to different sections of the Ukrainian fiscal code

It should be mentioned that other researchers' scientific works on similar topics mainly focus on the study of methods and algorithms for classification, translation, and clustering. Unlike other studies, this article is devoted to the development of methods and algorithms for determining the similarity of the content of pro-lingual unstructured text documents.

One of the problems noted by researchers in improving object recognition is the availability of prepared large datasets required for training neural networks. The authors of this article propose a methodology for preparing, processing, and analyzing multilingual similar text documents to determine their similarity in content. The model proposed by the authors was tested by comparing the texts of fiscal codes of different countries. For the sake of purity of the experiment, the authors chose German texts rather than English-language documents, which is a fairly common practice.

1. Methodology

The solution to this problem requires deep pre-processing of relevant information and unstructured data, the so-called sub-deep (in the sense of deep learning) processing.

When processing unstructured data, sub-deep processing includes Understanding, Cleaning, Data Transformation, Feature Engineering, Splitting DataSet, Data Augmentation.

Let us present our methodology in the form of a step-by-step algorithm, considering the case of comparing the similarity of the texts of the fiscal codes of Ukraine and Germany.

Ukrainian and German fiscal codes, in addition to having different structures and sizes are also offered in different formats.

1.1. Data preparation. Ukrainian and German fiscal codes, in addition to having different structures and sizes are also offered in different formats.

Ukrainian fiscal code is an HTML document, where different sections are represented as `<span>` tags with the specific class names:

```
<span class=rvts15>РОЗДІЛ І. ЗАГАЛЬНІ ПОЛОЖЕННЯ</span></p>
<p class=rvps2><a name=n131></a>
<span class=rvts9>Стаття 1.</span> Сфера дії Податкового кодексу України</p>
```

Using the regular expressions, we split the whole document into small segments – one segment is one article of the code with all the related sub-articles. Then, using the Python library BeautifulSoup, we extract the text from HTML.

```
from bs4 import BeautifulSoup
def strip_html(html):
    return BeautifulSoup(html, 'html.parser').get_text(separator=' ')
```

We have also excluded a small number of articles which consisted of the tables or other non-textual information.

To parse the PDF file of the German fiscal code we used PyMuPDF library and its "fitz" Python package. This PDF file has a table of contents (ToC) – so we tried to utilize the built-in "get_toc()" function to get the references to different sections of the fiscal code. But the ToC of this PDF file is not a properly structured ToC in terms of the PDF format and only has visual resemblance. So, to properly split the sections of the file we had to write our own ToC parser, and then our own extractor of the fragments from the ToC.

Finally we used the fitz builtin "get_text()" method to get the textual representation of the articles, skipping tables and other non-textual information similar to what we did for the Ukrainian fiscal code.

As a result, we were able to have similarly structured and sized articles of both fiscal codes, which we saved in the JSON format for the convenience of our analysis.

1.2. Research design. One of the popular approaches to analysis of multilingual texts is based on the usage of the Multilingual Universal Sentence Embeddings (Basystiuk, & Melnykova, 2022). We used an open, pre-trained model published by Google, available at the TensorFlow hub (Basystiuk et al., 2021).

As we can see on Kaggle (Boyko et al., 2019; Yang et al., 2019), this model has gained popularity, but to see if it can be applied to our use-case and our data we performed an initial test and checked how it would behave on the samples from our texts by comparing vector distances which we calculated using the embedding from the model with our perception of what articles are close to each other:

```
import tensorflow as tf
import tensorflow_hub as hub
from sklearn.metrics.pairwise import cosine_similarity
embed = hub.load("https://www.kaggle.com/models/google/universal-sentence-encoder/frameworks/TensorFlow2/
variations/multilingual/versions/2")

sentences = [
    '1.1. Податковий кодекс України регулює відносини...',
    'A commercial business is an independent...',
    'Платниками податку є: 212.1.1. Особа, постійне представництво...',
    '(1) The Federal Central Tax Office assigns every taxpayer...',
    '(2) For the purposes of this subsection...',
    '...'
]

embeddings = embed(sentences)
similarities = cosine_similarity(embeddings)
```

Finally, to evaluate the approach at scale, we used the heat map visualization from the "seaborn" Python package. We are also excluding points below some threshold to make the visualization less cluttered:

```
import seaborn as sns
import matplotlib.pyplot as plt

similarities = cosine_similarity(embed(de), embed(ua))
df = pd.DataFrame(similarities).round(2)

significant = df[corr>=.30]
plt.figure(figsize=(12,8))
sns.heatmap(significant, cmap="Greens")
```

1.3. Data analysis. After performing analysis and applying the visualization to a subset of the articles we've seen a visualization shown on Fig. 2. This result has immediately raised our concerns.

On the visualization we see recognizable horizontal and vertical lines – a clear indication that some articles of one fiscal code correlate with many articles from other fiscal code and vice versa. It is a clear indication that the model picked up some trivial textual similarities, but not the similarities in meanings.

After reviewing the data again, we realized that our versions of Ukrainian and German fiscal codes also included references and dates of the changes of the codes and by removing them we've seen some improvement shown on Fig. 3. Nevertheless, the pattern of horizontal and vertical lines was still there and we continued analyzing our data.

We were including the name of the article in the text of the article, but because every name of the Ukrainian article included "Стаття" as a part of the name – that could impact the result too. In the end, we decided to remove the article names from the article text, because even if they correlate in different fiscal codes, we're interested in the text of the articles itself, because we want to assess not only the structure but also the meaning of the fiscal codes.

The pattern was still there, so we decided to also remove all the references to other articles of the fiscal codes because they were represented as numbers and were also influencing the similarity score.

Finally, after running our analysis and visualization again, we've seen improvements on the subset data, but when we applied everything to the whole data set we've seen a picture shown on the Fig. 4. There was a clear vertical line on the chart, but after manual review we didn't find the correlation the chart showed us.

So we decided to make a more radical move and used machine translation to bring both fiscal codes to the same language. We used Google Translate API for the translation, it supports Ukrainian and works relatively well. We had to write a script to split articles into the segments to fit Google API limitation on the request length, but other than that we didn't experience problems with the API. Because Google charges money for the API access, we translated the initial, non-cleaned version of the Ukrainian fiscal code so that we're able to repeat any post processing without calling the Google Translation API again. But the tests showed that our initial cleanup made sense for the translated text as well, so we repeated it in full for the translated version.

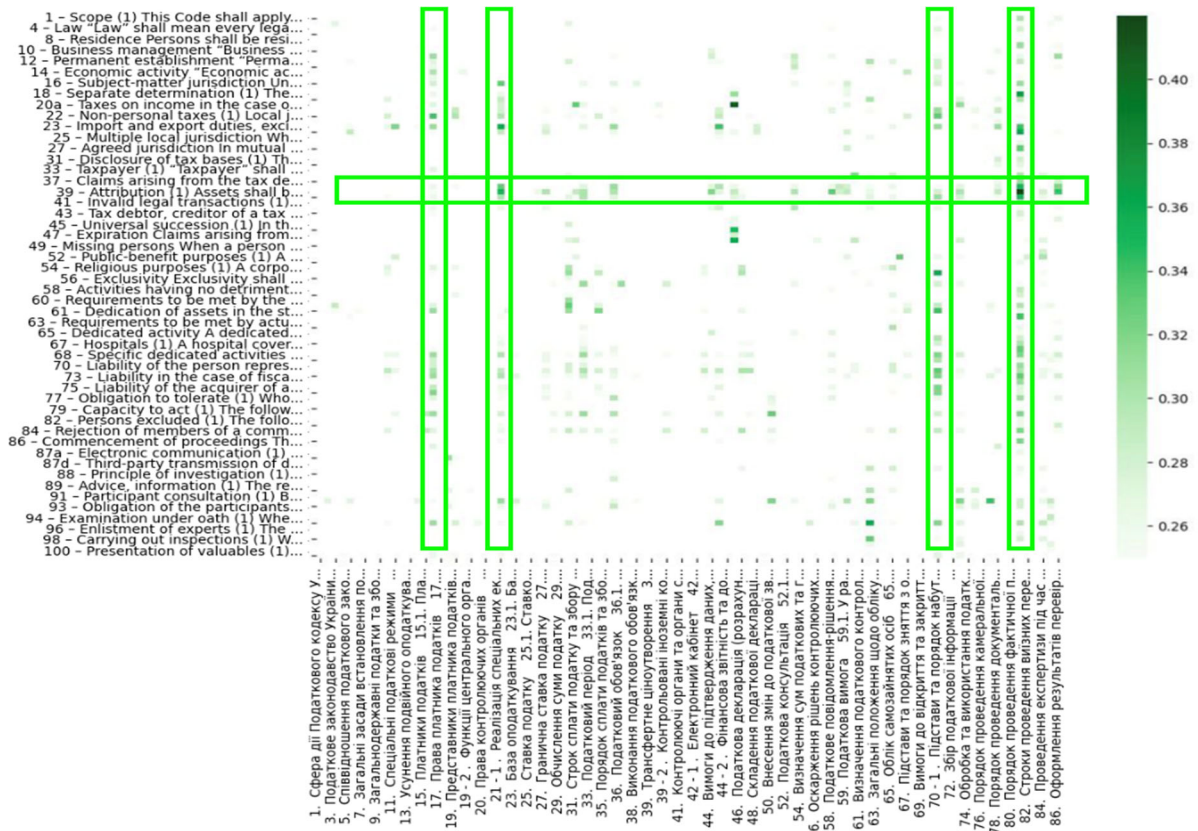


Fig. 2. Cosine similarity of the subset of articles from the Ukrainian and German fiscal codes

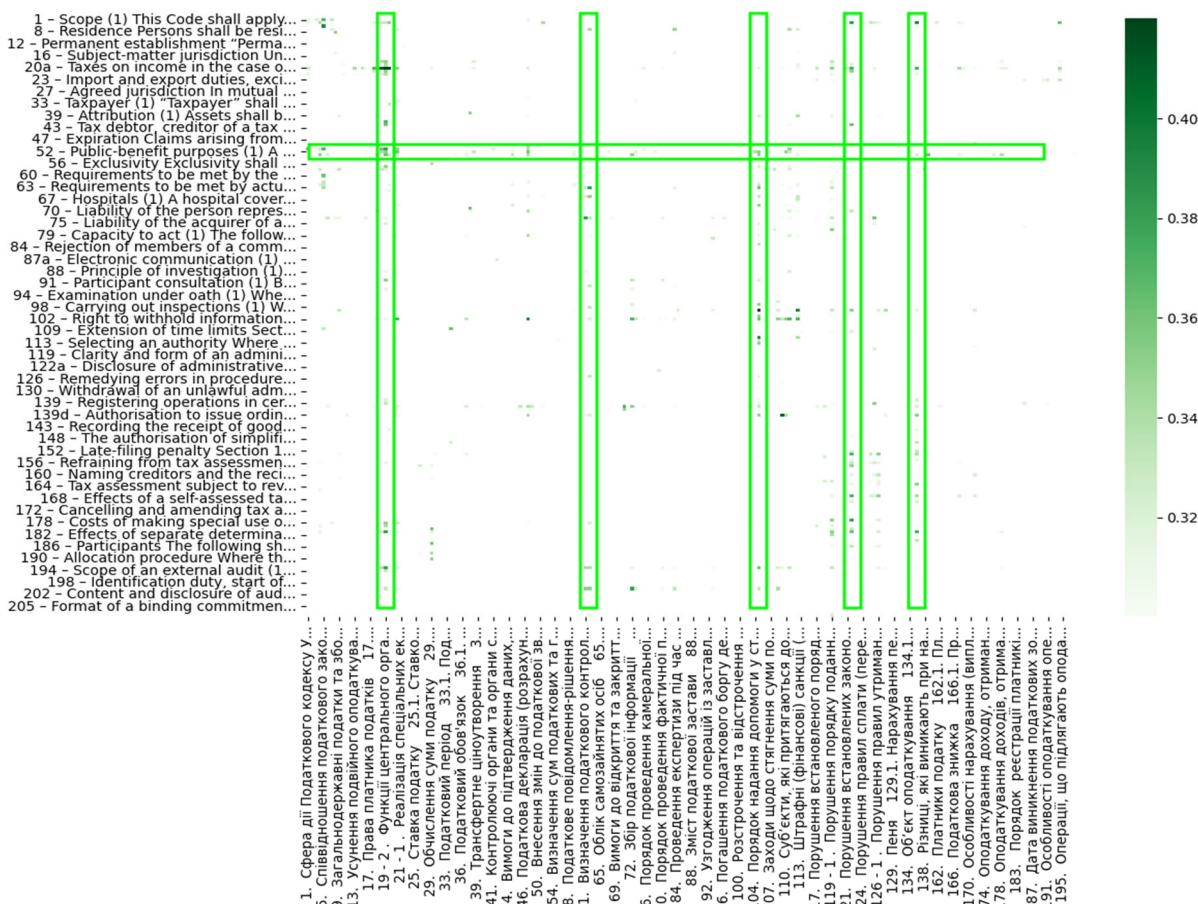


Fig. 3. Cosine similarity on the subset of data after removing references to changes in the Ukrainian and German fiscal codes



Fig. 4. A clear line pattern after cleaning up titles and references to other articles

2. Results. Finally, after applying the translation to English, increasing the correlation threshold $\text{significant} = \text{df}[\text{corr}] \geq .70$

and repeating the analysis and visualization, we've seen the picture without line patterns as shown on Fig.5. To manually assess the results, we're dropping rows and columns that have correlation below our threshold

`table = significant.dropna(how='all').dropna(axis='columns', how='all').fillna("")`

and as shown in Table 1. we can see a clear relation of articles, for example, "Local taxes include ..." of the Ukrainian fiscal code correspond very well to the German "Non-personal taxes Local jurisdiction ..." article.

Table 1

	6 When setting the tax, the followin...	9 Local taxes include:\n property t...	11 The Verkhovna Rada of Ukraine esta...	27 The unit of measurement of the tax...	31 Absolute (specific) is the tax rat...	33 Tax relief - provided by the tax a...	40 The duty of the taxpayer to calcul...	42 Fulfillment of the tax obligation ...
1 – Taxes, ancillary tax payments\n "T...					0.72			
16 – Taxes on income in the case of c...								
18 – Non-personal taxes\n Local jurisdi...		0.7	0.7			0.7		
28 – Taxpayer\n "Taxpayer" shall mean a...							0.73	
30 – Claims arising from the tax debto...							0.71	
36 – Tax debtor, creditor of a tax reb...								

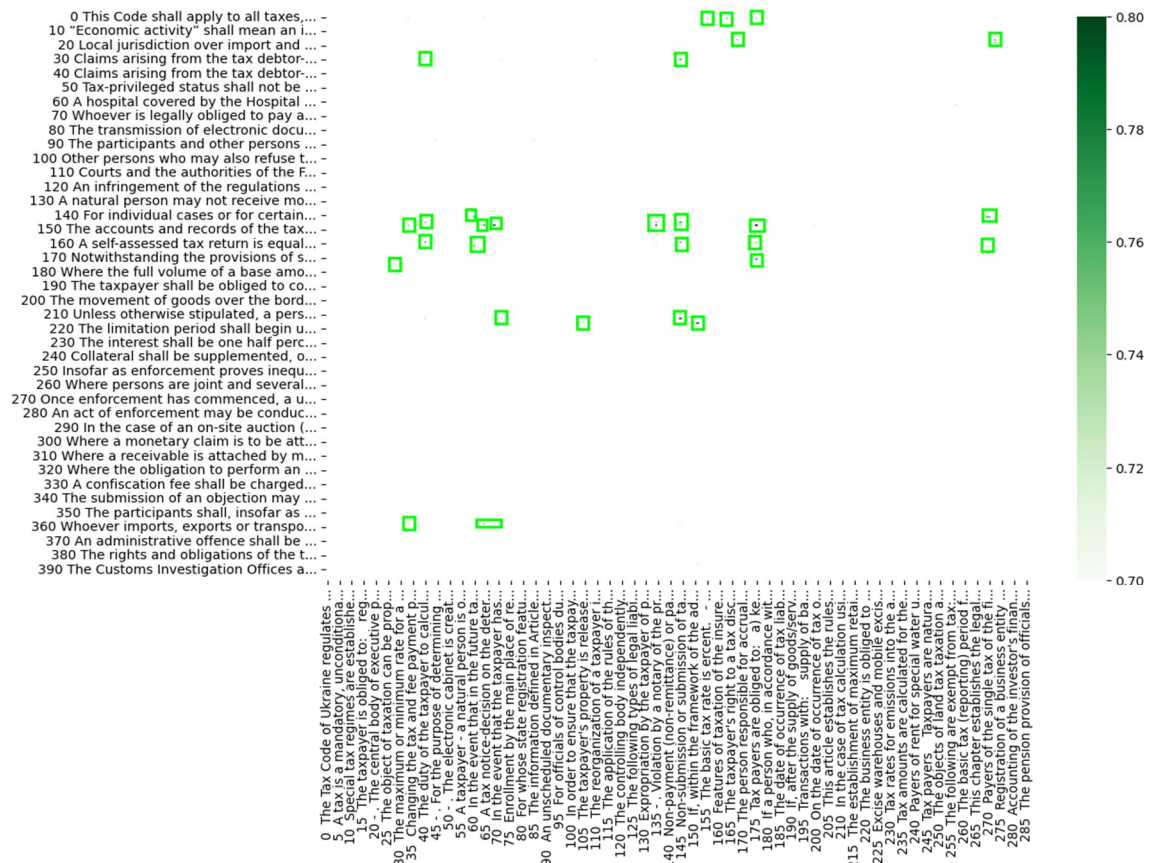


Fig. 5. No line patterns, only individual dots after translating both fiscal codes to English

Collecting all the steps we can summarize our method as following:

1. Analyze document formats and structures. Select appropriate method to convert it to data frames – for example, html and pdf processing
2. Decide what is considered a unit in every document – single article of the tax code in our case
3. Choose embeddings and metric, test it on samples from both documents – for example, Universal Sentence Embeddings with cosine metric
4. Analyze and visualize correlations, cleanup, iterate. Our research showed some unexpected correlations, which we were able to explain and eliminate with the further cleanup of our data
5. Finally, prepare the correlation table which shows us the similarities between the documents

While the proposed approach allows us to find the relation between the most similar articles of the two documents in the straightforward way, it can also be extended to highlight the most distant parts of the documents suggesting places where two fiscal codes differ a lot.

Discussion and conclusions

During our research, we were convinced once again of the importance of the well cleaned and prepared data (Lakhno et al., 2022; Savchenko et al., 2024; Boyko, 2021; Black, Kueper, & Williamson, 2023) – we had to parse different formats of data, replace irrelevant references, article numbers and other non-textual information. While we share our approach to data cleanup and normalization in the Data Cleanup section of this article, it might become less relevant for other types of documents and might limit the ability to replicate our results without the adaptation of the cleanup techniques.

We have also met the limitations of the available Multilingual Universal Sentence Embeddings model. While the model doesn't officially list Ukrainian support, our initial tests gave us a lot of hope that it will work without the translation. At the same time, the usage of translation opens access to a big variety of models and embeddings to evaluate similarity.

Ability to self-test and visualize an intermediate result was an important part of the research, because it allowed to iterate ideas, ensure the data cleanup was performed well, and the limitations of the selected model were addressed by introduction of the translation step.

As a result, we have offered a method of legislative documents analysis applied to the Ukrainian Fiscal Code. In the research we analyzed similarities with the German fiscal code, offering a detailed step-by-step replication guide to apply this method to the fiscal codes of other EU member states or other countries.

The analysis allows us to identify the positive aspects and shortcomings of the Fiscal Code of Ukraine and to develop expert recommendations for bringing the Code in line with European standards.

The continuation of the research initiated in this paper is the development of algorithms for standardizing or bringing to a common denominator the regulatory framework of the European Union, which is quite relevant given the processes of European integration and globalization that are currently being actively developed and implemented.

Authors' contribution: Olena Skliarenko – statement of the task, conceptualization, methodology, intelligence analysis, writing – viewing and editing; Anatolii Pashko – statement of the task, conceptualization, methodology, intelligence analysis, writing – viewing and editing; Leonid Lytvynenko – statement of the task, conceptualization, methodology, software, spelling – original draft; Yanina Kolodinska – data validation, formal analysis, spelling – original draft.

References

- Basystiuk O., & Melnykova N. (2022). Multimodal speech recognition based on audio and text data. *Bulletin of Khmelnytsky National University: Technical Sciences*, 313(5), 22–25 [in Ukrainian]. [Басистюк, О., & Мельникова, Н. (2022). Мультимодальне розпізнавання мовлення на основі звукових і текстових даних]. <https://www.doi.org/10.31891/2307-5732-2022-313-5-22-25>
- Basystiuk O., Shakhovska N., Bilynska V., Syvokon O., Shamuratov O., & Kuchkovskiy V. (2021) The Developing of the System for Automatic Audio to Text Conversion. *IT&AS'2021: Symposium on Information Technologies & Applied Sciences*, March 5–6, 2021, (pp. 1–8). Bratislava, Slovak Republic. <https://ceur-ws.org/Vol-2824/paper1.pdf>
- Black, J. E., Kueper, J. K., Williamson, T. S. (2023). An introduction to machine learning for classification and prediction. *Family Practice*, 40(1), 200–204.
- Boyko, N. (2021). Research into machine learning algorithms for the construction of multimodal data classification problems. *Computational Problems of Electrical Engineering*, 11, 1–11.
- Boyko, N., Mochurad, L., Parpan, U., & Basystiuk, O. (2019). Usage of Machine-based Translation Methods for Analyzing Open Data in Legal Cases. In *Proc. of the Intl Workshop on Cyber Hygiene (CybHyg-2019) co-located with 1st International Conference on Cyber Hygiene and Conflict Management in Global Information Networks* (p. 328–338). Kyiv, Ukraine. <https://ceur-ws.org/Vol-2654/paper26.pdf>
- Golomoziy, V., Mishura, Yu., Izarova, I., & Ianevych, T. (2023). Processing Big Data of Court Decisions. *Baltic Journal of Modern Computing*, 11. <https://doi.org/10.22364/bjmc.2023.11.4.04>
- Ianevych, T., Golomoziy, V., Mishura, Yu., & Izarova, I. (2023). Comparison of 2D convolutions and dense neural networks for natural language processing models with multi-sentence input. *Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics and Mathematics*, 20–29. <https://doi.org/10.17721/1812-5409.2023/2.3>
- Lakhno, V. A., Kasatkin, D. Y., Skliarenko, O. V., & Kolodinska, Y. O. (2022). Modeling and Optimization of Discrete Evolutionary Systems of Information Security Management in a Random Environment. *Machine Learning and Autonomous Systems. Smart Innovation, Systems and Technologies*, 269, 9–22. Springer, Singapore. https://doi.org/10.1007/978-981-16-7996-4_2
- Peng, Li, Xi, Rao, Jennifer, Blase et al. (2021). *CleanML: A Study for Evaluating the Impact of Data Cleaning on ML Classification Tasks*. arXiv. <https://doi.org/10.48550/arXiv.1904.09483>
- Savchenko, Y., Yarovy, R., Kolodinska, Y., Levchenko, S. (2024). The role of artificial intelligence and machine learning under martial law. *Measuring and computing devices in technological processes*, 2, 213–216. [in Ukrainian] <https://doi.org/10.31891/2219-9365-2024-78-25>
- Shakhovska, N., Basystiuk, O., Shakhovska, K. (2019). Development of the speech-to-text chatbot interface based on Google API. *CEUR Workshop Proceedings*, vol. 2386, 212–221.
- Yang, Y., Cer, D., Ahmad, A. et al. (2019). *Multilingual Universal Sentence Encoder for Semantic Retrieval*. arXiv. <https://doi.org/10.48550/arXiv.1907.04307>

Отримано редакцією журналу / Received: 05.09.24

Прорецензовано / Revised: 27.11.24

Схвалено до друку / Accepted: 07.03.25

Олена СКЛЯРЕНКО¹, канд. фіз.-мат. наук, доц.

ORCID ID: 0000-0001-6555-1223

e-mail: olena.skliarenko@e-u.edu.ua

Анатолій ПАШКО², д-р фіз.-мат. наук, проф.

ORCID ID: 0000-0001-6944-8477

e-mail: anatoliipashko@knu.ua

Леонід ЛИТВИНЕНКО¹, канд. техн. наук

ORCID ID: 0000-0002-0828-383X

e-mail: leonid.lytvynenko@e-u.edu.ua

Яніна КОЛОДІНСКА¹, ст. викл.

ORCID ID: 0000-0002-3330-7565

e-mail: yanina.kolodinska@e-u.edu.ua

¹Приватний вищий навчальний заклад "Європейський університет", Київ, Україна

²Київський національний університет імені Тараса Шевченка, Київ, Україна

МОДЕЛІ ГЛИБОКОГО НАВЧАННЯ ДЛЯ РОЗВ'ЯЗАННЯ ЗАДАЧІ ПОРІВНЯННЯ НЕСТРУКТУРОВАНОЇ ТЕКСТОВОЇ ІНФОРМАЦІЇ

Присвячено застосуванню сучасних методів глибокого навчання, таких як універсальні багатомовні векторні представлення речень, для аналізу та порівняння схожості змісту неструктурованих текстових документів. Розроблено алгоритм визначення подібності багатомовних текстів за допомогою глибоких нейронних мереж. Авторами запропоновано безсуперечливий метод для швидкого аналізу та порівняння складних документів подібної тематики, написаних у різних країнах та різними мовами. Представлена методологія передбачає попереднє оброблення документів, їхнє переформатування для порівняння та застосування методів глибокого навчання для виявлення текстової схожості. Алгоритм відпрацьовано на прикладі здійснення порівняльного аналізу схожості за змістом текстів Податкового кодексу України з німецьким фіскальним кодексом. Український фіскальний код, поданий у форматі HTML, було оброблено за допомогою бібліотеки аналізу HTML документів і регулярних виразів, тоді як німецький фіскальний код у форматі PDF потребував спеціального аналізу неструктурованого змісту. У процесі дослідження для підвищення точності й подолання виявлених обмежень універсальних багатомовних векторних представлень український документ було перекладено англійською мовою за допомогою Google Translate API. Розв'язання поставленої задачі потребує глибокого попереднього оброблення даних, що було здійснено у пропонованому дослідженні. Наведено також покроковий алгоритм представленої методології, що приводить до формування кореляційної таблиці, за якою визначається схожість між подібними різномовними текстовими документами. Крім того, автори доповнюють кожен крок своїми міркуваннями про те, як інтерпретувати проміжні результати та розуміти, чи варто коригувати метод відповідно до потреб або проблем конкретних документів, як-от застосування машинного перекладу чи розширеного попереднього оброблення.

Авторами запропоновано алгоритм для застосування зазначеного методу до фіскальних кодексів країн – членів Європейського Союзу або інших країн. У дослідженні виявлені й описані обмеження цього підходу, а також продемонстровано проведення перевірки проміжних результатів та здійснено доповнення методу додатковими можливостями задля підвищення його надійності. Крім цього, надано рекомендації щодо можливостей застосування запропонованої методології, зокрема розроблення алгоритмів для стандартизації певних документів країн Європейського Союзу.

Ключові слова: глибоке навчання, універсальні векторні представлення речень, фінансові документи, фіскальний код, регулярні вирази.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

УДК 004.023

DOI: <https://doi.org/10.17721/1812-5409.2025/1.22>

Олександр ФРАНКІВ, асп.
ORCID ID: 0009-0005-5264-358X
e-mail: o.frankiv@ukma.edu.ua

Національний університет "Києво-Могилянська академія", Київ, Україна

МАШИННЕ НАВЧАННЯ В ПОКРАЩЕННІ ВІЗУАЛІЗАЦІЇ ПРОСТОРОВОЇ МОДЕЛІ АРХІТЕКТУРИ ПРОГРАМИ

Візуальна естетичність та обчислювальна ефективність є однаково важливими аспектами у контексті створення якісного представлення даних для подальшого аналізу. Гармонійне поєднання цих характеристик забезпечує не лише зручність сприйняття, але й оптимізує процеси оброблення даних, що є критично важливим у сучасному програмному забезпеченні.

У пропонованій статті розкрито новий комбінований підхід до розміщення графів у просторі. Особливу увагу приділено просторовим моделям архітектури програмного забезпечення, що є ключовим інструментом для візуалізації складних взаємозв'язків між компонентами. Застосування графових нейронних мереж як спеціалізованої евристики – центральний елемент зазначеного підходу. Завдяки використанню методів машинного навчання, запропонований підхід дає змогу поліпшити результати візуалізації та підвищити обчислювальну ефективність.

Використання графової нейронної мережі забезпечує адаптивність і здатність моделі враховувати специфічні особливості графу. Це, у комбінації із силовим алгоритмом, сприяє збереженню високого рівня естетичності візуального представлення без значного збільшення ресурсних витрат. Отже, новий метод пропонує практичне рішення для ефективного поєднання візуальної естетики й обчислювальної ефективності, що є важливим кроком у вдосконаленні аналізу просторових моделей архітектури програмного забезпечення.

Ключові слова: *нейронна мережа, машинне навчання, граф, графова нейронна мережа, графовий згортковий оператор (GCNConv), архітектура програмного забезпечення, автоматична візуалізація архітектури програмного забезпечення, часова складність алгоритму.*

Класифікація відповідно до AMS 2020: QA76.9.A43.

Вступ

Забезпечення якості продукту є важливим аспектом виробництва програмного забезпечення. Нині існує велика кількість інструментів, що дають змогу оптимізувати певні частини цього процесу. Зважаючи на стрімкий розвиток галузі, створення нових і розвиток наявних інструментів залишається актуальним.

Серед великої кількості рішень, що дають змогу автоматизувати аналіз якості програмного коду, важливим залишається напрям аналізу саме архітектурних рішень. Це, зокрема, передбачає створення інструментів, що не вимагають значних додаткових затрат для створення спеціальної документації. Раніше, у роботах (Франків, 2023; Франків, & Глибовець, 2024) описано розроблений нами програмний комплекс для автоматичної візуалізації архітектури програмного модуля, спираючись лише на вихідний код. Аналізатор створює візуальну модель архітектури програми у вигляді графа, вузлами якого позначено структурні елементи програми, а ребрами – зв'язки між ними. За допомогою візуалізатора комплекс створює візуально естетичне розміщення цього графа у тривимірному просторі. За допомогою спеціального iOS-застосунку цей граф можна відобразити в доповненій реальності, що забезпечує природне сприйняття структурованої інформації.

Проблема автоматичного створення прийнятного для сприйняття розміщення графа в просторі є складною з огляду на те, що запропонований метод має працювати для графів різних форм і розмірів, забезпечувати відповідність загальноприйнятим критеріям естетичності й обраховуватись із прийнятною швидкістю. Підвищення швидкодії візуалізації є особливо затребуваним у контексті створення розміщень на пристроях з обмеженими ресурсами, зокрема й мобільних.

У пропонованій статті розглянуто спосіб застосування алгоритмів машинного навчання для створення спеціалізованої евристики, що дає змогу поліпшити швидкодію наявного рішення без негативного впливу на рівень естетичності розміщення графа в просторі. Описано підхід до створення цієї евристики в загальному випадку, що відкриває можливість створення аналогів для різних технологій аналізованих проєктів. Досліджено доцільність використання спеціалізованих евристик для графів різних розмірів.

1. Розміщення графів у просторі

Натепер задача відображення графів у просторі досліджена добре. Існує велика кількість способів представлення, що відповідають різним запитам користувача. Варто зауважити, що залежно від кількості, а головне – природи даних, поданих у формі графа, різні види візуалізації мають свої переваги й недоліки.

Оскільки ми розглядаємо задачу в контексті створення подання для аналізу архітектури програми, важливо визначити основні вимоги до застосованого підходу. По-перше, ваги ребер відіграють ключову роль у відображенні міри взаємної залежності компонентів архітектури, а отже, мають обов'язково враховуватись. По-друге, зважаючи на розмір програм і, як наслідок, потенційно великий розмір графів, бажано, щоб обраний підхід мав невелику часову складність. По-третє, критичною є висока естетичність подання, що допоможе не лише відобразити кластери вузлів, але й зберегти відносні відстані між ними.

З огляду на конкретні потреби, серед класичних методів візуалізації очевидним є вибір на користь силового алгоритму, що дає змогу досягнути порівняно кращого сприйняття інформації користувачем (Pohl, Schmitt, & Diehl, 2009). Важливою його перевагою є можливість працювати зі зваженими графами.

За силовим алгоритмом (Fruchterman, & Reingold, 1991) граф варто розглядати як фізичну модель, у якій вузлам графа відповідають позитивні точкові заряди, що відштовхуються один від одного за законом Кулона, а ребрам –

пружини, що їх пов'язують і створюють силу протидії за законом Гука. У класичному варіанті на початковому етапі заряди розміщуються в одній точці, або ж частіше випадковим чином, а далі в циклі відбувається постійний обрахунок нового положення системи відповідно до застосування всіх наявних у системі сил.

Серед недоліків силового алгоритму варто згадати високу часову складність обчислень $O(N^2)$ для однієї ітерації, що зумовлена високою часовою складністю обходу графа. З іншого боку, через використання випадкових евристик кількість ітерацій зовнішнього циклу не є фіксованою, що, зі свого боку, призводить до нестабільності та проблеми локальних мінімумів.

Існують оптимальніші аналоги класичного алгоритму. Для обрахування сил взаємодії для великої кількості тіл, наприклад, пропонують розглядати простір як дерево (Barnes, & Hut, 1986). Двовимірний простір ділять на чотири частини (тривимірний на вісім) рекурсивно до моменту, коли в кожному підпросторі не залишиться лише один вузол. Тоді обхід графа зводиться до обходу дерева і часова складність однієї ітерації становить $O(N * \log N)$. Однак, оскільки зміщення кожного вузла в алгоритмі можливе лише в межах його власного підпростору, то з'являється висока залежність від першопочаткового розміщення вузлів. Це може негативно впливати на естетичність подання (див. рис. 1) і робить цей метод малопридатним для застосування в контексті нашої задачі.

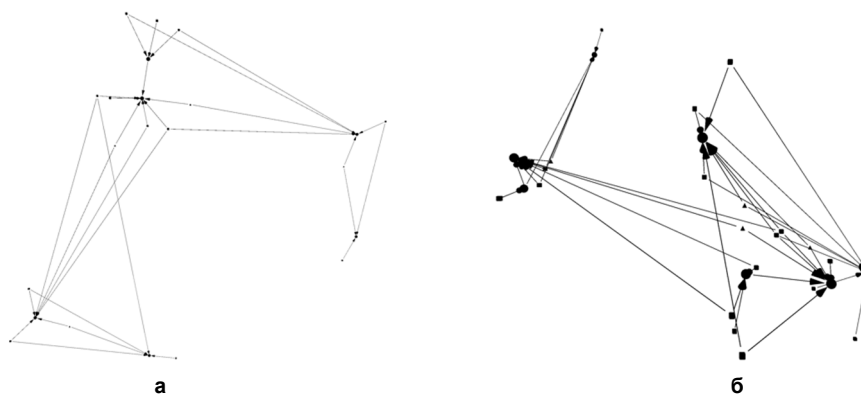


Рис. 1. Результат розміщення за допомогою класичного силового алгоритму (а) та алгоритму оптимізованого обрахування взаємодій (б)

У дослідженнях (Walshaw, 2001; Didimo, Liotta, & Romeo, 2011) автори пропонують зваженіші підходи до початкового розміщення графа в просторі, однак ці методи мають обмеження і не є універсальними, оскільки складно передбачити властивості графів, побудованих з реальних даних.

Абсолютно іншим підходом до задачі візуалізації графів є використання алгоритмів машинного навчання. Головні виклики в цій сфері – забезпечення високої естетичності розміщення, ефективність й універсальність (Wang et al., 2024).

Серед більш універсальних методів можна вирізнити node2vec (Grover, & Leskovec, 2016), force2vec (Rahman, Sujon, & Azad, 2020) та DeepDrawing (Wang et al., 2020). Кожен із наведених методів дає змогу отримати естетичні представлення графів, що все ж не будуть абсолютно відповідати результату роботи силового алгоритму і, зокрема, вимогам нашої задачі.

Складність застосування node2vec полягає в тому, що в його основі лежить генерація випадкових проходів по графу. Для отримання високої якості візуалізації необхідне тонке налаштування параметрів ймовірності повернення та ймовірності розширення, що складно зробити для графів, які потенційно можуть бути абсолютно різними за структурою.

Застосування force2vec до нашої задачі не вирішує проблеми залежності від початкового розміщення графа. Окрім того, класичний метод не працює з вагами ребер.

DeepDrawing передбачає роботу із графами лише фіксованого розміру, що робить цей метод не достатньо універсальним для наших потреб.

З огляду на описані вище вимоги до візуалізації й аналіз наявних, зокрема комбінованих (Huang et al., 2024), підходів було ухвалено рішення запропонувати універсальний метод, що передбачає неявне задання спеціалізованої евристики на основі аналізу реальних даних та її застосування для початкового розміщення в контексті класичного силового алгоритму. Це дасть змогу забезпечити високу якість візуалізації, а також підвищити стабільність алгоритму, на відміну від використання випадкових евристик, та потенційно зменшити кількість "дорогих" ітерацій зовнішнього циклу.

2. Використання машинного навчання для створення спеціалізованої евристики

Створення спеціалізованої евристики для розміщення графа у просторі в явному вигляді є складною задачею. З одного боку, детальний аналіз графа з метою визначення особливостей його топології передбачає обхід графа, що, власне, і є слабким місцем класичного силового алгоритму. З іншого боку, з огляду на те, що реальна робота візуалізатора передбачає розміщення графів довільної конфігурації (зумовлено довільною природою архітектури аналізованих програм), зробити явно задану евристику універсальною досить складно.

Натепер використання машинного навчання й нейронних мереж є перспективним способом неявного програмування знань про домен, отриманих унаслідок аналізу реальних даних. Ми висунули припущення про те, що за умови використання спеціалізованого апаратного забезпечення завдяки використанню нейронної мережі, побудованої за допомогою аналізу природних даних, можна значно зменшити час виконання силового алгоритму. Це можливо за одночасного виконання двох умов:

- Використання отриманої евристики зменшить кількість необхідних ітерацій обходу графа на деяку вагому кількість ΔI .

• Час, необхідний на обчислення початкового розміщення за допомогою нейронної мережі, буде меншим, ніж потрібний час на виконання ΔI ітерацій обходу графа.

Для перевірки цього припущення було підготовано колекцію даних та запропоновано архітектуру простої нейронної мережі. Ми провели тренування для отримання моделей і порівняння продуктивності виконання силового алгоритму з випадковою евристикою та новою, спеціалізованою.

3. Збір та підготовка даних

Евристика, що відображає певний досвід у розміщенні графа архітектури програми, є чутливою до конкретних технологій, що використовуються під час її побудови. Саме тому автоматизація збору даних є важливою. Вона дає змогу легко отримати нові дані для аналізу в майбутньому, якщо спосіб побудови програм конкретного призначення зміниться з появою нових технологій та підходів.

Окрім того, ми автоматизували збір проєктів зі сховища програмного забезпечення з відкритим вихідним кодом GitHub. Оскільки оригінальна імплементація нашого комплексу дає змогу аналізувати вихідний код програм, написаних мовою Swift, то і в запропонованому дослідженні ми перевіряли наше припущення на них.

Процес побудови графів архітектури ми також автоматизували шляхом створення спеціальної утиліти аналізу файлів проєкту, що автоматично готує необхідні параметри та формує команду аналізатору. Також було вдосконалено наявний інтерфейс аналізатора та візуалізатора з наголосом на тому, щоб внаслідок його роботи можна було в явному вигляді отримати ваги ребер, адже вони безпосередньо впливають на результат розміщення у просторі.

У контексті підготовки візуальних представлень для навчання важливою проблемою є вибір умови зупинки силового алгоритму. До найчастіше використовуваних належить зупинка за умови досягнення максимальної кількості ітерацій, порогу максимального зміщення всіх вузлів, або порогу потенціальної енергії системи (Fruchterman, & Reingold, 1991).

Очевидною перевагою зупинки за досягненням максимальної кількості ітерацій є визначеність часу роботи алгоритму. З іншого боку, такий підхід має два серйозні недоліки. По-перше, стала кількість ітерацій абсолютно не гарантує досягнення якомога більш естетичного положення розміщення, що відповідає стану рівноваги фізичної моделі. По-друге, такий підхід зовсім не є універсальним, адже не враховує особливості розмірів і топології аналізованих графів. До того ж, розроблення функцій отримання очікуваного числа ітерацій для конкретного графа передбачатиме детальний і складний аналіз, або ж вагомий спрощення.

Досягнення порогу зміщення вузлів або енергії системи є універсальнішими підходами, оскільки спираються саме на фізичну модель. Очевидним недоліком обох є невизначеність часу виконання алгоритму. На противагу ці підходи дають змогу отримати природні з позиції фізичної моделі і, як наслідок, естетичніші розміщення. Зупинка за умови досягнення порогу енергії системи дає змогу в загальному випадку досягнути глобального мінімуму, а отже, найбільш якісного розміщення. Однак обрахунок енергії системи є витратнішим, а сам критерій чутливий до визначеного порогу й потребує тонкого налаштування. Зупинка за умови досягнення порогу зміщення вузлів може відбутися у точці локального мінімуму, однак сама умова обраховується швидше. У цьому підході визначити поріг значно простіше, бо, по суті, потрібно просто визначити мінімальну відстань зміщення вузлів, що сприймається незначною і такою, що нею можна знехтувати.

Ми обрали комбіновану умову зупинки з використанням порогу зміщення вузлів, а також максимальною кількістю ітерацій. Використання додаткового обмеження за кількістю ітерацій обумовлене тим, що аналізовані графи можуть бути зовсім різними за розмірами, а обрахунок часу їхнього розміщення може значно різнитись. До того ж, оскільки метою є створення евристики для початкового розміщення, тенденції поведінки фізичних моделей саме на певній кількості перших ітерацій алгоритму є найбільш універсальними.

Також у загальному випадку граф архітектури програми може бути не суцільним, тому досягнення повного спокою системи може бути неможливим. Щоб вирішити цю проблему, на розміщення графу накладено додаткову умову – воно має вміщуватись у підпросторі фіксованого розміру.

У підсумку процес підготовки даних є модульним і охоплює три етапи: збір проєктів і початковий аналіз структури проєкту; побудову зваженого графа з розміщенням в просторі; генерацію представлення для навчання. Процес повністю автоматизований. Перевикористання розробленого інструменту для проєктів, що застосовують інші технології, не викликає труднощів і, по суті, зводиться до автоматизації початкового аналізу структури проєкту й імплементації аналізатора для конкретної мови й технології.

4. Використання графових нейронних мереж

Складність застосування нейронних мереж для задачі візуалізації довільного зваженого графа визначається тим, що у нашому випадку вузли графа не містять жодних метаданих, корисних для побудови евристики. По суті, єдиною інформацією, що впливає на розміщення в просторі, є наявність ребер між певними вузлами та їхні ваги. Саме тому підходи, що спираються на попередню кластеризацію вузлів за певними ознаками, у цьому випадку не можна застосувати.

Важливим аспектом під час підбору технологій і розроблення архітектури запропонованого рішення був намір зробити його крос-платформним, а саме – забезпечити можливість обчислення на мобільних пристроях. Тоді за умови досягнення достатньої продуктивності новоствореної евристики, візуалізатор можна було б перемістити з комп'ютерної частини програмного комплексу до мобільного додатка-переглядача (див. рис. 2). Це забезпечить гнучкість і можливість створення якісніших взаємодій користувача з візуальним представленням програми, наприклад, побудову представлення певної частини загального графа відповідно до конкретних потреб.

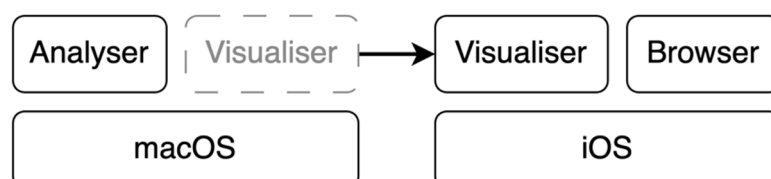


Рис. 2. Схема переміщення візуалізатора в межах програмного комплексу

З огляду на це в основу моделі нейронної мережі було покладено графовий згортковий оператор GCNConv, запропонований у (Kipf, & Welling, 2016). Серед переваг цього оператора варто звернути увагу на можливість роботи з вагами ребер, а також повну сумісність внутрішньої реалізації з CoreML, що, зі свого боку, дає змогу конвертувати модель для роботи в застосунках для macOS й iOS із забезпеченням сумісності зі спеціальним апаратним забезпеченням.

Архітектура моделі є послідовністю кількох шарів GCNConv (див. рис. 3). Перший шар приймає на вхід набір вузлів з їхніми властивостями (V) та набір ребер (E) з їхніми вагами (W). Варто зауважити, що, оскільки інформативних метаданих вузлів графів у нашій задачі не мають, то для забезпечення сумісності як властивість вузла було вирішено використати степінь вузла графу. Розмір даних на виході першого шару позначаємо `hidden_dim` (hidden layer dimension). Ця розмірність загалом може бути довільною, однак достатньою для оброблення властивостей всіх вузлів графа, а отже, не меншою за їхню кількість. У ході аналізу згенерованих даних було встановлено, що кількість вузлів для кожного з графів не перевищувала 2000, тому значення `hidden_dim` було встановлено відповідно.

Останній шар приймає на вхід дані розмірності `hidden_dim`, а повертає набір тривимірних векторів (X), що позначає набір точок у тривимірному просторі – координати кожного вузла у відповідному розміщенні графа.

Між першим й останнім шаром може бути довільна кількість допоміжних шарів, які можуть додати гнучкості та точності моделі. Розмірність вхідних і вихідних даних цих шарів однакова – `hidden_dim`. В оригінальному дослідженні (Kipf, & Welling, 2016) автори зазначають, що для їхньої задачі позитивні результати вдалось отримати з використанням дво- та тривірневої моделі. Очевидним є те, що збільшення кількості шарів моделі збільшує її розмір і час обчислення, а отже, несприятливо впливає на можливість портування. З іншого боку, передбачити оптимальну кількість допоміжних шарів з урахуванням різноманітності вхідних даних досить складно. Саме тому в контексті експерименту було підготовано три споріднені моделі: двовірневу (L2) – без додаткових шарів, тривірневу (L3) – з одним прихованим шаром, а також чотирирівневу (L4) – із двома прихованими шарами.

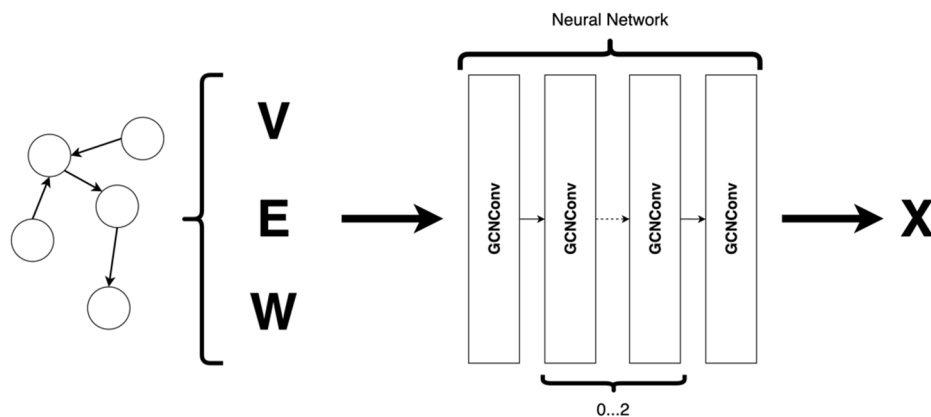


Рис. 3. Принцип роботи спеціалізованої евристики

Як функцію втрат під час навчання було використано середньоквадратичну похибку попарних відстаней між передбаченими й еталонними координатами. Бажане розміщення, тобто набір еталонних координат, обчислювали за допомогою силового алгоритму, описаного вище. З огляду на нестабільність алгоритму, очевидним є те, що такі координати можуть представляти не ідеальне розміщення, однак з точністю до створення евристики для цього ж алгоритму такий підхід є виправданим.

Як функцію активації було використано ReLU, що дало змогу враховувати нелінійні залежності у даних, зберігаючи ефективність обчислень.

Під час навчання мережі значення помилки спадало, однак залишалось високим. Ймовірно, це зумовлено великою різницею між розмірами графів. Через це складно вважати запропонований підхід ефективним для самостійного використання з метою передбачення координат усіх вузлів одразу.

Отримані моделі мереж було конвертовано в CoreML за допомогою спеціального набору утиліт `coremltools`. Унаслідок конвертування розміри готових до використання файлів *.mlpackage становили 50 кБ для L2, 8,1 МБ для L3 та 16,1 МБ для L4.

Отже, ми створили новий візуалізатор, який приймає на вхід модель мережі та зважений граф, застосовує її для обчислення початкового розміщення вузлів, а потім використовує силовий алгоритм для генерації розміщення в просторі.

5. Перевірка евристики й інтерпретація результатів

Для перевірки продуктивності роботи візуалізатора зі спеціалізованою евристикою створено спеціальну програму, що дає змогу фіксувати кількість ітерацій силового алгоритму під час побудови розміщення, а також час, витрачений на його обчислення.

Експеримент із заміру продуктивності було проведено на комп'ютері Apple MacBook M1 Pro з 10 ядрами CPU, 14 ядрами GPU та 16 ядрами Neural Engine. Графи було отримано автоматично, що допомогло в дослідженні особливостей роботи запропонованого рішення для задач різних розмірів і конфігурацій.

Для кожного з візуалізаторів було проведено по п'ять побудов кожного зразка. Візуалізатори сформовано на основі класичного силового алгоритму, однак кожен використовує власну евристику. Default застосовує випадкову евристику, L2, L3 та L4 використовують нашу спеціалізовану евристику, імплементовану у формі нейронної мережі з відповідною кількістю рівнів.

У загальній таблиці (дод. 1) наведено частину результатів проведення побудов розміщень графів архітектури для різних програм. Для кожного експерименту надано інформацію про кількість вузлів $N(V)$ і ребер $N(E)$ у графі, а також мінімальне, максимальне та середнє значення для кількості ітерацій I та часу t в секундах до зупинки алгоритму.

Важливою перевагою використання запропонованої евристики стала стабілізація кількості ітерацій, причому незалежно від глибини конкретно взятої нейронної мережі. Це, відповідно, робить виконання алгоритму передбачуваним і дає змогу краще оцінити очікуваний час завершення.

У контексті підвищення ефективності алгоритму за допомогою евристики вдалось дослідити певні особливості залежності I та t від комбінації розмірів графа $N(V)$, $N(E)$ та конкретної моделі нейронної мережі. Зауважимо, що тут і далі розподіл графів на малі, середні та великі є умовним і стосується, радше, зручності ілюстрації конкретних результатів (табл. 1).

Таблиця 1

Результати застосування евристики для малих графів
з $N(V) < 50, N(E) < 50$

	N(V)	N(E)	min(I)	max(I)	avg(I)	min(t)	max(t)	avg(t)
Sample 6								
Default	24	21	8550	10000	9516,67	0,20	0,23	0,22
L2	24	21	5860	5860	5860,00	0,38	0,40	0,39
L3	24	21	7675	7675	7675,00	0,50	0,59	0,53
L4	24	21	10000	10000	10000,00	0,63	0,64	0,64
Sample 13								
Default	15	19	950	1373	1141,67	0,01	0,06	0,03
L2	15	19	875	875	875,00	0,25	0,25	0,25
L3	15	19	2282	2282	2282,00	0,34	0,38	0,36
L4	15	19	1117	1117	1117,00	0,36	0,38	0,37

У табл. 1 добре видно, що для малих графів з $N(V) < 50, N(E) < 50$ добре працює евристика на базі L2. Однак попри те, що значення I значно менше і в середньому, і навіть для мінімального випадку, значення t виходить значно більшим. Такий результат пояснюємо тим, що час на виконання ΔI ітерацій для невеликого розміру вхідних даних незрівнянно менший, ніж час, необхідний для обчислення евристики. Отже, у такому випадку її використання не є ефективним.

Із табл. 2 випливає, що для графів більшого розміру в контексті значення I ефективні показники мають евристики L2 та L3. Що стосується мінімізації часу виконання, то зі збільшенням розмірів графа час, необхідний для виконання ΔI ітерацій, стає більшим і, як наслідок, застосування евристики демонструє певну ефективність.

Таблиця 2

Результати застосування евристики для середніх графів
з $N(V) \in (50, 100), N(E) \in (50, 100)$

	N(V)	N(E)	min(I)	max(I)	avg(I)	min(t)	max(t)	avg(t)
Sample 1								
Default	91	97	3833	10000	6092,33	0,57	1,49	0,91
L2	91	97	3537	3537	3537,00	0,74	0,75	0,74
L3	91	97	5416	5416	5416,00	1,08	1,10	1,09
L4	91	97	5324	5324	5324,00	1,16	1,18	1,17
Sample 12								
Default	68	78	8834	10000	9611,33	0,85	0,95	0,92
L2	68	78	10000	10000	10000,00	1,16	1,29	1,21
L3	68	78	5735	5735	5735,00	0,81	0,89	0,84
L4	68	78	10000	10000	10000,00	1,28	1,33	1,31

Як видно з табл. 3 для графів більшого розміру використання евристики L3 забезпечує значне спадання значень I та t . Зважаючи на ці дані, можна стверджувати, що застосування спеціалізованої евристики для графів, починаючи з такого розміру, є виправданим і скорочує середній час, необхідний для виконання алгоритму, приблизно до 50–60 % у середньому і приблизно до 40 %, якщо порівнювати максимальні значення.

Таблиця 3

Результати застосування евристики для середніх графів
з $N(V) \in (100, 300), N(E) \in (100, 300)$

	N(V)	N(E)	min(I)	max(I)	avg(I)	min(t)	max(t)	avg(t)
Sample 11								
Default	152	131	4802	10000	7683,33	1,92	3,98	3,07
L2	152	131	10000	10000	10000,00	4,19	4,21	4,20
L3	152	131	3728	3728	3728,00	1,77	1,78	1,78
L4	152	131	9311	9311	9311,00	4,10	4,11	4,11
Sample 42								
Default	129	138	4668	10000	7008,33	1,34	2,87	2,01
L2	129	138	4691	4691	4691,00	1,56	1,58	1,57
L3	129	138	1901	1901	1901,00	0,82	0,84	0,83
L4	129	138	3389	3389	3389,00	1,29	1,35	1,32

Варто зауважити, що для деяких графів середніх розмірів, як видно з табл. 4, евристика на основі L4 працює краще в контексті мінімізації I , однак через розмір моделі використання цієї мережі вимагає більше часу, а тому використання L4 не зменшує t порівняно з L3.

Таблиця 4

Результати застосування евристики для середніх графів
з $N(V) \in (50, 100), N(E) \in (100, 300)$

	N(V)	N(E)	min(I)	max(I)	avg(I)	min(t)	max(t)	avg(t)
Sample 31								
Default	89	101	7069	10000	9023,00	1,14	1,49	1,37
L2	89	101	7870	7870	7870,00	1,35	1,39	1,37
L3	89	101	6406	6406	6406,00	1,20	1,22	1,21
L4	89	101	5972	5972	5972,00	1,22	1,24	1,23
Sample 5								
Default	86	103	3754	9341	6103,67	0,51	1,28	0,83
L2	86	103	8285	8285	8285,00	1,33	1,35	1,34
L3	86	103	5500	5500	5500,00	1,02	1,04	1,03
L4	86	103	5278	5278	5278,00	1,07	1,10	1,09

Цікавим є також те, що хоча використання евристики на основі L3 не завжди гарантує прискорення виконання, її все ж варто розглядати як інструмент стабілізації, адже різниці Δt та ΔI для максимальних значень можуть бути відчутними.

З отриманих результатів також можна зробити припущення, що для графів більшого розміру ефективнішими можуть бути евристики, побудовані на основі мереж з більшою кількістю рівнів. Приклад 35 (див. табл. 5) ілюструє ефективність застосування L4 для графів з кількістю вузлів та ребер понад 300. Варто зауважити, що для більших за розміром графів обмеження кількості ітерацій може виявитись занадто сильним і негативно впливати на результати замірів, адже впорядкування більшої за розміром системи природно вимагає більшої кількості ітерацій. З огляду на це детальніше дослідження поведінки евристики для великих графів варто проводити у більшому часовому вікні.

Таблиця 5

Результати застосування евристики для великих графів
з $N(V) > 300, N(E) > 300$

	N(V)	N(E)	min(I)	max(I)	avg(I)	min(t)	max(t)	avg(t)
Sample 35								
Default	313	370	5082	10000	6855,67	8,33	16,38	11,24
L2	313	370	10000	10000	10000,00	16,57	16,83	16,68
L3	313	370	10000	10000	10000,00	16,62	16,69	16,66
L4	313	370	6759	6759	6759,00	11,51	11,68	11,60

Загалом маємо зазначити, що приблизно для третини тестових зразків досягти порогу зміщення вузлів у межах 10 000 ітерацій не вдалось. Однією з причин є, власне, великий розмір вхідних даних. Іншою причиною можуть бути особливості будови графа, як-от розрідженість чи наявність ізольованих частин.

Спираючись на отримані результати, можна зробити висновок, що для достатньо великих графів є сенс застосовувати запропонований підхід.

Дискусія і висновки

У цій статті запропоновано спосіб використання машинного навчання, а саме нейронних мереж, для створення спеціалізованих евристик для силового алгоритму, що застосовують для розміщення графових моделей архітектури програми в просторі.

Збирання реальних даних і процес створення евристик повністю автоматизовано так, що з невеликими адаптаціями його можна з легкістю використати для довільних технологій та мов програмування.

Для мови програмування Swift для iOS/macOS застосунків було створено спеціальний візуалізатор, що використовує запропоновані евристики. На основі проведених досліджень наочно продемонстровано ефективність застосування цього підходу для достатньо великих графів, що стабілізує процес роботи алгоритму, а також може значно зменшувати кількість часу обчислень.

Описано доцільність використання нейронних мереж з різною кількістю рівнів залежно від розмірів графу. Детальніше дослідження цієї проблеми залишається предметом майбутньої роботи.

Список використаних джерел

- Франків, О. О. (2023). Використання доповненої реальності для візуалізації архітектур програмних модулів. *Наукові записки НаУКМА*, 5, 26–30. <https://doi.org/10.18523/2617-3808.2022.5.26-30>
- Франків, О. О., & Глибовець, М. М. (2024). Автоматизована візуалізація компонентів архітектури програми для мови Swift. *Кібернетика та системний аналіз*, 60(6), 190–198. <https://doi.org/10.1007/s10559-024-00737-9>
- Barnes, J., & Hut, P. (1986). A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324(6096), 446–449. <https://doi.org/10.1038/324446a0>
- Didimo, W., Liotta, G., & Romeo, S. A. (2011). Topology-driven force-directed algorithms. In U. Brandes, & S. Cornelsen (Eds.), *Graph Drawing. GD 2010. Lecture Notes in Computer Science*, Vol. 6502 (pp. 165–176). Springer. https://doi.org/10.1007/978-3-642-18469-7_15
- Fruchterman, T. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software – Practice & Experience*, 21, 1129–1164. <https://doi.org/10.1002/spe.4380211102>
- Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 855–864). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939754>
- Huang, R., Gai, H., Jing, R., & Wan, F. (2024). Word spectral visualization based on Fruchterman-Reingold optimized ForceAtlas2. *Journal of Electrical Systems*, 20(2), 7–15.
- Kipf, T. N., & Welling, M. (2016). *Semi-supervised classification with graph convolutional networks*. arXiv. <https://doi.org/10.48550/arXiv.1609.02907>

- Pohl, M., Schmitt, M., & Diehl, S. (2009). Comparing the readability of graph layouts using eye-tracking and task-oriented analysis. In *CompAesth 09: Workshop on Computational Aesthetics* (pp. 49–56). <https://doi.org/10.2312/COMPAESTH/COMPAESTH09/049-056>
- Rahman, M. K., Sujon, M. H., & Azad, A. (2020, November). Force2Vec: Parallel force-directed graph embedding. In *2020 IEEE International Conference on Data Mining* (pp. 442–451). IEEE. <https://doi.org/10.1109/ICDM50108.2020.00056>
- Walshaw, C. (2001). A multilevel algorithm for force-directed graph drawing. In J. Marks (Ed.), *Graph Drawing. GD 2000. Lecture Notes in Computer Science, Vol. 1984* (pp. 171–182). Springer. https://doi.org/10.1007/3-540-44541-2_17
- Wang, X., Yen, K., Hu, Y., & Shen, H. W. (2024). SmartGD: A GAN-based graph drawing framework for diverse aesthetic goals. *IEEE Transactions on Visualization and Computer Graphics*, 30, 5666–5678 <https://doi.ieeecomputersociety.org/10.1109/TVCG.2023.3306356>
- Wang, Y., Jin, Z., Wang, Q., Cui, W., Ma, T., & Qu, H. (2020). DeepDrawing: A deep learning approach to graph drawing. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 676–686. <https://doi.ieeecomputersociety.org/10.1109/TVCG.2019.2934798>

References

- Barnes, J., & Hut, P. (1986). A hierarchical O (N log N) force-calculation algorithm. *Nature*, 324(6096), 446–449. <https://doi.org/10.1038/324446a0>
- Didimo, W., Liotta, G., & Romeo, S. A. (2011). Topology-driven force-directed algorithms. In U. Brandes, & S. Cornelsen (Eds.), *Graph Drawing. GD 2010. Lecture Notes in Computer Science, Vol. 6502*. (pp. 165–176). Springer. https://doi.org/10.1007/978-3-642-18469-7_15
- Frankiv, O. O. (2023). Using augmented reality for visualizing architectures of software modules. *NaUKMA Research Papers. Computer Science*, 5, 26–30 [in Ukrainian]. <https://doi.org/10.18523/2617-3808.2022.5.26-30>
- Frankiv, O. O., & Glybovets, M. M. (2024). Automated visualization of program architecture components for the Swift language. *Cybernetics and Systems Analysis*, 60(6), 190–198 [in Ukrainian]. <https://doi.org/10.1007/s10559-024-00737-9>
- Fruchterman, T. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software – Practice & Experience*, 21, 1129–1164. <https://doi.org/10.1002/spe.4380211102>
- Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 855–864). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939754>
- Huang, R., Gai, H., Jing, R., & Wan, F. (2024). Word spectral visualization based on Fruchterman-Reingold optimized ForceAtlas2. *Journal of Electrical Systems*, 20(2), 7–15.
- Kipf, T. N., & Welling, M. (2016). *Semi-supervised classification with graph convolutional networks*. arXiv. <https://doi.org/10.48550/arXiv.1609.02907>
- Pohl, M., Schmitt, M., & Diehl, S. (2009). Comparing the readability of graph layouts using eye-tracking and task-oriented analysis. In *CompAesth 09: Workshop on Computational Aesthetics* (pp. 49–56). <https://doi.org/10.2312/COMPAESTH/COMPAESTH09/049-056>
- Rahman, M. K., Sujon, M. H., & Azad, A. (2020, November). Force2Vec: Parallel force-directed graph embedding. In *2020 IEEE International Conference on Data Mining* (pp. 442–451). IEEE. <https://doi.org/10.1109/ICDM50108.2020.00056>
- Walshaw, C. (2001). A multilevel algorithm for force-directed graph drawing. In J. Marks (Ed.), *Graph Drawing. GD 2000. Lecture Notes in Computer Science, Vol. 1984* (pp. 171–182). Springer. https://doi.org/10.1007/3-540-44541-2_17
- Wang, X., Yen, K., Hu, Y., & Shen, H. W. (2024). SmartGD: A GAN-based graph drawing framework for diverse aesthetic goals. *IEEE Transactions on Visualization and Computer Graphics*, 30, 5666–5678 <https://doi.ieeecomputersociety.org/10.1109/TVCG.2023.3306356>
- Wang, Y., Jin, Z., Wang, Q., Cui, W., Ma, T., & Qu, H. (2020). DeepDrawing: A deep learning approach to graph drawing. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 676–686. <https://doi.ieeecomputersociety.org/10.1109/TVCG.2019.2934798>

Отримано редакцію журналу / Received: 01.02.25

Прорецензовано / Revised: 10.03.25

Схвалено до друку / Accepted: 12.03.25

Oleksandr FRANKIV, PhD Student

ORCID ID: 0009-0005-5264-358X

e-mail: o.frankiv@ukma.edu.ua

National University of Kyiv-Mohyla Academy, Kyiv, Ukraine

MACHINE LEARNING IN ENHANCING VISUALIZATION OF THE SPATIAL SOFTWARE ARCHITECTURE MODEL

Visual aesthetics and computational efficiency are equally important aspects in the context of creating high-quality data representations for further analysis. A harmonious combination of these characteristics not only enhances the ease of perception but also optimizes data processing workflows, which is critically important in modern software systems.

This paper proposes a novel combined approach for spatial graph placement. Special attention is given to spatial models of software architecture, which serve as key tools for visualizing complex relationships between components. The use of graph neural networks as a specialized heuristic forms the central element of this approach. Leveraging machine learning methods, the proposed solution improves visualization outcomes while enhancing computational efficiency.

The application of a graph neural network ensures adaptability and enables the model to account for the specific features of the graph. In combination with a force-directed algorithm, this allows for maintaining a high level of visual aesthetics without significant increases in resource consumption. Thus, the new method offers a practical solution for effectively combining visual aesthetics with computational efficiency, representing an important step forward in enhancing the analysis of spatial models in software architecture.

Keywords: neural network, machine learning, graph, graph neural network, graph convolutional operator (GCNConv), software architecture, automatic visualization of software architecture, time complexity of an algorithm.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

Додаток 1

Загальні результати ефективності використання спеціалізованої евристики

	N(V)	N(E)	min(l)	max(l)	avg(l)	min(t)	max(t)	avg(t)
Sample 1								
Default	91	97	3833	10000	6092,33	0,57	1,49	0,91
L2	91	97	3537	3537	3537,00	0,74	0,75	0,74
L3	91	97	5416	5416	5416,00	1,08	1,10	1,09
L4	91	97	5324	5324	5324,00	1,16	1,18	1,17
Sample 2								
Default	62	91	8492	10000	9497,33	0,67	0,80	0,75
L2	62	91	10000	10000	10000,00	1,00	1,03	1,02
L3	62	91	10000	10000	10000,00	1,06	1,07	1,07
L4	62	91	10000	10000	10000,00	1,14	1,16	1,15

Sample 3								
Default	56	56	2919	7569	5219,00	0,23	0,61	0,42
L2	56	56	8457	8457	8457,00	0,93	0,95	0,94
L3	56	56	10000	10000	10000,00	1,11	1,17	1,14
L4	56	56	4458	4458	4458,00	0,75	0,76	0,76
Sample 4								
Default	50	57	8445	10000	9481,67	0,49	0,58	0,54
L2	50	57	8409	8409	8409,00	0,68	0,71	0,70
L3	50	57	8402	8402	8402,00	0,74	0,77	0,76
L4	50	57	9058	9058	9058,00	0,85	0,89	0,87
Sample 5								
Default	86	103	3754	9341	6103,67	0,51	1,28	0,83
L2	86	103	8285	8285	8285,00	1,33	1,35	1,34
L3	86	103	5500	5500	5500,00	1,02	1,04	1,03
L4	86	103	5278	5278	5278,00	1,07	1,10	1,09
Sample 6								
Default	24	21	8550	10000	9516,67	0,20	0,23	0,22
L2	24	21	5860	5860	5860,00	0,38	0,40	0,39
L3	24	21	7675	7675	7675,00	0,50	0,59	0,53
L4	24	21	10000	10000	10000,00	0,63	0,64	0,64
Sample 7								
Default	1807	2260	10000	10000	10000,00	556,34	557,41	557,04
L2	1807	2260	10000	10000	10000,00	556,57	557,05	556,81
L3	1807	2260	10000	10000	10000,00	570,83	573,00	571,83
L4	1807	2260	10000	10000	10000,00	577,32	579,72	578,92
Sample 8								
Default	41	39	6000	10000	8666,67	0,23	0,37	0,32
L2	41	39	10000	10000	10000,00	0,57	0,59	0,58
L3	41	39	10000	10000	10000,00	0,64	0,65	0,64
L4	41	39	10000	10000	10000,00	0,71	0,71	0,71
Sample 9								
Default	343	462	1584	10000	7194,67	3,12	19,67	14,15
L2	343	462	10000	10000	10000,00	19,87	19,89	19,88
L3	343	462	10000	10000	10000,00	19,98	20,01	20,00
L4	343	462	10000	10000	10000,00	20,08	20,12	20,10
Sample 10								
Default	91	111	6709	10000	8540,67	1,32	1,95	1,67
L2	91	111	10000	10000	10000,00	2,26	2,82	2,53
L3	91	111	7451	7451	7451,00	1,79	1,81	1,80
L4	91	111	5917	5917	5917,00	1,59	1,59	1,59
Sample 11								
Default	152	131	4802	10000	7683,33	1,92	3,98	3,07
L2	152	131	10000	10000	10000,00	4,19	4,21	4,20
L3	152	131	3728	3728	3728,00	1,77	1,78	1,78
L4	152	131	9311	9311	9311,00	4,10	4,11	4,11
Sample 12								
Default	68	78	8834	10000	9611,33	0,85	0,95	0,92
L2	68	78	10000	10000	10000,00	1,16	1,29	1,21
L3	68	78	5735	5735	5735,00	0,81	0,89	0,84
L4	68	78	10000	10000	10000,00	1,28	1,33	1,31
Sample 13								
Default	15	19	950	1373	1141,67	0,01	0,06	0,03
L2	15	19	875	875	875,00	0,25	0,25	0,25
L3	15	19	2282	2282	2282,00	0,34	0,38	0,36
L4	15	19	1117	1117	1117,00	0,36	0,38	0,37
Sample 14								
Default	91	121	3792	7282	5583,67	0,67	1,28	0,98
L2	91	121	1434	1434	1434,00	0,48	0,49	0,48
L3	91	121	4061	4061	4061,00	1,01	1,03	1,02
L4	91	121	3305	3305	3305,00	0,94	0,97	0,95
Sample 15								
Default	68	73	5537	10000	7539,33	0,48	0,87	0,66
L2	68	73	5534	5534	5534,00	0,71	0,77	0,73
L3	68	73	7768	7768	7768,00	0,97	1,00	0,98
L4	68	73	7162	7162	7162,00	0,97	0,98	0,97
Sample 16								
Default	105	113	4636	8346	5977,33	0,89	1,61	1,15
L2	105	113	4046	4046	4046,00	0,99	0,99	0,99
L3	105	113	1589	1589	1589,00	0,59	0,62	0,60
L4	105	113	3070	3070	3070,00	0,97	0,99	0,98

Sample 17								
Default	28	27	4083	8518	6224,67	0,13	0,30	0,22
L2	28	27	1739	1739	1739,00	0,27	0,30	0,28
L3	28	27	770	770	770,00	0,30	0,33	0,31
L4	28	27	5011	5011	5011,00	0,57	0,66	0,62
Sample 18								
Default	824	955	10000	10000	10000,00	111,19	111,28	111,23
L2	824	955	10000	10000	10000,00	111,47	111,53	111,50
L3	824	955	10000	10000	10000,00	111,51	111,60	111,56
L4	824	955	10000	10000	10000,00	111,68	111,76	111,72
Sample 19								
Default	26	20	3862	9699	6257,67	0,13	0,35	0,23
L2	26	20	10000	10000	10000,00	0,56	0,60	0,58
L3	26	20	4898	4898	4898,00	0,38	0,44	0,41
L4	26	20	3153	3153	3153,00	0,43	0,48	0,45
Sample 20								
Default	45	41	6383	10000	7808,67	0,33	2,05	0,98
L2	45	41	4007	4007	4007,00	0,40	0,54	0,45
L3	45	41	10000	10000	10000,00	0,73	0,77	0,75
L4	45	41	10000	10000	10000,00	0,82	0,83	0,82
Sample 21								
Default	10	9	1539	7589	5516,33	0,08	0,26	0,20
L2	10	9	2799	2799	2799,00	0,28	0,33	0,31
L3	10	9	3971	3971	3971,00	0,35	0,43	0,39
L4	10	9	2085	2085	2085,00	0,40	0,48	0,45
Sample 22								
Default	4	3	2729	5501	4366,67	0,09	0,20	0,14
L2	4	3	5509	5509	5509,00	0,37	0,43	0,39
L3	4	3	2728	2728	2728,00	0,30	0,33	0,32
L4	4	3	8953	8953	8953,00	0,64	0,68	0,66
Sample 23								
Default	18	11	4969	6712	5984,33	0,06	0,09	0,08
L2	18	11	2144	2144	2144,00	0,23	0,23	0,23
L3	18	11	2518	2518	2518,00	0,29	0,31	0,30
L4	18	11	7114	7114	7114,00	0,37	0,43	0,41
Sample 24								
Default	20	14	7616	10000	8835,00	0,15	0,32	0,22
L2	20	14	7195	7195	7195,00	0,43	0,46	0,44
L3	20	14	2931	2931	2931,00	0,38	0,46	0,41
L4	20	14	4842	4842	4842,00	0,49	0,49	0,49
Sample 25								
Default	512	524	10000	10000	10000,00	41,88	42,30	42,10
L2	512	524	10000	10000	10000,00	41,53	42,48	41,96
L3	512	524	10000	10000	10000,00	41,64	42,73	42,29
L4	512	524	10000	10000	10000,00	41,63	41,70	41,66
Sample 26								
Default	493	635	10000	10000	10000,00	56,47	57,13	56,76
L2	493	635	10000	10000	10000,00	57,29	57,47	57,35
L3	493	635	10000	10000	10000,00	57,07	58,24	57,55
L4	493	635	10000	10000	10000,00	46,91	56,73	53,28
Sample 27								
Default	16	14	7586	10000	9111,67	0,20	0,32	0,25
L2	16	14	8852	8852	8852,00	0,42	0,43	0,43
L3	16	14	10000	10000	10000,00	0,55	0,58	0,56
L4	16	14	10000	10000	10000,00	0,57	0,68	0,62
Sample 28								
Default	87	114	3432	10000	6456,00	0,67	1,84	1,20
L2	87	114	8418	8418	8418,00	1,82	2,03	1,94
L3	87	114	7220	7220	7220,00	1,67	1,68	1,67
L4	87	114	10000	10000	10000,00	2,29	2,79	2,53
Sample 29								
Default	49	56	10000	10000	10000,00	0,52	0,57	0,55
L2	49	56	10000	10000	10000,00	0,75	0,77	0,76
L3	49	56	10000	10000	10000,00	0,81	0,83	0,82
L4	49	56	10000	10000	10000,00	0,89	0,91	0,90
Sample 30								
Default	35	46	3650	10000	6308,00	0,13	0,39	0,24
L2	35	46	3376	3376	3376,00	0,38	0,39	0,39
L3	35	46	10000	10000	10000,00	0,69	0,72	0,70
L4	35	46	7253	7253	7253,00	0,67	0,68	0,68

Sample 31								
Default	89	101	7069	10000	9023,00	1,14	1,49	1,37
L2	89	101	7870	7870	7870,00	1,35	1,39	1,37
L3	89	101	6406	6406	6406,00	1,20	1,22	1,21
L4	89	101	5972	5972	5972,00	1,22	1,24	1,23
Sample 32								
Default	46	46	9983	10000	9994,33	0,55	0,56	0,55
L2	46	46	10000	10000	10000,00	0,80	0,87	0,83
L3	46	46	10000	10000	10000,00	0,86	0,88	0,87
L4	46	46	10000	10000	10000,00	0,95	0,96	0,96
Sample 33								
Default	106	92	2986	5714	4514,00	0,56	1,08	0,85
L2	106	92	3449	3449	3449,00	0,86	0,87	0,87
L3	106	92	1396	1396	1396,00	0,54	0,54	0,54
L4	106	92	4878	4878	4878,00	1,26	1,27	1,27
Sample 34								
Default	30	27	10000	10000	10000,00	0,30	0,30	0,30
L2	30	27	10000	10000	10000,00	0,55	0,58	0,56
L3	30	27	10000	10000	10000,00	0,61	0,64	0,62
L4	30	27	10000	10000	10000,00	0,70	0,71	0,70
Sample 35								
Default	313	370	5082	10000	6855,67	8,33	16,38	11,24
L2	313	370	10000	10000	10000,00	16,57	16,83	16,68
L3	313	370	10000	10000	10000,00	16,62	16,69	16,66
L4	313	370	6759	6759	6759,00	11,51	11,68	11,60
Sample 36								
Default	56	65	2645	10000	5680,67	0,19	0,66	0,39
L2	56	65	4869	4869	4869,00	0,53	0,56	0,54
L3	56	65	6693	6693	6693,00	0,72	0,73	0,72
L4	56	65	5185	5185	5185,00	0,65	0,73	0,69
Sample 37								
Default	623	1148	10000	10000	10000,00	69,74	69,84	69,78
L2	623	1148	10000	10000	10000,00	69,72	69,98	69,85
L3	623	1148	10000	10000	10000,00	69,67	70,36	69,92
L4	623	1148	10000	10000	10000,00	70,38	70,49	70,43
Sample 38								
Default	64	49	2378	4474	3282,33	0,20	0,36	0,27
L2	64	49	4336	4336	4336,00	0,55	0,58	0,57
L3	64	49	2832	2832	2832,00	0,50	0,51	0,51
L4	64	49	2228	2228	2228,00	0,51	0,54	0,53
Sample 39								
Default	244	245	2304	4471	3543,00	2,27	4,40	3,49
L2	244	245	3628	3628	3628,00	3,79	3,84	3,82
L3	244	245	2303	2303	2303,00	2,59	2,75	2,65
L4	244	245	10000	10000	10000,00	10,27	10,52	10,37
Sample 40								
Default	63	67	3378	10000	6604,67	0,27	0,78	0,52
L2	63	67	3430	3430	3430,00	0,48	0,49	0,48
L3	63	67	10000	10000	10000,00	1,07	1,08	1,07
L4	63	67	10000	10000	10000,00	1,14	1,15	1,14
Sample 41								
Default	179	231	4681	6025	5266,67	2,62	3,37	2,95
L2	179	231	10000	10000	10000,00	5,75	5,83	5,79
L3	179	231	10000	10000	10000,00	5,83	5,84	5,84
L4	179	231	10000	10000	10000,00	5,89	5,91	5,90
Sample 42								
Default	129	138	4668	10000	7008,33	1,34	2,87	2,01
L2	129	138	4691	4691	4691,00	1,56	1,58	1,57
L3	129	138	1901	1901	1901,00	0,82	0,84	0,83
L4	129	138	3389	3389	3389,00	1,29	1,35	1,32
Sample 43								
Default	14	11	10000	10000	10000,00	0,29	0,33	0,31
L2	14	11	10000	10000	10000,00	0,49	0,55	0,51
L3	14	11	10000	10000	10000,00	0,52	0,62	0,57
L4	14	11	10000	10000	10000,00	0,66	0,71	0,68
Sample 44								
Default	361	531	10000	10000	10000,00	23,11	23,34	23,26
L2	361	531	10000	10000	10000,00	22,53	22,87	22,70
L3	361	531	10000	10000	10000,00	23,38	23,55	23,45
L4	361	531	10000	10000	10000,00	23,35	23,65	23,53

UDC 510.64

DOI: <https://doi.org/10.17721/1812-5409.2025/1.23>

Oksana SHKILNIAK, PhD (Phys. & Math.), Assoc. Prof.

ORCID ID: 0000-0003-4139-2525

e-mail: oksana.sh@knu.ua

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

FIRST-ORDER TRANSITIONAL MODAL LOGICS OF QUASIARY PREDICATES WITH EQUALITY

In this paper we study new classes of program-oriented logical formalisms of the modal type – first-order transitional modal logics (TML) of quasiary predicates with equality. Modal logics are successfully used for describing and modeling various subject areas. Epistemic logics are used to describe artificial intelligence, information, and expert systems. Temporal logics are used for modeling complex dynamic systems, program specification and verification. Traditional modal logics are based on classical predicate logic. However, such logic insufficiently accounts for the incompleteness, partiality and structuredness of information about the subject area. These limitations necessitate the development of new, program-oriented modal logical formalisms; this underscores the relevance of the proposed paper. We consider pure first-order TML of partial quasiary predicates without monotonicity condition and enriched with equality predicates. Two types of such logics can be distinguished: $TML^{Q=}$ with strong equality predicates $\equiv_x$, and $TML^{Q\neq}$ with weak equality predicates $\neq_{xy}$. A feature of these logics is the use of the extended renomination composition. For quantifier elimination in $TML^{Q=}$, special total indicator predicates Ez are used, while in $TML^{Q\neq}$ the partial indicator predicates are $\neq_{xx}$. Transitional modal systems (TMS) form the semantic basis of TML. For $TML^{Q=}$ and $TML^{Q\neq}$ important varieties of such systems are specified: general ($GMS^{Q=}$ and $GMS^{Q\neq}$), temporal ($TmMS^{Q=}$ and $TmMS^{Q\neq}$), and multimodal ($MMS^{Q=}$ and $MMS^{Q\neq}$). We describe semantic models and languages of these TMS and examine the properties related to equality predicates. Four types of consequence relations in TMS and logical consequence relations on sets of formulas specified with states are considered: irrefutability (IR), truth (T), falsity (F), and strong (TF). The properties of replacement of equals in $TML^{Q=}$ and $TML^{Q\neq}$ are described. The properties of logical consequence relations on sets of formulas specified with states are the semantic basis for the further construction of the corresponding sequent type calculi.

Key words: modal logic, transitional modal system, predicate, equality, logical consequence.

AMS 2020 classification: 62H30, 34C60.

Introduction

The development of information technologies has led to the emergence and practical application of new branches of mathematical logic. Among them, modal logics hold a special place. The concepts and tools of modal logics are successfully used for describing and modeling various subject areas, artificial intelligence systems, and information and software systems (see, for example, Blackburn, van Benthem, & Wolter, 2006; Abramsky, Gabbay, & Maibaum, 1993–2000). Epistemic logics are used to describe artificial intelligence, information, and expert systems, while temporal logics are used for modeling complex dynamic systems, program specification and verification (see Bjorner, & Henson, 2008; Goranko, 2023; Fisher, 2011; Kröger, & Merz, 2008). Modal logics, especially temporal ones, are widely used in model checking (see e. g. Baier, & Katoen, 2008). Traditional modal logics (see Cocchiarella, & Freund, 2008; Fitting, & Mendelsohn, 2023) are based on the classical predicate logic. However, such logic insufficiently accounts for the incompleteness, partiality and structuredness of information about the subject area. These limitations necessitate the development of new, program-oriented modal logical formalisms; this underscores the relevance of the proposed paper.

The capabilities of traditional modal logics and program-oriented composition nominative logics of partial quasiary predicates, described, in particular, in works (Shkilniak, O., & Shkilniak, S., 2023), are combined in composition nominative modal logics (CNML). An important class of CNML is transitional modal logics (TML), which reflect the aspect of change and development of subject areas. These logics have been studied, in particular, in work (Shkilniak, O., 2015). Important subclasses of TML studied in these works include general TML (GML), temporal TML (TmML) and multimodal TML (MML), with epistemic TML types further distinguished within MML.

Traditional systems of modal, temporal and epistemic logics can naturally be considered within TML.

The goal of this paper is to investigate new classes of program-oriented modal logics – pure first-order TML of partial quasiary predicates without monotonicity (equitonicity) restriction and enriched with equality predicates. Two types of such predicates can be distinguished (see Shkilniak, 2019): the weak equality $\neq_{xy}$ and the strong equality $\equiv_{xy}$. Pure first-order TML without monotonicity condition are designated as TML^Q . In work (Shkilniak, 2015), TML^Q without equality predicates have been considered. In work (Shkilniak, O., & Shkilniak S., 2024) TML^Q with strong equality predicates have been studied. In this paper we propose TML^Q with weak equality predicates. TML^Q with strong equality predicates will be designated as $TML^{Q=}$, and TML^Q with weak equality predicates will be designated as $TML^{Q\neq}$.

For quantifier elimination in the logics of non-monotone predicates, special predicates are needed to indicate the presence of a component with a corresponding name in the input data. The use of predicates-indicators is a characteristic feature of TML^Q . Total predicates-indicators determine the presence or absence of a component with a given name, while partial predicates-indicators only detect the presence of such a component. Total predicates-indicators Ez were used in works (Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016; Shkilniak, 2019; Shkilniak, 2015). It should be noted that the predicates Ez can be expressed (see Shkilniak, 2019) as $\exists y \equiv_{xy}$. However, they appear at the renominative level, so for $TML^{Q=}$, it is appropriate to define them explicitly as special 0-ary compositions. At the same time, partial predicates-indicators are already present in $TML^{Q\neq}$ as the $\neq_{xx}$ predicates.

Another feature of the studied TML^Q with equality is the presence of extended renominations, described in (Shkilniak, O., & Shkilniak S., 2023); they were used in (Shkilniak, O., & Shkilniak, S., 2024). Note that in the modal logics considered in (Shkilniak, 2015), traditional renominations were used.

The present paper describes semantic models and languages of TML^Q with equality and examines the properties of these logics related to equality predicates. A number of consequence relations in TML^Q and logical consequence relations on sets of formulas specified with states are considered, and the properties of replacement of equals in $TML^{Q=}$ and $TML^{Q\equiv}$ are determined. The properties of these relations form the semantic basis for the further construction of corresponding sequent type calculi for TML^Q with equality.

Concepts not defined here are interpreted in the sense of works (Shkilniak, 2015; Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016; Shkilniak, 2019; Shkilniak, O., & Shkilniak, S., 2023; Shkilniak, O., & Shkilniak, S., 2024).

1. Transitional modal systems

Transitional modal logics (TML) is a very important class of CNML, they describe transitions between states of the world and thus reflect the aspect of change and evolution of subject areas. Transitional modal system (TMS) is the basic concept of TML. It is an object (see Shkilniak, O., 2015) $M = (Cms, Ds, Dns)$, where:

- Cms is a composition modal system (CMS), it defines semantic aspects of the world;
- Ds is a descriptive system, it defines a set of standard descriptions – formulas of the corresponding TML language;
- Dns is a denotation system, it determines values of standard descriptions on semantic models: an interpretation mapping is used for this.

CMS have the form $Cms = (St, R, Pr, C)$, where:

- St is a set of states of the world;
- R is a set of relations on St of the form $R \subseteq St \times St$; they can be interpreted as state transition relations, hence the name;
- Pr is a set of predicates on states of the world;
- C is a set of compositions on Pr .

Pure first-order TMS (TMS of quantifier level) will be designated as TMS^Q . Here St is specified as a set of algebraic structures (systems) $\alpha = (A_\alpha, Pr_\alpha)$, where A_α is a set of basic data of the state α , Pr_α is a set of quasiary predicates $\forall A_\alpha \rightarrow \{T, F\}$, called *predicates of the state* α . Let us call *global* the predicates $\forall A \rightarrow \{T, F\}$, where $A = \bigcup_{\alpha \in S} A_\alpha$.

V-A-quasiary predicate is (see Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016) a partial function of the form $Q : \forall A \rightarrow \{T, F\}$, where V is a set of subject names (variables), A is a set of subject values (basic data), $\{T, F\}$ is the set of truth values, $\forall A$ is a set of all V-A-nominative sets. V-A-nominative set (V-A-NS) is defined as a single-valued function of the form $d : V \rightarrow A$. V-A-NS is presented as a set of pairs $[v_1 \mapsto a_1, \dots, v_n \mapsto a_n, \dots]$, where $v_i \in V$, $a_i \in A$, $v_i \neq v_j$ when $i \neq j$.

For V-A-NS we introduce an operation $\|_{-Z}$, where $Z \subseteq V$: $d \|_{-Z} = \{v \mapsto a \in d \mid v \notin Z\}$. Also a parametric operation of the extended renomination is specified $r_{\bar{x}, \perp}^{\bar{v}, \bar{u}} : \forall A \rightarrow \forall A$ (see Shkilniak O., & Shkilniak S., 2023).

Each V-A-quasiary predicate Q can be determined by two sets:

- the *truth* domain $T(Q) = \{d \in \forall A \mid Q(d) = T\}$;
- the *falsity* domain $F(Q) = \{d \in \forall A \mid Q(d) = F\}$.

Predicate Q is *single-valued*, or *P-predicate*, if $T(Q) \cap F(Q) = \emptyset$.

The class of V-A-quasiary single-valued predicates will be denoted by PrP^{V-A} .

For modal logics we specifically consider single-valued predicates.

Predicate Q is:

- *valid (irrefutable)*, if $F(Q) = \emptyset$;
- *satisfiable*, if $T(Q) \neq \emptyset$.

In the PrP^{V-A} class, there are 3 constant predicates denoted by T , F and $\perp$:

- Q is identically true, if $T(Q) = \forall A$ and $F(Q) = \emptyset$;
- Q is identically false, if $T(Q) = \emptyset$ and $F(Q) = \forall A$;
- Q is totally undefined, if $T(Q) = \emptyset$ and $F(Q) = \emptyset$.

For single-valued functions (and predicates), the concept of monotonicity is specified as equitonicity – preservation of the already accepted value after data expansion (see Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016).

P-predicate Q is *equitone*, or *PE-predicate*, if $(Q(d) \downarrow \text{ and } d \subseteq d') \Rightarrow Q(d') \downarrow = Q(d)$.

Subject name $x \in V$ is *unessential* for the predicate Q , if $d \|_{-x} = h \|_{-x} \Rightarrow Q(d) = Q(h)$.

The set of compositions in TMS is defined by the basic *logical* compositions and basic *modal* compositions.

Basic logical compositions for TMS^Q are logical connectives $\neg$ and $\vee$, renomination $R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}$ and existential quantification $\exists x$.

For TMS^Q with equality, equality predicates are added as specific 0-ary compositions. Two types of equality predicates are considered: weak equality $=_{xy}$ and strong equality $\equiv_{xy}$. Thus, TMS^Q with predicates $\equiv_{xy}$ will be called $TMS^{Q=}$, and TMS^Q with predicates $=_{xy}$ will be called $TMS^{Q\equiv}$.

Other compositions like $\&$, $\rightarrow$, $\forall x$ can be definable from these.

Basic compositions $\neg$, $\vee$, $R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}$, $\exists x$ and derivative compositions $\&$, $\rightarrow$, $\forall x$ are defined in (Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016; Shkilniak, O., & Shkilniak, S., 2023).

Let us specify two types of equality predicates according to work (Shkilniak, 2019):

- weak (up to definability) equality predicates $=_{\{x, y\}}$;
- strong (strict) equality predicates $\equiv_{\{x, y\}}$.

We define $=_{\{x, y\}}$ and $\equiv_{\{x, y\}}$ through their truth and falsity domains:

$$\begin{aligned} T(=_{\{x, y\}}) &= \{d \mid d(x) \downarrow, d(y) \downarrow \text{ and } d(x) = d(y)\}, \\ F(=_{\{x, y\}}) &= \{d \mid d(x) \downarrow, d(y) \downarrow \text{ and } d(x) \neq d(y)\}; \\ T(\equiv_{\{x, y\}}) &= \{d \mid d(x) \downarrow, d(y) \downarrow \text{ and } d(x) = d(y)\} \cup \{d \mid d(x) \uparrow \text{ and } d(y) \uparrow\}, \\ F(\equiv_{\{x, y\}}) &= \{d \mid d(x) \downarrow, d(y) \downarrow, d(x) \neq d(y)\} \cup \{d \mid d(x) \downarrow, d(y) \uparrow \text{ or } d(x) \uparrow, d(y) \downarrow\}. \end{aligned}$$

In case x and y match, predicates $=_{\{x,y\}}$ and $\equiv_{\{x,y\}}$ become $=_{\{x\}}$ and $\equiv_{\{x\}}$.

Further on, predicates $=_{\{x,y\}}$, $=_{\{x\}}$ and $\equiv_{\{x,y\}}$, $\equiv_{\{x\}}$ will be denoted more traditionally by $=_{xy}$, $=_{xx}$ and $\equiv_{xy}$, $\equiv_{xx}$ respectively. From the definitions implies that pairs $=_{xy}$ and $=_{yx}$, $\equiv_{xy}$ and $\equiv_{yx}$ represent the same predicates.

Predicates $\equiv_{xy}$ and $\equiv_{xx}$ are total and non-monotone (non-equitone), while predicates $=_{xy}$ and $=_{xx}$ are partial and equitone.

Special predicates-indicators which detect whether a component with a corresponding name has a value in the input data are needed for quantifier elimination. Total predicates-indicators can distinguish either the subject variable *has* an assigned value or *not*, while partial predicates-indicators can only *ascertain* that a value is assigned.

Total predicates-indicators Ez are non-monotone and defined (see Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016) as follows:

$$T(Ez) = \{d \mid d(z) \downarrow\};$$

$$F(Ez) = \{d \mid d(z) \uparrow\}.$$

$TMS^{Q=}$ don't need separate predicates-indicators as predicates $=_{zz}$ *already* work that way:

$$T(=_{zz}) = \{d \mid d(z) \downarrow\} = T(Ez);$$

$$F(=_{zz}) = \{d \mid d(z) \downarrow \text{ and } d(z) \neq d(z)\} = \emptyset.$$

Equality predicates and predicates-indicators can be interpreted as special 0-ary parametric compositions.

We have (see Shkilniak, S., 2019) $Ez = \exists y \equiv_{zy}$, thus predicates Ez are derivative in $TMS^{Q=}$. However, Ez can be introduced on the renominative level, hence it is convenient to specify them explicitly as 0-ary compositions.

Therefore, we obtain the following sets of basic logical compositions: $\{\neg, \vee, R_{x,\perp}^{\bar{v},\bar{u}}, \exists x, \equiv_{xy}, Ez\}$ for $TMS^{Q=}$ and $\{\neg, \vee, R_{x,\perp}^{\bar{v},\bar{u}}, \exists x, =_{xy}\}$ for $TMS^{Q=}$. Their properties are described in (Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016; Shkilniak, S., 2019; Shkilniak, O., & Shkilniak, S., 2023).

Important special cases of TMS are (see Shkilniak, 2015):

- *general* TMS (GMS), with $R = \{\triangleright\}$ and a basic modal composition $\Box$ ("necessary");
- *temporal* TMS (TmMS), with $R = \{\triangleright\}$ and basic modal compositions $\Box\uparrow$ ("it will always be the case") i $\Box\downarrow$ ("it has always been the case");
- *multimodal* TMS (MMS), with $R = \{\triangleright_i \mid i \in I\}$ and basic modal compositions M_i , $i \in I$ in which each $\triangleright_i \in R$ is matched with a corresponding modal composition M_i ; each M_i is basically $\Box$ with respect to its relation $\triangleright_i$.

Derivative compositions are also usually defined. For GMS, $\Diamond$ ("possibly"): $\Diamond P$ means $\neg\Box\neg P$.

For TmMS, $\Diamond\uparrow$ ("it will sometime be the case") and $\Diamond\downarrow$ ("it was sometime the case"): $\Diamond\uparrow P$ means $\neg\Box\uparrow\neg P$, $\Diamond\downarrow P$ means $\neg\Box\downarrow\neg P$.

The following types of TMS^Q with equality can be obtained. For $TMS^{Q=}$, we have $GMS^{Q=}$, $TmMS^{Q=}$, $MMS^{Q=}$; for $TMS^{Q=}$, we have $GMS^{Q=}$, $TmMS^{Q=}$, $MMS^{Q=}$.

2. Languages of transitional modal systems

Let us describe the language of $GMS^{Q=}$. The alphabet: a set V of subject names (variables); a set Ps of predicate symbols (language signature); a set of basic logical compositions' symbols $\{\neg, \vee, R_{x,\perp}^{\bar{v},\bar{u}}, \exists x, \equiv_{xy}, Ex\}$; a set of basic modal compositions' symbols (modal signature) $M_s = \{\Box\}$. The set Fr of language formulas is determined inductively:

$$Fa) Ps \subseteq Fr,$$

$$Fc) \{Ex \mid x \in V\} \subseteq Fr \text{ and } \{\equiv_{xy} \mid x, y \in V\} \subseteq Fr,$$

$$FpR) \Phi, \Psi \in Fr \Rightarrow \neg\Phi \in Fr, \vee\Phi\Psi \in Fr, \text{ and } R_{x,\perp}^{\bar{v},\bar{u}}\Phi \in Fr;$$

$$F\exists) \Phi \in Fr \Rightarrow \exists x\Phi \in Fr,$$

$$F\Box) \Phi \in Fr \Rightarrow \Box\Phi \in Fr.$$

Formulas of the type $p \in Ps$, Ex , $\equiv_{xy}$ are *atomic*.

We will use traditional derivative compositions and thus write $\Phi \& \Psi$, $\Phi \rightarrow \Psi$, $\forall x\Phi$, $\Diamond\Phi$ in place of formulas $\vee\neg\Phi\Psi$, $\neg\neg\Phi\neg\Psi$, $\neg\exists x\neg\Phi$, $\neg\Box\neg\Phi$.

Sets of guaranteed to be unessential names for formulas are defined by the function $v : Fr \rightarrow 2^V$ (see Shkilniak, 2019; Shkilniak, O., & Shkilniak, S., 2023).

A $GMS^{Q=}$ type is determined by an extended signature $\sigma = (Ps, v)$ and properties of the relation $\triangleright$.

An interpretation mapping Im is specified for formulas for TMS^Q . compositions' symbols are interpreted as corresponding compositions. For basic predicates, we have $Im : PS \times ST \rightarrow PR$, with condition $IT(P, \alpha) \in PR_\alpha$. the mapping Im is continued to $FT \times ST \rightarrow PR$ as follows:

$$IpR) Im(\neg, \alpha) = \neg(Im(\Phi, \alpha)); Im(\vee\Phi\Psi, \alpha) = \vee(Im(\Phi, \alpha), Im(\Psi, \alpha)); Im(R_{x,\perp}^{\bar{v},\bar{u}}(\Phi), \alpha) = R_{x,\perp}^{\bar{v},\bar{u}}(Im(\Phi, \alpha));$$

$$F\exists) \Phi \in Fr \Rightarrow \exists x\Phi \in Fr,$$

$$F\Box) \Phi \in Fr \Rightarrow \Box\Phi \in Fr.$$

Formulas of the type $p \in Ps$, Ex , $\equiv_{xy}$ are *atomic*.

We will use traditional derivative compositions and thus write $\Phi \& \Psi$, $\Phi \rightarrow \Psi$, $\forall x\Phi$, $\Diamond\Phi$ in place of formulas $\vee\neg\Phi\Psi$, $\neg\neg\Phi\neg\Psi$, $\neg\exists x\neg\Phi$, $\neg\Box\neg\Phi$.

Sets of guaranteed to be unessential names for formulas are defined by the function $v : Fr \rightarrow 2^V$ (see Shkilniak, 2019; Shkilniak, O., & Shkilniak, S., 2023).

A $GMS^{Q=}$ type is determined by an extended signature $\sigma = (Ps, v)$ and properties of the relation $\triangleright$.

An interpretation mapping Im is specified for formulas for TMS^Q . compositions' symbols are interpreted as corresponding compositions. For basic predicates, we have $Im: PS \times ST \rightarrow PR$, with condition $IT(P, \alpha) \in PR_\alpha$. the mapping Im is continued to $FT \times ST \rightarrow PR$ as follows:

$$\begin{aligned} \text{IpR)} \quad Im(\neg, \alpha) &= \neg(Im(\Phi, \alpha)); \quad Im(\vee \Phi \Psi, \alpha) = \vee(Im(\Phi, \alpha), Im(\Psi, \alpha)); \quad Im(R_{x, \perp}^{\bar{v}, \bar{u}}(\Phi), \alpha) = R_{x, \perp}^{\bar{v}, \bar{u}}(Im(\Phi, \alpha)); \\ \text{I}\exists) \quad Im(\exists x \Phi, \alpha)(d) &= \begin{cases} T, & \text{if } Im(\Phi, \alpha)(d \parallel_{-x} \cup x \mapsto a) = T \text{ for some } a \in A_\alpha, \\ F, & \text{if } Im(\Phi, \alpha)(d \parallel_{-x} \cup x \mapsto a) = F \text{ for any } a \in A_\alpha, \\ & \text{else undefined;} \end{cases} \\ \text{I}\Box) \quad Im(\Box \Phi, \alpha)(d) &= \begin{cases} T, & \text{if } Im(\Phi, \delta)(d) = T \text{ for any } \delta \in St : \alpha \triangleright \delta, \\ F, & \text{if exists } \delta \in St : \alpha \triangleright \delta \text{ and } Im(\Phi, \delta)(d) = F, \\ & \text{else undefined.} \end{cases} \end{aligned}$$

Given for $\alpha \in St$ there is no β such that $\alpha \triangleright \beta$, then for each $d \in {}^V A$ we define $Im(\Box \uparrow \Phi, \alpha)(d) \uparrow$.

The interpretation mapping Im can be specified in the similar manner for the derivative formulas $\Phi \& \Psi$, $\Phi \rightarrow \Psi$, $\forall x \Phi$, $\Diamond \Phi$.

Predicates which are values of formulas that do not use modalities (i.e. contain no symbols from Ms) belong to predicates of states. Predicates which are values of modalized formulas are global.

The language of $GMS^{Q=}$ is defined similarly. It has a different set of basic logical compositions' symbols: $\{\neg, \vee, R_{x, \perp}^{\bar{v}, \bar{u}}, \exists x, =_{xy}\}$.

The rule Fc for language formulas has a form $\{=_{xy} \mid x, y \in V\} \subseteq Fr$; the interpretation mapping is modified accordingly.

The languages of $TmMS^{Q=}$ and $TmMS^{Q=}$ are also specified similarly to the languages of $GMS^{Q=}$ and $GMS^{Q=}$.

The set of basic modal compositions' symbols is $Ms = \{\Box \uparrow, \Box \downarrow\}$. In the definition of language formulas Fm , the rule $F\Box$ changes to:

$$F\Box \uparrow \downarrow) \quad \Phi \in Fr \Rightarrow \Box \uparrow \Phi \in Fr \text{ and } \Box \downarrow \Phi \in Fr.$$

Short notation $\Diamond \uparrow \Phi$ AND $\Diamond \downarrow \Phi$ can be used instead of $\neg \Box \uparrow \neg \Phi$ AND $\neg \Box \downarrow \neg \Phi$.

Interpretation mapping Im should take into account time direction:

$$\begin{aligned} \text{I}\Box \uparrow \downarrow) \quad Im(\Box \uparrow \Phi, \alpha)(d) &= \begin{cases} T, & \text{if } Im(\Phi, \delta)(d) = T \text{ for any } \delta \in St : \alpha \triangleright \delta, \\ F, & \text{if exists } \delta \in St : \alpha \triangleright \delta \text{ and } Im(\Phi, \delta)(d) = F, \\ & \text{else undefined.} \end{cases} \\ Im(\Box \downarrow \Phi, \alpha)(d) &= \begin{cases} T, & \text{if } Im(\Phi, \delta)(d) = T \text{ for any } \delta \in St : \delta \triangleright \alpha, \\ F, & \text{if exists } \delta \in St : \delta \triangleright \alpha \text{ and } Im(\Phi, \delta)(d) = F, \\ & \text{else undefined.} \end{cases} \end{aligned}$$

Given for $\alpha \in St$ there is no β such that $\alpha \triangleright \beta$, then for each $d \in {}^V A$ we define $Im(\Box \uparrow \Phi, \alpha)(d) \uparrow$.

Given for $\alpha \in St$ there is no β such that $\beta \triangleright \alpha$, then for each $d \in {}^V A$ we define $Im(\Box \downarrow \Phi, \alpha)(d) \uparrow$.

Similarly, the languages of $MMS^{Q=}$ and $MMS^{Q=}$ are defined in the same fashion as the languages of $GMS^{Q=}$ and $GMS^{Q=}$.

The set of basic modal compositions' symbols changes to $Ms = \{M_i \mid i \in I\}$. In the definition of language formulas Fm , the rule $F\Box$ changes to $FM) \quad \Phi \in Fr \Rightarrow M_i \Phi \in Fr$.

For interpretation mapping Im , for modalities we have:

$$\text{IM)} \quad Im(M_i \Phi, \alpha)(d) = \begin{cases} T, & \text{if } Im(\Phi, \delta)(d) = T \text{ for any } \delta \in St : \alpha \triangleright_i \delta, \\ F, & \text{if exists } \delta \in St : \alpha \triangleright_i \delta \text{ and } Im(\Phi, \delta)(d) = F, \\ & \text{else undefined.} \end{cases}$$

Given for $\alpha \in St$ there is no β such that $\alpha \triangleright_i \beta$, then for each $d \in {}^V A$ we define $Im(M_i \Phi, \alpha)(d) \uparrow$.

In case the interpretation Im is defined explicitly, TMS can be represented as $\mathbf{M} = (St, \mathbf{R}, A, Im)$.

Predicate $Im(\Phi, \alpha)$ which is a value of the formula Φ in the state α is denoted by Φ_α .

Formula Φ is *valid (irrefutable) on state* α (denoted $\alpha \models \Phi$), if Φ_α is a valid (irrefutable) predicate.

Formula Φ is *valid in TMS* $\mathbf{M}$ (denoted $\mathbf{M} \models \Phi$), if for each $\alpha \in St$ the predicate Φ_α is valid.

Let $\mathcal{M}$ be a TMS class of a given type. Formula Φ is $\mathcal{M}$ -*valid* (denoted $\mathcal{M} \models \Phi$), if $\mathbf{M} \models \Phi$ for all $\mathbf{M} \in \mathcal{M}$.

Depending on conditions imposed on the relation $\triangleright$, different classes of TMS^Q can be specified. Traditionally, cases where $\triangleright$ can be reflexive, symmetric or transitive are considered. In case $\triangleright$ is reflexive/transitive/symmetric, then the corresponding symbol R/T/S will appear in the TMS^Q name. Thus, the following classes can be obtained:

$$\begin{aligned} &R\text{-}GMS^{Q=}, T\text{-}GMS^{Q=}, S\text{-}GMS^{Q=}, RT\text{-}GMS^{Q=}, RS\text{-}GMS^{Q=}, TS\text{-}GMS^{Q=}, RTS\text{-}GMS^{Q=}, \\ &R\text{-}GMS^{Q=}, T\text{-}GMS^{Q=}, S\text{-}GMS^{Q=}, RT\text{-}GMS^{Q=}, RS\text{-}GMS^{Q=}, TS\text{-}GMS^{Q=}, RTS\text{-}GMS^{Q=}, \\ &R\text{-}TmMS^{Q=}, T\text{-}TmMS^{Q=}, S\text{-}TmMS^{Q=}, RT\text{-}TmMS^{Q=}, RS\text{-}TmMS^{Q=}, TS\text{-}TmMS^{Q=}, RTS\text{-}TmMS^{Q=}, \\ &R\text{-}TmMS^{Q=}, T\text{-}TmMS^{Q=}, S\text{-}TmMS^{Q=}, RT\text{-}TmMS^{Q=}, RS\text{-}TmMS^{Q=}, TS\text{-}TmMS^{Q=}, RTS\text{-}TmMS^{Q=}. \end{aligned}$$

For MMS, the situation is more complicated. Given that all $\triangleright_i$ have the same type, we can obtain pure $MMS^{Q=}$ classes similar to the corresponding $GMS^{Q=}$ classes. However, more complex mixed types are also possible: for example, transitive $\triangleright_1$, symmetric $\triangleright_2$, reflexive and transitive $\triangleright_3$ etc.

Let us consider how the value of $\Phi_\delta(d)$ can be set in case $d \notin {}^V A_\delta$. Two types of TMS are specified (see also Shkilniak, O., & Shkilniak, S., 2024): with *strong* condition of definedness on states and with *general* condition of definedness on states. The *strong* condition requires $\Phi_\delta(d) \uparrow$ for $d \notin {}^V A_\delta$. Hence: $\Phi_\delta(d) \downarrow \Rightarrow d \in {}^V A_\delta$. Then $(\Box \Phi)_\alpha(d) = T$ implies that $d \in {}^V A_\delta$ for all δ such that $\alpha \triangleright \delta$. This means that basic objects (basic data) cannot disappear when transitioning to a successor state, which imposes

strong restrictions on semantic models. Equitonicity of predicates with modalities is also violated. Therefore, this condition turns out too restrictive.

The *general* condition of definedness on states looks more organic. Let d_δ denote the name set $[V \mapsto a \in d \mid a \in A_\delta]$. For non-modalized formulas for all $d \in {}^V A$ and $\delta \in St$ we specify $\Phi_\delta(d) = \Phi_\delta(d_\delta)$. Given $d \in {}^V A_\delta$, we obtain $\Phi_\delta(d) = \Phi_\delta(d_\delta)$. This means that predicates on states δ "perceive" only such components $v \mapsto a$ that $a \in A_\delta$. Moreover, in TMS with the general condition of definedness on states basic modal compositions preserve equitonicity. This is an additional reason to concentrate on TMS with such condition of definedness on states.

Further on, we will consider the case of GMS^Q . Similar statements can be formulated for $TmMS$ and MMS .

Let us consider the interaction of modalities with renominations and quantifiers.

Symbols of modal compositions Ms can be carried through renominations:

Theorem 1. $R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}(\Box\Phi)(d) = \Box(R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}(\Phi))(d)$ for arbitrary $\Phi \in Fr$, $d \in {}^V A$.

Corollary 1. Formulas $R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}(\Box\Phi) \rightarrow \Box(R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}(\Phi))$ and $\Box(R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}(\Phi)) \rightarrow R_{\bar{x}, \perp}^{\bar{v}, \bar{u}}(\Box\Phi)$ are valid.

The interaction of modal compositions and quantifiers is more complex. For modalities $\Box$ and $\Diamond$ and quantifiers $\exists x$ and $\forall x$ we get:

Theorem 2 (see also Shkilniak, O., & Shkilniak, S., 2024). In a *GMS of equitone predicates*:

- 1) formulas $\Box\exists x\Phi \rightarrow \exists x\Box\Phi$, $\forall x\Box\Phi \rightarrow \Box\forall x\Phi$, $\forall x\Diamond\Phi \rightarrow \Diamond\forall x\Phi$, $\Diamond\exists x\Phi \rightarrow \exists x\Diamond\Phi$ are refutable;
- 2) formulas $\exists x\Box\Phi \rightarrow \Box\exists x\Phi$, $\Box\forall x\Phi \rightarrow \forall x\Box\Phi$, $\Diamond\forall x\Phi \rightarrow \forall x\Diamond\Phi$, $\exists x\Diamond\Phi \rightarrow \Diamond\exists x\Phi$ are valid (irrefutable).

At the same time, formulas $\exists x\Box\Phi \rightarrow \Box\exists x\Phi$, $\Box\forall x\Phi \rightarrow \forall x\Box\Phi$, $\Diamond\forall x\Phi \rightarrow \forall x\Diamond\Phi$, $\exists x\Diamond\Phi \rightarrow \Diamond\exists x\Phi$ are refutable in GMS^Q without monotonicity restrictions.

3. Properties of GMS^Q with equality

Let us consider properties of equality predicates in GMS^Q .

Assertion 1. 1) Formulas $\equiv_{xy}$ and $\Box\equiv_{xy}$ are refutable in $GMS^{Q=}$ with 1-element states;

2) Formulas $=_{xy}$ and $\Box=_{xy}$ are irrefutable in $GMS^{Q=}$ with 1-element states.

Proof. Let us specify such $GMS^{Q=}$. We define $St = \{\alpha, \beta\}$, $R = \{\alpha \triangleright \beta\}$, $A_\alpha = A_\beta = \{a\}$. Let us take $d = [x \mapsto a]$. Then $d_\alpha(y) \uparrow$ and $d_\beta(y) \uparrow$, hence $(\equiv_{xy})_\alpha(d) = F$ and $(\equiv_{xy})_\beta(d) = F$, therefore $(\Box\equiv_{xy})_\alpha(d) = F$. So, $\alpha \not\models \equiv_{xy}$ and $\alpha \not\models \Box\equiv_{xy}$.

For 1-element interpretations with $A = \{a\}$ for each $d \in {}^V \{a\}$ impossible $d(x) \downarrow$, $d(y) \downarrow$ and $d(x) \neq d(y)$, hence impossible $=_{xy}(d) = F$. Therefore, $=_{xy}$ and $\Box=_{xy}$ are irrefutable in $GMS^{Q=}$ with 1-element states.

Note that formulas $x = y$ are irrefutable under 1-element interpretations in classical logic as well.

For equality predicates we have $F(\equiv_{xx}) = \emptyset$ and $F(=_{xx}) = \emptyset$. We obtain

Assertion 2. 1) $M \models \equiv_{xx}$ and $M \models \Box\equiv_{xx}$ for each $GMS^{Q=}$ M ;

2) $M \models =_{xx}$ and $M \models \Box=_{xx}$ for each $GMS^{Q=}$ M .

Proof. Predicate $=_{xx}$ is partial, therefore formulas $=_{xx}$ and $\Box=_{xx}$ cannot be interpreted as a constant predicate T.

Predicate $\equiv_{xx}$ is a constant T, hence formula $\equiv_{xx}$ is always interpreted as T. At the same time, formula $\Box\equiv_{xx}$ is not always interpreted as T. Indeed, let us consider $GMS^{Q=}$ with $St = \{\alpha, \beta\}$ and $R = \{\alpha \triangleright \beta\}$. As there does not exist a state η such that $\beta \triangleright \eta$, therefore $(\Box\equiv_{xx})_\beta(d) \uparrow$ for each d .

Theorem 3. Formulas $=_{xy} \rightarrow \Box=_{xy}$ and $\Box=_{xy} \rightarrow =_{xy}$ are irrefutable in $GMS^{Q=}$.

Proof. Let us prove that $=_{xy} \rightarrow \Box=_{xy}$ by contradiction. Suppose $=_{xy} \rightarrow \Box=_{xy}$ is refuted in a $GMS^{Q=}$ $M = (St, R, A, Im)$. Then for some $\alpha \in St$ and $d \in {}^V A$ we have $(=_{xy})_\alpha(d_\alpha) = F$, hence $(=_{xy})_\alpha(d_\alpha) = T$ and $(\Box=_{xy})_\alpha(d_\alpha) = F$. The latter implies that there exists $\beta \in St$ such that $\alpha \triangleright \beta$ and $(=_{xy})_\beta(d_\beta) = F$, hence $d_\beta(x) \downarrow$, $d_\beta(y) \downarrow$ and $d_\beta(x) \neq d_\beta(y)$. Therefore, there exist $a, b \in A_\beta$ such that $a \neq b$, $d_\beta(x) = a$ and $d_\beta(y) = b$, hence $d(x) = a$ and $d(y) = b$.

$(=_{xy})_\alpha(d_\alpha) = T$ implies that $d_\alpha(x) \downarrow$, $d_\alpha(y) \downarrow$ and $d_\alpha(x) = d_\alpha(y)$. Given $d_\alpha(x) \downarrow$ and $d_\alpha(y) \downarrow$, we have $d_\alpha(x) = d_\alpha(y) = a$ and $d_\alpha(y) = d_\alpha(y) = b$. However, $a \neq b$, which contradicts $d_\alpha(x) = d_\alpha(y)$.

Let us prove the irrefutability of $\Box=_{xy} \rightarrow =_{xy}$. Assume the opposite: $\Box=_{xy} \rightarrow =_{xy}$ is refuted in a $GMS^{Q=}$ $M = (St, R, A, Im)$. Then for some $\alpha \in St$ and $d \in {}^V A$ we have $(\Box=_{xy})_\alpha(d_\alpha) = F$, hence $(\Box=_{xy})_\alpha(d_\alpha) = T$ and $(=_{xy})_\alpha(d_\alpha) = F$. From $(\Box=_{xy})_\alpha(d_\alpha) = T$ we obtain: for all $\beta \in St$ such that $\alpha \triangleright \beta$, we have $(=_{xy})_\beta(d_\beta) = T$, therefore for all such β we have $d_\beta(x) \downarrow$, $d_\beta(y) \downarrow$ and $d_\beta(x) = d_\beta(y)$. Such β that $\alpha \triangleright \beta$ should exist, else $(\Box=_{xy})_\alpha(d) \uparrow$ for all $d \in {}^V A$. Let us fix such a β . Then there exists $a \in A_\beta$ such that $d_\beta(x) = d_\beta(y) = a$, hence $d(x) = d(y) = a$. From $(=_{xy})_\alpha(d_\alpha) = F$ we obtain $d_\alpha(x) \downarrow$, $d_\alpha(y) \downarrow$ and $d_\alpha(x) \neq d_\alpha(y)$. From $d_\alpha(x) \downarrow$ and $d_\alpha(y) \downarrow$ we obtain $d_\alpha(x) = d_\alpha(y) = a$ and $d_\alpha(y) = d_\alpha(y) = b$. However, $d(x) = d(y) = a$, hence $d_\alpha(x) = d_\alpha(y) = a$, which contradicts $d_\alpha(x) \neq d_\alpha(y)$.

Corollary 2. Formulas $=_{xx} \rightarrow \Box=_{xx}$ and $\Box=_{xx} \rightarrow =_{xx}$ are irrefutable in $GMS^{Q=}$.

For $GMS^{Q=}$, the situation is different (see (Shkilniak O. and Shkilniak S., 2024)):

Theorem 4. Formulas $\equiv_{xy} \rightarrow \Box\equiv_{xy}$ and $\Box\equiv_{xy} \rightarrow \equiv_{xy}$ are refutable in $GMS^{Q=}$.

The following statement is similar to Theorem 4, but formulated for predicates-indicators Ex .

Theorem 5. Formulas $Ex \rightarrow \Box Ex$ and $\Box Ex \rightarrow Ex$ are refutable in $GMS^{Q=}$.

Proof. Let us consider $GMS^{Q=}$ with $St = \{\alpha, \beta\}$, $R = \{\alpha \triangleright \beta\}$, $A_\alpha = \{a\}$, $A_\beta = \{b\}$. Let $d = [x \mapsto a, y \mapsto b] \Rightarrow d_\alpha = [x \mapsto a]$ and $d_\beta = [y \mapsto b]$. Hence $d_\alpha(x) = a$ and $d_\beta(x) \uparrow \Rightarrow Ex_\alpha(d) = Ex_\alpha(d_\alpha) = T$ and $Ex_\beta(d) = Ex_\beta(d_\beta) = F \Rightarrow (\Box Ex)_\alpha(d) = F$. Therefore, $(Ex \rightarrow \Box Ex)_\alpha(d) = F \Rightarrow \alpha \not\models Ex \rightarrow \Box Ex$.

Let $h = [x \mapsto b, y \mapsto a] \Rightarrow d_\alpha = [y \mapsto a]$ and $d_\beta = [x \mapsto b]$. Hence $d_\alpha(x) \uparrow$ and $d_\beta(x) = b \Rightarrow Ex_\alpha(d) = Ex_\alpha(d_\alpha) = F$ and $Ex_\beta(d) = Ex_\beta(d_\beta) = T \Rightarrow (\Box Ex)_\alpha(d) = T$. Therefore, $(\Box Ex \rightarrow Ex)_\alpha(d) = F \Rightarrow \alpha \not\models \Box Ex \rightarrow Ex$.

At the same time, everything is different for partial predicates-indicators $=_{xx}$ (see Corollary 2).

Theorem 6. Formula $\equiv_{xy} \& \Box \equiv_{xz} \rightarrow \Box \equiv_{yz}$ is refutable in $GMS^{Q=}$.

Proof. Let us consider $GMS^{Q=}$ with $St = \{\alpha, \beta\}$, $R = \{\alpha \triangleright \beta\}$, $A_\alpha = \{a\}$, $A_\beta = \{b, c\}$. Let $d = [x \mapsto c, y \mapsto b, z \mapsto c, s \mapsto a] \Rightarrow d_\alpha = [s \mapsto a]$ and $d_\beta = [x \mapsto c, y \mapsto b, z \mapsto c]$. Hence $d_\alpha(x) \uparrow$, $d_\alpha(y) \uparrow$; $d_\beta(x) = c$, $d_\beta(y) = b$, $d_\beta(z) = c$.

Therefore, $(\equiv_{xy})_\alpha(d) = (\equiv_{xy})_\alpha(d_\alpha) = T$; $(\equiv_{xz})_\beta(d) = (\equiv_{xz})_\beta(d_\beta) = T$, $(\equiv_{yz})_\beta(d) = (\equiv_{yz})_\beta(d_\beta) = F$.

Thence we obtain: $(\equiv_{xz})_\beta(d) = T$ and $\alpha \triangleright \beta \Rightarrow (\Box \equiv_{xz})_\alpha(d) = T$; $(\equiv_{yz})_\beta(d) = F$ and $\alpha \triangleright \beta \Rightarrow (\Box \equiv_{yz})_\alpha(d) = F$.

Therefore, $(\equiv_{xy} \& \Box \equiv_{xz} \rightarrow \Box \equiv_{yz})_\alpha(d) = F \Rightarrow \alpha \not\models \equiv_{xy} \& \Box \equiv_{xz} \rightarrow \Box \equiv_{yz}$.

Theorem 7. Formula $\equiv_{xy} \& \Box \equiv_{xz} \rightarrow \Box \equiv_{yz}$ is irrefutable in $GMS^{Q=}$.

Proof. Let us assume the opposite: $\equiv_{xy} \& \Box \equiv_{xz} \rightarrow \Box \equiv_{yz}$ is refuted in a $GMS^{Q=}$ $M = (St, R, A, Im)$. Then for some $\alpha \in St$ and $d \in {}^V A$ we have $(\equiv_{xy} \& \Box \equiv_{xz} \rightarrow \Box \equiv_{yz})_\alpha(d_\alpha) = F$, hence $(\equiv_{xy})_\alpha(d_\alpha) = T$, $(\Box \equiv_{xz})_\alpha(d_\alpha) = T$ and $(\Box \equiv_{yz})_\alpha(d_\alpha) = F$. From $(\Box \equiv_{yz})_\alpha(d_\alpha) = F$ we obtain: there exists $\beta \in St$ such that $\alpha \triangleright \beta$ and $(\equiv_{yz})_\beta(d_\beta) = F$, thence $d_\beta(y) \downarrow$, $d_\beta(z) \downarrow$ and $d_\beta(y) \neq d_\beta(z)$. Therefore, $a, b \in A_\beta$ should exist such that: $a \neq b$, $d_\beta(y) = a$ and $d_\beta(z) = b$, hence $d(y) = a$ and $d(z) = b$.

$(\Box \equiv_{xz})_\alpha(d_\alpha) = T$ implies that for all $\gamma \in St$ such that $\alpha \triangleright \gamma$, we have $(\equiv_{xz})_\gamma(d_\gamma) = T$, in particular, for a state β we have $(\equiv_{xz})_\beta(d_\beta) = T$. Hence $d_\beta(x) \downarrow$, $d_\beta(z) \downarrow$ and $d_\beta(x) = d_\beta(z)$, which implies $d(x) = d(z)$. However, $d(z) = b$, therefore $d(x) = b$.

$(\equiv_{xy})_\alpha(d_\alpha) = T$ implies $d_\alpha(x) \downarrow$, $d_\alpha(y) \downarrow$ and $d_\alpha(x) = d_\alpha(y)$, thence $d(x) = d(y)$. However, $d(x) = b$ and $d(y) = a$, and $a \neq b$. We came to contradiction.

Therefore as we see, there are significant differences in semantic properties of $GMS^{Q=}$ and $GMS^{Q=}$.

4. Logical consequence relations on a set of formulas specified with states

We define logical consequence relations in TMS on a set of formulas specified with names of states (or states of formulas). The specification indicates the state of the world on which the formula is considered. A formula specified with a state is denoted by Φ^α , where Φ is a formula of the language, $\alpha \in S$ is its specification, S is a set of names of states of the world.

Let Σ be a set of formulas specified with states with a specifications' set S .

We will call a set Σ *consistent* with TMS $M = (St, R, A, Im)$, provided that an injection of S into St is defined.

For sets of formulas specified with states, the notation $\Gamma_M \models \Delta$ assumes the consistency of Γ and Δ with TMS M .

On sets of formulas specified with states, let us introduce (see Shkilniak, O., & Shkilniak, S., 2024) the relations of irrefutability (*IR*), truth (*T*), falsity (*F*), and strong (*TF*) logical consequence. Such relations are specified similarly to the corresponding relations in the traditional logics of quasiary predicates (see Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016; Shkilniak, 2019).

Let Δ and Γ be sets of formulas specified with states.

Δ is an *IR*-consequence of Γ in a consistent with them TMS M (denoted $\Gamma_M \models_{IR} \Delta$), if for all $d \in {}^V A$:

$\Phi_\alpha(d) = T$ for all $\Phi^\alpha \in \Gamma \Rightarrow \Psi_\beta(d) \neq F$ for some $\Psi^\beta \in \Delta$.

Δ is a *T*-consequence of Γ in a consistent with them TMS M (denoted $\Gamma_M \models_T \Delta$), if for all $d \in {}^V A$:

$\Phi_\alpha(d) = T$ for all $\Phi^\alpha \in \Gamma \Rightarrow \Psi_\beta(d) = T$ for some $\Psi^\beta \in \Delta$.

Δ is a *F*-consequence of Γ in a consistent with them TMS M (denoted $\Gamma_M \models_F \Delta$), if for all $d \in {}^V A$:

$\Psi_\beta(d) = F$ for all $\Psi^\beta \in \Delta \Rightarrow \Phi_\alpha(d) = F$ for some $\Phi^\alpha \in \Gamma$

Δ is a *TF*-consequence of Γ in a consistent with them TMS M (denoted $\Gamma_M \models_{TF} \Delta$), if $\Gamma_M \models_T \Delta$ and $\Gamma_M \models_F \Delta$.

The relation of logical *-consequence with respect to a TMS of a certain type is introduced according to the following scheme (here * denotes one of the *IR*, *T*, *F*, *TF*):

Δ is a *logical *-consequence* of Γ with respect to a TMS of a type $\mathcal{M}$ (denoted $\Gamma_M \models^* \Delta$), if $\Gamma_M \models^* \Delta$ for all $M \in \mathcal{M}$.

The non-modal properties of the specified relations repeat the corresponding properties of the same-named relations for sets of formulas of the traditional logic of quasiary predicates (see Nikitchenko, Shkilniak, O., & Shkilniak, S., 2016; Shkilniak, O., & Shkilniak, S., 2023).

Let us consider the properties related to equality predicates.

The *symmetry* properties of equality follow from the definitions, because the pairs $\equiv_{xy}$ and $\equiv_{yx}$, $\equiv_{xy}$ and $\equiv_{yx}$ represent same predicates.

The *reflexivity* properties: predicates $\equiv_{xx}$ are constants T, while predicates $\equiv_{xx}$ are irrefutable.

The *transitivity* property for the weak equality is violated (see Shkilniak, S., 2019) for the *T*- and *F*-consequence relations:

$\equiv_{xy}, \equiv_{yz} \not\models_F \equiv_{xy}, \equiv_{yz}, \equiv_{xz}$ and $\neg \equiv_{xy}, \neg \equiv_{yz}, \neg \equiv_{xz} \not\models_T \neg \equiv_{xy}, \neg \equiv_{yz}$.

Indeed, for $d = [x \mapsto a, z \mapsto b]$ we have $d \notin F(\equiv_{xy}) \cup F(\equiv_{yz}) = F(\equiv_{xy} \& \equiv_{yz})$, while $d \in F(\equiv_{xz}) \subseteq F(\equiv_{xy} \& \equiv_{yz} \& \equiv_{xz})$.

Therefore, if the transitivity property for the weak equality is violated for the *T*- and *F*-consequence relations in $GMS^{Q=}$, then it is violated for the stronger *TF*-consequence relation. Hence, in $GMS^{Q=}$ these relations are *incorrect*.

We obtain

Assertion 3. For $GMS^{Q=}$, only the *IR*-consequence relation is non-degenerate and correct.

Properties of the transitivity of equality in $GMS^{Q=}$ and $GMS^{Q=}$ are represented as follows:

$T \models \equiv_{xy}^\alpha, \equiv_{yz}^\alpha, \Gamma_M \models^* \Delta \Leftrightarrow \equiv_{xy}^\alpha, \equiv_{yz}^\alpha, \equiv_{xz}^\alpha, \Gamma_M \models^* \Delta$;

$T \models \neg \equiv_{xy}^\alpha, \neg \equiv_{yz}^\alpha, \Gamma_M \models_{IR} \Delta \Leftrightarrow \neg \equiv_{xy}^\alpha, \neg \equiv_{yz}^\alpha, \neg \equiv_{xz}^\alpha, \Gamma_M \models_{IR} \Delta$.

Properties of substitution of equals are important for logics with equality. For $GMS^{Q=}$ they are specified as follows:

$\text{rpl} = \equiv_{xy}^\alpha, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma_M \models_{IR} \Delta \Leftrightarrow \equiv_{xy}^\alpha, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma_M \models_{IR} \Delta$;

$\text{rpr} = \equiv_{xy}^\alpha, \Gamma_M \models_{IR} R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Delta \Leftrightarrow \equiv_{xy}^\alpha, \Gamma_M \models_{IR} R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Delta$.

At the same time, the substitution of equals in $GMS^{Q=}$ is not straightforward. Let us consider some examples.

Example 1. $\models_{xy} \alpha, \Box \models_{xz} \alpha, \Box \models_{yz} \alpha \mid \models \Box \models_{yz} \alpha$ is true.

However, $\models_{xy} \alpha, \Box \models_{xz} \alpha \mid \models \Box \models_{yz} \alpha$ is refuted from the proof of Theorem 6: in GMS^Q for $d = [x \mapsto c, y \mapsto b, z \mapsto c, s \mapsto a]$ we have $(\models_{xy})_\alpha(d) = T, (\Box \models_{xz})_\alpha(d) = T, (\Box \models_{yz})_\alpha(d) = F$; hence $\models_{xy} \alpha, \Box \models_{xz} \alpha \mid \not\models_{IR} \Box \models_{yz} \alpha$, therefore $\models_{xy} \alpha, \Box \models_{xz} \alpha \mid \not\models \Box \models_{yz} \alpha$. However, $\Box \models_{xz} = R_x^V(\Box \models_{yz})$ and $\Box \models_{yz} = R_y^V(\Box \models_{xz})$, which implies $\models_{xy} \alpha, R_x^V(\Box \models_{yz})^\alpha \mid \not\models R_y^V(\Box \models_{xz})^\alpha$. But from $\models_{xy} \alpha, R_x^V(\Box \models_{yz})^\alpha, R_y^V(\Box \models_{xz})^\alpha \mid \models R_y^V(\Box \models_{yz})^\alpha$ follows that $\models_{xy} \alpha, R_x^V(\Phi)^\alpha, \Gamma \mid \models \Delta$ and $\models_{xy} \alpha, R_x^V(\Phi)^\alpha, R_y^V(\Phi)^\alpha \Gamma \mid \models \Delta$ are not equivalent.

As a result we obtain that the following statements are *not* equivalent ($*$ denotes one of the T, F, TF):

$$\begin{aligned} \models_{xy} \alpha, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma \mid \models \Delta \text{ and } \models_{xy} \alpha, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma \mid \models \Delta; \\ \models_{xy} \alpha, \Gamma \mid \models \Delta, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha \text{ and } \models_{xy} \alpha, \Gamma \mid \models \Delta, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha. \end{aligned}$$

We can see that validity of $\models_{xy} \alpha$ is not enough for the substitution of equals in GMS^Q .

Example 2. $\models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \Box \models_{xz} \alpha \mid \models \Box \models_{yz} \alpha$ is true.

Suppose the opposite: for a GMS $M = (St, R, A, Im)$ and $\alpha \in St, d \in {}^V A$ we have $(\models_{xy})_\alpha(d) = T, Ex^\alpha(d) = T, Ey^\alpha(d) = T, (\Box \models_{xz})_\alpha(d) = T$ and $(\Box \models_{yz})_\alpha(d) = F$; hence $d_\alpha(x) \downarrow = d_\alpha(y) \downarrow = a$ for some $a \in A \Rightarrow d(x) = d(y) = a$. Following $(\Box \models_{yz})_\alpha(d) = F$ we have $(\models_{yz})_\beta(d) = F$ for some $\beta: \alpha \triangleright \beta$.

Condition $(\Box \models_{xz})_\alpha(d) = T$ implies $(\models_{xz})_\beta(d) = T$, therefore two cases are possible:

1) $d_\beta(x) \uparrow$ and $d_\beta(z) \uparrow$; hence $d(x) \notin d_\beta$; from $d(x) = d(y) = a$ we get $d_\beta(y) \uparrow$, therefore $(\models_{yz})_\beta(d) = T$, which contradicts the condition $(\models_{yz})_\beta(d) = F$.

2) $d_\beta(x) \downarrow = d_\beta(z) \downarrow$; from $d(x) = d(y) = a$ we obtain $d_\beta(z) = d_\beta(x) = d_\beta(y) = a$, which contradicts the condition $(\models_{yz})_\beta(d) = F$.

We came to contradiction in both cases.

Therefore, the substitution of equals in GMS^Q requires an addition of conditions of Ex^α and Ey^α validity to the condition of validity of $\models_{xy} \alpha$.

Theorem 8. We obtain the following conditions for the substitution of equals in GMS^Q :

$$\begin{aligned} \models_{rL} \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma \mid \models \Delta &\Leftrightarrow \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma \mid \models \Delta; \\ \models_{rR} \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \Gamma \mid \models R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Delta &\Leftrightarrow \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \Gamma \mid \models R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Delta; \\ \models_{\neg rL} \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \neg R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma \mid \models \Delta &\Leftrightarrow \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \neg R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \neg R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Gamma \mid \models \Delta; \\ \models_{\neg rR} \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \Gamma \mid \models \neg R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Delta &\Leftrightarrow \models_{xy} \alpha, Ex^\alpha, Ey^\alpha, \Gamma \mid \models \neg R_{\bar{w}, \perp, x}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \neg R_{\bar{w}, \perp, y}^{\bar{v}, \bar{u}, z}(\Phi)^\alpha, \Delta. \end{aligned}$$

Properties of carrying modalities over renominations and modality elimination belong to the modal properties of logical consequence relations. For TMS^Q with equality, such properties basically repeat the corresponding properties for TMS^Q without equality (see Shkilniak, 2015).

In case of additional conditions imposed on the transition relation, the elimination properties for modalities will be modified accordingly. The properties for the relation $M \models_{IR}$ and dual properties for the relation $M \not\models_{IR}$ for the cases of transitive, reflexive or symmetric transition relation in GMS^Q can be found in work (Shkilniak, 2015).

The described properties of relations $M \models$ for sets of formulas induce analogous properties of the corresponding logical consequence relations $\models$ for a given type of TMS^Q .

Discussion and conclusions

In this paper we studied pure first-order transitional modal logics of partial non-monotonic quasiary predicates with equality which are program-oriented logical formalisms of the modal type. Two types of such logics were introduced: logics with strong equality predicates and with weak equality predicates. Depending on types of modal compositions and transition relations, classes of general, temporal and multimodal logics can be considered. Semantic models and languages of these logics were described with special attention to the properties of equality predicates. We determined a number of consequence relations in TMS and logical consequence relations on sets of formulas specified with states; the substitution of equals was described in detail.

Properties of these relations will be the semantic basis for constructing the corresponding sequent type calculi.

References

- Abramsky, S., Gabbay, D. M., & Maibaum, T. S. E. (Eds). (1993–2000). *Handbook of Logic in Computer Science* (Vol. 1–5). Oxford University Press.
- Baier, C., & Katoen, J. (2008). *Principles of Model Checking*. MIT Press.
- Björner, D., & Henson, M. C. (Eds). (2008). *Logics of Specification Languages, EATCS Series*. Springer.
- Blackburn, P., van Benthem, J., & Wolter, F. (Eds). (2006). *Handbook of Modal Logic*. Elsevier.
- Cocchiarella, N., & Freund, M. (2008). *Modal Logic. An Introduction to its Syntax and Semantics*. Oxford University Press.
- Fisher, M. (2011). *An Introduction to Practical Formal Methods Using Temporal Logic*. Somerset, NJ: Wiley.
- Fitting, M., & Mendelsohn, R. (2023). *First-Order Modal Logic*. Springer.
- Goranko, V. (2023). *Temporal Logics*. Cambridge University Press.
- Kröger, F. & Merz, S. (2008). *Temporal Logic and State Systems*. Springer-Verlag.
- Nikitichenko, M., Shkilniak, O., & Shkilniak, S. (2016). Pure first-order logics of quasiary predicates. *Problems in Programming*, 2–3, 73–86 [in Ukrainian].
- [Нікітченко, М., Шкілняк, О., & Шкілняк, С. (2016). Чисті першогопорядкові логіки квазіарних предикатів. *Проблеми програмування*, 2–3, 73–86].
- Shkilniak, O. (2015). Modal logics of non-monotone partial predicates. *Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics*, 3, 141–147 [in Ukrainian]. [Шкілняк, О. (2015). Модальні логіки немонотонних часткових предикатів. *Вісник Київського університету. Серія: Фізико-математичні науки*, 3, 141–147].
- Shkilniak, O., & Shkilniak, S. (2023). First-Order Sequent Calculi of Logics of Quasiary Predicates with Extended Renominations and Equality. *Proceedings of UkrPROG'2022, Kyiv, Ukraine. CEUR Workshop Proceedings (CEUR-WS.org)*, 3–18.
- Shkilniak, O., & Shkilniak, S. (2024). Modal Logics of Partial Quasiary Predicates with Equality. *Proceedings of DSMSI-23. Kyiv, Ukraine. CEUR Workshop Proceedings (CEUR-WS.org)*, 35–48.

Shkilniak, S. (2019). First-order composition-nominative logics with predicates of weak equality and of strong equality. *Problems in Programming*, 3, 28–44 [in Ukrainian]. [Шкільняк, С. (2019). Першопорядкові композиційно-номінативні логіки з предикатами слабкої та строгої рівності. *Проблеми програмування*, 3, 28–44].

Отримано редакцією журналу / Received: 05.07.24

Прорецензовано / Revised: 05.02.25

Схвалено до друку / Accepted: 06.03.25

Оксана ШКІЛЬНЯК, канд. фіз.-мат. наук, доц.

ORCID ID: 0000-0003-4139-2525

e-mail: oksana.sh@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

ПЕРШОПОРЯДКОВІ ТРАНЗИЦІЙНІ МОДАЛЬНІ ЛОГІКИ КВАЗІАРНИХ ПРЕДИКАТІВ З РІВНІСТЮ

Досліджено транзиційні модальні логіки (ТМЛ) квазіарних предикатів з рівністю. Модальні логіки успішно використовують для опису й моделювання різноманітних предметних областей. Епістемічні логіки застосовують для опису систем штучного інтелекту, інформаційних та експертних систем. З допомогою темпоральних логік описують динамічні системи, специфікації та верифікації програм. Традиційні модальні логіки базуються на класичній логіці предикатів. Проте ця логіка недостатньо враховує неповноту, частковість, структурованість інформації про предметну область. Такі обмеження ведуть до необхідності побудови нових, програмно-орієнтованих логік, що зумовлює актуальність пропонованої роботи. Ми розглядаємо чисті першопорядкові ТМЛ квазіарних предикатів, не обмежених умовою монотонності та збагачених предикатами рівності. Виділено два різновиди таких логік: $TML^{Q=}$ з предикатами строгої рівності $\equiv_{xy}$ та $TML^{Q=}$ з предикатами слабкої рівності $\approx_{xy}$. Характерною особливістю цих логік є використання композицій розширеної реномінації. Для елімінації кванторів у $TML^{Q=}$ використано спеціальні тотальні предикати-індикатори Ez , а в $TML^{Q=}$ у ролі часткових предикатів-індикаторів постають $\approx_{zz}$. Семантичною базою ТМЛ є транзиційні модальні системи (ТМС). Для $TML^{Q=}$ та $TML^{Q=}$ виділено найважливіші різновиди цих систем: загальні ($GMS^{Q=}$ і $GMS^{Q=}$), темпоральні ($TmMS^{Q=}$ і $TmMS^{Q=}$) та мультимодальні ($MMS^{Q=}$ і $MMS^{Q=}$). Описано семантичні моделі та мови цих різновидів ТМС, особливу увагу приділено властивостям, пов'язаним із предикатами рівності. Розглянуто чотири типи відношень логічного наслідку для множин, специфікованих станами формул у цих ТМС, за неспростовністю (IR), істинністю (T), хибністю (F) та сильний (TF). Описано особливості заміни рівних у $TML^{Q=}$ та $TML^{Q=}$. Властивості відношень логічного наслідку для множин специфікованих станами формул є семантичною основою подальшої побудови для досліджених логік відповідних числень секвенційного типу.

Ключові слова: модальна логіка, транзиційна модальна система, предикат, рівність, логічний наслідок.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

СУЧАСНА ФІЗИКА

УДК 538.9

DOI: <https://doi.org/10.17721/1812-5409.2025/1.24>

Ярослав ЧОЛІЙ, начальник сектору

e-mail: redox.sparrowhawk@gmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

Микола МАКАРЕЦЬ, д-р фіз.-мат. наук, проф.

ORCID ID: 0009-0005-8329-4079

e-mail: mmv_phys@knu.ua

Київський національний університет імені Тараса Шевченка, Київ, Україна

ВПЛИВ МОДЕЛІ ПРУЖНИХ ІОН-АТОМНИХ ЗІТКНЕНЬ НА КУМУЛЯНТИ ПРОСТОРОВОГО РОЗПОДІЛУ ІМПЛАНТОВАНИХ ІОНІВ

Розраховано кумулянти функції просторового розподілу іонів, імплантованих у тверде тіло, у моделі, де парні пружні іон-атомні зіткнення відбуваються на відстані, не меншій за міжатомну. Їхні значення знайдено для кумулянтів від першого по шостий порядок для кількох сотень пар іон – мішень і встановлено, що така розділеність пружних зіткнень змінює їхню величину порівняно зі значеннями, отриманими в моделі неперервних зіткнень. Відхилення можуть досягати десятків відсотків залежно від порядку кумулянта та відношення мас іона і атома мішені. Вони максимальні за енергії іонів порядку кількох кеВ, монотонно спадають з її зростанням приблизно до 100–300 кеВ, а за вищої енергії – перебувають у межах 1–2 %. Це зумовлено малим шляхом іона між послідовними пружними зіткненнями порівняно зі страгглінгами (розкид розподілу навколо максимуму) у поздовжньому і поперечному напрямі, що зростають зі збільшенням енергії іона.

Ключові слова: швидкі іони, тверде тіло, модель зіткнень, кумулянти розподілу.

Класифікація відповідно до AMS 2020: 81V45.

Вступ

Рівняння для функції просторового розподілу швидких іонів, імплантованих в аморфну мішень, було вперше отримано в межах теорії Ліндхарда – Шарфа – Шіотта (ЛШШ) (Lindhard, Scharf, & Schiott, 1963). Автори вважали, що іон втрачає енергію неперервно на збудження електронної підсистеми, та у парних пружних зіткненнях з екранованими атомними ядрами, які відбуваються також неперервно. Останнє припущення неможливо реалізувати в разі чисельного моделювання траєкторій швидких іонів у твердому тілі (напр., Ziegler, Biersack, & Littmark, 1985; Ziegler, 2024), тому тут парні пружні іон-атомні зіткнення відділяють одне від іншого, розігруючи випадкову довжину пробігу іонів між зіткненнями.

Щоб позбутися припущення про неперервні пружні зіткнення у теоретичному описі, автори (Bowyer, Ashworth, & Oven, 1992; Oven, Bowyer, & Ashworth, 1993; Ashworth, Bowyer, & Oven, 1995) розвинули модель, у якій пробіг іона між зіткненнями скінченний, випадковий і залежить від атомної структури мішені. Отримані інтегральні рівняння для моментів розподілу автори звели до диференціальних рівнянь першого порядку і розв'язали їх чисельно, урахувавши на кожній ітерації невідомі похідні моментів за початкової енергії іона. Моменти просторового розподілу вдалося дослідити у такий спосіб до енергії іонів меншої 1 МеВ, доки запропонований метод залишався стійким. Встановлено, що атомна будова мішені (газо-, рідиноподібна) може змінювати значення вищих моментів розподілу у кілька разів в області енергії до 100 кеВ. У роботі (Макарець, & Чолій, 2013) показано, що згадані інтегральні рівняння переходять у рівняння ЛШШ, якщо довжину вільного пробігу іона між двома послідовними пружними зіткненнями спрямувати до нуля.

У дослідженнях (Макарець, & Сторчака, 2001; Ілюїна, & Макарець, 2004; 2010), у межах теорії ЛШШ, отримали інтегральні рівняння для кумулянтів функції просторового розподілу імплантованих іонів і втрат енергії, зв'язок яких із моментами розподілу відомий. Інтегральні рівняння розв'язувались чисельно у широкому діапазоні нерелятивістських енергій іонів. Зауважимо, що в (Макарець, & Сторчака, 2001; Ілюїна, & Макарець, 2004; 2010) замість кумулянтів третього і вищих порядків, які швидко зростають і є енергією іонів, проаналізовано безрозмірні величини – так звані скісності, ексцеси першого, другого і вищих порядків. Вони задають відхилення розподілу від нормального і є відношеннями кумулянтів вищого порядку до кумулянтів другого порядку у відповідному степені. У роботі (Макарець, & Чолій, 2013) у межах моделі (Bowyer, Ashworth, & Oven, 1992; Oven, Bowyer, & Ashworth, 1993; Ashworth, Bowyer, & Oven, 1995) були отримані рівняння для кумулянтів функції просторового розподілу імплантованих іонів.

Метою пропонованої роботи є розрахунок кумулянтів просторового розподілу іонів, імплантованих у тверде тіло, у рамках моделі, де парні пружні іон-атомні зіткнення відокремлені одне від одного. Друга мета – дослідити зміну кумулянтів, викликану цим наближенням, порівняно з результатами теорії ЛШШ.

1. Рівняння і чисельний метод

Нехай іон з енергією E стартує з початку координат уздовж осі OZ усередині ізотропного твердого тіла. Імовірність його зупинки в околі $d\vec{r}$ навколо точки $\vec{r}$ позначимо як $\Pi(\vec{r}, E)$. У роботах (Bowyer, Ashworth, & Oven, 1992; Макарець, & Чолій, 2013) показано, що ця функція задовольняє рівняння:

$$\begin{aligned} \Pi(\vec{r}, E) = & \int_0^{\infty} \int_{T_{\min}}^{T_{\max}} \int_0^{2\pi} \Pi(\vec{r}', E') \Theta(E' - E_s) f(l) \frac{d\phi d\sigma_n(E - T_e, T)}{2\pi \sigma_{tot}} dl + \\ & + \int_0^{\infty} \int_{T_{\min}}^{T_{\max}} \delta(\vec{r}' - \vec{k}l) \Theta(E_s - E') f(l) \frac{d\sigma_n(E - T_e, T)}{\sigma_{tot}} dl, \end{aligned} \quad (1)$$

де $\vec{r}' = A(\vec{r} - \vec{k}l)$ – точка зупинки іона в новій системі координат, що зміщена вздовж старої осі OZ на випадковий шлях вільного пробігу l до пружного зіткнення і повернута на випадкові кути ϕ навколо OZ та пружного розсіяння

© Чолій Ярослав, Макарець Микола, 2025

θ навколо нової осі OY' , $A = A(\varphi, \theta)$ – матриця цих поворотів, $E' = E - T_e - T_n$ – енергія іона після пружного зіткнення із втратою енергії T_n , $T_e \equiv T_e(l)$ – втрата енергії на неперервне збудження електронної підсистеми вздовж l , $\Theta(x)$, – функція Хевісайда, E_s – найменша енергія іона, нижче якої іон вважається нерухомим (у роботах (Bowyer, Ashworth, & Oven, 1992; Oven, Bowyer, & Ashworth, 1993; Ashworth, Bowyer, & Oven, 1995) вона дорівнювала 5 eV), $f(l)dl$ – імовірність пружного розсіювання іона на атомі мішені в околі dl після вільного пробігу довжиною l , $d\sigma_n = d\sigma_n(E - T_e, T_n)$ – диференціальний переріз пружного розсіювання іона з енергією $E - T_e$ на атомі мішені у разі втрати енергії dT_n навколо T_n , σ_{tot} – так званий повний переріз пружного розсіювання в моделі парних пружних зіткнень у мішені з концентрацією атомів n_0 , T_{min} – мінімальна енергія, яку може втратити іон у парному пружному зіткненні з атомом мішені (саме її існування робить "повний" переріз скінченим), $T_{max} = \gamma(E - T_e)$ – максимальна енергія, яку може втратити іон у парному пружному лобовому зіткненні з атомом мішені, $\gamma = 4\mu/(1+\mu)^2$, $\mu = M_1/M_2$ – відношення атомних мас іона M_1 й атома мішені M_2 , $\delta(x)$ – дельта-функція Дірака.

В (1) уведено функцію розподілу довжини вільного пробігу іона $f(l)$, що визначається атомною структурою мішені. Наприклад, згідно з роботами (Bowyer, Ashworth, & Oven, 1992; Oven, Bowyer, & Ashworth, 1993; Ashworth, Bowyer, & Oven, 1995), у моделях рідиноподібного $f(l) = \delta(l - \Lambda)$ і газоподібного середовищ $f(l) = \exp(-l/\Lambda)/\Lambda$, де Λ – точна, або середня відстань між пружними зіткненнями, відповідно. Її моменти k -го порядку (середнє значення – $k = 1$, квадратичне відхилення – $k = 2$ і т. д.) пропорційні Λ^k .

У наступних перетвореннях (1) буде використано так званий сумарний переріз усіх пружних розсіювань іона із енергією $E - T_e$, у яких він втратив енергію від якоїсь заданої нами T_m до максимальної T_{max} . Цей переріз був уведений у роботі (Макарець, & Сторчака, 2001) як

$$\sigma_n(E - T_e, T_m) = - \int_{T_m}^{\gamma(E - T_e)} d\sigma_n(E - T_e, T'), \quad (2)$$

де враховано, що $d\sigma_n(E - T_e, T')$ спадає зі зростанням T' . Із (2) випливає, що сумарний переріз усіх пружних розсіювань іона спадає до нуля, якщо мінімальна втрачена енергія $T_m \rightarrow T_{max} = \gamma(E - T_e)$, а коли $T_m \rightarrow 0$, то він стає повним і прямує до нескінченності. Модель парних пружних зіткнень у мішені з концентрацією атомів n_0 справедлива, коли прицільна відстань зіткнення менша половини міжатомної відстані $p_{max} \leq a/2 \equiv 1/2n_0^{1/3}$, що дає верхню межу для сумарного перерізу парних пружних зіткнень $\sigma_n(E, T_m) \approx a^2$.

У моделі рідиноподібної мішені з концентрацією n_0 довжина вільного пробігу Λ визначається з умови $\Lambda\sigma_{tot} = 1/n_0$, де σ_{tot} – повний переріз пружного розсіювання кульок. Вільний пробіг можливий, якщо він більший міжатомної відстані, тобто якщо $\Lambda \geq a = 1/n_0^{1/3}$, звідки випливає, що $\sigma_{tot} \leq a^2$. Ми використали сумарний переріз $\sigma_n(E - T_e, T_m)$, щоб з умови

$$\Lambda\sigma_n(E - T_e, T_m) = 1/n_0 \quad (3)$$

знайти $T_m(\Lambda)$. Переріз розсіювання спадає зі збільшенням енергії іона, тому із (3) можна знайти найменшу енергію іонів у кожному зіткненні, вище якої вони будуть парними і рівняння (1) не матиме внутрішньої суперечності. Її знаходили при чисельному розв'язку рівнянь (6), які отримані далі, і вона виявилася значно більшою за уведену вище енергію зупинки іона T_s . Звідси випливає, що за малих енергій іонів наближення парних зіткнень вимагає корекції. Для цього ми використали так звану muffin tin модель.

Рівняння (2, 3) враховують вільний пробіг Λ між пружними зіткненнями, тому енергія наступного пружного зіткнення зменшується на роботу $\Lambda S_e(E)$, де $S_e(E)$ – енергія, втрачена іоном на збудження електронної підсистеми мішені на одиниці шляху, так зване електронне гальмування. Цим вони відрізняються від рівнянь моделі неперервних пружних зіткнень, але граничний перехід $\Lambda \rightarrow 0$ переводить їх у рівняння ЛШШ (Ilyina, & Makarets, 2004; 2010). Зважаючи на це, зміна кумулянтів розподілу в цій моделі розділених пружних зіткнень порівняно з моделлю ЛШШ буде зумовлена внеском вільного пробігу Λ й електронного гальмуванням $S_e(E)$.

Перший доданок у правій частині (1) відмінний від нуля для всіх іонів з енергією $E - T_e > E_s$, тобто для тих, що мають більш ніж одне пружне зіткнення. А другий – у разі опромінення іонами низької енергії, коли $E - T_e \leq E_s$ і вони зупиняються не далі, ніж після першого зіткнення. Тоді їх просторовий розподіл буде дельта-функцією Дірака із пробігом l . Проте вже за $E > E_s + T_e(l) + T_n$ другий доданок в (1) дорівнюватиме нулю. Рівняння для кумулянтів розподілу іонів з енергією, що значно перевищують цю межу, отримані в (Макарець, & Чолій, 2013).

Розподіл $\Pi(\vec{r}, E)$ має осьову симетрію відносно OZ , тому його можна подати такими виразами:

$$\Pi(\vec{r}, E) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \exp(-i\vec{\omega} \cdot \vec{r} + \Phi(\vec{\omega}, E)) d\vec{\omega}, \quad (4)$$

де $\Phi(\vec{\omega}, E)$ – друга характеристична функція розподілу, що виражається через кумулянти $\kappa_{i,j}(E)$:

$$\Phi(\vec{\omega}, E) = \sum_{N=1, n=0}^{\infty, [N/2]} i^N \kappa_{2n, N-2n}(E) \frac{(u^2 + v^2)^n \omega^{N-2n}}{(2n)!(N-2n)!}, \quad (5)$$

де $\vec{\omega} = (u, v, \omega)$ – вектор Фур'є-простору, $N = i + j$ – порядок кумулянта $\kappa_{i,j}$, а $[x]$ – ціла частина x .

Ці вирази підставили в (1), розклали обидві частини в ряди Тейлора за $(u^2 + v^2)$ та за ω , проінтегрували їх за кутом φ і прирівняли відповідні коефіцієнти рядів. Унаслідок цього отримали системи лінійних інтегральних неоднорідних рівнянь для кумулянтів $\kappa_{i,j}(E)$ однакового порядку N . У їхні праві частини входять кумулянти порядку $N-1, N-2, \dots, 1$, які потрібно знайти, починаючи з $N=1$. Кумулянт $\kappa_{0,0}(E) \equiv 1$, тому що це є нормуванням функції розподілу.

Отже, у випадку осьової симетрії функція розподілу іонів, імплантованих у тверде тіло, звелася до системи лінійних неоднорідних інтегральних рівнянь. Ураховуючи лінійність рівнянь, ми замінили кумулянти на їхні лінійні комбінації, які назвали ерзакумулянтами і позначили як $\chi_{i,j}(E)$ (Макарець, & Чолій, 2013). Це діагоналізує ліві частини рівнянь і надає для ерзакумулянтів незв'язані між собою лінійні неоднорідні інтегральні рівняння.

У цих рівняннях, окрім інтегралу за $d\sigma_n(E - T_e, T_n)$, є інтеграл за dl , який легко обчислити в моделі рідиноподібної мішені. Після переходу до безрозмірних довжин $z = Z/R^*$, $\rho = R/R^*$, $\lambda = \Lambda/R^*$ та енергій $\varepsilon = E/E^*$, $\tau_e = T_e/E^*$, $\tau = T_n/E^*$, $\tau_{\min} = T_{\min}/E^*$, де R^* і E^* – масштаби довжини і енергії (Ilyina, & Makarets, 2004; 2010) та громіздких викладок, було отримано ланцюжок лінійних неоднорідних інтегральних рівнянь:

$$\begin{aligned} \frac{\chi_{01} - \chi'_{01}}{\lambda} - \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{01}'' P_1\} d\tau &= 1, \\ \frac{\chi_{02} - \chi'_{02}}{\lambda} - \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{02}'' P_2\} d\tau &= \frac{\tilde{\chi}_{01}'^2 - \chi_{01}^2}{\lambda} + \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{01}''^2 P_2 + 2\chi_{01}'' \lambda P_1\} d\tau, \\ \frac{\chi_{20} - \chi'_{20}}{\lambda} - \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{20}'' P_0\} d\tau &= \frac{\tilde{\chi}_{01}'^2 - \chi_{01}^2}{\lambda} + \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{01}''^2 P_0 + 2\chi_{01}'' \lambda P_1\} d\tau, \\ \frac{\chi_{03} - \chi'_{03}}{\lambda} - \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{03}'' P_3\} d\tau &= \frac{\tilde{\chi}_{01}'^3 - \chi_{01}^3 + 3\tilde{\chi}_{01}' \chi_{02}' - 3\chi_{02} \chi_{01}}{\lambda} + \\ &+ \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{01}'' (\chi_{01}''^2 + 3\chi_{02}'') P_3 + 3(\chi_{01}''^2 + \chi_{02}'') \lambda P_2 + 3\chi_{01}'' \lambda^2 P_1\} d\tau, \\ \frac{\chi_{21} - \chi'_{21}}{\lambda} - \frac{1}{\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{\chi_{21}'' P_1\} d\tau &= \frac{3(\tilde{\chi}_{01}'^3 - \chi_{01}^3) + 4(\tilde{\chi}_{01}' \chi_{02}' - \chi_{01} \chi_{02}) + 5(\tilde{\chi}_{01}' \chi_{20}' - \chi_{01} \chi_{20})}{3\lambda} + \\ &+ \frac{1}{3\gamma} \int_0^{\gamma'} \tilde{\sigma}_n(\varepsilon - \tau_e, \tau) \frac{\partial}{\partial \tau} \{(3\chi_{01}''^3 + 9\chi_{01}'' \lambda^2 + 4\chi_{01}'' \chi_{02}'' + 5\chi_{01}'' \chi_{20}'') P_1 + 4(\chi_{01}''^2 + \chi_{02}'') \lambda P_2 + 5(\chi_{20}'' + \chi_{01}''^2) \lambda\} d\tau, \end{aligned} \quad (6)$$

де $\chi'_{i,j} = \chi_{i,j}(\varepsilon - \tau_e)$, $\chi''_{i,j} = \chi_{i,j}(\varepsilon - \tau_e - \tau)$, $\tilde{\chi}'_{0,1} = \chi_{0,1}(\varepsilon - \tau_e) - \lambda$, $P_j \equiv P_j(\cos(\theta))$ – поліном Лежандра j -го порядку, а кут розсіяння $\theta \equiv \theta(\tau)$ залежить від енергії τ , втраченої у пружному зіткненні, $\tilde{\sigma}_n(\varepsilon - \tau_e, \tau)$ – сумарний переріз пружного розсіяння іона виражений в одиницях квадрата радіуса екранування у потенціалі Бірсака – Циглера (див. Ziegler, Biersack, & Littmark, 1985; Ziegler, 2024).

Рівняння для кумулянтів з $N > 3$ занадто громіздкі і не наведені, але чисельні розв'язки були знайдені для ланцюжка із 15-ти рівнянь для кумулянтів до шостого порядку включно. Важливо те, що праві частини рівнянь для ерзакумулянтів порядку $N \geq 1$ залежать від ерзакумулянтів лише меншого порядку. Ці рівняння потрібно розв'язувати послідовно: спочатку для 1-го порядку, потім – 2-го і т. д.

Зауважимо, що у рівняннях (6), якщо $\tau \rightarrow 0$, усі підінтегральні функції спадають до нуля швидше, ніж сумарний переріз $\tilde{\sigma}_n(\varepsilon - \tau_e, \tau)$ зростає до нескінченності, тому нижню межу в інтегралах за τ покладено рівною нулю замість $\tau_{\min}$, як в (1). Повний переріз $\tilde{\sigma}_{tot}$, що входить в (1) і залишається в (6), шукали згідно з рівнянням (3).

Втрати енергії іона на збудження електронного газу на шляху Λ знаходили в лінійному наближенні $T_e = \Lambda S_e(E)$, де $S_e(E)$ – електронне гальмування. Граничний перехід $\lambda \rightarrow 0$ у (6), у нульовому порядку за λ , переводить їх у

рівняння, отримані у роботах (Ilyina, & Makarets, 2004; 2010). Точніші розрахунки T_e з виразу $T_e(E) = \Lambda S_e(E - T_e(E))$ надають малу поправку до цього результату. Тому розв'язки (6) відрізняються від отриманих у роботах (Ilyina, & Makarets, 2004; 2010) у першому порядку за степенями електронного гальмування за $T_e \ll E$. Ця умова виконується, коли довжина вільного пробігу іона Λ мала, порівняно зі страгглінгом проєктивного пробігу, існування якого зумовлено пружними зіткненнями і який задає розкид розподілу зупинених іонів навколо максимуму.

Для розв'язку (6) було використано чисельний метод, розвинений у роботах (Ilyina, & Makarets, 2004; 2010), у якому інтервал енергії іонів від $T_{\min}$ до 1 GeV ділили на частини, а кожне інтегральне рівняння зводили до системи неоднорідних лінійних алгебраїчних рівнянь для восьми невідомих коефіцієнтів розкладу $\chi_{i,j}(E)$ за степенями E .

2. Результати й обговорення

Чисельні розрахунки ерзацкумулянтів $\chi_{i,j}(E)$ були проведені для кількох сотень пар іон-мішень із відношенням маси іона до маси атома мішені $\mu \in (0.04 \div 25)$, які зустрічаються під час імплантації в метали та напівпровідники. Для коректного порівняння результатів відділених пружних зіткнень, отриманих у цій моделі, із результатами, одержаними в моделі неперервних зіткнень (Ilyina, & Makarets, 2004; 2010), використали потенціал Бірсака – Циглера для іон-атомної взаємодії та електронне гальмування із (Ziegler, Biersack, & Littmark, 1985; Ziegler, 2024). Знайдені залежності $\chi_{i,j}(E)$ в діапазоні енергії іонів від 1 кеВ до 1 GeV, виявилися близькими до знайдених у роботі (Ilyina, & Makarets, 2004; 2010) у межах теорії ЛШШ, а їхні відхилення – залежними від енергії та відношення маси іона до маси атома мішені. Потім знайшли кумулянти $\kappa_{i,j}(E)$, через які виражається функція просторового розподілу імплантованих іонів

$$\Pi(\vec{r}, E) = \frac{1}{2\Delta R_p (\pi\Delta R_\perp)^2} \int_0^\infty \int_0^\infty s \exp \left[-\frac{s^2 + t^2}{2!} + \frac{1}{4!} (t^4 Ex_\parallel + 6s^2 t^2 \beta_{2,2} + s^4 Ex_\perp) - \dots \right] \times \\ \times \cos \left[tZ - \frac{1}{3!} (t^3 Sk + t s^2 \beta_{2,1}) + \dots \right] J_0(sR_\perp) dt ds, \quad (7)$$

де $R_p \equiv \kappa_{0,1}$, $\Delta R_p \equiv \sqrt{\kappa_{0,2}}$, $\Delta R_\perp \equiv \sqrt{\kappa_{2,0}}$ – проєктивний пробіг, проєктивний і поперечний страгглінги іона – величини розмірності довжини, та безрозмірні $Z = (R_p - z)/\Delta R_p$ і $R_\perp = \sqrt{x^2 + y^2}/\Delta R_\perp$ – поздовжня, центрована на проєктивний пробіг, і поперечна радіальна координати точки зупинки іона, нормовані на відповідні страгглінги, $Sk \equiv \kappa_{0,3}/(\Delta R_p)^3$ – поздовжня скісність, $Ex_\parallel \equiv \kappa_{0,4}/(\Delta R_p)^4$ і $Ex_\perp \equiv \kappa_{4,0}/(\Delta R_\perp)^4$ – поздовжній і поперечний ексцеси, $\beta_{2,1} \equiv \kappa_{2,1}/(\Delta R_\perp)^2 \Delta R_p$, $\beta_{2,2} \equiv \kappa_{2,2}/(\Delta R_p \Delta R_\perp)^2$ – характеристики кореляції розподілу у поздовжньому і поперечному напрямках, $J_0(x)$ – функція Бесселя 1-го порядку нуля.

Якщо в (7) продовжити ряди в косинусі й експоненті, то з'являться доданки з кумулянтами вищого порядку: $Sk_n \equiv \kappa_{0,(2n+1)}/(\Delta R_p)^{2n+1}$, $Ex_{\parallel,n} \equiv \kappa_{0,(2n+2)}/(\Delta R_p)^{2n+2}$, $Ex_{\perp,n} \equiv \kappa_{(2n+2),0}/(\Delta R_\perp)^{2n+2}$ – поздовжні скісності, поздовжній і поперечний ексцеси порядку $n = 2, 3, \dots$, та з коефіцієнтами $\beta_{i,j} \equiv \kappa_{i,j}/(\Delta R_\perp)^i (\Delta R_p)^j$, де $i, j \neq 0$, що описують кореляцію розподілу у поперечному і поздовжньому напрямках. Вплив кожного з кумулянтів на форму розподілу було досліджено в роботі (Ilyina, & Makarets, 2004).

Розрахунки показали, якщо енергія іона становить 1 кеВ, то відношення вільного пробігу іона до страгглінгу проєктивного пробігу $\Lambda/\Delta R_p$ змінюється від 0.5 до 0.05 для зазначених вище відношень мас іона й атома мішені μ . Зі збільшенням енергії воно спадає і за $E > 1$ МеВ стає меншим на порядок. Відношення T_e/E приблизно на порядок менше за $\Lambda/\Delta R_p$ і має аналогічну поведінку. Тому вплив вільного пробігу іона Λ й електронного гальмування $S_e(T_e)$ на цьому шляху на кумулянти розподілу проявляється за енергії іонів, меншій від кількох сотень кеВ і залежить від μ .

Перше рівняння в (6) описує проєктивний пробіг іона $R_p(E)$, його права частина стала, тому його розв'язок показує вплив лише Λ і $S_e(T_e)$ на ліву частину рівняння. Оскільки ліві частини (6) відрізняються лише порядком полінома Лежандра, то це показує оцінку впливу Λ і $S_e(T_e)$ на ліві частини всіх рівнянь ланцюжка. Розрахунки виявили, що за енергій іонів від кількох сотень кеВ до 1 GeВ проєктивні пробіги перевищують результати роботи (Ilyina, & Makarets, 2004) менше ніж на 1 % для всіх комбінацій іон-мішень. Ця величина зберігається і за зменшення енергії іонів до 1 кеВ, якщо $\mu \gg 1$, або $\mu \ll 1$, а для $\mu \sim 1$, наприклад, для пари іон-мішень $\text{Ni}^+ \rightarrow \text{Fe}$, де $\mu \approx 1$, відхилення досягає 5 %. Отже, зміна лівих частин ланцюжка рівнянь (6) мала.

Розв'язки ланцюжка рівнянь (6) для страгглінгів $\Delta R_p(E)$ і $\Delta R_\perp(E)$ поведуться аналогічно до проєктивного пробігу, але відхилення зростають. Із рис. 1 видно, що поздовжній страгглінг у межах ≈ 1 % збігається з результатами (Ilyina, & Makarets, 2004) для комбінацій іон-мішень з $\mu < 1$, але для іонів, важчих за атом мішені, як у комбінаціях $\text{Nb}^+ \rightarrow \text{Fe}$ із $\mu \approx 1.7$ та $\text{Er}^+ \rightarrow \text{Si}$ із $\mu \approx 6$ він стає меншим приблизно на 4 % і на 9 % за енергії в кілька кеВ. Розрахунки показали, що відхилення поперечних страгглінгів мають різний знак, не перевищують 4 % і можуть змінювати свій знак з енергією іона.

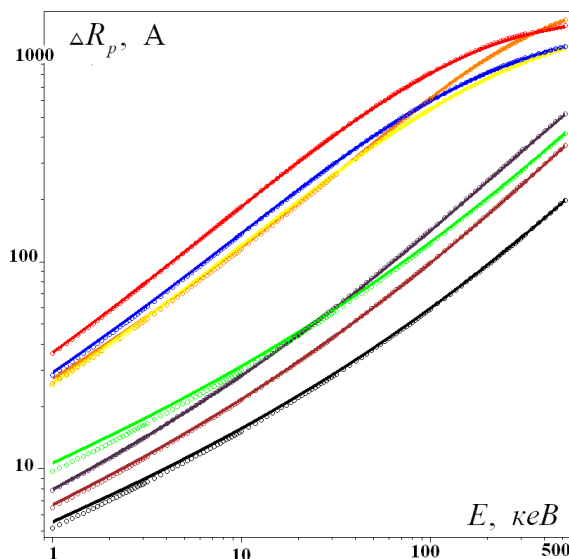


Рис. 1. Поздовжній страгглінг $\Delta R_p(E)$. Лінії – розрахунок згідно з роботами (Ilyina, & Makarets, 2004; 2010), кружечки – розрахунок із цих робіт у моделі відокремлених парних зіткнень.

Лінії та кружечки для різних комбінацій іон-мішень мають такий колір і відношення мас іона й атома мішені μ :

B to Si – червоний, 0.38; Er to Si – зелений, 5.96; C to Ti – синій, 0.25; N to Ti – помаранчевий, 0.29; Nb to Fe – коричневий, 1.66; Ni to Fe – фіолетовий, 1.05; C to W – жовтий, 0.065; W to Cu – чорний, 2.89

У роботі (Ilyina, & Makarets, 2004) було встановлено, що скісності, ексцеси та коефіцієнти $\beta_{i,j}(E)$ знакозмінні, а їхні значення малі порівняно з одиницею за енергії іонів меншій від кількох сотень кеВ. З огляду на це вільний пробіг іона Λ може вплинути на поведінку цих кумулянтів за низької енергії, зокрема на положення їхніх нулів.

На рис. 2 наведено графіки поперечного ексцесу, розраховані згідно з теорією ЛШШ (Ilyina, & Makarets, 2004), та для рідиноподібної моделі (6) для тих самих комбінацій іон-мішень що і на рис. 1. Видно, що для низьких енергій результати в моделі відокремлених пружних зіткнень на десятки відсотків менші, ніж у теорії ЛШШ для всіх комбінацій іон-мішень, а зі збільшенням енергії їхня різниця зменшується для всіх μ . Причиною їхнього зменшення за низьких енергій іонів є вплив довжини вільного пробігу Λ і втрат енергії на електронне гальмування T_e , що роблять функцію розподілу більш пологою біля максимуму. Зі зростанням енергії іонів цей вплив зменшується, оскільки вони стають малими порівняно зі страгглінгом та енергією іона.

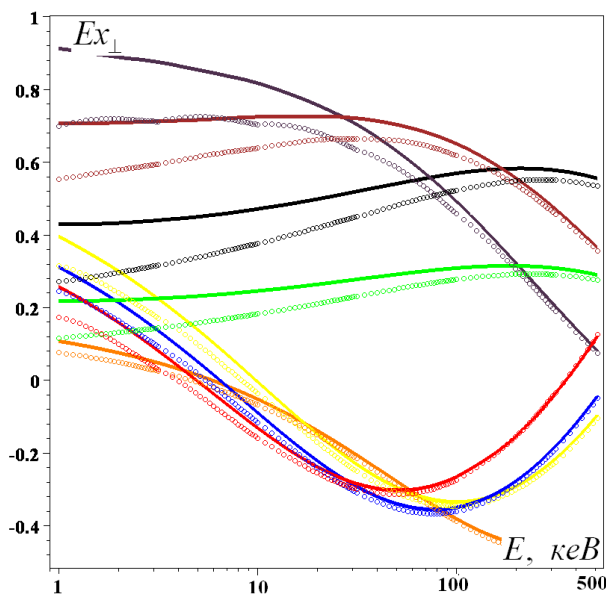


Рис. 2. Поперечний ексцес для тих самих комбінацій іон-мішень, що і на рис. 1. Лінії та кружечки для різних комбінацій іон-мішень мають такий колір і відношення мас іона й атома мішені μ :

B to Si – червоний, 0.38; Er to Si – зелений, 5.96; C to Ti – синій, 0.25; N to Ti – помаранчевий, 0.29; Nb to Fe – коричневий, 1.66; Ni to Fe – фіолетовий, 1.05; C to W – жовтий, 0.065; W to Cu – чорний, 2.89

Інші кумулянти четвертого порядку – поздовжній ексцеси та коефіцієнт $\beta_{2,2}(E)$ мають аналогічну поведінку для тих самих комбінацій іон-мішень. А кумулянти п'ятого і шостого порядку відхиляються від результатів роботи (Ilyina, & Makarets, 2004) в обидві сторони до 100 % за енергії до кількох сотень кеВ. Розрахунки показали, що зі збільшенням порядку кумулянтів зазначені відхилення за низьких енергій зростають.

Аналіз результатів розрахунку п'ятнадцяти кумулянтів розподілу до шостого порядку включно для кількох сотень комбінацій іон-мішень показав: якщо енергія вища приблизно за 500 кеВ, то вони збігаються з результатами роботи (Ilyina, & Makarets, 2004), де не враховано довжини вільного пробігу Λ . Причиною такої поведінки кумулянтів є малі значення врахованих поправок у механізмі гальмування іонів: скінченний шлях між пружними зіткненнями $\Lambda \ll \Delta R_p(E)$, втрати енергії $T_e(\Lambda) \ll E$. З огляду на це вони впливають на кумулянти розподілу лише в області низьких енергій, причому зі зростанням порядку кумулянта верхня межа цієї області теж зростає. Це викликано тим, що довжина вільного пробігу $\Lambda > a$ у цій моделі зменшує середні втрати енергії на шляху, що веде до збільшення проєктивного пробігу іона.

Дискусія і висновки

Розрахунки кумулянтів просторового розподілу іонів, імплантованих в ізотропну мішень, показали, що скінчений шлях іона між двома послідовними пружними зіткненнями з атомами мішені Λ змінює їхні значення порівняно з моделлю неперервного гальмування. Зокрема, для енергії іонів до кількох сотень кеВ кумулянти просторового розподілу в межах рідиноподібної моделі можуть відхилятися від результатів теорії ЛШШ в обидві сторони, залежно від відношення мас іона й атома мішені μ .

Для проєктивного пробігу і страгглінгів поздовжнього та поперечного пробігу (кумулянти 1-, 2-го порядків) ці зміни можуть досягати 5 % за енергій іонів у кілька кеВ і спадають до величини меншої 1 % зі зростанням енергії до кількох сотень кеВ. Вони залежать від відношення мас іона й атома мішені.

Кумулянти вищого порядку (скісності, ексцеси, коефіцієнти $\beta_{i,j}$) для цих енергій не перевищують одиниці і мало впливають на форму розподілу. Оскільки поздовжній і поперечний страгглінги вимірюються в експерименті, то їхнє порівняння з розрахунками дає змогу виявити вплив на них довжини вільного пробігу іона між пружними зіткненнями.

Якщо енергія іонів більша за кілька сотень кеВ, то всі кумулянти розподілу збігаються з результатами розрахунків у межах теорії ЛШШ, у якій довжина вільного пробігу нескінченно мала. Причиною цього є мале значення відношення Λ до страгглінгів пробігу іонів і це відношення спадає зі зростанням їхньої енергії.

Досліджений вплив довжини вільного пробігу іона між пружними зіткненнями на кумулянти розподілу імплантованих іонів притаманний всім ізотропним мішеням з випадково розміщеними атомами, де швидкі іони втрачають енергію в бінарних пружних зіткненнях з атомами та у неперервному електронному гальмуванні між ними.

У роботі використано рідинну модель мішені, у якій усі відстані між атомами однакові і дорівнюють a , а ймовірність пружного розсіяння іона на атомі мішені після вільного пробігу l задається дельта-функцією Дірака $f(l) = \delta(x - \Lambda)$ і має лише один параметр Λ . Для складнішої моделі середовища $f(l)$ залежатиме від кількох параметрів розподілу атомів у мішені, які матимуть порядок міжтомної відстані. Тоді рівняння аналогічні (6) будуть залежати від них. А тому їхній внесок у кумулянти розподілу буде близьким за величиною до внеску величини Λ . Звідси випливає, що встановлена вище поведінка кумулянтів просторового розподілу іонів, імплантованих у рідиноподібне середовище, характерна для всіх моделей середовища з випадково розміщеними атомами, з якими іон під час руху взаємодіє парно.

Внесок авторів: Ярослав Чолій – налагодження чисельного методу, розрахунки, обговорення, висновки; Микола Макарець – постановка задачі, огляд літературних джерел, аналітичні викладки, обговорення, формулювання висновків.

Подяки, джерела фінансування. Дослідження виконано за підтримки гранту Міністерства освіти і науки України для перспективного розвитку наукового напрямку "Математичні науки та природничі науки" в Київському національному університеті імені Тараса Шевченка.

Список використаних джерел

- Макарець, М. В., & Сторчака, С. Н. (2001). Новий метод розрахунку розподілу імплантованих іонів. 1. Алгоритм та семиінваріанти. *Український фізичний журнал*, 46(4), 486–494.
- Макарець, М. В., & Чолій, Я. В. (2013). Рівняння для кумулянтів просторового розподілу імплантованих іонів. *Вісник Київського університету. Серія: Фізико-математичні науки*, 4, 255–259.
- Ashworth, D. G., Bowyer, M. D. J., & Oven, R. (1995). The derivation and moments solution of approximate transport equations for the implantation of ions into amorphous targets. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 100B(4), 471–482. [https://doi.org/10.1016/0168-583X\(95\)00347-9](https://doi.org/10.1016/0168-583X(95)00347-9)
- Bowyer, M. D. J., Ashworth, D. G., & Oven, R. (1992). Central moments of ion implantation distributions derived by the backward Boltzmann transport equation compared with Monte Carlo simulations. *Journal of Physics D: Applied Physics*, 25(11), 1619. <https://doi.org/10.1088/0022-3727/25/11/009>
- Ilyina, V. V., & Makarets, M. V. (2004). The cumulants analysis of the implanted ions distribution. *Ukrainian journal of Physics*, 49(8), 815–823.
- Ilyina, V. V., & Makarets, M. V. (2010). Distribution of energy losses by fast ions along their propagation paths in solids. *Ukrainian journal of Physics*, 55(2), 235–242.
- Lindhard, J., Scharf M., & Schiott H. E. (1963). Range concepts and heavy ion ranges. *Matematisk-fysiske Meddelelser udgivet af Det Kongelige Danske Videnskabernes Selskab*, 33(14), 1–43.
- Oven, R., Bowyer, M. D. J., & Ashworth, D. G. (1993). A transport equation model that is consistent with Monte Carlo simulations of ion implantation into infinite targets. *Journal of Physics: Condensed Matter*, 5(14), 2157. <https://doi.org/10.1088/0953-8984/5/14/013>
- Ziegler, J. (2024). *Interactions of ions with matter*. www.srim.org.
- Ziegler, J. F., Biersack, J. P., & Littmark, U. (1985). *The Stopping and Ranges of Ions in Solids*. Pergamon Press.

References

- Ashworth, D. G., Bowyer, M. D. J., & Oven, R. (1995). The derivation and moments solution of approximate transport equations for the implantation of ions into amorphous targets. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 100B(4), 471–482. [https://doi.org/10.1016/0168-583X\(95\)00347-9](https://doi.org/10.1016/0168-583X(95)00347-9)

- Bowyer, M. D. J., Ashworth, D. G., & Oven, R. (1992). Central moments of ion implantation distributions derived by the backward Boltzmann transport equation compared with Monte Carlo simulations. *Journal of Physics D: Applied Physics*, 25(11), 1619. <https://doi.org/10.1088/0022-3727/25/11/009>
- Ilyina, V. V., & Makarets, M. V. (2004). The cumulants analysis of the implanted ions distribution. *Ukrainian journal of Physics*, 49(8), 815–823.
- Ilyina, V. V., & Makarets, M. V. (2010). Distribution of energy losses by fast ions along their propagation paths in solids. *Ukrainian journal of Physics*, 55(2), 235–242.
- Lindhard, J., Scharf M., & Schiott H. E. (1963). Range concepts and heavy ion ranges. *Matematisk-fysiske Meddelelser udgivet af Det Kongelige Danske Videnskabernes Selskab*, 33(14), 1–43.
- Makarets, M. V., & Storchaka, S. N. (2001). A new method for calculating the distribution of implanted ions. 1. Algorithm and semi-invariants. *Ukrainian journal of Physics*, 46(4), 486–494 [in Ukrainian].
- Makarets, M. V., & Choliy, Y. V. (2013). *Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics and Mathematics*, 4, 255–259 [in Ukrainian].
- Oven, R., Bowyer, M. D. J., & Ashworth, D. G. (1993). A transport equation model that is consistent with Monte Carlo simulations of ion implantation into infinite targets. *Journal of Physics: Condensed Matter*, 5(14), 2157. <https://doi.org/10.1088/0953-8984/5/14/013>
- Ziegler, J. (2024). *Interactions of ions with matter*. www.srim.org.
- Ziegler, J. F., Biersack, J. P., & Littmark, U. (1985). *The Stopping and Ranges of Ions in Solids*. Pergamon Press.

Отримано редакцією журналу / Received: 28.01.25

Прорецензовано / Revised: 25.02.25

Схвалено до друку / Accepted: 12.03.25

Yaroslav CHOLIY, Head of Sector
e-mail: redox.sparrowhawk@gmail.com
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Mykola MAKARETS, DSc (Phys. & Math.), Prof.
ORCID ID: 0009-0005-8329-4079
e-mail: mmv_phys@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

INFLUENCE OF THE MODEL OF ELASTIC ION-ATOMIC COLLISIONS ON THE CUMULANTS OF THE SPATIAL DISTRIBUTION OF IMPLANTED IONS

Cumulants of the spatial distribution function of ions implanted in solid have been calculated. Equation for distribution function takes into account a finite ion path between of elastic collisions with the target atoms. Comparison with results of Lindhard-Scharff-Schiott theory shown that cumulants of these distributions are differ each from other for ions energies lower than a few hundred keV, and at higher energies they are coincide. This is due to the small length of ion path between collisions in compared with its longitudinal and lateral stragglings.

Keywords: fast ions, solid, collision model, distribution cumulants.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

Наукове видання



ВІСНИК
КИЇВСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ ІМЕНІ ТАРАСА ШЕВЧЕНКА
ФІЗИКО-МАТЕМАТИЧНІ НАУКИ

Випуск 1(80)

Оригінал-макет виготовлено Видавничо-поліграфічним центром "Київський університет"

Автори опублікованих матеріалів несуть повну відповідальність за підбір, точність наведених фактів, цитат, економіко-статистичних даних, відповідної галузевої термінології, власних імен та інших відомостей. Редколегія залишає за собою право скорочувати та редагувати подані матеріали.

Opinions, statements, accuracy of the quotations, economic and statistical data, terminology, proper names and other information are made on the responsibility of the authors. The Editorial Board reserves the right to shorten and edit the submitted materials.



Формат 60x84^{1/8}. Обл. вид. арк. 23,2. Ум. друк. арк. 22,0. Наклад 300. Зам. № 225-11390.
Гарнітура Arial. Папір офсетний. Друк офсетний. Вид. №К1.
Підписано до друку 24.06.25

Видавець і виготовлювач
ВПЦ "Київський університет"
Б-р Тараса Шевченка 14, м. Київ, 01030
☎ (38044) 239 32 22; (38044) 239 31 58; (38044) 239 31 28
e-mail: vpc@knu.ua; vpc_div.chief@univ.net.ua; redaktor@univ.net.ua
http: vpc.knu.ua

Свідоцтво суб'єкта видавничої справи ДК № 1103 від 31.10.02