

Харківський національний університет радіоелектроніки

Факультет інформаційно-аналітичних технологій та менеджменту

Кафедра прикладної математики

Рівень вищої освіти другий (магістерський)

Спеціальність 124 Системний аналіз

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системний аналіз і управління

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри ПМ _____

(підпис)

“ 06 ” листопада 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Школіну Олегу Вікторовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Застосування нейронної мережі U-net для розв'язання крайових задач для рівняння Гельмгольца

затверджена наказом по університету від 2 листопада 2023 р. № 1277 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 10 січня 2024 р.

3. Вихідні дані до роботи нейронна мережа для розв'язання диференціального рівняння Гельмгольца

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Системний аналіз предметної області

2. Вибір і обґрунтування методу розв'язання

3. Програмна реалізація

4. Результати обчислювального експерименту

5. Аналіз можливих застосувань

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій _____

1. Актуальність теми роботи _____

2. Постановка задачі _____

3. Системний аналіз предметної області _____

4. Метод чисельного аналізу _____

5. Результати обчислювального експерименту _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Підбір та вивчення технічної літератури за темою роботи	6 – 12 листопада 2023 р.	виконано
2	Вибір та обґрунтування методу	13 – 26 листопада 2023 р.	виконано
3	Розробка алгоритму і програми	27 листопада – 10 грудня 2023 р.	виконано
4	Проведення аналітичних досліджень та розрахунків	11 грудня – 24 грудня 2023 р.	виконано
5	Робота над текстом пояснювальної записки	25 грудня 2023 р. – 9 січня 2024 р.	виконано
6	Представлення роботи на рецензію в ЕК	10 січня 2024 р.	виконано

Дата видачі завдання 6 листопада 2023 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Гибкіна Н.В.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 58 с., 7 табл., 18 рис., 1 додаток, 14 джерел.

РІВНЯННЯ ГЕЛЬМГОЛЬЦА, НЕЙРОННА МЕРЕЖА, U-NET, МЕТОД КРИЛОВА, GMRES.

Об'єкт дослідження – крайова задача для диференціального рівняння Гельмгольца.

Мета роботи – дослідження застосування нейромережових технологій для розв'язання крайових задач для рівняння Гельмгольца.

Методи дослідження – методи чисельного аналізу, нейронна мережа U-net.

В кваліфікаційній роботі проведено дослідження можливості розв'язання рівняння Гельмгольца – еліптичного рівняння у частинних похідних із заданими крайовими умовами, за допомогою якого виконується моделювання розповсюдження електромагнітних хвиль у просторі, за допомогою нейронних мереж. Загальний підхід до розв'язання крайової задачі з рівнянням Гельмгольца полягає у зведенні його до системи лінійних рівнянь за допомогою спектрального методу Фур'є. В кваліфікаційній роботі досліджуються сучасні підходи до прискорення збіжності ітераційних алгоритмів розв'язання подібних систем лінійних рівнянь, зокрема нейромережові технології, а саме нейронна мережа U-net. Мета полягає в тому, щоб якомога швидше розв'язати велику систему лінійних рівнянь, яка є дискретизацією диференціального рівняння Гельмгольца, за допомогою попередньо обумовленого ітераційного методу Крилова, такого як Flexible GMRES. Проведено обчислювальні експерименти та проаналізовано ефективність запропонованого підходу.

ABSTRACT

Introductory note: 58 pages, 7 tables, 18 figures, 1 appendix, 14 sources.

HELMHOLTZ EQUATION, NEURAL NETWORK, U-NET, KRYLOV METHOD, GMRES.

Object of research – The boundary value problem for the Helmholtz differential equation.

Purpose of work – Research on the use of neural network technologies for solving boundary value problems for the Helmholtz equation.

Methods of research – Numerical analysis, networks U-net.

The qualification work investigates the possibility of solving the Helmholtz equation, an elliptic partial differential equation with specified boundary conditions, which is used to model the propagation of electromagnetic waves, using networks. The general approach to solving a boundary value problem with the Helmholtz equation is to reduce it to a system of linear equations using the spectral Fourier method. The qualification work investigates the modern approaches of accelerating the convergence of iterative algorithms for solving such systems of linear equations, in particular neural network technologies, namely the U-net neural network. The goal is to solve a large system of linear equations, which is a discretization of the Helmholtz differential equation, as quickly as possible using a predefined Krylov iterative method such as Flexible GMRES. Computational experiments are carried out and the effectiveness of the proposed approach is analyzed.

ЗМІСТ

	С.
Перелік скорочень, умовних познач, одиниць і термінів	8
Вступ	9
1 Системний аналіз предметної області та постановка задач дослідження	11
1.1 Системний аналіз проблеми математичного моделювання поширення електромагнітних хвиль	11
1.1.1 Вербальна модель системи.....	11
1.1.2 Морфологічний опис системи	11
1.1.3 Функціональна модель системи.....	13
1.1.4 Інформаційна модель	16
1.2 Аналіз сценаріїв вирішення проблеми розв’язання крайової задачі для рівняння Гельмгольца, що описує поширення електромагнітних хвиль ...	18
1.2.1 Модель аналізу проблеми.....	18
1.2.2 Оцінювання вектора пріоритетів незадоволеностей методом аналізу ієрархій.....	19
1.2.3 Модель вирішення проблеми.....	23
1.3 Змістовна та формальна постановки задачі	24
1.3.1 Змістовна постановка задачі	24
1.3.2 Формальна постановка задачі	25
1.4 Постановка задач дослідження	26
2 Вибір та обґрунтування методу розв’язання	27
2.1 Крайова задача для рівняння Гельмгольца.....	27
2.2 Диференціювання за допомогою дискретного перетворення Фур’є.....	29
2.3 Ідеально підібраний шар PML	32
2.4 Алгоритм GMRES	35
2.5 Огляд досліджень за темою роботи.....	36
2.6 Навчання без вчителя оберненого оператора.....	37
2.6.1 Фізично-інформована нейронна мережа	38

	7
2.6.2 DeepONet.....	40
2.6.3 Нейронний оператор Фур'є.....	42
2.6.4 U-net.....	43
Висновки за розділом 2	45
3 Програмна реалізація	46
3.1 Мова програмування Python	46
3.2 Дискретне перетворення Фур'є	47
3.3 Обчислення похідних.....	47
3.4 Алгоритм GMRES	48
Висновки за розділом 3	49
4 Результати обчислювального експерименту та їх аналіз	50
4.1 Обчислювальний експеримент	50
Висновки за розділом 4	52
Висновки	53
Перелік джерел посилання	54
Додаток А Лістинг програми	56

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАК, ОДИНИЦЬ І ТЕРМІНІВ

GMRES – Generalized minimal residual method

FNO – Fourier neural operator

DeepONet – Deep operator network

PML – Perfectly Matched Layer

ВСТУП

Дослідження різноманітних фізичних процесів приводить до необхідності розв'язування крайових задач для диференціальних рівнянь з частинними похідними. Зазвичай такі задачі не можуть бути розв'язані точно, тому актуальним є застосування методів їх чисельного розв'язування та застосування для цього сучасних інформаційних та обчислювальних технологій.

У даній кваліфікаційній роботі ми будемо застосовувати запропонований підхід до розв'язання крайової задачі для диференціального рівняння Гельмгольца з граничною умовою поглинання, що моделює поширення електромагнітних хвиль, шляхом дискретизації початкової неперервної задачі та зведення її до систем лінійних рівнянь.

Відомо, що дискретизована лінійна система стає погано обумовленою зі збільшенням хвильового числа в рівнянні Гельмгольца, що впливає на швидкість збіжності ітераційних алгоритмів. Вирішенням цієї проблеми зазвичай є введення попередньо обумовленої матриці, метою якої є зменшення числа обумовленості системи [11]. Передумова залежить від форми області та параметрів рівняння, таких як хвильове число, а це означає, що її потрібно перераховувати кожного разу, коли деякі з них змінюються.

Для чисельного обчислення передумов запропоновано велику кількість методів, але вони вимагають великих обчислювальних ресурсів. Тому виникає питання щодо можливості уникнення цих перерахунків. У роботі розглядається використання у якості передумови для цієї цілі нейронної мережі в гнучкому ітераційному алгоритмі GMRES, причому ідея полягає у тому, щоб навчити її лише один раз, а потім використовувати її без перенавчання для різних розмірів області та параметрів рівняння.

Мета і завдання кваліфікаційної роботи.

Об'єкт дослідження – крайова задача для диференціального рівняння Гельмгольца.

Предмет дослідження – розв'язання крайових задач для рівняння Гельмгольца.

Мета роботи – дослідження застосування нейромережових технологій для розв’язання крайових задач для рівняння Гельмгольца.

Наведемо перелік задач, які планується виконати для досягнення поставленої мети:

- сформулювати крайову задачу для рівняння Гельмгольца;
- дискретизувати рівняння Гельмгольца на заданій сітці та отримати систему лінійних рівнянь;
- натренувати нейронну мережу для апроксимації оберненого оператора Гельмгольца;
- провести обчислювальні експерименти для алгоритму GMRES з передумовою на різних сітках дискретизації;
- виконати аналіз отриманих результатів, зокрема порівняти кількість ітерацій необхідних для збіжності алгоритму для різних вхідних умов.

Методи дослідження. У кваліфікаційній роботі використовуються нейронні мережі типу U-net, DeepONet, FNO.

Публікації. Результати, отримані у кваліфікаційній роботі, було представлено на 27-му Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» (м. Харків, 10-12 травня 2023 р.) [14].

1 СИСТЕМНИЙ АНАЛІЗ ПРОБЛЕМИ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ ПОШИРЕННЯ ЕЛЕКТРОМАГНІТНИХ ХВИЛЬ ТА ПОСТАНОВКА ЗАДАЧ ДОСЛІДЖЕННЯ

1.1 Системний аналіз проблеми математичного моделювання поширення електромагнітних хвиль

1.1.1 Вербальна модель системи

Об'єкт аналізу: крайова задача для диференціального рівняння Гельмгольца.

Предмет аналізу: розв'язання крайових задач для рівняння Гельмгольца.

Точка зору: дослідник.

Призначення: моделювання поширення електромагнітних хвиль у природних середовищах.

Ціль: отримання чисельних розв'язків крайових задач для рівняння Гельмгольца, що описують поширення електромагнітних хвиль у природних середовищах. Розв'язання крайових задач для рівняння Гельмгольца в прикладних галузях допомагає проектувати та оптимізувати системи, покращувати якість моделювання, проводити точну діагностику та дослідження, що робить це важливим інструментом для багатьох сучасних технологій і наукових досліджень.

1.1.2 Морфологічний опис системи

Розглянемо систему «Дослідження процесу поширення електромагнітних хвиль». Морфологічна модель передбачає виконання опису структури, меж та зовнішнього середовища.

Почнемо з розгляду моделі «чорна скринька». Вона дозволяє зробити спрощене подання системи, що дає змогу аналізувати поведінку системи без докладного знання її внутрішньої структури.

Для системи «Дослідження процесу поширення електромагнітних хвиль» модель «чорна скринька» подана на рисунку 1.1. Вона представляє собою спрощену математичну або чисельну модель, яка акцентується на вхідних та вихідних параметрах процесу без опису його внутрішньої структури. Входом до моделі «чорна скринька» є диференціальне рівняння Гельмгольца з крайовими умовами та значення параметрів у кожній точці простору. Виходом «чорної скриньки» є чисельний розв'язок заданої крайової задачі, який відображує залежність компонент електромагнітних хвиль у просторі.

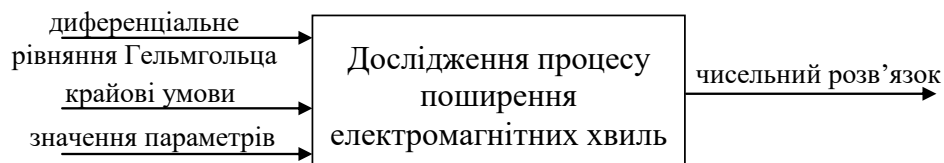


Рисунок 1.1 – Модель «чорна скринька»

Зовнішнім середовищем є умови та фактори, які впливають на роботу процесу. Для системи «Дослідження процесу поширення електромагнітних хвиль» об'єктами зовнішнього середовища є дослідник, обчислювальна техніка, теорія та методи математичної фізики та чисельних методів.

Межею системи «Дослідження процесу поширення електромагнітних хвиль» є простір, в якому користувач-дослідник моделює процес поширення електромагнітних хвиль за допомогою рівняння Гельмгольца, розробляє алгоритм розв'язання крайових задач для рівняння Гельмгольца та аналізує отримані результати.

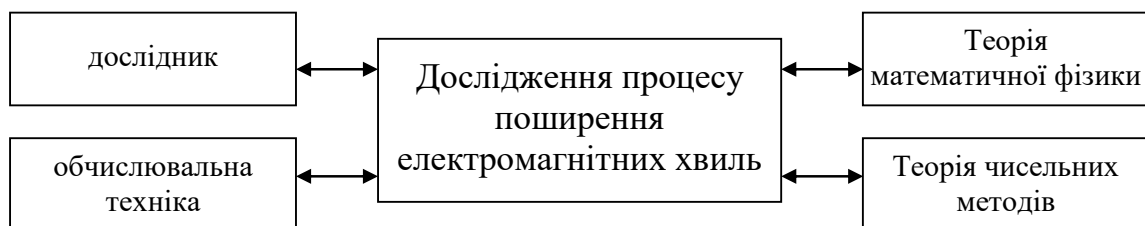


Рисунок 1.2 – Модель зовнішнього середовища системи

1.1.3 Функціональна модель системи

Функціональна модель системи «Дослідження процесу поширення електромагнітних хвиль» включає послідовність дій, метою яких є розв'язання крайових задач для рівняння Гельмгольца, що описують моделювання поширення електромагнітних хвиль у природних середовищах, з метою отримання чисельних розв'язків цих задач.

Розв'язання крайової задачі для рівняння Гельмгольца включає пошук функції, яка задовольняє рівняння і граничні умови. Розглянемо основні функціональні компоненти цієї системи:

- введення даних, зокрема, параметри середовища, крайові умови;
- побудова математичної моделі – рівняння Гельмгольца;
- дискретизація диференціального рівняння;
- розв'язання отриманої лінійної системи рівнянь;
- аналіз результатів;
- візуалізація рішення.

Для наочного зображення функціональної моделі використовується методологія IDEF0 – методологія для моделювання функцій та їх взаємозв'язків всередині системи. Найзагальніший опис системи, її функціонування в цілому та взаємодія з зовнішнім середовищем подається за допомогою першої діаграми в ієрархії діаграм IDEF0, яка називається контекстною діаграмою.

Контекстна діаграма IDEF0 для системи «Дослідження процесу поширення електромагнітних хвиль» наведена на рисунку 1.3. На вхід цієї системи подається диференціальне рівняння, значення параметрів, крайові умови; засобами управління є теорія математичної фізики та теорія чисельних методів; механізмами системи є дослідник та програмне забезпечення; на виході маємо чисельний розв'язок задачі.

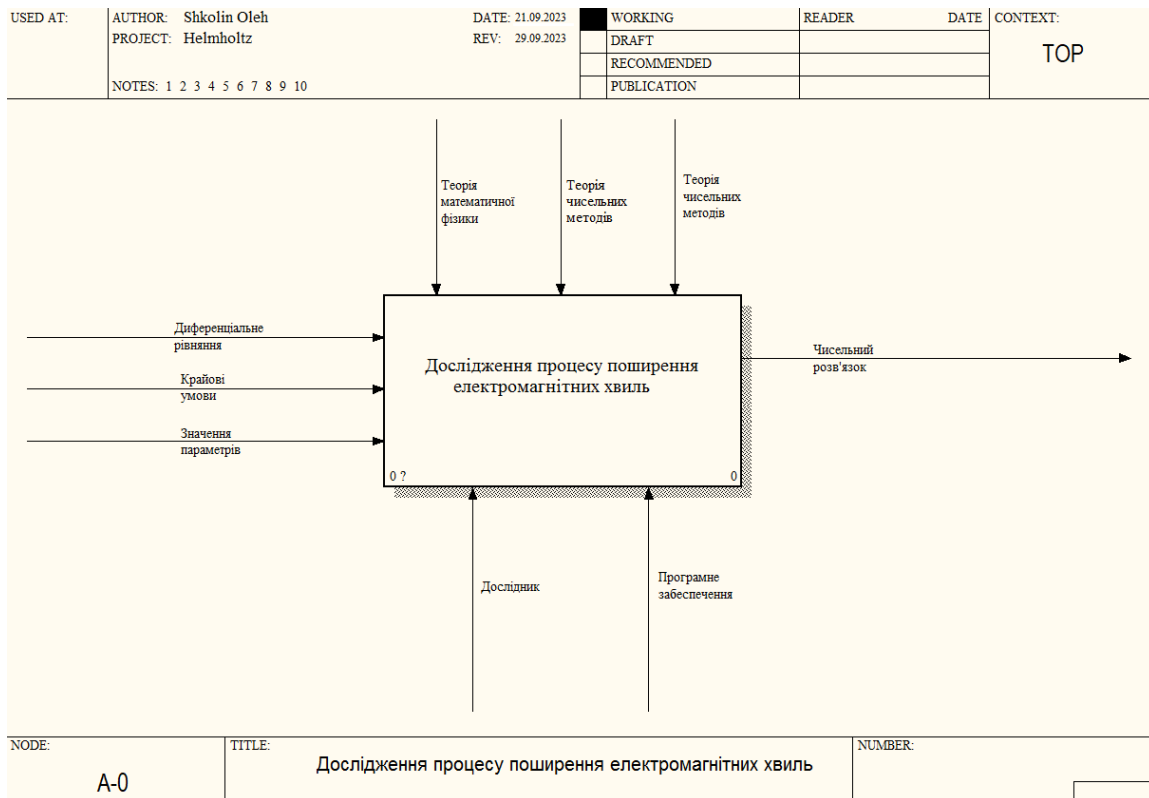


Рисунок 1.3 – Контекстна діаграма (рівень А-0)

Наступним кроком після загального опису системи виконуємо декомпозицію першого рівня на великі фрагменти, і далі – декомпозицію другого (і наступних) рівня окремих фрагментів на більш дрібні до досягнення потрібного рівня деталізації опису. Глибина декомпозиції процесів в системі визначається дослідником таким чином, щоб з одного боку, отримати достатній рівень деталізації, а з іншого – не перевантажувати модель несуттєвою інформацією.

Декомпозиція контекстної діаграми системи «Дослідження процесу поширення електромагнітних хвиль» містить основні функції системи (рис. 1.4). Декомпозиція другого рівня подає подальшу деталізацію однієї з функцій системи – «Виконати дискретизацію диференціального рівняння» та відображує функції, які реалізуються на цьому етапі (рис. 1.5).

Для документування та опису процесів, які відбуваються в системі, використовується методологія IDEF3. IDEF3-діаграма дозволяє візуально відобразити взаємодію функціональних блоків у системі та потоки даних і управління між ними.

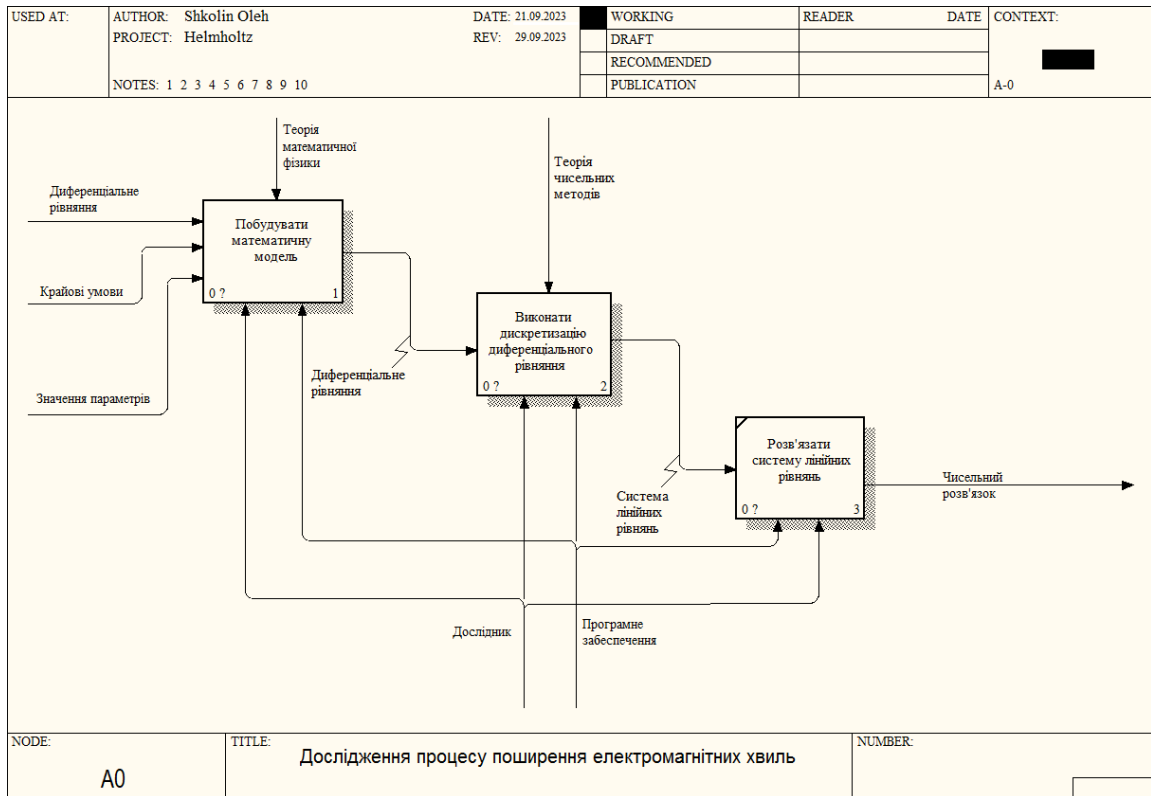


Рисунок 1.4 – Декомпозиція роботи «Дослідження процесу поширення електромагнітних хвиль»: рівень A0

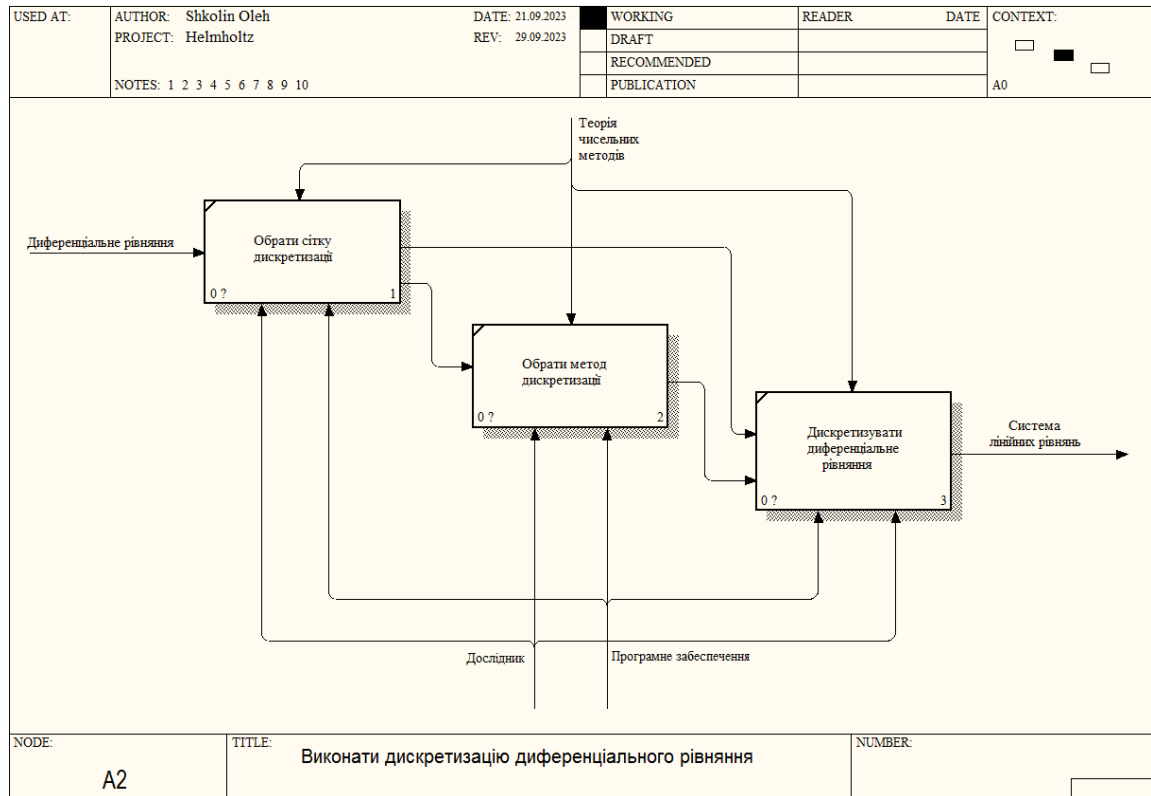


Рисунок 1.5 – Декомпозиція роботи «Виконати дискретизацію диференціального рівняння»: рівень A1

Перша діаграма в ієрархії IDEF3 для системи «Дослідження процесу поширення електромагнітних хвиль» наведена на рисунку 1.6. Вона складається з таких блоків: «побудувати математичну модель», «виконати дискретизацію диференціального рівняння» та «розв’язати систему лінійних рівнянь».

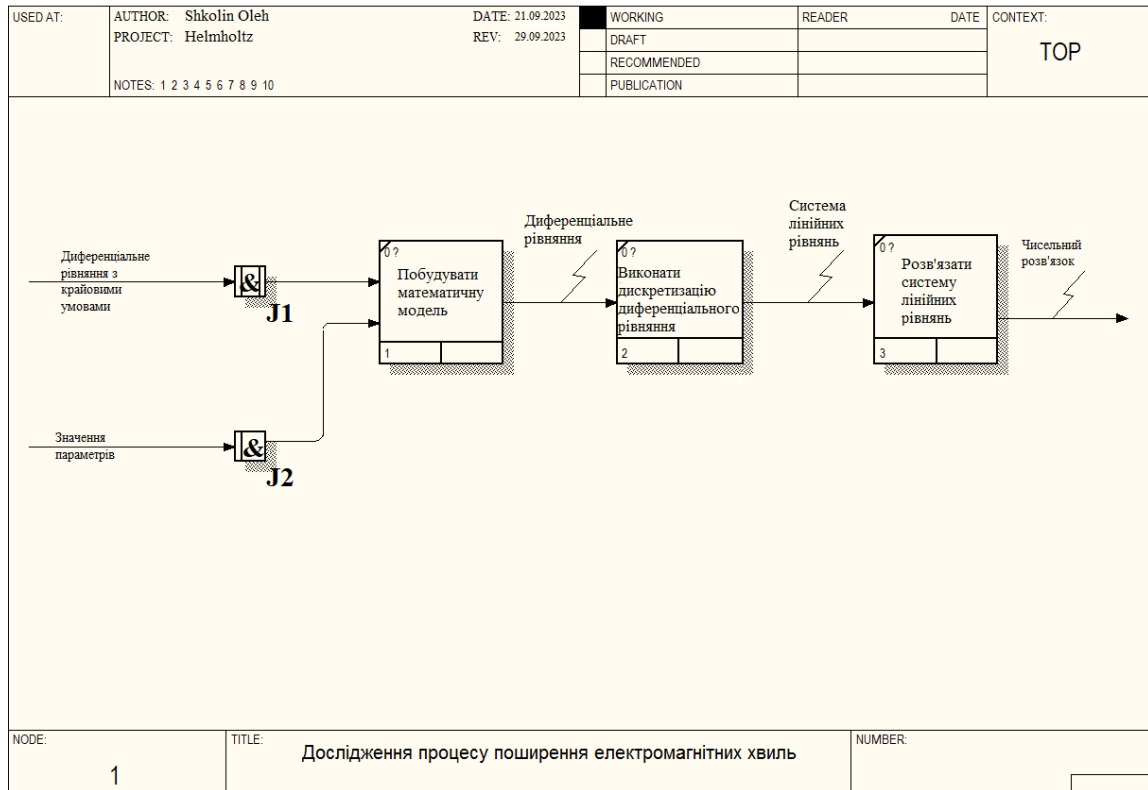


Рисунок 1.6 – Опис роботи «Дослідження процесу поширення електромагнітних хвиль»: рівень A0 (в нотації IDEF3)

1.1.4 Інформаційна модель

Інформаційна модель відображує структуру, потоки інформації та взаємозв'язки між ними всередині системи. Вона описує, як дані організовано, як їх передають і як використовують у межах системи. Інформаційна модель системи «Дослідження процесу поширення електромагнітних хвиль» може включати такі компоненти, дані, процеси, необхідні для проведення досліджень, як:

- вхідні дані: рівняння Гельмгольца, параметри, крайові умови;

- методи математичної фізики та чисельні методи;
- інформація про обчислювальні ресурси, за допомогою яких виконується розв’язання задачі та чисельне моделювання;
- вихідна інформація у вигляді розв’язку задачі, аналізу результатів та висновків про досліджуваний процес.

Для графічного подання потоків даних інформаційної моделі системи використовуються DFD-діаграми (Data Flow Diagram), які дозволяють наочно продемонструвати, як переміщуються та обробляються дані всередині системи, як взаємодіють різні компоненти системи.

DFD-діаграма інформаційної моделі системи «Дослідження процесу поширення електромагнітних хвиль» наведена на рисунку 1.7. На ній бачимо, що потоки інформації передаються від дослідника, який і займається моделюванням процесу, розв’язанням задачі та аналізом результатів. Після розв’язання задачі отримаємо звіт про поведінку процесу поширення електромеханічних хвиль із заданими параметрами.

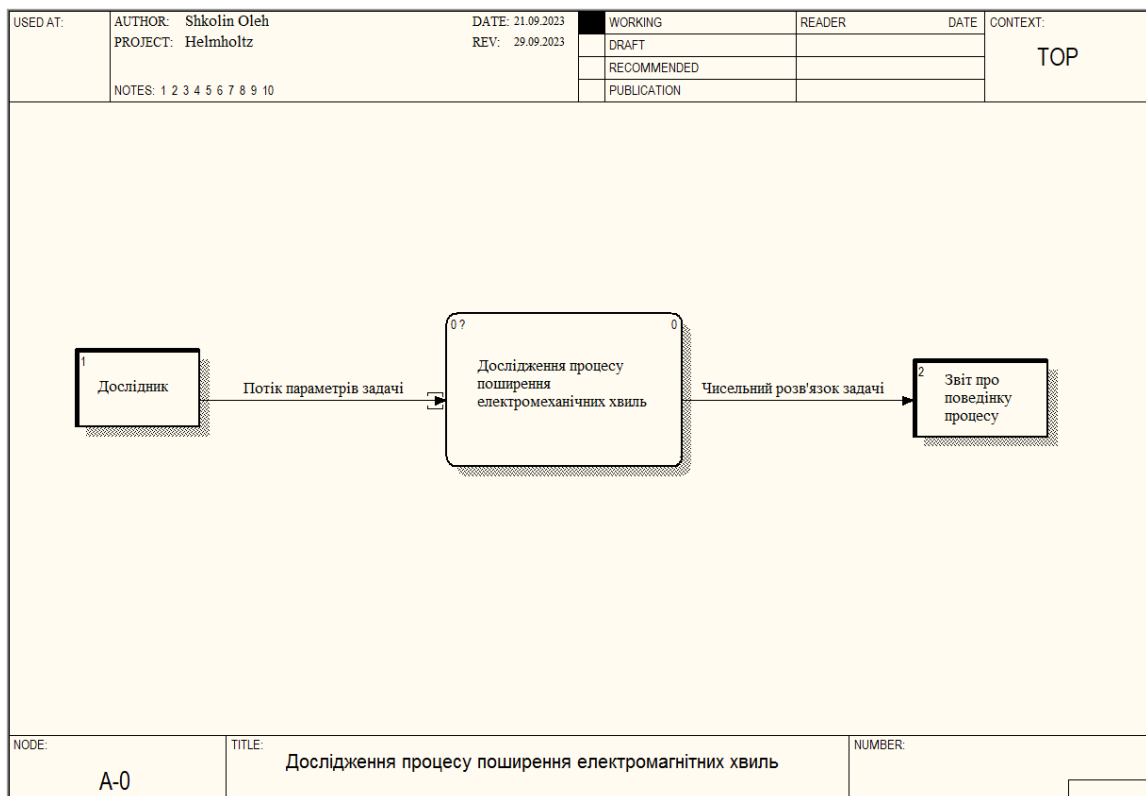


Рисунок 1.7 – DFD-діаграма

1.2 Аналіз сценаріїв вирішення проблеми розв'язання крайової задачі для рівняння Гельмгольца, що описує поширення електромагнітних хвиль

1.2.1 Модель аналізу проблеми

Тепер розглянемо проблему розв'язання крайової задачі для рівняння Гельмгольца, що описує поширення електромагнітних хвиль. Нас цікавитиме вибір з певної множини альтернатив такого методу для розв'язання цієї задачі, який би за певними критеріями був найефективнішим. Тому потрібно сформулювати критерії, за якими ми будемо порівнювати можливі методи розв'язання задачі:

- критерій 1 (K1): складність реалізації алгоритму;
- критерій 2 (K2): масштабованість алгоритму для різних сіток;
- критерій 3 (K3): чутливість алгоритму до зміни параметрів системи;
- критерій 4 (K4): обчислювальна складність етапу навчання;
- критерій 5 (K5): обчислювальна складність при застосуванні.

Серед альтернатив, з яких за заданими критеріями обиратимемо метод для розв'язання задачі, проаналізуємо наступні:

- альтернатива 1 (A1): метод підпросторів Крилова;
- альтернатива 2 (A2): нейронна мережа U-net;
- альтернатива 3 (A3): комбінація методу Крилова з U-net.

На рисунку 1.8 наведено ієрархічну структуру для моделі вибору методу для вирішення проблеми розв'язання крайової задачі для рівняння Гельмгольца, що описує поширення електромагнітних хвиль. Фокус цієї ієрархії – це безпосередньо проблема розв'язання поставленої задачі, перший рівень ієрархії містить критерії, за якими обирається метод, другий рівень ієрархії – це варіанти, з яких обираємо метод розв'язання.

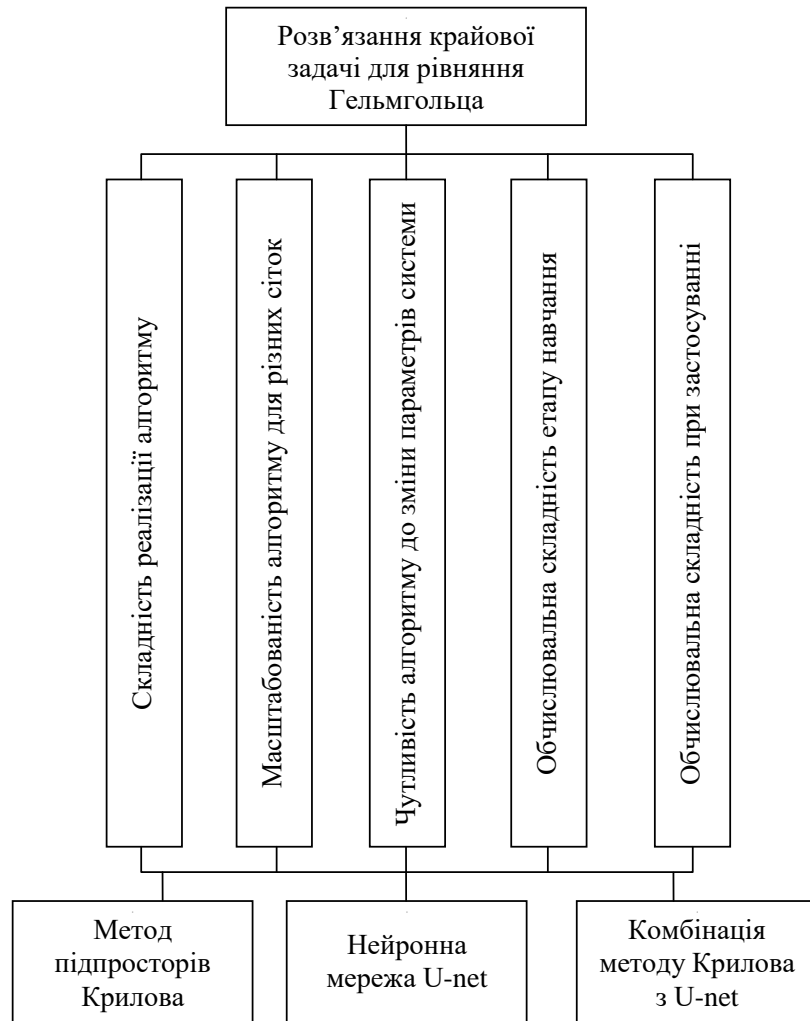


Рисунок 1.8 – Ієрархічна структура аналізу проблеми вибору методу для розв'язання крайової задачі для рівняння Гельмгольца

1.2.2 Оцінювання вектора пріоритетів незадоволеностей методом аналізу ієрархій

Для вибору методу застосуємо метод аналізу ієрархії, відповідно до якого побудуємо матрицю парних порівнянь критеріїв (таблиця 1.1) та матриці парних порівнянь методів за визначеними критеріями (таблиці 1.2 – 1.6).

Таблиця 1.1 – Матриця парних порівнянь критеріїв

Критерії оцінювання	К1	К2	К3	К4	К5	Оцінки компонентів	Вектор пріоритетів
К1	1	$\frac{1}{5}$	$\frac{1}{5}$	5	$\frac{1}{5}$	0,53	0,09
К2	5	1	1	5	1	1,90	0,32
К3	5	1	1	5	1	1,90	0,32
К4	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	1	2	0,44	0,07
К5	5	1	1	$\frac{1}{2}$	1	1,20	0,20
Усього						5,97	

За таблицею 1.1 обчислимо:

$$- \text{індекс узгодженості } IU = \frac{5,85 - 5}{5 - 1} \approx 0,21;$$

$$- \text{відносну узгодженість } BU = \frac{0,21}{1,12} \approx 0,19 = 19\%.$$

Тут значення 1,12 – випадкова узгодженість для матриці п'ятого порядку.

Оскільки відносна узгодженість не перебільшує значення 0,2, то можемо зробити висновок, що ми правильно побудували матрицю парних порівнянь критеріїв.

За даними таблиці 1.1 випишемо вектор локальних пріоритетів критеріїв відносно проблеми вибору методу розв'язання:

$$\vec{p}^K = (0,09; 0,32; 0,32; 0,07; 0,20).$$

Таблиця 1.2 – Матриця попарних порівнянь за першим критерієм

K1	A1	A2	A3	Оцінки компонентів	Вектор пріоритетів
A1	1	$\frac{1}{2}$	$\frac{1}{4}$	0,50	0,13
A2	2	1	$\frac{1}{5}$	0,74	0,19
A3	4	5	1	2,71	0,68
Усього				3,95	

За таблицею 1.2 обчислимо:

– індекс узгодженості $IY = \frac{3,09 - 3}{3 - 1} \approx 0,05$;

– відносну узгодженість $VY = \frac{0,05}{0,58} \approx 0,09 = 9\%$.

Тут значення 0,58 – випадкова узгодженість для матриці третього порядку.

Таблиця 1.3 – Матриця попарних порівнянь за другим критерієм

K2	A1	A2	A3	Оцінки компонентів	Вектор пріоритетів
A1	1	$\frac{1}{5}$	$\frac{1}{5}$	0,34	0,09
A2	5	1	$\frac{1}{3}$	1,19	0,30
A3	5	3	1	2,46	0,61
Усього				3,99	

За таблицею 1.3 обчислимо:

– індекс узгодженості $IY = \frac{3,14 - 3}{3 - 1} \approx 0,07$;

– відносну узгодженість $ВУ = \frac{0,07}{0,58} \approx 0,12 = 12\%$.

Таблиця 1.4. Матриця попарних порівнянь за третім критерієм

К3	А1	А2	А3	Оцінки компонентів	Вектор пріоритетів
А1	1	$\frac{1}{5}$	$\frac{1}{7}$	0,31	0,07
А2	5	1	$\frac{1}{4}$	1,08	0,24
А3	7	4	1	3,04	0,69
Усього				4,43	

За таблицею 1.4 обчислимо:

– індекс узгодженості $IУ = \frac{3,12 - 3}{3 - 1} \approx 0,06$;

– відносну узгодженість $ВУ = \frac{0,06}{0,58} \approx 0,10 = 10\%$.

Таблиця 1.5 – Порівняння за четвертим критерієм

К4	А1	А2	А3	Оцінки компонентів	Вектор пріоритетів
А1	1	$\frac{1}{3}$	$\frac{1}{5}$	0,41	0,10
А2	3	1	$\frac{1}{3}$	1,00	0,26
А3	5	3	1	2,47	0,64
Усього				3,88	

За таблицею 1.5 обчислимо:

$$- \text{індекс узгодженості } IU = \frac{3,04 - 3}{3 - 1} \approx 0,02;$$

$$- \text{відносну узгодженість } VU = \frac{0,02}{0,58} \approx 0,03 = 3\%.$$

Таблиця 1.6 – Порівняння за п'ятим критерієм

K5	A1	A2	A3	Оцінки компонентів	Вектор пріоритетів
A1	1	$\frac{1}{2}$	$\frac{1}{3}$	0,55	0,16
A2	2	1	$\frac{1}{3}$	0,87	0,25
A3	3	3	1	2,08	0,59
Усього				3,50	

За таблицею 1.6 обчислимо:

$$- \text{індекс узгодженості } IU = \frac{3,05 - 3}{3 - 1} \approx 0,03;$$

$$- \text{відносну узгодженість } VU = \frac{0,03}{0,58} \approx 0,05 = 5\%.$$

1.2.3 Модель вирішення проблеми

З використанням отриманих у попередньому пункті результатів (а саме вектори пріоритетів для критеріїв та для альтернатив за окремими критеріями) виконаємо остаточні розрахунки (таблиця 1.7) і на їх основі зробимо висновок щодо вирішення проблеми вибору методу вибору розв'язання крайової задачі для рівняння Гельмгольца.

Таблиця 1.7 – Розрахунок вектору узагальнених пріоритетів

Критерій Альтернатива	K1	K2	K3	K4	K5	Узагальнені пріоритети
A1	0,13	0,09	0,07	0,10	0,16	0,10
A2	0,19	0,30	0,24	0,26	0,25	0,26
A3	0,68	0,61	0,69	0,64	0,59	0,64

Відповідно до значень останнього стовпця таблиці 1.7 у якості найефективнішого за сформульованими критеріями методу розв'язання крайової задачі для рівняння Гельмгольца обираємо третю альтернативу (A3) – комбінацію методу Крилова з U-net.

1.3 Змістовна та формальна постановка задачі

1.3.1 Змістовна постановка задачі

Розглядається задача моделювання розповсюдження електромагнітних хвиль у просторі. Математичною моделлю цього процесу є рівняння Гельмгольца – еліптичне рівняння у частинних похідних із заданими крайовими умовами.

Безпосереднє розв'язання цього рівняння у загальному випадку пов'язано з певними складнощами, тому використовується підхід, що полягає у дискретизації початкової неперервної задачі і зведенні диференціального рівняння Гельмгольца до системи лінійних рівнянь за допомогою спектрального методу Фур'є. Розв'язання цієї системи лінійних рівнянь може бути виконано за допомогою ітеративних алгоритмів, але воно стикається з певними труднощами, викликаними розміром системи, зокрема, потребує значного часу на розрахунки.

У кваліфікаційній роботі досліджуються сучасні підходи до прискорення збіжності ітераційних алгоритмів розв'язання таких систем лінійних рівнянь, такі як нейромережеві технології, а саме нейронна мережа U-net, що добре себе зарекомендувала в різних задачах глибинного навчання.

1.3.2 Формальна постановка задачі

У кваліфікаційній роботі розглядається задача розв'язання рівняння Гельмгольца з крайовими умовами випромінювання Зоммерфельда на нескінченності:

$$(\nabla^2 + k^2(\mathbf{x}))u(\mathbf{x}) = \rho(\mathbf{x}), \quad (1.1)$$

де $\mathbf{x} \in \mathbb{R}^d$ – точка у d -вимірній області;

∇^2 – оператор Лапласа;

$k(\mathbf{x}) = \frac{\omega}{c(\mathbf{x})}$, $k: \mathbb{R}^d \rightarrow \mathbb{R}$, – функція хвильового числа;

$\rho: \mathbb{R}^d \rightarrow \mathbb{R}$ – функція джерела;

$u: \mathbb{R}^d \rightarrow \mathbb{C}$ – комплексне акустичне хвильове поле;

$c: \mathbb{R}^d \rightarrow \mathbb{R}_+$ – функція розподілу швидкості звуку;

$\omega \in \mathbb{R}_+$ – кутова частота джерела.

Вважається, що $c(x)$ є неоднорідною в обмеженій розглядуваній області, а поза нею вона однорідна і дорівнює c_0 .

Після дискретизації задача (1.1) перетворюється на задачу розв'язання системи лінійних рівнянь

$$A(\hat{\mathbf{c}})\hat{\mathbf{u}} = \hat{\boldsymbol{\rho}}, \quad (1.2)$$

де $A(\hat{\mathbf{c}}) \in \mathbb{C}^{n \times n}$ – дискретизований оператор Гельмгольца, який залежить від дискретизації розподілу швидкості звуку $\hat{\mathbf{c}} \in \mathbb{R}_+^n$;

$\hat{\mathbf{u}} \in \mathbb{C}^n$ – дискретизація розв'язку;

$\hat{\boldsymbol{\rho}}$ – дискретизація джерела;

$n \in \mathbb{N}$ – число точок дискретизації.

Систему (1.2) пропонується розв'язувати ітераційним алгоритмом із сім'ї методів підпросторів Крилова GMRES. Але чисельне розв'язання задачі (1.2) ускладнене через наявність у рівнянні хвильового числа $k(\mathbf{x})$, значення якого впливають на обумовленість отриманої після дискретизації системи лінійних рівнянь.

У кваліфікаційній роботі досліджується можливість прискорення роботи алгоритму GMRES та зменшення ітерацій, необхідних для розв'язання отриманої системи лінійних рівнянь, шляхом використання нейромережових технологій, зокрема, нейронної мережі U-net в якості оператора передумови.

1.4 Постановка задач дослідження

Метою даної кваліфікаційної роботи є дослідження застосування нейронних мереж для прискорення ітераційного алгоритму GMRES, застосованого до системи лінійних рівнянь, що була отримана після дискретизації оператора Гельмгольца.

Наведемо перелік задач, які планується виконати у ході виконання кваліфікаційної роботи:

- сформулювати крайову задачу для рівняння Гельмгольца;
- дискретизувати рівняння Гельмгольца на заданій сітці та отримати систему лінійних рівнянь;
- натренувати нейронну мережу для апроксимації оберненого оператора Гельмгольца;
- провести обчислювальні експерименти для алгоритму GMRES з передумовою на різних сітках дискретизації;
- виконати аналіз отриманих результатів, зокрема порівняти кількість ітерацій необхідних для збіжності алгоритму для різних вхідних умов.

2 ВИБІР ТА ОБҐРУНТУВАННЯ МЕТОДУ РОЗВ'ЯЗАННЯ

2.1 Крайова задача для рівняння Гельмгольца

Розглянемо диференціальне рівняння Гельмгольца

$$(\nabla^2 + k^2(\mathbf{x}))u(\mathbf{x}) = \rho(\mathbf{x}), \quad (2.1)$$

де $\mathbf{x} \in \mathbb{R}^d$ – точка у d -вимірній області;

∇^2 – оператор Лапласа;

$k: \mathbb{R}^d \rightarrow \mathbb{R}$ – функція хвильового числа;

$\rho: \mathbb{R}^d \rightarrow \mathbb{R}$ – функція джерела;

$u: \mathbb{R}^d \rightarrow \mathbb{C}$ – комплексне акустичне хвильове поле.

Хвильове число визначається як $k(\mathbf{x}) = \frac{\omega}{c(\mathbf{x})}$, де $c: \mathbb{R}^d \rightarrow \mathbb{R}_+$ – функція

розподілу швидкості звуку, а $\omega \in \mathbb{R}_+$ – кутова частота джерела.

У нашій роботі ми будемо розв'язувати рівняння (2.1) з урахуванням умови випромінювання Зоммерфельда на нескінченності. Вважається, що $c(\mathbf{x})$ є неоднорідною в обмеженій розглядуваній області, а поза нею вона однорідна і дорівнює c_0 . Умова випромінювання Зоммерфельда задовольняється за допомогою методу ідеально узгодженого шару [3], ідея якого полягає у зміні похідних. Щоб розв'язати отримане рівняння чисельно, нам потрібно його дискретизувати. Згідно з [13] будемо використовувати швидке перетворення Фур'є для дискретизації оператора Гельмгольца, застосовуючи теорему диференціювання [12], яка дає зв'язок між перетворенням Фур'є функції $f: \mathbb{R} \rightarrow \mathbb{R}$ (або \mathbb{C}) та її похідною f' у вигляді

$$\mathcal{F}\{f'\}(\omega) = i\omega\mathcal{F}\{f\}(\omega), \quad (2.2)$$

де $\mathcal{F}\{f\}$ – неперервне перетворення Фур'є функції f . Після дискретизації задача (2.1) перетворюється на задачу розв'язання лінійної системи

$$A(\hat{c})\hat{u} = \hat{\rho}, \quad (2.3)$$

де $A(\hat{c}) \in \mathbb{C}^{n \times n}$ – дискретизований оператор Гельмгольца, який залежить від дискретизації розподілу швидкості звуку $\hat{c} \in \mathbb{R}_+^n$;

$\hat{u} \in \mathbb{C}^n$ – дискретизація розв'язку;

$\hat{\rho}$ – дискретизація джерела;

$n \in \mathbb{N}$ – число точок дискретизації.

Відомо, що обумовленість системи (2.3) погіршується зі збільшенням хвильових чисел $k(\mathbf{x})$. Тому пропонується розв'язувати систему (2.3) ітераційним алгоритмом із сім'ї методів підпросторів Крилова GMRES.

Метою роботи є прискорення алгоритму GMRES за допомогою оператора передумови $M: \mathbb{C}^n \rightarrow \mathbb{C}^n$, побудованого за допомогою нейронної мережі, де оператор M певним чином апроксимує $A(c)^{-1} \in \mathbb{C}^{n \times n}$. Для того, щоб надати повну картину, ми повинні деталізувати процес дискретизації за допомогою дискретного перетворення Фур'є, використовуючи PML та GMRES з передумовою.

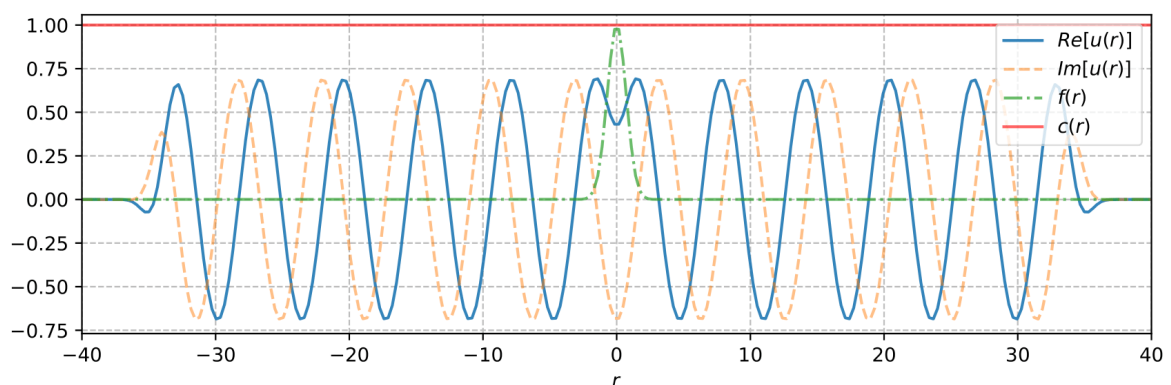


Рисунок 2.1 – Розв'язок одновимірного диференціального рівняння для 256 точок зі сталою $c = 1$

2.2 Диференціювання за допомогою дискретного перетворення Фур'є

Розглянемо операцію диференціювання, яку виконує перетворення Фур'є. Його аналогія в дискретному просторі буде корисною для дискретизації оператора Гельмгольца. Почнемо з означень прямого та оберненого перетворення Фур'є.

Означення 2.1 ([12]). Нехай f – абсолютно інтегровна функція, визначена для всіх $t \in \mathbb{R}$, тобто

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty. \quad (2.4)$$

Тоді існує відображення $\mathcal{F}\{f\}: \mathbb{R} \rightarrow \mathbb{C}$, яке називається перетворенням Фур'є f і визначається як

$$\mathcal{F}\{f\}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \quad (2.5)$$

а обернене перетворення задається як

$$\mathcal{F}^{-1}\{f\}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) e^{i\omega t} d\omega.$$

За допомогою наведеного означення сформулюємо теорему про диференціювання, дискретний варіант якої буде корисним у подальшому.

Теорема 2.1 (Теорема про диференціювання). Нехай $f(t)$ – диференційовна функція для всіх $t \in \mathbb{R}$, а $f'(t)$ – її похідна, причому вони абсолютно інтегровні. Тоді маємо

$$\mathcal{F}\{f'\}(\omega) = i\omega \mathcal{F}\{f\}(\omega). \quad (2.6)$$

Перейдемо до дискретного випадку і пояснимо зв'язок між перетворенням Фур'є та його дискретним варіантом, а потім адаптуємо результат теореми 2.1. Спочатку дамо означення прямому та оберненому дискретному перетворенню Фур'є.

Означення 2.2. Нехай $\hat{\mathbf{f}} = \{\hat{f}_0, \dots, \hat{f}_{N-1}\}$ – послідовність дійсних або комплексних чисел. Тоді дискретне перетворення Фур'є $\hat{\mathbf{f}}$ визначається як

$$\mathcal{F}\{\hat{\mathbf{f}}\}(n) = \sum_{k=0}^{N-1} \hat{f}_k e^{-i2\pi nk/N} \quad n \in \{0, \dots, N-1\}, \quad (2.7)$$

а обернене перетворення задається як

$$\mathcal{F}^{-1}\{\hat{\mathbf{f}}\}(k) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{f}_n e^{i2\pi nk/N}, \quad k \in \{0, \dots, N-1\}. \quad (2.8)$$

Використовуючи теорему 2.1 запишемо еквівалент співвідношення (2.6) у дискретному випадку.

Твердження 2.1. Нехай $\hat{\mathbf{f}} = \{f_0, \dots, f_{N-1}\}$ – дискретизація диференційовної функції $f(t)$ на $t \in [0, T]$ така, що $f(t)$ та $f'(t)$ абсолютно інтегровні, де $\hat{f}_k = f(kT/N)$ для $k \in \{0, \dots, N-1\}$. Тоді маємо

$$\mathcal{F}\{f'\}(\omega_n) \approx i\omega_n \mathcal{F}\{\hat{\mathbf{f}}\}(n), \quad (2.9)$$

де $\omega_n = \frac{2\pi n}{T}$ для $n \in \{0, \dots, N-1\}$.

Доведення цього твердження впливає із заміни правої частини в (2.6) на її дискретний варіант і завдання дискретних частот ω_n , як у (2.7).

Твердження 2.2. Нехай $\hat{\mathbf{f}} = \{\hat{f}_0, \dots, \hat{f}_{N-1}\}$ – дискретизація диференційовної функції $f(t)$ на $t \in [0, T]$ така, що $f(t)$ та $f'(t)$ абсолютно інтегровні, де

$\hat{f}_k = f(kT/N)$. Нехай $\hat{\mathbf{f}}' = \{\hat{f}'_0, \dots, \hat{f}'_{N-1}\}$ є дискретизацією похідної $f'(t)$, де $\hat{f}'_k = f'(kT/N)$. Тоді ми можемо апроксимувати $\hat{\mathbf{f}}'$ прямим та оберненим перетворенням Фур'є на $\hat{\mathbf{f}}$ як

$$\hat{\mathbf{f}}' \approx \mathcal{F}^{-1} \{i\omega \mathcal{F}\{\hat{\mathbf{f}}\}\}, \quad (2.10)$$

де $\omega = \{\omega_0, \dots, \omega_{N-1}\}$, $\omega_n = 2\pi n/T$.

Доведення цього твердження здійснюється шляхом застосування неперервного оберненого перетворення Фур'є зліва та його дискретної змінної справа у (2.9).

Зауважимо, що практична реалізація вимагає корекції значення T для заданого N . Розглянемо діапазон $[0, T]$ для $T > 0$. Тоді остання точка, в якій обчислюється функція f , буде $\frac{(N-1)T}{N}$, що означає, що ми не включаємо $f(T)$.

Також зауважимо, що у твердженні 2.2 було розглянуто N різних частот $\omega = \{\omega_0, \dots, \omega_{N-1}\}$ для диференціювання, що впливає з наївної заміни неперервного перетворення Фур'є на його дискретний варіант у (2.6). На практиці, однак, всі частоти ω_n , для яких $n > N/2$ для парних N і $n > (N-1)/2$ для непарних N не існують, оскільки $\omega_{N/2}$ (або $\omega_{(N-1)/2}$) є найвищою частотою, яку ми можемо визначити (частота Найквіста). Це впливає з теореми відліків Найквіста-Шеннона.

Тепер ми можемо використовувати цей інструмент для обчислення оператора Гельмгольца на послідовності дискретизованих точок $\hat{\mathbf{f}}$, які задовольняють умовам теореми 2.1.

Отже, ми спробували пояснити підхід до диференціювання з використанням перетворення Фур'є. Його теоретичне обґрунтування наведено у Твердженні 2.2. Зауважимо, що це не математична теорема з повним доведенням, а скоріше інтуїтивний нарис для пояснення реалізації на Python. Причиною цього є те, що хоча теорема диференціювання добре вивчена і

представлена у багатьох книгах з перетворень Фур'є, її адаптацію до дискретного випадку важко знайти в літературі.

2.3 Ідеально підібраний шар PML

Розглянемо поняття ідеально підібраного шару (англ. Perfectly Matched Layer, PML) [3].

Perfectly Matched Layer або PML – це підхід для усічення обчислювальної області в чисельних методах моделювання відкритих крайових задач. Іншими словами, ми хочемо змоделювати поглинаючу граничну умову й уникнути прямого використання умови випромінювання Зоммерфельда на нескінченності (2). Вона була введена Беренгером в роботі [1], де ідея полягала в тому, щоб замінити поглинаючу граничну умову поглинаючим примежовим шаром, як показано на рисунку 2.2.

Проблема цього підходу полягає в тому, що хвилі відбиваються на переході між матеріалами, що також стосується і поглинаючого матеріалу. Замість відбиття від межі решітки відбувається відбиття від межі поглинача. У роботі [1] показано, що спеціальне поглинаюче середовище можна сконструювати так, щоб хвилі не відбивалися на межі розділу.

Почнемо розгляд з розв'язку $u(\mathbf{x})$ одновимірного рівняння Гельмгольца (2.1) у нескінченному просторі. Розглянемо усікання задачі в додатному напрямку (від'ємний напрямок розглядатиметься аналогічно). Процес усікання можна пояснити за допомогою наступних трьох кроків.

Крок 1: аналітичне продовження. У нескінченному просторі розв'язки та рівняння аналітично продовжуються до $\mathbf{x} \in \mathbb{C}^n$, який перетворює коливальні хвилі в експоненціально затухаючі хвилі без відбиття за межами досліджуваної області.

Крок 2: перетворення координат. У нескінченному просторі виконується перетворення координат, щоб виразити $\mathbf{x} \in \mathbb{C}^n$ як функцію дійсної координати.

Крок 3: усічення області. Усікають область в нових дійсних координатах в межах області, де розв’язок розпадається. Оскільки розв’язок розпадається саме там, то за умови, що ми відсікаємо його на достатньо великій відстані, немає значення, яку граничну умову буде використано.

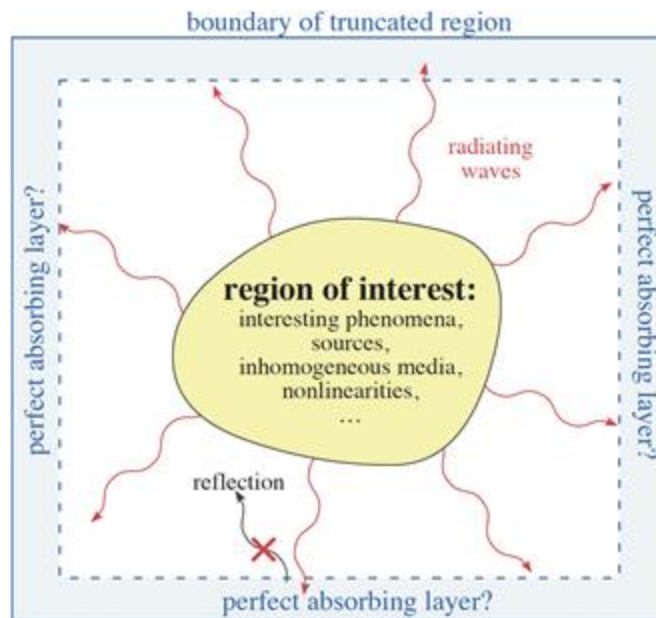


Рисунок 2.2 – Поглинаючий шар поглинає вихідні хвилі без відбиття від краю поглинача

За певних припущень аналітичний розв’язок у нескінченному просторі є суперпозицією плоских хвиль e^{ikx} для $x \in \mathbb{R}$, $k > 0$. Оскільки він є аналітичним, ми можемо продовжити оцінювати розв’язок при $x \in \mathbb{C}$. У цьому випадку, якщо уявна частина зростає лінійно, то розв’язок буде спадати експоненціально, як показано на рисунку 2.3. Причина в тому, що $e^{ik(\text{Re}\{x\}+i\text{Im}\{x\})} = e^{ik\text{Re}\{x\}}e^{-k\text{Im}\{x\}}$ експоненціально спадає при $k > 0$ зі збільшенням $\text{Im}\{x\}$. Наприклад, на рисунку 2.3 видно, що для $\text{Re}\{x\} > 0$ уявна частина лінійно зростає, що призводить до експоненціального спадання розв’язку. Зазначимо, що розв’язок для $\text{Re}\{x\} < 0$ залишається незмінним, а це означає, що PML діє не лише як поглинаючий матеріал, а й як невідбиваючий поглинаючий матеріал.

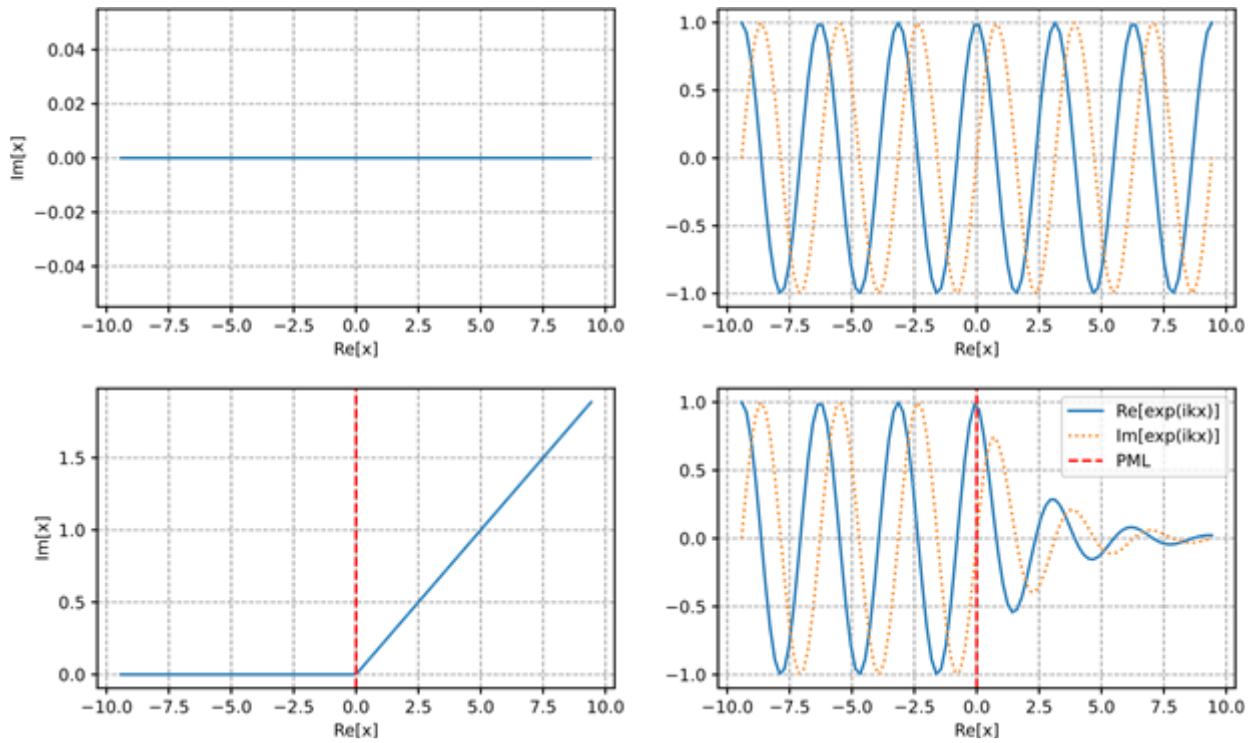


Рисунок 2.3 –

Зверху: дійсна та уявна частина вихідного коливального розв'язку e^{ikx} (праворуч) відповідає x вздовж дійсної осі в комплексній площині (ліворуч).

Внизу: Замість цього ми можемо оцінити цю аналітичну функцію вздовж деформованого контуру в комплексній площині: тут (ліворуч) ми деформуємо його для збільшення вздовж уявної осі для $\text{Re}\{x\} > 0$. Розв'язок e^{ikx} (праворуч) є незмінним для $\text{Re}\{x\} < 0$, але експоненціально спадає для $\text{Re}\{x\} > 0$, де контур деформовано, що відповідає "поглинаючій" області.

Щоб забезпечити перетворення координат, позначимо розширення $x \in \mathbb{R}$ як $\tilde{x} \in \mathbb{C}$ і залишимо позначення x для дійсної частини. Таким чином, маємо, що

$$\tilde{x}(x) = x + if(x), \quad (2.11)$$

де $f: \mathbb{R} \rightarrow \mathbb{R}$ – це спосіб перетворення уявної частини.

Пряме використання комплексної координати \tilde{x} є незручним, тому ми робимо заміну змінних, щоб повернутися до дійсних координат. Для

повернення достатньо змінити похідні у вигляді $\partial\tilde{x} = \left(1 + i \frac{df}{dx}\right) \partial x$.

Дотримуючись [3], зручно позначити $\frac{df}{dx} = \frac{\sigma_x(x)}{\omega}$ для деякої функції

$\sigma_x : \mathbb{R}^d \rightarrow \mathbb{R}$, щоб отримати незалежне від частоти поглинання. Тоді весь процес може бути концептуально підсумований єдиним перетворенням вихідного

$$\text{диференціального рівняння } \frac{\partial}{\partial\tilde{x}} \rightarrow \frac{1}{1 + i \frac{\sigma_x(x)}{\omega}} \frac{\partial}{\partial x}.$$

Після виконання PML-перетворення вихідної задачі розв'язки залишаються незмінними в розглядуваній області і експоненціально затухають у зовнішніх областях. Це означає, що ми можемо усікати обчислювальну область при будь-якому достатньо великому x .

2.4 Алгоритм GMRES

Опишемо алгоритм GMRES з передумовою [11]. Розглянемо систему (2.3) для фіксованого $\hat{\mathbf{c}}$ і позначимо $A = A(\hat{\mathbf{c}})$. Тоді алгоритм GMRES з правильною передумовою ґрунтується на розв'язанні

$$AM(\mathbf{y}) = \hat{\boldsymbol{\rho}}, \quad \hat{\mathbf{u}} = M(\mathbf{y}), \quad (2.12)$$

де $M : \mathbb{C}^n \rightarrow \mathbb{C}^n$ – оператор-передумова, тобто він певним чином апроксимує A^{-1} . Нову змінну \mathbf{y} не потрібно явно обчислювати, що продемонстровано у алгоритмі GMRES. Передумова застосовується до векторів базису Крилова V_m у рядках 3 та 11 алгоритму.

Зауважимо, що оператор $M : \mathbb{C}^n \rightarrow \mathbb{C}^n$ у (2.12) використано в узагальненому вигляді, у припущенні, що він може бути лінійним (тоді $M(\mathbf{y}) = M\mathbf{y}$ для $M \in \mathbb{C}^{n \times n}$) або нелінійним, що не є класичним варіантом,

представленим у [11]. Перша причина полягає в тому, що в цій роботі ми використовуємо нелінійну нейронну мережу як передумову, що не змінює алгоритм, за винятком рядків 3 та 11, де замість множення матриць записане загальне операторне застосування. Друга причина полягає в тому, щоб уникнути плутанини з Flexible GMRES, яка виникає при спробі проаналізувати причини використання Flexible GMRES, коли M є фіксованим, але нелінійним. У Flexible GMRES поведінка передумови змінюється з ітераціями, і алгоритм потребує додаткової матриці Z_m для зберігання модифікованого базису простору пошуку. Оскільки нам не потрібна Z_m , тому що M фіксований (але, можливо, нелінійний), ми не використовуватимемо Flexible GMRES.

2.5 Огляд досліджень за темою роботи

Розглянемо дослідження за темою роботи. Відправною точкою стала робота [13], де автори навчають нейронну мережу U-net розв'язувати двовимірну задачу Гельмгольца. Ідея авторів полягає в тому, щоб отримати нейронну мережу, яку можна використовувати в ітераційній схемі з фіксованою точкою

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \text{U-net}_\theta(\mathbf{u}_k, \mathbf{r}_k), \quad (2.13)$$

де $\mathbf{u}_k \in \mathbb{C}^n$ – k -те наближення дискретизованого розв'язку;

$\mathbf{r}_k = A(\mathbf{c})\mathbf{u}_k - \mathbf{p}$ – дискретизований залишок для деякого $\mathbf{c} \in \mathbb{R}_+^n$, що є параметрами нейронної мережі;

$$\mathbf{p} \in \mathbb{R}^n.$$

З огляду на це, перше питання, яке ми ставимо, полягає в тому, чи можемо ми використовувати мережу U-net як передумову. Результати, отримані в [14], показують, що це можливо, що також підтверджується в розділі 4. Обидва варіанти дають приблизно однакові результати:

$$M(\mathbf{v}_k) := \text{U-net}_\theta(\mathbf{u}_k, \mathbf{v}_k), \quad M(\mathbf{v}_k) := \text{U-net}_\theta(0, \mathbf{v}_k). \quad (2.13)$$

Ця передумова значно зменшує кількість ітерацій GMRES, але водночас збільшує обчислювальні витрати на кожну ітерацію. Незважаючи на те, що обчислення M добре розпаралелюється, що зменшує час обчислень, порівняння абсолютного енергоспоживання залишається нез'ясованим.

Друге питання, яке ми ставимо, полягає в тому, чи можна спростити процес навчання мережі U-net і її архітектуру. Процес навчання U-net в [13] є дуже нестабільним через диференціювання множини ітерацій з фіксованою точкою і вимагає деяких "магічних" констант, які не пояснюються в роботі. Архітектура мережі U-net також містить прихований стан, який теоретично допомагає вловити залежність між ітераціями, але робить архітектуру більш складною для розуміння та реалізації. Щоб вирішити ці проблеми, ми спробували спростити всі частини і апроксимувати A^{-1} безпосередньо за допомогою

$$\min_{\theta} \|X - AU\text{-net}_\theta(X)\|_2^2, \quad (2.14)$$

де $X \in \mathbb{C}^{N \times n}$, а N – розмір набору даних.

Мережа U-net, розглянута в розділі 4, не є єдиним типом архітектури, який може бути використаний. Було також розглянуто DeepONet та FNO, які є нейронними операторами [4]. Хоча ідея (2.14) проста і легко реалізується, питання розподілу вхідних даних X не є очевидним, оскільки ми не знаємо заздалегідь структуру \mathbf{v}_j на кроці 3 Алгоритму GMRES.

2.6 Навчання без вчителя оберненого оператора

Спробуємо апроксимувати обернений до лінійного оператора $A(\mathbf{c}) \in \mathbb{C}^{n \times n}$ для $\mathbf{c} \in \mathbb{R}_+^n$ за допомогою нейронної мережі. Мета полягає в тому, щоб знайти

архітектуру, яка є адаптивною до розміру сітки n і узагальненою до змінних $\mathbf{c} \in \mathbb{R}_+^n$. В ідеалі, після навчання на фіксованій області, нейронну мережу можна було б застосувати до більшої області, приймаючи \mathbf{c} як вхідні дані.

Ми почали дослідження з фізично-інформованої нейронної мережі (PINN) [9], яка призначена для апроксимації розв'язку будь-якого диференціального рівняння і навчається без вчителя. Результатом є навчена нейронна мережа, яка задовольняє диференціальному рівнянню і може бути оцінена в будь-якій точці області. Хоча нейронну мережу потрібно перенавчати щоразу, коли змінюються параметри рівняння або граничні умови, це був перший крок, який дав корисні ідеї, що представлені нижче.

Наступним кроком у дослідженні стало розширення PINN під назвою DeepONet [6], у якому параметри рівняння включено до вхідних даних моделі. Однак цей підхід все ще залежить від дискретизації, що означає перенавчання моделі при зміні області.

Наступним кроком дослідження став нейронний оператор [4]. Ідея, як і у випадку DeepONet, полягає у використанні параметрів рівняння, граничних умов та інших змінних як вхідних даних для нейронної мережі. Головна відмінність полягає в тому, що нейронний оператор є інваріантним до дискретизації. Існує кілька відомих архітектур, таких як нейронний оператор Фур'є [5] та U-net [10], які будуть розглянуті нижче.

2.6.1 Фізично-інформована нейронна мережа

Розглянемо загальне диференціальне рівняння

$$L(u) = f, \quad (2.15)$$

де L – диференціальний оператор.

Нейронна мережа u_θ прагне апроксимувати u так, щоб норма залишку $\|L(u_\theta) - f\|$ була мінімальною відносно параметрів моделі θ . Таким чином, нейронна мережа стає фізично-інформованою, коли мінімізуються втрати на основі диференціального рівняння (рисунок 2.4). У нашому випадку працюємо з оператором Гельмгольца $L = \nabla^2 + k^2$, де L можна розглядати як неперервний або дискретизований оператор. Для обчислення норми залишку в першому випадку можна використати алгоритм автоматичного диференціювання, який дозволяє обчислити $L(u_\theta)(x)$ для будь-якого x . Інший варіант – обчислити u_θ на фіксованій сітці і отримати \hat{u}_θ , яке потім використовується для апроксимації оператора Гельмгольца.

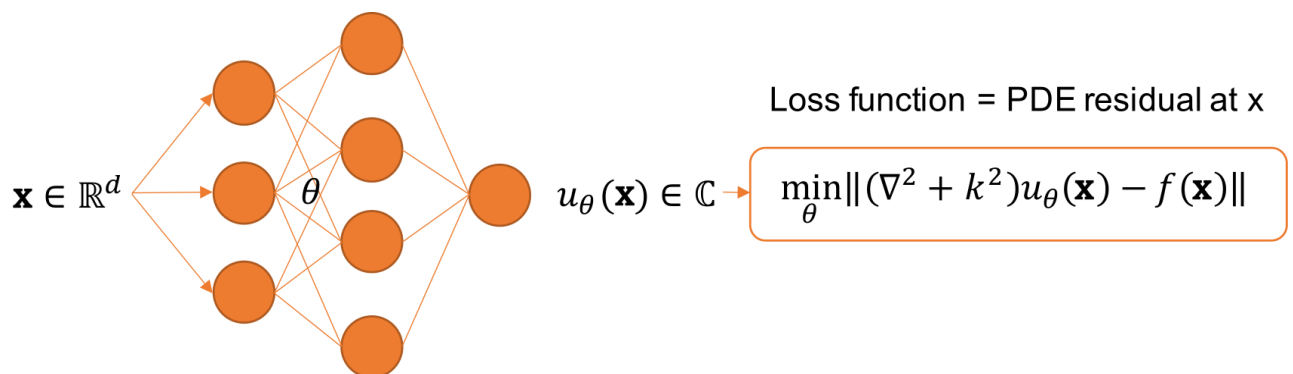


Рисунок 2.4 – Фізично-інформована концепція нейронної мережі, застосована до оператора Гельмгольца

Множення виконується з векторами, отриманими після повнозв'язного шару з синусоїдальною активаційною функцією. Періодична активація є вирішальною через періодичну природу розв'язку. Можна змінити активацію і перезапустити навчання, щоб побачити різницю.

Як зазначалося вище, цей підхід не є корисним для побудови оператора передумови через перенавчання при зміні параметрів рівняння, а також через те, що немає можливості подати на вхід вектор. Однак ми навчилися використовувати фізичні втрати для навчання без вчителя нейронної мережі, що буде корисним у подальших дослідженнях.

2.6.2 DeepONet

Розглянемо наступний крок у навчанні операторів, який полягає у розширенні фізично-інформованих нейронних мереж. Ідея роботи [6], де автори представляють DeepONet, полягає в апроксимації диференціального оператора L нейронною мережею замість апроксимації розв'язку u диференціального рівняння, як в PINN. Концепція DeepONet представлена на рисунку 2.5.

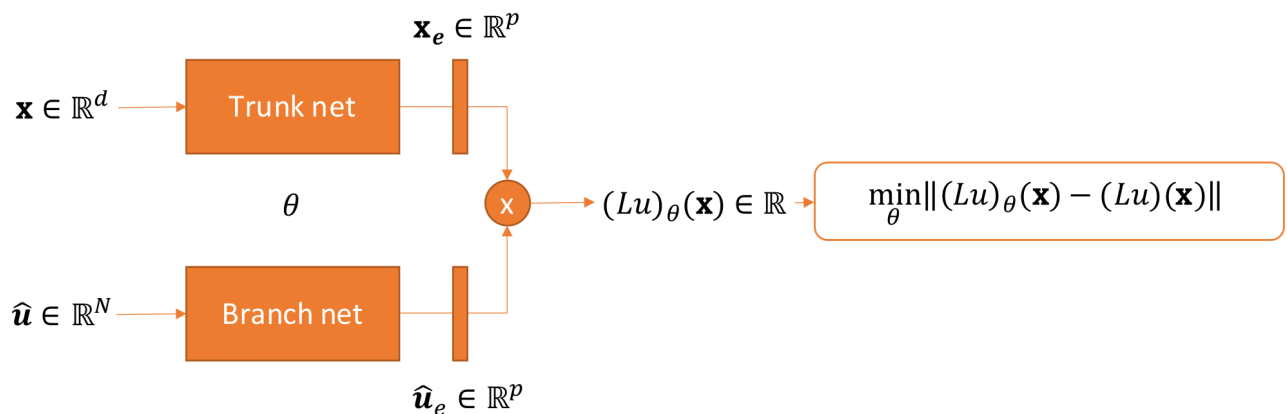


Рисунок 2.5 – Концепція DeepONet

На рисунку 2.5 через $(Lu)_\theta$ позначається нейронна мережа з параметрами θ , що навчаються, яка апроксимує (Lu) , а p – це розмір векторного вкладання. Позначення Trunk net та Branch net взято з [6].

Адаптуємо цю ідею до нашої задачі. Нам необхідно наблизити обернений оператор Гельмгольца за допомогою нейронної мережі, щоб використовувати його як передумову. Отже, маємо $L = (\nabla^2 + k^2)^{-1}$, а на вхід Branch net подаємо $\hat{\rho}$. Потім ми переходимо від навчання з вчителем до навчання без вчителя, застосовуючи обернену до L^{-1} величину до залишку. Позначимо $(L\rho)_\theta(\mathbf{x}) = u_\theta(\mathbf{x})$. Тоді:

$$\begin{aligned} L^{-1}((L\rho)_\theta(\mathbf{x}) - (L\rho)(\mathbf{x})) &= (L^{-1}u_\theta)(\mathbf{x}) - \rho(\mathbf{x}) = \\ &= (\nabla^2 + k^2)u_\theta(\mathbf{x}) - \rho(\mathbf{x}). \end{aligned} \quad (2.16)$$

Також, оскільки вхідні вектори для передумови M є комплексними в рядку 3 Алгоритму GMRES, потрібно використовувати $\hat{\rho} \in \mathbb{C}^N$. Модифікацію початкової ідеї показано на рисунку 2.6. Мережу Trunk net у ній пропонується для спрощення замінити на MMLP як для PINN, а Branch net – на MLP.

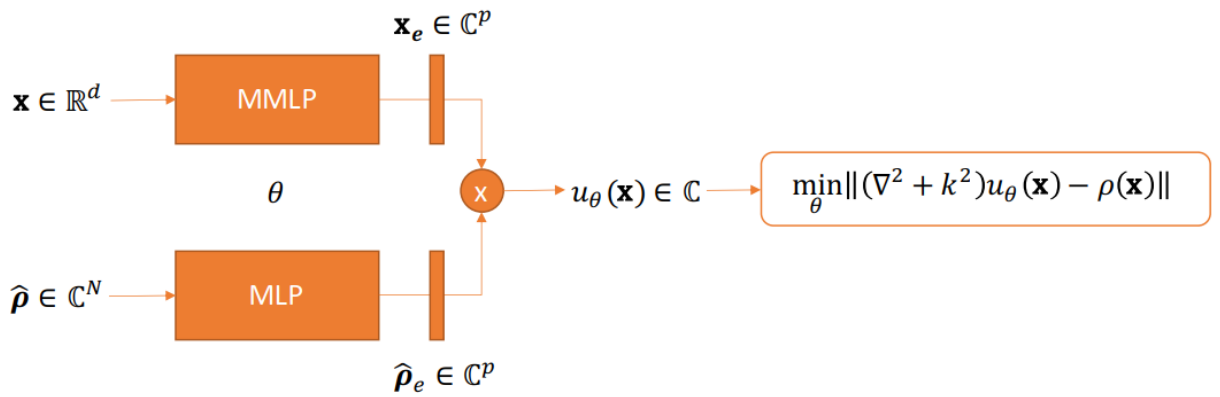


Рисунок 2.6 – Модифікована концепція DeepONet

Розглянемо обмеження моделі. Помічено, що якщо використовувати DeepONet як передумову для $p < N$ у повних ітераціях GMRES, то вектори з передумовного базису Крилова будуть лінійно залежними, що є першим обмеженням.

Зауваження. Нехай $X = \{\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(N-1)}\} \in \mathbb{R}^{d \times N}$ – множина точок дискретизації в області, а $\hat{\rho} \in \mathbb{C}^N$ – фіксована права частина. Нехай $X_e = \{\mathbf{x}_e^{(0)}, \dots, \mathbf{x}_e^{(N-1)}\} \in \mathbb{R}^{p \times N}$ – множина вкладень для X , нехай $p < N$, і нехай $\hat{\rho}_e \in \mathbb{C}^p$ – вкладення для фіксованої правої частини $\hat{\rho}$. Нехай $\hat{\mathbf{u}}_\theta = \{u_\theta^{(0)}, \dots, u_\theta^{(N-1)}\} \in \mathbb{C}^N$ – множина оцінок нейронної мережі на X для фіксованого $\hat{\rho}$. Тоді розмірність лінійного простору, якому належить $\hat{\mathbf{u}}_\theta$, не перевищує p .

2.6.3 Нейронний оператор Фур'є

Розглянемо концепцію нейронних операторів для вивчення диференціальних операторів відповідно до [4, 5]. Вона являє собою відображення між нескінченновимірними функціональними просторами, але на практиці розглядається відображення між дискретними просторами. Архітектура нейронних операторів дозволяє їм бути моделями, інваріантними до дискретизації. Це означає, що після навчання на фіксованій сітці їх можна застосовувати до більшої або меншої сітки. Добре відомою інваріантною до дискретизації нейронною мережею є згорткова нейронна мережа, де термін «сітка» означає пікселі зображення.

Архітектура нейронних операторів складається з L шарів, кожен з яких включає згорткову частину, зміщення та нелінійну функцію активації. Найважливішою частиною є згортка, яка забезпечує інваріантність дискретизації. Цю операцію можна виконати різними способами. Наприклад, показано, що множення в частотній області еквівалентне згортці у просторовій області [7]. Цей факт використовується в архітектурі, яка розглядається нижче.

Нейронний оператор Фур'є (FNO) – це нейромережева архітектура, яка поєднує потужність нейронних мереж з математичним апаратом аналізу Фур'є. Він призначений для розв'язання складних задач математичного моделювання та імітаційного моделювання в науці та інженерії. FNO використовує можливості репрезентативного навчання нейронних мереж та властивості спектрального аналізу перетворень Фур'є. Нижче наведено опис його роботи.

Етап 1: навчання за допомогою нейронних мереж. FNO починається з розпізнавання основної проблеми за допомогою нейронної мережі. Воно передбачає вилучення значущих особливостей і закономірностей з даних. Нейронна мережа вчиться фіксувати складні взаємозв'язки, присутні в даних, які можуть відображати просторово-часову поведінку або інші складні явища.

Етап 2: використання перетворення Фур'є. Після того, як нейронна мережа зафіксувала патерни, FNO застосовує перетворення Фур'є –

математичний метод, який розкладає функцію на набір синусоїдальних базисних функцій. Це перетворення зміщує точку зору з просторово-часової області в частотну. Таким способом FNO намагається знайти простіші, більш структуровані представлення вивчених закономірностей.

Етап 3: вивчення правил взаємодії: У частотній області FNO має на меті визначити правила взаємодії між найвищими частотами або режимами. Ці правила взаємодії можна розглядати як основні керуючі рівняння, які диктують еволюцію системи. Нейронна мережа вивчає, як ці частоти взаємодіють, ефективно вивчаючи математичні оператори, що описують динаміку системи.

Крок 4: зворотне перетворення Фур'є. Після того, як правила взаємодії в частотній області вивчені, FNO виконує зворотне перетворення Фур'є. Цей крок переводить спрощені, структуровані представлення назад у вихідну просторово-часову область, в результаті чого отримується прогнозований розв'язок, який відображає складну поведінку системи.

Таким чином, нейронний оператор Фур'є заповнює прогалину між нейронними мережами та методами математичного аналізу. Він об'єднує можливості навчання нейронних мереж зі спектральними можливостями аналізу Фур'є, щоб забезпечити точні та ефективні в обчислювальному плані розв'язання складних задач математичного моделювання, що виникають в наукових дослідженнях та інженерних додатках. Як було зазначено вище, ключова ідея полягає в тому, що ми усікаємо високі частоти і застосовуємо до них лінійний оператор, що навчається, перш ніж повернутися до просторової області за допомогою оберненого перетворення Фур'є.

2.6.4 U-net

Архітектура U-Net – це тип згорткової нейронної мережі (CNN), спеціально розроблений для задач сегментації зображень у комп'ютерному зорі [10]. Архітектура U-Net названа на честь своєї U-подібної форми. Вона особливо ефективна для задач, де необхідно сегментувати зображення та

ідентифікувати певні об'єкти або області на ньому. Сегментація зображення передбачає присвоєння мітки кожному пікселю зображення, щоб віднести його до певного об'єкта або фону. Розглянемо складові частини U-Net.

1. Contracting Path (Encoder) – верхня частина U, що називається контуром скорочення або кодером. Вона складається з низки шарів згортки та об'єднання. Ці шари зменшують просторові розміри вхідного зображення, одночасно збільшуючи кількість каналів ознак. Це допомагає мережі захоплювати різні рівні інформації та виокремлювати ознаки в різних масштабах.

2. Bottleneck – шар внизу U, що називається «вузьким місцем». В ньому мережа вловлює найбільш абстрактні та консолідовані ознаки вхідного зображення.

3. Expansive Path (Decoder) – нижня частина U, що називається розширювальним контуром або декодером. Він відповідає за поступове збільшення вибірки елементів до початкового розміру зображення. Кожен крок у декодері передбачає збільшення вибірки карти об'єктів і поєднання її з картою об'єктів з відповідного шару з контуру скорочення за допомогою процесу, який називається «пропуск з'єднань».

4. Skip Connections – пропуск з'єднань: Пропускні з'єднання є важливим компонентом архітектури U-Net. Вони з'єднують карти об'єктів з контуру скорочення, з відповідними шарами розширювального контуру. Ці зв'язки допомагають зберегти дрібнозернисту просторову інформацію, яка може бути втрачена в процесі зменшення вибірки. Об'єднуючи інформацію з різних масштабів, мережа може краще локалізувати об'єкти і видавати точніші результати сегментації.

5. Final Layer – останній шар. Останній шар мережі зазвичай складається зі згорткового шару, який створює карту сегментації. На цій карті кожному пікселю зображення присвоюється мітка, яка вказує, до якого об'єкта або категорії він належить.

Сила архітектури U-Net полягає в її здатності обробляти як контекстну інформацію (через контур скорочення), так і точну локалізацію (через контур розширення). Пропускні з'єднання відіграють важливу роль у досягненні цього

балансу. U-Net широко використовується в аналізі медичних зображень, де такі завдання, як сегментація органів або пухлин на медичних знімках є критично важливими. Її гнучкість і ефективність у вирішенні завдань сегментації зробила її популярним вибором для різних додатків аналізу зображень за межами медичної галузі. Ця архітектура також може бути використана для побудови передумов.

Висновки за розділом 2

В розділі 2 було розглянуто декілька способів конструювання апроксимації оберненого диференціального оператора Гельмгольца за допомогою нейронних мереж, зокрема, фізично проінформованої нейронної мережі PINN, DeepONet та U-net.

Було встановлено, що нейронна мережа PINN може апроксимувати обернений оператор Гельмгольца тільки при фіксованих параметрах рівняння. Зі зміною параметрів потрібно знову тренувати модель, що є не бажаним результатом.

Нейронна мережа DeepONet має більше можливостей для створення узагальненого оберненого оператора, який не залежить від зміни параметрів рівняння. Проте концепція лінійної комбінації виходів двох підмереж має приховані недоліки, які накладають обмеження на розмірність лінійного простору вихідних даних, що є критичним для застосування ітеративного метода Крилова.

Нейронна мережа U-net потенційно має найкращі можливості для створення узагальненого оберненого оператора, який не залежить як від параметрів рівняння, так і може бути адаптований для різних сіток.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ

3.1 Мова програмування Python

Python є високорівневою, інтерпретованою мовою програмування загального призначення. Відомий своєю читабельністю і простотою синтаксису, Python намагається робити код якнайбільш зрозумілим. Python використовується як інтерпретована мова, що дозволяє виконувати код без попередньої компіляції. Код написаний на Python може використовуватися на різних платформах без змін, завдяки вбудованій портативності. Python підтримує об'єктно-орієнтоване програмування, дозволяючи створювати та використовувати класи та об'єкти. Python має велику стандартну бібліотеку, яка включає в себе різноманітні модулі і пакети для різних задач. Є багато сторонніх бібліотек і фреймворків, що розширюють можливості мови.

Python широко використовується у веб-розробці, аналізі даних, штучному інтелекті, наукових дослідженнях, автоматизації систем та багатьох інших галузях. Інтерактивність використовується через інтерпретатор Python, який дозволяє тестувати код один рядок за один раз. Jupyter Notebooks – популярний інструмент для виконання коду, вивчення даних та створення звітів.

Python надає безліч бібліотек і фреймворків для створення і навчання нейронних мереж. Серед найпопулярніших інструментів для роботи з нейронними мережами на Python є бібліотеки TensorFlow, PyTorch, Keras тощо. Вибір конкретного фреймворка визначається потребами розробника, його рівнем досвіду та уподобаннями.

Python має велику та активну спільноту розробників. Документація, форуми і багато інших ресурсів допомагають новачкам та досвідченим розробникам. Багато великих технологічних компаній використовують Python у своїх проектах через його ефективність та продуктивність.

Python – це мова, яка стала популярною завдяки своїй простоті, читабельності та можливостям, які підтримують різні галузі програмування.

3.2 Дискретне перетворення Фур'є

Нехай $f = \{f_0, \dots, f_{N-1}\}$ – послідовність дійсних або комплексних чисел.

Тоді дискретне перетворення Фур'є f' визначається як

$$\mathcal{F}\{f\}(n) = \sum_{k=0}^{N-1} f_k e^{-i2\pi nk/N}, n \in \{0, \dots, N-1\}, \quad (3.1)$$

а обернене перетворення дорівнює

$$\mathcal{F}^{-1}\{f\}(k) = \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{i2\pi nk/N}, k \in \{0, \dots, N-1\}. \quad (3.2)$$

Приклад коду на Python для виконання цих перетворень наведено у додатку А.

3.3 Обчислення похідних

Розглянемо обчислення похідної засобами Python. Як тестову функцію для обчислення похідної ми обрали гауссову функцію, оскільки вона задовольняє умовам, за яких може бути застосована теорема 2.1, тобто є абсолютно інтегрованою функцією. Приклад коду на Python для обчислення похідної наведено у додатку А.

Ми можемо використовувати цей інструмент для обчислення оператора Гельмгольца на послідовності дискретизованих точок, які задовольняють

умовам теореми 2.1. Приклад коду на Python для дискретизації оператора Гельмгольца наведено у додатку А.

3.4 Алгоритм GMRES

Розглянемо алгоритм GMRES з правильною передумовою [11], відповідно до того, як він був описаний у пункті 2.4. Нижче наведено алгоритмізовану послідовність дій цього методу.

Алгоритм GMRES.

1: Обчислити $\mathbf{r}_0 = \hat{\mathbf{p}} - \mathbf{A}\mathbf{u}_0$, $\beta = \|\mathbf{r}_0\|^2$ та $\mathbf{v}_1 = \frac{\mathbf{r}_0}{\beta}$.

2: Для $j = 1, \dots, m$:

3: обчислити $\mathbf{w} := \mathbf{A}\mathbf{M}(\mathbf{v}_j)$;

4: для $i = 1, \dots, j$ виконати:

5: $h_{i,j} := \langle \mathbf{w}, \mathbf{v}_i \rangle$;

6: $\mathbf{w} := \mathbf{w} - h_{i,j} \mathbf{v}_i$;

7: **кінець циклу**;

8: обчислити $h_{j+1,j} = \|\mathbf{w}\|_2$ та $\mathbf{v}_{j+1} = \frac{\mathbf{w}}{h_{j+1,j}}$;

9: визначити $\mathbf{V}_m := [\mathbf{v}_1, \dots, \mathbf{v}_m]$, $\bar{\mathbf{H}}_m = \{h_{i,j}\}_{1 \leq i \leq j+1; 1 \leq j \leq m}$;

10: **кінець циклу**.

11: Обчислити $\mathbf{q}_m = \arg \min_{\mathbf{q}} \|\beta \mathbf{e}_1 - \bar{\mathbf{H}}\mathbf{q}\|_2$, та $\mathbf{u}_m = \mathbf{u}_0 + \mathbf{M}(\mathbf{V}_m \mathbf{q}_m)$.

12: Якщо тест на зупинку виконано, то **Stop**,

13: інакше встановити $\mathbf{u}_0 = \mathbf{u}_m$ і повернутись на 1.

14: Кінець **Якщо**.

Лістинг програми, яка реалізує даний алгоритм мовою Python, наведено у додатку А.

Висновки за розділом 3

В цьому розділі ми розглянули переваги високорівневої мови програмування Python, розглянули програмні реалізації обчислення оператора Гельмгольца, диференціювання за допомогою спектрального перетворення Фур'є, а також алгоритму GMRES для розв'язання системи лінійних рівнянь. Лістинг програм наведено у додатках А та Б.

Код для навчання нейронних мереж не приводиться через великий об'єм тексту.

4 РЕЗУЛЬТАТИ ОБЧИСЛЮВАЛЬНОГО ЕКСПЕРИМЕНТУ ТА ЇХ АНАЛІЗ

4.1 Обчислювальний експеримент

Проведемо обчислювальні експерименти з використанням досліджуваних у кваліфікаційній роботі нейронних мереж.

Візуалізуємо норму залишків для GMRES з попередньою передумовою та без неї на рисунку 4.1.

Можна побачити швидку збіжність з попередньою умовою DeepONet порівняно з повним GMRES без попередньої умови. Однак цей результат отримано лише для $p = 2N$, де $N = 128$ – кількість точок дискретизації. Якщо ми спробуємо зменшити p , то побачимо проблеми у збіжності, що частково можна пояснити зауваженням, наведеним наприкінці п. 2.6.2.

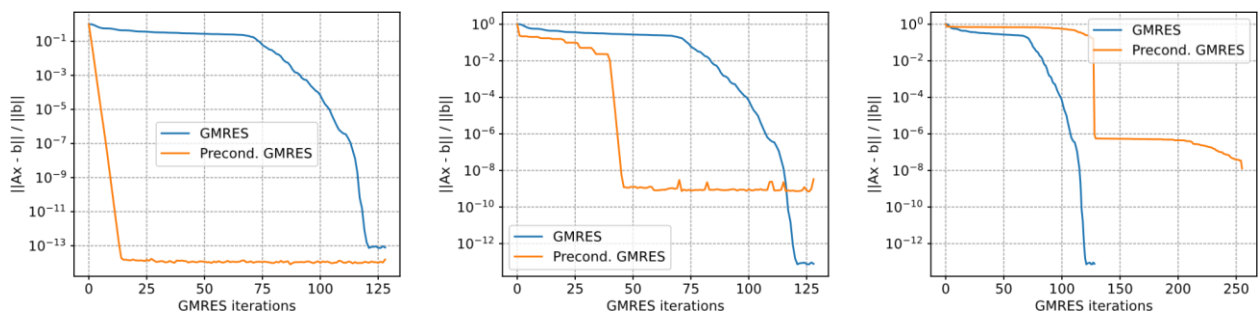


Рисунок 4.1 Непередумовлені та передумовлені DeepONet норми залишків GMRES під час ітерацій для $p \in \{2N, N, N/2\}$, $N = 128$ та 4 шарів у мережі Trunk net

Інша проблема виникає, коли ми зменшуємо кількість шарів для Trunk net в DeepONet. Відповідний результат проілюстровано на рисунку 4.2. Емпіричні спостереження сингулярних значень базису лінійного простору, оброблені через DeepONet, мають розподіл, який робить вектори майже лінійно залежними в 32-бітній арифметиці.

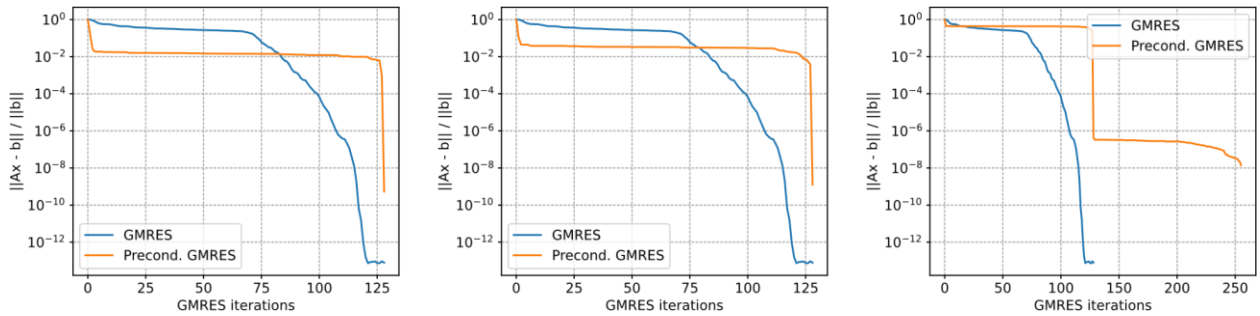


Рисунок 4.2 Непередумовлені та передумовлені DeepONet

норми залишків GMRES під час ітерацій

для $p \in \{2N, N, N/2\}$, $N = 128$ та 2 шарів у мережі Trunk net

Можна зробити висновок, що подальша робота з моделлю DeepONet та її застосування для прискорення GMRES недоцільна. Після проведених експериментів з архітектурою FNO таких проблем не спостерігається через принципіально інший підхід до побудови моделі. На рисунку 4.3 зображені норми залишків GMRES, попередньо обумовлені FNO, що навчалась на сітці 64 на 64 (помаранчева крива) і застосовувалася як попередня умова до сітки 512 на 512 (фіолетова крива). Ми можемо зробити висновок, що ця архітектура може бути застосована для прискорення алгоритму, проте узагальнююча здатність сильно деградує зі збільшенням вузлів в сітці.

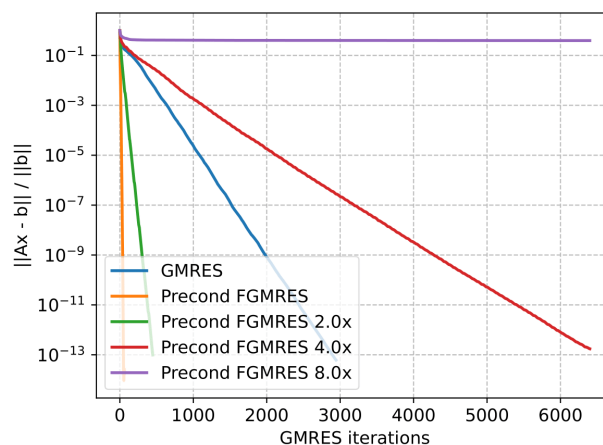


Рисунок 4.3 Норми залишків GMRES, попередньо обумовлені FNO,

що навчалась на сітці 64 на 64 (помаранчева крива)

і застосовувалася як попередня умова до сітки 512 на 512 (фіолетова крива)

Далі, спробуємо застосувати архітектуру U-net для прискорення алгоритму. Перевага U-Net полягає в її здатності обробляти як контекстну інформацію (через контур скорочення), так і точну локалізацію (через контур розширення). Пропускні з'єднання відіграють важливу роль у досягненні цього балансу. Ця архітектура може бути використана для побудови передумов.

На рисунку 4.4 ми чітко можемо спостерігати, що, на відміну від FNO, узагальнююча здатність U-net деградує набагато менше, що дає змогу ефективно застосовувати її як оператор передумови в алгоритмі GMRES.

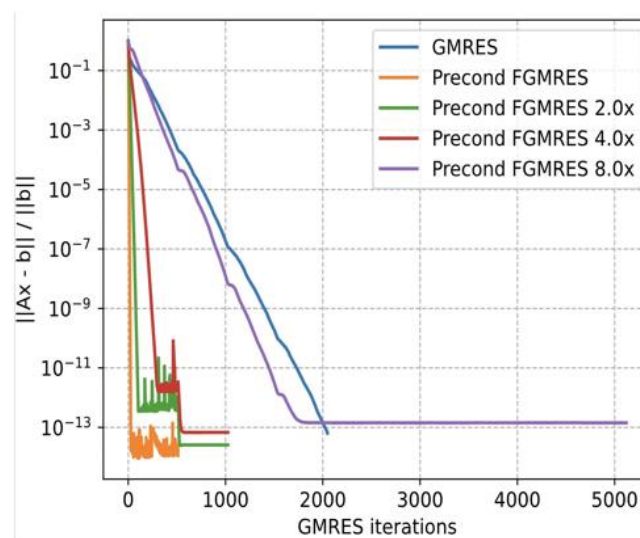


Рисунок 4.4 Норми залишків GMRES, попередньо обумовлені U-net, що навчалася на сітці 64 на 64 і застосовувалася як попередня умова до сітки 512 на 512

Висновки за розділом 4

Проведені чисельні експерименти показують, що U-net є найкращою нейронною мережею для використання в якості оператора передумови при використанні методу лінійних підпросторів Крылова для розв'язання систем лінійних рівнянь великих розмірностей.

ВИСНОВКИ

У роботі було проаналізовано кілька методів побудови наближення оберненого диференціального оператора Гельмгольца за допомогою нейронних мереж, таких як фізично проінформована нейронна мережа PINN (Physics Informed Neural Network), DeepONet та U-net. Виявлено, що нейронна мережа PINN може апроксимувати обернений оператор Гельмгольца лише за фіксованих параметрів рівняння, що вимагає повторного навчання моделі при зміні параметрів, а це є небажаним наслідком.

Нейронна мережа DeepONet має більше можливостей для створення узагальненого оберненого оператора, який не залежить від зміни параметрів рівняння. Однак концепція лінійної комбінації виходів двох підмереж має приховані недоліки, що обмежують розмірність лінійного простору вихідних даних, що є критичним для застосування ітеративного метода Крилова.

Нейронна мережа U-net потенційно має кращі можливості для створення узагальненого оберненого оператора, який не залежить від параметрів рівняння і може адаптуватися до різних сіток. Проведені чисельні експерименти продемонстрували, що U-net є найкращою нейронною мережею для використання як оператора передумови при застосуванні методу лінійних підпросторів Крилова для розв'язання систем лінійних рівнянь великих розмірностей.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Berenger J. P. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114 (2), 10 1994.
2. Gilliam D. Mathematics 5342 discrete Fourier transform. 11, 1998. 16 p.
3. Johnson S. Notes on perfectly matched layers (pmls). 08 2021. 18 p. (<https://arxiv.org/pdf/2108.05348.pdf>)
4. Kovachki N., Li Z., Liu B., Azizzadenesheli K., Bhattacharya K., Stuart A., Anandkumar A. Neural operator: Learning maps between function spaces. 08 2021. 52 p. (<https://arxiv.org/pdf/2108.08481v1.pdf>)
5. Li Z., Kovachki N., Azizzadenesheli K. Liu B., Bhattacharya K., Stuart A., Anandkumar A. Fourier neural operator for parametric partial differential equations. 10 2020. 16 p. (<https://arxiv.org/pdf/2010.08895v1.pdf>)
6. Lu L., Jin P., Pang G., Zhang Z., Karniadakis G.. Learning nonlinear operators via deepnet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 03 2021. P. 218-229.
7. McGillem C. D., Cooper G. R.. Continuous and discrete signal and system analysis. New York : Holt Rinehart & Winston. 1984. 400 p.
8. Nocedal J., Wright S. J. Numerical optimization pages. Springer Series in Operations Research and Financial Engineering. London : Springer Nature, 2006. 664 p.
9. Raissi M., Perdikaris P., Karniadakis G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 2019. P. 686–707.
10. Ronneberger O., Fischer P., Brox T.. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of LNCS, p. 234–241. Springer, 2015. (<https://arxiv.org/pdf/1505.04597.pdf>).
11. Saad Y. Iterative Methods for Sparse Linear Systems. *Society for Industrial and Applied Mathematics*, 2003. 567 p. (https://www-users.cse.umn.edu/~saad/IterMethBook_2ndEd.pdf).

12. Smith J. O. III. Mathematics of the Discrete Fourier Transform (DFT): With Audio Applications. North Charleston : W3K Publishing. 2007. 322 p.
13. Stanziola A., Arridge S. R., Cox B. T., Treeby B. E. A Helmholtz equation solver using unsupervised learning: Application to transcranial ultrasound. *Journal of Computational Physics*, 441:110430, 2021. 23 p.
14. Школін О. В. Застосування згорткових нейронних мереж для розв'язання диференціального рівняння Гельмгольца // 27-й Міжнародний молодіжний форум «Радіоелектроніка і молодь у ХХІ столітті». Зб. матеріалів форуму. Т. 7. Харків: ХНУРЕ. 2023. С.231-232.