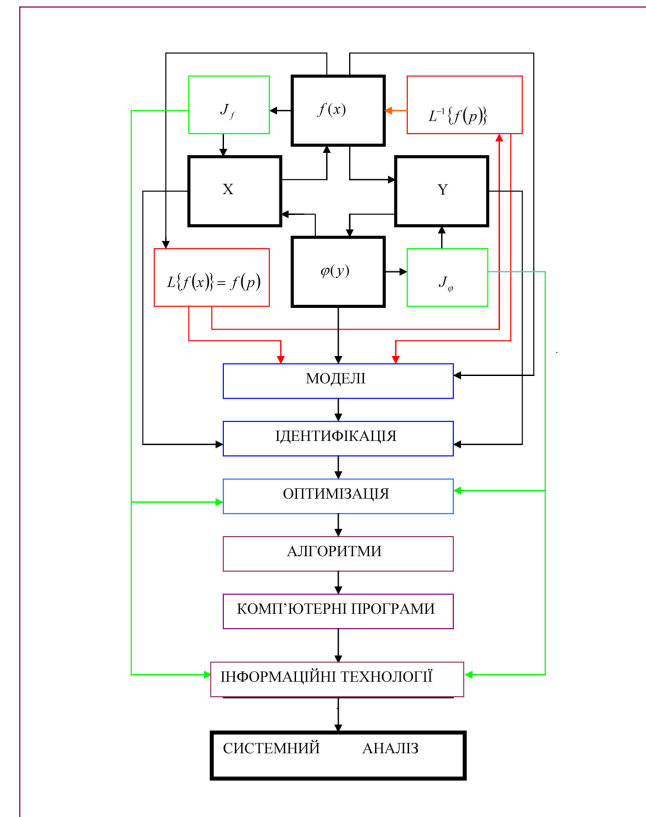


Б. І. Мокін, В. Б. Мокін, О. Б. Мокін

## ФУНКЦІОНАЛЬНИЙ АНАЛІЗ, адаптований до прикладних задач в галузі інформаційних технологій



Міністерство освіти і науки України  
Вінницький національний технічний університет

**Б. І. Мокін, В. Б. Мокін, О. Б. Мокін**

**ФУНКЦІОНАЛЬНИЙ АНАЛІЗ,  
адаптований до прикладних задач  
в галузі інформаційних технологій**

**Навчальний посібник**

Вінниця  
ВНТУ  
2020

**УДК 517.98**

**М74**

Рекомендовано Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України як навчальний посібник для студентів закладів вищої освіти, що спеціалізуються в галузі інформаційних технологій (протокол № 7 від 27.02.2020 р.).

Рецензенти:

**В. Я. Данилов**, доктор технічних наук, професор (НТУУ «КПІ ім. Сікорського»)

**В. І. Клочко**, доктор педагогічних наук, професор (ВНТУ)

**О. С. Макаренко**, доктор фіз.-мат. наук, професор (НТУУ «КПІ ім. Сікорського»)

**Мокін, Б. І.**

М74 Функціональний аналіз, адаптований до прикладних задач в галузі інформаційних технологій : навчальний посібник / Б. І. Мокін, В. Б. Мокін, О. Б. Мокін. – Вінниця : ВНТУ, 2020. – 192 с.

ISBN 978-966-641-796-4

В навчальному посібнику викладені основи функціонального аналізу, адаптованого до розв'язання прикладних задач в галузі інформаційних технологій, що використовують математичні моделі, закони керування та методи оптимізації.

Навчальний посібник рекомендується для студентів, що спеціалізуються в ІТ-галузі за спеціальностями 124 – «Системний аналіз» та 126 – «Інформаційні системи та технології».

**УДК 517.98**

## ЗМІСТ

<b>ВСТУП</b> .....	5
<b>Розділ 1 МНОЖИНИ І МЕТРИЧНІ ПРОСТОРИ ТА ЇХ ХАРАКТЕРИСТИКИ</b> .....	12
1.1 Множини, підмножини та їх характеристики .....	12
1.2 Метричні простори та їх характеристики .....	17
1.3 Лінійний і банахів метричні простори та їх характеристики .....	22
1.4 Завдання для самоперевірки .....	24
<b>Розділ 2 ГІЛЬБЕРТОВІ ПРОСТОРИ, ЇХ ХАРАКТЕРИСТИКИ ТА АПРОКСИМАЦІЯ ФУНКЦІЙ В НИХ</b> .....	25
2.1 Гільбертові простори та їх характеристики.....	25
2.2 Ортонормовані підмножини у гільбертових просторах.....	27
2.3 Апроксимація безперервних функцій в гільбертових просторах на основі ортогональних послідовностей.....	30
2.4 Методика Грама – Шмідта побудови систем ортонормованих послідовностей в гільбертових просторах.....	33
2.5 Завдання для самоперевірки .....	38
<b>Розділ 3 МІРА ЖОРДАНА І ЛЕБЕГА ДЛЯ МНОЖИН І ПРОСТОРІВ ТА ІНТЕГРАЛИ РІМАНА, СТІЛТЬЄСА І ЛЕБЕГА</b> .....	39
3.1 Міра Жордана для множин і просторів.....	39
3.2 Міра Лебега для множин і просторів .....	43
3.3 Інтеграл Рімана.....	47
3.4 Інтеграл Стільєса .....	48
3.5 Інтеграл Лебега.....	51
3.6 Завдання для самоперевірки .....	53
<b>Розділ 4 ФУНКЦІОНАЛИ, ЇХ ВАРІАЦІЇ ТА МЕТОДИ ПОШУКУ БЕЗУМОВНИХ ЕКСТРЕМУМІВ</b> .....	54
4.1 Функціонали, що використовуються в прикладних задачах системного аналізу .....	54
4.2 Екстремум функції .....	55
4.3 Суть варіаційного числення. Рівняння Ейлера.....	61
4.4 Достатні умови існування екстремуму функціонала на екстремалі .....	67
4.5 Узагальнення найпростішої задачі варіаційного числення для пошуку екстремалей з рухомими кінцями.....	71
4.6 Задачі, що приводять до екстремалей зі зломами.....	75
4.7 Пошук екстремумів функціоналів, що пов'язують між собою декілька невідомих функцій та їх перших похідних.....	79
4.8 Пошук екстремумів функціоналів, що залежать від старших похідних невідомої функції .....	82
4.9 Завдання для самоперевірки .....	84
<b>Розділ 5 ВАРІАЦІЙНІ МЕТОДИ ДОСЛІДЖЕННЯ ФУНКЦІОНАЛІВ НА УМОВНИЙ ЕКСТРЕМУМ</b> .....	86
5.1 Метод невизначених множників Лагранжа як найбільш загальний для дослідження функціоналів на умовний екстремум.....	86
5.2 Ізопериметрична задача пошуку екстремалей функціоналів.....	92
5.3 Задача Майєра .....	95
5.4 Пошук екстремалей функціоналів за наявності обмежень у вигляді нерівностей .....	98
5.5 Прямі методи пошуку екстремалей функціоналів в гільбертових просторах .....	102
5.6 Завдання для самоперевірки .....	104

Розділ 6 ОПЕРАТОРИ ТА ЇХ ОСНОВНІ ХАРАКТЕРИСТИКИ.....	106
6.1 Оператор, його лінійність, норма та ряд інших характеристик .....	106
6.2 Обернений оператор та резольвента і спектр оператора.....	109
6.3 Критерії існування однозначного та коректного розв'язків операторних рівнянь .....	111
6.4 Група операторів .....	113
6.5 Метод стиснених відображень.....	114
6.6 Завдання для самоперевірки .....	117
Розділ 7 ПРИКЛАДНІ АСПЕКТИ ТЕОРІЇ ОПЕРАТОРІВ.....	118
7.1 Застосування метода стиснених відображень в теорії алгебраїчних рівнянь .....	118
7.2 Застосування метода стиснених відображень для доведення існування єдиного розв'язку диференціальних та інтегральних рівнянь.....	121
7.3 Метод стиснених відображень в задачі доведення існування розв'язку інтегрального рівняння Фредгольма та побудова його резольвенти .....	123
7.4 Приклади розв'язання операторних рівнянь .....	128
7.5 Завдання для самоперевірки .....	133
Розділ 8 СПЕЦІАЛЬНІ ОПЕРАТОРИ ТА ЇХ ПРИКЛАДНІ ЗАСТОСУВАННЯ.....	134
8.1 Прямий та обернений оператори Лапласа .....	134
8.2 Прямий та обернений оператори Фур'є.....	144
8.3 Різницеві оператори та оператор Деча для відображення дискретних функцій, заданих на нульмірній множині, в простір неперервних функцій .....	150
8.4 Авторегресійні оператори в задачах відображення часових рядів .....	167
8.5 Приклади реалізації процесів застосування спеціальних операторів.....	180
8.6 Завдання для самоперевірки .....	188
Список використаної літератури.....	190

## ВСТУП

Як правило, програма навчальної дисципліни «Вища математика» для нематематичних спеціальностей закладів вищої освіти традиційно охоплює, з різним ступенем, деталізації п'ять математичних дисциплін – лінійну алгебру, математичний аналіз, теорію ймовірностей, функції комплексної змінної та векторну алгебру. Але для ефективного засвоєння деяких спеціальних дисциплін багатьох сучасних спеціальностей інженерії та спеціальностей галузі інформаційних технологій такого математичного підґрунтя вже недосить, а тому в навчальні плани цих спеціальностей входять й інші математичні дисципліни, серед яких для ІТ-спеціальностей важливу роль відіграє математична дисципліна, яка носить назву «Функціональний аналіз» і фактично є, образно кажучи, надбудовою, або «другим поверхом» навчальної дисципліни «Математичний аналіз», що дає право деяким авторам, які написали навчальні посібники з функціонального аналізу, назвати ці посібники, кілька з яких наведені в списку використаної нами літератури, «Додатковими главами математичного аналізу».

З навчальної дисципліни «Математичний аналіз» відомо, що *функція* – це закон, за яким одній числовій множині ставиться у відповідність інша числова множина.

Графічно це виглядає так, як показано на рис. В.1.

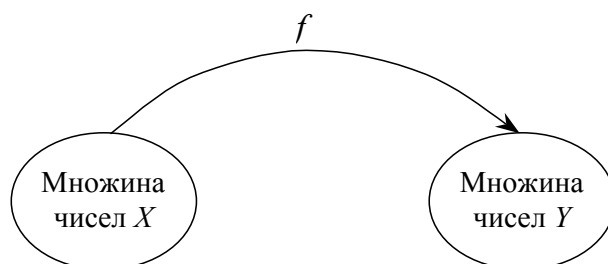


Рисунок В.1 – Графічна інтерпретація поняття функції

При цьому множину  $X$  називають областю задавання функції, а множину  $Y$  – областю значень цієї функції.

Умовно функцію найчастіше записують так

$$y = f(x), \quad x \in X, \quad y \in Y, \quad (1)$$

або так

$$y = y(x), \quad x \in X, \quad y \in Y, \quad (2)$$

де  $\in$  – символ належності елемента до множини.

Якщо кожному числу  $x \in X$  функція  $f$  задає у відповідність лише одне число  $y \in Y$ , то така функція називається *однозначною*, а якщо кожному числу  $x \in X$  функція  $f$  задає у відповідність два або більше чисел  $y \in Y$ , то така функція називається *багатозначною*.

Функцію можна задавати у вигляді таблиці, графіком або однією чи декількома формулами. Приклади наведені на рис. В.2.

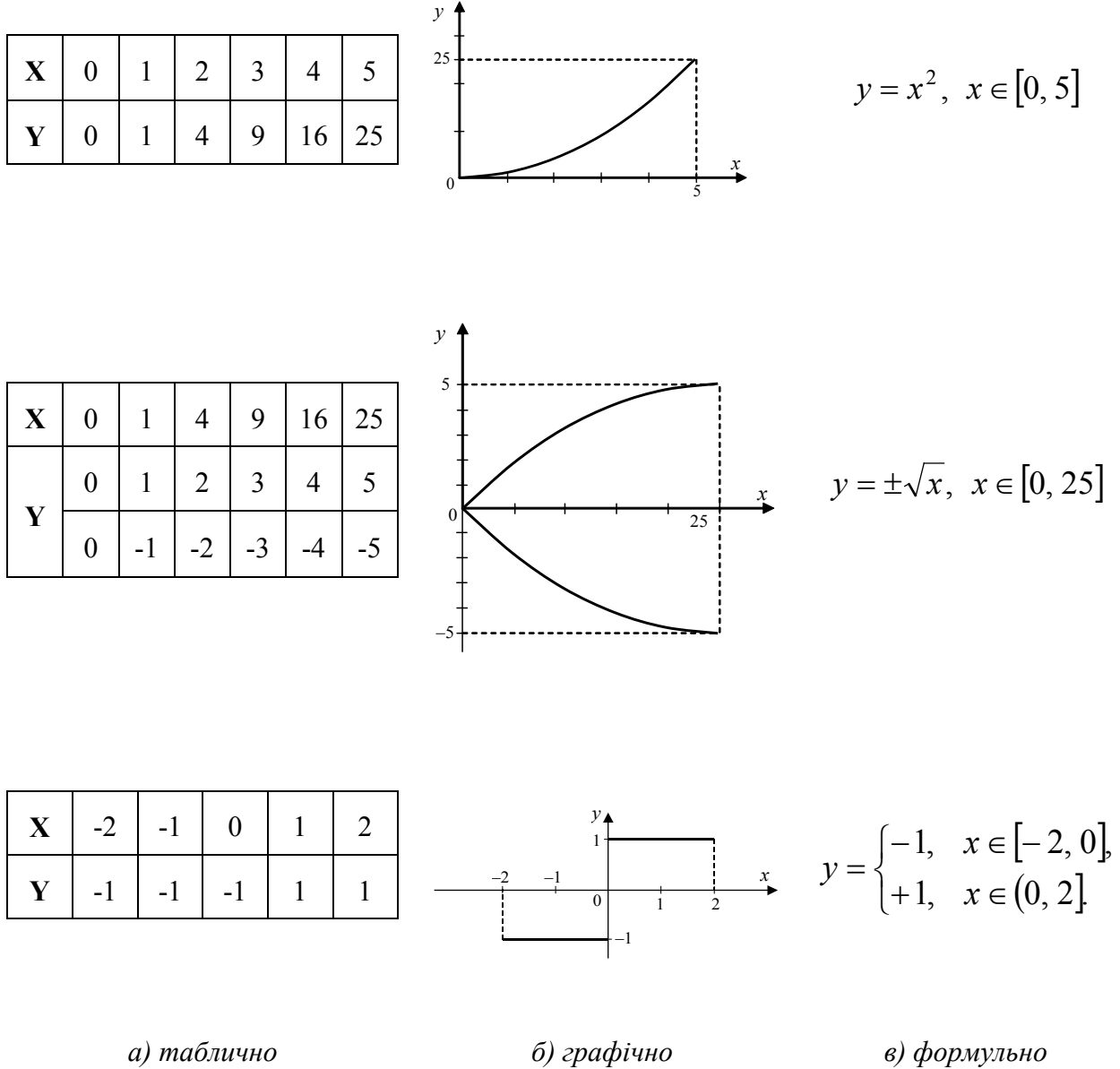


Рисунок В.2 – Приклади задавання функції

Функція, графік якої не має розривів, належить до класу *неперервних*, а неперервна функція, графік якої не містить зломів, а тому має неперервну першу похідну, належить до класу *гладких* (рис. В.3, а).

Неперервна функція, графік якої має злони, а тому її похідна – розриви 1-го роду, належить до класу *кусково-гладких* (рис. В.3, б).

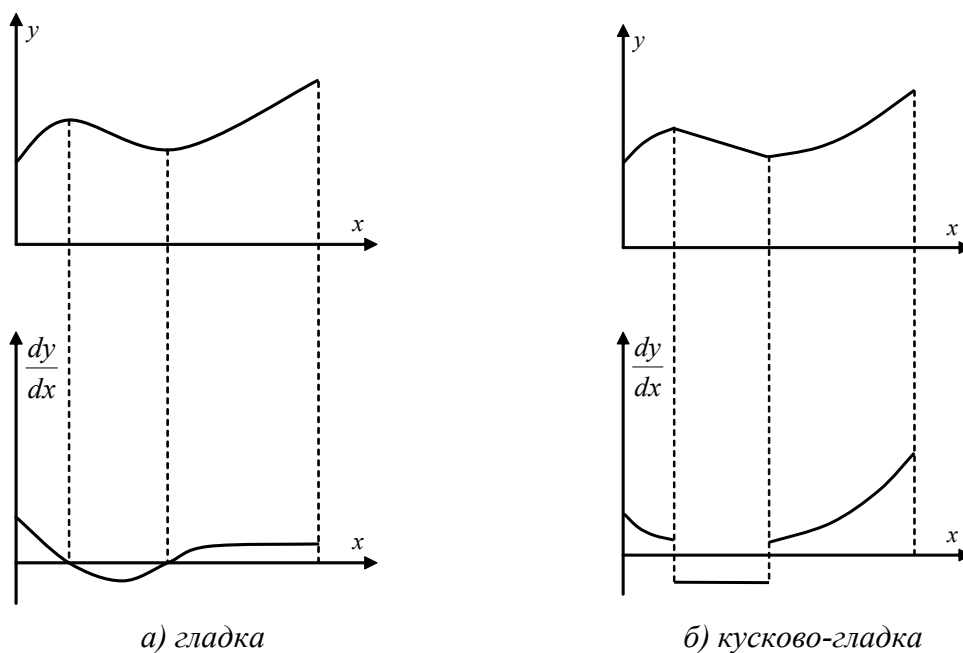


Рисунок В.3 – Приклади неперервних функцій

З тієї ж таки навчальної дисципліни «Математичний аналіз» відомо, що *функціонал* – це закон, за яким множині функцій ставиться у відповідність множина чисел. Графічно це виглядає так, як показано на рис. В.4.

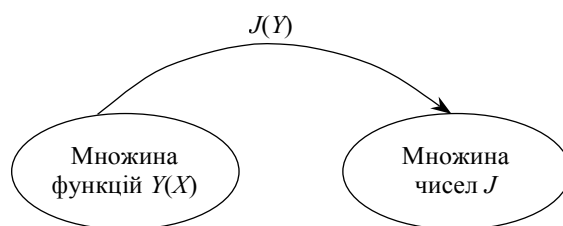


Рисунок В.4 – Графічна інтерпретація поняття функціонала

При цьому множину функцій  $Y(X)$  називають областю задання функціонала, а множину чисел  $J$  – областю його значень.

Умовно функціонал найчастіше записують так

$$J_y = J(y(x), x), \quad x \in X, \quad y(x) \in Y(X), \quad J_y \in J. \quad (3)$$

Прикладами функціоналів можуть бути визначені інтеграли:

$$J_y = \int_a^b y(x) dx \quad (4)$$

або

$$J_y^f = \int_a^b f(x, y) dx, \quad (5)$$

або



$$J_y^F = \int_a^b F(x, y, y') dx, \quad (6)$$

в яких  $f(x, y)$  – математичний вираз, який є конструкцією з незалежної змінної  $x$  та її функції  $y(x)$ , а  $F(x, y, y')$  – математичний вираз, який є конструкцією з незалежної змінної  $x$ , її функції  $y(x)$  та першої похідної  $y'(x)$  від цієї функції; при цьому відрізок  $[a, b]$  є областю задавання функції  $y(x)$ , тобто  $x \in [a, b]$ .

Покажемо, як можна побудувати функціонал.

На рис. В.5 показано дві криві, які з'єднують точки площини  $(0, 0)$  та  $(1, 1)$ .

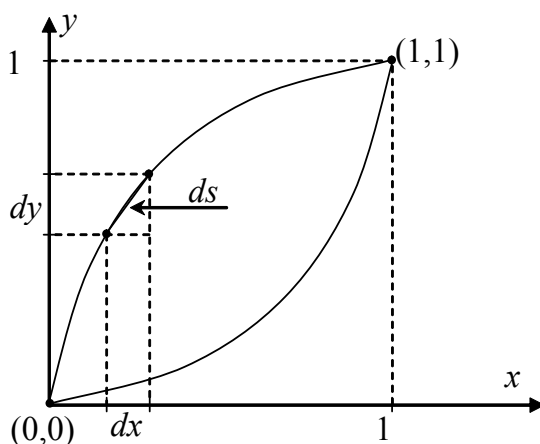


Рисунок В.5 – Графічна інтерпретація побудови функціонала

Побудуємо формулу, за допомогою якої можна знайти довжину такої кривої.

З прямокутного трикутника зі сторонами  $dx$  і  $dy$  (рис. В.5) знайдемо, що нескінченно малий елемент  $ds$  довжини  $L$  кривої, який є гіпотенузою цього трикутника, дорівнює

$$ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \cdot dx, \quad (7)$$

а сама довжина кривої  $L$  визначається інтегралом

$$L = \int_0^1 ds, \quad (8)$$

який після підстановки в нього виразу (7) набуває вигляду

$$L = \int_0^1 \sqrt{1 + (y')^2} dx. \quad (9)$$

Порівнюючи вираз (9) з функціоналом (6), бачимо, що цей вираз теж є функціоналом, для якого

$$\begin{cases} [a, b] = [0; 1]; & x \in [0; 1]; \\ F(x, y, y') = \sqrt{1 + (y')^2}; \\ J_y^F = L. \end{cases} \quad (10)$$

Як другий приклад функціонала можна навести інтеграл

$$S = \int_0^T V(t) dt, \quad (11)$$

в якому  $S$  – це шлях, пройдений автомобілем за час  $T$  зі швидкістю  $V(t)$ .

У цьому випадку

$$\begin{cases} [a, b] = [0; T]; & t \in [0; T]; \\ y(x) = V(t); \\ J_y = S. \end{cases} \quad (12)$$

Як третій приклад функціонала наведемо вираз, за допомогою якого знаходяться втрати активної електричної потужності  $\Delta P$  на активному опорі  $R$ , по якому тече струм  $I(t)$ . Як відомо ще з шкільного курсу фізики, цей вираз має вигляд

$$\Delta P = \frac{1}{T} \int_0^T I^2(t) R dt. \quad (13)$$

У цьому випадку

$$\begin{cases} [a, b] = [0; T]; & t \in [0; T]; \\ f(x, y) = I^2(t) R; \\ J_y^f = \Delta P. \end{cases} \quad (14)$$

Отже, ми вже встановили, що функція задає закон, за яким кожному елементу з однієї числової множини ставиться у відповідність якийсь елемент із іншої або цієї ж самої числової множини, а функціонал задає закон, за яким кожному елементу з множини функцій ставиться у відповідність якийсь елемент із множини чисел.

Тож виникає запитання: «А чи не можна знайти закон, за яким кожному елементу з множини функцій ставиться у відповідність якийсь елемент із іншої або цієї ж самої множини функцій?»

Відповідь на це питання позитивна, а математичне поняття, яке характеризує такий закон, в математиці назвали *оператором*.

Наприклад, між множиною неперервних на  $[a, b]$  функцій  $f(t), t \in [a, b]$  і множиною похідних від цих функцій існує однозначна відповідність, яка задається оператором диференціювання  $D = \frac{d}{dt}$ , наприклад, функції

$$y = t^2 \quad (15)$$

відповідає похідна

$$\frac{dy}{dt} = 2t, \quad (16)$$

яка теж є функцією цієї ж змінної  $t$ , а функції

$$y = e^{-3t^2} \quad (17)$$

відповідає похідна

$$\frac{dy}{dt} = -6t e^{-3t^2}, \quad (18)$$

яка теж є функцією цієї ж змінної  $t$ .

Зіставляючи наведені вище означення та приклади з програмою навчальної дисципліни «Математичний аналіз», ми можемо пересвідчитись у тому, що ця математична дисципліна присвячена, в основному, вивченню властивостей функцій та операцій з ними, основними з яких є диференціювання та інтегрування. І зовсім не приділяється уваги в ній вивченню властивостей функціоналів і операторів як самостійних математичних об'єктів. Почали вивчати ці математичні об'єкти спочатку в «додаткових главах математичного аналізу», які в процесі накопичення нових знань про властивості функціоналів і операторів оформились у вигляді окремої математичної дисципліни, яка отримала назву «Функціональний аналіз» і внесена в перелік обов'язкових математичних дисциплін для студентів ІТ-спеціальностей.

Математична дисципліна «Функціональний аналіз», яка вивчається студентами математичних спеціальностей в університетах, являє собою сукупність понять і теорем, які ці поняття об'єднують в єдину математичну структуру, а тому вона містить в собі на 90 відсотків матеріал, присвячений формулюванням і доведенням цих теорем. У той же час студентам ІТ-спеціальностей, як правило, доводити ці теореми немає сенсу, бо для них більш важливим є вміння використовувати доведений математиками матеріал в практичних застосуваннях. Тож і ми побудували наш навчальний посібник з використанням матеріалу, присвяченого, в основному, викладенню основних понять і результатів доведення теорем функціонального аналізу та їх застосуванню при розв'язанні прикладних задач, що виникають перед ІТ-спеціалістами, не відволікаючи студентів доведенням цих теорем від набуття компетентностей з практичного застосування основ функціонального аналізу у своїй діяльності. Доведення теорем нами наводиться лише у випадках, коли без нього не вдається викласти суть того чи іншого поняття або взаємозв'язку між використовуваними поняттями. Програмний матеріал дисципліни нами викладено у восьми розділах, перший з яких присвячено множинам, метричним просторам

та їх характеристикам; другий – просторам Гільберта та апроксимації функцій в них; третій – теорії міри та інтегралам Рімана, Лебега і Стільтєса; четвертий – функціоналам і методам пошуку їх безумовних екстремумів; п'ятий – методам пошуку умовних екстремумів функціоналів; шостий – теорії операторів; в сьомому розділі викладені прикладні аспекти застосування операторів, а у восьмому розділі дана характеристика і рекомендації із застосування кількох спеціальних операторів, таких як прямий і обернений оператори Лапласа, прямий і обернений оператори Фур'є, оператор Z-перетворення Деча та низка авторегресійних операторів, які широко використовуються в системному аналізі та прикладних інформаційних технологіях.

Завершимо цей короткий вступ до функціонального аналізу зауваженням, що при викладенні матеріалу авторами були використані, тією чи іншою мірою, всі літературні джерела, наведені в бібліографії, але без характерної для монографій і статей конкретизації, що і звідки конкретно взято. А деякий матеріал, взятий з навчальних посібників, написаних самими авторами для забезпечення вивчення деяких конкретних спеціальних дисциплін, в яких використовуються поняття функціонального аналізу і який ми у цьому нашому навчальному посібнику використовуємо для демонстрації розв'язання конкретних прикладних задач, ми наводимо і без купюр, і без взяття їх в цитатні лапки та без посилання на самих себе.

Що ж стосується відмінностей даного нашого навчального посібника від інших навчальних посібників з функціонального аналізу, то вони полягають, у першу чергу, в іншому структуруванні навчального матеріалу та його відборі, обумовленому тим, що наш навчальний посібник орієнтовано на розв'язання саме тих прикладних задач, що постають перед спеціалістами з інформаційних технологій в процесі синтезу ними математичних моделей реальних процесів, а також в процесі розроблення алгоритмів їх оптимізації та синтезу систем оптимального керування цими процесами, які є етапами системного аналізу і служать підґрунтям для створення комп'ютерних програм, що реалізують конкретні інформаційні технології.

# Розділ 1

## МНОЖИНИ І МЕТРИЧНІ ПРОСТОРИ ТА ЇХ ХАРАКТЕРИСТИКИ

### 1.1 Множини, підмножини та їх характеристики

Математики домовились під поняттям **множина** розуміти сукупність об'єктів якоїсь природи, які прийнято називати елементами. **Множина** вважається заданою, якщо є відомими всі її елементи та правило, згідно з яким ці елементи відносять до цієї **множини**.

Елементами **множини** можуть бути, наприклад, усі автомобілі, що зареєстровані у Україні, або усі річки, що течуть Україною, або усі числа натурального ряду на числовій осі, або усі дійсні числа, що розміщені на відрізку  $[0,1]$  числової осі, або усі неперервні функції, аргументи яких задані на цьому відрізку числової осі, – і кожен, хто читає ці рядки, може продовжувати наводити свої приклади **множин**.

Варто відзначити, що поняття **множина** належить до числа первинних понять математичної науки, які не можуть бути визначеними через якісь інші, більш прості поняття.

Математики домовились **множини** позначати великими літерами латинського або грецького алфавітів, а їх елементи позначати малими літерами з цих же алфавітів, наприклад,  $A, B, X, Z, E, \Phi, \Omega, \Psi$  – це **множини**, а  $a, b, x, z, \varepsilon, \phi, \omega, \psi$  – це елементи. На факт належності елемента  $x$  **множині**  $X$  вказує символічний запис  $x \in X$ , а неналежності – символічний запис  $x \notin X$ . **Множина** зі скінченною кількістю елементів називається **скінченною множиною**, а **множина** з нескінченною кількістю елементів називається **нескінченною множиною**. Прикладом **скінченної множини** є **множина** автомобілів, зареєстрованих в Україні, а прикладом **нескінченної множини** є **множина** дійсних чисел на відрізку  $[0,1]$  числової осі. Якщо елементами **множини**  $A$  є скінченна числова послідовність із  $n$  членів, то символічно це може бути записано як  $A = \{a_i\}, i = 1, 2, \dots, n$ . Якщо ж елементами цієї **множини**  $A$  є нескінченна числова послідовність членів, то символічно це може бути записано як  $A = \{a_i\}, i = 1, 2, 3, \dots$ . **Множина**, яка не містить жодного елемента, називається **пустою множиною** і позначається символом  $\emptyset$  чи  $O$ , який не потрібно ототожнювати з числом «нуль».

Якщо **множини**  $A$  і  $B$  складаються з однакових елементів, то вони вважаються рівними між собою, про що свідчить запис  $A = B$ . Якщо ж у **множині**  $A$  входять не всі елементи **множини**  $B$ , то **множину**  $A$  називають **підмножиною множини**  $B$ , про що свідчить запис  $A \subset B$ . Наприклад, на числовій осі **множина** раціональних чисел  $R$ , кожне з яких, як відомо, утворюється відношенням двох цілих чисел, є **підмножиною множини** дійсних чисел  $Z$ . Якщо ми не впевнені в тому, що **підмножина**  $A$  **множини**  $B$  містить в собі елементів менше, ніж **множина**  $B$ , то записуємо це так:  $A \subseteq B$ .

При об'єднанні двох **множин**  $A$  і  $B$  утворюється нова **множина**  $M$ , яку називають її **сумою** і яка містить у собі всі елементи обох цих **множин**, причому кожен однаковий елемент обох **множин** в їх **суму**  $M$  входить як один елемент – символічно записується це так:

$$M = A \cup B \quad (1.1)$$

Наприклад, якщо  $A$  і  $B$  – це числові **множини**, де

$$A = \{1,2,3,4,5\}, \quad B = \{4,5,6,7,8\}, \quad (1.2)$$

то, згідно з (1.1), матимемо

$$M = A \cup B = \{1,2,3,4,5\} \cup \{4,5,6,7,8\} = \{1,2,3,4,5,6,7\} \quad (1.3)$$

При перерізі двох **множин**  $A$  і  $B$  утворюється нова **множина**  $P$ , яку називають їх **перерізом** і яка містить у собі лише ті елементи обох цих **множин**, які є однаковими, причому кожен із цих однакових елементів обох **множин** в їх **переріз**  $P$  входить як один елемент – символічно записується це так:

$$P = A \cap B \quad (1.4)$$

Для наведених в умовах попереднього прикладу числових **множин** (1.2), згідно з (1.4), матимемо

$$P = A \cap B = \{1,2,3,4,5\} \cap \{4,5,6,7,8\} = \{4,5\} \quad (1.5)$$

Для суми і перерізу **множин**  $A$ ,  $B$ ,  $C$  справедливими є властивості: асоціативності

$$(A \cup B) \cup C = A \cup (B \cup C), \quad (1.6)$$

$$(A \cap B) \cap C = A \cap (B \cap C), \quad (1.7)$$

комутативності

$$A \cup B = B \cup A, \quad (1.8)$$

$$A \cap B = B \cap A, \quad (1.9)$$

дистрибутивності

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \quad (1.10)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C) \quad (1.11)$$

В справедливості співвідношень (1.6)–(1.11) легко переконатись, якщо на листі паперу нанести ці **множини** у вигляді кругів, що перетинаються, і розглянути операції суми і перерізу спочатку у послідовності, записаній у лівій частині цих рівнянь, а потім у послідовності, записаній у правій частині цих рівнянь, пам'ятаючи про те, що спочатку виконуються операції над **множинами**, поміщеними в дужки.

Звертаємо увагу на те, що в разі, якщо замість **множин** у співвідношення (1.6)–(1.10) поставимо числа, а операції суми і перерізу множин замінимо на алгебраїчні операції підсумовування та множення чисел, то вказані рівності підтвердяться, а якщо те ж саме проробити зі співвідношенням (1.11), то для чисел воно не виконуватиметься, що є наочним свідченням того, що операції з **множинами** не тотожні алгебраїчним операціям з числами.

Особливо наочно ця відмінність проявляється в операціях суми та перерізу **множини**  $A$  з самою собою та зі своєю **підмножиною**  $B$ , для яких справедливими є співвідношення

$$A \cup A = A, \quad (1.12)$$

$$A \cap A = A, \quad (1.13)$$

$$A \cup B = A, \quad (1.14)$$

$$A \cap B = B \quad (1.15)$$

Співвідношення (1.12), (1.13) є справедливими за визначенням операцій суми та перерізу *множин*, а у справедливості співвідношень (1.14), (1.15) легко пересвідчитись графічно, накладаючи круг, що символізує *підмножину*  $B$ , на круг, що символізує *множину*  $A$ .

*Множина*  $Q$ , яка складається з елементів *множини*  $A$ , що не входять у *множину*  $B$ , називається *різницею* цих *множин* і позначається як  $A - B$  або  $A \setminus B$ , тобто

$$Q = A - B = A \setminus B \quad (1.16)$$

Цілком очевидно, що в загальному випадку

$$A - B \neq B - A \quad (1.17)$$

Наприклад, для числових *множин* (1.2)

$$A - B = \{1, 2, 3\}, \quad (1.18)$$

$$B - A = \{6, 7\} \quad (1.19)$$

Якщо *множина*  $A$  є *підмножиною* *множини*  $B$ , то *різницю*  $B - A$  називають *доповненням* *множини*  $A$  до *множини*  $B$  і символічно позначають як  $C_B A$ , тобто

$$C_B A = B - A \quad (1.20)$$

Наприклад, *множина*  $\bar{R}$  ірраціональних чисел на числовій осі є *доповненням* *множини*  $R$  раціональних чисел до *множини* дійсних чисел  $Z$ , тобто

$$\bar{R} = C_Z R = Z - R \quad (1.21)$$

Якщо *множини*  $A_i, i = 1, 2, \dots, n$  є *підмножинами* *множини*  $A$ , то справедливими є співвідношення

$$C_A A_1 \cup C_A A_2 \cup \dots \cup C_A A_n = C_A (A_1 \cap A_2 \cap \dots \cap A_n), \quad (1.22)$$

$$C_A A_1 \cap C_A A_2 \cap \dots \cap C_A A_n = C_A (A_1 \cup A_2 \cup \dots \cup A_n), \quad (1.23)$$

в справедливості яких легко пересвідчитись графічно, наприклад, для  $n = 3$ , якщо *множину*  $A$  зобразити на рисунку у вигляді квадрата, в який вписано три круги, що зображають *підмножини*  $A_1, A_2, A_3$ .

Важливою характеристикою *множин* є їх *еквівалентність*, згідно з якою *множини*  $A, B$  вважаються *еквівалентними*, якщо кожному елементу  $a \in A$  за якимось, заздалегідь обумовленим правилом, ставиться у відповідність один єдиний елемент  $b \in B$ , а кожному елементу  $b \in B$  ставиться у відповідність один єдиний елемент  $a \in A$ . Наприклад, еквівалентними є *множина*  $A$  легкових автомобілів приватної власності, кожний із яких зареєстровано лише на одного власника у певному населеному пункті, і *множина*  $B$  людей, які є власниками цих автомобілів. Правилком, за яким встановлюється *еквівалентність* цих *множин*, є запис органом реєстрації автомобілів прізвища власника в паспорті автомобіля.

А для того, щоб порівнювати між собою *нееквівалентні* *множини*, вводиться поняття їх *потужності*, яка для *множини*  $A$  символічно записується як  $\bar{A}$  та іноді

називається також його **кардинальним числом**, але ми будемо використовувати цю характеристику у нашому навчальному посібнику лише у варіанті **потужності**, яку будемо встановлювати за чимось спільним, що має місце в усіх **множинах, еквівалентних** тому, яке ми розглядаємо. Очевидно, що спільним у **скінченних множинах** різної природи є лише кількість їх елементів, а тому, якщо **множина A** має  $n$  елементів, а **множина B** має  $t$  елементів і при цьому  $n > t$ , то ми констатуємо, що **множина A** має **потужність** більшу, ніж **множина B**.

Але виникає запитання: «А як порівнювати між собою за **потужністю нескінченні множини**, кожна з яких має нескінченну кількість елементів?».

Досліджуючи на числовій осі різні нескінченні числові послідовності, математики встановили, що з усіх нескінченних послідовностей найшвидше наближається до нескінченності **натуральний ряд N**, оскільки кожне його наступне число дорівнює попередньому, збільшеному на одиницю, і при цьому при формуванні цього ряду пропускаються всі дійсні числа, які містяться на числовій осі у кожній такій одиниці. А тому домовились вважати **натуральний ряд**, який являє собою нескінченний ряд чисел, **нескінченною множиною найменшої потужності, яку домовились позначати символічно малою латинською літерою a**, тобто,

$$\overline{N} = a, \quad (1.24)$$

і всі інші **нескінченні множини** порівнювати між собою, виходячи з того, як вони співвідносяться за **потужністю** з **потужністю натурального ряду**, визначеною співвідношенням (1.24), а всі **нескінченні множини з потужністю натурального ряду** називати **зліченими множинами**, оскільки кожному їх елементу можна приписати індекс, який дорівнює відповідному числу натурального ряду, за рахунок чого кожен їх елемент можна **полічити**.

І першим фактом, який встановили математики після цієї домовленості стосовно **потужності натурального ряду**, це те, що **потужність множини Z дійсних чисел**, заданих на відрізку  $[0,1]$ , **є більшою за a**.

Доведення цього факту просте – якщо скласти послідовність  $x_1, x_2, x_3, \dots$  таких дійсних чисел на відрізку  $[0,1]$  числової осі, щоб кожне наступне число було втричі меншим попереднього, і розділити цей відрізок  $[0,1]$  на три рівні сегменти шириною  $\Delta_1$  кожен, то принаймні в один із цих сегментів точка  $x_1$  не потрапить. Розділимо той сегмент, в який точка  $x_1$  не потрапила, теж на три рівні сегменти шириною  $\Delta_2$  кожен, і виберемо той з них, в який не потрапила точка  $x_2$ . За цим алгоритмом будемо продовжувати цей процес до нескінченності. В результаті ми на відрізку  $[0,1]$  числової осі отримаємо **зліченну множину**  $\Delta_1, \Delta_2, \Delta_3, \dots$  елементами якої є все менші і менші сегменти відрізка та поряд з якою на цьому ж відрізку існує задана нами попередньо **зліченна множина чисел**  $x_1, x_2, x_3, \dots$ , жодне із яких не попадає в жоден із цих сегментів. А це означає, що дійсних чисел на відрізку  $[0,1]$  числової осі є більше ніж чисел натурального ряду на усій числовій осі, що дозволяє зробити висновок, що **потужність нескінченної множини дійсних чисел на відрізку  $[0,1]$  є більшою потужності натурального ряду, який є зліченною множиною**.

**Потужність нескінченної множини дійсних чисел на відрізку  $[0,1]$**  математики **вирішили назвати потужністю континууму**, а символічно **позначати малою латинською літерою c**, отже, справедливою є нерівність

$$c > a. \quad (1.25)$$



Більше того, математики встановили, що для потужностей  $c$ , а справедливою є не лише нерівність (1.25), а й **рівність**

$$c = 2^a. \quad (1.26)$$

Для доведення рівності (1.26) використаємо метод математичної індукції за таким алгоритмом: розглянемо послідовно, які потужності  $n_0, n_1, n_2, n_3$  матимуть множини, породжувані скінченними множинами  $A_0 = \{O\}, A_1 = \{a_1\}, A_2 = \{a_1, a_2\}, A_3 = \{a_1, a_2, a_3\}$ , якщо у кожен з породжених множин як елементи ввести ще й усі можливі підмножини, породжені елементами цих скінченних множин. Пам'ятаючи про те, що число комбінацій  $C_n^m$  із  $n$  елементів по  $m$ , як це відомо ще з курсу шкільної математики, визначається за співвідношенням

$$C_n^m = \frac{n!}{m!(n-m)!}, \quad (1.27)$$

знайдемо, що

$$\begin{cases} n_0 = C_0^0 = \frac{0!}{0!0!} = 1 = 2^0, \\ n_1 = C_1^0 + C_1^1 = \frac{1!}{0!1!} + \frac{1!}{1!0!} = 1 + 1 = 2 = 2^1, \\ n_2 = C_2^0 + C_2^1 + C_2^2 = \frac{2!}{0!2!} + \frac{2!}{1!1!} + \frac{2!}{2!0!} = 1 + 2 + 1 = 4 = 2^2, \\ n_3 = C_3^0 + C_3^1 + C_3^2 + C_3^3 = \frac{3!}{0!3!} + \frac{3!}{1!2!} + \frac{3!}{2!1!} + \frac{3!}{3!0!} = 1 + 3 + 3 + 1 = 8 = 2^3. \end{cases} \quad (1.28)$$

За ідеологією методу математичної індукції зі співвідношень (1.28) витікає, що в разі, якщо скінченна множина  $A_n = \{a_1, a_2, a_3, \dots, a_n\}$  має  $n$  елементів, то потужність  $n_n$  породженої нею множини, яка містить в собі як елементи й усі можливі підмножини цієї породної множини, дорівнюватиме

$$n_n = 2^n. \quad (1.29)$$

Із співвідношення (1.29) витікає, що потужність  $n_n$  множини, породженої множиною потужності  $n$ , за рахунок внесення в породжену множину всіх можливих підмножин, складених із елементів породної множини, дорівнює двійці в степені, яким є потужність породної множини. А звідси висновок, що, якщо потужність породної зліченної множини дорівнює  $a$ , то потужність породженої нею нескінченної множини дійсних чисел на відріжку  $[0, 1]$ , елементами якої є усі можливі підмножини з елементів цієї породної множини, дорівнюватиме двійці в степені  $a$ , що і доводить справедливість рівності (1.26).

А тепер давайте повернемося до виразу (1.21), згідно з яким множина дійсних чисел  $Z$  на відріжку  $[0, 1]$  числової осі, є сумою підмножини  $R$  раціональних чисел та підмножини  $\bar{R}$  ірраціональних чисел, заданих на цьому ж відріжку.

Як відомо, кожне раціональне число можна подати у вигляді відношення двох цілих чисел, причому, якщо це раціональне число менше одиниці, то його чисельником завжди є ціле число, яке є меншим за число, що стоїть у знаменнику. Оскільки цілі числа є елементами натурального ряду, який є зліченною множиною потужності  $a$ , то ці числа і в чисельнику, і в знаменнику можна полічити, а тому підмножина раціональних чисел на

відрізку  $[0, 1]$  числової осі теж є зліченною множиною потужності  $a$ . А з цього факту витікає два наслідки, перший з яких засвідчує, що підмножина ірраціональних чисел на вказаному відрізку є нескінченною множиною потужності континууму  $c$ , бо лише за рахунок цієї підмножини множина дійсних чисел на вказаному відрізку матиме потужність  $c$ , що нами уже показано вище відносно множини дійсних чисел, виходячи з іншого посилання. А другим наслідком є твердження, що в разі додавання до множини потужності континууму будь-якої зліченної підмножини потужність їх суми залишається рівною  $c$ .

А далі звертаємо увагу на те, що всі одиничні відрізки на числовій осі, що знаходяться між сусідніми натуральними числами, можна полічити, приписавши кожному з них індекс, що дорівнює натуральному числу, розміщеному на правій границі кожного такого одиничного відрізка, тож підмножина одиничних відрізків, розміщених між натуральними числами на числовій осі, є зліченною множиною потужності  $a$ , яка є меншою потужності континууму  $c$  одиничного відрізка  $[0, 1]$  числової осі. Тож на основі цього твердження можна зробити важливий висновок, що і вся вісь дійсних чисел є множиною потужності континууму  $c$ .

Але, як ми уже показали вище, множина, яка породжується об'єднанням усіх можливих підмножин породної множини певної потужності, має потужність, що дорівнює двійці в степені, що дорівнює потужності породної множини. А з цього факту висновок: потужність  $\bar{f}$  множини усіх функцій  $f(x)$ , аргумент  $x$  яких задано на відрізку  $[0, 1]$  (або і на усій числовій осі) потужності континууму  $c$ , дорівнює двійці в степені  $c$ , тобто,

$$\bar{f} = 2^c \quad (1.30)$$

Ось на цьому ми і завершимо розгляд матеріалу підрозділу, присвяченого множинам, підмножинам та їх характеристикам, який нам потрібен буде при викладенні основ функціонального аналізу. А бажаючих дізнатись більше у цьому напрямку математики відсилаємо до навчальних посібників з теорії множин чи з функціонального аналізу, за якими цей матеріал засвоюють студенти математичних спеціальностей університетів.

## 1.2 Метричні простори та їх характеристики

Множина

$$\Omega = \{x, y, z, \dots, u, v, \dots\} \quad (1.31)$$

елементів якоїсь природи називається **метричним простором**, якщо кожній упорядкованій парі елементів  $x, y \in \Omega$  поставлено у відповідність невід'ємне число  $\rho(x, y)$ , яке називають **метрикою простору**  $\Omega$ , якщо це число задовольняє три аксіоми метрики:

1) аксіому тотожності

$$\rho(x, y) = 0 \quad (1.32)$$

тоді і лише тоді, коли

$$x = y; \quad (1.33)$$

2) аксіому симетрії

$$\rho(x, y) = \rho(y, x); \quad (1.34)$$

3) аксіому трикутника

$$\rho(x, y) + \rho(y, z) \geq \rho(x, z). \quad (1.35)$$

Розглядаючи ці аксіоми бачимо, що *метрика*  $\rho(x, y)$  *простору*  $\Omega$  *задає відстань між елементами*  $x, y$  *цього простору.*

*Елементи метричного простору називають точками.*

### Приклади

1. Для тривимірного, звичного для нас, евклідового простору  $E_3$  відстань між точками  $x = \{x_1, x_2, x_3\}$  і  $y = \{y_1, y_2, y_3\}$  ( $x, y \in E_3$ ) визначається формулою

$$\rho(x, y) = \sqrt{\sum_{i=1}^3 (x_i - y_i)^2}. \quad (1.36)$$

2. Для множини  $C[0, 1]$  неперервних функцій  $x(t), y(t), \dots$ , заданих на відрізку  $t \in [0, 1]$ , відстань між елементами  $x(t)$  і  $y(t)$  задається формулою

$$\rho(x, y) = \max_t |x(t) - y(t)|. \quad (1.37)$$

Якщо  $X$  – довільний метричний простір, то кажуть, що послідовність

$$\{x_n\} \subset X \quad (1.38)$$

збігається до точки  $x_0 \in X$ , якщо при  $n \rightarrow \infty$

$$\rho(x_n, x_0) \rightarrow 0, \quad (1.39)$$

або в іншому записі

$$\lim_{n \rightarrow \infty} x_n = x_0. \quad (1.40)$$

Про послідовність  $\{x_n\}$ , яка збігається до деякої точки  $x_0$ , кажуть, що вона є обмеженою.

**Сукупність точок**  $x$  метричного простору  $X$  таких, що

$$\rho(x, a) < r, \quad (1.41)$$

де  $a$  – фіксована точка в  $X$ ,  $r$  – фіксоване дійсне число, *називають відкритою кулею* радіуса  $r$  з центром в точці  $a$  і позначають  $S(a, r)$ , а таких, що

$$\rho(x, a) \leq r - \quad (1.42)$$

замкнутою кулею і позначають  $\overset{\circ}{S}(a, r)$ .

### Приклади

1. В просторі  $E_3$  відкрита куля  $S(a, r)$  – це звичайна куля радіуса  $r$  з центром в точці  $a$  з декартовими координатами  $\{a_1, a_2, a_3\}$ .

2. В просторі  $C[0, 1]$  замкнута куля  $\overset{\circ}{S}(x_0, r)$  – це сукупність усіх функцій  $x(t)$ , графіки яких (рис. 1.1) не виходять зі смуги  $2r$ , обмеженої кривими  $x_0(t) - r$ ,  $x_0(t) + r$ .

*Точку  $b \in X$  називають точкою множини  $M \subset X$ , якщо будь-який окіл точки  $b$  містить хоча б одну точку множини  $M$ , що відрізняється від  $b$ .* Нагадаємо, що запис  $b \in X$  означає, що  $b$  є точкою простору  $X$ , а запис  $M \subset X$  означає, що множина  $M$  є підмножиною простору  $X$ , тобто, що  $X$  містить в собі і множину  $M$ .

*Під околom якоїсь точки простору розуміють відкриту кулю з центром у цій точці.*

*Сукупність усіх граничних точок множини  $M$  називають похідною множиною множини  $M$  і позначають  $M'$ .*

*Якщо множина  $M$  містить у собі і всі свої граничні точки, то вона є замкнутою.*

*Операцію об'єднання множини  $M$  із похідною множиною  $M'$*

$$\overset{\circ}{M} = M \cup M' \quad (1.43)$$

*називають замиканням множини  $M$ .*

Зауважимо, що множина натуральних чисел  $N$  хоч і нескінченна, але є замкнутою, оскільки  $N'$  є пустою множиною  $\emptyset$ , а тому

$$\overset{\circ}{N} = N \cup N' = N \cup \emptyset = N. \quad (1.44)$$

*Кожну точку множини  $M$  метричного простору  $X$ , яка не є граничною, називають ізольованою.*

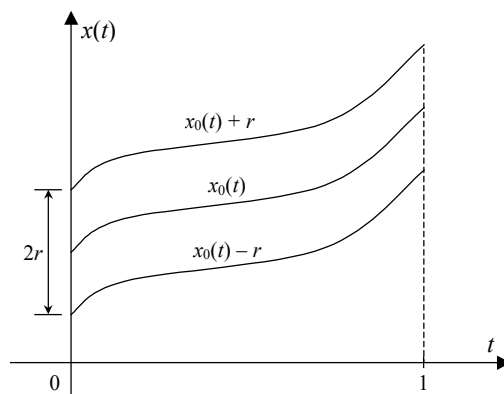


Рисунок 1.1 – Приклад замкнутої кулі  $\overset{\circ}{S}(x_0, r)$  в метричному просторі  $C[0, 1]$   
**Якщо множина  $M$  замкнута і не містить ізольованих точок, тобто,**

$$M = M', \quad (1.45)$$

то вона називається *досконалою*.

**Приклад:** сфера – це замкнута множина точок, рівновіддалених від центра.

**Множину**  $P$  метричного простору  $X$  *називають щільною* в множині  $M \subset X$ , якщо

$$\overset{\circ}{P} \supset M. \quad (1.46)$$

**Множину**  $P$ , яка є щільною в самому просторі  $X$ , *називають всюди щільною*.

**Приклад:** множина раціональних чисел на числовій осі.

**Метричний простір**  $X$  *називають сепарабельним*, якщо в ньому існує зліченна, всюди щільна множина.

**Приклад:** числова вісь, простір  $C[0, 1]$ .

**Множину**  $N \subset X$  *називають ніде не щільною*, якщо в будь-якій кулі  $S(a, r)$  простору  $X$  знайдеться інша куля

$$S(a^*, r^*) \subset S(a, r), \quad (1.47)$$

яка не містить жодної точки множини  $N$ .

**Приклад:** множина  $N$  натуральних чисел.

Нехай дано метричний простір  $X$ , і нехай має місце послідовність точок  $\{x_n\}$  в цьому просторі, яка збігається до точки  $x_0 \in X$ . Тоді при  $n \rightarrow \infty$  справедливим буде і вираз (1.39), і вираз

$$\rho(x_{n+p}, x_0) \rightarrow 0, \quad (1.48)$$

для будь-якого  $p > 0$ . А нерівність трикутника (1.35) з використанням виразів (1.39) та (1.48) набуває вигляду

$$\rho(x_{n+p}, x_0) + \rho(x_n, x_0) \geq \rho(x_{n+p}, x_n). \quad (1.49)$$

А із виразів (1.39), (1.48) і (1.49) в силу нерівності трикутника для метрики наслідком матимемо вираз

$$\rho(x_{n+p}, x_n) \rightarrow 0. \quad (1.50)$$

Якщо ж для якоїсь послідовності  $\{x_n\} \subset X$  виконується умова (1.50), то її *називають фундаментальною послідовністю* або послідовністю, що збігається в собі, або *послідовністю Коші*.

Якщо в метричному просторі  $X$  будь-яка послідовність  $\{x_n\} \subset X$ , що збігається в собі, збігається і до деякої граничної точки  $x_0$ , яка є елементом цього ж простору, тобто  $x_0 \in X$ , то *цей простір*  $X$  *називають повним*.

**Приклад повного простору:** простір  $C[0, 1]$ .

**Приклад неповного простору:** простір дійсних багаточленів

$$p_n(t) = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \dots + a_1 \cdot t + a_0, \quad (1.51)$$

визначених на відрізку  $t \in [0, 1]$  з метрикою

$$\rho(p_n, q_m) = \max_t |p_n(t) - q_m(t)|. \quad (1.52)$$

Для підтвердження нагадаємо, що послідовність  $\{p_n(t)\}$ , де

$$P_n(t) = \sum_{k=0}^n \frac{1}{k!} \cdot t^k, \quad (1.53)$$

що належить до класу (1.52), рівномірно збігається в собі, але не має границі у просторі багаточленів, оскільки її границею

$$\lim_{n \rightarrow \infty} P_n(t) = e^t \quad (1.54)$$

є трансцендентна функція  $e^t$ , яка не є точкою в просторі дійсних багаточленів.

**Множину  $M \subset X$  називають компактною**, якщо із будь-якої послідовності  $\{x_n\} \subset M$  елементів цієї множини можна виділити збіжну підпослідовність

$$\{x_{n_i}\} \subset M. \quad (1.55)$$

Якщо граничні точки  $x_{0_i}$  для всіх  $\{x_{n_i}\}$  належать множині  $M$ , тобто,  $x_{0_i} \in M$ , то **множину  $M$  називають компактною в собі**.

Зрозуміло, що множина компактна в собі, якщо вона просто компактна і замкнута.

Якщо весь **метричний простір  $X$**  є множиною компактною в собі, то його **називають компактном**.

**Приклад:** відрізок  $[0, 1]$  є компактна в собі множина на числовій осі.

Встановлення властивості компактності множини є задачею непростою, а тому для різних просторів розроблені **ознаки компактності**.

Наприклад, для метричного простору  $C[0, 1]$  має місце **критерій компактності Арчела**, який формулюється так: для того, щоб множина функцій

$$Q = \{x(t)\} \subset C[0, 1] \quad (1.56)$$

була компактною, необхідно і достатньо, щоб функції  $x(t)$  цієї множини  $Q$  були рівномірно обмежені та рівностепеневно неперервні.

При цьому функції  $x(t)$  є рівномірно обмеженими, якщо виконується умова

$$\begin{cases} |x(t)| \leq K, \\ K > 0, \\ \forall x(t) \subset Q, \end{cases} \quad (1.57)$$

і є рівностепеневно неперервними, якщо для будь-якого  $\varepsilon > 0$  знайдеться таке  $\delta > 0$ , що

$$|x(t^*) - x(t^{**})| < \varepsilon \text{ при } |t^* - t^{**}| < \delta \quad (1.58)$$

одразу для всіх  $x(t) \in Q$ , незалежно від знаходження точок  $t^*$ ,  $t^{**}$  на  $[0, 1]$ . Нагадаємо, що символ  $\forall$  означає «для будь-якого».

**Завершимо** цей підрозділ ми **теоремою Вейєрштрасса**, яку наведемо без доведення: функція  $f(x)$ , яка є неперервною на компактній в собі множині  $Q$  метричного простору  $X$ , є обмеженою на множині  $Q$  і досягає на цій множині своєї точної верхньої та точної нижньої границь.

### 1.3 Лінійний і банахів метричні простори та їх характеристики

В попередньому підрозділі ми вже встановили, що метричний простір  $X$  є повним, якщо будь-яка його фундаментальна послідовність збігається до елемента цього ж простору.

**Метричний простір  $X$  називається лінійним, якщо на ньому визначені операції додавання і множення на скаляр, котрі задовольняють такі умови:**

$$1) \quad x^* + x^{**} = x^{**} + x^*, \quad \forall x^*, x^{**} \in X; \quad (1.59)$$

$$2) \quad (x^* + x^{**}) + x^{***} = x^* + (x^{**} + x^{***}), \quad \forall x^*, x^{**}, x^{***} \in X; \quad (1.60)$$

$$3) \quad x + 0 = x, \quad \forall x \in X, 0 \in X, \quad (1.61)$$

де елемент  $0$  є нулем множини  $X$ ;

$$4) \quad \text{для } \forall x^* \in X \text{ знайдеться } x^{**} \in X \text{ такий, що}$$

$$x^* + x^{**} = 0, \quad (1.62)$$

де елемент  $x^{**}$  є елементом, протилежним елементу  $x^*$ ;

$$5) \quad 1 \cdot x = x, \quad \forall x \in X; \quad (1.63)$$

$$6) \quad \alpha \cdot (\beta \cdot x) = (\alpha \cdot \beta) \cdot x, \quad \forall x \in X \text{ і } \forall \alpha, \beta; \quad (1.64)$$

$$7) \quad (\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x, \quad \forall x \in X \text{ і } \forall \alpha, \beta; \quad (1.65)$$

$$8) \quad \alpha \cdot (x^* + x^{**}) = \alpha \cdot x^* + \alpha \cdot x^{**}, \quad \forall x^*, x^{**} \in X \text{ і } \forall \alpha. \quad (1.66)$$

**Лінійний метричний простір  $X$  називається нормованим, якщо  $\forall x \in X$  може бути поставлене у відповідність деяке невід'ємне число  $\|x\|$ , яке називається нормою і яке задовольняє такі умови:**

$$1) \quad \|x\| = 0 \text{ тоді і тільки тоді, коли}$$

$$x = 0; \quad (1.67)$$

$$2) \quad \|\alpha \cdot x\| = |\alpha| \cdot \|x\|, \quad \alpha - \text{скаляр}; \quad (1.68)$$

$$3) \quad \|x^* + x^{**}\| \leq \|x^*\| + \|x^{**}\|, \quad \forall x^*, x^{**} \in X. \quad (1.69)$$

Нескладно переконатись в тому, що норма  $\|x\|$  – це відстань від елемента  $x$  до нульового елемента множини  $X$ .

**Приклади норм:**

1) для простору  $C[0, 1]$

$$\|x(t)\| = \max_{t \in [0, 1]} |x(t)| \quad (1.70)$$

або

$$\|x(t)\| = \sup_{t \in [0, 1]} |x(t)|; \quad (1.71)$$

2) для евклідового простору  $E_n$  розмірності  $n$

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2}, \quad (1.72)$$

де  $x = \{x_1, x_2, \dots, x_n\}$ ,  $x \in E_n$ .

Зрозуміло, що для будь-якого лінійного нормованого простору  $X$  є справедливим співвідношення

$$\|x^* - x^{**}\| = \rho(x^*, x^{**}), \quad (1.73)$$

де  $x^*, x^{**} \in X$ .

**Повний лінійний нормований простір називається банаховим (за прізвищем математика, який досліджував цей простір) і позначається як  $B$ -простір.**

Зрозуміло, що простори  $C[0, 1]$  і  $E_n$  є банаховими.

Відзначимо, що норма в  $B$ -просторі може бути введена по-різному, аби лише відповідала умовам (1.67), (1.68), (1.69).

Наприклад, в просторі неперервних на відрізку  $t \in [0, 1]$  функцій  $x(t)$  норму можна ввести не лише у вигляді (1.70), але й у вигляді



$$\|x\| = \int_0^1 |x(t)| dt. \quad (1.74)$$

Такий  $B$ -простір називають простором Лебега і позначають  $L[0, 1]$ , щоб відрізнити від простору  $C[0, 1]$  тих же функцій, але з нормою (1.70).

Для простору  $E_n$  як норму можна використовувати не лише співвідношення (1.72), але і більш загальне

$$\|x\| = \left( \sum_{i=1}^n x_i^p \right)^{\frac{1}{p}}, \quad p > 0. \quad (1.75)$$

Зрозуміло, що (1.72) збігається з (1.75) лише для  $p = 2$ .

#### 1.4 Завдання для самоперевірки

1. Визначте поняття «множина» і наведіть приклади множин.
2. Що собою являють сума, переріз і різниця множин?
3. Які множини є еквівалентними?
4. Визначте поняття «потужність множини»
5. Чому дорівнює потужність натурального ряду чисел?
6. Чому дорівнює потужність множини раціональних чисел?
7. Чому дорівнює потужність множини дійсних чисел?
8. Чому дорівнює потужність множини ірраціональних чисел?
9. Дайте означення зліченної множини
10. Як співвідноситься потужність зліченної множини з потужністю континууму?
11. Що собою являє метричний простір?
12. Що таке метрика простору і які умови вона має задовольняти? Наведіть приклади метрик.
13. Як виглядає відкрита куля в метричному просторі  $C[0, 1]$ ?
14. Які множини є повними і замкнутими?
15. Які множини є досконалими? Наведіть приклад.
16. Дайте означення фундаментальної послідовності в метричному просторі.
17. Який метричний простір є повним? Наведіть приклади повного і неповного метричних просторів.
18. Що таке компактна множина, множина компактна в собі, компакт?
19. Як визначається лінійний метричний простір?
20. Що таке норма простору? Наведіть приклади норм.
21. Дайте означення банахового простору.

## Розділ 2 ГІЛЬБЕРТОВІ ПРОСТОРИ, ЇХ ХАРАКТЕРИСТИКИ ТА АПРОКСИМАЦІЯ ФУНКЦІЙ В НИХ

### 2.1 Гільбертові простори та їх характеристики

*Банахів простір зі скалярним добутком елементів називають гільбертовим простором (за прізвищем математика, який його вивчав) і позначають як  $H$ -простір.*

$H$ -простір може бути скінченновимірним і нескінченновимірним.

Скалярний добуток елементів  $f, g \in H$  записують у вигляді  $(f, g)$  або  $\langle f, g \rangle$ , а для комплексних елементів  $(f, \bar{g})$  або  $\langle f, \bar{g} \rangle$ , де  $\bar{g}$  – спряжене до  $g$  комплексне число. Скалярний добуток повинен задовольняти такі умови:

$$1) \quad \langle f, g \rangle = \langle g, f \rangle, \quad (2.1)$$

а для комплексних елементів –

$$\langle f, \bar{g} \rangle = \langle g, \bar{f} \rangle; \quad (2.2)$$

$$2) \quad \langle \alpha \cdot f, g \rangle = \alpha \cdot \langle f, g \rangle, \quad (2.3)$$

де  $\alpha$  – скаляр; а для комплексних елементів –

$$\langle \alpha \cdot f, \bar{g} \rangle = \alpha \cdot \langle g, \bar{f} \rangle; \quad (2.4)$$

$$3) \quad \langle f, \alpha \cdot g \rangle = \alpha \cdot \langle f, g \rangle, \quad (2.5)$$

де  $\alpha$  – скаляр; а для комплексних елементів –

$$\langle f, \overline{\alpha \cdot g} \rangle = \bar{\alpha} \cdot \langle g, \bar{f} \rangle; \quad (2.6)$$

$$4) \quad \langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle; \quad (2.7)$$

$$5) \quad \langle f, g_1 + g_2 \rangle = \langle f, g_1 \rangle + \langle f, g_2 \rangle; \quad (2.8)$$

$$6) \quad \langle f, f \rangle > 0, \text{ якщо } f \neq 0. \quad (2.9)$$

З означення норми випливає, що для Н-простору

$$\|f\| = \sqrt{\langle f, f \rangle}. \quad (2.10)$$

**Найчастіше Н-простір розглядають в двох реалізаціях.**

**1 Простір  $l_2$  всіх злічених упорядкованих послідовностей  $x \in l_2$ ,**

$$x = \{x_1, x_2, \dots, x_n, \dots\} \quad (2.11)$$

таких, що мають властивість

$$\sum_{i=1}^{\infty} x_i^2 < \infty. \quad (2.12)$$

Для елементів  $x^*, x^{**} \in l_2$ :

$$\text{а) } \rho(x^*, x^{**}) = \sqrt{\sum_{i=1}^{\infty} (x_i^* - x_i^{**})^2}; \quad (2.13)$$

$$\text{б) } \|x^*\| = \sqrt{\sum_{i=1}^{\infty} (x_i^*)^2}; \quad (2.14)$$

$$\text{в) } \|x^* - x^{**}\| = \rho(x^*, x^{**}); \quad (2.15)$$

$$\text{г) } \langle x^*, x^{**} \rangle = \sum_{i=1}^{\infty} x_i^* \cdot x_i^{**}; \quad (2.16)$$

$$\text{д) } \|x^*\| = \sqrt{\langle x^*, x^* \rangle} \quad (2.17)$$

З цих співвідношень випливає, що  $l_2$ -простір – це узагальнення евклідового  $E_n$ -простору при  $n \rightarrow \infty$ .

$l_2$ -простір інколи називають координатним гільбертовим простором.

**2. Простір  $L_2[a, b]$  функцій  $f(t)$  з інтегрованим квадратом,** тобто, для яких

$$\int_a^b f^2(t) dt < \infty. \quad (2.18)$$

Для  $f(t), g(t) \in L_2[a, b]$  є справедливими такі співвідношення:

$$\text{а) } \rho(f, g) = \sqrt{\int_a^b (f(t) - g(t))^2 dt}; \quad (2.19)$$

$$\text{б) } \|f\| = \sqrt{\int_a^b f^2(t) dt}; \quad (2.20)$$

$$\text{в) } \|f - g\| = \rho(f, g); \quad (2.21)$$

$$\text{г) } \langle f, g \rangle = \int_a^b f(t) \cdot g(t) dt. \quad (2.22)$$

Окремо випишемо *дві знамениті нерівності*:  
*нерівність Коші–Мінковського*

$$\|f + g\| \leq \|f\| + \|g\|, \quad (2.23)$$

*нерівність Буняковського–Шварца*

$$|\langle f, g \rangle| \leq \|f\| \cdot \|g\|, \quad (2.24)$$

перша з яких стає очевидною після запису її лівої частини згідно з виразом (2.20) у вигляді

$$\|f + g\| = \sqrt{\int_a^b (f(t) + g(t))^2 dt} = \sqrt{\int_a^b (f^2(t) + 2f(t)g(t) + g^2(t)) dt}, \quad (2.25)$$

з якого витікає, що при від’ємних значеннях однієї з функцій, які входять в їх подвійний добуток, сума квадратів цих функцій зменшується, а тому права частина виразу (2.25), якою задається ліва частина виразу (2.23), стає меншою правої частини виразу (2.23), а друга нерівність стає очевидною, якщо згадати, що геометрично визначений інтеграл (2.22) задає площу під кривою  $f(t)g(t)$ , яка на відрізках значень аргументу  $t$ , на яких одна з функцій, що входять в цей добуток, має від’ємні значення, отримує знак «мінус», зменшуючи загальну площу, що визначається інтегралом (2.22), яким задається ліва частина виразу (2.24), у той час як права частина виразу (2.24), в якій фігурують лише інтеграли від модулів цих функцій, такого зменшення не зазнає.

## 2.2 Ортонормовані підмножини у гільбертових просторах

Розглянемо функціональний гільбертовий простір  $H[a, b]$  такий, що  $x(t), y(t) \in H[a, b], t \in [a, b]$ .

Нехай скалярний добуток  $\langle x, y \rangle$  функцій  $x(t)$  і  $y(t)$  дорівнює нулю, тобто,

$$\langle x, y \rangle = \int_a^b x(t) \cdot y(t) dt = 0. \quad (2.26)$$

**Якщо для функцій  $x(t), y(t) \in H[a, b]$  виконується умова (2.26), то про них кажуть, що вони є ортогональними на  $[a, b]$ .**

Слід зауважити, що для комплексних елементів  $x, y$  гільбертового простору  $H$  скалярний добуток  $\langle x, y \rangle$  має вигляд

$$\langle x, y \rangle = \int_a^b x(t) \cdot \overline{y(t)} dt, \quad (2.27)$$

де символом  $\overline{y(t)}$  позначена комплексна функція, спряжена з  $y(t)$ .

Зрозуміло, що для комплексних елементів  $H$ -простору при перемноженні суттєву роль відіграє порядок запису співмножників у їх скалярному добутку, але якщо ці елементи є ортогональними, то порядок запису значення не має, тобто, в цьому випадку для них виконується умова

$$\int_a^b x(t) \cdot \overline{y(t)} dt = \int_a^b \overline{x(t)} \cdot y(t) dt = 0. \quad (2.28)$$

Нехай ми маємо в гільбертовому просторі  $H[a, b]$  скінченновимірну чи нескінченну послідовність функцій  $\{\varphi_k(t)\}$  таких, що

$$\{\varphi_k(t)\} \subset H[a, b], \quad t \in [a, b]. \quad (2.29)$$

**Якщо для цієї послідовності  $\{\varphi_k(t)\}$  виконується умова**

$$\langle \varphi_k, \varphi_m \rangle = \int_a^b \varphi_k(t) \cdot \varphi_m(t) dt = 0, \quad k \neq m, \quad (2.30)$$

**то цю послідовність називають ортогональною.**

**Якщо для ортогональної послідовності  $\{\varphi_k(t)\} \subset H[a, b]$  виконується умова**

$$\int_a^b \varphi_k^2(t) dt = 1, \quad (2.31)$$

**то цю послідовність називають ортонормованою.**

Послідовність  $\{\varphi_k(t)\} \subset H[a, b]$  називають ортогональною з вагою  $w(t)$ , якщо існує така функція  $w(t) \in H[a, b]$ , яка забезпечує умову

$$\int_a^b \varphi_k(t) \cdot \varphi_m(t) \cdot w(t) dt = 0, \quad k \neq m. \quad (2.32)$$

Зрозуміло, що послідовність  $\{\sqrt{w(t)} \cdot \varphi_k(t)\} \subset H[a, b]$  є просто ортогональною.

Підмножина ортогональних функцій  $\{\varphi_k(t)\} \subset H[a, b]$  є повною в  $H$ -просторі, якщо в ньому не існує відмінної від нуля функції, що була б ортогональною до якоїсь із функцій цієї послідовності.

*Послідовність функцій  $\{\varphi_k(t)\} \subset H[a, b]$  називають замкнутою в  $H$ -просторі, якщо для  $\forall f(t) \in H[a, b]$  і для  $\forall \varepsilon > 0$  можна побудувати таку лінійну комбінацію з функцій  $\varphi_k(t)$ , взятих з вагою  $\lambda_k$ , щоб виконувалась умова*

$$\|f(t) - \lambda_1 \cdot \varphi_1(t) - \lambda_2 \cdot \varphi_2(t) - \dots - \lambda_k \cdot \varphi_k(t) - \dots\| < \varepsilon. \quad (2.33)$$

Це означає, що з похибкою, котра не перевищує  $\varepsilon$ , функцію  $f(t) \in H[a, b]$  на відріжку  $[a, b]$  можна подати у вигляді

$$f(t) \cong \sum_{k=1}^N \lambda_k \cdot \varphi_k(t), \quad (2.34)$$

де  $N$  може бути як кінцевим цілим числом, так і нескінченністю.

*В теорії гільбертових просторів існує поняття спряженого простору. Воно введено у такий спосіб:  $H$ -простір  $L_{p^*}[1 \leq p^* < \infty)$  вважається спряженим з  $H$ -простором  $L_p[1 \leq p < \infty)$ , якщо виконується умова*

$$\frac{1}{p} + \frac{1}{p^*} = 1. \quad (2.35)$$

Математиками встановлено, що кожна послідовність  $\{\varphi_k(t)\} \subset H[a, b]$ , яка є замкнутою в просторі

$$H[a, b] = L_p[1 \leq p < \infty), \quad (2.36)$$

одночасно буде повною у спряженому з ним просторі  $L_{p^*}[1 \leq p^* < \infty)$ . І навпаки, кожна послідовність  $\{\varphi_k(t)\} \subset H[a, b]$ , яка є повною в просторі (2.36), одночасно буде замкнутою в спряженому з ним просторі  $L_{p^*}[1 \leq p^* < \infty)$ .

Звертаємо увагу на те, що в найбільш вживаному варіанті  $H$ -простору, яким є  $L_2[a, b]$  – це простір функцій, інтегрованих зі своїм квадратом (див. (2.18)), поняття замкнутості та повноти збігаються, оскільки при  $p = 2$ , згідно з умовою (2.35)  $p^* = 2$ , тобто, простір  $L_{p^*}$ , спряжений з  $L_p$ , для  $L_2$  збігається з самим  $L_2$ .

### 2.3 Апроксимація неперервних функцій в гільбертових просторах на основі ортогональних послідовностей

*Під апроксимацією неперервних функцій будемо розуміти процес знаходження у вибраному просторі аналітичного опису функцій, заданої елементами якоїсь множини, котра може і не бути підмножиною вибраного простору.* Наприклад, апроксимація багаточленом функції, заданої у вигляді таблиці.

Нехай  $\{\varphi_k(t)\} \subset H[a, b]$ ,  $t \in [a, b]$  – деяка повна послідовність ортонормованих функцій, яка є замкнутою у цьому просторі.

Нехай  $H[a, b] = L[a, b]$  – це  $H$ -простір функцій  $f(t) \in L[a, b]$ , для яких виконується умова

$$\int_a^b |f(t)| dt < \infty, \quad (2.37)$$

а метрика  $\rho(f_1, f_2)$  задається співвідношенням

$$\rho(f_1, f_2) = \int_a^b |f_1(t) - f_2(t)| dt. \quad (2.38)$$

Припустимо, що ряд

$$\sum_k \lambda_k \cdot \varphi_k(t), \quad (2.39)$$

де  $\lambda_k$  – деякий, поки що невідомий нам скаляр, збігається рівномірно до деякої функції  $f(t) \in L[a, b]$ . Це означає, що для  $\forall \varepsilon > 0$  існує таке  $m$ , що для  $\forall t \in [a, b]$  і  $\forall n \geq m$  виконується співвідношення

$$\int_a^b \left| f(t) - \sum_{k=0}^n \lambda_k \cdot \varphi_k(t) \right| dt < \varepsilon, \quad (2.40)$$

з якого при  $n \rightarrow \infty$  випливає, що

$$f(t) = \sum_{k=0}^{\infty} \lambda_k \cdot \varphi_k(t). \quad (2.41)$$

Для визначення вагових коефіцієнтів  $\lambda_k, k = \overline{0, \infty}$  помножимо обидві частини рівняння (2.41) на  $\varphi_j(t)$  і проінтегруємо результат в межах від « $a$ » до « $b$ ». В результаті отримаємо:

$$\int_a^b f(t) \cdot \varphi_j(t) dt = \sum_{k=0}^{\infty} \lambda_k \cdot \int_a^b \varphi_k(t) \cdot \varphi_j(t) dt. \quad (2.42)$$

Оскільки  $\{\varphi_k(t)\} \subset L[a, b]$  є послідовністю ортонормованою, то для неї виконуються співвідношення (2.30) і (2.31). З урахуванням цього з (2.42) маємо

$$\lambda_j = \int_a^b f(t) \cdot \varphi_j(t) dt, \quad j = 0, 1, 2, \dots \quad (2.43)$$

*Вагові коефіцієнти  $\lambda_j$  називають коефіцієнтами Фур'є, а їх повну послідовність  $\{\lambda_j\}$  називають спектром Фур'є розкладу функції  $f(t) \in L[a, b]$  за ортонормованою системою функцій  $\{\varphi_j(t)\} \subset L[a, b]$ .*

Вимога (2.40) рівномірної збіжності ряду (2.39) до функції  $f(t)$  є так званою «сильною вимогою збіжності».

Але виявляється, що в  $H$ -просторі  $L_2[a, b]$  сильна збіжність еквівалентна «збіжності в середньому», яка є більш слабкою вимогою і може бути записана у вигляді

$$\lim_{n \rightarrow \infty} \int_a^b \left[ f(t) - \sum_{k=0}^n \lambda_k \cdot \varphi_k(t) \right]^2 dt = 0. \quad (2.44)$$

Розглянемо процес апроксимації функції  $f(t) \in L_2[a, b]$  в  $H$ -просторі  $L_2[a, b]$  за допомогою ортонормованої послідовності  $\{\varphi_k(t)\} \subset L_2[a, b]$ .

Задачу апроксимації можна звести в цьому випадку до такого підбору коефіцієнтів  $C_k$  часткової суми

$$S_n(t) = \sum_{k=0}^n C_k \cdot \varphi_k(t) \quad (2.45)$$

в  $H$ -просторі  $L_2[a, b]$ , щоб ця сума наближалась до функції  $f(t) \in L_2[a, b]$  з похибкою, що не перевищує задану, тобто, щоб

$$\|f(t) - S_n(t)\| = \sqrt{\int_a^b [f(t) - S_n(t)]^2 dt} \rightarrow \min_{C_k}. \quad (2.46)$$

Але, щоб знайти  $\min_{C_k}$  виразу (2.46), треба розв'язати систему рівнянь



$$\frac{\partial E}{\partial C_k} = 0, \quad k = \overline{0, n}, \quad (2.47)$$

де

$$\begin{aligned} E &= \int_a^b [f(t) - S_n(t)]^2 dt = \\ &= \int_a^b f^2(t) dt - 2 \cdot \int_a^b f(t) \cdot S_n(t) dt + \int_a^b [S_n(t)]^2 dt. \end{aligned} \quad (2.48)$$

Підставляючи (2.45) у (2.48), матимемо

$$E = \int_a^b f^2(t) dt - 2 \cdot \sum_{k=0}^n C_k \cdot \int_a^b f(t) \cdot \varphi_k(t) dt + \int_a^b \left( \sum_{k=0}^n C_k \cdot \varphi_k(t) \right)^2 dt. \quad (2.49)$$

З урахуванням (2.31) і (2.32) співвідношення (2.49) можна переписати так:

$$E = \int_a^b f^2(t) dt - 2 \cdot \sum_{k=0}^n C_k \cdot \lambda_k + \sum_{k=0}^n C_k^2. \quad (2.50)$$

Підставляючи (2.50) у систему рівнянь (2.47), отримаємо:

$$-2 \cdot \lambda_k + 2 \cdot C_k = 0, \quad k = \overline{0, n}. \quad (2.51)$$

З рівнянь (2.51) випливає, що

$$C_k = \lambda_k. \quad (2.52)$$

Тож, щоб часткова сума  $S_n(t)$  апроксимувала функцію  $f(t)$  з заданою точністю, необхідно за коефіцієнти  $C_k$  вибирати коефіцієнти Фур'є  $\lambda_k$  функції  $f(t)$ .

Підставляючи (2.52) у (2.50), матимемо:

$$E = \int_a^b f^2(t) dt - \sum_{k=0}^n \lambda_k^2 \geq 0. \quad (2.53)$$

Зрозуміло, що

$$\lim_{n \rightarrow \infty} E = 0. \quad (2.54)$$

З виразів (2.53), (2.54) випливає, що

$$\int_a^b f^2(t) dt = \sum_{k=0}^{\infty} \lambda_k^2. \quad (2.55)$$



нулю лише в тому випадку, коли його доповнити ваговою функцією  $w(t) \in H[a, b]$ , тобто, поклавши, що

$$\int_a^b \varphi_i(t) \cdot \varphi_j(t) \cdot w(t) dt = \delta_{ij}, \quad i = \overline{1, n}, \quad j = \overline{1, n}, \quad (2.59)$$

де

$$\delta_{ij} = \begin{cases} 1 & \text{при } i = j, \\ 0 & \text{при } i \neq j \end{cases} \quad (2.60)$$

символ Кронекера.

На першому кроці методики Грама–Шмідта покладемо, що

$$\varphi_1(t) = \frac{f_1(t)}{\sqrt{\int_a^b f_1^2(t) \cdot w(t) dt}}. \quad (2.61)$$

Підставляючи (2.61) у співвідношення (2.59) при  $i = j = 1$  бачимо, що, задаючи  $\varphi_1(t)$  у вигляді (2.61), ми одразу отримуємо перший член послідовності  $\{\varphi_k(t)\}$  у нормованому вигляді.

На другому кроці методики Грама–Шмідта задамо функцію

$$\psi_2(t) = a_{21} \cdot \varphi_1(t) + f_2(t), \quad (2.62)$$

до якої висуватимемо вимоги, щоб вона була ортогональною до  $\varphi_1(t)$ . В цьому випадку буде справедливим рівняння

$$\int_a^b \psi_2(t) \cdot \varphi_1(t) \cdot w(t) dt = 0, \quad (2.63)$$

підставляючи в яке вираз (2.62), отримаємо

$$a_{21} \cdot \int_a^b \varphi_1^2(t) \cdot w(t) dt + \int_a^b f_2(t) \cdot \varphi_1(t) \cdot w(t) dt = 0. \quad (2.64)$$

З урахуванням того, що

$$\int_a^b \varphi_1^2(t) \cdot w(t) dt = 1, \quad (2.65)$$

із виразу (2.64) матимемо

$$a_{21} = -\int_a^b f_2(t) \cdot \varphi_1(t) \cdot w(t) dt. \quad (2.66)$$

Підставляючи (2.66) у (2.62), отримаємо функцію  $\psi_2(t)$ , яка буде ортогональною до  $\varphi_1(t)$ .

Для отримання ортонормованої функції  $\varphi_2(t)$ , яка стане другим членом послідовності  $\{\varphi_k(t)\}$ , пов'яжемо функцію  $\varphi_2(t)$  з функцією  $\psi_2(t)$  у такий спосіб

$$\varphi_2(t) = \frac{\psi_2(t)}{\sqrt{\int_a^b \psi_2^2(t) \cdot w(t) dt}}. \quad (2.67)$$

Підставляючи (2.67) у (2.59) при  $i=1, j=2$  та при  $i=2, j=2$ , легко переконатись, що функція  $\varphi_2(t)$ , якщо вона задовольняє умови (2.62), (2.66) і (2.67), буде і ортогональною до  $\varphi_1(t)$ , і нормованою.

Узагальнюючи, можна стверджувати, що для кроку з номером  $k$  справедливими будуть співвідношення

$$\psi_k(t) = a_{k1} \cdot \varphi_1(t) + a_{k2} \cdot \varphi_2(t) + \dots + a_{k(k-1)} \cdot \varphi_{k-1}(t) + f_k(t), \quad k = \overline{2, n}, \quad (2.68)$$

$$a_{kj} = -\int_a^b f_k(t) \cdot \varphi_j(t) \cdot w(t) dt, \quad k = \overline{2, n}, \quad (2.69)$$

$$\varphi_k = \frac{\psi_k(t)}{\sqrt{\int_a^b \psi_k^2(t) \cdot w(t) dt}}, \quad k = \overline{2, n}. \quad (2.70)$$

Як бачимо з наведених кроків, процедура Грама–Шмідта є рекурентною, тобто, кожний наступний член послідовності  $\{\varphi_k(t)\}$  знаходиться з використанням інформації, отриманої на попередніх кроках.

Різними математиками для базових функцій

$$f_k(t) = t^k, \quad k = \overline{0, n} \quad (2.71)$$

отримано різноманітні системи ортонормованих поліномів для різних функцій ваги та інтервалів ортогоналізації. Тому немає необхідності кожний раз, коли необхідно апроксимувати функцію  $f(t) \in H[a, b]$  за допомогою ортонормованої послідовності  $\{\varphi_k(t)\} \subset H[a, b]$ , самому будувати цю послідовність. Достатньо вибрати одну з тих,

що побудовані іншими, скориставшись довідником з вищої математики чи посібником з математичної теорії обробки результатів експериментів.

Наведемо приклади інтервалів ортогоналізації, вагових функцій і нормувальних множників найбільш розповсюджених систем ортогональних поліномів (табл. 2.1).

Таблиця 2.1 – Приклади інтервалів ортогоналізації, вагових функцій і нормувальних множників найбільш розповсюджених систем ортогональних поліномів

Поліноми	Інтервал ортогоналізації	Вагова функція $W(t)$	Нормувальний множник
Лежандра $P_k(t)$	$t \in [-1, 1]$	1	$\int_{-1}^1 (P_k(t))^2 dt = \frac{2}{2k+1}$
Чебишова I $T_k(t)$	$t \in [-1, 1]$	$(1-t^2)^{-\frac{1}{2}}$	$\int_{-1}^1 (T_k(t))^2 (1-t^2)^{-\frac{1}{2}} dt = \begin{cases} \frac{\pi}{2}, & k \neq 0 \\ \pi, & k = 0 \end{cases}$
Чебишова II $U_k(t)$	$t \in [-1, 1]$	$(1-t^2)^{\frac{1}{2}}$	$\int_{-1}^1 (U_k(t))^2 (1-t^2)^{\frac{1}{2}} dt = \frac{\pi}{2}$
Лагерра $L_k(t)$	$t \in [0, \infty)$	$e^{-t}$	$\int_0^{\infty} (L_k(t))^2 e^{-t} dt = 1$
Лагерра приєднані $L_k^{(i)}(t)$	$t \in [0, \infty)$	$t^i \cdot e^{-t}$	$\int_0^{\infty} (L_k^{(i)}(t))^2 t^i e^{-t} dt = \frac{(k+1)!}{k!}$
Ерміта $H_k(t)$	$t \in (-\infty, \infty)$	$e^{-t^2}$	$\int_{-\infty}^{\infty} (H_k(t))^2 e^{-t^2} dt = 2^k \cdot k! \cdot \sqrt{\pi}$

Таким чином, для того, щоб здійснити апроксимацію функції  $f(t) \in H[a, b]$ ,  $t \in [a, b]$  за допомогою ортонормованої системи поліномів  $\{\varphi_k(t)\} \subset H[a, b]$ , необхідно, виходячи з інтервалу ортогоналізації  $[a, b]$  та зручності вагової функції  $w(t)$ , вибрати ту чи іншу ортонормовану систему поліномів з довідника та знайти співвідношення для загального члена вибраної системи, розкриваючи яке для  $k = 1, 2, 3, \dots, n$  отримати таку кількість її членів, якої достатньо для забезпечення заданої точності апроксимації.

Для прикладу наведемо перші три члени кількох ортонормованих послідовностей з наведених в таблиці 2.1:

– для поліномів Лежандра:

$$P_0(t) = 1, P_1(t) = t, P_2(t) = \frac{1}{2} \cdot (3t^2 - 1); \quad (2.72)$$

– для поліномів Чебишова:

$$T_0(t) = 1, T_1(t) = t, T_2(t) = 2t^2 - 1; \quad (2.73)$$

$$U_0(t) = 1, U_1(t) = 2t, U_2(t) = 4t^2 - 1; \quad (2.74)$$

– для поліномів Лагерра:

$$L_0(t) = 1, L_1(t) = 1 - t, L_2(t) = 1 - 2t + \frac{t^2}{2}. \quad (2.75)$$

Зауважимо, що всі системи ортонормованих поліномів із числа наведених у таблиці 2.1 мають дві спільні властивості:

- 1) всі вони отримані, виходячи з базових функцій (2.71);
- 2) всі вони є розв'язками гіпергеометричного диференціального рівняння 2-го порядку, яке має вигляд

$$t \cdot (t-1) \cdot y'' + [\xi - (\eta + \nu + 1) \cdot t] \cdot y' - \eta \cdot \nu \cdot y = 0. \quad (2.76)$$

Гаусс знайшов розв'язок рівняння (2.76) у вигляді ряду

$$y(t) = 1 + \frac{\eta \cdot \nu}{\xi} \cdot \frac{t}{1!} + \frac{\eta \cdot (\eta + 1) \cdot \nu \cdot (\nu + 1)}{\xi \cdot (\xi + 1)} \cdot \frac{t^2}{2!} + \dots \quad (2.77)$$

Залежно від значень коефіцієнтів  $\eta, \nu, \xi$  і відрізка  $[a, b]$ , на якому знаходиться розв'язок, із ряду (2.77) можна отримати всі ортонормовані поліноми з числа наведених у таблиці 2.1.

Крім того, можна ще отримати функції Якобі–Каутца і функції Куммера. Але ці функції, в свою чергу, виражаються через такі спеціальні функції, як гамма-функції і функції Бесселя, тому, аби не ускладнювати викладки, ми їх вирішили в цьому навчальному посібнику не наводити.

На завершення цього підрозділу наведемо деякі міркування стосовно вибору кількості членів в ортонормованих рядах (2.45) при апроксимації функцій  $f(t) \in H[a, b]$  частковими сумами  $S_n(t)$ .

В прикладному аналізі функції  $f(t)$ , які необхідно апроксимувати, як правило, отримують шляхом вимірювань. А тому замість функції  $f(t)$  ми маємо справу з функціями

$$f^*(t) = f(t) + h(t), \quad (2.78)$$

в яких  $h(t)$  – «шум» вимірювань.

Тож, ортонормований ряд, яким ми будемо апроксимувати  $f^*(t)$ , матиме дві складові: регулярну і шумову. Коефіцієнти  $\lambda_k$  регулярної складової при  $k \rightarrow \infty$  будуть зменшуватись до нуля, а коефіцієнти шумової складової, починаючи з деякого значення  $k$ , при подальшому збільшенні ( $k \rightarrow \infty$ ) практично перестають змінюватись.

Важливою задачею є визначення такого значення  $k = n$  в частковій сумі  $S_n(t)$ , завдяки якому можна відкинути «високочастотну» частину спектра Фур'є функції  $f^*(t)$ , що обумовлена наявністю в цій функції шумової складової  $h(t)$ .

Доведено, що якщо функція  $f(t)$  є абсолютно неперервною (тобто,  $f(t) \in C[a, b]$  і має обмежену похідну), то порядок спадання її коефіцієнтів Фур'є  $\lambda_k$  є функцією, оберненою до номера коефіцієнта, тобто,  $O\left(\frac{1}{k}\right)$ .

Тож, в околі коефіцієнта Фур'є  $\lambda_k$ , для якого наведене вище твердження порушується, ряд  $\sum_{k=0}^{\infty} \lambda_k \cdot \varphi_k(t)$  можна обривати, перетворюючи його на часткову суму  $S_n(t)$ , що відображає лише регулярну складову функції  $f(t)$ , яка апроксимується цим рядом.

## 2.5 Завдання для самоперевірки

1. Дайте означення гільбертового простору. Наведіть приклади гільбертових просторів та співвідношень, за допомогою яких записуються скалярні добутки в цих просторах.
2. Які умови має задовольняти скалярний добуток у гільбертовому просторі?
3. Як пов'язані між собою метрика і норма, норма і скалярний добуток в гільбертовому просторі?
4. Що собою являє ортонормована послідовність функцій в гільбертовому просторі?
5. Яка ортогональна послідовність в гільбертовому просторі є повною? Яка є замкнутою? В якій реалізації гільбертового простору повнота і замкнутість ортогональної послідовності є еквівалентами?
6. Що собою являє спряжений гільбертів простір?
7. Дайте означення процесу апроксимації функції.
8. Як апроксимувати неперервну функцію в просторі  $L[a, b]$ ?
9. Як апроксимувати неперервну функцію в просторі  $L_2[a, b]$ ?
10. Виведіть рівність Парсеваля. Дайте її геометричну інтерпретацію.
11. Що собою являє спектр Фур'є неперервної функції в гільбертовому просторі?
12. Для чого використовується методика Грама–Шмідта?
13. Як побудувати ортогональну послідовність в гільбертовому просторі? Наведіть конкретну процедуру, знайшовши перших два члени послідовності.
14. Які ортогональні послідовності на базі степеневих функцій ви знаєте?
15. Що собою являють гіпергеометричне диференціальне рівняння і його розв'язок? Що встановив Гаусс, аналізуючи цей розв'язок?
16. Як відсікти шумову складову при апроксимації функції в гільбертовому просторі?

### Розділ 3

## МІРА ЖОРДАНА І ЛЕБЕГА ДЛЯ МНОЖИН І ПРОСТОРІВ

### ТА ІНТЕГРАЛИ РІМАНА, СТИЛТЬЄСА І ЛЕБЕГА

#### 3.1 Міра Жордана для множин і просторів

На усій числовій осі чи на її відрізку  $[a,b]$ , яким є і відрізок  $[0,1]$ , на яких задано певну множину, кожен елемент цієї множини відображається точкою, яка, як відомо, має нульову довжину. У зв'язку з цим виникають два взаємопов'язані запитання: «А яку ж довжину відрізка числової осі з власною довжиною  $b-a$  чи відрізка з одиничною довжиною займає задана на цьому відрізку множина? І як виміряти довжину тієї частини відрізка, яку займають точки цієї множини?».

Математики дали вичерпні відповіді на ці запитання, ввівши спочатку поняття міри Жордана для множин точок, заданих на відрізку числової осі, на частині площини двовимірного простору та на частині об'єму тривимірного простору, а потім увівши поняття більш узагальненої міри Лебега та розповсюдивши її і на функції, що задані на цьому відрізку, на цій частині площини чи на цьому об'ємі.

Отже, почнемо вивчення торії міри для множин і просторів з розгляду поняття міри, введеного Жорданом. Для цього розглянемо числову множину  $E$ , задану на відрізку  $[a,b]$ , обмеженому цілими числами, завдяки чому матимемо

$$b - a = k, \quad (3.1)$$

де  $k$  – ціле число. Розділимо відрізок  $b-a$  на  $k$  рівних частин, які назвемо сегментами 1-го рангу, кожен із яких, вочевидь, має одиничну довжину. Підрахуємо кількість  $L_1$  сегментів 1-го рангу, в кожен із яких потрапила хоча б одна точка множини  $E$ , а також підрахуємо кількість  $l_1$  сегментів 1-го рангу, кожен із яких виявився суцільно заповненим точками множини  $E$ . Цілком очевидно, що отримаємо нерівність

$$L_1 > l_1, \quad (3.2)$$

оскільки суцільно заповнені точками множини  $E$  сегменти є одночасно і сегментами, в які потрапила хоча б одна точка цієї множини. Звертаємо увагу на те, що, завдяки тому, що сегменти 1-го рангу мають одиничну довжину, то кількості  $L_1$  та  $l_1$  задають, відповідно, сумарну довжину сегментів 1-го рангу, в які потрапила хоча б одна точка множини  $E$ , що задана на відрізку  $[a,b]$ , та сумарну довжину сегментів 1-го рангу, суцільно заповнених точками цієї множини. Нехай для прикладу ми отримали

$$L_1 = 5, \quad l_1 = 0. \quad (3.3)$$

А далі розіб'ємо кожен сегмент 1-го рангу на дві рівні частини і отримаємо на відрізку  $[a,b]$   $2k$  сегментів 2-го рангу. Після цього підрахуємо кількість  $L_2^*$  сегментів 2-го рангу, в кожен із яких потрапила хоча б одна точка множини  $E$ , а також підрахуємо кількість  $l_2^*$  сегментів 2-го рангу, кожен із яких виявився суцільно заповненим точками множини  $E$ . Цілком очевидно, що отримаємо нерівність



$$L_2^* > l_2^*, \quad (3.4)$$

знову ж таки, оскільки суцільно заповнені точками множини  $E$  сегменти 2-го рангу є одночасно і сегментами 2-го рангу, в які потрапила хоча б одна точка цієї множини. Звертаємо увагу на те, що завдяки тому, що сегменти 1-го рангу мають одиничну довжину, кожному з яких для отримання сегментів 2-го рангу ми ділили на 2, то для отримання, відповідно, сумарної довжини  $L_2$  сегментів 2-го рангу, в які потрапила хоча б одна точка множини  $E$ , що задана на відрізку  $[a, b]$ , та сумарної довжини  $l_2$  сегментів 2-го рангу, суцільно заповнених точками цієї множини, кількості  $L_2^*$  та  $l_2^*$ , отримані для цих сегментів, потрібно ділити на 2, тобто,

$$L_2 = \frac{L_2^*}{2}, \quad l_2 = \frac{l_2^*}{2}. \quad (3.5)$$

Нехай, для прикладу, ми отримали

$$L_2^* = 8, \quad l_2^* = 2, \quad (3.6)$$

Тоді, згідно з виразами (3.5), отримаємо

$$L_2 = 4, \quad l_2 = 1. \quad (3.7)$$

А далі розіб'ємо кожен сегмент 2-го рангу на дві рівні частини і отримаємо на відрізку  $[a, b]$   $4k$  сегментів 3-го рангу. Після цього підрахуємо кількість  $L_3^*$  сегментів 3-го рангу, в кожен із яких потрапила хоча б одна точка множини  $E$ , а також підрахуємо кількість  $l_3^*$  сегментів 3-го рангу, кожен із яких виявився суцільно заповненим точками множини  $E$ . Цілком очевидно, що отримаємо нерівність

$$L_3^* > l_3^*, \quad (3.8)$$

знову ж таки, оскільки суцільно заповнені точками множини  $E$  сегменти 3-го рангу є одночасно і сегментами 3-го рангу, в які потрапила хоча б одна точка цієї множини. Знову ж таки, звертаємо увагу на те, що завдяки тому, що сегменти 1-го рангу мають одиничну довжину, кожному з яких для отримання сегментів 2-го рангу ми ділили на 2, кожен із яких ми, у свою чергу, знову ділили на 2, то для отримання, відповідно, сумарної довжини  $L_3$  сегментів 3-го рангу, в які потрапила хоча б одна точка множини  $E$ , що задана на відрізку  $[a, b]$ , та сумарної довжини  $l_3$  сегментів 3-го рангу, суцільно заповнених точками цієї множини, кількості  $L_3^*$  та  $l_3^*$ , отримані для цих сегментів, потрібно ділити на 4, тобто,

$$L_3 = \frac{L_3^*}{4}, \quad l_3 = \frac{l_3^*}{4}. \quad (3.9)$$

Нехай, для прикладу, ми отримали

$$L_3^* = 12, \quad l_3^* = 8, \quad (3.10)$$

Тоді, згідно з виразами (3.9), отримаємо

$$L_3 = 3, \quad l_3 = 2. \quad (3.11)$$

Продовжуючи цей процес шляхом переходу до сегментів 4-го, 5-го і аж до  $n$ -го рангу, ми отримаємо дві послідовності чисел:

$$L_1, L_2, L_3, \dots, L_n; \quad (3.12)$$

$$l_1, l_2, l_3, \dots, l_n, \quad (3.13)$$

перша з яких (3.12), як видно з виразів (3.3), (3.7), (3.11), є спадною і називається верхньою послідовністю Жордана, а друга (3.13), як видно з тих же виразів (3.3), (3.7), (3.11), є висхідною і називається нижньою послідовністю Жордана. Очевидно, що і верхня і нижня послідовності Жордана є обмеженими, причому граничним елементом верхньої послідовності є число з нижньої послідовності, а граничним елементом нижньої послідовності, навпаки, є число з верхньої послідовності.

**Означення:** *граничне число, до якого прямує верхня послідовність Жордана, що отримана для множини  $E$ , заданої на відрізку  $[a, b]$  числової осі, називається зовнішньою мірою множини  $E$  за Жорданом і позначається  $mes^* E$ , а граничне число, до якого прямує нижня послідовність, називається внутрішньою мірою множини  $E$  за Жорданом і позначається  $mes_* E$ . Якщо зовнішня і внутрішня міри множини  $E$  за Жорданом збігаються, то множина  $E$  називається вимірною за Жорданом, а спільне значення його зовнішньої та внутрішньої мір за Жорданом*

$$mes^* E = mes_* E = mes E \quad (3.14)$$

*називається мірою множини  $E$  за Жорданом і позначається  $mes E$ .*

Із виразу (3.14) витікає, що множина  $E$  є вимірною за Жорданом лише тоді, коли різниця зовнішньої та внутрішньої мір множини  $E$  дорівнює нулеві, тобто,

$$mes^* E - mes_* E = \lim_{n \rightarrow \infty} (L_n - l_n) = 0. \quad (3.15)$$

Із означення понять множини та міри Жордана витікає, що:

1) якщо множини  $E_1, E_2$  є вимірними за Жорданом, то і множини  $E_1 \cup E_2, E_1 \cap E_2, E_1 - E_2$  теж є вимірними за Жорданом;

2) якщо множини  $E_1, E_2$  є вимірними за Жорданом і не мають спільних внутрішніх точок, то

$$mes(E_1 \cup E_2) = mes E_1 + mes E_2; \quad (3.16)$$

3) якщо множини  $E_1, E_2$  є вимірними за Жорданом і  $E_1 \supset E_2$ , то

$$\text{mes}(E_1 - E_2) = \text{mes } E_1 - \text{mes } E_2; \quad (3.17)$$

4) якщо множина  $E$  задана в обмеженій області площини, то, обмежуючи цю область прямокутником зі сторонами, що мають цілочислову довжину, розбиваючи цей прямокутник на одиничні квадрати, які визначатимемо як сегменти 1-го рангу, та застосовуючи процедуру, описану вище при послідовній сегментації відрізка числової осі, але в інтерпретації числового визначення площ цих сегментів, які містять в собі хоча б одну точку множини  $E$  та які суцільно заповнені цими точками, отримаємо знову ж таки верхню та нижню послідовності Жордана з чисел, що характеризуватимуть площі сегментів відповідних рангів, які визначатимуть зовнішню та внутрішню міри множини  $E$  за Жорданом, котрі в граничному варіанті приводять до визначення міри цієї множини  $E$  за Жорданом на площині, при цьому вимірні за Жорданом множини точок, заданих на площині, називають квадрикованими;

5) якщо множина  $E$  задана в обмеженій області тривимірного простору, то, поміщаючи цю область в прямокутний паралелепіпед зі сторонами, що мають цілочислову довжину, розбиваючи цей прямокутний паралелепіпед на одиничні куби, які визначатимемо як сегменти 1-го рангу, та застосовуючи процедуру, описану вище при послідовній сегментації відрізка числової осі, але в інтерпретації числового визначення об'ємів цих сегментів, які містять в собі хоча б одну точку множини  $E$  та які суцільно заповнені цими точками, отримаємо знову ж таки верхню та нижню послідовності Жордана з чисел, що характеризуватимуть об'єми сегментів відповідних рангів, які визначатимуть зовнішню та внутрішню міри множини  $E$  за Жорданом, котрі в граничному варіанті приводять до визначення міри цієї множини  $E$  за Жорданом у тривимірному просторі, при цьому вимірні за Жорданом множини точок, заданих у тривимірному просторі, називають кубовуваними.

А на завершення матеріалу, присвяченого мірі Жордана, визначимо міру Жордана зліченої множини раціональних чисел, заданих, наприклад, на відрізку  $[0,1]$  числової осі. Оскільки при будь-якій сегментації цього відрізка числової осі кожен сегмент матиме, принаймні, одну точку, яка є проекцією на вісь раціонального числа, то верхня послідовність Жордана буде послідовністю одиниць, оскільки

$$L_1 = 1, \quad L_2 = 1, \quad L_3 = 1, \dots, L_n = 1. \quad (3.18)$$

І у той же час при цій же сегментації даного відрізка числової осі ми не матимемо жодного сегмента, суцільно заповненого точками, оскільки всі раціональні числа «оточені» ірраціональними числами, множина яких має потужність континууму, що є набагато більшою потужності зліченої множини, а тому яким би малим ми не взяли сегмент числової осі, на ньому поряд із точками, які є проекціями раціональних чисел, знаходяться і точки, проекціями яких є ірраціональні числа. А тому нижня послідовність Жордана буде послідовністю нулів, оскільки

$$l_1 = 0, \quad l_2 = 0, \quad l_3 = 0, \dots, l_n = 0. \quad (3.19)$$

І оскільки верхня (3.18) і нижня (3.19) послідовності Жордана не збігаються до спільного граничного числа, то, згідно з означенням міри Жордана, множина раціональних чисел на числовій осі не є вимірною за Жорданом. І той же самий висновок можна зробити і щодо множини ірраціональних чисел.

Цей висновок змусив математиків шукати інший спосіб вимірювати множини і простори, який би розповсюджувався на всі, без винятку, випадки. І таку міру побудував Лебег. Теорію міри Лебега ми викладемо у наступному підрозділі.

### 3.2 Міра Лебега для множин і просторів

Як ми уже відзначали у попередніх підрозділах, потужність множини дійсних чисел і на всій числовій осі, і на її відрізку  $[a, b]$ , який можна розглядати і у варіанті відрізка  $[0, 1]$ , дорівнює потужності континууму  $c$ , а тому для зручності викладок, присвячених теорії міри Лебега, саме множини  $E$ , задану на відрізку  $[0, 1]$ , ми і розглянемо.

Нагадаємо, що мірою відрізка  $[a, b]$ , як і інтервалу  $(a, b)$ , є його довжина  $b - a$ , яка для відрізка  $[0, 1]$  та інтервалу  $(0, 1)$  дорівнює одиниці. Очевидно, що мірою суми інтервалів  $\alpha_i$ , що не перетинаються, але кожен із яких є підмножиною відрізка  $[0, 1]$ , буде сума  $\sum_i \alpha_i$  довжин цих інтервалів.

Нехай обмежена числова множина  $E$  є підмножиною одиничного відрізка  $[0, 1]$  числової осі, тобто,  $E \subset [0, 1]$ . Позначимо підмножину, яка доповнює множину  $E$  до одиничного відрізка числової осі, символом  $CE$ .

Конкретизуємо множину  $E$ , задавши її структуру у вигляді зкінченної

$$E = \{x_1, x_2, \dots, x_n\} \quad (3.20)$$

чи зліченної

$$E = \{x_1, x_2, \dots, x_n, \dots\} \quad (3.21)$$

числової послідовності.

Накриємо кожену точку  $x_i$  послідовностей (3.20) чи (3.21) інтервалом  $\alpha_i$ , що не перетинається з іншими інтервалами, якими накриваються інші точки цих послідовностей, та знайдемо суму  $\sum_i \alpha_i$  довжин цих інтервалів для кожної з цих послідовностей, які дорівнюватимуть, відповідно,

$$\sum_{i=1}^n \alpha_i = \alpha_1 + \alpha_2 + \dots + \alpha_n, \quad (3.22)$$

$$\sum_{i=1}^{\infty} \alpha_i = \alpha_1 + \alpha_2 + \dots + \alpha_n + \dots \quad (3.23)$$

Оскільки точки на числовій осі не мають довжини, то інтервали, якими вони накриваються в межах одиничного відрізка цієї осі, можуть бути дуже малими, а тому їх сума завжди буде мати нижню границю, яку позначимо як

$$m^* E = \inf \sum_i \alpha_i. \quad (3.24)$$

Цю нижню границю, що задається виразом (3.24) і позначається символом  $m^* E$ , названо *зовнішньою мірою множини  $E$* . Очевидно, що аналогічним чином можна визначити і *зовнішню міру  $m^* CE$  множини  $CE$* , яка доповнює множину  $E$  до відрізка  $[0, 1]$  з одиничною довжиною, на якому вони обидві задані.

Різницю між довжиною відрізка  $[0,1]$  та зовнішньою мірою  $m^*CE$  множини  $CE$ , яка доповнює множину  $E$  до цього відрізка  $[0,1]$

$$m_*E = 1 - m^*CE, \quad (3.25)$$

названо *внутрішньою мірою множини  $E$*  і позначено символом  $m_*E$ .

**Означення:** якщо зовнішня міра  $m^*E$  множини  $E$  та її внутрішня міра  $m_*E$  збігаються й дорівнюють одному і тому ж числу  $mE$ , тобто,

$$m^*E = m_*E = mE, \quad (3.26)$$

то множина  $E$  називається *вимірною за Лебегом*, а це число  $mE$  називається *мірою множини  $E$  за Лебегом*.

Очевидно, що з цього твердження, як наслідок, витікає, що і множина  $CE$ , яка доповнює множину  $E$  до одиничного відрізка на числовій осі, теж є вимірною за Лебегом, і для її міри  $mCE$  теж є справедливим вираз

$$m^*CE = m_*CE = mCE, \quad (3.27)$$

подібний до (3.26).

Для прикладу знайдемо міри Лебега для множин дійсних  $Z$ , раціональних  $R$  та ірраціональних  $\bar{R}$  чисел, заданих на відрізку  $[0,1]$ , які пов'язані між собою співвідношенням

$$Z = R \cup \bar{R}. \quad (3.28)$$

Відрізок  $[0,1]$  числової осі містить в собі щільно розміщені точки, які є проєкціями на нього дійсних чисел  $Z$ , тому мірою Лебега множини  $Z$ , яка має потужність континууму  $c$ , на цьому відрізку є його довжина, тобто,

$$mZ = 1. \quad (3.29)$$

Як ми уже знаємо, множина раціональних чисел  $R$  є зліченною множиною, тому її можна записати у вигляді нескінченної послідовності (3.21). Виберемо величину  $\varepsilon > 0$  і покриємо числа зліченної послідовності (3.21) зліченною послідовністю інтервалів (3.23), вибравши за перший інтервал величину  $\varepsilon$ , за другий – величину  $\frac{\varepsilon}{2}$ , яку можна подати і як  $\frac{\varepsilon}{2^1}$ , за третій – величину  $\frac{\varepsilon}{4}$ , яку можна подати як  $\frac{\varepsilon}{2^2}$ , за  $n$ -ий – величину  $\frac{\varepsilon}{2^{n-1}}$ . У цьому випадку вираз (3.23) можна переписати так:

$$\sum_{i=1}^{\infty} \alpha_i = \varepsilon + \frac{\varepsilon}{2^1} + \frac{\varepsilon}{2^2} + \dots + \frac{\varepsilon}{2^{n-1}} + \dots = \varepsilon \left( 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} + \dots \right) = \varepsilon \left( \frac{1}{1 - \frac{1}{2}} \right) = 2\varepsilon. \quad (3.30)$$

Звертаємо увагу на те, що в дужках у середній частині виразу (3.30) ми маємо суму членів спадної геометричної прогресії, яка, як відомо ще з шкільного курсу математики, дорівнює першому її члену, поділеному на одиницю, мінус знаменник, тобто мінус те число, на яке потрібно помножити попередній член геометричної прогресії, щоб отримати наступний.

Оскільки, як ми уже відмічали, точка на числовій осі не має ширини, то і величину  $\varepsilon$  можна задати в околі нуля, а тому, з урахуванням виразів (3.24) і (3.30), отримуємо

$$m^*R = \inf \sum_{i=1}^{\infty} \alpha_i = \lim_{\varepsilon \rightarrow 0} (2\varepsilon) = 0. \quad (3.31)$$

Множина  $\bar{R}$  ірраціональних чисел, заданих на відрізку  $[0,1]$ , яка доповнює множину раціональних чисел  $R$  до множини дійсних чисел  $Z$  на цьому одиничному відрізку числової осі, як ми уже знаємо з попередніх підрозділів, як і множина дійсних чисел  $Z$ , теж має потужність континууму  $c$ , наслідком чого є те, що мірою Лебега множини  $\bar{R}$  ірраціональних чисел теж є довжина відрізка, тобто,

$$m\bar{R} = 1 \quad (3.32)$$

А на основі виразів (3.26) і (3.27) та виразу (3.32) ми маємо право записати, що

$$m\bar{R} = m^*\bar{R} = m_*\bar{R} = 1 = m^*CR \quad (3.33)$$

З урахуванням виразу (3.33) з виразу (3.25) отримуємо

$$m_*R = 1 - m^*CR = 1 - 1 = 0. \quad (3.34)$$

Тож із виразів (3.26), (3.31) та (3.34) витікає, що міра Лебега множини раціональних чисел  $R$  на відрізку  $[0,1]$  дорівнює нулю, а у зв'язку зі зліченною множиною таких відрізків на усій числовій осі, нульвимірною ця множина є і на усій числовій осі.

А як наслідок із міркувань, наведених вище, витікає таке означення: **множина  $E$  є нульвимірною за Лебегом, якщо для будь-якого  $\eta > 0$  її можна покрити скінченною або зліченною системою інтервалів з сумарною довжиною, меншою  $\eta$ .**

Із цього означення витікає, що будь-яка множина, елементами якої є окремі точки, незалежно від того чи їх кількість визначається конкретним числом чи цих точок є стільки ж, скільки і чисел натурального ряду, належить до класу нульвимірних, тобто, до класу вимірних, але з мірою Лебега, що дорівнює нулю. А це, у свою чергу, означає, що якщо якась функція, що задана на відрізку  $[a,b]$ , має на цьому відрізку не більше ніж зліченну множину точок, в яких функція терпить розриви 1-го роду, то множина точок розриву цієї функції є нульвимірною.

Із означення понять множини та міри Лебега витікає, що:

1) якщо множини  $E_1, E_2$  є вимірними за Лебегом, то і множини  $E_1 \cup E_2, E_1 \cap E_2, E_1 - E_2$  теж є вимірними за Лебегом;

2) якщо множини  $E_1, E_2$  є вимірними за Лебегом і не мають спільних внутрішніх точок, то

$$m(E_1 \cup E_2) = m E_1 + m E_2; \quad (3.35)$$

3) якщо множини  $E_1, E_2$  є вимірними за Лебегом і  $E_1 \supset E_2$ , то

$$m(E_1 - E_2) = m E_1 - m E_2; \quad (3.36)$$

4) якщо множина  $E$  задана в обмеженій області площини, яка являє собою двовимірний простір, то поміщаючи кожну її точку  $x_i$  у відкритий прямокутник з площею  $\alpha_i^2$  та застосовуючи процедуру, описану вище для множини, точки якої розміщені на відрізку числової осі, дійдемо до чисел, що визначатимуть зовнішню та внутрішню міри множини  $E$  за Лебегом на площині, котрі приводять до визначення міри цієї множини  $E$  за Лебегом у двовимірному просторі;

5) якщо множина  $E$  задана в обмеженій області тривимірного простору, то поміщаючи кожну її точку  $x_i$  у відкритий куб об'ємом  $\alpha_i^3$  та застосовуючи процедуру, описану вище для множини, точки якої розміщені на відрізку числової осі, дійдемо до чисел, що визначатимуть зовнішню та внутрішню міри множини  $E$  за Лебегом у тривимірному просторі, котрі приводять до визначення міри цієї множини  $E$  за Лебегом у цьому просторі.

**Важливе зауваження:** після введення Лебегом свого означення поняття міри на множині чи у просторі практично в усіх математичних дослідженнях, пов'язаних з застосуванням поняття міри, використовується саме міра Лебега, тому математики домовились, застосовуючи термін «міра», не вказувати кожен раз, що це міра Лебега, а розуміючи, що це саме так і без згадування цього прізвища біля цього терміна. Але, якщо хтось використовує поняття міри за Жорданом, яке стало використовуватись рідко, то він прізвище Жордана при терміні «міра» зобов'язаний вказувати.

А тепер розповсюдимо поняття міри і на функції  $f(x)$  дійсної змінної, що визначені на вимірних множинах  $E$  свого аргументу  $x$ .

**Означення.** Функція  $f(x)$  дійсної змінної  $x$ , що визначена на вимірній множині  $E$  значень свого аргументу  $x$ , належить до класу вимірних функцій, якщо для будь-якого числа  $A$  підмножина  $E_A \subset E$ , для усіх елементів  $x$  якої  $f(x) > A$ , є вимірною.

Звертаємо увагу на те, що вимірність функції визначається через вимірність множини значень її аргументу. Тож, якщо навіть множина усіх значень аргументу функції буде мати міру «нуль», ця функція буде належати класу вимірних. А тому, як наслідок з такого означення вимірності функції, маємо:

1) функція  $f(x)$ , яка є неперервною на замкнутій обмеженій множині  $E \subset (-\infty, \infty)$  свого аргументу  $x$ , є вимірною на цій множині;

2) сума, різниця і добуток двох вимірних функцій  $f(x), g(x)$ , заданих на вимірній множині  $E$  свого аргументу  $x$ , є функцією вимірною на цій множині, а частка від ділення однієї вимірної функції на іншу вимірну функцією теж є вимірною функцією за винятком випадку, коли функція, що є дільником, в якійсь із точок  $x \in E$  дорівнює нулю;

3) решітчаста функція  $f[k \Delta x] = f(x)_{x=k\Delta x}$ , що породжується неперервною функцією  $f(x)$  шляхом урахування її значень лише на дискретній множині  $E = \{k\Delta x\}$  значень

аргументу  $x$ , де  $\Delta x$  – інтервал дискретизації, а  $k = 0, 1, 2, \dots, n$ , належить до класу вимірних функцій, оскільки вона визначена на нульвимірній множині  $E$ ;

4) якщо на вимірній множині  $E$  задані дві вимірні функції  $f(x), g(x)$ , то підмножина  $E_{>} \subset E$ , в точках  $x \in E_{>}$  якої  $f(x) > g(x)$ , теж є вимірною;

5) неперервна вимірна функція  $f(x)$ , що визначена на вимірній множині  $E$  свого аргументу  $x$ , має на цій множині  $C$ -властивість, якщо для будь-якого числа  $\varepsilon > 0$  існує така замкнута підмножина  $F \subset E$ , для якої справедливою є нерівність

$$m(E - F) < \varepsilon; \quad (3.37)$$

6) якщо функція  $f(x)$  визначена на вимірній множині  $E$  з мірою, меншою нескінченності, й в усіх точках множини  $E$ , за винятком, можливо, деякої нульвимірної підмножини, має значення менші нескінченності, то для вимірності цієї функції достатньо, щоб вона мала на цій множині  $C$ -властивість.

### 3.3 Інтеграл Рімана

Цей інтеграл є одним із основних понять математичного аналізу, який вивчається в курсі вищої математики за навчальним планом будь-якої спеціальності нематематичної спеціалізації, тому в навчальному посібнику з функціонального аналізу, який є надбудовою над математичним аналізом, ми його лише пригадаємо, аби потім зіставити з ним уведені далі уже за програмою функціонального аналізу інтегралі Стілтєса та Лебега.

Отож, нехай на числовій осі на відрізку  $[a, b]$  значень аргументу  $x$  задана неперервна і обмежена функція  $f(x)$ . Розіб'ємо цей відрізок точками  $x_0, x_1, x_2, \dots, x_n$ , причому так, щоб  $x_0 = a$ , а  $x_n = b$ , на  $n$  сегментів  $\Delta_i x, i = 1, 2, \dots, n$  таких, що

$$\Delta_i x = x_i - x_{i-1}, \quad i = 1, 2, \dots, n \quad (3.38)$$

Позначимо  $\xi_i$  внутрішню точку сегмента  $\Delta_i x$ , тобто нехай  $\xi_i \in \Delta_i x$ . Оскільки функція  $f(x)$  є неперервною і обмеженою, то на кожному сегменті  $\Delta_i x$  у якихось його внутрішніх точках чи на його границях вона набуватиме максимального  $M_i$  та мінімального  $m_i$  числового значення, тобто матимемо

$$m_i \leq f(\xi_i) \leq M_i, \quad i = 1, 2, \dots, n \quad (3.39)$$

Помножимо всі члени нерівності (3.39) на  $\Delta_i x$  і підсумуємо ці добутки, в результаті чого отримаємо

$$\sum_{i=1}^n m_i \Delta_i x \leq \sum_{i=1}^n f(\xi_i) \Delta_i x \leq \sum_{i=1}^n M_i \Delta_i x \quad (3.40)$$

Суми

$$\Sigma_{\max} = \sum_{i=1}^n M_i \Delta_i x, \quad (3.41)$$



$$\Sigma_{\min} = \sum_{i=1}^n m_i \Delta_i x \quad (3.42)$$

називають , відповідно, верхньою та нижньою сумами Дарбу. Цілком очевидно, що при зменшенні  $\Delta_i x$  до нуля верхній індекс  $n$  в сумах Дарбу наблизатиметься до нескінченності, а самі ці суми наблизатимуться зверху і знизу до спільної границі у вигляді числа

$$J =_{\Delta_i x} \lim_0 \sum_{i=1}^{n \rightarrow \infty} f(\xi_i) \Delta_i x, \quad (3.43)$$

яке називають інтегралом Рімана і позначають

$$J = \int_a^b f(x) dx. \quad (3.44)$$

Оскільки операція знаходження суми є лінійною, для якої справедливо, що

$$\sum_{i=1}^{\infty} f(\xi_i) \Delta_i x = \sum_{i=1}^n f(\xi_i) \Delta_i x + \sum_{i=n+1}^{\infty} f(\xi_i) \Delta_i x, \quad (3.45)$$

то і операція інтегрування теж є лінійною, для якої справедливим є те, що

$$\int_a^b f(x) dx = \int_a^{a_*} f(x) dx + \int_{a_*}^b f(x) dx, \quad (3.46)$$

якщо  $a_* \in [a, b]$ .

Виходячи з того, що кожна складова  $M_i \Delta_i x$ ,  $m_i \Delta_i x$  в сумах Дарбу (3.41), (3.42) задає площу прямокутника з основою  $\Delta_i x$  та висотою, відповідно,  $M_i$  та  $m_i$ , геометрично інтерпретувати інтеграл Рімана, який є визначеним інтегралом, що належить до класу функціоналів, можна як вираз, за допомогою якого визначається площа плоскої фігури, обмеженої знизу відрізком  $[a, b]$  числової осі, а зверху – графіком функції  $f(x)$  в межах від  $f(a)$  до  $f(b)$ .

Інші властивості визначених інтегралів детально розглядаються в курсі математичного аналізу, який вивчається студентами усіх спеціальностей, тому в даному навчальному посібнику ми на них зупиняємся не будемо, вважаючи їх такими, які уже відомі студентам, що почали вивчення курсу функціонального аналізу.

### 3.4 Інтеграл Стілтєса

Із попереднього підрозділу легко бачити, що інтеграл Рімана (3.44), який є одним із основних понять математичного аналізу, ми побудували, не використовуючи теорію міри, а скориставшись верхньою (3.41) і нижньою (3.42) сумами Дарбу, до яких застосували граничний перехід.

Виявляється, що у такий же спосіб можна побудувати ще один клас інтегралів, які ввійшли в математику під іменем голландського математика Стілтєса, котрий створив їх теорію у відповідь на запит теорії ймовірностей.

Основною відмінністю інтеграла Стільтєса від інтеграла Рімана є те, що в інтегралі Рімана інтегрування на відрізку  $[a, b]$  функції  $f(x)$  здійснюється з використанням приростів  $dx$  її аргументу на цьому ж відрізку числової осі, а в інтегралі Стільтєса інтегрування на відрізку  $[a, b]$  функції  $f(x)$  здійснюється з використанням приростів  $dg(x)$  іншої функції  $g(x)$ , заданої на цьому ж відрізку числової осі, а сама функція  $f(x)$ , що інтегрується, називається інтегрованою за функцією  $g(x)$  на відрізку  $[a, b]$  числової осі.

Символічно інтеграл Стільтєса записується так:

$$S = \int_a^b f(x) dg(x). \quad (3.47)$$

Алгоритм побудови інтеграла (3.47), запропонований Стільтєсом, відрізняється від алгоритму побудови інтеграла Рімана лише в деталях, оскільки спочатку, як і в алгоритмі Рімана, пропонується на числовій осі на відрізку  $[a, b]$  значень аргументу  $x$  задати неперервну і обмежену функцію  $f(x)$  та розбити цей відрізок точками  $x_0, x_1, x_2, \dots, x_n$  так, щоб  $x_0 = a$ , а  $x_n = b$ , на  $n$  таких сегментів  $\Delta_i x$ ,  $i = 1, 2, \dots, n$ , що

$$\Delta_i x = x_i - x_{i-1}, \quad i = 1, 2, \dots, n. \quad (3.48)$$

Якщо позначити символом  $\xi_i$  внутрішню точку сегмента  $\Delta_i x$ , то, оскільки функція  $f(x)$  є неперервною і обмеженою, на кожному сегменті  $\Delta_i x$  у якихось його внутрішніх точках чи на його границях вона набуватиме максимального  $M_i$  та мінімального  $m_i$  числових значень, тобто матимемо

$$m_i \leq f(\xi_i) \leq M_i, \quad i = 1, 2, \dots, n. \quad (3.49)$$

А далі розглянемо задану на цьому ж відрізку  $[a, b]$  обмежену функцію  $g(x)$ , на яку не накладаємо умову неперервності й для якої на кожному сегменті, заданому виразом (3.48), знайдемо прирости

$$\Delta_i g = g(x_i) - g(x_{i-1}), \quad i = 1, 2, \dots, n. \quad (3.50)$$

Помножимо усі члени нерівності (3.49) на  $\Delta_i g$  і підсумуємо ці добутки, в результаті чого отримаємо

$$\sum_{i=1}^n m_i \Delta_i g \leq \sum_{i=1}^n f(\xi_i) \Delta_i g \leq \sum_{i=1}^n M_i \Delta_i g. \quad (3.51)$$

Суми

$$\Sigma_{\max}^s = \sum_{i=1}^n M_i \Delta_i g = \sum_{i=1}^n M_i (g(x_i) - g(x_{i-1})), \quad (3.52)$$

$$\Sigma_{msn}^s = \sum_{i=1}^n m_i \Delta_i g = \sum_{i=1}^n m_i (g(x_i) - g(x_{i-1})) \quad (3.53)$$

називають, відповідно, верхньою та нижньою сумами Дарбу-Стільтєса. Цілком очевидно, що при зменшенні  $\Delta, x$  до нуля верхній індекс  $n$  в сумах Дарбу-Стільтєса наблизатиметься до нескінченності, а самі ці суми, в силу обмеженості приростів (3.50), наблизатимуться зверху і знизу до спільної границі у вигляді числа

$$S =_{\Delta, x} \lim_0 \sum_{i=1}^{n \rightarrow \infty} f(\xi_i) (g(x_i) - g(x_{i-1})), \quad (3.54)$$

яке називають інтегралом Стільтєса і записують у вигляді (3.47).

Якщо функція  $g(x)$  на відрізку  $[a, b]$  є не лише обмеженою, але і неперервною, то, як відомо із курсу математичного аналізу, її диференціал  $dg(x)$  можна записати у вигляді

$$dg(x) = \frac{dg}{dx} dx = g'(x) dx. \quad (3.55)$$

Підставляючи вираз (3.55) у вираз (3.47), отримаємо

$$S = \int_a^b f(x) g'(x) dx = \int_a^b F(f, g, g', x) dx, \quad (3.56)$$

де

$$F(f, g, g', x) = f(x) g'(x). \quad (3.57)$$

Звертаємо увагу на те, що інтеграл Стільтєса у вигляді (3.56) – це вже інтеграл Рімана від функції (3.57), за допомогою якого обчислюється площа під графіком функції (3.57) на відрізку  $[a, b]$ .

А якщо функція  $g(x)$  на відрізку  $[a, b]$  є лінійною, тобто

$$g(x) = x, \quad (3.58)$$

то, оскільки у цьому випадку

$$g'(x) = 1, \quad (3.59)$$

інтеграл Стільтєса (3.47) стає інтегралом Рімана (3.44), в чому легко пересвідчитись, підставивши вираз (3.59) у вираз (3.56).

Із самого визначення інтеграла Стільтєса витікає, що:

$$1) \quad \int_a^b dg(x) = g(b) - g(a), \quad (3.60)$$

$$2) \quad \int_a^b kf(x)dg(x) = k \int_a^b f(x)dg(x), \quad \forall k = const, \quad (3.61)$$

$$3) \quad \int_a^b f(x)dg(x) = \int_a^{a_*} f(x)dg(x) + \int_{a_*}^b f(x)dg(x), \quad \forall a_* \in [a, b], \quad (3.62)$$

$$4) \quad \int_a^b (f_1(x) \pm f_2(x))dg(x) = \int_a^b f_1(x)dg(x) \pm \int_a^b f_2(x)dg(x), \quad (3.63)$$

$$5) \quad \int_a^b f(x)d(g_1(x) \pm g_2(x)) = \int_a^b f(x)dg_1(x) \pm \int_a^b f(x)dg_2(x). \quad (3.64)$$

Завершуючи викладення цього матеріалу, зауважимо, що виразам (3.38) та (3.39) ми, повторно використовуючи їх, задали нові номери, відповідно, (3.48) та (3.49), виключно заради зручності посилань.

### 3.5 Інтеграл Лебега

Коли математики побачили, що існують функції, які не інтегруються за Ріманом, вони почали пошук такого узагальнення поняття визначеного інтеграла, за допомогою якого ці функції теж можна було б інтегрувати. І таке узагальнення вдалося здійснити Лебегу, який запропонував приріст координати, за якою здійснюється інтегрування функції  $y = f(x)$ , заданої на відрізку  $[a, b]$ , визначати не по осі аргументу  $x$ , а по функціональній осі  $y$ , адже у цьому випадку, навіть коли координата  $x$  задається на нульвимірній множині  $E$  зкінченної чи зліченної кількості точок на осі  $x$ , координата  $y$  буде елементом множини дійсних чисел на відрізку  $[m, M]$  осі  $y$ , мірою якого є його довжина, і лівою границею якого є дійсне число  $m$ , що є мінімальним значенням цієї функції на відрізку  $[a, b]$ , а правою границею є дійсне число  $M$ , що є максимальним значенням цієї функції на цьому ж відрізку  $[a, b]$ .

Конструював свій інтеграл Лебег, висунувши умову, щоб вимірна і обмежена нижнім  $m$  та верхнім  $M$  значеннями функція  $y = f(x)$ , що задана на відрізку  $[a, b]$  осі  $x$ , була визначеною на множині  $E$  з мірою

$$mE(m \leq y < M). \quad (3.65)$$

Він запропонував розбити відрізок  $[m, M]$  осі  $y$  точками

$$m = y_0, y_1, y_2, \dots, y_{i-1}, y_i, \dots, y_n = M \quad (3.66)$$

на  $n$  сегментів  $[y_{i-1}, y_i]$ ,  $i = 1, 2, \dots, n$ , позначивши максимальний із них

$$\alpha = \max[y_{i-1}, y_i], \quad (3.67)$$

і скласти суму

$$\Sigma = \sum_{i=1}^n y_{i-1} mE_i(y_{i-1} \leq y < y_i), \quad (3.68)$$

використавши при цьому те, що, згідно з властивістю (3.35) міри

$$mE(m \leq y < M) = \sum_{i=1}^n mE_i(y_{i-1} \leq y < y_i). \quad (3.69)$$

**Границю суми (3.68) при наближенні навіть максимального із сегментів, визначеного виразом (3.67), до нуля**

$$L = {}_\alpha \underline{\lim}_0 \Sigma = {}_\alpha \underline{\lim}_0 \sum_{i=1}^n y_{i-1} mE_i(y_{i-1} \leq y < y_i), \quad (3.70)$$

**яка має всі властивості інтеграла, Лебег увів в математику як нове трактування визначеного інтеграла, який інші математики назвали в його честь інтегралом Лебега, а суму (3.68) назвали інтегральною сумою Лебега на множині  $E$  з мірою, визначеною виразом (3.65).**

Якщо, скориставшись властивістю монотонності міри, розписати  $mE_i(y_{i-1} \leq y < y_i)$  для  $i$ -го сегмента так

$$mE_i(y_{i-1} \leq y < y_i) = mE_i(y_i > y) + mE_i(y \geq y_{i-1}) = mE_i(y_i > y) - mE_i(y_{i-1} \geq y), \quad (3.71)$$

і підставити вираз (3.71) у вираз (3.70), то отримаємо інтеграл Лебега у вигляді

$$L = {}_\alpha \underline{\lim}_0 \sum_{i=1}^n y_{i-1} mE_i(y_{i-1} \leq y < y_i) = {}_\alpha \underline{\lim}_0 \sum_{i=1}^n y_{i-1} \{mE_i(y_i > y) - mE_i(y_{i-1} \geq y)\}. \quad (3.72)$$

Оскільки міра  $mE_i$  є монотонною функцією, позначимо її символом  $g$ -координати  $u$ , тоді можна записати, що

$$\begin{cases} mE_i(y_i > y) = g(y_i), \\ mE_i(y_{i-1} \geq y) = g(y_{i-1}), \\ \Delta_i g = g(y_i) - g(y_{i-1}) \end{cases} \quad (3.73)$$

Підставляючи вирази (3.73) у (3.72), отримаємо

$$L = {}_\alpha \underline{\lim}_0 \sum_{i=1}^n y_{i-1} \{g(y_i) - g(y_{i-1})\} = {}_\alpha \underline{\lim}_0 \sum_{i=1}^n y_{i-1} \Delta_i g. \quad (3.74)$$

Аналізуючи праву частину виразу (3.74), бачимо, що це є інтеграл Стілтєса від функції  $y$  за функцією  $g(y)$  на відрізку  $[m, M]$ , який є мірою функції  $y$ , а тому вираз (3.74) можна записати і так:

$$L = \int_m^M y dg(y). \quad (3.75)$$

Зведенням інтеграла Лебега (3.70) до інтеграла Стілтєса (3.75) ми фактично довели, що границя інтегральної суми Лебега (3.68) існує, тож вираз (3.70) дійсно задовольняє вимоги, які має задовольняти визначений інтеграл.

Звертаємо увагу на те, що, якщо функція міри за координатою  $y$  на відрізку  $[m, M]$  в проєкції на відрізок  $[a, b]$  осі  $x$  є не лише монотонною, а й лінійною, тобто, якщо

$$g(y) = x, \quad (3.76)$$

то вираз (3.75) перетворюється на вираз

$$L = \int_m^M y dg(y) = \int_a^b f(x) dx. \quad (3.77)$$

Тобто, у цьому випадку інтегрування за Лебегом на відрізку  $[m, M]$  осі  $y$  і за Ріманом на відрізку  $[a, b]$  осі  $x$  дають один і той же результат.

А тому в подальшому при аналізі функціоналів, які, як показано у вступі, відіграють роль критеріїв оптимізації в практичних задачах ІТ-спеціальностей, і підінтегральні функції в яких задаються на відрізках дійсних чисел, що мають потужність континууму і міру, яка задається довжиною цих відрізків, ми будемо використовувати лише визначені інтеграли Рімана.

### 3.6 Завдання для самоперевірки

1. Дайте означення поняття «міра».
2. Як визначається міра Жордана і які її властивості?
3. Чому дорівнює міра Жордана множини раціональних та ірраціональних чисел?
4. Як визначається міра Лебега та які її властивості?
5. Що собою являє нульвимірна множина?
6. Дайте означення вимірної функції.
7. Запишіть вирази для нижньої і верхньої сум Дарбу.
8. Дайте означення інтеграла Рімана.
9. Запишіть вирази для нижньої і верхньої сум Дарбу-Стілтєса.
10. Дайте означення інтеграла Стілтєса.
11. За якої умови інтеграл Рімана можна отримати з інтеграла Стілтєса?
12. Дайте означення інтеграла Лебега.
13. Який зв'язок існує між інтегралами Лебега і Стілтєса?
14. За яких умов інтеграл Рімана і Лебега дають однаковий результат?

## Розділ 4

# ФУНКЦІОНАЛИ, ЇХ ВАРІАЦІЇ ТА МЕТОДИ ПОШУКУ БЕЗУМОВНИХ ЕКСТРЕМУМІВ

### 4.1 Функціонали, що використовуються в прикладних задачах системного аналізу

Визначення функціонала як закону, за яким кожній функції з деякої її множини ставиться у відповідність число з деякої множини чисел, та приклади функціоналів у вигляді визначених інтегралів, що використовуються в прикладних задачах як критерії оптимізації процесів в складних системах, нами наведені у вступній частині навчального посібника, а тому викладення матеріалу у цьому підрозділі ми почнемо з конкретизації саме тих типів функціоналів, які будемо використовувати в процесі викладення основ функціонального аналізу, адаптованого до прикладних задач в ІТ-галузі.

Нагадаємо, що завершальним етапом системного аналізу складного об'єкта є його оптимізація, яка полягає у пошуку таких параметрів цього об'єкта, за яких його вихідна координата чи його просторова структура характеризуються найкращим чином за показниками, вибраними як міра цієї «найкращості». Саме ці показники, внутрішніми складовими яких є функції, що характеризують просторові структури складних об'єктів чи процесів у них, а зовнішніми складовими є числа, якими характеризується «найкращість» цих просторових структур чи процесів, як правило, є функціоналами, серед яких в загальній методології функціонального аналізу в його застосунку у формі варіаційного числення найпоширенішими є:

1) функціонал

$$J_y^F = \int_a^b F(x, y, y') dx, \quad (4.1)$$

де  $F(x, y, y')$  – математичний вираз, який є конструкцією з незалежної змінної  $x$ , її функції  $y(x)$  та першої похідної  $y'(x)$  від цієї функції; при цьому відрізок  $[a, b]$  є областю задавання функції  $y(x)$ , тобто,  $x \in [a, b]$ ;

2) функціонал

$$J = \int_a^b F(x, y, y', y'', \dots, y^{(n)}) dx, \quad (4.2)$$

який пов'язує між собою не лише функцію  $y(x)$  та її похідну  $y'(x)$ , як у випадку (4.1), але і старші похідні  $y''(x), \dots, y^{(n)}(x)$ ;

3) функціонал

$$J = \int_a^b F(x, y, z, y', z') dx, \quad (4.3)$$

який пов'язує між собою функції

$$y(x), z(x), \quad (4.4)$$

що задають поверхню у звичному тривимірному просторі  $\{x, y, z\}$ , та їх похідні

$$y'(x), z'(x) \quad (4.5)$$

у цьому ж просторі;

4) функціонал

$$J = \int_a^b F(x, y_1, y_2, \dots, y_n; y'_1, y'_2, \dots, y'_n) dx, \quad (4.6)$$

який пов'язує множину функцій

$$\{y_1(x), y_2(x), \dots, y_n(x)\}, \quad (4.7)$$

що задає поверхню в  $n$ -вимірному просторі, та множину

$$\{y'_1(x), y'_2(x), \dots, y'_n(x)\} \quad (4.8)$$

перших похідних цих функцій у цьому ж просторі.

Дослідженням умов, за яких ці функціонали набувають екстремальних значень, виконується в рамках розділу функціонального аналізу, який носить назву варіаційного числення, а функції, на яких ці функціонали набувають екстремальних значень, математики домовились називати екстремалами.

Але, перш ніж розпочати дослідження умов, за яких функціонали набувають екстремальних значень, ми нагадаємо за яких умов, визначених у вже вивченому нами курсі математичного аналізу, екстремуму набувають функції, оскільки це дасть змогу, завдяки певним аналогіям між цими умовами, легше зрозуміти матеріал, присвячений умовам набуття екстремуму функціоналами.

## 4.2 Екстремум функції

Розглянемо дві функції  $f_1(x)$ ,  $f_2(x)$ , задані графіками в області  $x \in (-\infty, \infty)$  (рис. 4.1).

Нагадаємо, що для побудови графіка першої похідної від функції, заданої графічно, необхідно в кожній точці графіка функції визначити тангенс кута нахилу до осі  $x$  дотичної до цього графіка, проведеної в даній точці. Аналогічно, з використанням графіка першої похідної, будується графік другої похідної цієї функції.

З цих графіків видно, що, як у випадку мінімуму функції  $f_1(x)$  в точці  $x_{01}$ , так і у випадку максимуму функції  $f_2(x)$  в точці  $x_{02}$ , перші похідні  $f'_1(x)$ ,  $f'_2(x)$  в цих точках дорівнюють нулю. Тож цілком зрозумілим є той факт, що значення  $x_{01}$  і  $x_{02}$  необхідно знаходити шляхом розв'язання рівнянь



$$\begin{cases} f_1'(x) = 0, \\ f_2'(x) = 0. \end{cases} \quad (4.9)$$

І як розпізнати, де матимемо максимум, а де – мінімум, теж легко бачити з графіків, наведених на рис. 4.1.

Зрозуміло, що якщо

$$f_1''(x_{01}) > 0, \quad (4.10)$$

то в точці  $x_{01}$  має місце мінімум функції  $f_1(x)$ , а якщо

$$f_2''(x_{02}) < 0, \quad (4.11)$$

то в точці  $x_{02}$  має місце максимум функції  $f_2(x)$ .

Рівняння (4.9) задають необхідні умови існування екстремуму функції однієї змінної, а нерівності (4.10), (4.11) – достатні.

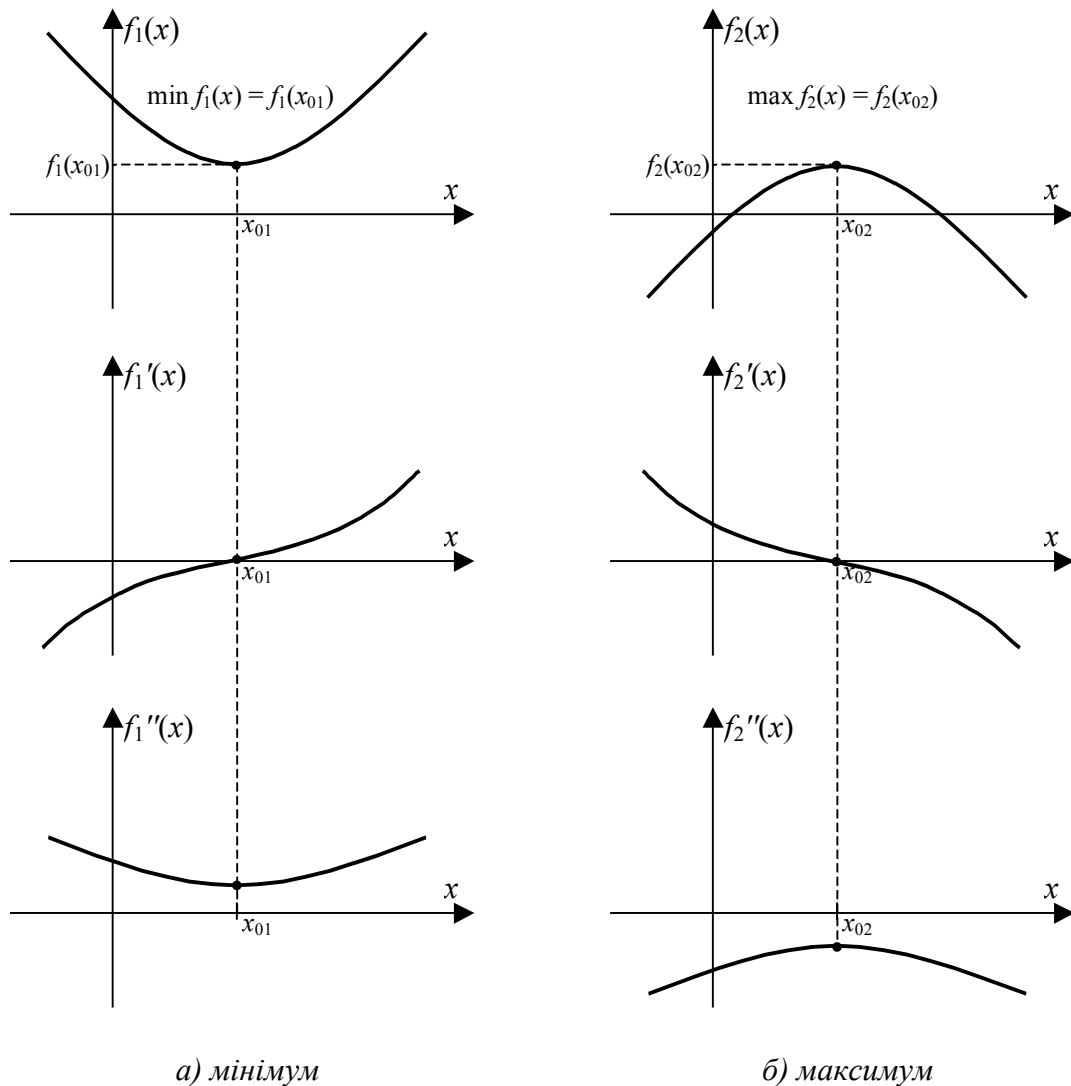


Рисунок 4.1 – Графічна інтерпретація умов існування екстремуму функції однієї незалежної змінної

Виникає запитання: «А чи завжди при виконанні умов (4.9) має місце екстремум функції однієї змінної?».

Виявляється, що не завжди. Це легко бачити з графіків функції  $f_3(x)$  та її першої похідної  $f_3'(x)$ , наведених на рис. 4.2. З цих графіків видно, що в точці  $x_{03}$ , незважаючи на виконання необхідних умов екстремуму, тобто, незважаючи на те, що

$$f_3'(x_{03}) = 0, \quad (4.12)$$

маємо і не максимум, і не мінімум функції, а її перегин.

Отже, умовою того, що в точці  $x_{03}$  маємо перегин функції, як видно з графіка другої похідної  $f_3''(x)$ , наведеного теж на рис. 4.2, є

$$f_3''(x_{03}) = 0. \quad (4.13)$$

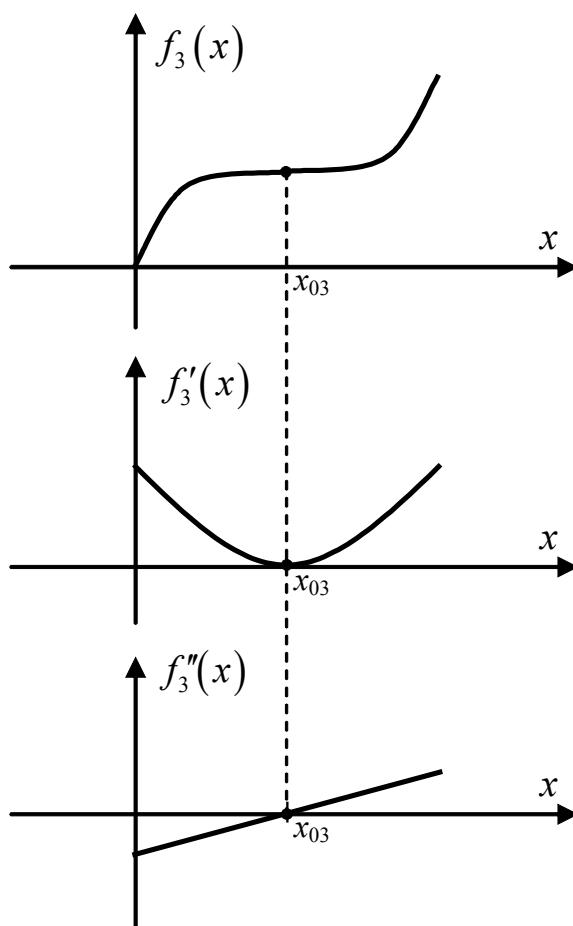


Рисунок 4.2 – Графічна інтерпретація умов відсутності екстремуму функції однієї незалежної змінної в точці, для якої виконуються необхідні умови його існування

А далі розглянемо ще один графік, наведений на рис. 4.3.

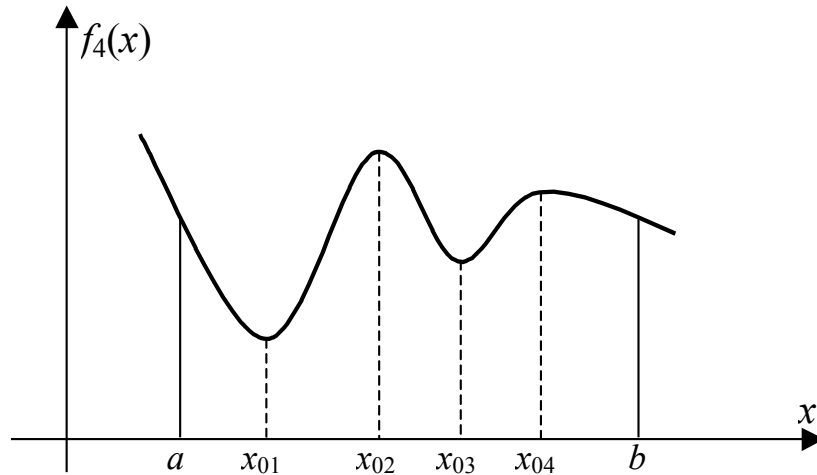


Рисунок 4.3 – Графік функції однієї змінної, яка має декілька мінімумів і максимумів

З цього графіка видно, що в області значень  $x \in [a, b]$  функція має дві точки мінімуму ( $f_4(x_{01}), f_4(x_{03})$ ) і дві точки максимуму ( $f_4(x_{02}), f_4(x_{04})$ ). При цьому значення функції  $f_4(x)$  в точці  $x_{01}$  є меншим за значення цієї ж функції в точці  $x_{03}$ , а значення функції  $f_4(x)$  в точці  $x_{02}$  є більшим за її значення в точці  $x_{04}$ .

В зв'язку з такою можливістю максимуми та мінімуми функції поділяють на глобальні та локальні. Зрозуміло, що в точці глобального мінімуму функція має найменше значення в області її задавання, а в точці глобального максимуму функція має найбільше значення в цій області.

Тож, досліджуючи функцію на екстремум, слід не забувати дослідити її в усіх кореневих точках, отримуваних з рівняння необхідних умов екстремуму, інакше є небезпека зупинитись на локальному екстремумі і пропустити глобальний.

А тепер розглянемо функцію  $f_{12}(x_1, x_2)$  двох незалежних змінних  $x_1$  та  $x_2$ , де  $x_1 \in [a, b]$ ,  $x_2 \in [c, d]$ .

Зрозуміло, що такою функцією можна описати поверхню гори, яка піднялася над площиною землі, обмеженою прямокутником зі сторонами « $ab$ » та « $cd$ » (рис. 4.4).

На цьому ж рисунку показані графіки функцій однієї незалежної змінної  $f_1(x_1, e)$  та  $f_2(g, x_2)$ , які є лініями перерізу поверхні гори з площинами, паралельними координатним площинам, за умови, що ці площини проходять через точку з координатами  $(g, e)$  на площині землі. Як бачимо, в функціях  $f_1(x_1, e)$ ,  $f_2(g, x_2)$  одна зі змінних перетворилась на параметр, який після того, як ми його задали чисельно, перетворюється на коефіцієнт. Після цього функція двох незалежних змінних стає функцією лише однієї незалежної змінної.

Задамося питанням: «Як знайти найвищу точку гори?».

З математичної точки зору це означає, що потрібно знайти максимум функції  $f_{12}(x_1, x_2)$ .

Зрозуміло, що для того, щоб знайти цей максимум, потрібно спочатку знайти його координати на площині  $x_1x_2$  (на рис. 4.4 – це точка  $(g, e)$ ).

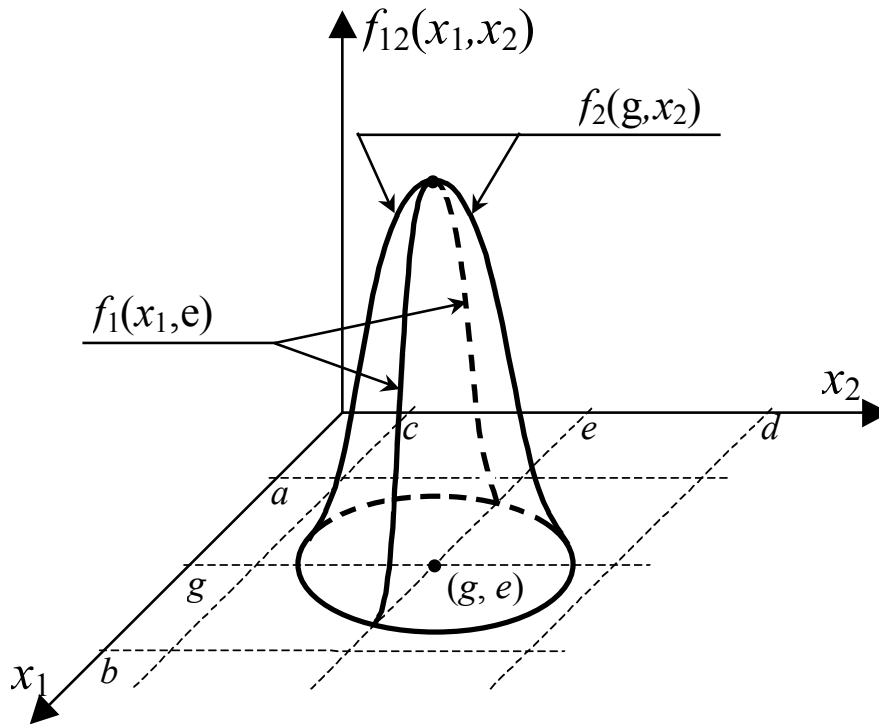


Рисунок 4.4 – Графічна інтерпретація функції  $f_{12}(x_1, x_2)$   
двох незалежних змінних  $x_1$  і  $x_2$ ,  $x_1 \in [a, b]$ ,  $x_2 \in [c, d]$

Ми вже знаємо, як знайти координату точки максимуму для функції однієї незалежної змінної: потрібно взяти похідну від цієї функції, прирівняти цю похідну до нуля і знайти корінь (чи корені) отриманого рівняння.

З викладеного вище сам собою напрашується спосіб пошуку координат точки максимуму функції  $f_{12}(x_1, x_2)$  двох незалежних змінних  $x_1, x_2$  – потрібно взяти перші частинні похідні від цієї функції за кожною з незалежних змінних, вважаючи другу незалежну змінну в цьому процесі параметром, прирівняти отримані перші частинні похідні нулю і сумісно розв'язати отриману систему двох рівнянь з двома невідомими. Цей розв'язок і дасть нам чисельне значення координат точки максимуму функції.

Формально цю умову можна записати так:

$$\begin{cases} \frac{\partial f_{12}(x_1, x_2)}{\partial x_1} = 0, \\ \frac{\partial f_{12}(x_1, x_2)}{\partial x_2} = 0. \end{cases} \quad (4.14)$$

⇓

$$\begin{cases} x_1 = g, \quad x_2 = e; \\ f_{12}(g, e) = \underset{\substack{[a, b] \\ [c, d]}}{\text{extremum}} f_{12}(x_1, x_2). \end{cases} \quad (4.15)$$

Але цілком зрозуміло, що функцією  $f_{12}(x_1, x_2)$  може описуватись не лише поверхня гори над площиною землі, але і поверхня провалля під площиною землі – саме тому у виразі (4.15) ми написали не  $\max$ , а  $\text{extr}$ .

Тож, із системи рівнянь (4.14) незрозуміло, що саме ми знайшли, розв'язавши їх, координати точки максимуму чи точки мінімуму.

Узагальнюючи викладки, які ми наводили щодо характеру екстремуму функції однієї незалежної змінної, на функцію двох незалежних змінних, можна стверджувати: для того, щоб розв'язок системи рівнянь (4.14) задавав координати точки максимуму функції  $f_{12}(x_1, x_2)$  достатньо сумісного виконання системи нерівностей відносно других частинних похідних цієї функції такого характеру

$$\begin{cases} \left. \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} \right|_{x_1=g, x_2=e} < 0, \\ \left. \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} \right|_{x_1=g, x_2=e} < 0, \end{cases} \quad (4.16)$$

а для того, щоб цей розв'язок задавав координати точки мінімуму функції  $f_{12}(x_1, x_2)$ , достатньо сумісного виконання системи нерівностей відносно других частинних похідних цієї функції такого характеру

$$\begin{cases} \left. \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} \right|_{x_1=g, x_2=e} > 0, \\ \left. \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} \right|_{x_1=g, x_2=e} > 0. \end{cases} \quad (4.17)$$

Якщо в умовах (4.16) або (4.17) одна з других частинних похідних дорівнює нулю, то точка  $x_1 = g$ ,  $x_2 = e$  лежить на прямій, по якій відбувається перегин поверхні, що описується функцією  $f_{12}(x_1, x_2)$ .

Якщо в умовах (4.16) або (4.17) обидві другі частинні похідні дорівнюють нулю, то точка  $x_1 = g$ ,  $x_2 = e$  лежить на плоскій ділянці поверхні, яка має місце на якійсь ділянці гори.

Якщо ж в умовах (4.16) або (4.17) одна друга частинна похідна більша за нуль, а інша – менша за нуль, то точка  $x_1 = g$ ,  $x_2 = e$  є сідловою – це нижня точка перевалу між двома горами або верхня точка перевалу між двома проваллями.

Завершуючи цей підрозділ, відзначимо, що якщо функція має більше двох незалежних змінних, то необхідні умови існування екстремуму відрізнятимуться від (4.14) лише більшою кількістю аналогічних рівнянь. Але достатні умови вже не можуть обмежуватись виконанням нерівностей (4.16), (4.17). Необхідно, щоб виконувались і ще деякі більш складні умови.

### 4.3 Суть варіаційного числення. Рівняння Ейлера

*Сукупність методів пошуку екстремумів функціоналів різних типів, із числа наведених нами у першому підрозділі цього розділу, становить суть варіаційного числення*, для розуміння основ якого розглянемо функціонал (4.1), який для зручності перепишемо під номером цього підрозділу

$$J_y^F = \int_a^b F(x, y, y') dx. \quad (4.18)$$

Але, перш ніж визначати умови, за яких функціонал (4.18) набуватиме екстремального значення, уточнимо, що ми розумітимемо під поняттями абсолютний та відносний екстремуми функціонала, та що ми будемо розуміти під поняттям слабкий відносний екстремум функціонала.

За аналогією з тим, як для функції ми визначали глобальний і локальний екстремуми, для функціонала визначається абсолютний і відносний екстремуми, перший з яких задає найбільше (чи найменше) значення функціонала на всій множині функцій, на якій цей функціонал задано, а другий задає найбільше (чи найменше) значення функціонала на підмножині близьких функцій, які є лише частиною від усієї множини функцій, на яких задано цей функціонал.

В свою чергу серед відносних екстремумів розрізняють сильний і слабкий.

Для уніфікації підходів домовились вважати, що на екстремалі  $f_e(x)$  досягається сильний відносний мінімум функціонала, якщо його значення на цій кривій у заданому діапазоні  $[a, b]$  значень аргументу  $x$  є меншим, ніж на усіх інших кривих  $f_i(x)$ ,  $i = 1, \dots, n$  даного класу функцій, відстань нульового порядку

$$\Delta_0 = \max |f_e(x) - f_i(x)|, \quad i = 1, \dots, n \quad (4.19)$$

до яких мала. Очевидно, що сильний відносний максимум функціонала буде досягатись на екстремалі з цього ж класу кривих, якщо його значення буде в заданому діапазоні значень аргументу найбільшим.

Якщо ж відносний мінімум (чи максимум) функціонала досягається на екстремалі, відстань першого порядку

$$\Delta_1 = \max |f'_e(x) - f'_i(x)|, \quad i = 1, \dots, n, \quad (4.20)$$

від якої до усіх інших кривих даного класу функцій є малою, то на цій екстремалі має місце слабкий відносний мінімум (чи максимум).

Зауважимо, що у виразі (4.20) символами  $f'_e(x)$ ,  $f'_i(x)$  позначені перші похідні від функцій  $f_e(x)$ ,  $f_i(x)$ .

Зрозуміло, що абсолютний екстремум є в той же час і відносним, а сильний відносний екстремум є в той же час і слабким.

А тому, якщо якась умова має виконуватись відносно слабого відносного екстремуму, то вона має бути справедливою і для сильного відносного екстремуму, і для абсолютного.

Розібравшись з наведеними вище поняттями, визначимо, які умови має задовольняти функція  $y = y(x)$ , щоб на ній мав місце слабкий відносний мінімум функціонала (4.18).

Для пошуку цих умов припустимо, що цей слабкий відносний мінімум функціонала (4.18) на функції  $y = y(x)$  досягається.

Це припущення дає нам право вважати, що значення функціонала (4.18) на будь-якій іншій функції, наприклад,  $y_\eta(x) = y(x) + \alpha \cdot \eta(x)$ , де  $\alpha$  – число, а  $\eta(x)$  – довільна гладка функція, для якої

$$\eta(a) = \eta(b) = 0, \quad (4.21)$$

не буде меншим, ніж на функції  $y = y(x)$ , тобто,

$$\Delta J_y^F = \int_a^b (x, y + \alpha\eta, y' + \alpha\eta') dx - \int_a^b F(x, y, y') dx = J_y^F(\alpha) - J_y^F \geq 0. \quad (4.22)$$

Оскільки визначений інтеграл після інтегрування і підстановки границь перетворюється на число, то значення приросту функціонала, який задається різницею (4.22) визначених інтегралів, залежатиме лише від значення параметра  $\alpha$ , тобто цей приріст стає неперервною функцією, в якій незалежною змінною виступає  $\alpha$

$$\Delta J_y^F = \Delta J_y^F(\alpha) = J_y^F(\alpha) - J_y^F. \quad (4.23)$$

Як відомо з курсу математичного аналізу, неперервну функцію в околі будь-якого значення її аргументу, зокрема нульового, можна розкласти в ряд Тейлора. Для функції (4.23) в околі точки  $\alpha = 0$  цей ряд матиме вигляд:

$$\Delta J_y^F(\alpha) = \alpha \cdot \frac{dJ_y^F(\alpha)}{d\alpha} + \frac{\alpha^2}{2!} \cdot \frac{d^2 J_y^F(\alpha)}{d\alpha^2} + \frac{\alpha^3}{3!} \cdot \frac{d^3 J_y^F(\alpha)}{d\alpha^3} + \dots. \quad (4.24)$$

Перший і другий члени ряду (4.24), тобто  $\alpha \cdot \frac{dJ_y^F(\alpha)}{d\alpha}$  і  $\frac{\alpha^2}{2} \cdot \frac{d^2 J_y^F(\alpha)}{d\alpha^2}$ , називають, відповідно, першою та другою **варіаціями функціонала**  $J_y^F$  і позначають  $\delta J_y^F$  та  $\delta^2 J_y^F$ .

Оскільки в другу варіацію  $\delta^2 J_y^F$  незалежна змінна  $\alpha$  входить в квадраті, то при значеннях  $\alpha$ , близьких до нуля, друга варіація стає набагато меншою за першу варіацію  $\delta J_y^F$ , в яку  $\alpha$  входить в першому степені. А це, в свою чергу, дає нам право вважати, що в околі точки  $\alpha = 0$

$$\Delta J_y^F(\alpha) \approx \delta J_y^F(\alpha). \quad (4.25)$$

З виразів (4.22), (4.23), (4.24) та (4.25) випливає, що

$$\alpha \cdot \frac{dJ_y^F(\alpha)}{d\alpha} \geq 0. \quad (4.26)$$

Зрозуміло, що для довільних значень  $\alpha$  (як більших, так і менших за нуль) вираз (4.26) в околі точки  $\alpha = 0$  виконується лише в одному випадку, коли

$$\frac{dJ_y^F(\alpha)}{d\alpha} = 0, \quad (4.27)$$

або

$$\frac{d}{d\alpha} \left( \int_a^b F(x, y + \alpha\eta, y' + \alpha\eta') dx \right) = 0. \quad (4.28)$$

Оскільки операції диференціювання та інтегрування є лінійними, то їх можна міняти місцями, тобто (4.28) можна переписати і так:

$$\int_a^b \frac{d}{d\alpha} F(x, y + \alpha\eta, y' + \alpha\eta') dx = 0. \quad (4.29)$$

Беручи похідну за  $\alpha$  в підінтегральному виразі (4.29), за правилом диференціювання складної функції будемо мати

$$\int_a^b \left( \frac{\partial F}{\partial y} \cdot \frac{d(y + \alpha\eta)}{d\alpha} + \frac{\partial F}{\partial y'} \cdot \frac{d(y' + \alpha\eta')}{d\alpha} \right) dx = 0, \quad (4.30)$$

або

$$\int_a^b \left( \frac{\partial F}{\partial y} \cdot \eta + \frac{\partial F}{\partial y'} \cdot \eta' \right) dx = 0. \quad (4.31)$$

Вираз (4.31) є незручним для аналізування, оскільки до його першої складової входить сама функція  $\eta(x)$ , а до другої – її похідна  $\eta'(x)$ . Спробуємо позбутись похідної  $\eta'(x)$ , взявши інтеграл від другої складової виразу частинами. Для спрощення записів будемо в подальшому частинні похідні позначати так:

$$\begin{cases} \frac{\partial F}{\partial x} = F_x, & \frac{\partial F}{\partial y} = F_y, & \frac{\partial F}{\partial y'} = F_{y'}, \\ \frac{\partial^2 F}{\partial y' \partial x} = F_{y'x}, & \frac{\partial^2 F}{\partial y' \partial y} = F_{y'y}, & \frac{\partial^2 F}{\partial y' \partial y'} = F_{y'y'} \end{cases} \quad (4.32)$$

Нехай в другій складовій частині інтеграла в рівнянні (4.31), тобто в інтегралі



$$\int_a^b \frac{\partial F}{\partial y'} \cdot \eta' dx = \int_a^b F_{y'} \cdot \eta' dx, \quad (4.33)$$

$$\begin{cases} u = F_{y'}, \\ dv = \eta' dx. \end{cases} \quad (4.34)$$

Тоді

$$\begin{cases} du = \frac{d}{dx} F_{y'} dx, \\ v = \int \frac{d\eta}{dx} dx = \int d\eta = \eta. \end{cases} \quad (4.35)$$

Виходячи з того, що

$$\int_a^b u dv = (uv) \Big|_a^b - \int_a^b v du, \quad (4.36)$$

інтеграл (4.33), з урахуванням (4.34) та (4.35), набуде вигляду

$$\int_a^b F_{y'} \cdot \eta' dx = (F_{y'} \cdot \eta) \Big|_a^b - \int_a^b \frac{d}{dx} F_{y'} \cdot \eta dx. \quad (4.37)$$

Беручи до уваги умову (4.21), з виразу (4.37), після підстановки границь у вираз  $(F_{y'} \cdot \eta) \Big|_a^b$ , будемо мати

$$\int_a^b F_{y'} \cdot \eta' dx = - \int_a^b \frac{d}{dx} F_{y'} \cdot \eta dx. \quad (4.38)$$

Враховуючи (4.38), рівняння (4.31) можна переписати так:

$$\int_a^b \left( F_{y'} \cdot \eta - \frac{d}{dx} F_{y'} \cdot \eta \right) dx = 0, \quad (4.39)$$

або

$$\int_a^b \left( F_{y'} - \frac{d}{dx} F_{y'} \right) \cdot \eta dx = 0. \quad (4.40)$$

Проаналізуємо отримане інтегральне рівняння (4.40). В ньому на гладку функцію  $\eta(x)$  накладена лише одна умова (4.21). В усьому іншому ця функція може бути довільною, зокрема і такою, як показано на рис. 4.5.

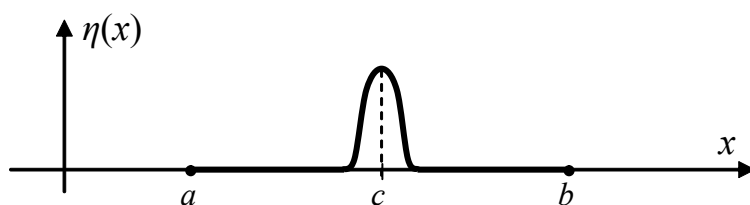


Рисунок 4.5 – Графік функції  $\eta(x)$

Якщо ми будемо «гладкий» імпульс, зображений на графіку функції  $\eta(x)$  на рис. 4.5 в околі точки  $x = c$ , переміщувати вліво або вправо по графіку аж до околів точок  $x = a$  або  $x = b$ , то, зрозуміло, що рівняння (4.40) може задовольнятися лише в одному випадку, коли вираз, який стоїть в дужках, буде дорівнювати нулю, тобто,

$$F_y - \frac{d}{dx} F_{y'} = 0. \quad (4.41)$$

*Це і є відоме рівняння Ейлера*, яке він вивів в 1744 році шляхом перетворення функціонала у функцію багатьох змінних з подальшою її мінімізацією і яке в класичному варіаційному численні може бути отримано з використанням леми Лагранжа.

На наш погляд, запропонований евристичний спосіб отримання рівняння Ейлера є більш простим для розуміння, хоча, звичайно, поступається класичним в строгості.

Резюмуючи все викладене вище, можна стверджувати, що для того, щоб функція  $y(x)$  доставляла слабкий відносний мінімум функціоналу (4.18), їй необхідно бути розв'язком рівняння Ейлера (4.41). **В цьому випадку функція  $y(x)$  називається екстремаллю функціонала.**

На використанні рівняння Ейлера в різних інтерпретаціях власне і базується варіаційне числення.

Розпишемо рівняння Ейлера, використавши формулу повної похідної від функції трьох змінних при диференціюванні складової  $\frac{d}{dx} F_{y'}$ . Отримаємо

$$F_y - \frac{\partial F_{y'}}{\partial x} \cdot \frac{dx}{dx} - \frac{\partial F_{y'}}{\partial y} \cdot \frac{dy}{dx} - \frac{\partial F_{y'}}{\partial y'} \cdot \frac{dy'}{dx} = 0, \quad (4.42)$$

або

$$F_y - F_{y'x} - F_{y'y} \cdot y' - F_{y'y'} \cdot y'' = 0. \quad (4.43)$$

З виразу (4.43) видно, що рівняння Ейлера є нелінійним диференціальним рівнянням другого порядку, для якого не існує єдиного способу розв'язання.

Важливим є випадок, коли функція  $F(\bullet)$  у функціоналі (4.18) явно не залежить від  $x$ , тобто

$$F(x, y, y') = F(y, y'). \quad (4.44)$$

В цьому випадку замість рівняння (4.43) матимемо рівняння

$$F_y - F_{y'y'} \cdot y' - F_{y'y'} \cdot y'' = 0, \quad (4.45)$$

яке множенням на  $y'$  легко перетворюється у рівняння

$$\frac{d}{dx}(F - y' \cdot F_{y'}) = 0. \quad (4.46)$$

В свою чергу, з рівняння (4.46) випливає, що

$$F - y' \cdot F_{y'} = C, \quad (4.47)$$

де  $C$  – константа.

**Отриманий у такий спосіб вираз (4.47) називають першим інтегралом рівняння Ейлера.** Він є диференціальним рівнянням першого порядку, яке явно не залежить від  $x$ , а тому може бути розв'язаним завжди.

Як буде показано далі, перший інтеграл рівняння Ейлера фігуруватиме і в багатьох інших цікавих співвідношеннях варіаційного числення.

В деяких задачах оптимізації рівняння Ейлера, яке є нелінійним диференціальним рівнянням другого порядку, зручно подавати у вигляді системи двох диференціальних рівнянь першого порядку, для отримання якої вводиться нова змінна

$$p = F_{y'}. \quad (4.48)$$

Підставляючи (4.48) у (4.41), отримаємо

$$F_y = \frac{dp}{dx}. \quad (4.49)$$

Введемо функцію  $H$ , яка задає перший інтеграл рівняння Ейлера, тобто,

$$H = F - y' \cdot F_{y'}. \quad (4.50)$$

Підставивши (4.48) у (4.50), отримаємо

$$H = F - y' \cdot p. \quad (4.51)$$

Диференціюючи функцію  $H$  за  $y$  і за  $p$ , отримаємо

$$\begin{cases} \frac{\partial H}{\partial y} = F_y, \\ \frac{\partial H}{\partial p} = -y', \end{cases} \quad (4.52)$$

Враховуючи те, що  $y' = \frac{dy}{dx}$ , та рівність (4.49), систему рівнянь (4.52) можна переписати так:

$$\begin{cases} \frac{\partial H}{\partial y} = \frac{dp}{dx}, \\ \frac{\partial H}{\partial p} = -\frac{dy}{dx}. \end{cases} \quad (4.53)$$

*Система рівнянь (4.53) є іншою формою подання рівняння Ейлера (4.41), яку математики назвали канонічною.*

В цій системі одні математики канонічними називають змінні  $H$  та  $p$ , а інші –  $y$  та  $p$ .

Завершимо цей підрозділ реплікою про те, що *задачу пошуку екстремалі  $y(x)$ , яка доставляє екстремум функціоналу (4.18) і є кривою, зацмленою на кінцях, в класичному варіаційному численні прийнято називати найпростішою.*

#### 4.4 Достатні умови існування екстремуму функціонала на екстремалі

*Наявність розв'язку рівняння Ейлера для екстремалі  $y(x)$  є умовою, необхідною для того, щоб на ній досягався мінімум чи максимум функціоналу (4.18).* Але, як і у випадку екстремуму функції, необхідні умови екстремуму функціонала обов'язково потрібно доповнити умовами достатніми, за допомогою яких розпізнаються як ті функції, на яких максимум чи мінімум функціоналу досягається, так і ті, на яких, незважаючи на виконання необхідних умов, функціонал екстремуму не досягає.

Для визначення достатніх умов екстремуму функціоналу (4.18) повернемося знову до розкладу в ряд Тейлора (4.24) приросту функціонала  $\Delta J_y^F(\alpha)$ .

Як уже встановлено в попередньому підрозділі, в околі точки  $\alpha = 0$  за рахунок виконання умови (4.27) перша варіація  $\delta J_y^F(\alpha)$  наближається до нуля, тому приріст функціонала визначається другою варіацією  $\delta^2 J_y^F(\alpha)$ , оскільки інші члени ряду Тейлора наближаються до нуля швидше за другу варіацію за рахунок більш високих, але близьких до нуля, степенів  $\alpha$ .

Зрозуміло, що у випадку мінімуму функціонала і приріст функціонала  $\Delta J_y^F(\alpha)$  в околі точки  $\alpha = 0$ , і його друга варіація  $\delta^2 J_y^F(\alpha)$  не будуть меншими нуля, оскільки в точці мінімуму значення цього функціонала є найменшим і будь-яке зміщення з цієї точки  $\alpha = 0$  або не приведе до зміни значення функціонала, або приведе до приросту його значення.

Тож, евристично, мінімум функціонала (4.18) досягається на екстремалі  $y(x)$ , для якої справедливим є вираз

$$\delta^2 J_y^F(\alpha) \geq 0, \quad (4.54)$$

де, нагадаємо,

$$\delta^2 J_y^F(\alpha) = \frac{\alpha^2}{2} \frac{d^2 J_y^F}{d\alpha^2}. \quad (4.55)$$

Розмірковуючи за аналогією, дійдемо висновку, що максимум функціонала (4.18) досягається на екстремалі  $y(x)$ , для якої справедливим є вираз

$$\delta^2 J_y^F(\alpha) \leq 0. \quad (4.56)$$

Оскільки у вирази (4.54), (4.56), згідно з виразом (4.55),  $\alpha$  входить в квадраті, то знаки цих нерівностей визначаються виключно знаками другої похідної від приросту функціонала, яку можна записати так:

$$\begin{aligned} \frac{d^2 J_y^F}{d\alpha^2} &= \frac{d^2}{d\alpha^2} \left( \int_a^b F(x, y + \alpha\eta, y' + \alpha\eta') dx \right) = \\ &= \int_a^b \frac{d^2}{d\alpha^2} F(x, y + \alpha\eta, y' + \alpha\eta') dx = \\ &= \int_a^b \left( F_{yy} \eta^2 + (F_{yy'} + F_{y'y}) \cdot \eta\eta' + F_{y'y'} (\eta')^2 \right) dx. \end{aligned} \quad (4.57)$$

Для того, щоб і в середньому члені в правій частині виразу (4.57) отримати квадрат функції  $\eta$ , що полегшить аналізування виразу (4.57) в цілому, візьмемо частинами інтеграл від цього середнього члена. Після підстановки значення отриманого інтегруванням частинами інтеграла в вираз (4.57), отримаємо

$$\frac{d^2 J_y^F(\alpha)}{d\alpha^2} = \int_a^b \left( \left( F_{yy} - \frac{d}{dx} F_{yy'} \right) \eta^2 + F_{y'y'} (\eta')^2 \right) dx. \quad (4.58)$$

Нагадаємо, що допоміжна функція  $\eta(x)$  є довільною, а тому й можна взяти її такою, як показано на рис. 4.6.

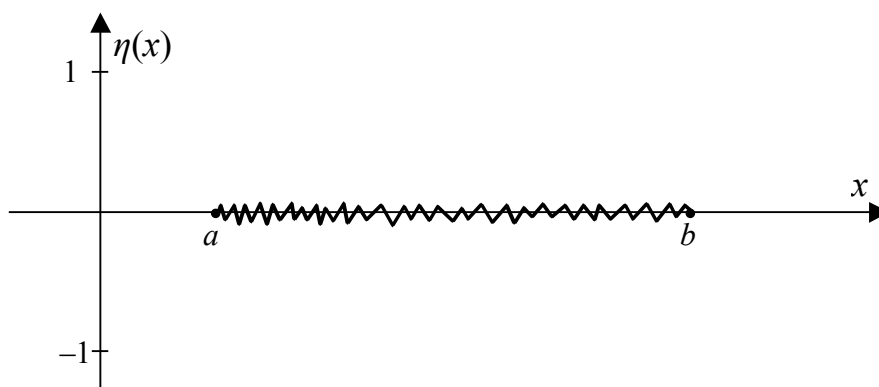


Рисунок 4.6 – Графік допоміжної функції  $\eta(x)$  для виразу (4.58)

При такому виборі  $\eta(x)$  матимемо чисельно близькі до нуля значення  $(\eta)^2$  і в сотні разів більші значення  $(\eta')^2$ . А тому знак виразу (4.58) буде повністю визначатись знаком коефіцієнта при  $(\eta')^2$ , тобто знаком при  $F_{y'y'}$ .

Виходячи з виразів (4.54), (4.56) і всього викладеного вище, можна стверджувати, що на екстремалі  $y(x)$  в межах  $x \in [a, b]$  досягається мінімум функціонала (4.18), якщо для всіх  $x \in [a, b]$  маємо

$$F_{y'y'} \geq 0, \quad (4.59)$$

і максимум, якщо –

$$F_{y'y'} \leq 0. \quad (4.60)$$

**Умови (4.59), (4.60) є достатніми умовами досягнення на екстремалі  $y(x)$  в межах  $x \in [a, b]$  екстремуму функціонала (4.18).**

**Ці умови часто називають умовами Лежандра**, за іменем математика, який їх вивів.

Для продовження аналізу достатніх умов досягнення на екстремалі  $y(x)$  в межах  $x \in [a, b]$  екстремуму функціонала (4.18) повернемося до розгорнутого запису рівняння Ейлера (4.43). Цей вираз можна переписати і так

$$y'' = \frac{F_y - F_{y'x} - F_{y'y} \cdot y'}{F_{y'y'}}. \quad (4.61)$$

З виразу (4.61) випливає, що функція  $y(x)$  для того, щоб претендувати на роль екстремалі, має в області  $x \in [a, b]$  мати другу похідну  $y''$ , а її перша похідна  $y'$  має задовольняти в деяких випадках ще й додаткові умови.

Із (4.61) зрозуміло, що при

$$F_{y'y'} > 0 \quad (4.62)$$

для мінімуму функціонала (4.18) та при

$$F_{y'y'} < 0 \quad (4.63)$$

для його максимуму, на першу похідну  $y'$  в області  $x \in [a, b]$  крім того, що вона має в цій області існувати, ніяких інших умов накладати не потрібно.

Але якщо в окремих точках області  $x \in [a, b]$

$$F_{y'y'} = 0, \quad (4.64)$$

то необхідно, щоб в цих же точках перша похідна  $y'$  чисельно збігалася зі значенням виразу  $\frac{F_y - F_{y'x}}{F_{y'y}}$ , тобто, щоб в цих точках виконувалась рівність

$$y' = \frac{F_y - F_{y'x}}{F_{y'y}}, \quad (4.65)$$

що впливає з необхідності мати, для забезпечення існування в області  $x \in [a, b]$  другої похідної  $y''$ , крім виконання умови (4.64), ще й виконання умови

$$F_y - F_{y'x} - F_{y'y} \cdot y' = 0. \quad (4.66)$$

Зрозуміло, що якщо рівність (4.64) виконується для всіх точок області  $x \in [a, b]$ , то для всіх точок цієї області має виконуватись і рівність (4.66).

Слід зазначити, що якщо рівність (4.64) виконується для всіх точок області  $x \in [a, b]$ , то це означає, що підінтегральна функція  $F(x, y, y')$  у функціоналі (4.18) залежить від першої похідної лінійно, тобто що цей функціонал має вигляд

$$J_y^F = \int_a^b (M(x, y) + N(x, y) \cdot y') dx. \quad (4.67)$$

**Функціонали такого виду називають виродженими.**

Для них рівняння Ейлера не є диференціальним, тому що

$$\begin{aligned}
F_y - \frac{d}{dx} F_{y'} &= \frac{\partial M}{\partial y} + y' \frac{\partial N}{\partial y} - \frac{d}{dx} (N(x, y)) = \\
&= \frac{\partial M}{\partial y} + \frac{\partial N}{\partial y} \cdot y' - \frac{\partial N}{\partial y} \cdot \frac{dy}{dx} - \frac{\partial N}{\partial x} \cdot \frac{dx}{dx} = \frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} = 0.
\end{aligned}
\tag{4.68}$$

З виразу (4.68) випливає, що рівняння Ейлера для вироджених функціоналів не містить похідних від екстремалі  $y(x)$ .

Розкриваючи в функціоналі (4.67) зовнішні дужки, його можна записати у вигляді

$$J_y^F = \int_a^b M dx + N dy, \tag{4.69}$$

тобто, він перетворюється на інтеграл від повного диференціала.

А, як відомо з курсу математичного аналізу, значення такого інтеграла не залежить від шляху інтегрування, тобто його значення на всіх функціях  $y(x)$ , на яких задається функціонал (4.69), є однаковими, що в нашому випадку, в свою чергу, означає, що екстремум такого функціонала досягається на будь-якій функції з області його задавання.

В разі, якщо функціонал (4.18) не є виродженим, тобто його не можна подати у вигляді (4.67), а умова (4.64) виконується, то це означає, як буде показано далі, що екстремум функціонала не може бути досягнутим в класі гладких функцій. В цьому випадку його треба шукати в класі кусково-гладких функцій, графіки яких мають зломи.

**Примітка.** Математики віднайшли приклади задач оптимізації, для яких недостатньо виконання умов Лежандра для підтвердження того, що мінімум чи максимум функціонала (4.18) досягається саме на отриманій з рівняння Ейлера екстремалі  $y(x)$ .

Для узагальнення і на такі задачі методики пошуку достатніх умов досягнення мінімуму чи максимуму на екстремалі  $y(x)$  математиками було введено поняття поля екстремалей і побудовано диференціальне рівняння Якобі, за розв'язком якого можна робити висновок про стан виконання достатніх умов існування екстремуму функціонала (4.18).

З виведенням рівняння Якобі та його аналізом можна ознайомитись в будь-якому достатньо повному підручнику з варіаційного числення.

Ми ж у нашому навчальному посібнику вирішили не приділяти уваги умовам Якобі тому, що в практичних задачах оптимізації динамічних об'єктів, характерних для IT-спеціальностей, умов Лежандра цілком досить для перевірки того, виконуються чи не виконуються достатні умови існування екстремуму на екстремалі  $y(x)$ , отриманій з рівняння Ейлера.

#### 4.5 Узагальнення найпростішої задачі варіаційного числення для пошуку екстремалей з рухомими кінцями

Всі викладки, що наведені в попередніх підрозділах, стосувалися задач пошуку екстремуму функціоналів в класі функцій, графіки яких зацмелені на кінцях.

Але існує досить багато задач оптимізації, в яких екстремалі необхідно шукати в класі функцій, кінці графіків яких є рухомими.



Математично така задача формулюється так: серед кривих, кінці яких можуть рухатись по двох інших кривих – позначимо їх  $\varphi(x)$  і  $\psi(x)$  (рис. 4.7), знайти криву  $y(x)$ , що доставляє мінімум функціоналу (4.18).

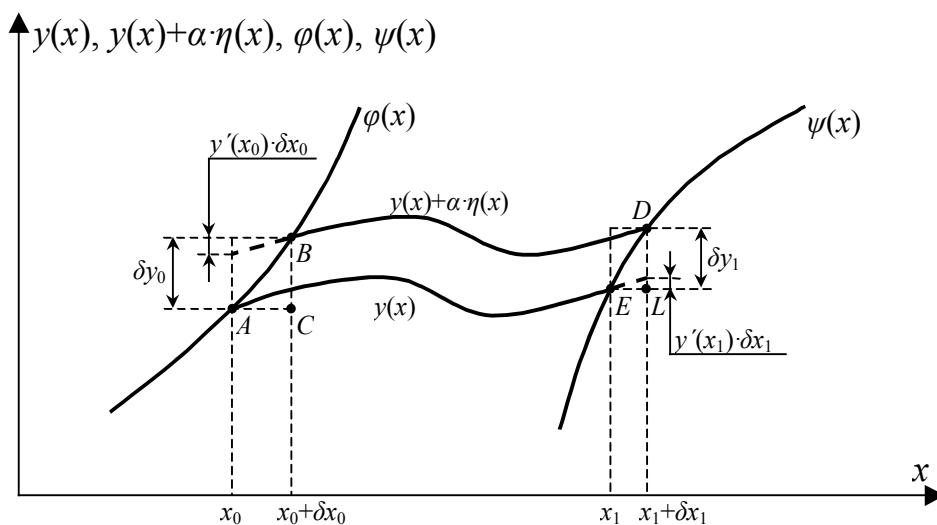


Рисунок 4.7 – Графічна інтерпретація задачі мінімізації функціоналу в області функцій з рухомими по кривих  $\varphi(x)$  і  $\psi(x)$  кінцями

Для розв’язання цієї задачі потрібно спочатку ввести у вираз для першої варіації функціоналу не лише складову, що викликана варіацією екстремалі  $y(x)$ , як це вже зроблено раніше, але і другу складову, обумовлену варіаціями кінців екстремалі вздовж функцій  $\varphi(x)$  і  $\psi(x)$ .

Зрозуміло, що приріст функціонала  $\Delta J_y^F(\alpha)$  при переході від функції  $y(x)$  до  $y(x) + \alpha\eta(x)$  може бути поданий так (рис. 4.7):

$$\begin{aligned} \Delta J_y^F(\alpha) &= J_y^F(y + \alpha\eta) - J_y^F(y) = \\ &= \int_{x_0 + \delta x_0}^{x_1 + \delta x_1} F(x, y + \alpha\eta, y' + \alpha\eta') dx - \int_{x_0}^{x_1} F(x, y, y') dx. \end{aligned} \quad (4.70)$$

Виділяючи як області інтегрування  $[x_0, x_0 + \delta x_0)$ ,  $[x_0, x_1]$ ,  $(x_1, x_1 + \delta x_1]$  і пам’ятаючи, що операція інтегрування є лінійною, інтеграл в правій частині виразу (4.70) можна записати і так:

$$\begin{aligned} J_y^F(\alpha) &= \int_{x_0}^{x_1} (F(x, y + \alpha\eta, y' + \alpha\eta') - F(x, y, y')) dx + \\ &+ \int_{x_1}^{x_1 + \delta x_1} F(x, y + \alpha\eta, y' + \alpha\eta') dx - \int_{x_0}^{x_0 + \delta x_0} F(x, y + \alpha\eta, y' + \alpha\eta') dx. \end{aligned} \quad (4.71)$$

Зрозуміло, що перша варіація  $\delta J$  приросту функціонала  $\Delta J_y^F(\alpha)$  (4.71) буде мати три складових  $\delta J^{(x_0, x_1)}$ ,  $\delta J^{(x_1, x_1 + \delta x_1)}$ ,  $\delta J^{(x_0, x_0 + \delta x_0)}$ , тобто,

$$\delta J = \delta J^{(x_0, x_1)} + \delta J^{(x_1, x_1 + \delta x_1)} - \delta J^{(x_0, x_0 + \delta x_0)}. \quad (4.72)$$

Розкладаючи в ряд Тейлора перший інтеграл у виразі (4.71) так, як це зроблено в підрозділі 4.3 (вираз (4.24)), для першої складової  $\delta J^{(x_0, x_1)}$  першої варіації  $\delta J$  отримаємо після інтегрування частинами

$$\delta J^{(x_0, x_1)} = \int_{x_0}^{x_1} \left( F_y - \frac{d}{dx} F_{y'} \right) \cdot \eta \, dx + \left( F_{y'} \cdot \eta \right) \Big|_{x_0}^{x_1}. \quad (4.73)$$

Звертаємо увагу на те, що в співвідношенні (4.38) член  $\left( F_{y'} \cdot \eta \right) \Big|_a^b$  випав. Це було обумовлено тим, що на функцію  $\eta(x)$  ми накладали умови (4.21). Оскільки в задачі, яку ми розглядаємо зараз, умови (4.21) знято, то і в (4.73) цей член зберігається.

Оскільки інтервали  $\delta x_0$  і  $\delta x_1$  малі, то визначені інтеграли (другий і третій у виразі (4.70)) на них можна наближено подати за формулою прямокутників, тобто

$$\int_{x_1}^{x_1 + \delta x_1} F(x, y + \alpha \eta, y' + \alpha \eta') \, dx \approx F \Big|_{x_1} \cdot \delta x_1, \quad (4.74)$$

$$\int_{x_0}^{x_0 + \delta x_0} F(x, y + \alpha \eta, y' + \alpha \eta') \, dx \approx F \Big|_{x_0} \cdot \delta x_0. \quad (4.75)$$

Але ж праві частини у виразах (4.74) і (4.75) це і є перші варіації функціоналів, що стоять в лівих частинах виразів. Тому, з урахуванням (4.73), (4.74), (4.75), ми можемо (4.72) записати так

$$\begin{aligned} \delta J = & \int_{x_0}^{x_1} \left( F_y - \frac{d}{dx} F_{y'} \right) \cdot \eta \, dx + F_{y'} \cdot \eta(x_1) - \\ & - F_{y'} \cdot \eta(x_0) + F \Big|_{x_1} \cdot \delta x_1 - F \Big|_{x_0} \cdot \delta x_0. \end{aligned} \quad (4.76)$$

А тепер звернемось до рис. 4.7. З наведених на ньому графіків і допоміжних побудов видно, що при  $\alpha = 1$  (для спрощення і без втрати змістовності) мають місце співвідношення:

$$\delta y_0 = y' \cdot \delta x_0 + \eta(x_0), \quad (4.77)$$

$$\delta y_1 = y' \cdot \delta x_1 + \eta(x_1). \quad (4.78)$$

Вирази (4.77) і (4.78) можна переписати і так:

$$\eta(x_0) = \delta y_0 - y' \cdot \delta x_0, \quad (4.79)$$

$$\eta(x_1) = \delta y_1 - y' \cdot \delta x_1. \quad (4.80)$$

Підставляючи (4.79) і (4.80) в (4.76), отримаємо:

$$\begin{aligned} \delta J = & \int_{x_0}^{x_1} \left( F_y - \frac{d}{dx} F_{y'} \right) \cdot \eta \, dx + \left( F - y' \cdot F_{y'} \right) \Big|_{x_1} \cdot \delta x_1 - \\ & - \left( F - y' \cdot F_{y'} \right) \Big|_{x_0} \cdot \delta x_0 + F_{y'} \Big|_{x_1} \cdot \delta y_1 - F_{y'} \Big|_{x_0} \cdot \delta y_0. \end{aligned} \quad (4.81)$$

Це і є вираз для першої варіації функціонала в задачі пошуку екстремалей з рухомими кінцями.

Припустимо, що ми розв'язали задачу і знайшли екстремаль, яка доставляє мінімум функціоналу (4.18).

В цьому випадку інтегральний член, який містить в собі рівняння Ейлера, у виразі (4.76) дорівнюватиме нулю, а перша варіація  $\delta J$  функціонала складатиме

$$\begin{aligned} \delta J = & \left( F - y' \cdot F_{y'} \right) \Big|_{x_1} \cdot \delta x_1 - \left( F - y' \cdot F_{y'} \right) \Big|_{x_0} \cdot \delta x_0 + \\ & + F_{y'} \Big|_{x_1} \cdot \delta y_1 - F_{y'} \Big|_{x_0} \cdot \delta y_0. \end{aligned} \quad (4.82)$$

Знову повернемося до рис. 4.7. При малих  $\delta x_0$ ,  $\delta x_1$  фігури ABC і EDL можна вважати прямокутними трикутниками, з яких

$$\begin{cases} \delta y_0 = \varphi'(x) \cdot \delta x_0, \\ \delta y_1 = \psi'(x) \cdot \delta x_1. \end{cases} \quad (4.83)$$

Підставляючи (4.83) у (4.82), отримаємо

$$\delta J = \left( F - (y' - \psi') \cdot F_{y'} \right) \Big|_{x_1} \cdot \delta x_1 - \left( F - (y' - \varphi') \cdot F_{y'} \right) \Big|_{x_0} \cdot \delta x_0. \quad (4.84)$$

Як нам уже відомо з попереднього матеріалу, на екстремалі перша варіація функціонала має дорівнювати нулю, тобто має виконуватись умова

$$\delta J = 0. \quad (4.85)$$

А виконання цієї умови, в свою чергу, може задовольнятися лише за умови, що нульовими стають значення коефіцієнтів при  $\delta x_0$  і  $\delta x_1$  у співвідношенні (4.84), тобто що

$$\begin{cases} (F - (y' - \psi') \cdot F_{y'})|_{x_1} = 0, \\ (F - (y' - \varphi') \cdot F_{y'})|_{x_0} = 0, \end{cases} \quad (4.86)$$

або в канонічних змінних –

$$\begin{cases} (H + \psi' \cdot p)|_{x_1} = 0, \\ (H + \varphi' \cdot p)|_{x_0} = 0. \end{cases} \quad (4.87)$$

**Умови, що задаються виразами (4.86) або (4.87), математики називають умовами трансверсальності.**

Вони дозволяють знайти точки перетину екстремалі  $y(x)$ , яка є розв'язком рівняння Ейлера, з лініями  $\varphi(x)$  і  $\psi(x)$ , котрі задають закони зміни рухомих кінців цієї екстремалі.

Оскільки рівняння Ейлера є диференціальним рівнянням другого порядку, то його загальний розв'язок, який задає множину екстремалей, містить дві довільні сталі.

Для конкретизації екстремалі ці сталі визначаються з рівнянь (4.86) або (4.87), кількість яких має дорівнювати кількості шуканих сталих.

#### 4.6 Задачі, що приводять до екстремалей зі зломами

Розглянемо функціонал

$$J = \int_1^3 y^4 (2 + y')^2 dx. \quad (4.88)$$

Оскільки в цей функціонал функція  $y(x)$  і функція від її похідної  $(2 + y')$  входять у парних степенях, то функціонал ніколи не може мати від'ємних значень. Тож найменшим його значенням може бути лише нуль.

Нульові значення функціонала (4.88), як легко бачити, досягаються або на функції

$$y = 0, \quad (4.89)$$

або на функції, яка є розв'язком рівняння

$$2 + y' = 0, \quad (4.90)$$

що утворюється від прирівнювання нулю виразу, який стоїть у дужках.

Розв'язком рівняння (4.90) є функція

$$y(x) = -2x + C, \quad (4.91)$$

що задає сімейство прямих ліній.

Нам уже відомо, що функції, які доставляють функціоналу мінімум, називаються екстремалами. Тож для функціонала (4.88) екстремалами є функції (4.89) і функції (4.91).

Нехай для екстремали задані граничні умови у вигляді

$$\begin{cases} y(1) = 0, \\ y(3) = -2. \end{cases} \quad (4.92)$$

Зрозуміло, що для екстремали (4.89) перша умова з (4.92) може бути виконана, а друга – ні.

Для екстремалей (4.91) перша умова може бути виконана лише на прямій

$$y(x) = -2x + 2, \quad (4.93)$$

а друга – на прямій

$$y(x) = -2x + 4, \quad (4.94)$$

тобто, ні в класі функцій (4.89), ні в класі функцій (4.91) немає жодної такої функції, графік якої проходив би через точки, що задані умовами (4.92).

Якщо функцію (4.94), значення константи  $C$  в якій отримано з другої граничної умови (4.92), прирівняти до нуля

$$-2x + 4 = 0, \quad (4.95)$$

то знайдемо, що її графік перетинає вісь  $x$  в точці  $x_0 = 2$  (рис. 4.8).

Тож із викладеного вище випливає, що екстремаллю функціонала (4.88) є функція

$$y(x) = \begin{cases} 0 & \text{при } x \in [1, 2], \\ -2x + 4 & \text{при } x \in [2, 3], \end{cases} \quad (4.96)$$

графік якої зображено на рис. 4.8 суцільною лінією.

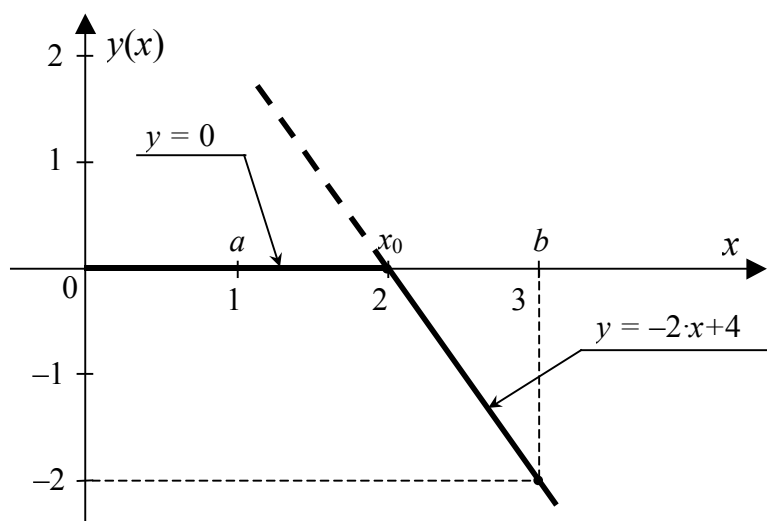


Рисунок 4.8 – Графік екстремалі функціонала (4.88) в області  $x \in [a, b]$ , де  $a = 1$ ,  $b = 3$

Як бачимо, графік цієї функції має злом в точці  $x_0 = 2$ , а сама функція належить до класу кусково-гладких на відрізку  $x \in [1, 3]$ . Легко переконатись, що для неї на всьому відрізку значень  $x \in [1, 3]$  виконується умова (4.64), а умова (4.66) виконується лише для точок  $x \in [1, 2]$ .

Оскільки рівняння Ейлера, як ми встановили раніше, є диференціальним рівнянням другого порядку, то його розв'язок має містити дві сталі, значення яких визначається або з заданих граничних умов для задач з нерухомим кінцями екстремалей, або з заданих умов трансверсальності, якщо ці кінці є рухомими. І в першому, і в другому випадках для визначення конкретних значень двох сталих потрібно було мати два рівняння.

Якщо ж екстремум функціонала досягається на кривій з точкою злому, то двох рівнянь для визначення сталих уже буде недостатньо, оскільки їх значення для екстремалі на ділянці до точки злому і після цієї точки будуть різними, адже різними є рівняння, які задають екстремаль на цих ділянках.

Зрозуміло, що для визначення сталих у розв'язку рівняння Ейлера для екстремалей з однією точкою злому  $x_0 \in [a, b]$  потрібно мати чотири рівняння. Два з них будуть задаватись граничними умовами, тобто значеннями екстремалі в точках  $x = a$  і  $x = b$

$$\begin{cases} y(a) = y_a, \\ y(b) = y_b, \end{cases} \quad (4.97)$$

а два інших потрібно визначити в точці злому  $x_0$  при наближенні до неї по екстремалі з лівого боку, тобто з боку  $(x_0 - 0)$ , та з правого боку, тобто з боку  $(x_0 + 0)$ .

Для отримання цих додаткових рівнянь в загальному вигляді для функціонала (4.18) при наявності однієї точки злому  $x_0$  екстремалі  $y(x)$  подамо функціонал (4.18) у вигляді

$$J = \int_a^{x_0} F(x, y, y') dx + \int_{x_0}^b F(x, y, y') dx. \quad (4.98)$$

Зрозуміло, що перша варіація  $\delta J$  функціонала  $J$  в цьому випадку дорівнюватиме сумі перших варіацій  $\delta J_1$  і  $\delta J_2$  від кожного з інтегралів, що входять до (4.98), тобто

$$\delta J = \delta J_1 + \delta J_2. \quad (4.99)$$

В загальному вигляді кожна з цих варіацій  $\delta J_1$ ,  $\delta J_2$  може бути записана згідно з виразом (4.81). Але, якщо в областях значень  $x \in [a, x_0]$ ,  $x \in [x_0, b]$  функція  $y(x)$  є екстремаллю функціонала (4.18), то інтегральні члени для кожної зі складових варіацій  $\delta J_1$ ,  $\delta J_2$  дорівнюватимуть нулю. Тому в точці злому  $x_0$  ці варіації, використовуючи вираз (4.84), можна записати так:

$$\delta J_1 = (F - y'F_{y'})\Big|_{x_0-0} \cdot \delta x_0 + F_{y'}\Big|_{x_0-0} \cdot \delta y_0; \quad (4.100)$$

$$\delta J_2 = (F - y'F_{y'})\Big|_{x_0+0} \cdot (-\delta x_0) + F_{y'}\Big|_{x_0+0} \cdot (-\delta y_0). \quad (4.101)$$

Зауважимо, що у виразах (4.100), (4.101) варіаціям аргументу  $\delta x_0$  і функції  $\delta y_0$  ми встановили знак «+» при їх прирості зліва направо та знизу вгору (по стрілках осей координат), а знак «-» – при їх прирості справа наліво та зверху вниз в околі точки  $x_0$  (проти напрямку стрілок осей координат).

Як ми вже встановили в попередніх підрозділах, на екстремалі виконується умова

$$\delta J = 0, \quad (4.102)$$

що для екстремалі зі зломом виглядатиме так:

$$\delta J_1 + \delta J_2 = 0. \quad (4.103)$$

Підставляючи (4.100) і (4.101) у (4.103), матимемо

$$\begin{aligned} & \left( (F - y'F_{y'})\Big|_{x_0-0} - (F - y'F_{y'})\Big|_{x_0+0} \right) \cdot \delta x_0 + \\ & + \left( F_{y'}\Big|_{x_0-0} - F_{y'}\Big|_{x_0+0} \right) \cdot \delta y_0 = 0. \end{aligned} \quad (4.104)$$

При довільних  $\delta x_0$  і  $\delta y_0$  рівність (4.104) виконується лише за умови рівності нулю коефіцієнтів, які стоять перед  $\delta x_0$  і  $\delta y_0$ , тобто за умови, що

$$\begin{cases} (F - y'F_{y'})|_{x_0-0} = (F - y'F_{y'})|_{x_0+0}, \\ F_{y'}|_{x_0-0} = F_{y'}|_{x_0+0} \end{cases} \quad (4.105)$$

або при застосуванні канонічних змінних  $H, p$

$$\begin{cases} H|_{x_0-0} = H|_{x_0+0}, \\ p|_{x_0-0} = p|_{x_0+0}. \end{cases} \quad (4.106)$$

*Умови (4.105) чи (4.106) називають умовами Вейєрштрасса–Ердмана. Вони і задають ті додаткові рівняння, які потрібно мати для однозначного визначення екстремалі  $y(x)$ , отриманої шляхом розв’язання рівняння Ейлера зліва і справа від точки злому  $x_0 \in [a, b]$ .*

Введенням цих умов ми розширили клас функцій, на яких може досягатись екстремум функціонала (4.18).

Якщо в попередніх підрозділах ми шукали екстремалі лише в класі гладких функцій, графіками яких є плавні криві, то тепер ми зможемо розв’язувати задачу мінімізації функціонала в більш широкому класі – класі кусково-гладких функцій, графіками яких є неперервні криві, що містять точки злому між плавними складовими графіка.

#### **4.7 Пошук екстремумів функціоналів, що пов’язують між собою декілька невідомих функцій та їх перших похідних**

Розглянемо задачу пошуку екстремуму функціонала

$$J = \int_a^b F(x, y_1, y_2, \dots, y_n; y'_1, y'_2, \dots, y'_n) dx, \quad (4.107)$$

який пов’язує множину функцій

$$\{y_1(x), y_2(x), \dots, y_n(x)\}, \quad (4.108)$$

що задає поверхню в  $n$ -вимірному просторі, та множину

$$\{y'_1(x), y'_2(x), \dots, y'_n(x)\} \quad (4.109)$$

перших похідних цих функцій у цьому ж просторі.

Зрозуміло, що якщо екстремум функціонала (4.107) існує, то його екстремалі знаходяться в множині (4.108).





$$J_{y_2}^F = \int_a^b F(x, C_1, y_2, C_3, C_4, \dots, C_n; C_{11}, y_2', C_{33}, C_{44}, \dots, C_{nn}) dx, \quad (4.115)$$

який теж формально не відрізняється від функціонала (4.18).

Мінімум такого функціонала, за аналогією з попереднім випадком, досягається на екстремалі  $y_2^*(x)$ , котра є розв'язком рівняння Ейлера

$$F_{y_2} - \frac{d}{dx} F_{y_2'} = 0. \quad (4.116)$$

Цей процес продовжуватимемо, аж поки не отримаємо екстремаль  $y_n^*(x)$ , яка є розв'язком рівняння Ейлера

$$F_{y_n} - \frac{d}{dx} F_{y_n'} = 0. \quad (4.117)$$

А тепер зведемо всі рівняння Ейлера, від (4.113) до (4.117), в одну систему

$$\begin{cases} F_{y_1} - \frac{d}{dx} F_{y_1'} = 0, \\ F_{y_2} - \frac{d}{dx} F_{y_2'} = 0, \\ \dots\dots\dots \\ F_{y_n} - \frac{d}{dx} F_{y_n'} = 0. \end{cases} \quad (4.118)$$

Зрозуміло, що множина функцій

$$\{y_1^*(x), y_2^*(x), \dots, y_n^*(x)\}, \quad (4.119)$$

яка є сумісним розв'язком системи рівнянь Ейлера (4.118), і буде множиною екстремалей функціонала (4.107), на якій він досягає екстремуму.

Як і у випадку функціонала (4.18), тим, що множина функцій (4.119) є розв'язком системи (4.118), задається лише необхідна умова існування екстремуму функціонала (4.107).

Для перевірки достатніх умов існування екстремуму функціонала (4.107) на множині екстремалей (4.119) потрібно, як і у випадку функціонала (4.18), переконатись, що виконуються умови Лежандра, які для однієї екстремалі мали вигляд (4.59) для мінімуму і (4.60) для максимуму, а для системи екстремалей (4.119) матимуть вигляд (для мінімуму)

$$\left\{ \begin{array}{l} F_{y_1 y_1'} \geq 0, \\ \left| \begin{array}{cc} F_{y_1 y_1'} & F_{y_1 y_2'} \\ F_{y_2 y_1'} & F_{y_2 y_2'} \end{array} \right| \geq 0, \\ \dots\dots\dots \\ \left| \begin{array}{cccc} F_{y_1 y_1'} & F_{y_1 y_2'} & \dots & F_{y_1 y_n'} \\ F_{y_2 y_1'} & F_{y_2 y_2'} & \dots & F_{y_2 y_n'} \\ \dots\dots\dots & \dots & \dots & \dots \\ F_{y_n y_1'} & F_{y_n y_2'} & \dots & F_{y_n y_n'} \end{array} \right| \geq 0. \end{array} \right. \quad (4.120)$$

Достатні умови існування максимуму функціонала (4.107) на множині екстремалей (4.119) за Лежандром будуть мати вигляд, аналогічний (4.120), але знаки нерівностей в них будуть протилежними.

Той, кого зацікавить виведення умов Лежандра для функціонала (4.107), може ознайомитись з його алгоритмом в будь-якому повному підручнику з варіаційного числення.

#### 4.8 Пошук екстремумів функціоналів, що залежать від старших похідних невідомої функції

Розглянемо функціонал

$$J = \int_a^b F(x, y, y', y'', \dots, y^{(n)}) dx, \quad (4.121)$$

який пов'язує між собою не лише функцію  $y(x)$  та її похідну  $y'(x)$ , як у випадку (4.18), але і старші похідні  $y''(x), \dots, y^{(n)}(x)$ .

Ще Ейлер в першій половині вісімнадцятого сторіччя довів, що функція  $y(x)$  буде екстремаллю функціонала (4.121) в тому випадку, якщо вона є розв'язком рівняння

$$F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} - \dots + (-1)^n \cdot \frac{d^n}{dx^n} F_{y^{(n)}} = 0. \quad (4.122)$$

Ейлер отримав це рівняння шляхом математичної індукції.

Він спочатку розглянув задачу пошуку функції, на якій функціонал (4.18) досягає мінімуму. Цей процес нами детально описаний у попередніх підрозділах. Потім Ейлер розглянув функціонал

$$J = \int_a^b F(x, y, y', y'') dx, \quad (4.123)$$

для якого, йдучи тим же шляхом з виділенням першої варіації, прирівнюванням її нулю і

взяттям другої частини інтеграла частинами, отримав умову, що функція  $y(x)$  є екстремаллю, якщо вона є розв'язком рівняння

$$F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} = 0. \quad (4.124)$$

Після цього Ейлер зробив висновок, що функціонал (4.121) досягатиме екстремуму на функціях, які є розв'язками рівняння (4.122). Розглянувши декілька прикладів з похідними з порядком, вищим другого, він переконався, що це так. А строге доведення того, що екстремаллю функціонала (4.121) є функція  $y(x)$ , яка задовольняє рівняння (4.122), здійснив Пуассон – ось чому це рівняння й увійшло у варіаційне числення з подвійною назвою – рівняння Ейлера-Пуассона.

Для функціонала (4.123) рівняння (4.122), яке скорочується до (4.124), є диференціальним рівнянням четвертого порядку. Його розв'язком буде функція  $y(x, C_1, C_2, C_3, C_4)$ , яка містить чотири довільні сталі  $C_1, C_2, C_3, C_4$ , для визначення яких потрібно мати чотири рівняння. Такими рівняннями будуть граничні умови

$$\begin{cases} y(a, C_1, C_2, C_3, C_4) = A_1, \\ y'(a, C_1, C_2, C_3, C_4) = A_2, \\ y(b, C_1, C_2, C_3, C_4) = B_1, \\ y'(b, C_1, C_2, C_3, C_4) = B_2. \end{cases} \quad (4.125)$$

Зрозуміло, що необхідною умовою екстремуму функціонала (4.121), в якому старшою похідною від невідомої функції  $y(x)$  є похідна третього порядку  $y'''(x)$ , буде умова, щоб ця функція задовольняла рівняння (4.122), яке в цьому випадку матиме шостий порядок. Шість довільних сталих цієї екстремалі потрібно буде знаходити з граничних умов типу (4.125) з тією різницею, що рівнянь уже буде шість і задаватимуть вони на границях не лише значення самої екстремалі та її першої похідної, як це було у випадку мінімізації функціонала (4.123), але й її другої похідної, тобто система рівнянь (4.125) доповниться ще рівняннями

$$\begin{cases} y''(a, C_1, C_2, C_3, C_4, C_5, C_6) = A_3, \\ y''(b, C_1, C_2, C_3, C_4, C_5, C_6) = B_3. \end{cases} \quad (4.126)$$

Ну і, звичайно, в усіх рівняннях (4.125) для цього випадку додадуться довільні сталі  $C_5$  і  $C_6$ .

Дуже простими для цієї задачі виявились умови Лежандра, за якими мінімум функціонала (4.121) на екстремалі  $y(x)$  відрізняють від максимуму. Вони є настільки простими, що в навчальному посібнику для ІТ-спеціальностей цю простоту не варто, на думку авторів посібника, затьмарювати громіздкими доведеннями.

Тож, для того, щоб на екстремалі  $y(x)$  досягався мінімум функціоналу (4.121), достатньо виконання умови

$$F_{y^{(n)}y^{(n)}} \geq 0, \quad (4.127)$$

а для максимуму –

$$F_{y^{(n)}y^{(n)}} \leq 0. \quad (4.128)$$

З виведенням умов (4.127), (4.128), за бажанням, можна ознайомитись в будь-якому повному підручнику з варіаційного числення, призначеному для студентів університетів математичних спеціальностей.

#### 4.9 Завдання для самоперевірки

1. Розкрийте поняття «функція» та способи її задавання.
2. Дайте означення неперервної, гладкої та кусково-гладкої функцій.
3. Розкрийте поняття «функціонал» і наведіть приклади його задавання.
4. Дайте графічну інтерпретацію умов існування екстремуму функції однієї незалежної змінної.
5. Як знайти координату точки екстремуму для функції однієї незалежної змінної?
6. Як розрізнити, де в точці екстремуму має місце мінімум функції однієї незалежної змінної, а де – максимум?
7. Які умови існування екстремуму функції однієї незалежної змінної є необхідними, а які – достатніми?
8. Як визначити відсутність екстремуму в точці, для якої необхідні умови існування екстремуму функції однієї незалежної змінної виконуються?
9. Що собою являють локальний та глобальний екстремуми функції?
10. Дайте графічну інтерпретацію функції двох незалежних змінних.
11. Як знайти координати точки екстремуму для функції двох незалежних змінних?
12. Як розпізнати, де в точці екстремуму має місце мінімум функції двох незалежних змінних, а де – максимум?
13. Які умови існування екстремуму функції двох незалежних змінних є необхідними, а які – достатніми?
14. Як визначити відсутність екстремуму в точці, для якої необхідні умови існування екстремуму функції двох незалежних змінних виконуються?
15. Як змістовно відрізнити абсолютний та відносний екстремуми функціонала? Наведіть приклад.
16. Як визначити відстань нульового і першого порядку між двома функціями?
17. Коли на екстремалі досягається сильний відносний екстремум функціонала? Коли досягається слабкий відносний екстремум?
18. Який послідовний логічний зв'язок між слабким, сильним та абсолютним екстремумами функціонала? Чи зберігається логічна послідовність у зворотному напрямку?
19. Виведіть рівняння Ейлера. Дайте його інтерпретацію.
20. Отримайте рівняння Ейлера у вигляді нелінійного диференціального рівняння другого порядку.
21. Отримайте перший інтеграл рівняння Ейлера. Дайте його інтерпретацію.
22. Виведіть рівняння Ейлера у канонічній формі.
23. Якими є достатні умови існування екстремуму функціонала в найпростішій задачі варіаційного числення? Як їх вивести і як з їх допомогою розрізнити максимум і мінімум функціонала?
24. Як відрізнити мінімум від максимуму функціонала, якщо  $F_{y'y'} = 0$ ?

25. Які функціонали є виродженими і що можна сказати про їх екстремалі?
26. Виведіть співвідношення для першої варіації функціонала з рухомими кінцями.
27. Отримайте умови трансверсальності. Поясніть, для чого вони потрібні.
28. Наведіть приклад задачі мінімізації функціонала, для якої  $F_{y,y'} = 0$ , а екстремаль має злом графіка.
29. Виведіть умови Вейерштрасса–Ердмана. Поясніть для чого вони потрібні.
30. Виведіть необхідні умови існування екстремуму функціонала, який пов'язує декілька невідомих функцій та їх перших похідних.
31. Якими є достатні умови існування екстремуму функціонала, що пов'язує декілька невідомих функцій та їх перших похідних?
32. Виведіть евристично (за Ейлером) рівняння Ейлера–Пуассона. Дайте його інтерпретацію.
33. Якими є достатні умови існування екстремуму функціонала, що пов'язує екстремаль з її старшими похідними? Як розрізнити максимум та мінімум функціонала в цьому випадку?

## Розділ 5

# ВАРІАЦІЙНІ МЕТОДИ ДОСЛІДЖЕННЯ ФУНКЦІОНАЛІВ НА УМОВНИЙ ЕКСТРЕМУМ

**5.1 Метод невизначених множників Лагранжа як найбільш загальний для дослідження функціоналів на умовний екстремум**

Нехай дано функціонал

$$J = \int_a^b F(x, y, z, y', z') dx \quad (5.1)$$

і необхідно знайти такі функції

$$y(x), z(x), \quad (5.2)$$

які доставляють екстремум функціоналу (5.1) за умови, що

$$\varphi(x, y, z) = 0, \quad (5.3)$$

тобто за умови, що всі точки кривої (5.2) лежать на поверхні (5.3).

Спробуємо знайти необхідні умови існування екстремуму функціонала (5.1) за наявності обмеження (5.3).

Припустимо, що ми знайшли такі функції  $y(x)$  і  $z(x)$ , які доставляють екстремум, наприклад, мінімум, функціоналу (5.1).

Додамо до функцій  $y(x)$  і  $z(x)$  варіації  $\delta y$  і  $\delta z$ , які задовольняють вимоги

$$\begin{cases} \delta y > 0 & \text{для } x \in [x_1, x_2], \\ \delta y = 0 & \text{для } x \notin [x_1, x_2], \\ \delta z > 0 & \text{для } x \in [x_1, x_2], \\ \delta z = 0 & \text{для } x \notin [x_1, x_2], \\ [x_1, x_2] \subset [a, b], \end{cases} \quad (5.4)$$

і знайдемо першу варіацію  $\delta J$  функціонала (5.1) при переході від кривої на поверхні (5.3), що описується функціями  $y(x)$ ,  $z(x)$ , до кривої на цій же поверхні, що описується функціями

$$\begin{cases} y(x) + \delta y, \\ z(x) + \delta z. \end{cases} \quad (5.5)$$

Зрозуміло, що

$$\delta J = \delta J_y + \delta J_z, \quad (5.6)$$

Тобто, перша варіація  $\delta J$  функціонала (5.1) дорівнюватиме сумі першої варіації  $\delta J_y$  цього ж функціонала за координатою  $y$  при незмінній координаті  $z$ , та першої варіації  $\delta J_z$  за координатою  $z$  при незмінній координаті  $y$ .

Очевидно, що в точці  $(x_0, y_0, z_0)$  мінімуму функціонала (5.1) справедливим буде співвідношення

$$\delta J = 0. \quad (5.7)$$

А це можливо в силу незалежності  $y$  і  $z$  одне від одного лише за умови, що

$$\begin{cases} \delta J_y = 0, \\ \delta J_z = 0. \end{cases} \quad (5.8)$$

Але, як ми вже переконались при виведенні рівняння Ейлера, перші варіації  $\delta J_y$  і  $\delta J_z$  можна подати у вигляді

$$\delta J_y = \int_a^b \left( F_y - \frac{d}{dx} F_{y'} \right) \cdot \delta y \cdot dx, \quad (5.9)$$

$$\delta J_z = \int_a^b \left( F_z - \frac{d}{dx} F_{z'} \right) \cdot \delta z \cdot dx. \quad (5.10)$$

Тож, підставляючи (5.9) і (5.10) у співвідношення (5.6), матимемо

$$\delta J = \int_a^b \left( F_y - \frac{d}{dx} F_{y'} \right) \cdot \delta y \cdot dx + \int_a^b \left( F_z - \frac{d}{dx} F_{z'} \right) \cdot \delta z \cdot dx = 0. \quad (5.11)$$

Нехай  $[x_1, x_2]$  – це малий окіл за координатою  $x$  точки  $(x_0, y_0, z_0)$ , в якій досягається мінімум функціонала (5.1), і  $x_0 \in [x_1, x_2]$ .

Тоді вирази

$$\sigma_1 = \int_{x_1}^{x_2} \delta y \cdot dx, \quad (5.12)$$

$$\sigma_2 = \int_{x_1}^{x_2} \delta z \cdot dx \quad (5.13)$$



задаватимуть малі прямокутні площинки  $\sigma_1$ ,  $\sigma_2$  на координатних площинах  $(x, y)$ ,  $(x, z)$  в околі точки  $x_0$  зі сторонами  $dx$ ,  $\delta y$  та  $dx$ ,  $\delta z$ , а вирази (5.9), (5.10) задаватимуть об'єми призм з основами  $dx$ ,  $\delta y$  та  $dx$ ,  $\delta z$ , які знаходяться під поверхнями

$$F_y - \frac{d}{dx} F_{y'}, \quad (5.14)$$

$$F_z - \frac{d}{dx} F_{z'}. \quad (5.15)$$

Але ми знаємо, що об'єм прямокутної призми з малою прямокутною площиною в основі можна знайти перемноженням площі основи на висоту цієї призми, що дає нам право переписати вираз (5.11) у вигляді

$$\delta J = \left( F_y - \frac{d}{dx} F_{y'} \right) \Big|_{x_0} \cdot \sigma_1 + \left( F_z - \frac{d}{dx} F_{z'} \right) \Big|_{x_0} \cdot \sigma_2 = 0. \quad (5.16)$$

Як ми вже відзначали, крива, що описується функціями (5.5), теж лежить на поверхні (5.3), тому для неї теж є справедливою рівність (5.3), яка набуває в цьому випадку вигляду

$$\varphi(x, y + \delta y, z + \delta z) = 0. \quad (5.17)$$

Проінтегруємо рівняння (5.3), (5.17) в області  $x \in [x_1, x_2]$ . Зрозуміло, що в результаті інтегрування ми отримаємо

$$\int_{x_1}^{x_2} \varphi(x, y + \delta y, z + \delta z) dx = 0, \quad (5.18)$$

$$\int_{x_1}^{x_2} \varphi(x, y, z) dx = 0. \quad (5.19)$$

Віднімемо рівняння (5.19) від (5.18). Отримаємо

$$\int_{x_1}^{x_2} (\varphi(x, y + \delta y, z + \delta z) - \varphi(x, y, z)) dx = 0. \quad (5.20)$$

Формально під інтегралом у рівнянні (5.20) стоїть приріст  $\Delta\varphi$  функції  $\varphi(x, y, z)$ , тобто

$$\Delta\varphi = \varphi(x, y + \delta y, z + \delta z) - \varphi(x, y, z), \quad (5.21)$$

який вона отримує при переході від точки поверхні з координатами  $(x, y, z)$  до точки цієї ж поверхні з координатами  $(x, y + \delta y, z + \delta z)$ .

Нагадаємо, що в околі точки  $x_0 \in [x_1, x_2]$  приріст функції дорівнює її першій варіації, тобто

$$\Delta \varphi = \delta \varphi. \quad (5.22)$$

Але

$$\delta \varphi = \varphi_y \cdot \delta y + \varphi_z \cdot \delta z. \quad (5.23)$$

Тож, з урахуванням виразів (5.21), (5.22), (5.23), рівняння (5.20) можна переписати так:

$$\int_{x_1}^{x_2} (\varphi_y \cdot \delta y + \varphi_z \cdot \delta z) dx = 0, \quad (5.24)$$

або так:

$$\int_{x_1}^{x_2} \varphi_y \cdot \delta y \cdot dx + \int_{x_1}^{x_2} \varphi_z \cdot \delta z \cdot dx = 0. \quad (5.25)$$

З урахуванням виразів (5.12), (5.13) об'єми прямокутних призм, що задаються інтегралами у виразі (5.25), можна записати так:

$$\begin{cases} \int_{x_1}^{x_2} \varphi_y \cdot \delta y \cdot dx \approx \varphi_y|_{x_0} \cdot \sigma_1, \\ \int_{x_1}^{x_2} \varphi_z \cdot \delta z \cdot dx \approx \varphi_z|_{x_0} \cdot \sigma_2. \end{cases} \quad (5.26)$$

Підставляючи вираз (5.26) у (5.25), в околі точки  $x_0 \in [x_1, x_2]$  матимемо

$$\varphi_y \cdot \sigma_1 = -\varphi_z \cdot \sigma_2, \quad (5.27)$$

звідки

$$\sigma_2 = -\frac{\varphi_y}{\varphi_z} \cdot \sigma_1. \quad (5.28)$$

Підставляючи вираз (5.28) у (5.16), в околі точки  $x_0 \in [x_1, x_2]$  матимемо

$$\left(F_y - \frac{d}{dx}F_{y'}\right) \cdot \sigma_1 - \frac{\varphi_y}{\varphi_z} \cdot \left(F_z - \frac{d}{dx}F_{z'}\right) \cdot \sigma_1 = 0. \quad (5.29)$$

З рівняння (5.29) випливає, що

$$\frac{F_y - \frac{d}{dx}F_{y'}}{\varphi_y} = \frac{F_z - \frac{d}{dx}F_{z'}}{\varphi_z}. \quad (5.30)$$

Оскільки всі операції до отримання співвідношень (5.30) ми проводили з використанням операції інтегрування за координатою  $x$ , то кожне співвідношення у виразі (5.30) ми можемо прирівняти до невідомої нам функції  $\lambda(x)$ , тобто

$$\begin{cases} \frac{F_y - \frac{d}{dx}F_{y'}}{\varphi_y} = \lambda(x), \\ \frac{F_z - \frac{d}{dx}F_{z'}}{\varphi_z} = \lambda(x). \end{cases} \quad (5.31)$$

В свою чергу, співвідношення (5.31) можна переписати так:

$$\begin{cases} F_y - \frac{d}{dx}F_{y'} + \lambda(x) \cdot \varphi_y = 0, \\ F_z - \frac{d}{dx}F_{z'} + \lambda(x) \cdot \varphi_z = 0. \end{cases} \quad (5.32)$$

Після всіх цих викладок ми можемо стверджувати, що для того, щоб функції (5.2) доставляли екстремум функціоналу (5.1) за наявності обмежень (5.3), необхідно, щоб вони були розв'язком рівнянь (5.32).

Виходячи з виразу (5.31), формально у співвідношеннях (5.32) потрібно було б поставити знак «мінус» при членах  $\lambda(x) \varphi_y$ ,  $\lambda(x) \varphi_z$ , але, оскільки функція  $\lambda(x)$  є поки що невизначеною, то при ній можна ставити будь-який знак. Чому нам зручно поставити «плюс» стане зрозумілим з викладок, що слідують далі. І почнемо ці викладки з того, що сконструємо функцію

$$L = F(x, y, z, y', z') + \lambda(x) \cdot \varphi(x, y, z). \quad (5.33)$$

В точці  $(x_0, y_0, z_0)$  мінімуму функціонала (5.1) справедливим є вираз (5.3), тож зрозуміло, що в цій точці

$$L = F. \quad (5.34)$$

А тому ми маємо право від пошуку необхідних умов існування мінімуму функціонала  $J$  (5.1) перейти до пошуку умов існування мінімуму функціонала  $J^L$ , де

$$J^L = \int_a^b L dx = \int_a^b (F + \lambda(x) \cdot \varphi(x, y, z)) dx. \quad (5.35)$$

Цей перехід дозволяє перевести задачу пошуку екстремалей, які доставляють умовний екстремум функціоналу (5.1), в задачу пошуку екстремалей, які доставляють безумовний екстремум функціоналу (5.35), яку ми вже розв'язувати вміємо, оскільки знаємо, що за відсутності обмежень мінімум функціонала досягається на функціях, які є розв'язком системи рівнянь Ейлера, отриманої з підінтегральної функції кількох змінних цього функціонала.

Цілком очевидно, що для нашого випадку ця система набуває вигляду:

$$\begin{cases} L_y - \frac{d}{dx} L_{y'} = 0, \\ L_z - \frac{d}{dx} L_{z'} = 0. \end{cases} \quad (5.36)$$

Знаходячи  $L_y$ ,  $L_z$ ,  $L_{y'}$ ,  $L_{z'}$ , бачимо, що система (5.36) тотожна системі (5.32).

Але два рівняння системи (5.36) або, що одне й те ж, системи (5.32) не дозволяють однозначно знайти три невідомі функції

$$y(x), z(x), \lambda(x). \quad (5.37)$$

Тож їх потрібно доповнити третім рівнянням, за яке взяти рівняння обмеження (5.3). В цьому випадку система визначальних рівнянь є замкнутою і такою, що дає однозначний розв'язок задачі пошуку умовного екстремуму.

**Функцію  $\lambda(x)$  Лагранж, який її вперше ввів, назвав невизначеним множником, а тому у варіаційне числення, яке є невід'ємною частиною функціонального аналізу, метод пошуку екстремалей, на яких функціонали набувають умовного екстремуму, ввійшов під назвою методу невизначених множників Лагранжа.**

**Зауваження.** Після того, як введенням функції  $L$  (5.33) ми перевели задачу на умовний екстремум у розряд найпростішої задачі варіаційного числення, для неї стають справедливими і введені нами раніше умови трансверсальності для екстремалей з рухомими кінцями, і умови Вейерштрасса–Ердмана для екстремалей зі зломами, і достатні умови Лежандра, за допомогою яких мінімум відрізняють від максимуму, і необхідні умови існування екстремуму для функціоналів, які залежать від старших похідних. Потрібно лише не забувати у співвідношеннях, якими ці умови визначаються, замість функції  $F$  та її похідних підставляти функцію  $L$  та її похідні. При цьому, якщо  $J$  має вигляд (4.107), а обмеження є аналогом (5.3), тобто

$$\varphi(x, y_1, y_2, \dots, y_n) = 0, \quad (5.38)$$

то функція  $L$  за структурою буде подібною до (5.33), тобто матиме вигляд

$$L = F(x, y_1, y_2, \dots, y_n, y'_1, y'_2, \dots, y'_n) + \lambda(x) \cdot \varphi(x, y_1, \dots, y_n). \quad (5.39)$$

Але якщо як обмеження виступає система рівнянь

$$\varphi_j(x, y_1, \dots, y_n) = C_j, \quad j = \overline{1, m}, \quad m \leq n, \quad (5.40)$$

то за  $L$  потрібно брати функцію

$$L = F(x, y_1, y_2, \dots, y_n, y'_1, y'_2, \dots, y'_n) + \sum_{j=1}^m \lambda_j(x) \cdot (\varphi_j(x, y_1, \dots, y_n) - C_j). \quad (5.41)$$

В цьому випадку систему рівнянь Ейлера (4.118), в яких замість  $F_{y_i}, F_{y'_i}, i = \overline{1, n}$  беруться  $L_{y_i}, L_{y'_i}, i = \overline{1, n}$ , необхідно розв'язувати разом з системою рівнянь (5.40), оскільки потрібно знайти не лише  $n$  екстремалей  $y_i(x)$ , але й  $m$  невизначених множників Лагранжа  $\lambda_j(x)$ .

## 5.2 Ізопериметрична задача пошуку екстремалей функціоналів

*В 1732 році Леонардом Ейлером було здійснено перший підхід до розв'язання задачі пошуку екстремалей функціоналів в умовах обмежень, теж заданих функціоналами, яка дістала назву ізопериметричної.* В 1744 році ним же був опублікований розв'язок цієї задачі в найбільш загальному вигляді.

Формулювалась ізопериметрична задача пошуку екстремалей функціоналів так: серед кривих однакової довжини, або, що одне й те ж, однакового периметра, знайти криву, яка обмежує найбільшу площу.

Математично це можна записати так: знайти криву  $y(x)$ , яка доставляє екстремум функціоналу (4.18), тобто, функціоналу

$$J^F = \int_a^b F(x, y, y') dx,$$

яким оцінюється площа заданої фігури, за умови, що інший функціонал

$$J^K = \int_a^b K(x, y, y') dx, \quad (5.42)$$

яким задається довжина периметра цієї фігури, має стале значення  $J_0^K$ , тобто

$$\int_a^b K(x, y, y') dx = J_0^K. \quad (5.43)$$

Розв'язував цю задачу Ейлер надзвичайно складним шляхом, який сьогодні може бути цікавим лише фахівцям з історії математики.

Ми наведемо розв'язання ізопериметричної задачі, що було отримано Лагранжем через 15 років у набагато простіший спосіб за допомогою методу невизначених множників.

Суть процесу розв'язання задачі Лагранжем така: якщо у функціоналі (5.42) відпустити верхню границю, то отримаємо інтеграл зі змінною верхньою границею, тобто

$$\psi(x) = \int_a^x K(x, y, y') dx, \quad (5.44)$$

похідна від якого  $\psi'(x)$ , як відомо з курсу математичного аналізу, дорівнюватиме підінтегральній функції, тобто

$$\psi'(x) = K(x, y, y'). \quad (5.45)$$

Рівняння (5.45) можна розглядати як обмеження виду (5.40) при  $j = 1$ .

З урахуванням цього обмеження, виходячи з (5.41), функцію Лагранжа для ізопериметричної задачі можна записати таким чином:

$$L = F(x, y, y') + \lambda(x) \cdot (\psi'(x) - K(x, y, y')). \quad (5.46)$$

Тепер ми можемо ізопериметричну задачу переформулювати так: знайти функції  $y(x)$  та  $\psi(x)$ , які доставляють мінімум функціоналу

$$J^L = \int_a^b L(x, y, \psi, y', \psi') dx. \quad (5.47)$$

Згідно з методом невизначених множників Лагранжа рівняння Ейлера для функціонала (5.47) матимуть вигляд

$$\begin{cases} L_y - \frac{d}{dx} L_{y'} = 0, \\ L_\psi - \frac{d}{dx} L_{\psi'} = 0. \end{cases} \quad (5.48)$$

Підставляючи вираз (5.46) в систему (5.48), отримаємо

$$\begin{cases} F_y - \lambda(x) \cdot K_y - \frac{d}{dx}(F_{y'} - \lambda(x) \cdot K_{y'}) = 0, \\ -\frac{d}{dx}\lambda(x) = 0. \end{cases} \quad (5.49)$$

З другого рівняння системи (5.49) знайдемо, що

$$-\lambda(x) = C_1. \quad (5.50)$$

Підставляючи вираз (5.50) в перше рівняння системи (5.49), отримаємо рівняння

$$F_y + C_1 \cdot K_y - \frac{d}{dx}(F_{y'} + C_1 \cdot K_{y'}) = 0, \quad (5.51)$$

яке в загальному вигляді є нелінійним диференціальним рівнянням другого порядку. Розв'язок цього рівняння  $y(x, C_1, C_2, C_3)$  залежатиме від трьох сталих  $C_1, C_2, C_3$ , наявність двох із яких ( $C_2, C_3$ ) обумовлена другим порядком диференціального рівняння, а третьої ( $C_1$ ) – підстановкою (5.50). Для їх визначення, крім рівняння з трьома невідомими  $C_1, C_2, C_3$ , яке отримаємо із умови (5.43) після підстановки у функцію  $K(x, y, y')$  загального розв'язку  $y(x, C_1, C_2, C_3)$  та інтегрування результату підстановки за  $x$  в межах від  $a$  до  $b$  у вигляді

$$Q(C_1, C_2, C_3) = J_0^K, \quad (5.52)$$

необхідно синтезувати ще два рівняння з цими ж невідомими, що нескладно здійснити, використавши граничні умови, які матимуть вигляд

$$\begin{cases} y(a, C_1, C_2, C_3) = y_a, \\ y(b, C_1, C_2, C_3) = y_b. \end{cases} \quad (5.53)$$

Розв'язавши систему трьох рівнянь (5.52), (5.53) з трьома невідомими  $C_1, C_2, C_3$ , ми знайдемо ті їх значення

$$\{C_1^*, C_2^*, C_3^*\}, \quad (5.54)$$

які, після підставлення їх у загальний розв'язок  $y(x, C_1, C_2, C_3)$  рівняння (5.51), перетворюють цей розв'язок в екстремаль

$$y(x, C_1^*, C_2^*, C_3^*), \quad (5.55)$$





наприклад,  $y_{m+1}(x), \dots, y_n(x)$ , конкретні значення, наприклад,  $C_{m+1}, \dots, C_n$ . Тоді система (5.58) перетвориться на систему

$$\begin{cases} \varphi_1(x, y_1, \dots, y_m; y'_1, \dots, y'_m; C_{m+1}, \dots, C_n) = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \varphi_m(x, y_1, \dots, y_m; y'_1, \dots, y'_m; C_{m+1}, \dots, C_n) = 0 \end{cases} \quad (5.60)$$

$m$  диференціальних рівнянь першого порядку з  $m$  невідомими, розв'язком якої буде система  $m$  функцій

$$\begin{cases} y_1(x, C_1, C_{m+1}, \dots, C_n), \\ y_2(x, C_2, C_{m+1}, \dots, C_n), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_m(x, C_m, C_{m+1}, \dots, C_n), \end{cases} \quad (5.61)$$

в якій значення сталим  $C_{m+1}, C_{m+2}, \dots, C_n$  ми самі задали попередньо, а сталі  $C_1, \dots, C_m$  визначаються з граничних умов

$$\begin{cases} y_1(a, C_1, C_{m+1}, \dots, C_n) = y_{10}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_m(a, C_m, C_{m+1}, \dots, C_n) = y_{m0}, \end{cases} \quad (5.62)$$

які задані умовами задачі.

При інших заданих нами значеннях  $C_{m+1}, \dots, C_n$  отримаємо іншу систему функцій (5.61), яка теж задовольнить систему рівнянь (5.58).

**Але можна поставити децю іншу задачу, а саме: в умовах дії обмежень, які задаються системою рівнянь (5.58), знайти такі значення функцій  $y_{m+1}(x), \dots, y_n(x)$ , котрі на границі області їх задавання досягали б екстремального значення. Саме таку задачу поставив Майєр і вказав шлях її розв'язання, тому вона і отримала назву задачі Майєра.**

Для її розв'язання вводять  $(n - m)$  нових змінних

$$\begin{cases} y'_{m+1} = u_1, \\ y'_{m+2} = u_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y'_n = u_{n-m}, \end{cases} \quad (5.63)$$

формують функціонал

$$J = \int_a^b (u_1 + u_2 + \dots + u_{n-m}) dx \quad (5.64)$$

і знаходять  $(n - m)$  екстремалей, які доставляють екстремум цьому функціоналу в умовах дії обмежень, заданих системою рівнянь (5.58).

Функція Лагранжа для цієї задачі матиме вигляд

$$L = (u_1 + u_2 + \dots + u_{n-m}) + \sum_{j=1}^{n-m} \lambda_j(x) \cdot \varphi_j(\bullet), \quad (5.65)$$

де  $(n - m)$  функцій  $\varphi_j(\bullet)$ ,  $j = \overline{1, n - m}$ , беруться із системи рівнянь (5.58).

Зрозуміло, що шукані екстремалі мають задовольняти систему рівнянь Ейлера–Лагранжа

$$\begin{cases} L_{u_1} - \frac{d}{dx} L_{u_1'} = 0, \\ L_{u_2} - \frac{d}{dx} L_{u_2'} = 0, \\ \dots\dots\dots \\ L_{u_{n-m}} - \frac{d}{dx} L_{u_{n-m}'} = 0. \end{cases} \quad (5.66)$$

Система (5.66) після підстановки в неї функції Лагранжа (5.65) набуває вигляду:

$$\begin{cases} 1 + \sum_{j=1}^{n-m} \lambda_j(x) \cdot \varphi_{ju_1}(\bullet) = 0, \\ 1 + \sum_{j=1}^{n-m} \lambda_j(x) \cdot \varphi_{ju_2}(\bullet) = 0, \\ \dots\dots\dots \\ 1 + \sum_{j=1}^{n-m} \lambda_j(x) \cdot \varphi_{ju_{n-m}}(\bullet) = 0. \end{cases} \quad (5.67)$$

Для однозначного розв'язання задачі Майєра потрібно визначити  $n$  функцій  $y_i(x)$  та  $(n - m)$  невизначених множників Лагранжа  $\lambda_j(x)$ , для чого потрібно мати  $(2n - m)$  рівнянь. Рівнянь у трьох системах (5.58), (5.63) та (5.67) якраз і є  $(2n - m)$ . Тож, розв'язуючи сумісно всі ці три системи рівнянь, отримуємо розв'язок задачі Майєра.

## 5.4 Пошук екстремалей функціоналів за наявності обмежень у вигляді нерівностей

В попередніх підрозділах цього розділу ми розв'язували задачу пошуку екстремалей функціоналів за умови, що обмеження мали вигляд одного або системи рівнянь, якими задавалась якась поверхня. Це означало, що крива екстремалі має лежати на цій поверхні.

Але ж при розв'язанні конкретних прикладних задач пошуку екстремалей функціоналів може трапитись випадок, коли екстремаль необхідно шукати не серед кривих, які лежать на обмежувальній поверхні, а серед кривих, які проходять через область, що обмежується деякою поверхнею.

В цьому випадку обмеження вже задаватимуться не рівнянням чи системою рівнянь, а нерівністю або системою нерівностей.

Для побудови алгоритму розв'язування цього класу задач розглянемо спочатку таку задачу: знайти екстремаль  $y(x)$  функціонала

$$J = \int_a^b F(x, y, y') dx \quad (5.68)$$

за умови, що ця функція  $y(x)$ , котра доставлятиме екстремум функціоналу (5.68), задовольнятиме нерівність

$$y(x) \geq \varphi(x), \quad (5.69)$$

де  $\varphi(x)$  – довільна крива, наприклад така, як показано на рис. 5.1.

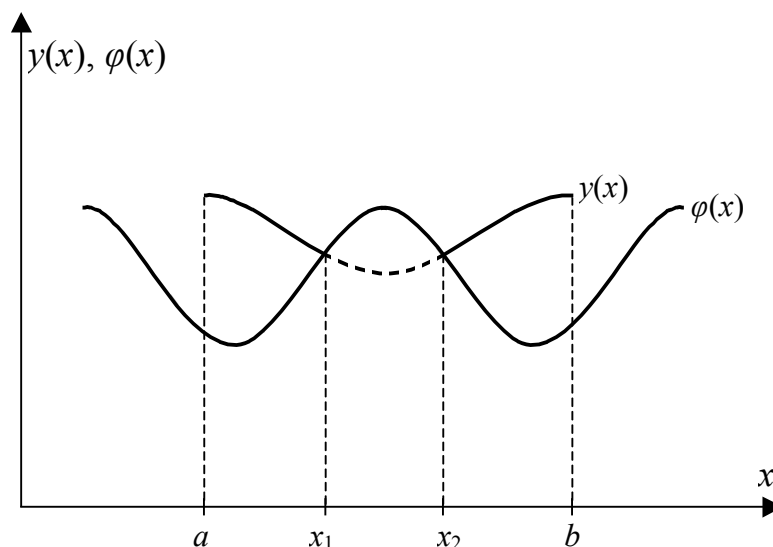


Рисунок 5.1 – Графічна інтерпретація обмеження в формі нерівності

З рис. 5.1 бачимо, що за відсутності обмеження (5.69) екстремаллю  $y(x)$  була б частина синусоїди в межах від  $x = a$  до  $x = b$  з ділянкою від  $x_1$  до  $x_2$ , показаною пунктирною лінією, включно.

Але за наявності обмеження (5.69) екстремум функціонала (5.68) може бути досягнутим лише на складеній кривій

$$y^*(x) = \begin{cases} y(x), & x \in [a, x_1], \\ \varphi(x), & x \in (x_1, x_2), \\ y(x), & x \in [x_2, b]. \end{cases} \quad (5.70)$$

Зрозуміло, що функцію  $y^*(x)$ , яка задається виразом (5.70), не можна отримати з рівняння Ейлера

$$F_y - \frac{d}{dx} F_{y'} = 0 \quad (5.71)$$

для функціонала (5.68), оскільки виведення цього рівняння базувалось на варіації функції  $y(x)$  в будь-якому напрямку зміни  $\delta y$ , а з рис. 5.1 видно, що на ділянці  $x \in (x_1, x_2)$  варіація функції  $y(x)$  може бути лише односторонньою, що входить в протиріччя з умовою виведення рівняння Ейлера.

Вихід із цієї ситуації можна знайти, якщо здійснити заміну змінних у такий спосіб.

Нехай

$$z^2 = y - \varphi, \quad (5.72)$$

звідки

$$y = z^2 + \varphi. \quad (5.73)$$

Диференціюючи вираз (5.73), отримаємо

$$y' = 2z \cdot z' + \varphi'. \quad (5.74)$$

В нових змінних  $(x, z)$  функціонал (5.68) набуде вигляду

$$J = \int_a^b F(x, z^2 + \varphi, 2z \cdot z' + \varphi') dx. \quad (5.75)$$

Введенням функції  $z(x)$  замість  $y(x)$  ми зняли обмеження (5.69) у вигляді нерівності, оскільки для функції  $z(x)$  воно виконується завжди.

Але у такий спосіб задачу пошуку екстремалі функціонала з обмеженням у вигляді нерівності ми звели до найпростішої задачі варіаційного числення – пошуку екстремуму функціонала (5.75) у відкритій області.

Як ми вже знаємо з попереднього, необхідні умови існування екстремуму функціонала (5.75) задаються рівнянням Ейлера

$$F_z - \frac{d}{dx} F_{z'} = 0. \quad (5.76)$$

З урахуванням виразів (5.73) і (5.74) для  $F_z, F_{z'}$  знайдемо, що

$$F_z = \frac{\partial F}{\partial z} = \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial z} + \frac{\partial F}{\partial y'} \cdot \frac{\partial y'}{\partial z} = F_y \cdot 2z + F_{y'} \cdot 2z', \quad (5.77)$$

$$F_{z'} = \frac{\partial F}{\partial z'} = \frac{\partial F}{\partial y'} \cdot \frac{\partial y'}{\partial z'} = F_{y'} \cdot 2z, \quad (5.78)$$

$$\frac{d}{dx} F_{z'} = \frac{d}{dx} (F_{y'} \cdot 2z) = 2z' \cdot F_{y'} + 2z \cdot \frac{d}{dx} F_{y'}. \quad (5.79)$$

Підставляючи вирази (5.77), (5.79) у вираз (5.76), отримаємо

$$2z \cdot F_y + 2z' \cdot F_{y'} - 2z' \cdot F_{y'} - 2z \cdot \frac{d}{dx} F_{y'} = 0 \quad (5.80)$$

або

$$2z \cdot \left( F_y - \frac{d}{dx} F_{y'} \right) = 0. \quad (5.81)$$

З рівняння (5.81) отримуємо два нових рівняння

$$z = 0, \quad (5.82)$$

$$F_y - \frac{d}{dx} F_{y'} = 0. \quad (5.83)$$

Розв'язком рівняння (5.82) є функція

$$y(x) = \varphi(x), \quad x \in (x_1, x_2), \quad (5.84)$$

яка є середньою складовою функції  $y^*(x)$ , заданої виразом (5.70), а розв'язком рівняння (5.83) є відрізки екстремалі  $y(x)$ , які є крайніми складовими функції  $y^*(x)$ .

Цими викладками ми довели, що екстремум функціонала (5.68) за обмежень (5.69) досягається на складеній кривій, однією частиною якої є екстремаль, а іншою – крива, яка є границею області обмеження.

Таким чином, щоб знайти загальний розв'язок поставленої на початку цього підрозділу задачі, необхідно:

- 1) скласти рівняння Ейлера (5.83) для функціонала (5.68);
- 2) розв'язати отримане рівняння Ейлера і знайти екстремаль  $y(x)$ ;
- 3) знайти точки перетину  $x_1, x_2$  екстремалі  $y(x)$  з функцією обмеження  $\varphi(x)$ , скориставшись рівнянням (5.84);
- 4) скласти функцію  $y^*(x)$ , яка задає загальний розв'язок задачі пошуку екстремалі функціонала з обмеженням у вигляді нерівності, за виразом (5.70).

Для конкретизації загального розв'язку  $y^*(x)$  задачі потрібно знайти дві сталі  $C_1, C_2$ , які з'являються в екстремалі  $y(x)$  в результаті подвійного інтегрування рівняння Ейлера. Зрозуміло, що якщо екстремаль  $y(x)$  в інтервалі  $(x_1, x_2)$  не має точки злому, то для визначення сталих  $C_1, C_2$  буде достатньо двох рівнянь граничних умов, тобто

$$\begin{cases} y(a) = y_a, \\ y(b) = y_b. \end{cases} \quad (5.85)$$

Якщо ж в інтервалі  $(x_1, x_2)$  екстремаль  $y(x)$  має точку злому, то тоді її сталі  $C_1, C_2$  на початковій  $[a, x_1]$  і кінцевій  $[x_2, b]$  ділянках матимуть різні значення, тобто фактично треба знаходити не дві сталі  $C_1, C_2$ , а чотири –  $C_1, C_2, C_3, C_4$ .

Тож для загального випадку, крім рівнянь (5.85), потрібно мати ще два додаткових рівняння, оскільки лише за допомогою чотирьох рівнянь можна визначити всі чотири сталі інтегрування, які увійдуть у функцію  $y^*(x)$ .

Для виведення цих додаткових рівнянь припустимо спочатку, що

$$x_2 = b, \quad (5.86)$$

тобто що функція  $y^*(x)$ , яка доставляє екстремум функціоналу (5.68) за умови (5.69), має лише дві верхні складові у виразі (5.70).

В цьому випадку функціонал (5.68) можна переписати так:

$$J = \int_a^{x_1} F(x, y, y') dx + \int_{x_1}^b F(x, \varphi(x), \varphi'(x)) dx. \quad (5.87)$$

Оскільки в точці  $x_1$  маємо точку переходу від  $y(x)$  до  $\varphi(x)$ , то в разі надання варіації цій точці, тобто переходу від  $x_1$  до  $x_1 + \delta x_1$ , матимемо і варіації складових функціоналів, що входять у вираз (5.87).

Для першого з цих функціоналів функція  $y(x)$  є екстремаллю, а функція  $\varphi(x)$  є рівнянням кривої, по якій «ковзає» кінець екстремалі. Для другого з цих функціоналів навпаки – екстремаллю є  $\varphi(x)$ , а  $y(x)$  – це рівняння кривої, по якій «ковзає» кінець екстремалі.

Тож справедливою буде рівність, що випливає з умов трансверсальності,

$$\left( F(x, y, y') - (y' - \varphi') \cdot F_{y'} \right) \Big|_{x_1} = \left( F(x, \varphi, \varphi') - (\varphi' - y') \cdot F_{\varphi'} \right) \Big|_{x_1}, \quad (5.88)$$

з якої, в свою чергу, випливає, що

$$\begin{cases} F(x_1, y(x_1), y'(x_1)) = F(x_1, \varphi(x_1), \varphi'(x_1)), \\ \left( y'(x_1) - \varphi'(x_1) \right) \cdot F_{y'} \Big|_{x_1} = \left( \varphi'(x_1) - y'(x_1) \right) \cdot F_{\varphi'} \Big|_{x_1}. \end{cases} \quad (5.89)$$

Обидва рівняння системи (5.89) можуть задовольнятися лише за умови, що

$$\begin{cases} y(x_1) = \varphi(x_1), \\ y'(x_1) = \varphi'(x_1). \end{cases} \quad (5.90)$$

Проробивши аналогічні викладки для точки  $x_2$ , можна отримати аналогічні рівняння

$$\begin{cases} y(x_2) = \varphi(x_2), \\ y'(x_2) = \varphi'(x_2). \end{cases} \quad (5.91)$$

Цих шести рівнянь, що задаються виразами (5.85), (5.90), (5.91), достатньо для того, щоб знайти і граничні точки  $x_1$ ,  $x_2$ , і сталі інтегрування  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ .

## 5.5 Прямі методи пошуку екстремалей функціоналів в гільбертових просторах

*Ми вже знаємо, як знаходити екстремалі функціоналів шляхом розв'язання рівнянь Ейлера, Ейлера–Лагранжа чи Ейлера–Пуассона.*

*Але поряд з цими методами існує ще один клас методів визначення екстремалей функціоналів, за допомогою яких ця процедура здійснюється прямою мінімізацією функціонала за умови, що екстремаль задається частковою сумою  $S_n(t)$  (2.45), в якій як функції  $\varphi_k(t)$  вибираються члени ортонормованої послідовності з числа тих, що розглянуті в підрозділі 2.3.*

*Методи цього класу називають прямими методами пошуку екстремалей функціоналів або наближеними.*

Одним із найбільш популярних методів в цьому класі є *метод Рітца*, запропонований на початку ХХ століття.

*Суть методу Рітца така.*

Нехай потрібно знайти екстремаль  $y(x)$ , яка мінімізує функціонал

$$J(y) = \int_a^b F(x, y, y') dx \quad (5.92)$$

за умови, що на границях області  $[a, b]$  маємо

$$\begin{cases} y(a) = y_a, \\ y(b) = y_b \end{cases} \quad (5.93)$$

і мають місце обмеження вигляду

$$g(x, y, y') = 0 \quad (5.94)$$

або

$$\int_a^b K(x, y, y') dx = J_{1_0}. \quad (5.95)$$

Будемо шукати екстремаль  $y(x)$  у вигляді

$$y(x) = \sum_{k=1}^n C_k \cdot \varphi_k(x), \quad (5.96)$$

де  $\varphi_k(x)$  – відомі ортонормовані поліноми з числа тих, що розглянуті в другому розділі, наприклад, поліноми Лагерра, Лежандра чи Чебишова, які ми свідомо вибрали та детально розглянули у підрозділі 2.3, виходячи зі зручності їх використання для розв'язання нашої задачі.

Вибором виразу (5.96) ми фактично зводимо задачу пошуку екстремалей функціонала до задачі визначення коефіцієнтів  $C_k, k = \overline{1, n}$ . Підставляючи вираз (5.96) у рівняння граничних умов (5.93), ми на першому етапі розв'язання задачі скорочуємо на два кількість невідомих коефіцієнтів  $C_k, k = \overline{1, n}$ , які нам потрібно знайти, щоб однозначно визначити екстремаль (5.96), котра доставляє мінімум функціоналу (5.92).

Це нам вдається зробити на першому етапі тому, що за допомогою двох рівнянь (5.93) можна два коефіцієнти, наприклад  $C_1$  і  $C_2$ , виразити через інші коефіцієнти  $C_k, k = \overline{3, n}$ .

Після цього шукана екстремаль (5.96) вже матиме вигляд

$$y = f_1(x, C_3, C_4, \dots, C_n). \quad (5.97)$$

Підставляючи вираз (5.97) у (5.94) чи у (5.95), ми отримаємо на другому етапі рівняння, за допомогою якого можна вилучити ще один коефіцієнт  $C_k$ , наприклад  $C_3$ , виразивши його через інші коефіцієнти  $C_k, k = \overline{4, n}$ .

Після цього шукана екстремаль (5.97) вже матиме вигляд

$$y = f_2(x, C_4, \dots, C_n). \quad (5.98)$$



Підставляючи вираз (5.98) у функціонал (5.92), обраховуючи функцію  $F(x, y, y')$  і беручи інтеграл, ми на третьому етапі отримаємо функцію

$$J(y) = f_3(C_4, \dots, C_n), \quad (5.99)$$

яка вже не містить в собі змінної  $x$  і буде функцією виключно лише коефіцієнтів  $C_k, k = \overline{4, n}$ .

Для пошуку оптимальних значень цих коефіцієнтів використовуємо на четвертому етапі стандартну методику пошуку екстремуму функції  $(n - 3)$  змінних  $C_k, k = \overline{4, n}$ .

Для цього частинні похідні  $\frac{\partial J(y)}{\partial C_k}, k = \overline{4, n}$  від (5.99), прирівняємо до нуля

$$\frac{\partial f_3(C_4, \dots, C_n)}{\partial C_k} = 0, \quad k = \overline{4, n}, \quad (5.100)$$

і розв'яжемо отриману систему алгебраїчних рівнянь (5.100) відносно  $C_k, k = \overline{4, n}$ .

Знайдені значення коефіцієнтів  $C_k, k = \overline{4, n}$  підставляємо у функцію (5.98). Це і буде екстремаль  $y(x)$ , яка в умовах дії обмежень (5.94) чи (5.95) та граничних умов (5.93) доставляє мінімум функціоналу (5.92).

На цьому ми завершуємо розгляд методів пошуку екстремалей функціоналів класу, окресленого в підрозділі 4.1, і перейдемо до аналізу властивостей та областей застосування операторів, які теж відносять до основних об'єктів дослідження в курсі функціонального аналізу.

## 5.6 Завдання для самоперевірки

1. В чому відмінність пошуку умовного екстремуму від безумовного?
2. Сформулюйте задачу Лагранжа і алгоритм її розв'язання з використанням невизначених множників.
3. Доведіть, що задачу мінімізації функціонала за наявності обмежень можна трансформувати у найпростішу задачу варіаційного числення відносно функціонала, в якій ці обмеження введені з використанням невизначених множників Лагранжа.
4. Запишіть алгоритм Лагранжа для функціонала, що залежить від кількох функцій та їх похідних, в умовах обмежень, визначених одним рівнянням, а також системою рівнянь.
5. Як формулюється ізопериметрична задача оптимізації і чому вона називається ізопериметричною?
6. Побудуйте алгоритм розв'язання ізопериметричної задачі оптимізації.
7. Як відрізнити мінімум функціонала від максимуму в задачі на умовний екстремум?
8. Сформулюйте задачу Майєра. В чому її особливість?
9. Як звести задачу Майєра до задачі Лагранжа?
10. Побудуйте алгоритм розв'язання задачі Майєра з використанням невизначених множників Лагранжа.
11. В чому суть принципу взаємності в ізопериметричній задачі?
12. Що собою являє екстремаль функціонала з обмеженнями у вигляді нерівностей?

13. Як перевести задачу на умовний екстремум функціонала при обмеженнях у вигляді нерівностей в задачу на його безумовний екстремум?
14. Як скласти рівняння для визначення точок припасування функцій і сталих інтегрування в задачі пошуку екстремуму функціонала з обмеженнями у вигляді нерівностей? Скільки потрібно таких рівнянь і чому?
15. В чому суть прямих методів пошуку екстремумів функціоналів? Суть методу Рітца.
16. В якому вигляді відшуковується екстремаль функціонала в методі Рітца? Які ортонормовані послідовності використовуються в цьому методі?
17. Розкрийте суть етапів розв'язання задачі пошуку екстремуму функціонала методом Рітца.

## Розділ 6

# ОПЕРАТОРИ ТА ЇХ ОСНОВНІ ХАРАКТЕРИСТИКИ

### 6.1 Оператор, його лінійність, норма та ряд інших характеристик

В четвертому і п'ятому розділах цього навчального посібника ми досліджували властивості функціоналів, що задають закони, за якими кожному елементу з множини функцій ставиться у відповідність якийсь елемент із множини чисел.

А тепер перейдемо до дослідження операторів, що задають закони, за якими кожному елементу з множини функцій ставиться у відповідність якийсь елемент із іншої або цієї ж самої множини функцій.

І розпочнемо вивчення цього розділу функціонального аналізу зі встановлення деяких найбільш вживаних характеристик та властивостей операторів.

**Оператор**  $A$ , який здійснює перетворення множини  $X$  у множину  $Y$ , **називається лінійним, якщо він, по-перше, є адитивним**, тобто для нього виконується співвідношення

$$A \cdot (x_1 + x_2) = A \cdot x_1 + A \cdot x_2, \quad \forall x_1, x_2 \in X, \quad (6.1)$$

**а по-друге, є неперервним**, тобто для нього виконується співвідношення

$$A \cdot x_n \rightarrow A \cdot x_0, \quad (6.2)$$

якщо  $x_n \rightarrow x_0, \quad \forall x_0, x_n \in X$ .

Для лінійного оператора  $A$  справедливо:

$$1) \quad A \cdot 0 = 0, \quad 0 \in X; \quad (6.3)$$

$$2) \quad A(-x) = -A \cdot x, \quad \forall x \in X; \quad (6.4)$$

$$3) \quad A(t \cdot x) = t \cdot A \cdot x, \quad \forall x \in X, t - \text{скаляр}. \quad (6.5)$$

**Якщо для оператора виконується співвідношення (6.5), то його називають однорідним.**

Однією з найважливіших властивостей оператора є його обмеженість на одиничній кулі  $S[0, 1] \subset X$ , для якої виконується

$$\|x\| \leq 1, \quad x \in S. \quad (6.6)$$

Обмеженість лінійного оператора  $A$  задають у такий спосіб.

Нехай існує така константа  $K$ , що

$$\|A \cdot x\| \leq K \cdot \|x\|, \quad \forall x \in X. \quad (6.7)$$

**Число**

$$K_0 = \sup_{\|x\| \leq 1} \|A \cdot x\| \quad (6.8)$$

*називають нормою оператора  $A$  і позначають*

$$K_0 = \|A\|. \quad (6.9)$$

Нагадаємо, що символ «Sup» означає «верхню границю» виразу, який стоїть з правого боку від цього символу. Під символом вказують множину елементів, на які діє оператор  $A$ .

З нерівності трикутника для норми випливає, що для  $\forall x \in S$  справедливо

$$\|A \cdot x\| \leq \|A\| \cdot \|x\|. \quad (6.10)$$

А тому можна стверджувати, що **множина лінійних операторів  $A$ , які перетворюють множину  $X$  у множину  $Y$** , що символічно позначається як

$$A: (X \rightarrow Y) \quad (6.11)$$

або

$$A \in (X \rightarrow Y), \quad (6.12)$$

**є лінійним нормованим простором  $(X \rightarrow Y)$ .**

Із співвідношень (6.7) ÷ (6.10) випливає, що для норми оператора виконується

$$\|A\| = \sup_{x \neq 0} \frac{\|A \cdot x\|}{\|x\|}. \quad (6.13)$$

**Дуже важлива теорема: якщо адитивний і однорідний оператор є обмеженим, то він є неперервним, а тому лінійним.**

Доведення.

Нехай  $x_n \rightarrow x_0$ ,  $\{x_n\} \subset X$ ,  $x_0 \in X$ .

Тоді, в силу адитивності та обмеженості,

$$\|A \cdot x_n - A \cdot x_0\| = \|A \cdot (x_n - x_0)\| \leq K \cdot \|x_n - x_0\|. \quad (6.14)$$

Але з того, що  $x_n \rightarrow x_0$ , випливає, що

$$\|x_n - x_0\| = 0. \quad (6.15)$$

Підставляючи вираз (6.15) у (6.14) отримаємо, що

$$\|A \cdot x_n - A \cdot x_0\| = 0. \quad (6.16)$$

Співвідношення (6.16) – це записана через норму властивість (6.2), яка визначає неперервність оператора.

Але, якщо оператор  $A$  є неперервним і адитивним, то він лінійний за означенням.

Часто математики оперують поняттям цілком неперервного оператора, задаючи його так: *лінійний оператор  $A$ , який перетворює деякий банаховий простір  $X$  у банаховий простір  $Y$ , називається цілком неперервним, якщо він перетворює будь-яку обмежену множину  $G \subset X$  у компакту множину  $M \subset Y$ .*

Слід зауважити, що це формулювання теж відповідає співвідношенню (6.2) і є лише іншим словесно-понятійним трактуванням введеного раніше поняття неперервності.

Але для гільбертового простору поняття цілком неперервного оператора можна ввести і з використанням скалярного добутку. Тобто в  $H$ -просторі це такий лінійний оператор  $A$  із множини  $(X \rightarrow Y)$ , для якого з того, що  $x_n \rightarrow x_0$ ,  $y_n \rightarrow y_0$ , випливає, що

$$\langle A \cdot x_n, y_n \rangle \rightarrow \langle A \cdot x_0, y_0 \rangle. \quad (6.17)$$

Завдяки наявності скалярного добутку в  $H$ -просторі виділяють спеціальний клас лінійних операторів, які називають симетризувальними. При цьому *оператор  $A \in (H \rightarrow H)$  називається симетризувальним, якщо для  $\forall x, y \in H$*

$$\langle A \cdot x, y \rangle = \langle x, A \cdot y \rangle. \quad (6.18)$$

Розглянемо  $\langle Ax, y \rangle$  за умови, що  $x$  є змінним елементом, а  $y$  – фіксованим елементом у  $H$ -просторі, тобто покладемо, що

$$\{x\} \subset H, \quad y \in H. \quad (6.19)$$

За умови (6.19) скалярний добуток  $\langle Ax, y \rangle$  є лінійним функціоналом від  $x$ , тобто маємо

$$F_y(x) = \langle A \cdot x, y \rangle. \quad (6.20)$$

Змінюючи елементи  $y \in H$ , ми будемо отримувати різні значення функціонала (6.20), які формально можна позначити символом  $y^*$ , тобто

$$y^* = F_y(x), \quad (6.21)$$

де  $y^* \in H$ .

Але, з іншого боку, формально можна записати, що

$$y^* = A^* y, \quad (6.22)$$

де  $A^*$  – це лінійний оператор, який перетворює елементи  $y \in H$  у елементи  $y^* \in H$ . Цей оператор  $A^*$  називають спряженим до оператора  $A$ .

Оскільки для елементів  $x$  і  $y^*$  теж можна визначити скалярний добуток як  $\langle x, y^* \rangle$ , то справедливим буде і співвідношення

$$\langle x, y^* \rangle = \langle x, A^* y \rangle, \quad (6.23)$$

яке отримуємо підстановкою у вираз  $\langle x, y^* \rangle$  співвідношення (6.22).

Але, якщо  $A$  є оператором симетризувальним, то з зіставлення виразів (6.18) і (6.23) випливає, що

$$A = A^*, \quad (6.24)$$

тобто, симетризувальний оператор  $A$  збігається зі своїм спряженим  $A^*$ . В зв'язку з цим симетризувальні оператори ще називають самоспряженими.

*Оператор  $A$ , для якого*

$$A \cdot A^* = A^* \cdot A = I, \quad (6.25)$$

де  $I$  – одиничний оператор, тобто такий, що

$$I \cdot x = x, \quad (6.26)$$

*називають унітарним.*

## 6.2 Обернений оператор та резольвента і спектр оператора

Нехай задано оператор  $A \in (X \rightarrow Y)$ .

*Оператор  $A^{-1} \in (Y \rightarrow X)$ , який задовольняє рівняння*

$$\begin{cases} A^{-1} \cdot (A \cdot x) = x, \quad \forall x \in X, \\ A \cdot (A^{-1} \cdot y) = y, \quad \forall y \in Y. \end{cases} \quad (6.27)$$

*називають оберненим до оператора  $A$ .* Зрозуміло, що рівняння (6.27) можуть задовольнятись лише за умови, що

$$A^{-1} \cdot A = A \cdot A^{-1} = I. \quad (6.28)$$

*Зауважимо, що під добутком будь-яких операторів ми розуміємо їх послідовне застосування до елемента, що стоїть з правого боку від них.*

Після введення оберненого оператора  $A^{-1}$  зрозуміло, що для розв'язання рівняння

$$A \cdot x = y \quad (6.29)$$

необхідно знайти  $A^{-1}$ , оскільки

$$\begin{aligned} A^{-1}(A \cdot x) &= A^{-1}y \\ \Downarrow \\ (A^{-1}A) \cdot x &= A^{-1}y \\ \Downarrow \\ I \cdot x &= A^{-1}y \\ \Downarrow \\ x &= A^{-1}y. \end{aligned} \quad (6.30)$$

Виникає питання: «А як же розв'язати рівняння

$$A \cdot x - \lambda \cdot x = y, \quad (6.31)$$

$$A \cdot x - \lambda \cdot x = 0, \quad (6.32)$$

де  $x \in X$ ,  $y \in Y$ ,  $A \in (X \rightarrow Y)$ , а  $\lambda$  – скаляр?»

Перепишемо (6.31), (6.32) так

$$(A - \lambda \cdot I) \cdot x = y, \quad (6.33)$$

$$(A - \lambda \cdot I) \cdot x = 0. \quad (6.34)$$

Нехай існує оператор  $R_\lambda$ , обернений до  $(A - \lambda \cdot I)$ , тобто

$$R_\lambda = (A - \lambda \cdot I)^{-1}. \quad (6.35)$$

В цьому випадку, перемножуючи рівняння (6.33) зліва на  $R_\lambda$ , отримаємо

$$x = (A - \lambda \cdot I)^{-1} \cdot y. \quad (6.36)$$

Вираз (6.36) і буде розв'язком рівняння (6.33).

**Обернений оператор  $R_\lambda$ , який визначається співвідношенням (6.35), називають резольвентою оператора  $A$  або розв'язувальним оператором для рівняння (6.33).**

Зрозуміло, що розв'язок рівняння (6.34) буде мати вигляд

$$x = (A - \lambda \cdot I)^{-1} \cdot 0 = 0. \quad (6.37)$$

*Ті значення параметра  $\lambda$ , які дозволяють мати резольвенту  $R_\lambda$ , називають регулярними значеннями оператора  $A$ . Всі інші значення параметра  $\lambda$ , які не є регулярними, складають спектр оператора  $A$ .*

*Як виявили дослідження математиків, існують такі значення  $\lambda$ , за яких однорідне рівняння (6.32) має розв'язок, відмінний від нульового, тобто для них не виконується (6.37). Такі значення  $\lambda$  називають характеристичними числами або власними значеннями оператора  $A$ . Зрозуміло, що вони є точками спектра цього оператора.*

Але слід зауважити, що до складу спектра оператора  $A$  можуть входити значення  $\lambda$ , які не є його характеристичними числами, тобто потужність спектра оператора є більшою потужності множини його характеристичних чисел.

### 6.3 Критерії існування однозначного та коректного розв'язків операторних рівнянь

Для пошуку розв'язків багатьох неоднорідних операторних рівнянь класу (6.31) корисною є теорема: нехай  $A \in (X \rightarrow X)$ , де  $X \in \mathbb{B}$ -простір. Нехай також існує деякий параметр  $\mu$ , який задовольняє умову

$$|\mu| < \frac{1}{\|A\|}. \quad (6.38)$$

За цих умов оператор  $(I - \mu A)$  має обернений оператор  $R_\mu$  і при цьому

$$R_\mu = (I - \mu \cdot A)^{-1} = I + \mu \cdot A + \mu^2 \cdot A^2 + \mu^3 \cdot A^3 + \dots \quad (6.39)$$

Зрозуміло, що  $R_\mu$  є теж резольвентою оператора  $A$ , унормованого до  $\lambda$ .

Доводиться (6.39) розкладенням  $R_\mu$  у степеневий ряд в околі точки  $\mu = 0$ .

**Наведемо деякі властивості резольвенти.**

1. Для множини  $D$  регулярних точок оператора  $A$

$$R_\lambda - R_\mu = (\lambda - \mu) \cdot R_\lambda \cdot R_\mu, \quad \forall \lambda, \mu \in D. \quad (6.40)$$

З рівняння (6.40) шляхом граничного переходу  $\lambda \rightarrow \mu$  отримаємо



$$\frac{d^k R_\lambda}{d\lambda^k} = k! \cdot R_\lambda^{k+1}. \quad (6.41)$$

2. Якщо оператор  $A$  є обмеженим, то весь його спектр лежить у крузі

$$\|\lambda\| \leq \|A\|, \quad (6.42)$$

а за межами цього круга, тобто при

$$\|\lambda\| > \|A\| \quad (6.43)$$

резольвенту  $R_\lambda$  можна розкласти у збіжний за нормою оператора ряд

$$R_\lambda = -\frac{1}{\lambda} \cdot \left( I + \frac{1}{\lambda} \cdot A + \frac{1}{\lambda^2} \cdot A^2 + \dots \right). \quad (6.44)$$

3. Радіус найменшого круга  $r_A$  з центром у початку координат, який містить весь спектр оператора  $A$ , називається **спектральним радіусом** і його можна визначити за формулою Гельфанда

$$r_A = \lim_{n \rightarrow \infty} \sqrt[n]{\|A^n\|}. \quad (6.45)$$

З виразу (6.45) випливає, що

$$r_A \leq \|A\|. \quad (6.46)$$

4. Нехай  $\lambda$  є спільною регулярною точкою двох замкнених лінійних операторів  $A$  і  $B$ . Якщо  $D_B \supset D_A$ , то

$$R_\lambda^B - R_\lambda^A = R_\lambda^B \cdot (A - B) \cdot R_\lambda^A. \quad (6.47)$$

Повернемось до рівняння (6.31) або, що одне і те ж, до рівняння (6.33). Стосовно нього можна зробити два твердження.

1. Рівняння (6.31) називають **таким, що має однозначний розв'язок, якщо відповідне йому однорідне рівняння**

$$A \cdot x = 0 \quad (6.48)$$

**має лише нульовий розв'язок.**

2. Рівняння (6.31) називають таким, що має коректний розв'язок, якщо для  $\forall x \in X$  виконуються співвідношення

$$\|x\| \leq K \cdot \|A \cdot x\|, \quad A \in (X \rightarrow X). \quad (6.49)$$

З першого твердження випливає, що для оператора  $A \in D_A$  існує лівий обернений оператор  $A^{-1}$ .

З другого твердження випливає, що оператор  $A^{-1}$  є обмеженим, а тому розв'язок рівняння (6.31) неперервно залежить від правої частини.

#### 6.4 Група операторів

Введемо поняття групи операторів.

*Під групою  $u(t)$  операторів  $A$  розуміють таку їх сукупність, якій притаманні властивості:*

$$1) \quad u(0) = I; \quad (6.50)$$

$$2) \quad \begin{aligned} u(t + \tau) &= u(t) \cdot u(\tau), \\ -\infty < t, \tau < \infty; \end{aligned} \quad (6.51)$$

$$3) \quad \frac{du}{dt} = A \cdot u. \quad (6.52)$$

Зрозуміло, що ці три властивості мають лише ті оператори  $A$ , які є обмеженими і задовольняють співвідношення

$$u(t) = e^{t \cdot A}. \quad (6.53)$$

З виразів (6.50), (6.52) та (6.53) випливає, що оператор  $A$  можна визначити як похідну від групи  $u(t)$  при  $t = 0$ .

Тому оператор  $A$  називають *породним оператором для групи  $u(t)$  або, що одне і те ж, інфінітезимальним*.

Розкладаючи експоненту у виразі (6.53) у степеневий ряд, отримаємо

$$u(t) = 1 + t \cdot A + \frac{t^2}{2!} \cdot A^2 + \frac{t^3}{3!} \cdot A^3 + \dots, \quad (6.54)$$

тобто група  $u(t)$  може бути заданою не лише через експоненту (6.53), а й у вигляді степеневого ряду (6.54).

## 6.5 Метод стиснених відображень

Цей метод є одним із ключових для багатьох застосунків функціонального аналізу в прикладних задачах ІТ-сфери, але, перш ніж викласти його суть, *наведемо теорему про єдину спільну точку* послідовності вкладених одна в одну замкнутих куль, яка базується на матеріалі, викладеному в підрозділі 1.2 з використанням виразів (1.41) та (1.42) і формулюється так: *послідовність вкладених одна в одну замкнених куль  $K_1 \supset K_2 \supset K_3 \supset \dots$ , які є підмножинами повного метричного простору  $R$ , таких що*

$$K_n = \{x_n, \rho(x, x_n) \leq r_n\}, \quad n = 1, 2, 3, \dots, \quad (6.55)$$

*і радіуси  $r_n$  яких наближаються до нуля при  $n \rightarrow \infty$ , мають єдину спільну точку  $x_0$ .*

Доведення цієї теореми базується на тому, що для послідовності (6.55), яким би ми малим не вибрали число  $\varepsilon > 0$ , завжди можна побудувати замкнуту кулю з радіусом  $r_k < \varepsilon$ , для якої справедливим буде вираз

$$\rho(x, x_k) \leq r_k, \quad (6.56)$$

з якого в силу повноти простору  $R$  витікає, що, які б дві точки цього простору ми не взяли, завжди можна буде побудувати кулю з радіусом, меншим відстані між цими точками, а тому послідовність  $\{x_n\}$  є фундаментальною, для якої справедливо

$$\lim_{n \rightarrow \infty} x_n = x_0. \quad (6.57)$$

А з виразів (6.55), (6.56), (6.57) і витікає, що з вкладенням кулі в кулю, радіус кожної з наступних із яких є усе меншим, ми прийдемо до єдиної спільної для усіх цих куль точки  $x_0$  простору  $R$ , в якому ця послідовність куль формується.

А тепер уявімо, що послідовність таких «куль»  $[a_i, b_i]$ ,  $i = 0, 1, 2, \dots, n$  ми формуємо на відрізку  $[a, b]$  числової осі  $x$  за умови, що

$$\begin{aligned} b &= b_0 > b_1 > b_2 > \dots > b_n > \dots, \\ a &= a_0 < a_1 < a_2 < \dots < a_n < \dots \end{aligned} \quad (6.58)$$

Згідно з вище сформульованою і доведеною теоремою ця послідовність, стискуючись, прямуватиме до єдиної спільної точки на числовій осі  $x$ . Тож, якщо ми визначимо закон  $f(x)$ , за яким при стисненні відрізка  $[a, b]$  числової осі  $x$  кожній точці  $x_i$ ,  $i = 0, 1, 2, \dots, n$  попереднього відрізка  $[a_i, b_i]$ ,  $i = 0, 1, 2, \dots, n$  відповідатиме точка  $x_i^*$ ,  $i = 1, 2, \dots, n$  наступного відрізка, вкладеного в попередній, то знайти цю єдину спільну точку можна, розв'язавши рівняння

$$x = f(x). \quad (6.59)$$

А далі випишемо означення: *відображення, задане в метричному просторі  $R$  оператором  $A$ , називається стисненим, якщо для  $x_1 \in R$  і  $x_2 \in R$  виконується нерівність*

$$\rho(Ax_1, Ax_2) \leq \theta \rho(x_1, x_2), \quad (6.60)$$

де 
$$0 < \theta < 1 \quad (6.61)$$

Далі доцільною для розгляду є *теорема Банаха*, яка, фактично, є розширенням на оператори вище доведеної теореми про єдину спільну точку послідовності вкладених одна в одну куль, заданих у повному метричному просторі, і яка формулюється так: *стиснене відображення оператором  $A$  повного метричного простору  $R$  в себе має одну і тільки одну нерухому точку.*

Для доведення цієї теореми покладемо, що оператор  $A$  діє на точку  $x_i \in R, i = 0, 1, 2, \dots, n$  так, що

$$x_1 = Ax_0, \quad x_2 = Ax_1, \dots, x_n = Ax_{n-1}, \dots \quad (6.62)$$

Покладемо також, що

$$\begin{aligned} \rho(x_1, x_2) &= \theta \rho(x_0, x_1), \\ \rho(x_2, x_3) &= \theta \rho(x_1, x_2) = \theta^2 \rho(x_0, x_1), \\ &\dots\dots\dots \\ \rho(x_{n-1}, x_n) &= \theta^{n-1} \rho(x_0, x_1), \end{aligned} \quad (6.63)$$

де  $\theta$  задається виразом (6.61).

Задамо  $q > n$ . Тоді, згідно з виразами (6.60) і (6.62), матимемо

$$\begin{aligned} \rho(x_n, x_q) &= \rho(Ax_{n-1}, Ax_{q-1}) \leq \theta \rho(x_{n-1}, x_{q-1}) = \theta \rho(Ax_{n-2}, Ax_{q-2}) \leq \theta^2 \rho(x_{n-2}, x_{q-2}) = \\ &= \theta^2 \rho(Ax_{n-3}, Ax_{q-3}) \leq \theta^3 \rho(x_{n-3}, x_{q-3}) = \dots = \theta^{n-1} \rho(Ax_0, Ax_{q-n}) \leq \theta^n \rho(x_0, x_{q-n}) \end{aligned} \quad (6.64)$$

На підставі виразу (1.35) для метрики  $\rho(x_0, x_{q-n})$  можна записати, що

$$\rho(x_0, x_{q-n}) = \rho(x_0, x_1) + \rho(x_1, x_2) + \dots + \rho(x_{q-(n-1)}, x_{q-n}) \quad (6.65)$$

Залишимо в нерівності (6.64) лише перший і останній члени, що лише підсилить її. В результаті такого кроку отримаємо, що

$$\rho(x_n, x_q) \leq \theta^n \rho(x_0, x_{q-n}). \quad (6.66)$$

Нехай

$$q, n \rightarrow \infty. \quad (6.67)$$

Підставивши вирази (6.63) у (6.65), а результат цієї підстановки у вираз (6.66), з урахуванням умови (6.13) отримаємо

$$\rho(x_n, x_q) \leq \theta^n \rho(x_0, x_1) (1 + \theta + \theta^2 + \dots + \theta^{q-n-1} + \dots) = \frac{\theta^n}{1-\theta} \rho(x_0, x_1) \quad (6.68)$$

Зауважимо, що при отриманні виразу (6.68) ми використали формулу для суми членів нескінченно спадної геометричної прогресії зі знаменником  $\theta$ .

Легко бачити, що зі зростанням  $n$  множник  $\theta^n$  у правій частині нерівності (6.68) внаслідок виконання умови (6.61) буде наближатись до нуля, що, в свою чергу, свідчить про те, що зі зростанням  $n, q$  і ліва частина цієї нерівності наближатиметься до нуля, тобто, що послідовність  $\{x_n\}$  є фундаментальною, а тому для загального члена

$$x_n = Ax_{n-1} \quad (6.69)$$

послідовності (6.62) при  $n \rightarrow \infty$  справедливою є рівність

$$x = Ax, \quad (6.70)$$

яка, згідно з виразом (6.59) теореми про існування нерухомої точки, підтверджує ту частину теореми Банаха, в якій стверджується, що стиснене відображення, яке здійснюється оператором  $A$  в повному метричному просторі  $\mathbf{R}$ , має нерухому точку  $x$ . І цілком очевидно, що ця нерухома точка є розв'язком операторного рівняння (6.70).

Залишилось довести ту частину теореми Банаха, в якій стверджується, що ця нерухома точка  $x$  є єдиною для оператора  $A$ .

Це доведення здійснимо від протилежного, тобто припустимо, що відображення, яке здійснюється оператором  $A$  у метричному просторі  $\mathbf{R}$ , має дві нерухомі точки  $x$  та  $x^*$ . Тоді для кожної з них буде справедливою рівність (6.70), тобто тоді і для точки  $x^*$  матимемо

$$x^* = Ax^*. \quad (6.71)$$

Але тоді, згідно з означенням поняття стисненого відображення та виразів (6.60), (6.70), (6.71), можна записати, що

$$\rho(x, x^*) = \rho(Ax, Ax^*) \leq \theta \rho(x, x^*). \quad (6.72)$$

Вираз (6.72) може виконуватись лише за умови, що

$$\theta = 1, \quad (6.73)$$

тобто, що подальшого стиснення цим оператором уже здійснити не можна. А це у свою чергу свідчить про те, що

$$x = x^*, \quad (6.74)$$

тобто, що ці точки займають одне і те ж місце на осі  $x$ , – це і є підтвердженням тієї частини теореми Банаха, в якій стверджується, що нерухома точка  $x$  є єдиною для оператора  $A$ .

Як ми покажемо у наступному розділі, присвяченому прикладним аспектам функціонального аналізу, з використанням методу стиснених відображень та вище доведеної теореми Банаха легко доводиться існування єдиного розв'язку таких класів операторних рівнянь, якими є алгебраїчні і диференціальні рівняння, та легко формуються алгоритми розв'язання такого класу операторних рівнянь, якими є інтегральні рівняння.

## 6.6 Завдання для самоперевірки

1. Дайте означення оператора. Наведіть приклади.
2. Який оператор є лінійним? Які властивості лінійного оператора ви знаєте?
3. Що таке норма оператора? Запишіть нерівність трикутника для норми оператора.
4. Що таке «цілком неперервний оператор»? Як можна визначити цілком неперервний оператор в гільбертовому просторі?
5. Який лінійний оператор є симетризувальним?
6. Що собою являє спряжений і самоспряжений оператори?
7. Який оператор називають унітарним?
8. Як визначити обернений оператор?
9. Що є резольвентою неоднорідного операторного рівняння?
10. Які значення оператора є регулярними, а які складають його спектр?
11. Що собою являють характеристичні числа або власні значення оператора?
12. Чи входять характеристичні числа оператора до множини його регулярних значень?
13. Як можна виразити резольвенту через ряд за степенями оператора?
14. Які властивості резольвенти оператора ви знаєте?
15. За яких умов неоднорідне операторне рівняння має коректний розв'язок?
16. Що розуміють під групою операторів, і які властивості такої групи ви знаєте?
17. В чому суть ідеї метода стиснених відображень оператора?
18. Сформулюйте і доведіть теорему Банаха про єдину нерухому точку стисненого відображення оператора.

## Розділ 7

### ПРИКЛАДНІ АСПЕКТИ ТЕОРІЇ ОПЕРАТОРІВ

#### 7.1 Застосування метода стиснених відображень в теорії алгебраїчних рівнянь

Розглянемо алгебраїчне рівняння

$$F(x) = 0, \quad (7.1)$$

задане на відрізку  $[a, b]$ , наприклад, рівняння

$$x^2 - 5x + 4 = 0 \quad (7.2)$$

задане на відрізку  $[0, 2]$ .

Приведемо рівняння (7.1) до вигляду (6.59), тобто, до вигляду

$$x = f(x), \quad (7.3)$$

яким для виразу (7.2) буде рівняння

$$x = \frac{1}{5}x^2 + \frac{4}{5}. \quad (7.4)$$

Накладемо на функцію  $f(x)$  умови, щоб вона на відрізку  $[a, b]$  була неперервною, диференційовною та щоб її значеннями були числа з цього ж відрізка, тобто, щоб

$$\forall x \in [a, b] \Rightarrow f(x) \in [a, b] \quad (7.5)$$

Для прикладу (7.4) умова (7.5) набуває вигляду

$$\forall x \in [0, 2] \Rightarrow \left( \frac{1}{5}x^2 + \frac{4}{5} \right) \in [0, 2] \quad (7.6)$$

Підстановкою значень аргументу  $x$  із відрізка  $[0, 2]$  у вираз (7.6) легко переконатись у тому, що умова (7.5) для функції (7.4) нашого прикладу виконується.

Накладемо на функцію  $f(x)$  ще одну умову, щоб її похідна  $f'(x)$  в усіх точках відрізка  $[a, b]$  за модулем була меншою одиниці, тобто, щоб

$$\forall x \in [a, b] \Rightarrow |f'(x)| \leq k < 1 \quad (7.7)$$

де

$$k = \max_{x \in [a, b]} |f'(x)|. \quad (7.8)$$

Для прикладу (7.4) матимемо:

$$f'(x) = \frac{d}{dx} \left( \frac{1}{5}x^2 + \frac{4}{5} \right) = \frac{2}{5}x, \quad (7.9)$$

$$k = \frac{\max_{x \in [0,2]} \left| \frac{2}{5}x \right|}{1} = \frac{4}{5} < 1 \quad (7.10)$$

Як бачимо з виразів (7.9), (7.10), умова (7.7) для нашого прикладу (7.4) виконується. Надамо відрізьку  $[a, b]$  статус метричного простору з метрикою

$$\rho(x_1, x_2) = |x_1 - x_2| \quad (7.11)$$

і припустимо, що оператор  $A$  здійснює відображення цього простору самого в себе, тобто, що вираз (7.3) набуває вигляду

$$Ax = f(x). \quad (7.12)$$

Доведемо, що відображення (7.12) є стисненим, а тому має єдину нерухому точку, яка і є коренем алгебраїчного рівняння (7.1).

Підставляючи вираз (7.12) у вираз для метрики (7.11), отримаємо

$$\rho(Ax_1, Ax_2) = |f(x_1) - f(x_2)|. \quad (7.13)$$

Якщо ми помножимо і розділимо праву частину виразу (7.13) на метрику  $|x_1 - x_2|$ , то отримаємо

$$\rho(Ax_1, Ax_2) = \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|} |x_1 - x_2| = |f'(\xi)| |x_1 - x_2|, \quad (7.14)$$

де

$$\xi \in [x_1, x_2] \subset [a, b]. \quad (7.15)$$

Для нашого прикладу (7.4) вираз (7.15) матиме вигляд

$$\xi \in [x_1, x_2] \subset [0, 2]. \quad (7.16)$$

З урахуванням (виразів (7.7), (7.8) та (7.15) вираз (7.14) набуває вигляду

$$\rho(Ax_1, Ax_2) = |f'(\xi)| |x_1 - x_2| \leq k |x_1 - x_2| = k\rho(x_1, x_2), \quad (7.17)$$



Оскільки, як показано вище,  $k < 1$  (для нашого прикладу  $k = \frac{4}{5} = 0,8$ , що витікає з виразу (7.10), то вираз (7.17) задає стиснене відображення, яке, згідно з теоремою Банаха, має єдину нерухому точку  $x^*$ .

Використовуючи вирази (6.57), (6.69) та (6.70), можна записати, що

$${}_n \underline{\lim}_{\infty} x_n = {}_n \underline{\lim}_{\infty} Ax_{n-1} = Ax^* = x^*. \quad (7.18)$$

Фактично виразом (7.18) ми задаємо алгоритм руху шляхом послідовних наближень від довільної точки  $x_0$  відрізка  $[a, b]$  до нерухомої точки  $x^*$ .

Для нашого прикладу (7.4) цей алгоритм, виходячи з виразів (7.3), (7.4) та (7.12), матиме вигляд

$${}_n \underline{\lim}_{\infty} x_n = {}_n \underline{\lim}_{\infty} \left( \frac{1}{5} x_{n-1}^2 + \frac{4}{5} \right). \quad (7.19)$$

Точність наближень на кожному кроці алгоритму (7.18) ми можемо оцінити за виразом (6.68), який для нашого випадку набуває вигляду

$$|x_n - x^*| \leq k^n \frac{|x_0 - x_1|}{1 - k}. \quad (7.20)$$

Покажемо, як працює цей алгоритм для нашого прикладу, заданого виразами (7.2)–(7.4).

Нехай  $x_0 = 0,5$ . Тоді, згідно з наведеним вище алгоритмом, матимемо

$$\begin{aligned} x_1 &= \frac{1}{5} x_0^2 + \frac{4}{5} = \frac{1}{5} (0,5)^2 + \frac{4}{5} = 0,85; \\ x_2 &= \frac{1}{5} x_1^2 + \frac{4}{5} = \frac{1}{5} (0,85)^2 + \frac{4}{5} = 0,9445; \\ x_3 &= \frac{1}{5} x_2^2 + \frac{4}{5} = \frac{1}{5} (0,9445)^2 + \frac{4}{5} = 0,9784; \\ x_4 &= \frac{1}{5} x_3^2 + \frac{4}{5} = \frac{1}{5} (0,9784)^2 + \frac{4}{5} = 0,9915; \\ x_5 &= \frac{1}{5} x_4^2 + \frac{4}{5} = \frac{1}{5} (0,9915)^2 + \frac{4}{5} = 0,9966 \end{aligned} \quad (7.21)$$

Легко бачити, що наближення ведуть нас до нерухомої точки  $x^* = 1$ , яка і є коренем рівняння (7.2) на відрізку  $[0, 2]$  числової осі. Але, водночас, слід пам'ятати і про те, що метод стиснених відображень за допомогою оператора  $A$  відображає метричний простір  $[0, 2]$  сам на себе, а тому він не дозволяє за один цикл віднайти і другий корінь рівняння (7.2), яким є число  $x^{**} = 4$ . Для визначення цієї нерухомої точки відображення оператором (7.4) необхідно розширювати межі пошуку так, щоб новий метричний простір  $[c, d]$  мав

границі за межами, визначеними метричному простору  $[a, b]$ , наприклад, переходити до розгляду метричного простору  $[1, 5]$  для рівняння (7.2).

## 7.2 Застосування методу стиснених відображень для доведення існування єдиного розв'язку диференціальних та інтегральних рівнянь

Розглянемо простір  $C$  неперервних функцій  $y = y(x)$ , заданих на відрізку  $x \in [a, b]$  числової осі, зі значеннями, заданими на відрізку  $y \in [M, N]$  числової осі, та метрикою

$$\rho(y, y_1) = \sup_{x \in [a, b]} |y - y_1|, \quad (7.22)$$

де  $y, y_1$  – точки цього простору.

Нехай у просторі  $C$  задане диференціальне рівняння

$$y' = f(x, y) \quad (7.23)$$

з початковою умовою

$$y(x_0) = y_0. \quad (7.24)$$

Накладемо на функцію  $f(x, y)$  умову, щоб вона була такою, що задовольняє умови Ліпшиця, тобто, щоб

$$|f(x, y_0) - f(x, y_1)| \leq L|y_0 - y_1|, \quad (7.25)$$

де  $(x, y_0), (x, y_1)$  – точки прямокутної області  $G$ , обмеженої відрізками  $x \in [a, b]$ ,  $y \in [M, N]$  на площині  $(x, y)$ , а  $L$  – константа, числове значення якої визначимо дещо пізніше.

Проінтегруємо диференціальне рівняння (7.23) в межах від  $x_0$  до  $x$

$$\int_{x_0}^x \frac{dy}{dx} dx = \int_{x_0}^x f(x, y) dx. \quad (7.26)$$

Матимемо

$$y(x) - y(x_0) = \int_{x_0}^x f(x, y) dx, \quad (7.27)$$

або

$$y = y_0 + \int_{x_0}^x f(x, y) dx. \quad (7.28)$$

Очевидно, що інтегральне рівняння (7.28) в операторному вигляді можна переписати і так

$$y = Ay. \quad (7.29)$$

Якщо ми доведемо, що оператор  $A$  здійснює стиснене відображення простору функцій  $C$  в себе, то тим самим ми доведемо, що це відображення задає єдину нерухому точку в цьому просторі, яка у функціональній інтерпретації є розв'язком інтегрального рівняння (7.28), яке, у свою чергу, є іншою формою запису диференціального рівняння (7.23), а тому це доведення буде одночасно і доведенням того, що у заданому просторі як інтегральне рівняння (7.28), так і диференціальне рівняння (7.23) мають розв'язок, який до того ж є єдиним.

Застосовуючи метод стиснених відображень, маємо право записати, що

$$\begin{aligned} \rho(y_1, y_2) &= \rho(Ay_0, Ay_1) = \frac{\sup_{x \in [a, b]} |Ay_0 - Ay_1|}{\sup_{x \in [a, b]} \int |f(x, y_0) - f(x, y_1)| dx} \leq \\ &\leq \frac{\sup_{x \in [a, b]} \int_{x_0}^x L |y_0 - y_1| dx}{\sup_{x \in [a, b]} \int |f(x, y_0) - f(x, y_1)| dx} = L |x - x_0| \frac{\sup_{x \in [a, b]} |y_0 - y_1|}{\sup_{x \in [a, b]} \int |f(x, y_0) - f(x, y_1)| dx} = \theta \rho(y_0, y_1), \end{aligned} \quad (7.30)$$

$$\text{де} \quad \theta = L |x - x_0| \quad (7.31)$$

Згідно з ідеологією методу стиснених відображень для того, щоб вираз (7.30) реалізовував цю ідеологію, потрібно, щоб виконувалась нерівність

$$\theta < 1. \quad (7.32)$$

Порівнюючи вирази (7.31) та (7.32), бачимо, що вираз (7.30) реалізовуватиме ідеологію методу стиснених відображень і спрямовуватиме послідовність

$$y_1 = Ay_0, \quad y_2 = Ay_1, \quad y_3 = Ay_2, \dots, \quad y_n = Ay_{n-1}, \dots \quad (7.33)$$

до єдиної нерухомої точки  $y$  області  $G$ , яка буде проекцією функції  $y = y(x)$  на метричний функціональний простір  $C$  та розв'язком інтегрального рівняння (7.28) у тому випадку, коли константа  $L$  буде задовольняти нерівність

$$|x - x_0| < \frac{1}{L}. \quad (7.34)$$

### 7.3 Метод стиснених відображень у задачі доведення існування розв'язку інтегрального рівняння Фредгольма та побудова його резольвенти

Інтегральне рівняння

$$\varphi(x) = f(x) + \lambda \int_a^b K(x, y) \varphi(y) dy, \quad (7.35)$$

яке зв'язує між собою в просторі неперервних функцій  $C$  неперервні функції  $\varphi(x)$  з нормою

$$\|\varphi(x)\| = \sup_{x \in [a, b]} |\varphi(x)| \quad (7.36)$$

з неперервними в точках площини  $(x, y)$ , обмеженої границями прямокутника  $[a \leq x, y \leq b]$ , функціями  $K(x, y)$  з нормою

$$\|K(x, y)\| = \sup_{[a \leq x, y \leq b]} |K(x, y)|, \quad (7.37)$$

називається рівнянням Фредгольма 2-го роду в честь математика, який його сконструював і вивчав властивості. Його ще називають неоднорідним рівнянням Фредгольма.

Якщо верхню границю інтеграла в рівнянні (7.35) покласти рівною  $x$ , то отримаємо інтегральне рівняння

$$\varphi(x) = f(x) + \lambda \int_a^x K(x, y) \varphi(y) dy, \quad (7.38)$$

яке зв'язує у тому ж просторі неперервних функцій  $C$  неперервні функції  $\varphi(x)$  з нормою (7.36) з тими ж неперервними функціями  $K(x, y)$  з нормою (7.37) і яке називають рівнянням Вольтерра 2-го роду – теж в честь математика, який таку його структуру запропонував і вивчав властивості. Це інтегральне рівняння називають неоднорідним рівнянням Вольтерра.

Якщо в рівняннях (7.35), (7.38) покласти

$$f(x) = 0, \quad (7.39)$$

то ці інтегральні рівняння називають, відповідно, рівняннями Фредгольма 1-го роду та рівняннями Вольтерра 1-го роду або, знову ж таки відповідно, однорідним рівнянням Фредгольма та однорідним рівнянням Вольтерра.

Оператор  $A$ , який в просторі  $C$  перетворює клас функцій  $\varphi$  в самих себе згідно з виразом

$$A\varphi = \lambda \int_a^b K(x, y) \varphi(y) dy, \quad (7.40)$$

називається оператором Фредгольма з ядром  $K(x, y)$ .

З урахуванням виразу (7.40) рівняння Фредгольма (7.35) можна переписати в операторному вигляді

$$\varphi = f + A\varphi, \quad (7.41)$$

а його  $n$ -та ітерація при послідовних наближеннях до розв'язку матиме вигляд

$$\varphi_n = f + A\varphi_{n-1}. \quad (7.42)$$

Очевидно, що оператор Вольтерра формально буде відрізнятися від оператора Фредгольма лише тим, що верхньою границею інтеграла в ньому буде не константа  $b$ , а незалежна змінна  $x$ , але, насправді, за рахунок цього ми матимемо потужність множини функцій, яку формує оператор Вольтерра, більшою потужності множини функцій, яку формує оператор Фредгольма. Такий висновок витікає на підставі матеріалу, викладеного у підрозділі 1.1. Але, оскільки властивості операторів Фредгольма і Вольтерра збігаються, то подальші викладки ми зосередимо на операторі Фредгольма. А цей оператор таких основних властивостей має три: по-перше, він лінійний, по-друге, він неперервний, по-третє, він обмежений, що достатньо і для доведення, що відображення ним у метричному просторі має єдину нерухому точку, яка є розв'язком рівняння Фредгольма, і для побудови резольвентного алгоритму отримання цього розв'язку. Отже,

1) оператор Фредгольма лінійний, тому що

$$\begin{aligned} A(\varphi_1 + \varphi_2) &= \lambda \int_a^b K(x, y)(\varphi_1(y) + \varphi_2(y))dy = \lambda \int_a^b K(x, y)\varphi_1(y)dy + \\ &+ \lambda \int_a^b K(x, y)\varphi_2(y)dy = A\varphi_1 + A\varphi_2 \end{aligned} \quad ; \quad (7.43)$$

2) оператор Фредгольма неперервний, тому що, якщо існує границя  $\varphi$  послідовності  $\{\varphi_n\}$ , тобто,

$$\lim_{n \rightarrow \infty} \varphi_n = \varphi, \quad (7.44)$$

де

$$\varphi_n(x) = f(x) + \lambda \int_a^b K(x, y)\varphi_{n-1}(y)dy = f + A\varphi_{n-1}, \quad (7.45)$$

то

$$\begin{aligned} \lim_{n \rightarrow \infty} \varphi_n &= \lim_{n \rightarrow \infty} \left( f(x) + \lambda \int_a^b K(x, y)\varphi_{n-1}(y)dy \right) = f(x) + \lambda \int_a^b K(x, y) \lim_{n \rightarrow \infty} \varphi_{n-1}(y)dy = \\ &= f(x) + \lambda \int_a^b K(x, y)\varphi(y)dy = f + A\varphi = \varphi \end{aligned} \quad ; \quad (7.46)$$

3) оператор Фредгольма обмежений, тому що

$$\begin{aligned}
\|A\varphi\| &= \frac{\sup}{[a \leq x, y \leq b]} |A\varphi| = \frac{\sup}{[a \leq x, y \leq b]} \left| \lambda \int_a^b K(x, y) \varphi(y) dy \right| \leq \\
&\leq \frac{\sup}{[a \leq x, y \leq b]} \lambda \int_a^b |K(x, y)| |\varphi(y)| dy = \\
&= \lambda \int_a^b \frac{\sup}{[a \leq x, y \leq b]} |K(x, y)| \frac{\sup}{[a \leq y \leq b]} |\varphi(y)| dy = \lambda \|K(x, y)\| \|\varphi(y)\| \int_a^b dy = \\
&= \lambda \|K(x, y)\| \|\varphi(y)\| (b-a)
\end{aligned} \tag{7.47}$$

Покажемо, що оператор Фредгольма здійснює стиснене відображення простору функцій  $\varphi$  самого в себе. Для цього застосуємо стандартний алгоритм методу стиснених відображень, але не щодо метрик, а щодо норм, згідно з яким матимемо

$$\begin{aligned}
\|\varphi_1, \varphi_2\| &= \frac{\sup}{[a \leq x \leq b]} |\varphi_1(x) - \varphi_2(x)| = \frac{\sup}{[a \leq x \leq b]} |(f + A\varphi_0) - (f + A\varphi_1)| = \\
&= \frac{\sup}{[a \leq x \leq b]} |A(\varphi_0 - \varphi_1)| = \frac{\sup}{[a \leq x \leq b]} \left| \lambda \int_a^b K(x, y) (\varphi_0(y) - \varphi_1(y)) dy \right| \leq \\
&\leq \lambda \int_a^b \frac{\sup}{[a \leq x, y \leq b]} |K(x, y)| \frac{\sup}{[a \leq y \leq b]} |\varphi_0(y) - \varphi_1(y)| dy = \\
&= \lambda \|K(x, y)\| \|\varphi_0, \varphi_1\| \int_a^b dy = \lambda \|K(x, y)\| (b-a) \|\varphi_0, \varphi_1\| = \theta \|\varphi_0, \varphi_1\|,
\end{aligned} \tag{7.48}$$

$$\begin{aligned}
\|\varphi_2, \varphi_3\| &= \frac{\sup}{[a \leq x \leq b]} |\varphi_2(x) - \varphi_3(x)| = \frac{\sup}{[a \leq x \leq b]} |(f + A(A\varphi_0)) - (f + A(A\varphi_1))| = \\
&= \frac{\sup}{[a \leq x \leq b]} |A^2(\varphi_0 - \varphi_1)| = \frac{\sup}{[a \leq x \leq b]} \left| \lambda \int_a^b K(x, y) \left[ \lambda \int_a^b K(y, \omega) (\varphi_0(\omega) - \varphi_1(\omega)) d\omega \right] dy \right| \leq \\
&\leq \lambda^2 \int_a^b \left\{ \frac{\sup}{[a \leq x, y \leq b]} |K(x, y)| \right\}^2 \frac{\sup}{[a \leq \omega \leq b]} |\varphi_0(\omega) - \varphi_1(\omega)| \int_a^b d\omega dy = \\
&= \lambda^2 \|K(x, y)\|^2 \|\varphi_0, \varphi_1\| (b-a) \int_a^b dy = \lambda^2 \|K(x, y)\|^2 (b-a)^2 \|\varphi_0, \varphi_1\| = \theta^2 \|\varphi_0, \varphi_1\|,
\end{aligned} \tag{7.49}$$

Продовжуючи, за аналогією, для послідовності (7.42) отримаємо

$$\begin{aligned}
\|\varphi_n, \varphi_{n+1}\| &= \frac{\sup}{[a \leq x \leq b]} |\varphi_n(x) - \varphi_{n+1}(x)| = \frac{\sup}{[a \leq x \leq b]} |(f + A(A^{n-1}\varphi_0)) - (f + A(A^{n-1}\varphi_1))| = \\
&= \frac{\sup}{[a \leq x \leq b]} |A^n(\varphi_0 - \varphi_1)| \leq \lambda^n \|K(x, y)\|^n (b-a)^n \|\varphi_0, \varphi_1\| = \theta^n \|\varphi_0, \varphi_1\|
\end{aligned} \tag{7.50}$$

У виразах (7.48), (7.49), (7.50)

$$\theta = \lambda \|K(x, y)\| (b-a). \tag{7.51}$$

Як ми пам'ятаємо, для того, щоб вирази (7.48), (7.49), (7.50) здійснювали стиснення, що веде до нерухокої точки, яка є розв'язком інтегрального рівняння (7.35), необхідно, щоб виконувалась нерівність

$$\theta < 1. \quad (7.52)$$

А оскільки норма  $\|K(x, y)\|$  ядра  $K(x, y)$  оператора Фредгольма і відрізок  $[a, b]$  числової осі, на якому здійснюється інтегрування в інтегральному рівнянні (7.35), визначаються умовами задачі, для якої синтезоване це рівняння, то з виразів (7.51), (7.52) витікає, що для того, щоб неоднорідне інтегральне рівняння Фредгольма (7.35) мало єдиний розв'язок  $\varphi(x)$ , необхідно вибрати параметр  $\lambda$ , виходячи зі співвідношення

$$\lambda < \frac{1}{\|K(x, y)\| (b - a)}. \quad (7.53)$$

Оскільки норма  $\|\varphi_n, \varphi_{n+1}\|$  задає відстань між двома точками  $\varphi_n$  і  $\varphi_{n+1}$  метричного простору  $C$  неперервних функцій з метрикою (7.22), а тому

$$\rho(\varphi_n, \varphi_{n+1}) = \sup_{[a \leq x \leq b]} |\varphi_n(x) - \varphi_{n+1}(x)| = \|\varphi_n, \varphi_{n+1}\|, \quad (7.54)$$

то вираз (7.50), який ми використали для того, щоб довести, що оператор Фредгольма здійснює стиснення відображень у заданому функціональному метричному просторі, можна використовувати також і для прийняття рішення, на якому етапі можна зупинити ітераційний процес

$$\begin{aligned} \varphi_0(x) &= f(x) = A^0 f, \\ \varphi_1(x) &= f(x) + \lambda \int_a^b K(x, y) \varphi_0(y) dy = f(x) + \lambda \int_a^b K(x, y) f(y) dy = A^0 f + A f, \\ \varphi_2(x) &= f(x) + \lambda \int_a^b K(x, y) \varphi_1(y) dy = A^0 f + A f + A^2 f, \\ &\dots\dots\dots, \\ \varphi_n(x) &= f(x) + \lambda \int_a^b K(x, y) \varphi_{n-1}(y) dy = \sum_{i=0}^n A^i f(x), \end{aligned} \quad (7.55)$$

обчислення наближень  $\varphi_n(x)$  до розв'язку  $\varphi(x)$  рівняння (7.35). Для цього потрібно задатись допустимим значенням  $\varepsilon$  похибки наближення і зупинити обчислення за виразами (7.55) на тому значенні  $n$ , для якого виконуватиметься нерівність

$$\|\varphi_n, \varphi_{n+1}\| < \varepsilon. \quad (7.56)$$

А на завершення цього підрозділу побудуємо резольвенту оператора Фредгольма. Перепишемо операторне рівняння (7.41) у вигляді

$$\varphi - A\varphi = f. \quad (7.57)$$

Операторне рівняння (7.57) можна подати і так:

$$(I - A)\varphi = f. \quad (7.58)$$

Визначимо обернений оператор  $(I - A)^{-1}$  і помножимо на нього зліва операторне рівняння (7.58). Отримаємо

$$(I - A)^{-1}(I - A)\varphi = (I - A)^{-1}f, \quad (7.59)$$

або

$$\varphi = (I - A)^{-1}f = R_{\Phi}f. \quad (7.60)$$

Згідно з визначенням, наведеним у підрозділі 6.2, оператор

$$R_{\Phi} = (I - A)^{-1} = \frac{I}{I - A} = I + A + A^2 + A^3 + \dots = \sum_{i=0}^{\infty} A^i \quad (7.61)$$

і є резольвентою операторного рівняння Фредгольма. А оскільки, визначаючи її, ми використали формулу для суми членів нескінченно спадної геометричної прогресії зі знаменником  $A$ , то існуватиме резольвента  $R_{\Phi}$  лише в тому випадку, коли

$$\|A\| < 1, \quad (7.62)$$

тобто, коли

$$\lambda < \frac{1}{\|K(x, y)\| (b - a)\|\varphi\|}, \quad (7.63)$$

що цілком узгоджується з виразом (7.53).

З урахуванням виразу (7.61) розв'язок (7.60) операторного рівняння (7.57) можна переписати і так

$$\varphi(x) = \sum_{i=0}^{\infty} A^i f(x) \quad (7.64)$$

А застосовуючи до останнього рівняння в системі наближень (7.55) граничний перехід, матимемо



$${}_n \underline{\lim}_{\infty} \varphi_n(x) = {}_n \underline{\lim}_{\infty} \sum_{i=0}^n A^i f(x) = \sum_{i=0}^{\infty} A^i f(x). \quad (7.65)$$

Із виразів (7.64) та (7.65) витікає, що

$${}_n \underline{\lim}_{\infty} \varphi_n(x) = \varphi(x). \quad (7.66)$$

Цим ми підтвердили, що в процесі послідовних наближень (7.55) ми обов'язково прийдемо до наближеного розв'язку рівняння Фредгольма 2-го роду, похибка якого не перевищуватиме задану нами величину  $\varepsilon$ .

#### 7.4 Приклади розв'язання операторних рівнянь

**Приклад перший**, яким ми продемонструємо, як застосовувати метод стиснених відображень для розв'язання операторного рівняння, що має форму інтегрального рівняння Фредгольма 2-го роду, заданого у вигляді (7.35).

Нехай

$$a = 0, \quad b = 1, \quad f(x) = x, \quad K(x, y) = x + y, \quad \varepsilon = 0,1. \quad (7.67)$$

Підставляючи умови (7.67) в вираз (7.35), матимемо конкретизоване інтегральне рівняння Фредгольма 2-го роду у вигляді

$$\varphi(x) = x + \lambda \int_0^1 (x + y) y dy, \quad (7.68)$$

Розпочати розв'язання цього рівняння методом стиснених відображень нам потрібно з визначення допустимого значення параметра  $\lambda$ , для чого слід використати нерівність (7.53), у праву частину якої необхідно підставити також і норму ядра оператора Фредгольма, числове значення якої визначається виразом (7.37). Для умов, заданих виразами (7.67), числове значення цієї норми дорівнюватиме

$$\|K(x, y)\| = \frac{\sup}{[a \leq x, y \leq b]} |K(x, y)| = \frac{\sup}{[0 \leq x, y \leq 1]} |x + y| = |1 + 1| = 2. \quad (7.69)$$

Підставляючи числові значення відповідних параметрів із виразів (7.67) та (7.69) у нерівність (7.53), матимемо

$$\lambda < \frac{1}{\|K(x, y)\| (b - a)} = \frac{1}{2 |1 - 0|} = \frac{1}{2} = 0,5. \quad (7.70)$$

Приймаємо

$$\lambda = 0,4. \quad (7.71)$$

Для наших умов ітераційний процес (7.55) наближень до розв'язку рівняння (7.68) матиме вигляд

$$\begin{cases} \varphi_0(x) = x, \\ \varphi_n(x) = x + 0,4 \int_0^1 (x+y)\varphi_{n-1}(y)dy, \quad n = 1, 2, 3, \dots \end{cases} \quad (7.72)$$

Зупинятимемо цей процес і оголошуватимемо наближеним розв'язком рівняння (7.68) останнє наближення, що задовольнятиме критерій (7.56), який для наших умов (7.67) матиме вигляд

$$\|\varphi_n, \varphi_{n+1}\| < 0,1 \quad n = 0, 1, 2, \dots \quad (7.73)$$

А далі приступаємо до ітерацій.

Нехай  $n = 1$ .

У цьому випадку з виразу (7.72) матимемо

$$\begin{aligned} \varphi_1(x) &= x + 0,4 \int_0^1 (x+y)ydy = x + 0,4x \int_0^1 ydy + 0,4 \int_0^1 y^2 dy = x + 0,4x \left( \frac{y^2}{2} \right) \Big|_0^1 + \\ &+ 0,4 \left( \frac{y^3}{3} \right) \Big|_0^1 = x + 0,4x \left( \frac{1}{2} \right) + 0,4 \left( \frac{1}{3} \right) = 1,2x + 0,133 \end{aligned} \quad (7.74)$$

Перевіримо отримане наближення  $\varphi_1(x)$  до розв'язку  $\varphi(x)$ , скориставшись критерієм (7.73), підставляючи в який значення  $\varphi_0(x)$  із виразу (7.72) та значення  $\varphi_1(x)$  із виразу (7.74) та розкриваючи норму на основі виразу (7.54), матимемо

$$\begin{aligned} \|\varphi_0, \varphi_1\| &= \sup_{[0 \leq x \leq 1]} |\varphi_0(x) - \varphi_1(x)| = \sup_{[0 \leq x \leq 1]} |x - 1,2x - 0,133| = \\ &= |(-0,2) \bullet 1 - 0,133| = 0,333 \end{aligned} \quad (7.75)$$

Порівнюючи чисельне значення виразу (7.75) з правою частиною виразу (7.73) при  $n = 0$ , бачимо, що нерівність не виконується, тож зупинити ітераційний процес на отриманні наближення (7.74) не можна.

А тому нехай тепер  $n = 2$ .

У цьому випадку із виразів (7.72) та (7.74) матимемо

$$\begin{aligned} \varphi_2(x) &= x + 0,4 \int_0^1 K(x, y)\varphi_1(y)dy = x + 0,4 \int_0^1 (x+y)(1,2y + 0,133)dy = x + 0,48x \int_0^1 ydy + \\ &+ 0,48 \int_0^1 y^2 dy + 0,053x \int_0^1 dy + 0,053 \int_0^1 ydy = x + 0,48x \left( \frac{y^2}{2} \right) \Big|_0^1 + 0,48 \left( \frac{y^3}{3} \right) \Big|_0^1 + 0,053x(y) \Big|_0^1 + \\ &+ 0,053 \left( \frac{y^2}{2} \right) \Big|_0^1 = x + 0,24x + 0,16 + 0,053x + 0,026 = 1,293x + 0,186 \end{aligned} \quad (7.76)$$

Перевіримо отримане наближення  $\varphi_2(x)$  до розв'язку  $\varphi(x)$ , скориставшись критерієм (7.73), підставляючи в який значення  $\varphi_1(x)$  із виразу (7.74) та значення  $\varphi_2(x)$  із виразу (7.76) та розкриваючи норму на основі виразу (7.54), матимемо

$$\begin{aligned} \|\varphi_1, \varphi_2\| &= \frac{\sup}{[0 \leq x \leq 1]} |\varphi_1(x) - \varphi_2(x)| = \frac{\sup}{[0 \leq x \leq 1]} |1,2x + 0,133 - 1,293x - 0,186| = \\ &= |(-0,093) \bullet 1 - 0,053| = 0,146 \end{aligned} \quad (7.77)$$

Порівнюючи числове значення виразу (7.77) з правою частиною виразу (7.73) при  $n=1$ , бачимо, що нерівність не виконується, тож зупиняти ітераційний процес на отриманні наближення (7.76) теж не можна.

Тож нехай тепер  $n=3$ .

У цьому випадку із виразів (7.72) та (7.76) матимемо

$$\begin{aligned} \varphi_3(x) &= x + 0,4 \int_0^1 K(x, y) \varphi_2(y) dy = x + 0,4 \int_0^1 (x+y)(1,293y + 0,186) dy = x + 0,516x \int_0^1 y dy + \\ &+ 0,516 \int_0^1 y^2 dy + 0,074x \int_0^1 dy + 0,074 \int_0^1 y dy = x + 0,516x \left( \frac{y^2}{2} \right) \Big|_0^1 + 0,516 \left( \frac{y^3}{3} \right) \Big|_0^1 + 0,074x(y) \Big|_0^1 + \\ &+ 0,074 \left( \frac{y^2}{2} \right) \Big|_0^1 = x + 0,258x + 0,172 + 0,074x + 0,037 = 1,332x + 0,209 \end{aligned} \quad (7.78)$$

Перевіримо отримане наближення  $\varphi_3(x)$  до розв'язку  $\varphi(x)$ , скориставшись критерієм (7.73), підставляючи в який значення  $\varphi_2(x)$  із виразу (7.76) та значення  $\varphi_3(x)$  із виразу (7.78) та розкриваючи норму на основі виразу (7.54), матимемо

$$\begin{aligned} \|\varphi_2, \varphi_3\| &= \frac{\sup}{[0 \leq x \leq 1]} |\varphi_2(x) - \varphi_3(x)| = \frac{\sup}{[0 \leq x \leq 1]} |1,293x + 0,186 - 1,332x - 0,209| = \\ &= |(-0,039) \bullet 1 - 0,023| = 0,062 \end{aligned} \quad (7.79)$$

Порівнюючи числове значення виразу (7.79) з правою частиною виразу (7.73) при  $n=2$ , бачимо, що нерівність виконується, тож ітераційний процес після отримання наближення (7.78) можна зупинити і вважати, що з похибкою, яка не перевищує задане умовами (7.67) значення 0,1, наближеним розв'язком інтегрального рівняння (7.68) є вираз

$$\varphi(x) \approx \varphi_3(x) = 1,332x + 0,209. \quad (7.80)$$

**Приклад другий**, яким ми ще раз продемонструємо, як застосовувати метод стиснених відображень для розв'язання операторного рівняння, що має форму інтегрального рівняння Фредгольма 2-го роду, заданого у вигляді (7.35), але тепер уже з використанням програмного середовища MathCAD [26], MATLAB [27], [28] або Python [29],[30],[31].

Нехай тепер

$$a = 0, \quad b = \pi, \quad f(x) = \sin x, \quad K(x, y) = e^{x+y}, \quad \varepsilon = 0,05. \quad (7.81)$$

Підставляючи умови (7.81) у вираз (7.35), матимемо конкретизоване інтегральне рівняння Фредгольма 2-го роду у вигляді

$$\varphi(x) = \sin x + \lambda \int_0^{\pi} e^{x+y} \sin y dy, \quad (7.82)$$

Розпочати розв'язання і цього рівняння методом стиснених відображень нам знову ж таки потрібно з визначення допустимого значення параметра  $\lambda$ , для чого слід використати нерівність (7.53), у праву частину якої необхідно підставити також і норму ядра оператора Фредгольма, числове значення якої визначається виразом (7.37). Для умов, заданих виразами (7.81), числове значення цієї норми дорівнюватиме

$$\|K(x, y)\| = \frac{\sup}{[a \leq x, y \leq b]} |K(x, y)| = \frac{\sup}{[0 \leq x, y \leq \pi]} |e^{-(x+y)}| = |e^{-(0+0)}| = 1. \quad (7.83)$$

Підставляючи числові значення відповідних параметрів із виразів (7.81) та (7.83) у нерівність (7.53), матимемо

$$\lambda < \frac{1}{\|K(x, y)\| (b-a)} = \frac{1}{1 |\pi - 0|} \approx \frac{1}{3,14} \approx 0,318. \quad (7.84)$$

Приймаємо

$$\lambda = 0,3. \quad (7.85)$$

Для наших умов ітераційний процес (7.55) наближень до розв'язку рівняння (7.82) матиме вигляд

$$\begin{cases} \varphi_0(x) = \sin x, \\ \varphi_n(x) = \sin x + 0,3 \int_0^{\pi} e^{-(x+y)} \varphi_{n-1}(y) dy, \quad n = 1, 2, 3, \dots \end{cases} \quad (7.86)$$

Зупинятимемо цей процес і оголошуватимемо наближеним розв'язком рівняння (7.82) останнє наближення, що задовольнятиме критерій (7.56), який для наших умов (7.81) матиме вигляд

$$\|\varphi_n, \varphi_{n+1}\| < 0,05 \quad n = 0, 1, 2, \dots \quad (7.87)$$

А далі приступаємо до ітерацій.

Нехай  $n = 1$ .

У цьому випадку з виразу (7.86) матимемо

$$\varphi_1(x) = \sin x + 0,3 \int_0^{\pi} e^{-(x+y)} \sin y dy = \sin x + 0,3e^{-x} \int_0^{\pi} e^{-y} \sin y dy. \quad (7.88)$$

Аналізуючи вираз (7.88), бачимо, що для визначення числового значення інтеграла в правій частині цього виразу нам потрібно використати якусь стандартну програму в якомусь пакеті прикладних програм (MathCAD, MATLAB або Python).

Використовуючи будь-який із цих ППП, знаходимо, що

$$\int_0^{\pi} e^{-y} \sin y dy \approx 0,25. \quad (7.89)$$

Підставляючи числове значення виразу (7.89) у вираз (7.88), матимемо

$$\varphi_1(x) = \sin x + 0,075e^{-x}. \quad (7.90)$$

Перевіримо отримане наближення  $\varphi_1(x)$  до розв'язку  $\varphi(x)$ , скориставшись критерієм (7.87), підставляючи в який значення  $\varphi_0(x)$  із виразу (7.86) та значення  $\varphi_1(x)$  із виразу (7.90) і розкриваючи норму на основі виразу (7.54), матимемо

$$\begin{aligned} \|\varphi_0, \varphi_1\| &= \sup_{[0 \leq x \leq \pi]} |\varphi_0(x) - \varphi_1(x)| = \sup_{[0 \leq x \leq \pi]} |\sin x - \sin x - 0,075e^{-x}| = \\ &= |-0,075e^{-0}| = 0,075 \end{aligned} \quad (7.91)$$

Порівнюючи чисельне значення виразу (7.91) з правою частиною виразу (7.87) при  $n=1$ , бачимо, що нерівність не виконується, тож зупиняти ітераційний процес на отриманні наближення (7.90) не можна.

Тож нехай тепер  $n=2$ .

У цьому випадку із виразів (7.86) та (7.90) матимемо

$$\begin{aligned} \varphi_2(x) &= \sin x + 0,3 \int_0^{\pi} e^{-(x+y)} (\sin y + 0,075e^{-y}) dy = \sin x + 0,3e^{-x} \left( \int_0^{\pi} e^{-y} \sin y dy + \right. \\ &0 + 0,075 \int_0^{\pi} e^{-2y} dy \left. \right) = \sin x + 0,3e^{-x} \left( 0,25 + \frac{0,075}{(-2)} (e^{-2y}) \Big|_0^{\pi} \right) = \sin x + 0,3e^{-x} (0,25 + \\ &+ \frac{0,075(0,001 - 1,0)}{(-2)}) = \sin x + 0,086e^{-x} \end{aligned} \quad (7.92)$$

Перевіримо отримане наближення  $\varphi_2(x)$  до розв'язку  $\varphi(x)$ , скориставшись критерієм (7.87), підставляючи в який значення  $\varphi_1(x)$  із виразу (7.90) та значення  $\varphi_2(x)$  із виразу (7.92) і розкриваючи норму на основі виразу (7.54), матимемо

$$\begin{aligned} \|\varphi_1, \varphi_2\| &= \sup_{[0 \leq x \leq \pi]} |\varphi_1(x) - \varphi_2(x)| = \sup_{[0 \leq x \leq \pi]} |\sin x + 0,075e^{-x} - \sin x - 0,086e^{-x}| = \\ &= |-0,011e^{-0}| = 0,011 \end{aligned} \quad (7.93)$$

Порівнюючи числове значення виразу (7.93) з правою частиною виразу (7.87) при  $n=1$ , бачимо, що нерівність виконується, тож ітераційний процес після отримання наближення (7.92) можна зупинити і вважати, що з похибкою, яка не перевищує задане умовами (7.81) значення 0,05, наближеним розв'язком інтегрального рівняння (7.82) є вираз

$$\varphi(x) \approx \varphi_2(x) = \sin x + 0,086e^{-x}. \quad (7.94)$$

### 7.5 Завдання для самоперевірки

1. Застосовуючи метод стиснення відображень, доведіть, що алгебраїчне рівняння має єдину нерухому точку на заданому відрізку числової осі.
2. Якщо алгебраїчне рівняння має кілька коренів, як віднайти усі його нерухомі точки?
3. Застосовуючи метод стиснення відображень, доведіть існування єдиної нерухомої точки диференціального оператора.
4. Як трансформувати інтегральний оператор в диференціальний?
5. Застосовуючи метод стиснення відображень, доведіть існування єдиної нерухомої точки оператора Фредгольма.
6. Що собою являє резольвента оператора Фредгольма?
7. Побудуйте алгоритм послідовних наближень до єдиної нерухомої точки оператора Фредгольма.
8. За яким критерієм визначається досягнення необхідної точності наближеного розв'язку неоднорідного інтегрального рівняння Фредгольма 2-го роду?
9. В чому принципова відмінність операторів Фредгольма і Вольтерра?
10. Як співвідносяться між собою потужності множин розв'язків інтегральних рівнянь Фредгольма та Вольтера?

## Розділ 8

### СПЕЦІАЛЬНІ ОПЕРАТОРИ ТА ЇХ ПРИКЛАДНІ ЗАСТОСУВАННЯ

#### 8.1 Прямий та обернений оператори Лапласа

З загального курсу вищої математики, який викладається студентам технічних ЗВО, відомо, що *за допомогою оператора Лапласа*

$$L\{f(t)\} = F(p) = \int_0^{\infty} f(t)e^{-pt} dt \quad (8.1)$$

*кожній неперервній функції  $f$  часу  $t$ , заданій на множині дійсних чисел, яка задовольняє умову  $f(t) = 0$  при  $t < 0$ , умови Діріхле і називається оригіналом, можна поставити у відповідність функцію  $F$  комплексної змінної  $p = \sigma + j\omega$ , яка називається зображенням оригіналу на комплексній площині. Ця відповідність записується у такий спосіб:*

$$f(t) \Leftrightarrow F(p). \quad (8.2)$$

Наприклад, часовій функції одиничного стрибка

$$x(t) = 1(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0, \end{cases} \quad (8.3)$$

на комплексній площині відповідає зображення

$$F(p) = \int_0^{\infty} 1(t)e^{-pt} dt = \int_0^{\infty} e^{-pt} dt = \frac{1}{-p} \cdot e^{-pt} \Big|_0^{\infty} = \frac{1}{p}, \quad (8.4)$$

або

$$1(t) \Leftrightarrow \frac{1}{p}. \quad (8.5)$$

Ще один приклад – експоненті  $e^{-\alpha t}$  при  $t \geq 0$  на комплексній площині відповідає зображення

$$F(p) = \int_0^{\infty} e^{-\alpha t} \cdot e^{-pt} dt = \int_0^{\infty} e^{-(p+\alpha)t} dt = \frac{1}{-(p+\alpha)} e^{-(p+\alpha)t} \Big|_0^{\infty} = \frac{1}{p+\alpha}. \quad (8.6)$$

Головною перевагою аналізу в області зображень  $F(p)$ , тобто на комплексній площині, порівняно з аналізом в області оригіналів  $f(t)$ , тобто у часовому просторі, є те, що за нульових початкових умов операції диференціювання  $\frac{d}{dt}$  оригіналу  $f(t)$  у часовому просторі відповідає операція перемноження на комплексну змінну  $p$  його зображення  $F(p)$  на комплексній площині, тобто

$$\frac{df}{dt} = \dot{f}(t) \Leftrightarrow p \cdot F(p), \quad (8.7)$$

оскільки

$$L\left\{\frac{df}{dt}\right\} = \int_0^{\infty} \frac{df}{dt} e^{-pt} dt = \left(f(t)e^{-pt}\right)\Big|_0^{\infty} - \int_0^{\infty} f(t)(-pe^{-pt})dt = p \int_0^{\infty} f(t)e^{-pt} dt = pF(p). \quad (8.8)$$

Звертаємо увагу на те, що при інтегруванні за виразом (8.8) ми використали відому методику інтегрування частинами, задаючи новим змінним  $u, v$  значення  $u = e^{-pt}$ ,  $dv = df$ , внаслідок чого матимемо  $v = f$ ,  $du = -pe^{-pt} dt$ , та нульову початкову умову  $f(0) = 0$ .

Застосовуючи цю ж методику, отримаємо, що

$$\frac{d^2 f}{dt^2} = \ddot{f}(t) \Leftrightarrow p^2 \cdot F(p) \quad (8.9)$$

та

$$\frac{d^n f}{dt^n} = f^{(n)}(t) \Leftrightarrow p^n \cdot F(p). \quad (8.10)$$

А операції інтегрування оригіналу  $f(t)$  у часовому просторі відповідає операція ділення на комплексну змінну  $p$  його зображення  $F(p)$  на комплексній площині, тобто

$$\int_0^t f(\tau) d\tau \Leftrightarrow \frac{F(p)}{p}, \quad (8.11)$$

оскільки

$$\begin{aligned} L\left\{\int_0^t f(\tau) d\tau\right\} &= \int_0^{\infty} \left(\int_0^t f(\tau) d\tau\right) e^{-pt} dt = \left(\frac{1}{-p} e^{-pt} \int_0^t f(\tau) d\tau\right)\Big|_0^{\infty} - \int_0^{\infty} \left(\frac{1}{-p} e^{-pt}\right) f(t) dt = \\ &= \frac{1}{p} \int_0^{\infty} f(t) e^{-pt} dt = \frac{1}{p} F(p) \end{aligned} \quad (8.12)$$



При інтегруванні за виразом (8.12) ми теж використали методику інтегрування частинами, задаючи новим змінним  $u, v$  значення  $u = \int_0^t f(\tau) d\tau$ ,  $dv = e^{-pt} dt$ , внаслідок чого матимемо  $v = \left( \frac{1}{-p} e^{-pt} \right)$ ,  $du = f(t) dt$ .

Застосовуючи цю ж методику, отримаємо, що

$$\underbrace{\int_0^t \dots \int_0^t f(\tau_1, \tau_2, \dots, \tau_n) d\tau_1 d\tau_2 \dots d\tau_n}_{n} \Leftrightarrow \frac{F(p)}{p^n}. \quad (8.13)$$

*Завдяки властивостям (8.7), (8.11) та їх наслідкам (8.10), (8.13) диференціальним та інтегральним рівнянням, записаним у часовому просторі, відповідають алгебраїчні рівняння на комплексній площині, розв'язувати які набагато простіше, оскільки цьому вчать ще у школі.*

Наприклад, диференціальному рівнянню в області оригіналів  $x(t)$ ,  $y(t)$

$$a_2 \frac{d^2 y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_1 \frac{dx(t)}{dt} + b_0 x(t) \quad (8.14)$$

на комплексній площині відповідає алгебраїчне рівняння

$$a_2 p^2 Y(p) + a_1 p Y(p) + a_0 Y(p) = b_1 p X(p) + b_0 X(p) \quad (8.15)$$

відносно зображень  $X(p)$  та  $Y(p)$ . Його розв'язком є функція  $Y(p)$ , яку з рівняння (8.15) можна визначити у такий спосіб:

$$Y(p) = \frac{b_1 p + b_0}{a_2 p^2 + a_1 p + a_0} X(p) \quad (8.16)$$

або

$$Y(p) = W(p) \cdot X(p), \quad (8.17)$$

де

$$W(p) = \frac{b_1 p + b_0}{a_2 p^2 + a_1 p + a_0}. \quad (8.18)$$

Якщо  $X(p)$  вважати зображенням сигналу  $x(t)$ , що надходить на вхід лінійної динамічної системи (ЛДС), математичну модель якої можна записати у вигляді (8.14), де

$y(t)$  є реакцією цієї системи на вхідний сигнал  $x(t)$ , то функцію  $W(p)$  можна інтерпретувати як передаточну (або передатну) функцію системи (рис. 8.1).

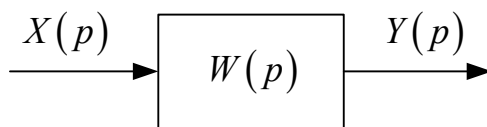


Рисунок 8.1 – Узагальнена структурна схема лінійної динамічної системи в області зображень

Як видно із виразів (8.16)–(8.18), передаточна функція  $W(p)$  не залежить від зовнішніх сигналів, які діють на систему, але однозначно характеризує її здатність передавати ці сигнали зі свого входу на вихід.

А тому ця функція є однією з найважливіших математичних моделей лінійних динамічних систем, для яких перетворення Лапласа (8.1) задає взаємно однозначну відповідність між оригіналами та їх зображеннями.

Із виразу (8.17) випливає, що, знаючи зображення  $X(p)$  вхідного сигналу  $x(t)$  та зображення  $Y(p)$  реакції системи  $y(t)$  на цей сигнал, передаточну функцію можна отримати, взявши їх відношення, тобто

$$W(p) = \frac{Y(p)}{X(p)}. \quad (8.19)$$

**Для оператора Лапласа, який визначається виразом (8.1), існує обернений оператор**

$$L^{-1}\{F(p)\} = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} F(p)e^{pt} dp = \frac{1}{2\pi j} \oint F(p)e^{pt} dp = f(t), \quad (8.20)$$

згідно з яким за відомим зображенням  $F(p)$  можна знайти оригінал  $f(t)$ , який, як правило, використовується лише для побудови таблиць відповідності між  $f(t)$  та  $F(p)$ , а в практиці аналізу частіше використовуються формули розкладання, отримані математиками шляхом застосування теореми лишків при інтегруванні у виразі (8.20), одна з яких — для некротних полюсів  $p_i$  зображення

$$Y(p) = \frac{C(p)}{D(p)}, \quad (8.21)$$

де  $C(p), D(p)$  – багаточлени за степенями  $p$  порядків  $m$  та  $n$ , відповідно, має вигляд

$$y(t) = \sum_{i=1}^n \frac{C(p_i)}{D'(p_i)} e^{p_i t}. \quad (8.22)$$

Нагадаємо, що  $p_i$  – це корені рівняння

$$D(p) = 0, \quad (8.23)$$

які називають полюсами виразу (8.21), а

$$D'(p_i) = \left. \frac{dD}{dp} \right|_{p=p_i}. \quad (8.24)$$

Наведемо приклад використання формули розкладання (8.22).  
Нехай маємо зображення невідомого нам оригіналу у вигляді

$$F(p) = \frac{2p+1}{p^3+5p^2+6p}. \quad (8.25)$$

Потрібно визначити його оригінал  $f(t)$ .

Очевидно, що для нашого прикладу

$$\begin{aligned} C(p) &= 2p+1, \\ D(p) &= p^3+5p^2+6p, \\ \frac{dD}{dp} &= 3p^2+10p+6. \end{aligned} \quad (8.26)$$

Знайдемо полюси зображення (8.25), тобто корені рівняння

$$p^3+5p^2+6p=0. \quad (8.27)$$

Привівши рівняння (8.27) до вигляду

$$p(p^2+5p+6)=0, \quad (8.28)$$

легко бачити, що полюсами зображення (8.25) є

$$\begin{cases} p_1 = 0, \\ p_2 = -2, \\ p_3 = -3. \end{cases} \quad (8.29)$$

Підставляючи (8.26) та (8.29) у формулу розкладання (8.22), отримаємо

$$\begin{aligned} f(t) &= \frac{C(p_1)}{D'(p_1)} e^{p_1 t} + \frac{C(p_2)}{D'(p_2)} e^{p_2 t} + \frac{C(p_3)}{D'(p_3)} e^{p_3 t} = \\ &= \frac{2 \cdot (0) + 1}{3 \cdot (0)^2 + 10 \cdot (0) + 6} e^{0 \cdot t} + \frac{2 \cdot (-2) + 1}{3 \cdot (-2)^2 + 10 \cdot (-2) + 6} e^{-2t} + \\ &+ \frac{2 \cdot (-3) + 1}{3 \cdot (-3)^2 + 10 \cdot (-3) + 6} e^{-3t} = \frac{1}{6} + \frac{3}{2} e^{-2t} - \frac{5}{3} e^{-3t}. \end{aligned} \quad (8.30)$$

Функція

$$f(t) = \frac{1}{6} + \frac{3}{2} e^{-2t} - \frac{5}{3} e^{-3t} \quad (8.31)$$

і є оригіналом зображення  $F(p)$ , заданого виразом (8.25). Ще раз нагадуємо, що оригінал визначається лише при значеннях  $t \geq 0$ .

Відзначимо, що в разі, якщо серед полюсів зображення (8.21), або, що одне і те ж, серед коренів рівняння (8.23)  $n$ -го порядку є кратний корінь, наприклад  $p_1$  кратності  $k$ , тобто, коли рівняння (8.23) набуває вигляду

$$(p - p_1)^k D_{n-k}(p) = 0, \quad (8.32)$$

то замість формули розкладання у вигляді (8.22) матимемо формулу розкладання у вигляді

$$y(t) = \frac{1}{(k-1)!} \frac{d^{(k-1)}}{dp^{(k-1)}} \left[ \frac{C(p)(p-p_1)^k e^{p t}}{D(p)} \right]_{p=p_1} + \sum_{i=2}^{n-1} \frac{C(p_i)}{D'(p_i)} e^{p_i t}, \quad (8.33)$$

або

$$y(t) = \frac{1}{(k-1)!} \frac{d^{(k-1)}}{dp^{(k-1)}} \left[ \frac{C(p) e^{p t}}{D_{n-k}(p)} \right]_{p=p_1} + \sum_{i=2}^{n-1} \frac{C(p_i)}{D'(p_i)} e^{p_i t} \quad (8.34)$$

Зв'яжемо передаточну функцію  $W(p)$  лінійної динамічної системи з її перехідною  $h(t)$  та імпульсною перехідною  $g(t)$  характеристиками.

**Перехідна характеристика  $h(t)$  системи є її реакцією на вхідний сигнал  $x(t)$  у вигляді одиничного стрибка.** Тобто у випадку, коли  $x(t)=1(t)$ , маємо  $y(t)=h(t)$ .

Графічна інтерпретація цього означення показана на рис. 8.2.

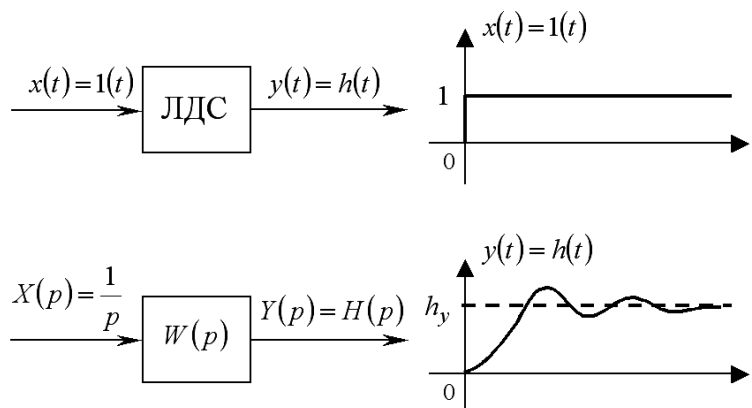


Рисунок 8.2 – Графік реакції  $h(t)$  лінійної динамічної системи на одиничний стрибок  $1(t)$

**Імпульсна перехідна або вагова характеристика  $g(t)$  системи є її реакцією на одиничний імпульсний вхідний сигнал  $x(t)$  у вигляді дельта-функції  $\delta(t)$ ,** для якої справедливо:

$$\delta(t) = \begin{cases} \infty, & t = 0, \\ 0, & t \neq 0, \end{cases} \quad (8.35)$$

$$\int_{-\infty}^{\infty} \delta(t) dt = \int_0^{\infty} \delta(t) dt = 1. \quad (8.36)$$

Із виразів (8.35), (8.36) випливає, що дельта-функція – це ідеалізація імпульсу одиничної площі з надзвичайно великою висотою і надзвичайно малою протяжністю (рис. 8.3).

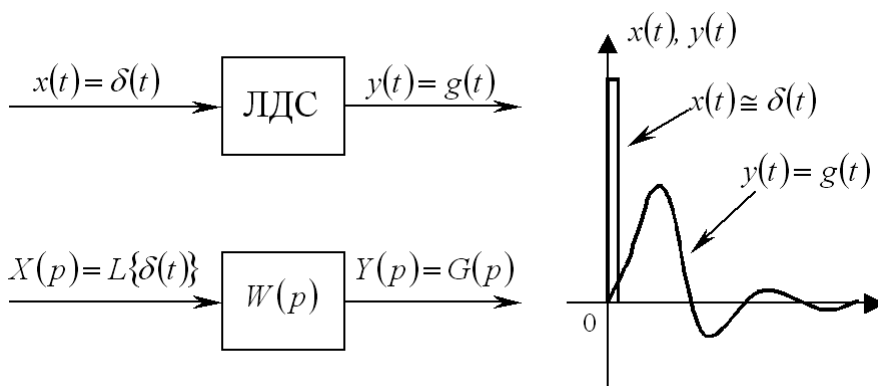


Рисунок 8.3 – Графік реакції  $g(t)$  лінійної динамічної системи ЛДС на одиничний імпульс  $\delta(t)$

Дуже важливим є те, що сигнал  $x(t)$ , який діє на вході лінійної динамічної системи з імпульсною перехідною характеристикою  $g(t)$  (рис. 8.4), та реакція системи  $y(t)$  на цей сигнал пов'язані між собою інтегралом згортки

$$y(t) = \int_0^{\infty} x(t-\tau) g(\tau) dt, \quad (8.37)$$

який відносять до класу інтегральних рівнянь Фредгольма, розглянутих нами раніше, і який має надзвичайно прозорий зміст – вихідний сигнал динамічної системи формується сумою реакцій на кожний імпульс вхідного сигналу під час подання цього вхідного сигналу у вигляді послідовності імпульсів з висотою, що дорівнює значенню вхідного сигналу у відповідний момент часу.

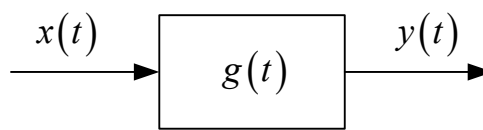


Рисунок 8.4 – Узагальнена структурна схема лінійної динамічної системи

Оскільки, згідно зі співвідношенням (8.17), зображення за Лапласом вихідного сигналу  $Y(p)$  лінійної динамічної системи є добутком передаточної функції цієї системи  $W(p)$  та зображення за Лапласом вхідного сигналу системи  $X(p)$ , яке в разі виконання (8.3), згідно з (8.5), дорівнюватиме  $\frac{1}{p}$ , то для зображення за Лапласом перехідної характеристики  $H(p)$  матимемо:

$$H(p) = \frac{W(p)}{p}, \quad (8.38)$$

що, в свою чергу, дає нам право записати

$$h(t) = L^{-1} \left\{ \frac{W(p)}{p} \right\}. \quad (8.39)$$

Легко бачити, що зображення за Лапласом дельта-функції дорівнює

$$L\{\delta(t)\} = \int_0^{\infty} \delta(t) e^{-pt} dt = e^{-p \cdot (t=0)} \int_0^{\infty} \delta(t) dt = e^{-p \cdot 0} \cdot 1 = 1. \quad (8.40)$$

Тож, якщо при

$$x(t) = \delta(t), \quad (8.41)$$

за визначенням

$$y(t) = g(t), \quad (8.42)$$

то з (8.17) випливає, що

$$G(p) = W(p). \quad (8.43)$$

Тобто, передаточна функція лінійної динамічної системи є зображенням за Лапласом її імпульсної перехідної характеристики, і навпаки

$$g(t) = L^{-1}\{W(p)\}. \quad (8.44)$$

Із співвідношень (8.38) і (8.43) випливає, що

$$G(p) = p \cdot H(p). \quad (8.45)$$

А це, в свою чергу, означає, що в області оригіналів справедливим є рівняння

$$g(t) = \frac{dh(t)}{dt}, \quad (8.46)$$

тобто, що імпульсну перехідну характеристику  $g(t)$  системи можна отримати диференціюванням її перехідної характеристики  $h(t)$ .

Підводячи підсумок викладеному вище, можна стверджувати, що математичну модель ЛДС у вигляді передаточної функції  $W(p)$  можна визначити, поділивши перетворену за Лапласом реакцію системи  $y(t)$  на перетворений за Лапласом вхідний сигнал  $x(t)$ . Цілком очевидно, що до перетворення за Лапласом і експериментально зафіксований вхідний сигнал  $x(t)$ , і експериментально зафіксовану реакцію  $y(t)$  системи на цей сигнал потрібно апроксимувати відповідними функціями аргументу  $t$ . Згідно з теоремою Вейерштрасса майже завжди це можна зробити за допомогою поліномів за степенями аргументу  $t$ , перетворення за Лапласом яких приводить до відношення поліномів за степенями аргументу  $p$ . Очевидним є також і те, що найбільш просто задача ідентифікації такої системи за цим алгоритмом буде розв'язуватись, якщо вхідним сигналом динамічної системи є одиничний стрибок  $1(t)$  або одиничний імпульс  $\delta(t)$ , з перетворених за Лапласом реакцій системи на кожен із яких одразу отримуємо передаточну функцію згідно зі співвідношеннями (8.38) або (8.43).

На завершення цього підрозділу покажемо, як, отримавши передаточну функцію  $W(p)$  ЛДС, побудувати математичну модель цієї системи у вигляді диференціального рівняння.

Нехай математична модель ЛДС на комплексній площині має вигляд

$$W(p) = \frac{b_m p^m + b_{m-1} p^{m-1} + \dots + b_1 p + b_0}{a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0}. \quad (8.47)$$

Підставивши значення  $W(p)$  із виразу (8.47) у вираз (8.17) та перенісши знаменник у ліву частину рівності, отримаємо

$$\begin{aligned} & (a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0) \cdot Y(p) = \\ & = (b_m p^m + b_{m-1} p^{m-1} + \dots + b_1 p + b_0) \cdot X(p). \end{aligned} \quad (8.48)$$

Розкриваючи дужки в рівності (8.48) і враховуючи те, що перемноженню зображення на  $p$  відповідає диференціювання оригіналу за  $t$ , приходимо до диференціального рівняння  $n$ -го порядку

$$\sum_{i=0}^n a_i \frac{d^i y}{dt^i} = \sum_{l=0}^m b_l \frac{d^l x}{dt^l}, \quad (8.49)$$

для якого виконується умова

$$m \leq n, \quad (8.50)$$

що обумовлена здатністю системи фізично реалізуватись під таку модель.

Цілком очевидно, що для отримання розв'язку диференціального рівняння  $n$ -го порядку його необхідно проінтегрувати  $n$  раз, що викликає появу  $n$  сталих інтегрування, для конкретизації яких потрібно знати в початковий момент часу не лише значення вихідної координати  $y(t)$ , а й значення у цей початковий момент усіх її похідних до  $(n-1)$  порядку включно, тобто початкові умови для рівняння (8.49) мають вигляд:

$$\left. \begin{aligned} & y(t)|_{t=0} = y(0), \\ & \frac{dy}{dt} \Big|_{t=0} = y^1(0), \\ & \dots\dots\dots, \\ & \frac{d^{n-1}y}{dt^{n-1}} \Big|_{t=0} = y^{n-1}(0). \end{aligned} \right\} \quad (8.51)$$

Слід відзначити, що для значної кількості ЛДС характерним є те, що всі елементи їхніх структур, які здатні запасати енергію, після відключення системи цю енергію



втрачають, що дає підставу перед новим включенням у роботу системи для її моделі у вигляді (8.49) вважати початкові умови нульовими, тобто в системі рівнянь (8.51) вважати усі праві частини такими, що дорівнюють нулю. Це одразу ж приводить при розв'язанні цього диференціального рівняння до переваг, обумовлених його трансформацією на комплексну площину за допомогою оператора Лапласа з застосуванням потім до отриманого зображення оберненого оператора Лапласа у вигляді однієї з форм теореми розкладання.

*Резюмуючи усе викладене вище можемо констатувати, що прямий оператор Лапласа реалізує закон відображення множини неперервних функцій дійсного аргументу, заданого на числовій осі, у множину неперервних функцій комплексного аргументу, заданого на комплексній площині, координати точок якої теж є дійсними числами. А обернений оператор Лапласа діє навпаки, реалізуючи обернений процес перетворення вказаних множин одна в одну.*

І завершимо ми розгляд прямого та оберненого операторів Лапласа посиланням на те, що у функціональному аналізі та пов'язаних з ним розділах математики теорія і практика застосування цих операторів носить назву операційного числення.

## 8.2 Прямий та обернений оператори Фур'є

*Для аналізу частотного спектра сигналу, описаного функцією  $x(t)$ , заданою у гільбертовому просторі, Фур'є запропонував використати оператор*

$$X(j\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt, \quad (8.52)$$

який множині функцій  $\{x(t)\}$  дійсного аргументу  $t$ , для яких виконуються умови Діріхле та умова обмеженості енергії

$$\int_{-\infty}^{\infty} x^2(t) dt < \infty, \quad (8.53)$$

ставить у відповідність множини функцій  $\{X(j\omega)\}$  уявного аргументу  $j\omega$ , в якому  $\omega$  є круговою частотою, що пов'язана з основною частотою  $f$  періодичних коливань сигналу  $x(t)$  співвідношенням

$$\omega = 2\pi f = \frac{2\pi}{T}, \quad (8.54)$$

де  $T$  – період коливання цього сигналу.

*А для встановлення відповідності між множиною функцій  $\{X(j\omega)\}$ , отриманих за допомогою оператора (8.52), і множиною функцій  $\{x(t)\}$  Фур'є запропонував обернений оператор*

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega) e^{j\omega t} d\omega. \quad (8.55)$$

А далі звертаємо увагу на те, що, як показано у попередньому підрозділі, достатньо повну інформацію про динаміку лінійної системи несе її передаточна функція  $W(p)$ , яка є функцією комплексної змінної  $p = \sigma + j\omega$ .

Якщо накласти умову

$$\sigma = 0, \quad (8.56)$$

то передаточна функція  $W(p)$  вироджується у функцію уявної частини  $j\omega$  комплексної змінної  $p$ , тобто в  $W(j\omega)$ .

До прикладу, нехай

$$W_1(p) = \frac{10}{5p + 1}. \quad (8.57)$$

За виконання умови (8.56) із (8.57) маємо

$$W_1(j\omega) = \frac{10}{1 + j5\omega}. \quad (8.58)$$

Функцію  $W_1(j\omega)$  шляхом очевидних перетворень можна привести до такої форми:

$$\begin{aligned} W_1(j\omega) &= \frac{10 \cdot (1 - j5\omega)}{(1 + j5\omega) \cdot (1 - j5\omega)} = \frac{10 - j50\omega}{1 + (5\omega)^2} = \\ &= \frac{10}{1 + 25\omega^2} - j \frac{50\omega}{1 + 25\omega^2} = P_1(\omega) + jQ_1(\omega), \end{aligned} \quad (8.59)$$

де

$$P_1(\omega) = \frac{10}{1 + 25\omega^2}, \quad (8.60)$$

$$Q_1(\omega) = -\frac{50\omega}{1 + 25\omega^2}. \quad (8.61)$$

Вирази (8.58)–(8.61) за своєю структурою справедливі для довільної функції  $W(p)$ . Оскільки як функція  $W(j\omega)$ , так і її складові  $P(\omega)$  та  $Q(\omega)$  у своїй структурі містять всі коефіцієнти передаточної функції  $W(p)$ , то можна стверджувати, що кожна з них несе однакову кількість інформації про динаміку системи, яка характеризується цією передаточною функцією. У цьому легко перекоонатись на наведеному прикладі.

З теорії комплексних чисел відомо, що при  $\omega = \omega^*$  комплексне число

$$W(j\omega^*) = P(\omega^*) + jQ(\omega^*) \quad (8.62)$$

можна записати не тільки в алгебраїчній формі (8.62), а й у показниковій

$$W(j\omega^*) = A(\omega^*) \cdot e^{j\varphi(\omega^*)}, \quad (8.63)$$

де

$$A(\omega^*) = \sqrt{P^2(\omega^*) + Q^2(\omega^*)}, \quad (8.64)$$

$$\varphi(\omega^*) = \operatorname{arctg} \frac{Q(\omega^*)}{P(\omega^*)}. \quad (8.65)$$

*Оскільки параметр  $\omega$  в усіх наведених вище виразах має чітко визначений фізичний сенс – це кругова частота, – то всі функції, що залежать від нього, називаються частотними характеристиками, з яких:*

*$P(\omega)$  – дійсна частотна характеристика (ДЧХ),*

*$Q(\omega)$  – уявна частотна характеристика (УЧХ),*

*$A(\omega)$  – амплітудна частотна характеристика (АЧХ),*

*$\varphi(\omega)$  – фазова частотна характеристика (ФЧХ),*

*$W(j\omega)$  – амплітудно-фазова частотна характеристика (АФЧХ).*

Для співвідношень (8.62)–(8.65), які пов'язують між собою частотні характеристики, слід додати ще два, які є очевидними для всіх, хто знайомий з теорією комплексних чисел, а саме:

$$P(\omega) = A(\omega) \cdot \cos \varphi(\omega), \quad (8.66)$$

$$Q(\omega) = A(\omega) \cdot \sin \varphi(\omega). \quad (8.67)$$

*Амплітудну  $A(\omega)$  та фазову  $\varphi(\omega)$  частотні характеристики динамічної системи легко отримати експериментально за допомогою спеціального вимірювального комплексу, який випускається серійно і містить у собі генератор синусоїдальних коливань змінної частоти, подвійний піковий вольтметр, фазометр-частотомір та пристрій для перетворення фізичних параметрів процесів на вході та виході системи в електричний сигнал і для їх узгодження. Зазначимо, що амплітудно-фазову частотну характеристику з експерименту можна побудувати тільки для стійких об'єктів.*

Експеримент з отримання частотних характеристик  $A(\omega)$  та  $\varphi(\omega)$  полягає в тому, що на вхід системи від генератора синусоїдальних коливань подаємо сигнал

$$x_1(t) = X_1(\omega_1) \sin \omega_1 t \quad (8.68)$$

з фіксованими частотою  $\omega_1 = 2\pi f_1$ , де  $f_1$  – реальна частота у герцах, та амплітудою  $X_1(\omega_1)$ .

Після закінчення перехідного процесу, тривалість якого визначається чотирма еквівалентними сталими часу цієї системи, на її виході встановлюються коливання

$$y_1(t) = Y_1(\omega_1) \cdot \sin(\omega_1 t + \varphi(\omega_1)). \quad (8.69)$$

Замірявши подвійним піковим вольтметром одночасно значення амплітуди  $X_1(\omega_1)$  сигналу на вході системи та  $Y_1(\omega_1)$  на її виході й взявши їх відношення, отримаємо значення амплітудної частотної характеристики системи на частоті  $\omega_1$ , тобто

$$A(\omega_1) = \frac{Y_1(\omega_1)}{X_1(\omega_1)}. \quad (8.70)$$

За допомогою фазометра-частотоміра на цій ж частоті  $\omega_1$  заміряємо зсув фаз  $\varphi(\omega_1)$ , який виникає між сигналами  $x_1(t)$  та  $y_1(t)$ . Він безпосередньо задає значення фазової частотної характеристики  $\varphi(\omega_1)$  на частоті  $\omega_1$ .

Дискретно змінюючи частоту вхідного сигналу від нуля до значень  $\omega_n$ , за яких можна вважати, що

$$Y(\omega_n) \cong 0, \quad (8.71)$$

і здійснюючи дії, що описані вище, отримаємо скінченну множину чисел

$$\{(\omega_0, A(\omega_0), \varphi(\omega_0)), (\omega_1, A(\omega_1), \varphi(\omega_1)), \dots, (\omega_n, A(\omega_n), \varphi(\omega_n))\}, \quad (8.72)$$

яка дискретно задає амплітудну  $A(\omega)$  та фазову  $\varphi(\omega)$  частотні характеристики системи.

Для отримання неперервних функцій  $A(\omega)$  та  $\varphi(\omega)$  залишається тільки здійснити апроксимацію послідовностей чисел

$$A(\omega_0), A(\omega_1), \dots, A(\omega_n), \quad (8.73)$$

$$\varphi(\omega_0), \varphi(\omega_1), \dots, \varphi(\omega_n) \quad (8.74)$$

за допомогою відомих інтерполяційних процедур, або використавши спеціально підібрані формули для  $A(\omega)$  і  $\varphi(\omega)$  та суму квадратів відхилень як критерій оптимізації процедури.

У разі використання першого шляху для отримання  $W(j\omega)$  необхідно підставити інтерполяційні вирази для  $A(\omega)$  та  $\varphi(\omega)$  у формулу (8.63), а потім за допомогою іншої стандартної процедури перейти від  $W(j\omega)$  до  $W(p)$ .

Другий шлях дозволяє одразу отримати формулу для  $W(p)$ .

Алгоритм побудови  $W(p)$  за цим підходом такий: спочатку висуваємо гіпотезу про те, що динаміка розглядуваної системи описується передаточною функцією

$$W_1(p) = \frac{b_1}{a_1 p + 1}. \quad (8.75)$$

Із формули (8.75) видно, що для ідентифікації системи необхідно знайти лише два коефіцієнти  $a_1$  та  $b_1$ .

Використовуючи формули (8.62)–(8.65) та (8.75), знаходимо, що

$$W_1(j\omega) = \frac{b_1}{a_1 j\omega + 1} = \frac{b_1 \cdot (1 - a_1 j\omega)}{(1 + a_1 j\omega) \cdot (1 - a_1 j\omega)}, \quad (8.76)$$

$$P_1(\omega) = \frac{b_1}{1 + a_1^2 \omega^2}, \quad (8.77)$$

$$Q_1(\omega) = -\frac{a_1 b_1 \omega}{1 + a_1^2 \omega^2}, \quad (8.78)$$

$$A_1(\omega) = \frac{b_1}{\sqrt{1 + a_1^2 \omega^2}}, \quad (8.79)$$

$$\varphi_1(\omega) = -\arctg(a_1 \omega). \quad (8.80)$$

Підставляючи у формулу (8.79) значення  $\omega_0$ ,  $A(\omega_0)$ ,  $\omega_1$ ,  $A(\omega_1)$  чи якісь інші дві пари з (8.73), складаємо систему двох рівнянь з двома невідомими  $a_1$ ,  $b_1$ . Розв'язавши цю систему чисельними методами, отримаємо оцінки коефіцієнтів  $a_1$ ,  $b_1$ .

А використовуючи значення  $\omega_0$ ,  $\varphi(\omega_0)$ ;  $\omega_1$ ,  $\varphi(\omega_1)$ , ...,  $\omega_n$ ,  $\varphi(\omega_n)$  із (8.74) та значення коефіцієнта  $a_1$ , знайденого на попередньому етапі, вичислимо оцінку

$$\sigma_1^2 = \sum_{i=0}^n (\varphi(\omega_i) + \arctg(a_1 \omega_i))^2 \quad (8.81)$$

суми квадратів відхилень експериментальних значень фазової частотної характеристики від її значень, що визначаються за вибраною моделлю (8.80).

Після цього висуваємо гіпотезу, що динаміка системи описується якоюсь іншою передаточною функцією і для її конкретного виду повторюємо всю описану вище процедуру.

У результаті отримаємо оцінку  $\sigma_2^2$ .

Якщо

$$\sigma_2^2 > \sigma_1^2, \quad (8.82)$$

то найкращу апроксимацію передаточної функції системи за експериментальними даними (8.73), (8.74) забезпечує формула (8.75). Якщо ж

$$\sigma_2^2 < \sigma_1^2, \quad (8.83)$$

то пошук потрібно продовжувати.

А далі звертаємо увагу на таке.

Ввівши умову (8.56), від перетворення за Лапласом, яке кожній неперервній функції, визначеній на часовій осі як оригінал, ставить у відповідність неперервну функцію на комплексній площині, ми, фактично, переходимо до перетворення Фур'є, яке відображає оригінали на модифіковану комплексну площину, оскільки реалізується не у вигляді (8.52), а у вигляді

$$W(j\omega) = \int_0^{\infty} g(t)e^{-j\omega t} dt, \quad (8.84)$$

що цілком правомірно, оскільки імпульсна перехідна функція  $g(t)$ , як показано в попередньому підрозділі, визначена лише для значень  $t \in [0, \infty)$ , а тому

$$\int_{-\infty}^0 g(t)e^{-j\omega t} dt = 0. \quad (8.85)$$

А модифікація комплексної площини полягає у тому, що кожна її точка задається не просто парою дійсних чисел, як при відображенні за Лапласом, а парою дійсних чисел, кожне з яких є значенням функцій кругової частоти.

При цьому вираз для оберненого оператора Фур'є не змінюється, тобто і у цьому випадку обернене перетворення Фур'є здійснюється за виразом (8.55).

Але практично обернене перетворення Фур'є спектральної функції, заданої у вигляді АФЧХ

$$W(j\omega) = A(\omega)e^{j\varphi(\omega)}, \quad (8.86)$$

простіше здійснювати не шляхом її інтегрування за виразом (8.55), а за алгоритмом

$$\{\omega_i, A(\omega_i)\} \rightarrow \{W_i(p)\} \rightarrow \{\sigma_i^2 \rightarrow \min\} \rightarrow \left[ W(p) = \frac{B(p)}{C(p)} \right] \rightarrow [D(p) = 0 \rightarrow \{p_k\}] \rightarrow \left[ \sum_k \frac{B(p_k)}{D'(p_k)} e^{p_k t} \right],$$

який використовує наведену вище процедуру пошуку передаточної функції  $W(p)$  за експериментально визначеними АЧХ  $A(\omega)$  та ФЧХ  $\varphi(\omega)$  та наведену у попередньому підрозділі процедуру еквівалентного подання оберненого оператора Лапласа однією з форм теореми розкладання.

*Резюмуючи викладене вище, можемо констатувати, що прямий оператор Фур'є реалізовує закон відображення неперервних функцій дійсного аргументу, заданого на числовій осі, у множину неперервних функцій комплексного аргументу, заданого на комплексній площині, координати точок якої теж є дійсними числами, але, на відміну від тих чисел, які задаються прямим оператором Лапласа, ці дійсні числа конкретизуються значеннями неперервних функцій кругової частоти. А обернений оператор Фур'є, як і обернений оператор Лапласа, діє навпаки, теж реалізуючи обернений процес перетворення вказаних множин одна в одну.*

### **8.3 Різницеві оператори та оператор Деча для відображення дискретних функцій, заданих на нульвимірній множині, в простір неперервних функцій**

Прямий оператор Лапласа, який нами розглянуто у першому підрозділі цього розділу, реалізовував закон відображення множини неперервних функцій дійсного аргументу, заданого на числовій осі, у множину неперервних функцій комплексного аргументу, заданого на комплексній площині, координати точок якої теж є дійсними числами. А обернений оператор Лапласа діяв навпаки, реалізуючи обернений процес перетворення вказаних множин одна в одну.

Прямий оператор Фур'є, який нами розглянуто у другому підрозділі цього розділу реалізовував закон відображення неперервних функцій дійсного аргументу, заданого на числовій осі, у множину неперервних функцій комплексного аргументу, заданого на комплексній площині, координати точок якої теж є дійсними числами, але, на відміну від тих чисел, які задаються прямим оператором Лапласа, ці дійсні числа конкретизуються значеннями неперервних функцій кругової частоти. А обернений оператор Фур'є, як і обернений оператор Лапласа, діяв навпаки, теж реалізуючи обернений процес перетворення вказаних множин одна в одну.

Але для аналізу процесів у дискретних системах необхідно використовувати апарат функцій дискретного аргументу, а тому математики задалися метою сконструювати такий оператор, який би зміг здійснити відображення множини функцій дискретного аргументу в множину функцій неперервного аргументу, а також сконструювати обернений оператор для оберненого відображення цих множин. Цю задачу розв'язав математик Деч, тож запропоновані ним прямий і обернений оператори заслуговують на те, щоб бути названими прямим і оберненим операторами Деча.

Але, перш ніж викласти матеріал, присвячений перетворенням множин функцій прямим та оберненим операторами Деча, згадаємо оператори, за допомогою яких із неперервних функцій неперервного аргументу синтезують функції дискретного аргументу.

Тож нехай  $f(t)$  є неперервною функцією неперервного аргументу, графік якої має вигляд, наведений на рис. 8.5, а.

Виберемо інтервал дискретності  $T$  і визначимо значення функції  $f(t)$  лише при значеннях аргументу  $t$ , кратних  $T$  (рис. 8.5, б).

У результаті цих дій отримаємо функцію  $f[kT]$  дискретного аргументу  $kT$  ( $k = 0, 1, 2, \dots$ ) (рис. 8.5, в), яку математики домовились називати решітчастою (або гратчастою).

Записується оператор відображення неперервної функції  $f(t)$  у решітчасту функцію  $f[kT]$  так:

$$f[kT] = \int_0^{\infty} f(t) \delta(t - kT) dt, \quad T > 0, \quad k = 0, 1, 2, \dots, \quad (8.87)$$

де  $\delta(t)$  – дельта-функція, визначення і властивості якої, згідно з виразами (8.35), (8.36), наведені у першому підрозділі цього розділу.

На практиці цей оператор реалізується у вигляді алгоритму, який виглядає так:

$$f[kT] = f(t)|_{t=kT}, \quad (8.88)$$

$$T > 0, \quad k = 0, 1, 2, \dots$$

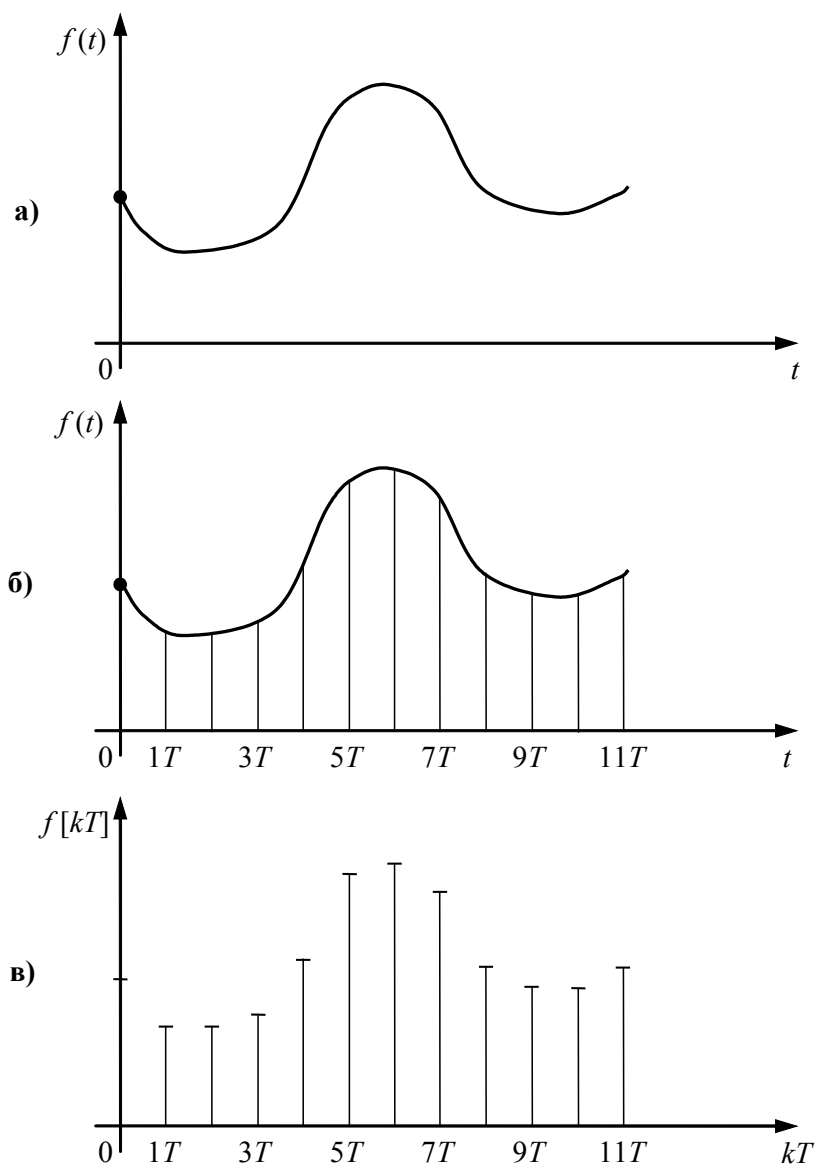


Рисунок 8.5 – Графіки неперервної  $f(t)$  і породженої нею решітчастої  $f[kT]$  функцій



Приклад 1. Нехай

$$f_1(t) = 2t - 4, \quad t \geq 0, \quad (8.89)$$

$$T = 0,5. \quad (8.90)$$

Згідно з алгоритмом (8.88) решітчаста функція  $f_1[kT]$ , яка є дискретним відображенням функції (8.89), за умови (8.90) має вигляд

$$f_1[kT] = f_1(t)|_{t=0,5k} = 2 \cdot [0,5k] - 4, \quad (8.91)$$

$$k = 0, 1, 2, \dots$$

Приклад 2. Нехай

$$f_2(t) = 3t^2 + 2, \quad t \geq 0, \quad (8.92)$$

$$T = 0,5.$$

Цій неперервній функції відповідає решітчаста, яка має вигляд

$$f_2[kT] = f_2(t)|_{t=0,5k} = 3 \cdot [0,5k]^2 + 2, \quad k = 0, 1, 2, \dots \quad (8.93)$$

Графіки решітчастих функцій (8.91), (8.93) наведені на рис. 8.6.

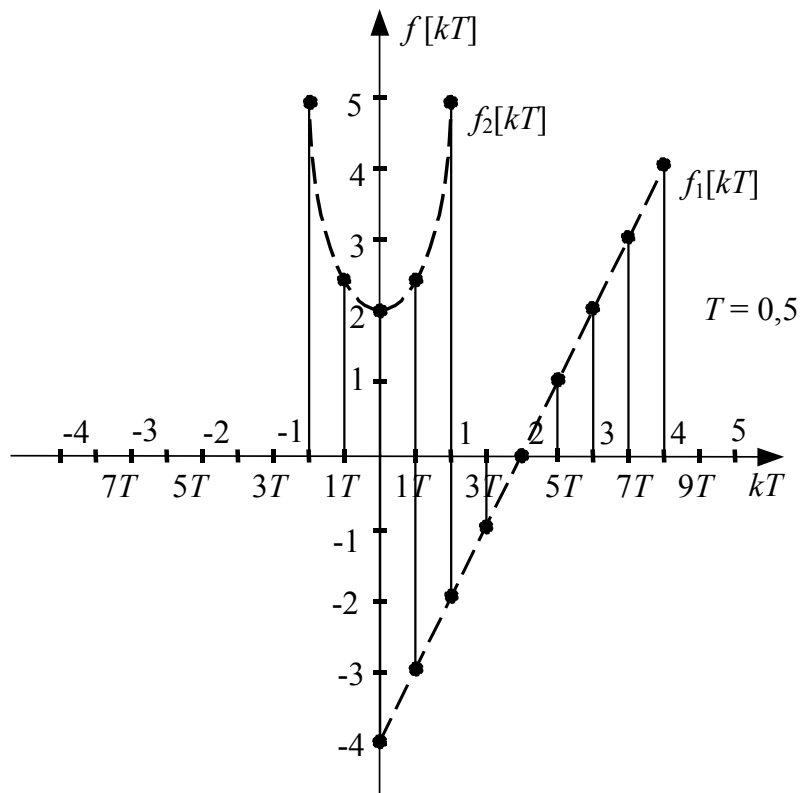


Рисунок 8.6 – Графіки решітчастих функцій  $f_1[kT] = 2[0,5k] - 4$  та  $f_2[kT] = 3[0,5k]^2 + 2$  при  $T = 0,5$

Із формул (8.91) та (8.93) легко бачити, що після вибору числового значення інтервалу дискретності  $T$  решітчаста функція  $f[kT]$  стає функцією лише аргументу  $k$ , а тому автори багатьох навчальних посібників і монографій записують її у вигляді не  $f[kT]$ , а  $f[k]$ . Роблячи так, слід пам'ятати, по-перше, що коефіцієнти у функціях  $f[k]$  і  $f[kT]$  не збігаються, а по-друге, що для відтворення неперервної функції  $f(t)$ , породженої оператором (8.87), необхідно порядковий номер  $k$ , який виступає як аргумент решітчастої функції  $f[k]$ , помножити на числове значення інтервалу дискретності  $T$  до виконання операції над ним. Для того, щоб не тримати цього у пам'яті, будемо зберігати період дискретності  $T$  в аргументі решітчастої функції  $f[kT]$ .

Як відомо з курсу математичного аналізу, швидкість зміни неперервної функції  $f(t)$  у кожній точці графіка характеризується значенням її похідної  $\frac{df}{dt}$  у цій точці, яка є неперервним оператором і числове значення якої у цій точці дорівнює тангенсу кута  $\alpha$  нахилу дотичної до графіка, проведеної через цю ж точку (рис. 8.7).

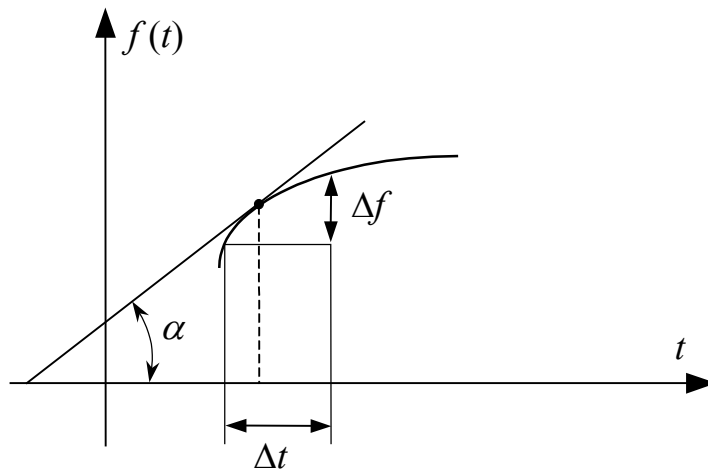


Рисунок 8.7 – Геометрична інтерпретація похідної  $\dot{f}(t)$  від функції  $f(t)$

Відомо, що похідна  $\frac{df}{dt}$  є оператором  $\dot{f}(t)$ , який визначається співвідношенням

$$\dot{f}(t) = \frac{df}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta f}{\Delta t}, \quad (8.94)$$

яке не може мати місця для решітчастих функцій.

Але оскільки сусідні значення аргументу  $kT$  решітчастої функції  $f[kT]$  відрізняються на  $T$ , то, сформувавши різницю

$$\begin{aligned} \Delta f[kT] &= f[(k+1)T] - f[kT], \\ k &= 0, 1, 2, \dots, \end{aligned} \quad (8.95)$$

ми можемо формально використати цю різницю  $\Delta f[kT]$  для характеристики швидкості зміни решітчастої функції  $f[kT]$  у точці з аргументом  $kT$ .

*Різницю  $\Delta f[kT]$  називають правою скінченною різницею першого порядку решітчастої функції  $f[kT]$  або правим різницеvim оператором першого порядку.*

За аналогією з (8.95) можна визначити праву скінченну різницю другого порядку

$$\Delta^2 f[kT] = \Delta f[(k+1)T] - \Delta f[kT], \quad (8.96)$$

$$k = 0, 1, 2, \dots$$

Очевидно, що ця різниця являє собою правий різницеvim оператор другого порядку.

Якщо у вираз (8.96) підставити значення  $\Delta f[kT]$  та  $\Delta f[(k+1)T]$ , визначені за формулою (8.95), отримаємо

$$\Delta^2 f[kT] = f[(k+2)T] - 2f[(k+1)T] + f[kT], \quad (8.97)$$

$$k = 0, 1, 2, \dots$$

Узагальнюючи (8.95) та (8.96), маємо

$$\Delta^n f[kT] = \Delta^{n-1} f[(k+1)T] - \Delta^{n-1} f[kT], \quad (8.98)$$

$$n = 1, 2, \dots; \quad k = 0, 1, 2, \dots$$

При розкритті формули (8.98) для конкретних  $n$  варто пам'ятати, що

$$\begin{cases} \Delta^0 f[kT] = f[kT], \\ \Delta^0 f[(k+1)T] = f[(k+1)T]. \end{cases} \quad (8.99)$$

*Відомо, що значення похідної  $\dot{f}(t)$  від неперервної функції  $f(t)$  у точці з аргументом  $t$  не залежить від того зліва чи справа наближається  $\Delta t$  до нуля (див. формулу (8.94)). Інша ситуація виникає при наближенні до точки з аргументом  $kT$  під час аналізу решітчастої функції  $f[kT]$ , а тому для характеристики швидкості її зміни крім правої скінченної різниці  $\Delta f[kT]$  першого порядку доцільно ввести ще й ліву скінченну різницю першого порядку за формулою*

$$\nabla f[kT] = f[kT] - f[(k-1)T], \quad (8.100)$$

$$k = 1, 2, 3, \dots$$

*Цю різницю називають лівим різницеvim оператором першого порядку.*

Нагадаємо, що символи « $\Delta$ » та « $\nabla$ » є грецькими літерами, відповідно, «дельта» та «набла».

На рис. 8.8 показано, наскільки суттєво можуть відрізнятися права  $\Delta f[kT]$  та ліва  $\nabla f[kT]$  скінченні різниці першого порядку решітчастої функції  $f[kT]$  при значенні аргументу  $kT = 4T$ .

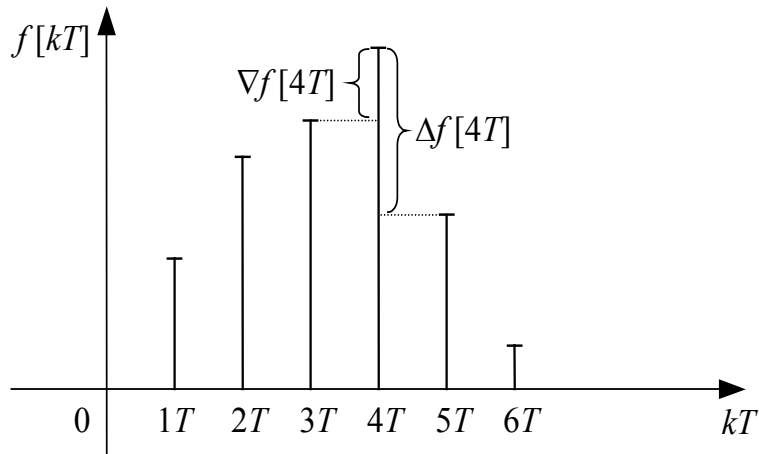


Рисунок 8.8 – Геометрична інтерпретація відмінності між правою  $\Delta f[kT]$  та лівою  $\nabla f[kT]$  скінченними різницями першого порядку у точці з аргументом  $kT = 4T$

За аналогією з (8.98) визначається і ліва скінченна різниця  $\nabla^n f[kT]$  порядку  $n$

$$\nabla^n f[kT] = \nabla^{n-1} f[kT] - \nabla^{n-1} f[(k-1)T], \quad (8.101)$$

$$n = 1, 2, \dots; \quad k = 1, 2, \dots,$$

де, як і в (8.99),

$$\begin{cases} \nabla^0 f[kT] = f[kT], \\ \nabla^0 f[(k-1)T] = f[(k-1)T]. \end{cases} \quad (8.102)$$

При  $n = 2$  з формули (8.101) отримаємо

$$\nabla^2 f[kT] = \nabla f[kT] - \nabla f[(k-1)T], \quad (8.103)$$

$$k = 2, 3, 4, \dots,$$

або, з урахуванням (8.100),

$$\begin{aligned}\nabla^2 f[kT] &= f[kT] - 2f[(k-1)T] + f[(k-2)T], \\ k &= 2, 3, 4, \dots\end{aligned}\quad (8.104)$$

Як приклад знайдемо  $\Delta f[kT]$ ,  $\Delta^2 f[kT]$  та  $\nabla f[kT]$ ,  $\nabla^2 f[kT]$  для решітчастих функцій  $f_1[kT]$ ,  $f_2[kT]$ , визначених формулами (8.91) та (8.93).

Для функції  $f_1[kT]$ :

$$\begin{aligned}\Delta f_1[kT] &= f_1[(k+1)T] - f_1[kT] = k+1 - 4 - (k-4) = 1, \\ k &= 0, 1, 2, \dots,\end{aligned}\quad (8.105)$$

$$\begin{aligned}\Delta^2 f_1[kT] &= f_1[(k+2)T] - 2f_1[(k+1)T] + f_1[kT] = \\ &= k+2 - 4 - 2((k+1) - 4) + (k-4) = \\ &= k+2 - 4 - 2k - 2 + 8 + k - 4 = 0, \\ k &= 0, 1, 2, \dots,\end{aligned}\quad (8.106)$$

$$\begin{aligned}\nabla f_1[kT] &= f_1[kT] - f_1[(k-1)T] = k - 4 - (k-1-4) = 1, \\ k &= 1, 2, 3, \dots,\end{aligned}\quad (8.107)$$

$$\begin{aligned}\nabla^2 f_1[kT] &= f_1[kT] - 2f_1[(k-1)T] + f_1[(k-2)T] = \\ &= k - 4 - 2((k-1) - 4) + (k-2-4) = \\ &= k - 4 - 2k + 2 + 8 + k - 2 - 4 = 0, \\ k &= 2, 3, 4, \dots\end{aligned}\quad (8.108)$$

Для функції  $f_2[kT]$ :

$$\begin{aligned}\Delta f_2[kT] &= f_2[(k+1)T] - f_2[kT] = 0,75(k+1)^2 + 2 - (0,75k^2 + 2) = \\ &= 0,75k^2 + 1,5k + 0,75 + 2 - 0,75k^2 - 2 = 1,5k + 0,75, \\ k &= 0, 1, 2, \dots,\end{aligned}\quad (8.109)$$

$$\begin{aligned}\Delta^2 f_2[kT] &= f_2[(k+2)T] - 2f_2[(k+1)T] + f_2[kT] = \\ &= 0,75(k+2)^2 + 2 - 2(0,75(k+1)^2 + 2) + 0,75k^2 + 2 = \\ &= 0,75k^2 + 3k + 3 + 2 - 1,5k^2 - 3k - 1,5 - 4 + 0,75k^2 + 2 = 1,5, \\ k &= 0, 1, 2, \dots,\end{aligned}\quad (8.110)$$

$$\begin{aligned}\nabla f_2[kT] &= f_2[kT] - f_2[(k-1)T] = 0,75k^2 + 2 - (0,75(k-1)^2 + 2) = \\ &= 0,75k^2 + 2 - 0,75k^2 + 1,5k - 0,75 - 2 = 1,5k - 0,75, \quad (8.111) \\ &k = 1, 2, 3, \dots,\end{aligned}$$

$$\begin{aligned}\nabla^2 f_2[kT] &= f_2[kT] - 2f_2[(k-1)T] + f_2[(k-2)T] = \\ &= 0,75k^2 + 2 - 2(0,75(k-1)^2 + 2) + 0,75(k-2)^2 + 2 = \\ &= 0,75k^2 + 2 - 1,5k^2 + 3k - 1,5 - 4 + 0,75k^2 - 3k + 3 + 2 = 1,5, \quad (8.112) \\ &k = 2, 3, 4, \dots\end{aligned}$$

Як видно з отриманих результатів, скінченні різниці першого порядку від лінійної решітчастої функції є константами, а другого порядку дорівнюють нулю. Нагадаємо, що і перша похідна від лінійної неперервної функції є константою, а друга похідна теж дорівнює нулю. Така ж аналогія спостерігається і між решітчастою та неперервною квадратичними функціями.

Оскільки аналогом похідної  $l$ -го порядку, де  $l = 1, 2, \dots, n$ , для решітчастої функції  $f[kT]$  є права  $\Delta^l f[kT]$  та ліва  $\nabla^l f[kT]$  скінченні різниці того ж порядку, то для решітчастих функцій  $x[kT]$  та  $y[kT]$  можна сконструювати рівняння в скінченних різницях  $\Delta^l y[kT]$  ( $l = 0, 1, 2, \dots, n$ ),  $\Delta^q x[kT]$  ( $q = 0, 1, 2, \dots, m$ ) або  $\nabla^l y[kT]$  ( $l = 0, 1, 2, \dots, n$ ),  $\nabla^q x[kT]$  ( $q = 0, 1, 2, \dots, m$ ), яке буде відображенням диференціального рівняння, синтезованого відносно функцій  $y(t)$ ,  $x(t)$  та їх похідних  $y^{(l)}(t)$  ( $l = 1, 2, \dots, n$ ),  $x^{(q)}(t)$  ( $q = 1, 2, \dots, m$ ).

Покажемо на прикладі як можна побудувати дискретне відображення диференціального рівняння.

Нехай маємо диференціальне рівняння

$$\frac{d^2 y}{dt^2} + 4 \frac{dy}{dt} + 2y = 2x \quad (8.113)$$

з початковими умовами

$$\begin{cases} y(0) = 0, \\ y'(0) = 1. \end{cases} \quad (8.114)$$

Нехай  $T$  – період дискретності функцій  $x(t)$ ,  $y(t)$  для породження решітчастих функцій  $x[kT]$ ,  $y[kT]$ .

Нагадаємо, що аргументи  $t$  і  $kT$  пов'язані між собою співвідношенням

$$t_k = kT. \quad (8.115)$$

Зрозуміло, що похідній  $\frac{dy}{dt}$ , яку маємо у рівнянні (8.113), в її дискретному відображенні відповідатиме вираз

$$\frac{\Delta y[kT]}{\Delta t_k} = \frac{\Delta y[kT]}{(k+1)T - kT} = \frac{\Delta y[kT]}{T}, \quad (8.116)$$

$$k = 0, 1, 2, \dots,$$

а другій похідній – вираз

$$\frac{\Delta^2 y[kT]}{(\Delta t_k)^2} = \frac{\Delta^2 y[kT]}{T^2}, \quad (8.117)$$

$$k = 0, 1, 2, \dots$$

Підставляючи (8.115)–(8.117) у рівняння (8.113), (8.114), отримаємо

$$\frac{\Delta^2 y[kT]}{T^2} + 4 \frac{\Delta y[kT]}{T} + 2y[kT] = 2x[kT], \quad (8.118)$$

$$k = 0, 1, 2, \dots,$$

$$\begin{cases} y[0] = 0, \\ \frac{\Delta y[0]}{T} = 1, \end{cases} \quad (8.119)$$

або

$$\Delta^2 y[kT] + 4T \Delta y[kT] + 2T^2 y[kT] = 2T^2 x[kT], \quad (8.120)$$

$$k = 0, 1, 2, \dots,$$

$$\begin{cases} y[0] = 0, \\ \Delta y[0] = T. \end{cases} \quad (8.121)$$

Зрозуміло, що чим меншим буде значення періоду дискретності  $T$ , тим менше будуть відрізнятися один від одного значення  $t_k$  та  $t_{k+1}$  або  $t_{k-1}$  і, як наслідок, ближчими до розв'язку  $y(t)$  диференціального рівняння (8.113) у точках  $t_k$  будуть розв'язки  $y[kT]$  рівняння в скінченних різницях (8.120).

Найбільш просто рівняння (8.120) розв'язується шляхом перетворення його в різницеве, яке містить у собі не скінченні різниці, а значення решітчастих функцій, взятих при декількох значеннях аргументу.

Для здійснення цього перетворення потрібно у рівняннях (8.120), (8.121) замість скінченних різниць  $\Delta y[kT]$ ,  $\Delta^2 y[kT]$  підставити їхні значення з формул (8.95), (8.97).

Здійснивши це, отримаємо

$$y[(k+2)T] - 2y[(k+1)T] + y[kT] + 4T(y[(k+1)T] - y[kT]) + 2T^2 y[kT] = 2T^2 x[kT],$$

$$k = 0, 1, 2, \dots,$$

$$\begin{cases} y[0] = 0, \\ y[1 \cdot T] - y[0] = T, \end{cases}$$

або

$$y[(k+2)T] = 2T^2 x[kT] - (4T - 2)y[(k+1)T] - (2T^2 - 4T + 1)y[kT],$$

$$k = 0, 1, 2, \dots,$$

$$\begin{cases} y[0] = 0, \\ y[1 \cdot T] = T. \end{cases}$$

Вибравши конкретне значення  $T$  та сформувавши з заданої функції  $x(t)$  решітчасту  $x[kT]$ , шляхом підстановки по черзі  $k=0$ , потім  $k=1$  і так далі у рівняння (8.124) отримаємо стільки значень розв'язку  $y[kT]$ , скільки нам треба.

У першому підрозділі цього розділу було показано, що лінійні диференціальні рівняння можна розв'язувати не лише безпосередньо, але й завдяки попередньому їх перетворенню в алгебраїчні рівняння комплексної змінної із застосуванням прямого та оберненого операторів Лапласа.

Покажемо, що аналогічну процедуру можна застосувати і до різницевих рівнянь.

Нехай  $y[kT]$  ( $k=0, 1, 2, \dots$ ) є решітчастою функцією, породженою неперервною функцією-оригіналом  $y(t)$  з періодом дискретності  $T$ , для якої, як відомо, справедливо:

$$y(t) = \begin{cases} y(t), & t \geq 0, \\ 0, & t < 0, \end{cases}$$

$$Y(p) = \int_0^{\infty} y(t) e^{-pt} dt.$$

Оскільки

$$\begin{cases} t_k = kT, \\ dt \cong T, \end{cases}$$

то інтеграл в операторі Лапласа (8.127) можна з деякою похибкою, яка зменшується зі зменшенням значення  $T$ , записати у вигляді суми

$$Y^*(p) \cong \sum_{k=0}^{\infty} y(t_k) e^{-pt_k} T,$$



або

$$Y^*(p) \cong T \sum_{k=0}^{\infty} y[kT] e^{-pkT}. \quad (8.130)$$

Введемо нову комплексну змінну  $q$  у такий спосіб

$$q = pT. \quad (8.131)$$

і запишемо суму, яка стоїть у правій частини виразу (8.130), у вигляді

$$Y(q) = \sum_{k=0}^{\infty} y[kT] e^{-qk}. \quad (8.132)$$

Формула (8.132) задає дискретний оператор Лапласа, який решітчасту функцію  $y[kT]$  відображає у функцію  $Y(q)$  масштабовану періодом дискретності  $T$  до значення  $q$  комплексної змінної  $p$ .

**Приклад.** У попередньому розділі показано, що зображенням за Лапласом експоненціальної функції

$$f(t) = e^{-\alpha t} \quad (8.133)$$

є функція

$$F(p) = \frac{1}{p + \alpha}. \quad (8.134)$$

Знайдемо дискретне зображення за Лапласом решітчастого аналога

$$f[kT] = e^{-\alpha kT} \quad (8.135)$$

цієї експоненти, використавши формулу (8.132).

$$\begin{aligned} F(q) &= \sum_{k=0}^{\infty} f[kT] e^{-qk} = \sum_{k=0}^{\infty} e^{-\alpha kT} \cdot e^{-qk} = \sum_{k=0}^{\infty} e^{-(\alpha T + q)k} = \\ &= 1 + e^{-(\alpha T + q)} + e^{-2(\alpha T + q)} + \dots = \\ &= \frac{1}{1 - e^{-(\alpha T + q)}} = \frac{1}{1 - e^{-\alpha T} \cdot e^{-q}} = \frac{e^q}{e^q - e^{-\alpha T}}. \end{aligned} \quad (8.136)$$

Звертаємо увагу на те, що у перетвореннях (8.136) використано формулу суми членів нескінченно спадної геометричної прогресії зі знаменником  $e^{-(\alpha T + q)}$ .

Порівнюючи формули (8.133) та (8.136) бачимо, що зображення неперервної експоненти, отримане за допомогою прямого оператора Лапласа, є набагато простішим за своєю структурою, аніж зображення її решітчастого аналога за допомогою дискретного оператора Лапласа. Для більш складних неперервних функцій громіздкість зображень їх решітчастих аналогів суттєво зростає.

*Ось чому нині відображення решітчастих функцій за допомогою дискретного оператора Лапласа повністю витіснене з практики аналізу дискретних систем відображенням на основі більш простого Z-оператора, запропонованого Дечем і отриманого ним заміною*

$$z = e^q \quad (8.137)$$

у формулі (8.132), що трансформувало її до вигляду

$$Z \{y[kT]\} = Y(z) = \sum_{k=0}^{\infty} y[kT] z^{-k} \quad (8.138)$$

*Оператор Деча (8.138) відображає решітчасті функції  $y[kT]$ , визначені на нульвимірному просторі дискретних значень аргументу  $kT$ , у точки круга одиничного радіуса на комплексній площині  $z$ , кожній з яких відповідає неперервна функція  $Y(z)$ .*

Прямою підстановкою виразу (8.137) у (8.136) знайдемо, що Z-перетворення решітчастої експоненти має вигляд

$$F(z) = \frac{z}{z - e^{-\alpha T}}. \quad (8.139)$$

Порівнюючи (8.139) з (8.134), бачимо, що вони мають однакову структуру, а саме: і у першому, і у другому випадках маємо дробово-раціональні функції, які є відношеннями поліномів у першому випадку за степенями  $p$ , а у другому – за степенями  $z$ .

Слід пам'ятати, що  $e^{-\alpha T}$  за заданих  $\alpha$  і  $T$  – це є число, наприклад, при  $\alpha = 1$  і  $T = 2$  – це 0,135.

Як ще один приклад знайдемо Z-перетворення одиничної решітчастої функції  $1[kT]$ ,  $k = 0, 1, \dots$

Оскільки одиничну функцію  $1(t)$  можна отримати з експоненти  $e^{-\alpha T}$ , поклавши  $\alpha = 0$ , то одинична решітчаста функція  $1[kT]$  є теж окремим випадком решітчастої експоненти при  $\alpha = 0$ . А тому для отримання Z-перетворення  $1[kT]$  достатньо покласти у формулі (8.139)  $\alpha = 0$ . Таким чином

$$Z \{1[kT]\} = 1(z) = \frac{z}{z - 1}. \quad (8.140)$$

У справедливості формули (8.140) легко пересвідчитись і за допомогою виразу (8.138) для оператора Деча, підставляючи в нього  $y[kT] = 1[kT]$  і знаходячи суму членів нескінченно спадної геометричної прогресії зі знаменником  $z^{-1}$ .

Оскільки оператор Деча (8.138) отримано з лінійного оператора Лапласа шляхом двох лінійних підстановок, то всі властивості оператора Лапласа справедливі і для Z-оператора, наприклад

$$Z \{f_1[kT] + f_2[kT]\} = Z \{f_1[kT]\} + Z \{f_2[kT]\}, \quad (8.141)$$

$$Z \{\alpha f_1[kT]\} = \alpha Z \{f_1[kT]\}, \quad (8.142)$$

де  $\alpha$  — число, значення якого не залежить від аргументу  $kT$  функції  $f_1[kT]$ .

Під час відображення різницевих рівнянь з використанням Z-оператора доводиться знаходити Z-зображення зсунутих вправо чи вліво значень решітчастих функцій, відносно яких ці рівняння складені, а тому побудуємо алгоритми їх знаходження.

Спочатку розглянемо зсув вправо

$$Z \{y[(k+1)T]\} = \sum_{k=0}^{\infty} y[(k+1)T] z^{-k}. \quad (8.143)$$

Нехай  $k+1 = m$ , тоді

$$\begin{aligned} Z \{y[(k+1)T]\} &= \sum_{m=1}^{\infty} y[mT] z^{-(m-1)} = z \cdot \sum_{m=1}^{\infty} y[mT] z^{-m} = \\ &= z \cdot \left( \sum_{m=0}^{\infty} y[mT] z^{-m} - y[0] \right) = z \cdot (Y(z) - y[0]). \end{aligned} \quad (8.144)$$

Аналогічно

$$Z \{y[(k+2)T]\} = \sum_{k=0}^{\infty} y[(k+2)T] z^{-k}. \quad (8.145)$$

Тепер нехай  $m = k+2$ , тоді

$$\begin{aligned} Z \{y[(k+2)T]\} &= \sum_{m=2}^{\infty} y[mT] z^{-(m-2)} = \\ &= z^2 \cdot \left( \sum_{m=0}^{\infty} y[mT] - y[0] - y[1 \cdot T] z^{-1} \right) = \\ &= z^2 \cdot (Y(z) - y[0] - y[1 \cdot T] z^{-1}). \end{aligned} \quad (8.146)$$

Узагальнюючи (8.144) та (8.146), отримуємо для довільного  $n$

$$Z \{y[(k+n)T]\} = z^n \cdot (Y(z) - y[0] - y[1 \cdot T]z^{-1} - \dots - y[(n-1)T]z^{-(n-1)}). \quad (8.147)$$

Тепер розглянемо зсув вліво

$$Z \{y[(k-1)T]\} = \sum_{k=0}^{\infty} y[(k-1)T] z^{-k}. \quad (8.148)$$

Нехай  $k-1 = m$ , тоді

$$\begin{aligned} Z \{y[(k-1)T]\} &= \sum_{m=-1}^{\infty} y[mT] z^{-(m+1)} = \\ &= z^{-1} \cdot \left( y[-1 \cdot T] \cdot z + \sum_{m=0}^{\infty} y[mT] z^{-m} \right) = z^{-1} \cdot (0 + Y(z)) = z^{-1} \cdot Y(z). \end{aligned} \quad (8.149)$$

Виводячи формулу (8.149), використано властивість оригіналу дорівнювати нулю за від'ємних значеннях аргументу.

Аналогічно

$$Z \{y[(k-2)T]\} = \sum_{k=0}^{\infty} y[(k-2)T] z^{-k}. \quad (8.150)$$

Нехай  $k-2 = m$ , тоді

$$\begin{aligned} Z \{y[(k-2)T]\} &= \sum_{m=-2}^{\infty} y[mT] z^{-(m+2)} = \\ &= z^{-2} \cdot \left( y[-2 \cdot T] \cdot z^2 + y[-1 \cdot T] \cdot z^1 + \sum_{m=0}^{\infty} y[mT] z^{-m} \right) = \\ &= z^{-2} \cdot (0 + 0 + Y(z)) = z^{-2} \cdot Y(z). \end{aligned} \quad (8.151)$$

Узагальнюючи (8.149) та (8.151), отримуємо для довільного  $n$

$$Z \{y[(k-n)T]\} = z^{-n} \cdot Y(z). \quad (8.152)$$

Порівнюючи вирази (8.147) і (8.152), приходимо до висновку, що з використанням  $Z$ -оператора доцільніше розв'язувати різницеві рівняння, які конструюються з диференціальних шляхом застосування лівих скінченних різниць.

Якщо ж різницеве рівняння сконструйоване з диференціального за допомогою правих скінченних різниць, то його доцільніше розв'язувати рекурентною (крок за кроком) процедурою, аналогічною (8.124).

На конкретному прикладі розглянемо процедуру  $Z$ -перетворення різницевого рівняння та алгоритм його розв'язання у цій області за допомогою оберненого  $Z^{-1}$  оператора Деча, для якого справедливим є те, що  $Z Z^{-1} = I$ , де  $I$  – одиничний оператор.

Нехай різницеве рівняння має вигляд

$$y[kT] + 2y[(k-1)T] + 4y[(k-2)T] = 2x[kT] - x[(k-1)T], \quad (8.153)$$

$$k = 2, 3, 4, \dots,$$

з початковими умовами

$$\begin{cases} y[-1 \cdot T] = 0, \\ y[-2 \cdot T] = 0. \end{cases} \quad (8.154)$$

Знайдемо  $Z$ -перетворення рівняння (8.153) за початкових умов (8.154).

Записуючи, спочатку формально

$$Z \{y[kT] + 2y[(k-1)T] + 4y[(k-2)T]\} = Z \{2x[kT] - x[(k-1)T]\} \quad (8.155)$$

та враховуючи лінійні властивості  $Z$ -перетворення (8.141), (8.142), отримаємо

$$\begin{aligned} Z \{y[kT]\} + 2Z \{y[(k-1)T]\} + 4Z \{y[(k-2)T]\} = \\ = 2Z \{x[kT]\} - Z \{x[(k-1)T]\}. \end{aligned} \quad (8.156)$$

Із виразу (8.156), з урахуванням (8.138), (8.149), (8.151), впливає:

$$Y(z) + 2z^{-1}Y(z) + 4z^{-2}Y(z) = 2X(z) - z^{-1}X(z), \quad (8.157)$$

або

$$Y(z) = \frac{2 - z^{-1}}{1 + 2z^{-1} + 4z^{-2}} X(z), \quad (8.158)$$

або

$$Y(z) = W(z) X(z), \quad (8.159)$$

де

$$W(z) = \frac{2 - z^{-1}}{1 + 2z^{-1} + 4z^{-2}}. \quad (8.160)$$

– дискретна передаточна функція динамічної системи, математичну модель якої в дискретному часі можна подати різницеvim рівнянням (8.153).

Нехай  $x[kT] = 1[kT]$ , тоді, згідно з (8.140),

$$X(z) = \frac{z}{z-1} = \frac{1}{1-z^{-1}}. \quad (8.161)$$

Підставляючи вираз (8.161) у (8.158), отримаємо

$$Y(z) = \frac{2 - z^{-1}}{(1 + 2z^{-1} + 4z^{-2}) \cdot (1 - z^{-1})} = \frac{2 - z^{-1}}{1 + z^{-1} + 2z^{-2} - 4z^{-3}}. \quad (8.162)$$

**Формально, для знаходження оригіналу за відомим зображенням необхідно використовувати обернений оператор, тобто**

$$y[kT] = Z^{-1}\{Y(z)\} \quad (8.163)$$

Для того, що усвідомити, якою є найпростіша форма оберненого оператора Деча  $Z^{-1}\{Y(z)\}$ , розпишемо спочатку суму у формулі Z-перетворення в (8.138), в результаті чого отримаємо

$$Y(z) = y[0] + y[1 \cdot T] z^{-1} + y[2 \cdot T] z^{-2} + y[3 \cdot T] z^{-3} + \dots \quad (8.164)$$

Із виразу (8.164) бачимо, що коефіцієнти при степенях  $z^{-k}$  ( $k = 0, 1, \dots$ ) і є значеннями решітчастої функції  $y[kT]$ , яка є оригіналом для Z-зображення  $Y(z)$ .

А до форми (8.164) дробово-раціональна функція, що стоїть у правій частині виразу (8.162), яка є відношенням двох багаточленів, приводиться діленням чисельника на знаменник і взяттям такої кількості членів ряду, яка необхідна за умовами розв'язання.

Застосуємо описану процедуру до виразу (8.162).

$$\begin{array}{r}
 2 - z^{-1} \\
 (-) \quad 2 + 2z^{-1} + 4z^{-2} - 8z^{-3} \\
 \hline
 -3z^{-1} - 4z^{-2} + 8z^{-3} \\
 (-) \quad -3z^{-1} - 3z^{-2} - 6z^{-3} + 12z^{-4} \\
 \hline
 -z^{-2} + 14z^{-3} - 12z^{-4} \\
 (-) \quad -z^{-2} - z^{-3} - 2z^{-4} + 4z^{-5} \\
 \hline
 15z^{-3} - 10z^{-4} - 4z^{-5} \\
 (-) \quad 15z^{-3} + 15z^{-4} + 30z^{-5} - 60z^{-6} \\
 \hline
 -25z^{-4} - 34z^{-5} + 60z^{-6} \\
 (-) \quad -25z^{-4} - 25z^{-5} - 50z^{-6} + 100z^{-7} \\
 \hline
 -9z^{-5} + 110z^{-6} - 100z^{-7}.
 \end{array}$$

Зупинившись на цьому кроці, вираз для  $Y(z)$ , який задано виразом (8.162), можна записати так

$$Y(z) = 2 - 3z^{-1} - z^{-2} + 15z^{-3} - 25z^{-4} + \frac{-9z^{-5} + 110z^{-6} - 100z^{-7}}{1 + z^{-1} + 2z^{-2} - 4z^{-3}}. \quad (8.165)$$

Порівнюючи вирази (8.164) та (8.165), знаходимо п'ять перших значень решітчастої функції  $y[kT]$ , яка є оригіналом для Z-зображення  $Y(z)$ , визначеного виразом (8.162).

$$\begin{aligned} y[0] &= 2, \\ y[1 \cdot T] &= -3, \\ y[2 \cdot T] &= -1, \\ y[3 \cdot T] &= 15, \\ y[4 \cdot T] &= -25. \end{aligned} \tag{8.166}$$

Оскільки процес ділення багаточлена на багаточлен легко алгоритмізується, то на комп'ютері дуже швидко можна отримати яку завгодно сукупність значень решітчастого оригіналу на тому відрізку дискретних значень аргументу, який нас цікавить.

*На цьому прикладі ми показали, яким чином кожній неперервній функції  $Y(z)$ , образом якої є точка на комплексній площині в крузі одиничного радіуса, можна поставити у відповідність решітчасту функцію  $y[kT]$ ,  $k = 0, 1, 2, \dots$  дискретного аргументу на осі дійсних чисел, тобто показали, як практично реалізується обернений оператор Деча.*

#### 8.4 Авторегресійні оператори в задачах відображення часових рядів

Під час системного аналізу динамічних процесів, що мають випадковий характер, та створенні інформаційних технологій, придатних для реалізації цього аналізу, значна роль відводиться прогнозуванню розвитку цих процесів у часі.

На сьогоднішній день найбільш ефективними математичними моделями, за допомогою яких можна прогнозувати розвиток процесів, є ті, під час побудови яких використовуються часові ряди.

Нагадаємо, що **часовий ряд** – це сукупність значень випадкового процесу, взятих через однакові проміжки часу  $t$ . Позначимо цю сукупність символом  $z_t$ .

Фактично  $z_t$  – це є дискретний у часі випадковий процес.

Задача прогнозу полягає в тому, що, знаючи значення процесу у момент  $t$ , необхідно спрогнозувати його значення в момент  $t+l$ , де  $l$  – час упередження. Щоб відрізнити прогнозне значення процесу від дійсного, дійсне значення часового ряду у момент  $t+l$  позначають символом  $z_{t+l}$ , а прогнозне – символом  $z_t(l)$ .

Зрозуміло, що точно спрогнозувати значення випадкового процесу, яким є часовий ряд, неможливо в принципі, а тому прогноз здійснюють, домагаючись мінімуму якогось функціонала, вибраного за критерій адекватності прогнозної моделі.

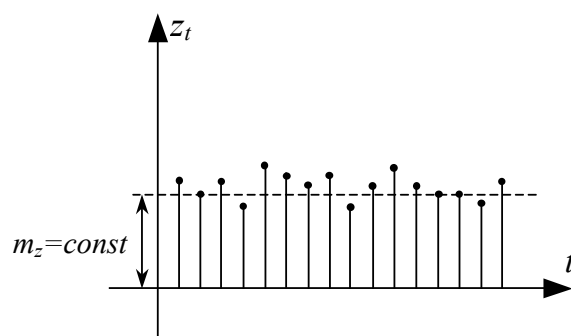
Якщо значення  $t$  є малим (1, 2 кроки), то одним із таких критеріїв може бути дисперсія відхилення  $z_t(l)$  від  $z_{t+l}$ , яка має бути для оптимальної моделі прогнозу мінімальною, тобто

$$E \left\{ \left( z_{t+l} - z_t(l) \right)^2 \right\} \rightarrow \min, \tag{8.167}$$

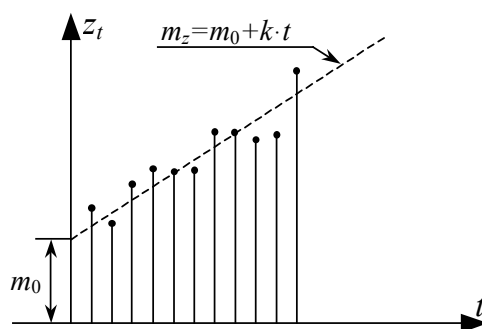
де  $E$  – символ операції математичного очікування.



Як і будь-який інший випадковий процес, часовий ряд  $z_t$  може бути стаціонарним (рис. 8.9, а) або нестаціонарним (рис. 8.9, б).



а)



б)

Рисунок 8.9 – Графіки реалізації стаціонарного (а) та нестаціонарного (б) часових рядів

Для стаціонарного часового ряду характерною є рівновага його значень  $z_t$  біля середнього значення  $m_z$ , яке є константою, як показано на рис. 8.9, а.

Для нестаціонарного часового ряду ковзне середнє значення  $m_z(t)$  процесу є функцією часу  $t$ , як показано на рис. 8.9, б.

**Введемо низку корисних операторів**, які будуть потрібними у подальшому.

**1. Оператор  $B$  зсуву назад на одну одиницю часу**

$$z_{t-1} = Bz_t. \quad (8.168)$$

Зрозуміло, що, згідно з виразом (8.168), справедливим є і вираз

$$z_{t-2} = Bz_{t-1}. \quad (8.169)$$

Підставляючи значення  $z_{t-1}$  із виразу (8.168) в (8.169), отримаємо

$$z_{t-2} = B(Bz_t) = B^2 z_t. \quad (8.170)$$

Узагальнюючи вираз (8.170), маємо

$$z_{t-m} = B^m z_t. \quad (8.171)$$

**2. Оператор  $F$  зсуву вперед на одну одиницю часу**

$$z_{t+1} = Fz_t. \quad (8.172)$$

Зрозуміло, що, згідно з виразом (8.172), справедливим є і вираз

$$z_{t+2} = Fz_{t+1}. \quad (8.173)$$

Підставляючи значення  $z_{t+1}$  із виразу (8.172) в (8.173), отримаємо

$$z_{t+2} = F(Fz_t) = F^2 z_t. \quad (8.174)$$

Узагальнюючи вираз (8.174), маємо

$$z_{t+m} = F^m z_t. \quad (8.175)$$

**3. Різницевий оператор  $\nabla$  із зсувом назад на одну одиницю часу**

$$\nabla z_t = z_t - z_{t-1}. \quad (8.176)$$

Підставляючи значення  $z_{t-1}$  із виразу (8.168) в (8.176), отримаємо

$$\nabla z_t = z_t - z_{t-1} = z_t - Bz_t = (1 - B)z_t. \quad (8.177)$$

Із виразу (8.177) випливає, що

$$\nabla = 1 - B. \quad (8.178)$$

**4. Різницевий оператор  $\Delta$  із зсувом вперед на одну одиницю часу**

$$\Delta z_t = z_{t+1} - z_t. \quad (8.179)$$

Підставляючи значення  $z_{t+1}$  із виразу (8.172) в (8.179), отримаємо

$$\Delta z_t = Fz_t - z_t = (F - 1)z_t. \quad (8.180)$$

Із виразу (8.180) випливає, що

$$\Delta = F - 1. \quad (8.181)$$

### 5. Оператор суми $S$

$$Sz_t = z_t + z_{t-1} + z_{t-2} + \dots = \sum_{j=0}^{\infty} z_{t-j}. \quad (8.182)$$

Підставляючи значення  $z_{t-m}$  із виразу (8.171) у (8.182), отримаємо

$$Sz_t = (1 + B + B^2 + \dots)z_t = \frac{1}{1-B}z_t = (1-B)^{-1}z_t. \quad (8.183)$$

Звертаємо увагу на те, що під час виведення співвідношення (8.183) було використано формулу суми членів нескінченно спадної геометричної прогресії зі знаменником  $B$ , який, за умови розгляду його як числа та умови збіжності ряду (8.182), має бути меншим одиниці.

Із співвідношень (8.178) та (8.183) випливає, що

$$S z_t = \nabla^{-1} z_t, \quad (8.184)$$

або

$$S = \nabla^{-1}. \quad (8.185)$$

**Отже, оператор суми є оберненим різницевому оператору зі зсувом назад.**

Далі нагадаємо, що **послідовність некорельованих і розподілених нормально випадкових імпульсів  $a_t$  з нульовим середнім значенням і дисперсією**

$$\sigma_a^2 = const \quad (8.186)$$

**називають дискретним білим шумом.**

Спробуємо використати імпульси білого шуму  $a_t$  для побудови моделі часового ряду  $z_t$  у такий спосіб

$$z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots, \quad (8.187)$$

де  $\mu$  – рівень відліку (середнє значення) часового ряду  $z_t$ , а  $\psi_k$ ,  $k = 1, 2, \dots$  – коефіцієнти ваги імпульсів білого шуму, з якими вони входять до суми (8.187).

Здійснимо операцію центрування часового ряду  $z_t$ , віднявши від  $z_t$  середнє значення  $\mu$ . Для центрованого часового ряду

$$\tilde{z}_t = z_t - \mu. \quad (8.188)$$

Із виразу (8.187) отримаємо

$$\tilde{z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (8.189)$$

Використовуючи для імпульсів  $a_{t-m}$  співвідношення (8.171), із (8.189) маємо

$$\tilde{z}_t = a_t + \psi_1 B a_t + \psi_2 B^2 a_t + \dots = \left(1 + \psi_1 B + \psi_2 B^2 + \dots\right) \cdot a_t. \quad (8.190)$$

Позначимо

$$\Psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \quad (8.191)$$

З урахуванням (8.191) співвідношення (8.190) можна записати таким чином

$$\tilde{z}_t = \Psi(B) a_t. \quad (8.192)$$

**Вираз  $\Psi(B)$  у вигляді (8.191) являє собою оператор фільтра, який** перетворює послідовність імпульсів білого шуму  $a_t$  у часовий ряд з заданими властивостями (рис. 8.10), тобто, **ставить у відповідність дискретній стохастичній функції з однієї нульвимірної множини іншу дискретну стохастичну функцію з іншої нульвимірної множини.**

Коефіцієнти оператора фільтра  $\psi_k$ ,  $k = 1, 2, \dots$  підбираються у процедурі мінімізації критерію (8.167) при  $l = 0$ .

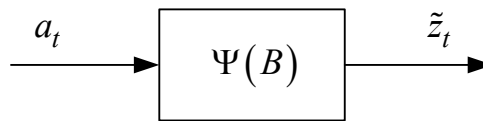


Рисунок 8.10 – Структурна схема лінійного фільтра

В моделі лінійного фільтра (8.189) значення часового ряду  $z_t$  визначаються через зважену суму поточного та попередніх імпульсів білого шуму  $a_t$ .

Характерною особливістю оператора  $\Psi(B)$  фільтра, заданого виразом (8.191), є те, що теоретично він має нескінченну кількість членів, що створює певні незручності у разі його практичного використання.

Тому цікавою виявилась пропозиція будувати модель часового ряду  $z_t$  на основі скінченної множини потужністю  $q$  зважених імпульсів білого шуму  $a_t$  у вигляді

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \quad (8.193)$$

Оскільки у співвідношенні (8.193) використовується  $q$  попередніх значень білого шуму  $a_{t-i}$ ,  $i = \overline{1, q}$ , які зважено віднімаються від поточного імпульсу  $a_t$ , то це співвідношення фактично задає «ковзне середнє» яке «зміщується» вздовж послідовності  $a_t$  зі зростанням  $t$ , утримуючи в собі при «зміщенні» одну і ту ж кількість членів.

Застосовуючи до імпульсів  $a_{t-m}$  ідеологію співвідношення (8.171), із виразу (8.193) отримуємо

$$\tilde{z}_t = a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \dots - \theta_q B^q a_t, \quad (8.194)$$

або

$$\tilde{z}_t = \Theta(B) a_t, \quad (8.195)$$

де  $\Theta(B)$  — оператор ковзного середнього

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (8.196)$$

порядку  $q$ , який фактично теж є оператором фільтра, але з обмеженою кількістю складових, тобто «укороченим» оператором фільтра.

Співвідношення (8.195) задає модель стаціонарного часового ряду  $z_t$ , у якій використовується оператор ковзного середнього (8.196), а тому математики домовились цю модель називати моделлю ковзного середнього порядку  $q$  (скорочено: модель КС( $q$ )).

**Покажемо, як оператор фільтра пов'язаний з оператором авторегресії.**

З філософської точки зору модель регресії – це модель «з оглядкою назад, у бік, звідки прийшов»; тобто це модель, яка встановлює значення якоїсь координати процесу у даний момент часу за її незалежними складовими, визначеними у попередній момент. Кількість врахованих складових визначає порядок регресії.

Виходячи з даного трактування, **модель авторегресії – це модель, яка встановлює значення якоїсь координати процесу у даний момент часу на основі своїх попередніх значень. Кількість врахованих попередніх значень визначає порядок авторегресії.**

Для центрованого часового ряду  $\tilde{z}_t$  модель авторегресії порядку  $p$  (скорочено: АР( $p$ )) можна записати у вигляді

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t, \quad (8.197)$$

де  $a_t$  – імпульс білого шуму, означення якого наведено вище.

З урахуванням співвідношення (8.171) вираз (8.197) можна переписати так

$$\tilde{z}_t - \phi_1 B \tilde{z}_t - \phi_2 B^2 \tilde{z}_t - \dots - \phi_p B^p \tilde{z}_t = a_t,$$

або

$$\Phi(B) \tilde{z}_t = a_t, \quad (8.198)$$

де  $\Phi(B)$  — **оператор авторегресії порядку  $p$** , який має вигляд

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (8.199)$$

Скористаємось далі тотожністю

$$\Phi(B) \Phi^{-1}(B) = \Phi^{-1}(B) \Phi(B) = I, \quad (8.200)$$

в якій  $I$  – одиничний оператор; оператор  $\Phi^{-1}(B)$  – обернений до оператора  $\Phi(B)$ .

Помноживши на  $\Phi^{-1}(B)$  зліва рівняння (8.198), отримаємо

$$\Phi^{-1}(B)\Phi(B)\tilde{z}_t = \Phi^{-1}(B)a_t, \quad (8.201)$$

або (з урахуванням (8.200))

$$I\tilde{z}_t = \Phi^{-1}(B)a_t. \quad (8.202)$$

Оскільки перемноження на одиничний оператор результату не змінює, то вираз (8.202) можна записати і так

$$\tilde{z}_t = \Phi^{-1}(B)a_t. \quad (8.203)$$

Порівнюючи вираз (8.203) з (8.191), можна стверджувати, що

$$\Psi(B) = \Phi^{-1}(B). \quad (8.204)$$

Отже, синтезувавши за реалізацією досліджуваного часового ряду  $z_t$  оператор авторегресії  $\Phi(B)$ , що зробити, як буде показано далі, нескладно, та визначивши оператор  $\Phi^{-1}(B)$ , обернений до  $\Phi(B)$ , що теж досить просто, тим самим одночасно визначаємо і оператор  $\Psi(B)$  лінійного фільтра, який з білого шуму  $a_t$  формує часовий ряд  $z_t$  з заданими властивостями.

Звертаємо увагу на те, що у цьому випадку  $\Psi(B)$  знаходиться без мінімізації критерію (8.167) при  $l = 0$ , про що йшла мова вище.

Під час розв'язання задачі ідентифікації моделі часового ряду  $z_t$  на основі оператора авторегресії порядку  $p$  необхідно визначити  $p + 2$  невідомих, якими є коефіцієнти  $\phi_1, \phi_2, \dots, \phi_p$  оператора  $\Phi(B)$ , середнє значення  $\mu$  процесу  $z_t$  та дисперсія  $\sigma_a^2$  білого шуму  $a_t$ .

Про те, як розв'язати цю задачу, мова піде після того, як ми **визначимось з поняттями автоковаріації та автокореляції часового ряду.**

**Автоковаріацією  $\gamma_k$  часового ряду  $z_t$  з затримкою  $k$  називають вираз**

$$\gamma_k = \text{cov}\{z_t, z_{t+k}\} = E\{(z_t - \mu) \cdot (z_{t+k} - \mu)\}, \quad (8.205)$$

в якому  $E$  – символ обчислення математичного очікування від виразу, що стоїть у фігурних дужках.

Зрозуміло, що

$$\gamma_0 = E\{(z_t - \mu)^2\} = \sigma_z^2 \quad (8.206)$$

– дисперсія часового ряду  $z_t$ .

Для отримання статистичної оцінки  $\gamma_k^*$  автоковаріації  $\gamma_k$ , визначеної виразом (8.205), використовують вираз

$$\gamma_k^* = \frac{1}{N-k} \sum_{t=1}^{N-k} (z_t - \mu) \cdot (z_{t+k} - \mu). \quad (8.207)$$

**Автоковаріація  $\gamma_k$  характеризує ступінь лінійного зв'язку між значеннями часового ряду  $z_t$  та  $z_{t+k}$ .**

Зрозуміло, що

$$\begin{cases} |\gamma_k| \leq \gamma_0, \\ \gamma_k = \gamma_{-k}. \end{cases} \quad (8.208)$$

**Автокореляцією  $\rho_k$  часового ряду  $z_t$  із затримкою  $k$  називають вираз**

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{E\{(z_t - \mu) \cdot (z_{t+k} - \mu)\}}{E\{(z_t - \mu)^2\}}. \quad (8.209)$$

Для довільної автокореляції  $\rho_k$  справедливими є такі співвідношення:

$$\begin{cases} |\rho_k| \leq \rho_0, \\ \rho_0 = 1, \\ \rho_k = \rho_{-k}. \end{cases} \quad (8.210)$$

**Вся можлива сукупність  $\gamma_k$  є автоковаріаційною функцією часового ряду  $z_t$ . Вона належить до класу решітчастих функцій. Аналогічно, сукупність всіх значень  $\rho_k$  задає автокореляційну функцію часового ряду  $z_t$ .**

Приклад графіка автокореляційної функції  $\rho_k$  наведено на рис. 8.11.

А тепер повернемося до сформульованої уже вище задачі визначення коефіцієнтів оператора авторегресії, тобто отримаємо відповідь на запитання: «**А як же визначити коефіцієнти авторегресії моделі часового ряду у формі  $AR(p)$ ?**».

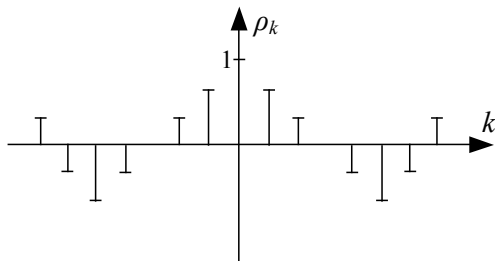


Рисунок 8.11 – Один із можливих графіків автокореляційної функції  $\rho_k$ ,  $k = \overline{-N, N}$

**Покажемо, що відповідь на це запитання дають рівняння Юла – Уокера.**

Для їх синтезу спочатку помножимо вираз (8.192) на  $\tilde{z}_{t-k}$ . У результаті цього отримаємо

$$\tilde{z}_{t-k}\tilde{z}_t = \phi_1\tilde{z}_{t-k}\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-k}\tilde{z}_{t-2} + \dots + \phi_p\tilde{z}_{t-k}\tilde{z}_{t-p} + \tilde{z}_{t-k}a_t. \quad (8.211)$$

Зробимо у виразі (8.211) заміну дискретної змінної, поклавши

$$t - k = \lambda. \quad (8.212)$$

Отримаємо

$$\tilde{z}_\lambda\tilde{z}_{\lambda+k} = \phi_1\tilde{z}_\lambda\tilde{z}_{\lambda+k-1} + \phi_2\tilde{z}_\lambda\tilde{z}_{\lambda+k-2} + \dots + \phi_p\tilde{z}_\lambda\tilde{z}_{\lambda+k-p} + \tilde{z}_\lambda a_{\lambda+k}. \quad (8.213)$$

Знайдемо математичне очікування  $E$  від обох частин рівняння (8.213). Отримаємо

$$E\{\tilde{z}_\lambda\tilde{z}_{\lambda+k}\} = \phi_1E\{\tilde{z}_\lambda\tilde{z}_{\lambda+k-1}\} + \phi_2E\{\tilde{z}_\lambda\tilde{z}_{\lambda+k-2}\} + \dots + \phi_pE\{\tilde{z}_\lambda\tilde{z}_{\lambda+k-p}\} + E\{\tilde{z}_\lambda a_{\lambda+k}\}. \quad (8.214)$$

Враховуючи вирази (8.188) та (8.205), із виразу (8.214) маємо

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \dots + \phi_p\gamma_{k-p} \quad (8.215)$$

для всіх  $k$  від 1 до  $p$ .

Але при  $k = 0$ , з урахуванням виразу (8.208), отримаємо ще одне рівняння

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \dots + \phi_p\gamma_p + \sigma_a^2. \quad (8.216)$$

Відсутність результатів обчислення математичного очікування  $E\{\tilde{z}_\lambda \cdot a_{\lambda+k}\}$  у рівняннях (8.215) та їх присутність у рівнянні (8.216) у вигляді дисперсії білого шуму  $\sigma_a^2$  пояснюється тим, що, згідно з властивостями білого шуму, кожний його імпульс корелюється (взаємпов'язується) лише сам із собою і зовсім не корелюється ні з яким іншим, навіть розміщеним у часі поряд. Тож

$$E\{\tilde{z}_\lambda \cdot a_{\lambda+k}\} = \begin{cases} 0, & \text{при } k \neq 0 \\ \sigma_a^2, & \text{при } k = 0 \end{cases}. \quad (8.217)$$

З рівняння (8.216), з урахуванням виразу (8.206), маємо

$$\sigma_a^2 = \sigma_z^2 - \phi_1\gamma_1 - \phi_2\gamma_2 - \dots - \phi_p\gamma_p. \quad (8.218)$$

Тож, якщо для реалізації стаціонарного часового ряду  $z_t$  довжиною  $N$  вже обчислені оціночні середні значення  $\mu$  і дисперсія  $\sigma_z^2$  за відомими формулами





Для розв'язання системи (8.222) спочатку визначаємо матрицю  $M^{-1}$ , яка є оберненою до матриці  $M$ . Потім множимо матричне рівняння (8.222) зліва на  $M^{-1}$ . У результаті цього отримуємо

$$M^{-1}M \cdot \Phi = M^{-1}\rho, \quad (8.223)$$

або

$$I \cdot \Phi = M^{-1}\rho, \quad (8.224)$$

і остаточно

$$\Phi = M^{-1}\rho. \quad (8.225)$$

*Рівняння (8.221), (8.222) носять назву рівнянь Юла – Уокера. Їх розв'язок у вигляді (8.225) дозволяє за попередньо обрахованими автокореляціями  $\rho_k$ ,  $k = \overline{1, p}$  часового ряду  $z_t$  визначити вектор  $\Phi(p \times 1)$  коефіцієнтів  $\phi_1, \phi_2, \dots, \phi_p$  оператора авторегресії у моделі  $AR(p)$ .*

В моделі часового ряду  $z_t$  на основі авторегресії порядку  $p$  у формуванні поточного значення ряду бере участь лише один поточний імпульс білого шуму  $a_t$ . Природно припустити, що якщо від цього імпульсу  $a_t$  відняти зважену суму  $q$  попередніх значень білого шуму, то отримаємо модель, яка буде враховувати більш «тонкі» моменти випадкового процесу і більш адекватно віддзеркалюватиме його властивості, оскільки крім авторегресії ця модель буде враховувати ще й ковзне середнє процесу.

Така модель часового ряду  $z_t$  носить назву моделі авторегресії – ковзного середнього порядку  $(p, q)$  (скорочено: модель АРКС  $(p, q)$ ) і має вигляд

$$\begin{aligned} \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + \\ &+ a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \end{aligned} \quad (8.226)$$

Перенісши всі члени з  $\tilde{z}_{t-i}$ ,  $i = \overline{1, p}$  у ліву частину рівняння (8.226) та виконавши вже відомі перетворення, отримаємо рівняння

$$\Phi(B)\tilde{z}_t = \Theta(B)a_t, \quad (8.227)$$

в якому оператори  $\Phi(B)$  та  $\Theta(B)$  визначаються виразами (8.195), (8.196).

Рівняння (8.227) і є основною формою моделі часового ряду  $z_t$  на основі АРКС  $(p, q)$ .

Для ідентифікації цієї моделі потрібно визначити  $p + q + 2$  невідомих, якими є коефіцієнти  $\phi_i$ ,  $i = \overline{1, p}$  оператора  $\Phi(B)$ , коефіцієнти  $\theta_j$ ,  $j = \overline{1, q}$  оператора  $\Theta(B)$ , середнє значення  $\mu$  процесу  $z_t$  та дисперсія  $\sigma_a^2$  білого шуму  $a_t$ .

**Всі моделі часових рядів, що побудовані вище, базувались на умові стаціонарності цих рядів. Але у повсякденному житті постійно стикаємося і з нестаціонарними випадковими процесами.** Наприклад, це процеси пуску або гальмування будь-якого технологічного обладнання, яке реалізує технологічний процес стохастичного характеру.

*Покажемо, що такі нестационарні випадкові процеси, які при їх дискретизації перетворюються на часові ряди, можна досить адекватно описувати за допомогою моделі, в якій закладені оператори авторегресії – проінтегрованого ковзного середнього.*

Для їх синтезу припустимо, що у моделі АРКС( $p, q$ ), поданій виразом (8.227), оператор  $\Phi(B)$  має  $d$  кратних коренів, що дорівнюють одиниці.

У цьому випадку, згідно з теоремою Вієтта, оператор  $\Phi(B)$  можна записати у вигляді

$$\Phi(B) = (1 - B)^d \cdot (1 - \phi_1^* B - \phi_2^* B^2 - \dots - \phi_l^* B^l), \quad (8.228)$$

де

$$d + l = p. \quad (8.229)$$

Позначимо

$$\Phi^*(B) = 1 - \phi_1^* B - \phi_2^* B^2 - \dots - \phi_l^* B^l. \quad (8.230)$$

З урахуванням виразів (8.228) та (8.230) рівняння (8.227) можна переписати так

$$\Phi^*(B) \cdot (1 - B)^d \tilde{z}_t = \Theta(B) a_t. \quad (8.231)$$

Оскільки, згідно з виразом (8.178),

$$(1 - B)^d = \nabla^d, \quad (8.232)$$

тобто,  $(1 - B)^d$  є різницеvim оператором із зсувом назад порядку  $d$ , то рівняння

$$(1 - B)^d z_t = \nabla^d z_t \quad (8.233)$$

задає нову змінну  $w_t$ , яка пов'язана з  $z_t$  співвідношенням

$$w_t = \nabla^d z_t. \quad (8.234)$$

Підставляючи вираз (8.234) у рівняння (8.231), отримуємо

$$\Phi^*(B) w_t = \Theta(B) a_t. \quad (8.235)$$

Очевидно, що вираз (8.235) задає модель АРКС( $l, q$ ) відносно  $w_t$ , яке можна переписати і так

$$\begin{aligned} w_t &= \phi_1^* w_{t-1} + \phi_2^* w_{t-2} + \dots + \phi_l^* w_{t-l} + \\ &+ a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \end{aligned} \quad (8.236)$$

Рівняння (8.234), (8.235) задають модель нестационарного часового ряду  $z_t$  у вигляді авторегресії – проінтегрованого ковзного середнього порядку  $(l, q, d)$ . Скорочено: модель АРПКС $(l, q, d)$ .

Звертаємо увагу на те, що вже перша різниця  $\nabla z_t$  значень будь-якого нестационарного часового ряду  $z_t$  має менший ступінь нестационарності, ніж сам часовий ряд  $z_t$ . Ще менший ступінь нестационарності матиме друга різниця  $\nabla^2 z_t$ , яка є різницею перших різниць  $\nabla(\nabla z_t)$  цього часового ряду  $z_t$ .

Підвищуючи порядок  $d$  різниці  $\nabla^d z_t$ , рано чи пізно дійдемо до такого її значення  $w_t$ , яке вже являтиме собою стаціонарний часовий ряд відносно  $w_t$ . На рис. 8.12 наведена графічна інтерпретація цього факту.

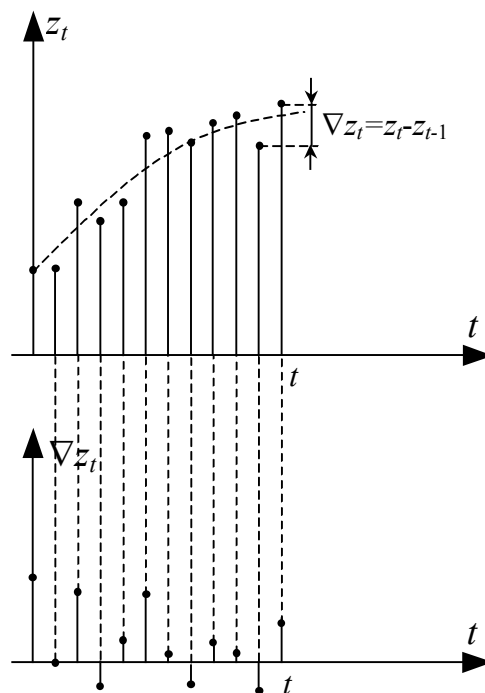


Рисунок 8.12 – Графічна інтерпретація перетворення нестационарного часового ряду  $z_t$  у стаціонарний часовий ряд для його різниці  $\nabla z_t$

Зрозуміло, що з моделі АРПКС $(l, q, d)$  (8.234), (8.235) при  $d = 0$  отримаємо модель АРКС $(p, q)$ .

Дамо пояснення чому у назві моделі АРПКС $(l, q, d)$  має місце слово «проінтегроване» стосовно ковзного середнього.

Нагадаємо, що оберненим оператором для  $\nabla$  є оператор суми  $S$  (8.185). Тому, отримавши  $w_t$  із рівняння (8.235), для переходу до часового ряду  $z_t$  необхідно координату  $w_t$  просумувати  $d$  разів, оскільки, помножуючи рівняння (8.234) зліва на  $\nabla^{-d}$ , маємо

$$\nabla^{-d} w_t = \nabla^{-d} \nabla^d z_t, \quad (8.237)$$

або

$$\nabla^{-d} w_t = I \cdot z_t, \quad (8.238)$$

звідки, з урахуванням (8.185), матимемо

$$z_t = \nabla^{-d} w_t = S^d w_t. \quad (8.239)$$

Зрозуміло, що найскладнішим завданням під час використання моделі АРПКС( $l, q, d$ ) є визначення числового значення параметра інтегрування  $d$ , або, в інших термінах, визначення кількості різниць, які потрібно послідовно взяти від нестационарного часового ряду  $z_t$ , щоб перетворити його у стаціонарний ряд відносно якоїсь різниці цього ряду. Очевидно, що його потрібно розв'язувати шляхом підстановки значення  $z_t$ , отриманого за виразом (8.239), та експериментального значення цієї координати в критеріальний функціонал (8.167) і пошуку мінімального значення цього функціоналу в наближеній інтерпретації.

## 8.5 Приклади реалізації процесів застосування спеціальних операторів

**Першим розглянемо приклад застосування прямого (8.1) і оберненого (8.20) операторів Лапласа,** які широко використовуються при аналізі процесів в лінійних динамічних системах.

Нехай процес в лінійній динамічній системі, на вхід якої надходить сигнал  $x(t)$  і реакцією якої на цей сигнал є вихідна координата  $y(t)$ , описується диференціальним рівнянням

$$\frac{d^4 y}{dt^4} + 10 \frac{d^3 y}{dt^3} + 35 \frac{d^2 y}{dt^2} + 50 \frac{dy}{dt} + 24y = 2 \frac{dx}{dt} + x, \quad (8.240)$$

яке є математичною моделлю цього процесу в часовій області за нульових початкових умов (8.51).

І нехай нам потрібно з'ясувати, який характер буде мати реакція  $y(t)$  цієї динамічної системи на вхідний сигнал

$$x(t) = 5e^{-t}. \quad (8.241)$$

Як відомо, за нульових початкових умов класичним методом диференціальне рівняння (8.240) розв'язане бути не може, оскільки система рівнянь для визначення констант інтегрування при кожній із експонент інших розв'язків, окрім нульового, не матиме. Але це диференціальне рівняння (8.240) легко розв'язується, якщо за допомогою прямого оператора Лапласа трансформувати його на комплексну площину, тобто, якщо здійснити перетворення

$$L\left\{\frac{d^4 y}{dt^4}\right\} + L\left\{10 \frac{d^3 y}{dt^3}\right\} + L\left\{35 \frac{d^2 y}{dt^2}\right\} + L\left\{50 \frac{dy}{dt}\right\} + L\{24y\} = L\left\{2 \frac{dx}{dt}\right\} + L\{x\}, \quad (8.242)$$

що допустимо, оскільки і диференціальне рівняння, і оператор Лапласа є лінійними. А далі, застосовуючи вираз (8.10) і виносячи за дужки  $Y(p)$  та  $X(p)$ , операторне рівняння (8.241), приведемо до вигляду

$$(p^4 + 10p^3 + 35p^2 + 50p + 24)Y(p) = (2p + 1)X(p), \quad (8.243)$$

який легко трансформується у вираз

$$Y(p) = \frac{2p + 1}{p^4 + 10p^3 + 35p^2 + 50p + 24} X(p). \quad (8.244)$$

Перетворюючи за Лапласом сигнал (8.242) з використанням виразу (8.6) за умови, що  $\alpha = 1$ , матимемо

$$X(p) = \frac{5}{p + 1}. \quad (8.245)$$

А підставляючи вираз (8.245) у (8.244), отримаємо

$$Y(p) = \frac{2p + 1}{p^4 + 10p^3 + 35p^2 + 50p + 24} \cdot \frac{5}{p + 1} \quad (8.246)$$

або

$$Y(p) = \frac{10p + 5}{p^5 + 11p^4 + 45p^3 + 85p^2 + 74p + 24} = \frac{C(p)}{D(p)} \quad (8.247)$$

Легко бачити, що вираз (8.247) є перетвореною за Лапласом реакцією нашої динамічної системи, визначеної на комплексній площині, на вхідний сигнал (8.241). А тому для знаходження цієї реакції у часовій області, тобто для визначення  $y(t)$ , необхідно до виразу (8.247) застосувати обернений оператор Лапласа, який теж є лінійним, а тому справедливим буде вираз

$$L^{-1}\{Y(p)\} = L^{-1}\left\{\frac{10p + 5}{p^5 + 11p^4 + 45p^3 + 85p^2 + 74p + 24}\right\} = L^{-1}\left\{\frac{C(p)}{D(p)}\right\} = y(t) \quad (8.248)$$

Обернений оператор Лапласа у виразі (8.248) будемо застосовувати у вигляді теореми розкладання (8.22), якщо всі полюси виразу (8.248), тобто корені рівняння (8.23), яке для виразу (8.248) матиме вигляд

$$p^5 + 11p^4 + 45p^3 + 85p^2 + 74p + 24 = 0, \quad (8.249)$$

будуть різними числами (дійсними або комплексними), або у вигляді (8.34), якщо серед коренів рівняння (8.249) виявляться і кратні.

Тож наступним кроком в алгоритмі застосування оберненого оператора Лапласа є визначення полюсів виразу (8.248), або, що одне і те ж, визначення коренів рівняння (8.249). Цілком очевидно, що для розв'язання цього рівняння 5-го порядку необхідно застосувати відповідну програму в якомусь пакеті прикладних програм, наприклад в ППП MathCAD чи MALAB або Python.

Застосовуючи один із цих пакетів, знайдемо, що коренями рівняння (8.249) будуть: кратний корінь

$$p_1 = -1 \quad (8.250)$$

з кратністю  $k = 2$  і три простих корені

$$p_2 = -2, \quad p_3 = -3, \quad p_4 = -4. \quad (8.251)$$

Тож, оскільки маємо серед коренів рівняння (8.249) і один кратний, то обернений оператор Лапласа будемо застосовувати у формі (8.34), для чого нам потрібно спочатку знайти многочлен  $D_{n-k}(p)$ , який для наших умов

$$n = 5, \quad k = 2 \quad (8.252)$$

та, згідно з виразом (8.32), матиме вигляд

$$D_3(p) = \frac{p^5 + 11p^4 + 45p^3 + 85p^2 + 74p + 24}{(p+1)^2} = p^3 + 9p^2 + 26p + 24. \quad (8.253)$$

Крім того, згідно з виразами (8.24) та (8.248), матимемо

$$D'(p) = \frac{dD}{dp} = 5p^4 + 44p^3 + 135p^2 + 170p + 74. \quad (8.254)$$

З урахуванням виразів (8.250)–(8.254) для нашого прикладу обернений оператор Лапласа у формі (8.34) набуває вигляду

$$y(t) = \frac{d}{dp} \left[ \frac{(10p+5)e^{p_1 t}}{p^3 + 9p^2 + 26p + 24} \right]_{p=p_1} + \sum_{i=2}^4 \frac{10p_i + 5}{5p_i^4 + 44p_i^3 + 135p_i^2 + 170p_i + 74} e^{p_i t}. \quad (8.255)$$

Беручи похідну з виразу (8.255) та розписуючи суму, матимемо

$$\begin{aligned} y(t) = & \frac{[10e^{p_1 t} + (10p_1 + 5)te^{p_1 t}](p_1^3 + 9p_1^2 + 26p_1 + 24) - (10p_1 + 5)e^{p_1 t}(3p_1^2 + 18p_1 + 26)}{(p_1^3 + 9p_1^2 + 26p_1 + 24)^2} + \\ & + \frac{(10p_2 + 5)e^{p_2 t}}{5p_2^4 + 44p_2^3 + 135p_2^2 + 170p_2 + 74} + \\ & + \frac{(10p_3 + 5)e^{p_3 t}}{5p_3^4 + 44p_3^3 + 135p_3^2 + 170p_3 + 74} + \frac{(10p_4 + 5)e^{p_4 t}}{5p_4^4 + 44p_4^3 + 135p_4^2 + 170p_4 + 74} \end{aligned} \quad (8.256)$$

А підставляючи у вираз (8.256) числові значення полюсів із виразів (8.250), (8.251) та здійснюючи відповідні розрахунки, отримаємо

$$y(t) = \frac{5}{36}(23 - 6t)e^{-t} - \frac{15}{2}e^{-2t} + \frac{25}{4}e^{-3t} - \frac{35}{18}e^{-4t}. \quad (8.257)$$

Приклади застосування операторів Фур'є і Деча наведені нами в підрозділах 8.2 та 8.3 при викладенні матеріалу, пов'язаного з цими операторами, а тому у цьому підрозділі ми не будемо їм приділяти увагу додатково, а як ще один приклад, який вважатимемо другим, розглянемо задачу синтезу оператора авторегресії з використанням експериментально визначених значень часового ряду.

Цей приклад запозичено нами з нашого ж навчального посібника з математичних методів ідентифікації динамічних систем.

Отже, нехай 10 значень вихідної координати  $y$  об'єкта, зафіксованих нами через однакові проміжки часу в процесі нормальної експлуатації, виявились такими, як показано в таблиці 8.1.

Таблиця 8.1

$t$	1	2	3	4	5	6	7	8	9	10
$y_t$	20	50	40	20	30	50	10	40	50	20

Побудуємо модель цього часового ряду, придатну для прогнозу наступних значень, на основі авторегресій 1-го, 2-го та 3-го порядків, що мають вигляд

$$z_t = \phi_1 z_{t-1} + a_t, \quad (8.258)$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t, \quad (8.259)$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + a_t, \quad (8.260)$$

де

$$z_t = y_t - m_y = y_t - \mu \quad (8.261)$$

Перш ніж розв'язувати задачу ідентифікації цих моделей і вибору з них кращої, розрахуємо всі необхідні для цього параметри часового ряду, заданого таблицею 8.1, скориставшись виразами (8.219), (8.220) та (8.207), (8.210).

Отже,

$$\mu = m_y = \frac{1}{10} \sum_{t=1}^{10} y_t = \frac{1}{10} (20 + 50 + 40 + 20 + 30 + 50 + 10 + 40 + 50 + 20) = 33, \quad (8.262)$$

$$\begin{aligned} \gamma_z[0] = \sigma_z^2 = \sigma_y^2 = \frac{1}{9} \sum_{t=1}^{10} (y_t - m_y)^2 = \frac{1}{9} [(-13)^2 + (17)^2 + (7)^2 + (-13)^2 + (-3)^2 + \\ + (17)^2 + (-23)^2 + (7)^2 + (17)^2 + (-13)^2] = 221, \end{aligned} \quad (8.263)$$

$$\begin{aligned} \gamma_z[1] = \frac{1}{8} \sum_{t=1}^9 (y_t - m_y)(y_{t+1} - m_y) = \frac{1}{8} [(-13) \cdot 17 + 17 \cdot 7 + 7 \cdot (-13) + \\ + (-13) \cdot (-3) + (-3) \cdot 17 + 17 \cdot (-23) + (-23) \cdot 7 + 7 \cdot 17 + 17 \cdot (-13)] = -105, \end{aligned} \quad (8.264)$$

$$\begin{aligned} \gamma_z[2] = \frac{1}{7} \sum_{t=1}^8 (y_t - m_y)(y_{t+2} - m_y) = \frac{1}{7} [(-13) \cdot 7 + 17 \cdot (-13) + 7 \cdot (-3) + \\ + (-13) \cdot 17 + (-3) \cdot (-23) + 17 \cdot 7 + (-23) \cdot 17 + 7 \cdot (-13)] = -116, \end{aligned} \quad (8.265)$$

$$\begin{aligned} \gamma_z[3] = \frac{1}{6} \sum_{t=1}^7 (y_t - m_y)(y_{t+3} - m_y) = \frac{1}{6} [(-13) \cdot (-13) + 17 \cdot (-3) + \\ + 7 \cdot 17 + (-13) \cdot (-23) + (-3) \cdot 7 + 17 \cdot 17 + (-23) \cdot (-13)] = -173, \end{aligned} \quad (8.266)$$

$$\rho_z[1] = \frac{\gamma_z[1]}{\gamma_z[0]} = \frac{-105}{221} = -0,475, \quad (8.267)$$



$$\rho_z[2] = \frac{\gamma_z[2]}{\gamma_z[0]} = \frac{-116}{221} = -0,525, \quad (8.268)$$

$$\rho_z[3] = \frac{\gamma_z[3]}{\gamma_z[0]} = \frac{-173}{221} = -0,783. \quad (8.269)$$

Тепер ми маємо всі необхідні дані для визначення матриць

$$M, \phi, \rho, \quad (8.270)$$

з використанням яких здійснюється ідентифікація моделей авторегресії (8.258)–(8.260).

Оскільки для моделі (8.258)  $q = 1$ , то ці матриці для неї будуть мати вигляд

$$M = [1], \quad \phi = [\phi_1], \quad \rho = [\rho_z[1]]. \quad (8.271)$$

Підставляючи (8.271) у (8.225) та враховуючи те, що у даному випадку  $M^{-1} = M$ , матимемо

$$\phi_1 = \rho_z[1], \quad (8.272)$$

а з урахуванням виразу (8.267) –

$$\phi_1 = -0,475. \quad (8.273)$$

Оскільки для моделі (8.259)  $q = 2$ , то матриці (8.270) для неї будуть мати вигляд

$$M = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}, \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}, \quad \rho = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}, \quad (8.274)$$

а з урахуванням виразів (8.267), (8.268) –

$$M = \begin{bmatrix} 1 & -0,475 \\ -0,475 & 1 \end{bmatrix}, \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}, \quad \rho = \begin{bmatrix} -0,475 \\ -0,525 \end{bmatrix} \quad (8.275)$$

Оскільки для моделі (8.260)  $q = 3$ , то матриці (8.270) для неї будуть мати вигляд

$$M = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}, \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}, \quad \rho = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} \quad (8.276)$$

а з урахуванням виразів (8.267), (8.268), (8.269) –

$$M = \begin{bmatrix} 1 & -0,475 & -0,525 \\ -0,475 & 1 & -0,475 \\ -0,525 & -0,475 & 1 \end{bmatrix}, \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}, \quad \rho = \begin{bmatrix} -0,475 \\ -0,525 \\ -0,783 \end{bmatrix} \quad (8.277)$$

Зіставляючи вирази (8.258) та (8.272), бачимо, що для ідентифікації моделі авторегресії (8.258) нам залишилось лише віднайти дисперсію  $\sigma_a^2$  «білого шуму»  $a(t)$ , з якого комп'ютерна програма формуватиме імпульси  $a_t$  для цієї моделі. Цю дисперсію ми знайдемо з виразу (8.218), підставивши в який вирази (8.263), (8.264) і (8.273), отримаємо

$$\sigma_a^2 = \gamma_z[0] - \phi_1 \gamma_z[1] = 221 - (-0,475)(-105) = 171,125 \approx 171. \quad (8.278)$$

Тож у цьому випадку модель стохастичної складової  $z_t$  вихідної координати  $y_t$  матиме вигляд

$$z_t = \phi_1 z_{t-1} + a_t = -0,475 z_{t-1} + a_t, \quad (8.279)$$

а значення самої вихідної координати  $y_t$  знайдемо з виразу (8.261), який після підстановки в нього значення  $\mu$  з виразу (8.262) набуде вигляду

$$y_t = 33 + z_t. \quad (8.280)$$

Підставляючи у вираз (8.280) значення  $y_{10}$  із таблиці 8.1, знайдемо, що

$$z_{10} = y_{10} - 33 = 20 - 33 = -13. \quad (8.281)$$

І якщо ми хочемо спрогнозувати значення  $y_{11}$  вихідної координати об'єкта, то, спочатку, згенерувавши комп'ютерною програмою, в яку закладаємо  $\sigma_a^2 = 171$ , в момент часу, який настає одразу після  $t = 10$ , тобто, після того, як ми підраховали  $z_{10}$ , і якому ми надаємо значення  $t = 11$ , імпульс «білого шуму»  $a_{11}$  – нехай це значення дорівнює середньоквадратичному відхиленню  $\sigma_a = \sqrt{\sigma_a^2} \approx 13$  – підставляємо значення  $z_{10}$  та  $a_{11}$  у вираз (8.279) і отримуємо числове значення  $z_{11}$ , яке у нашому випадку дорівнюватиме

$$z_{11} = -0,475 z_{10} + a_{11} = -0,475(-13) + 13 = 19,175 \approx 19. \quad (8.282)$$

А потім отримане числове значення  $z_{11}$  підставляємо у вираз (8.280) і знаходимо, що

$$y_{11} = 33 + z_{11} = 33 + 19 = 52. \quad (8.283)$$

Аналогічно, використовуючи знайдені  $z_{11}, y_{11}$  та генеруючи наступний імпульс «білого шуму», ми можемо спрогнозувати в момент часу  $t = 10$  і значення  $z_{12}$  та  $y_{12}$ , а потім, використовуючи уже отримані  $z_{12}$  та  $y_{12}$ , спрогнозувати і  $z_{13}, y_{13}$  однак, зрозуміло, що чим більш віддалені від  $t = 10$  моменти часу ми будемо використовувати, тим точність прогнозу зменшуватиметься.

Але перш ніж використовувати ідентифіковану у вигляді авторегресії 1-го порядку модель (8.279) для прогнозування наступних значень вихідної координати об'єкта, потрібно пересвідчитись, що ця модель задає достатню точність прогнозу. А для цього потрібно, скориставшись матрицями (8.275), ідентифікувати модель часового ряду, заданого таблицею 8.1, у вигляді авторегресії 2-го порядку (8.259) і порівняти ступені точності прогнозування за обома моделями.

А процес ідентифікації моделі авторегресії 2-го порядку почнемо з визначення матриці  $M^{-1}$ , оберненої до матриці  $M$ , заданої виразом (8.275).

Нагадаємо, що для обернення матриці потрібно замінити кожний елемент матриці, яку обертають, на його алгебраїчне доповнення, котре являє собою визначник, отриманий з матриці, яку обертають, викресленням рядка і стовпця, на перетині яких стоїть цей елемент, помножений на «мінус одиницю» у степені, що дорівнює сумі номерів викреслених рядка і стовпця, та поділений на визначник матриці, яку обертають. Перш, ніж, виходячи з цього визначення, віднайти обернену матрицю  $M^{-1}$  для матриці  $M$ , заданої виразом (8.275), знайдемо спочатку визначник  $\Delta_M$  матриці  $M$ . Очевидно, що

$$\Delta_M = \begin{vmatrix} 1 & -0,475 \\ -0,475 & 1 \end{vmatrix} = 1 \cdot 1 - (-0,475) \cdot (-0,475) = 0,775 \quad (8.284)$$

Тож, згідно з наведеним вище визначенням, обернену матрицю  $M^{-1}$  для матриці  $M$ , заданої виразом (8.275), матимемо у вигляді

$$M^{-1} = \begin{bmatrix} 1 \cdot (-1)^{1+1} \cdot \frac{1}{0,775} & -0,475 \cdot (-1)^{1+2} \cdot \frac{1}{0,775} \\ -0,475 \cdot (-1)^{2+1} \cdot \frac{1}{0,775} & 1 \cdot (-1)^{2+2} \cdot \frac{1}{0,775} \end{bmatrix} = \begin{bmatrix} 1,290 & 0,613 \\ 0,613 & 1,290 \end{bmatrix}. \quad (8.285)$$

А підставляючи у матричне рівняння (8.225) із виразу (8.285) матрицю  $M^{-1}$  та з виразу (8.275) матриці  $\phi$  і  $\rho$ , знайдемо, що

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} 1,290 & 0,613 \\ 0,613 & 1,290 \end{bmatrix} \cdot \begin{bmatrix} -0,475 \\ -0,525 \end{bmatrix} = \begin{bmatrix} 1,290 \cdot (-0,475) + 0,613 \cdot (-0,525) \\ 0,613 \cdot (-0,475) + 1,290 \cdot (-0,525) \end{bmatrix} = \begin{bmatrix} -0,935 \\ -0,968 \end{bmatrix}. \quad (8.286)$$

В свою чергу, підставляючи значення  $\phi_1, \phi_2$  із виразу (8.286) в рівняння (8.259), отримаємо модель для стохастичної складової  $z_t$  у вигляді авторегресії II-го порядку

$$z_t = -0,935 z_{t-1} - 0,968 z_{t-2} + a_t, \quad (8.287)$$

в якому імпульси  $a_t$  формуватимуться комп'ютерною програмою, виходячи з числового значення дисперсії «білого шуму», визначеної з виразу (8.218) після підстановки в нього значень параметрів із виразів (8.263), (8.264), (8.265), (8.286), тобто, виходячи з числового значення

$$\sigma_a^2 = 221 - (-0,935)(-105) - (-0,968)(-116) = 10,537. \quad (8.288)$$

Оскільки дисперсія «білого шуму» у випадку використання моделі авторегресії 2-го порядку є суттєво меншою дисперсії «білого шуму», яку ми мали для моделі авторегресії 1-го порядку, то це означає, що для формування стохастичної складової за виразом (8.287) будуть використовуватись суттєво менші імпульси «білого шуму», а тому прогноз буде точнішим. Як і у попередньому випадку для прогнозування значення  $z_{11}$  сформуємо імпульс  $a_{11}$ , значення якого дорівнює середньоквадратичному відхиленню; у даному

випадку це  $\sigma_a = \sqrt{\sigma_a^2} \approx 3$ . Підставивши це значення  $a_{11}$  у вираз (8.287) разом з обрахованими з використанням даних із таблиці 8.1 значеннями  $z_{t-1} = z_{10} = y_{10} - \mu = 20 - 33 = -13$ ,  $z_{t-2} = z_9 = y_9 - \mu = 50 - 33 = 17$ , спрогнозуємо, що

$$z_{11} = -0,935 \cdot (-13) - 0,968 \cdot 17 + 3 = -1,3, \quad (8.289)$$

а тому

$$y_{11} = \mu + z_{11} = 33 - 1,3 = 31,7. \quad (8.290)$$

А для закріплення викладеного матеріалу здійснимо процес ідентифікації моделі прогнозу у вигляді авторегресії 3-го порядку, що має вигляд (8.260).

Зберігаючи ту ж послідовність дій, що і у попередніх двох випадках, ми у цьому, третьому, випадку не будемо робити пояснень, а лише наведемо результати розрахунків.

Отже,

$$\Delta_M = \begin{vmatrix} 1 & -0,475 & -0,525 \\ -0,475 & 1 & -0,475 \\ -0,525 & -0,475 & 1 \end{vmatrix} = 1 \cdot 1 \cdot 1 + (-0,475) \cdot (-0,475) \cdot (-0,525) +$$

$$+ (-0,475) \cdot (-0,475) \cdot (-0,525) - (-0,525) \cdot 1 \cdot (-0,525) - (-0,475) \cdot (-0,475) \cdot 1 -$$

$$- (-0,475) \cdot (-0,475) \cdot 1 = 0,25, \quad (8.291)$$

$$M^{-1} = \frac{1}{0,25} \begin{bmatrix} (-1)^{1+1} \begin{vmatrix} 1 & -0,475 \\ -0,475 & 1 \end{vmatrix} & (-1)^{1+2} \begin{vmatrix} -0,475 & -0,475 \\ -0,525 & 1 \end{vmatrix} & (-1)^{1+3} \begin{vmatrix} -0,475 & 1 \\ -0,525 & -0,475 \end{vmatrix} \\ (-1)^{2+1} \begin{vmatrix} -0,475 & -0,525 \\ -0,475 & 1 \end{vmatrix} & (-1)^{2+2} \begin{vmatrix} 1 & -0,525 \\ -0,525 & 1 \end{vmatrix} & (-1)^{2+3} \begin{vmatrix} 1 & -0,475 \\ -0,525 & -0,475 \end{vmatrix} \\ (-1)^{3+1} \begin{vmatrix} -0,475 & -0,525 \\ 1 & -0,475 \end{vmatrix} & (-1)^{3+2} \begin{vmatrix} 1 & -0,525 \\ -0,475 & -0,475 \end{vmatrix} & (-1)^{3+3} \begin{vmatrix} 1 & -0,475 \\ -0,475 & 1 \end{vmatrix} \end{bmatrix} =$$

$$= \begin{bmatrix} 3,1 & 2,9 & 3 \\ 2,9 & 3 & 2,9 \\ 3 & 2,9 & 3,1 \end{bmatrix}, \quad (8.292)$$

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} 3,1 & 2,9 & 3 \\ 2,9 & 3 & 2,9 \\ 3 & 2,9 & 3,1 \end{bmatrix} \cdot \begin{bmatrix} -0,475 \\ -0,525 \\ -0,783 \end{bmatrix} = \begin{bmatrix} -5,346 \\ -5,222 \\ -5,374 \end{bmatrix} \quad (8.293)$$

$$z_t = -5,346 z_{t-1} - 5,222 z_{t-2} - 5,374 z_{t-3} + a_t, \quad (8.294)$$

$$\sigma_a^2 = 221 - (-5,346)(-105) - (-5,222)(-116) - (-5,374)(-173) \approx -1876 \quad (8.295)$$

Оскільки в разі використання моделі (8.294) потрібно реалізовувати «білий шум» з від'ємною дисперсією, що в принципі неможливо, то це означає, що модель авторегресії 3-го порядку не може бути використана як модель стохастичної складової  $z_t$  при прогнозуванні майбутніх членів часового ряду  $y_t$ , який відображає вихідну координату

об'єкта. Тож найкращою моделлю  $z_t$  для розглянутого нами випадку буде модель авторегресії 2-го порядку, яка показана виразом (8.287).

А завершити цей матеріал ми хочемо попередженням, що імпульс «білого шуму» генеруватиметься комп'ютерною програмою на основі таблиці випадкових чисел в межах від  $-3\sigma_a$  до  $+3\sigma_a$ , тож ці значення можуть мати як знаки «плюс», так і знаки «мінус» і суттєво відрізнятись за модулем від значення одного середньоквадратичного відхилення, взятого нами в прикладах для того, щоб можна було здійснити порівняння прогнозу при однакових умовах.

Ну і, звичайно ж, при обробленні великих масивів значень часового ряду з метою синтезу оператора авторегресії потрібно всі ті розрахунки, які ми здійснили задля того, щоб наочно продемонструвати їх суть і структуру, виконувати в якомусь програмному середовищі, наприклад, MathCAD, MATLAB чи Python.

## 8.6 Завдання для самоперевірки

1. Що собою являє оператор Лапласа?
2. Які ви знаєте властивості оператора Лапласа?
3. Які головні переваги аналізу в області зображень?
4. Що собою являють відображення за Лапласом похідної від неперервної функції та інтеграла від неї?
5. Як визначити передаточну функцію системи, якщо відоме диференціальне рівняння, котре описує процеси у цій системі?
6. Дайте означення перехідної та імпульсної перехідної характеристик лінійної динамічної системи.
7. Що собою являє одиничний стрибок, і який графік він має?
8. Що собою являє одиничний імпульс і які його властивості ви знаєте?
9. Як пов'язані зображення за Лапласом перехідної та імпульсної перехідної характеристик системи з передаточною функцією цієї системи?
10. Доведіть, що імпульсна перехідна характеристика системи є похідною від її перехідної характеристики.
11. Як, маючи передаточну функцію лінійної динамічної системи, відтворити диференціальне рівняння, яке описує процеси у цій системі?
12. Що собою являє обернений оператор Лапласа? Які форми його реалізації ви знаєте?
13. Як визначити оригінал за відомим його зображенням?
14. Що собою являють прямий і обернений оператори Фур'є?
15. Як пов'язані між собою прямі оператори Лапласа і Фур'є?
16. Яким чином, маючи передаточну функцію лінійної динамічної системи, отримати її частотні характеристики: дійсну, уявну, амплітудну, фазову, амплітудно-фазову?
17. Які частотні характеристики і за допомогою яких приладів можна знайти експериментально?
18. Як, маючи експериментально визначені окремі значення частотних характеристик, синтезувати передаточну функцію системи?
19. Як, знаючи передаточну функцію системи, синтезувати її диференціальний оператор?
20. Дайте означення решітчастої функції, наведіть приклад із побудовою графіка.
21. Які різницеві оператори є аналогом похідної для решітчастих функцій?
22. Чому недостатньо лише правого різницевого оператора першого порядку для характеристики швидкості зміни решітчастої функції? Який вихід?
23. Чим відрізняється рівняння в скінченних різницях від різницевого рівняння?
24. Як за диференціальним рівнянням побудувати різницеве? Наведіть приклад.
25. Що собою являє Z-оператор Деча для відображення решітчастих функцій і чому виникає в ньому потреба? Наведіть приклад Z-відображення решітчастої функції.
26. Наведіть властивості оператора Деча.

27. У чому полягає перевага використання лівих різницевих операторів під час застосування  $Z$ -оператора для аналізу дискретних систем?
28. Як здійснити аналіз процесів у дискретній динамічній системі за допомогою математичної моделі?
29. Як отримати дискретну передаточну функцію системи, що в часовій області описується різницевиими рівняннями?
30. Що собою являє обернений оператор Деча, і яку практичну реалізацію його визнаєте? Наведіть приклад отримання решітчастого оригіналу.
31. Що собою являє «часовий ряд»? Наведіть приклади стаціонарного та нестаціонарного часових рядів.
32. Дайте означення операторів зсуву назад та вперед, різницевого оператора та оператора суми.
33. Що собою являє оператор лінійного фільтра?
34. Що собою являє білий шум? Які його основні властивості?
35. Що таке регресія і авторегресія? Яку форму мають їх оператори?
36. Синтезуйте модель часового ряду у формі авторегресії.
37. Як пов'язані між собою оператор лінійного фільтра та оператор авторегресії для часового ряду?
38. Що собою являють ковзне середнє часового ряду і оператор ковзного середнього?
39. Дайте означення автоковаріації і автокореляції часового ряду. Як знайти їх числові оцінки? Які основні їх властивості ви знаєте?
40. Для чого потрібні і як виводяться рівняння Юла – Уокера?
41. Як розв'язати рівняння Юла – Уокера?
42. За реалізацією часового ряду синтезуйте його модель у формі авторегресії з ідентифікацією цієї моделі
43. Побудуйте модель часового ряду у формі авторегресії – ковзного середнього.
44. Яким чином можна нестаціонарний часовий ряд трансформувати у стаціонарний?
45. Побудуйте модель нестаціонарного часового ряду у формі авторегресії – проінтегрованого ковзного середнього.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Бабак В. П. Сигнали і спектри : навчальний посібник. / В. П. Бабак, А. Я. Білецький, А. М. Гуржій. – Київ : Книжне видавництво Національного авіаційного університету, 2005. – 492 с.
2. Бессонов Л. А. Теоретические основы электротехники. Электрические цепи / Л. А. Бессонов. – М. : Высш. шк., 1978. – 528 с.
3. Бокс Дж. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс. – Вып. 1. – М. : Мир, 1974. – 408 с.
4. Бокс Дж. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс. – Вып. 2. – М. : Мир, 1974. – 197 с.
5. Бронштейн И. Н. Справочник по математике для инженеров и учащихся вузов / И. Н. Бронштейн, К. А. Семендяев. – М. : Наука, 1967. – 608 с.
6. Демидович Б. П. Основы вычислительной математики / Б. П. Демидович, И. А. Марон. – 3-е изд. – М. : Наука, 1966. – 664 с.
7. Деруссо П. Пространство состояний в теории управления (для инженеров) / П. Деруссо, Р. Рой, Ч. Клоуз. – М. : Наука, 1970. – 620 с.
8. Деч Г. Руководство к практическому применению преобразования Лапласа и Z-преобразования / Г. Деч. – М. : Наука, 1971. – 288 с.
9. Изерман Р. Цифровые системы управления / Р. Изерман. – М. : Мир, 1984. – 541 с.
10. Згуровський М. З. Основи системного аналізу. / М. З. Згуровський, Н. Д. Панкратова. – К. : Видавнична група ВНУ, 2007. – 546 с.
11. Колмогоров А. Н. Элементы теории функций и функционального анализа. / А. Н. Колмогоров, С. В. Фомин. – 7-е изд. – М. : Физматлит, 2004. – 572 с.
12. Макаров И. П. Дополнительные главы математического анализа / И. П. Макаров. М. : Просвещение, 1968. – 312 с.
13. Математическая статистика / [под ред. А. М. Дина]. – М. : Высш. шк., 1975. – 398 с.
14. Мокін Б. І. Математичні методи ідентифікації динамічних систем : навчальний посібник. / Б. І. Мокін, В. Б. Мокін, О. Б. Мокін. – Вінниця : ВНТУ, 2010. – 260 с.
15. Мокін Б. І. Теорія автоматичного керування, методологія та практика оптимізації : навчальний посібник. / Б. І. Мокін, О. Б. Мокін. – Вінниця : ВНТУ, 2013. – 210 с.
16. Николаев А. Г. Функциональный анализ : учебное пособие. / А. Г. Николаев, Т. В. Рвачева, А. И. Соловьев. – Харьков : ХАИ, 2008. – 164 с.
17. Петров Ю. П. Вариационные методы теории оптимального управления. / Ю. П. Петров. – М.-Л. : Энергия, 1965. – 220 с.
18. Соболев В. И. Лекции по дополнительным главам математического анализа / В. И. Соболев. – М. : Наука, 1968. – 288 с.
19. Треногин В. А. Функциональный анализ. / В. А. Треногин. – М.: Наука, 1980. – 496 с.
20. Фильчаков П. Ф. Справочник по высшей математики / Фильчаков П. Ф. – Киев : Наукова думка, 1974. – 743 с.
21. Функціональний аналіз : навчальний посібник. / С. А. Ус. – Дніпропетровськ : Національний гірничий університет, 2013. – 236 с. – Режим доступу : [ir.nmu.org.ua/handle/123456789/3496/CD266.pdf?sequence=1](http://ir.nmu.org.ua/handle/123456789/3496/CD266.pdf?sequence=1)
22. Цлаф Л. Я. Вариационное исчисление и интегральные уравнения / Цлаф Л. Я. – Москва : Наука, 1966. – 176 с.
23. Шилов Г. Е. Интеграл, мера и производная на линейных пространствах. / Г. Е. Шилов, Фан Дык Тинь. – М. : Наука, 1967. – 192 с.
24. Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление / Эльсгольц Л. Э. – Москва : Наука, 1965. – 424 с.

25. Янг Л. Лекции по вариационному исчислению и теории оптимального управления / Янг Л. – Москва : Мир, 1974. – 488 с.
26. Макаров Е. Г. MathCAD : учебный курс / Е. Г. Макаров. – СПб. : Питер, 2009. – 384 с.
27. Верлань А. Ф. Моделювання систем керування в середовищі MATLAB : навчальний посібник / А. Ф. Верлань, І. О. Горошко, Д. Е. Контрерас, В. А. Федорчук, В. Ф. Юзвенко – К. : ЦКІС АПНУ, 2002. – 68 с
28. Лазарев Ю. Моделирование процессов и систем в MATLAB : учебный курс / Ю. Лазарев. – СПб. : Питер; Киев : Издательская группа BHV, 2005. – 51 с.
29. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5108753/>
30. <http://scipy-lectures.org/packages/sympy.html#integration>
31. [https://drive.google.com/open?id=1csncEhe5s9z\\_bkd4tkzNXX6UDnp3jKYv](https://drive.google.com/open?id=1csncEhe5s9z_bkd4tkzNXX6UDnp3jKYv)



*Навчальне видання*

**Борис Іванович Мокін  
Віталій Борисович Мокін  
Олександр Борисович Мокін**

**ФУНКЦІОНАЛЬНИЙ АНАЛІЗ,  
адаптований до прикладних задач в галузі інформаційних технологій**

Навчальний посібник

Рукопис оформив Б. Мокін

Редактор В. Дружиніна

Оригінал-макет виготовив О. Ткачук

Підписано до друку 20.05.2020.  
Формат 29,7×42¼. Папір офсетний.  
Гарнітура Times New Roman.  
Друк різнографічний. Ум. друк. арк. 11,52.  
Наклад 50 (1-й запуск 1–21) пр. Зам. № 2020-050.

Видавець та виготовлювач  
інформаційний редакційно-видавничий центр.  
ВНТУ, ГНК, к. 114.  
Хмельницьке шосе, 95,  
м. Вінниця, 21021.  
Тел. (0432) 65-18-06.  
**press.vntu.edu.ua;**  
*E-mail:* kivc.vntu@gmail.com.  
Свідоцтво суб'єкта видавничої справи  
серія ДК № 3516 від 01.07.2009 р.