Київський національний університет імені Тараса Шевченка

В. А. Львов
А. О. Косогор
О. В. Барабанов

ПРОСТО ПРО СКЛАДНЕ

ОСНОВИ ТА ЗАСТОСУВАННЯ ЛІНІЙНОЇ АЛГЕБРИ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

В. А. Львов
А. О. Косогор
О. В. Барабанов

ПРОСТО ПРО СКЛАДНЕ
ОСНОВИ ТА ЗАСТОСУВАННЯ ЛІНІЙНОЇ АЛГЕБРИ

Підручник

КИЇВський ВПЦ УНІВЕРСИТЕТ 2024

УДК 517:512.64(075.8)
Л89

Рецензенти:
д-р фіз.-мат. наук, доц. Євген Бондаренко,
канд. фіз.-мат. наук, доц. Маргарита Разумова
(Київський національний університет імені Тараса Шевченка)

Рекомендовано до друку вченою радою
факультету радіофізики, електроніки та комп'ютерних систем
(протокол № 4 від 25 жовтня 2023 року)

Ухвалено науково-методичною радою
Київського національного університету імені Тараса Шевченка
(протокол № 01-24 від 25 січня 2024 року)

Львов В. А.

Л89 Просто про складне: основи та застосування лінійної алгебри : підручник / В. А. Львов, А. О. Косогор, О. В. Барабанов. – К. : ВПЦ "Київський університет", 2024. – 223 с.

ISBN 978-966-933-245-5

Зазначено, що лінійна алгебра логічно випливає з алгебри геометричних векторів і полегшує розуміння та викладення інших розділів математики. У спрощеній формі пояснено суть аксіом лінійної алгебри та її важливих понять – лінійного простору, скалярного добутку елементів простору, лінійного оператора, білінійної та квадратичної форми. Показано, як ці поняття використовують для пояснення основних властивостей матриць та їхніх визначників, розв'язання алгебраїчних і диференціальних рівнянь, вивчення коливальних та хвильових процесів, дослідження стійкості стаціонарних станів динамічних систем до дії різних чинників, які збурюють ці стани. Пояснено зв'язок лінійної алгебри з квантовою механікою, рядом та перетворенням Фур'є, гармонічним аналізом сигналів.

Для студентів і викладачів природничих та інженерно-технічних факультетів університетів.

УДК 517:512.64(075.8)

ISBN 978-966-933-245-5

ПЕРЕДМОВА

Підручник призначений для студентів, які не планують бути професійними математиками, а також для молодих математиків, яким доручено викладати окремі розділи алгебри, математичного аналізу та теорії диференціальних рівнянь студентам природничих, фізико-технічних та інженерних факультетів.

На думку авторів, основні ідеї лінійної алгебри є важливою частиною наукового світогляду фахівців із фізико-технічних наук та інформаційних технологій. Безпосередньою мотивацією для написання підручника стало те, що на запитання "Що таке лінійна алгебра?" більшість старшокурсників та випускників нематематичних факультетів відповідають: "Це той розділ математики, у якому йдеться про сім або вісім аксіом". Отримати більш змістовну відповідь зазвичай не вдається. Для виправлення цієї ситуації в підручнику показано, що ті аксіоми, про які пам'ятають старшокурсники і випускники нематематичних факультетів, по-перше, логічно випливають з відомої випускникам середніх шкіл векторної алгебри і, по-друге, поширюються на алгебраїчні дії з іншими математичними об'єктами. Завдяки цьому засвоєння *основ* лінійної алгебри полегшує вивчення багатьох розділів вищої алгебри, математичного аналізу, теорії диференціальних рівнянь та фізико-технічних наук.

Назва підручника акцентує на *застосуванні* лінійної алгебри, адже, спираючись на основні поняття лінійної алгебри, викладачі можуть краще пояснити студентам, а студенти – легше засвоїти важливі поняття з інших наукових дисциплін (напр., принцип суперпозиції розв'язків лінійного диференціального рівняння та суть представлення функції рядом Фур'є). Окрім цього, лінійна алгебра допомагає науковцям та інженерам виконувати важливі практичні завдання (напр., виявляти нестійкість стаціонарного стану фізичної чи інженерної системи, зумовлену впливом на цю систему зовнішніх або внутрішніх чинників, здатних змінити її стан).

Додатковою мотивуючою обставиною до створення підручника стало те, що лінійна алгебра є формальною основою квантової механіки. Саме тому в ньому у спрощений спосіб пояснено такі поняття, як лінійний оператор, матричні елементи оператора, комутатор двох операторів, простір функцій, базис у просторі функцій, а також надано посилання на всесвітньо відомий підручник із квантової механіки, у якому докладно пояснено й обґрунтовано принциповий зв'язок квантової механіки з лінійною алгеброю.

Щоб полегшити розуміння навчального матеріалу, певну його частину викладено не у звичній для підручників з математики формі *теорем*, а у формі тверджень, які не виділяються з тексту, а обґрунтовуються у спрощений спосіб і підкріплюються посиланнями на більш повні підручники з математики. На відміну від цього, основні *означення*, які бажано запам'ятовувати, важливі *зауваження*, про які також не слід забувати, і *приклади* розв'язання типових задач виділені з тексту та пронумеровані, щоб уможливити посилання на них у наступних розділах підручника і в такий спосіб вказати на зв'язок лінійної алгебри з тими розділами вищої математики, у яких явно чи неявно застосовуються основи лінійної алгебри.

Читаючи підручник, варто звернути увагу на те, що відрізняє його від більшості інших навчальних видань з лінійної алгебри, а саме:

✓ пояснення поняття лінійної незалежності елементів та поняття вимірності лінійного простору на прикладі простору геометричних векторів (підрозд. 1.4);

✓ просте пояснення подібності між геометричними векторами та числовими функціями (підрозд. 1.8);

✓ структуровані означення понять "матриця" та "лінійний простір" (підрозд. 2.1 та 3.1, відповідно);

✓ простий розгляд поняття "ортонормований базис у просторі функцій" (п. 3.4.2);

✓ означення та розгляд властивостей визначника матриці за допомогою символів Леві-Чивіти (підрозд. 4.1 – 4.3);

✓ пояснення того, чому саме лінійні оператори Ерміта (а не якісь інші) важливі для квантової механіки (підрозд. 6.2, 6.4);

✓ приклад 6.2 у підрозд. 6.1 та приклад того, як квадратичні форми застосовуються для вивчення коливань систем із декількома ступенями вільності в п. 7.3.3;

✓ застосування загальних принципів теорії лінійних диференціальних рівнянь до вивчення вільних та вимушених коливань, динамічної сприйнятливості коливальної системи та явища резонансу (пп. 8.3.3, 8.4.3 та 8.4.4);

✓ простий розгляд загального інтеграла хвильового рівняння та доведення закону дисперсії плоских хвиль (підрозд. 8.5);

✓ графіки інтегральних кривих у близькому околі точок спокою, розраховані та побудовані за допомогою спеціалізованих програм для персонального комп'ютера (підрозд. 10.4, 10.5);

✓ простий приклад дослідження стаціонарних станів такої динамічної системи, яка описується нелінійним звичайним диференціальним рівнянням 2-го порядку (підрозд. 10.6);

✓ просте пояснення основних принципів наближеного представлення функцій сумою гармонік та *кількісний* розгляд похибки такого представлення (підрозд. 11.3);

✓ графіки спектральної щільності деяких функцій, розраховані та побудовані за допомогою спеціалізованих програм для персонального комп'ютера (підрозд. 11.7).

РОЗДІЛ 1
Алгебра геометричних векторів як підґрунтя лінійної алгебри

1.1. Уявлення про геометричні вектори

Кажуть, що людство існує у *фізичному просторі*. Словосполученнями "фізичний простір", "точка простору" та "відстань між точками простору" люди називають інтуїтивні поняття, які у давні часи не мали формального означення. Але вже в ті часи практично важливою справою було визначення відстаней між фізичними об'єктами на поверхні землі, тобто геометрія. Як відомо зі шкільних курсів математики та фізики, *геометричні вектори* – це відрізки прямої лінії між двома точками фізичного простору. Одну з цих точок називають *початком* вектора, а другу – його *кінцем*. Кінець геометричного вектора позначають стрілочкою.

Важливість геометричних векторів для сучасної науки полягає в тому, що вони є "повноважними представниками" багатьох фізичних величин, наприклад, швидкостей і прискорень фізичних тіл, сил, що діють на фізичне тіло, напруженості електричного або магнітного поля та багатьох інших. Такі фізичні величини називають *векторними величинами*. У фізиці простором зазвичай називають сукупність векторних величин і при цьому використовують такі поняття, як простір швидкостей, простір сил тощо.

Означення 1.1. Довжину геометричного вектора називають його модулем.

Щоб відрізняти вектор від його модуля, позначатимемо вектор жирною літерою, а його модуль – або цією ж літерою, розташованою між двома вертикальними рисками, або італіком. Наприклад, модуль вектора $\mathbf{a}$ позначатимемо або $|\mathbf{a}|$, або a.

Означення 1.2. Вектор, довжина якого дорівнює нулю, є точкою на площині або в просторі; позначатимемо його символом $\mathbf{0}$. Орт – це вектор, довжина якого дорівнює одиниці; позначатимемо його символом $\mathbf{e}$, якщо не буде обумовлене інше позначення.

Означення 1.3. Геометричні вектори називають *колінеарними*, якщо вони лежать на паралельних прямих, і – *компланарними*, якщо вони розташовані в паралельних площинах.

Зауваження 1.1. Колінеарні вектори можуть бути напрямлені або в один і той самий бік, або в протилежні боки.

Означення 1.4. Геометричні вектори називають рівними один одному, якщо вони мають однакову довжину, колінеарні та напрямлені в один і той самий бік.

Зауваження 1.2. З наданого означення рівності векторів випливає, що паралельний перенос вектора в просторі не змінює цього вектора і не порушує тих математичних рівностей, які він задовольняє. Вектори, рівність яких означена в такий спосіб, називають *вільними векторами*. Надалі розглядатимемо саме такі вектори.

У *векторній алгебрі* виконуються операції додавання векторів один до одного та операція множення вектора на число. Числа називають *скалярами*.

1.2. Сума векторів та добуток вектора на число

Означення 1.5. Сумою геометричних векторів називають вектор, який задовольняє правило трикутника, представлене в графічній формі на рис. 1, *а*, де вектори $\mathbf{a}$ та $\mathbf{b}$ зображені відрізками зі стрілочками на кінцях векторів. Сумою цих векторів *за означенням* є вектор $\mathbf{a}+\mathbf{b}$, також представлений на рис. 1, *а*[1].

---

[1] Рис. 1.1, *а* показує, що сумою векторів **a** та **b** називають вектор **a** + **b**, початок якого збігається з початком вектора **a**, кінець – з кінцем вектора **b**, якщо початок вектора **b** збігається з кінцем вектора **a**.

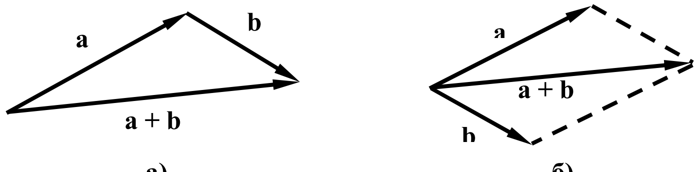

Рис. 1.1. Графічне означення суми векторів

Зауваження 1.3. Представлене на рис. 1, *б* правило паралелограма еквівалентне правилу трикутника. Воно використовується при додаванні векторів, зведених до спільного початку.

Зауваження 1.4. З правила трикутника та правила паралелограма випливає, що додавання вектора $\mathbf{0}$ до будь-якого вектора $\mathbf{a}$ не змінює вектора $\mathbf{a}$, тобто справджується рівність $\mathbf{a}+\mathbf{0}=\mathbf{a}$.

Означення 1.6. Добутком вектора $\mathbf{a}$ на число λ називають вектор, довжина якого дорівнює добутку $|\lambda||\mathbf{a}|$, а напрямок збігається з напрямком вектора $\mathbf{a}$, якщо $\lambda>0$, і протилежний йому, якщо $\lambda<0$.

Зауваження 1.5. З цього означення випливає, що множення вектора $\mathbf{a}$ на число –1 змінює напрямок вектора на протилежний, не змінюючи його довжини. Тому пишуть $(-1)\mathbf{a}=-\mathbf{a}$ і називають вектор $-\mathbf{a}$ протилежним до вектора $\mathbf{a}$. Суму вектора $\mathbf{b}$ і вектора $-\mathbf{a}$ називають різницею векторів і вважають, що віднімання векторів не є принципово новою операцією, оскільки вона зводиться до додавання векторів та множення вектора на число:

$$\mathbf{b}-\mathbf{a}=\mathbf{b}+(-1)\mathbf{a}. \tag{1.1}$$

У ході практично важливих розрахунків необхідно використовувати такі ***властивості алгебраїчних дій з векторами***:

1) $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (комутативність додавання векторів);

2) $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ (асоціативність додавання векторів);

3) $\mu(\lambda\mathbf{a})=(\mu\lambda)\mathbf{a}$ (асоціативність множення вектора на число);

4) $(\mu+\lambda)\mathbf{a}=(\mu\mathbf{a}+\lambda\mathbf{a})$ (дистрибутивність множення вектора на число відносно числа);

5) $\lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}$ (дистрибутивність множення вектора на число відносно вектора).

Наявність властивостей 1) – 5) можна перевірити, виходячи з *означень* 1.4 та 1.5, шляхом досить простих геометричних побудов. Для ілюстрації цього твердження розглянемо побудови, що доводять дві з цих властивостей.

На рис. 1.2 показано побудову, що доводить комутативність додавання векторів. На цьому рисунку перші доданки суми векторів показані штрихованими, другі – штрих-пунктирними, а суми доданків – суцільними лініями. Вектори, зображені суцільними лініями, мають однакові модулі, лежать на паралельних прямих і напрямлені в один і той самий бік, а отже, ці вектори дорівнюють один одному.

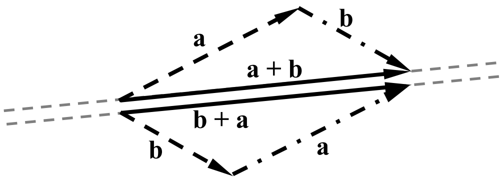

Рис. 1.2. Графічне доведення комутативності суми векторів

На рис. 1.3 показано побудову, що доводить дистрибутивність множення вектора на число відносно вектора: трикутники *ABC* та *ADE* подібні, тому що довжини сторін *AB* та *AC* трикутника *ABC* пропорційні довжинам сторін *AD* та *AE* трикутника *ADE* і кут між сторонами *AB* та *AC* дорівнює куту між сторонами *AD* та *AE*; коефіцієнт пропорційності дорівнює λ, а отже, $AE = \lambda \cdot AC$, тобто $\lambda\mathbf{a} + \lambda\mathbf{b} = \lambda(\mathbf{a} + \mathbf{b})$. Доведення властивостей 2), 3) та 4) наведено в підручниках [1, 2].

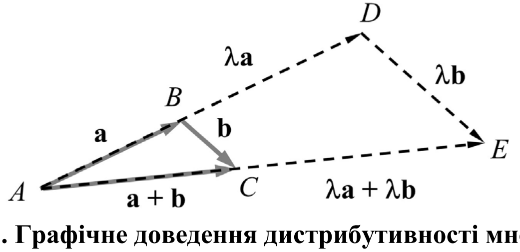

Рис. 1.3. Графічне доведення дистрибутивності множення вектора на число відносно числа

1.3. Метод координат

Для природничих наук та інженерії надзвичайно важливим є кількісний аналіз тих явищ, що описуються векторними величинами. Такий аналіз значною мірою спирається на *метод координат*, який дозволяє розв'язувати практично важливі задачі, що стосуються векторних величин, за допомогою математичних операцій із числами.

У багатьох випадках у методі координат використовується прямокутна система координат на площині або в просторі, яку називають декартовою системою координат на честь видатного французького вченого Рене Декарта (1596–1650). Прямокутна система координат загальновідома зі шкільного курсу математики. Для кращого розуміння того, що викладено далі, доцільно згадати про складові частини прямокутної системи та її застосування в методі координат.

На рис. 1.4 показано прямокутну систему координат у просторі. Вона складається з трьох координатних осей – осі абсцис *OX*, осі ординат *OY* та осі аплікат *OZ*, що перетинаються у спільному початку координат *O*. Уздовж цих осей напрямлені орти $\mathbf{e}_x$, $\mathbf{e}_y$ та $\mathbf{e}_z$.

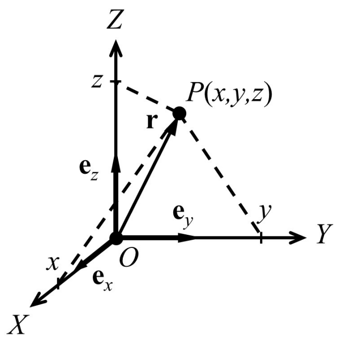

Рис. 1.4. Проєкції точки простору на осі прямокутної системи координат
(штрихові лінії перпендикулярні до координатних осей)

Розглянемо довільну точку P фізичного простору. Опускаючи перпендикуляри з цієї точки на осі координат, знаходимо точки, які є проєкціями точки P. Вектор $\mathbf{r}$, початок якого лежить на початку координат, а кінець – у точці P, називають радіус-вектором цієї точки. Якщо вектор $\mathbf{r}$ утворює гострі кути з осями координат, то координатами $\{x, y, z\}$ точки P у прямокутній системі координат називають відстані між проєкціями цієї точки на координатні осі та початком координат (рис. 1.4). Координати точки P водночас є координатами її радіус-вектора. Якщо кут між вектором $\mathbf{r}$ і якоюсь з координатних осей тупий, то відповідна координата цього вектора і точки P дорівнює відстані між проєкцією точки на цю вісь зі знаком мінус. Отже, координати векторів можуть набувати як додатних, так і від'ємних значень.

Означення 1.7. Сукупність ортів $\mathbf{e}_x$, $\mathbf{e}_y$ та $\mathbf{e}_z$ називають базисом у тривимірному просторі.

Зауваження 1.6. Працюючи з вільними векторами, зручно вважати, що початки всіх ортів лежать на початку координат. Якщо не буде обумовлено інше, вважатимемо, що й початок кожного іншого геометричного вектора також лежить на початку координат.

Означення 1.8. Якщо не обумовлено інше, координатами геометричного вектора називають взяті з належними знаками відстані від початку координат до проєкцій кінця вектора на осі OX, OY та OZ.

Координати вектора $\mathbf{a}$, що відповідають його проєкціям на осі OX, OY, OZ, позначають символами a_x, a_y, a_z, відповідно. Координата a_i ($i = x, y, z$) є додатним числом, якщо кут між вектором $\mathbf{a}$ та ортом $\mathbf{e}_i$ гострий, і від'ємним числом, якщо цей кут тупий. Якщо вектор $\mathbf{a}$ лежить у площині XOY, то з правила паралелограма, яке в цьому разі стає правилом прямокутника, випливає вираз

$$\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y . \tag{1.2}$$

У загальному випадку вектор $\mathbf{a}$ не лежить у жодній із координатних площин. Тоді, двічі застосувавши правило прямокутника, легко побудувати спочатку вектор $a_x\mathbf{e}_x + a_y\mathbf{e}_y$, а потім вектор

$$\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z, \tag{1.3}$$

(див. рис. 1.5). Вираз (1.3) називають розкладом вектора $\mathbf{a}$ по базисних векторах декартової системи координат. Отже, вектор $\mathbf{a}$ є сумою координатних ортів, помножених на координати вектора (рис. 1.5).

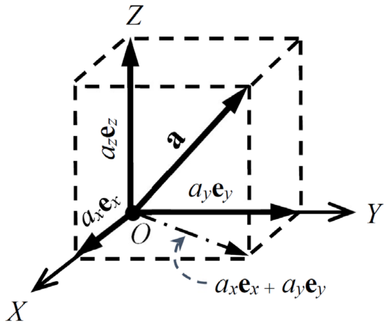

Рис. 1.5. Розклад вектора по векторах базису

Розклади векторів по векторах базису та правила 1) – 5) дають змогу виконувати операції з векторами, не вдаючись до геометричних побудов. Щоб скоротити формули, покажемо це на прикладі векторів, які лежать у площині XOY: $\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y$, $\mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y$.

Спочатку розглянемо операцію додавання векторів. Покажемо, що координати суми векторів дорівнюють сумам координат векторів, що додаються. Спочатку використаємо асоціативність (властивість 2)), а потім комутативність (властивість 1)) операції додавання векторів:

$$\mathbf{a} + \mathbf{b} = (a_x\mathbf{e}_x + a_y\mathbf{e}_y) + (b_x\mathbf{e}_x + b_y\mathbf{e}_y) =$$
$$= a_x\mathbf{e}_x + (a_y\mathbf{e}_y + b_x\mathbf{e}_x) + b_y\mathbf{e}_y =$$
$$= a_x\mathbf{e}_x + (b_x\mathbf{e}_x + a_y\mathbf{e}_y) + b_y\mathbf{e}_y.$$

Ще раз використавши властивість 2), а потім – властивість 4), знаходимо розклад суми векторів по базисних ортах:

$$\mathbf{a}+\mathbf{b}=(a_x+b_x)\mathbf{e}_x+(a_y+b_y)\mathbf{e}_y\,. \qquad (1.4)$$

Числа у дужках є координатами суми векторів $\mathbf{a}$ та $\mathbf{b}$.

Тепер розглянемо операцію множення вектора на число. Покажемо, що в результаті множення вектора на число координати вектора помножуються на це число. Використаємо дистрибутивність множення вектора на число відносно вектора (властивість 5)):

$$\lambda\mathbf{a}=\lambda(a_x\mathbf{e}_x+a_y\mathbf{e}_y)=\lambda(a_x\mathbf{e}_x)+\lambda(a_y\mathbf{e}_y)\,.$$

Урахувавши асоціативність множення вектора на число (властивість 3)), одержуємо

$$\lambda\mathbf{a}=(\lambda a_x)\mathbf{e}_x+(\lambda a_y)\mathbf{e}_y\,. \qquad (1.5)$$

Числа у дужках є координатами добутку вектора $\mathbf{a}$ на число λ.

1.4. Лінійна комбінація та лінійна незалежність векторів

Правила обчислення координат суми векторів та добутку вектора на число були знайдені без використання геометричних побудов. Отже, ці правила застосовні не лише для прямокутної системи координат, а й для будь-якої координатної системи, у якій справедливі вирази подібні до розкладу вектора по ортах базису (вирази (1.2), (1.3)). Ці розклади є окремими випадками загального поняття "лінійна комбінація векторів". Розглянемо це поняття.

Вищенаведені означення операцій з векторами уможливлюють побудову вектора

$$\boldsymbol{l}=\alpha\mathbf{a}+\beta\mathbf{b}+\gamma\mathbf{c}+\ldots+\rho\mathbf{q}\,. \qquad (1.6)$$

Означення 1.9. Вектор $\boldsymbol{l}$ називають *лінійною комбінацією* векторів $\mathbf{a},\mathbf{b},\mathbf{c},\ldots,\mathbf{q}$. Числа α, β, γ, … ρ називають *коефіцієнтами* лінійної комбінації векторів. Лінійну комбінацію називають *тривіальною*, якщо всі її коефіцієнти дорівнюють нулю, і *нетривіальною*, якщо хоча б один коефіцієнт не дорівнює нулю.

Цілком очевидно, що тривіальна лінійна комбінація будь-яких векторів дорівнює вектору $\mathbf{0}$. А нетривіальна лінійна комбінація може як дорівнювати, так і не дорівнювати вектору нульової довжини, залежно від того, з яких векторів вона складається і чому дорівнюють її коефіцієнти.

Означення 1.10. Вектори $\mathbf{a}, \mathbf{b}, \mathbf{c}, ..., \mathbf{q}$ називають *лінійно незалежними*, якщо з них неможливо утворити *нетривіальну* лінійну комбінацію, що дорівнює вектору $\mathbf{0}$, тобто не існує таких коефіцієнтів α, β, γ, ... ρ, які б задовольнили умову $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} + ... + \rho\mathbf{q} = \mathbf{0}$, і при цьому не всі вони дорівнювали б нулю.

Важливим прикладом лінійно незалежних векторів є орти прямокутної системи координат, оскільки з теореми Піфагора випливає, що довжина лінійної комбінації цих ортів, тобто вектора $\boldsymbol{l} = \alpha\mathbf{e}_x + \beta\mathbf{e}_y + \gamma\mathbf{e}_z$, дорівнює кореню квадратному з його координат: $|\boldsymbol{l}| = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$. Ця довжина дорівнює нулю *лише* тоді, коли всі три коефіцієнти лінійної комбінації дорівнюють нулю.

1.5. Умова лінійної залежності векторів та її наслідки

З означення 1.10 випливає така умова лінійної залежності векторів: вектори $\mathbf{a}, \mathbf{b}, \mathbf{c}, ..., \mathbf{q}$ лінійно залежні, якщо існують такі величини коефіцієнтів α, β, γ, ... ρ, які *не всі* дорівнюють нулю і при цьому задовольняють рівняння

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} + ... + \rho\mathbf{q} = \mathbf{0}. \tag{1.7}$$

Ця умова має такі наслідки, які дозволяють поширити поняття базисних векторів на функції і деякі інші математичні об'єкти та застосувати метод координат для розв'язання таких задач, у яких ці об'єкти зустрічаються. Перелічимо ці наслідки, обґрунтувавши їх геометрично.

Наслідок 1.1. Два *неколінеарні* вектори лінійно незалежні.

Щоб довести справедливість цього наслідку, достатньо розглянути будь-які вектори $\mathbf{a}$ та $\mathbf{b}$ і зауважити, що умова лінійної залежності цих векторів $\alpha\mathbf{a}+\beta\mathbf{b}=\mathbf{0}$ еквівалентна рівності $\mathbf{a}=-(\beta/\alpha)\mathbf{b}$, яка показує, що вектор $\mathbf{a}$ є добутком вектора $\mathbf{b}$ на число $-(\beta/\alpha)$. Із геометричного за його суттю означення 1.6 випливає, що для виконання цієї рівності вектор $\mathbf{b}$ має бути колінеарним з $\mathbf{a}$.

У справедливості наслідку 1.1 можна упевнитись також, додаючи вектори $\alpha\mathbf{a}$ та $\beta\mathbf{b}$ за правилом трикутника (рис. 1.6). Згідно з цим правилом нетривіальна лінійна комбінація $\alpha\mathbf{a}+\beta\mathbf{b}$ є стороною AC трикутника ABC, побудованого на векторах $\alpha\mathbf{a}$ та $\beta\mathbf{b}$.

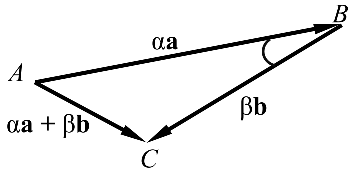

Рис. 1.6. Лінійна комбінація двох неколінеарних векторів

Із рис. 1.6 цілком очевидно, що довжина сторони трикутника AC, а отже, і довжина вектора $\alpha\mathbf{a}+\beta\mathbf{b}$, не дорівнює нулю, коли не дорівнює нулю кут між векторами $\alpha\mathbf{a}$ та $\beta\mathbf{b}$.

Наслідок 1.2. Вектор $\mathbf{c}$, компланарний з *неколінеарними* векторами $\mathbf{a}$ та $\mathbf{b}$, може бути розкладений по цих векторах, тобто існують числа c_1 та c_2, які задовольняють рівність

$$\mathbf{c}=c_1\mathbf{a}+c_2\mathbf{b}\,. \tag{1.8}$$

Цей наслідок умови лінійної залежності векторів може бути обґрунтований геометрично (рис. 1.7). На рис. 1.7, *а* показано, що лінійна комбінація $\alpha\mathbf{a}+\beta\mathbf{b}$ може відрізнятися від вектора $\mathbf{c}$ як за напрямком, так і за величиною. Зменшуючи або збільшуючи коефіцієнт β за незмінної величини коефіцієнта α, можна наближати напрямок лінійної комбінації $\alpha\mathbf{a}+\beta\mathbf{b}$ до напрямку вектора $\mathbf{a}$, або віддаляти її напрямок від напрямку цього вектора. У випадку, зображеному на рис. 1.7, *а*, зменшуючи коефіцієнт β, можна наближати напрямок лінійної комбінації до напрямку вектора $\mathbf{a}$ доти, доки вона не стане колінеарною з вектором $\mathbf{c}$. Після цього можна поступово зменшувати величину коефіцієнтів α та β, не змінюючи відношення α/β, що приведе до поступового зменшення довжини вектора $\alpha\mathbf{a}+\beta\mathbf{b}$, без зміни його напрямку. За певної величини коефіцієнтів $\alpha = c_1$ та $\beta = c_2$ не лише напрямки, але й модулі векторів $\alpha\mathbf{a}+\beta\mathbf{b}$ та $\mathbf{c}$, стануть однаковими, тобто справдиться рівність (1.8).

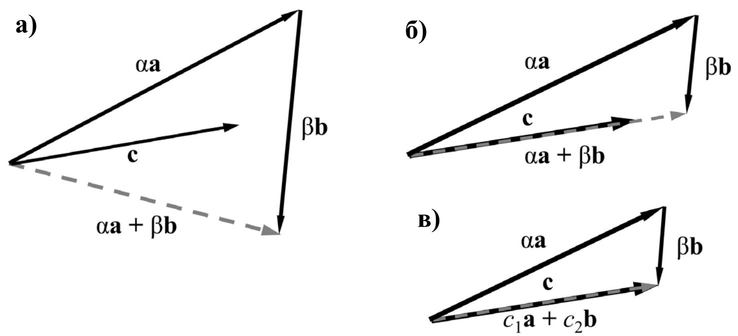

Рис. 1.7. Геометричне обґрунтування лінійної залежності трьох компланарних векторів:
(а) довільна лінійна комбінація векторів **a** та **b**;
(б) лінійна комбінація векторів **a** та **b,** колінеарна з вектором **c**;
(в) розклад вектора **c** по векторах **a** та **b**

Наслідок 1.3. Будь-який вектор $\mathbf{d}$ може бути розкладений по некомпланарних векторах $\mathbf{a}$, $\mathbf{b}$ та $\mathbf{c}$, тобто існують числа d_1 d_2 та d_3, які задовольняють рівність

$$\mathbf{d} = d_1\mathbf{a} + d_2\mathbf{b} + d_3\mathbf{c}. \tag{1.9}$$

Обґрунтуємо цей наслідок геометрично. Розглянемо три некомпланарні вектори $\mathbf{a}$, $\mathbf{b}$ та $\mathbf{c}$ довільної довжини, не вважаючи їх попарно перпендикулярними. Зведемо їх до спільного початку – точки O. Якщо четвертий вектор, $\mathbf{d}$, лежить у площині двох із цих трьох векторів ($\mathbf{a}, \mathbf{b}$ або $\mathbf{a}, \mathbf{c}$ або $\mathbf{b}, \mathbf{c}$), то він може бути розкладений по цих векторах (наслідок 1.2). У загальному випадку вектор $\mathbf{d}$ не лежить у жодній із цих площин. Із кінця вектора $\mathbf{d}$ проведемо пряму лінію, *паралельну з вектором* $\mathbf{c}$ (рис. 1.8). Ця лінія перетне площину векторів $\mathbf{a}, \mathbf{b}$ у певній точці P. Вектор, початок якого лежить у точці P, а кінець збігається з кінцем вектора $\mathbf{d}$, є колінеарним з вектором $\mathbf{c}$, а отже, він дорівнює добутку вектора $\mathbf{c}$ на якесь число γ. Побудуємо вектор $\mathbf{OP}$. Цей вектор лежить у площині векторів $\mathbf{a}, \mathbf{b}$, і згідно з наслідком 1.2 він є їхньою лінійною комбінацією: $\mathbf{OP} = \alpha\mathbf{a} + \beta\mathbf{b}$. Вектори $\mathbf{OP}$, $\gamma\mathbf{c}$ та $\mathbf{d}$ утворюють трикутник і задовольняють правило додавання векторів $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = \mathbf{d}$ (рис. 1.8). Звідси випливає, що числа α, β та γ є коефіцієнтами розкладу вектора $\mathbf{d}$ по векторах $\mathbf{a}, \mathbf{b}$ та $\mathbf{c}$: $\alpha = d_1$, $\beta = d_2$, $\gamma = d_3$.

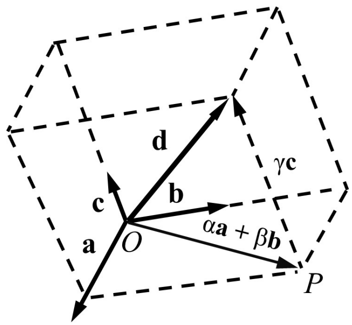

Рис. 1.8. Геометричне обґрунтування лінійної залежності чотирьох некомпланарних векторів

Зауважимо, що наслідок 1.2 і рівняння (1.8) були обґрунтовані геометрично для будь-яких трьох неколінеарних векторів на площині, а наслідок 1.3 та рівняння (1.9) – для будь-яких геометричних векторів у фізичному просторі. Оскільки ці рівняння тотожно перетворюються на умови лінійної залежності векторів $c_1\mathbf{a} + c_2\mathbf{b} - \mathbf{c} = \mathbf{0}$, $d_1\mathbf{a} + d_2\mathbf{b} + d_3\mathbf{c} - \mathbf{d} = \mathbf{0}$, то з наслідків 1.2 та 1.3 випливає важливий висновок.

Висновок 1.1. На площині можна побудувати не більше двох лінійно незалежних векторів, а у фізичному просторі – не більше трьох.

Зауваження 1.7. Оскільки для визначення довжини та напрямку вектора на площині достатньо задати два коефіцієнти розкладу (1.8) по неколінеарних векторах, то сукупність геометричних векторів на площині називають двовимірним простором. Фізичний простір з тих самих причин називають тривимірним (див. вираз (1.9)). Згідно зі зробленим висновком *вимірність кожного з двох розглянутих просторів дорівнює максимальній кількості лінійно незалежних векторів у цьому просторі.* Цей факт та Висновок 1.1 дуже важливий для аксіоматичної будови лінійної алгебри, тому що він дозволив означити поняття вимірності лінійного простору як максимальної кількості лінійно незалежних елементів цього простору (про це див. далі, розд. 3).

Зауваження 1.8. Порівняння виразів (1.3) та (1.9) показує, що поняття базису в тривимірному просторі можна поширити з випадку прямокутної системи координат на той випадок, коли кути між базисними векторами не дорівнюють 90° і модулі цих векторів не дорівнюють одиниці. Але базиси, утворені попарно перпендикулярними ортами, зазвичай найбільш зручні для розв'язання практично важливих задач. Такі базиси називають ортонормованими. Більше того, наслідки 1.1 та 1.2 дозволили поширити поняття базисних векторів і координат вектора на інші математичні об'єкти.

1.6. Скалярний добуток векторів. Обчислення довжини вектора та його координат

Означення 1.11. Скалярним добутком вектора $\mathbf{a}$ на вектор $\mathbf{b}$ називають добуток модулів цих векторів, помножений на кут між ними. Пишуть

$$\mathbf{ab} = |\mathbf{a}||\mathbf{b}|\cos(\mathbf{a}\,\hat{}\,\mathbf{b}), \qquad (1.10)$$

де $\mathbf{ab}$ – скалярний добуток векторів, $(\mathbf{a}\,\hat{}\,\mathbf{b})$ – кут між ними.

Скалярний добуток векторів – це число, тому можна сказати, що *скалярний добуток – це математичне правило, за яким кожній парі векторів відповідає число.* У випадку геометричних векторів це правило сформульовано у формі означення 1.1. Як буде показано далі, сформулювавши інші правила скалярного добутку, можна поширити поняття скалярного добутку на інші математичні об'єкти.

З означення скалярного добутку випливає, що проєкцію будь-якого вектора на напрямок іншого вектора можна обчислити за допомогою формули (1.10). Щоб це довести, розглянемо орт $\mathbf{e}_b$, спрямований уздовж напрямку вектора $\mathbf{b}$. Цей орт пов'язаний з вектором $\mathbf{b}$ рівнянням

$$\mathbf{b} = |\mathbf{b}|\,\mathbf{e}_b\,. \tag{1.11}$$

У справедливості рівняння (1.11) легко впевнитися, узявши до уваги, що $|\mathbf{e}_b| = 1$, тому права частина цього рівняння є вектором, модуль якого дорівнює $|\mathbf{b}||\mathbf{e}_b| = |\mathbf{b}|$, а отже, вектори у правій та лівій частинах рівняння однакові не лише за напрямком, а й за довжиною. Із рівняння (1.11) випливає, що

$$\mathbf{e}_b = \mathbf{b}/|\mathbf{b}|\,. \tag{1.12}$$

Із рівнянь (1.10) – (1.12) легко знайти формулу, яка зводить обчислення проєкції вектора $\mathbf{a}$ на напрямок будь-якого вектора $\mathbf{b}$ до обчислення його скалярних добутку на орт (1.12):

$$\mathbf{a}\mathbf{e}_b = |\mathbf{a}|\cos(\mathbf{a}\hat{\ }\mathbf{b})\,. \tag{1.13}$$

Розглянемо два геометричні вектори $\mathbf{a}$ і $\mathbf{b}$, сумістивши початок вектора $\mathbf{b}$ з кінцем вектора $\mathbf{a}$, і спроєктуємо ці вектори на напрямок вектора $\mathbf{c} = \mathbf{a} + \mathbf{b}$ (рис. 1.9).

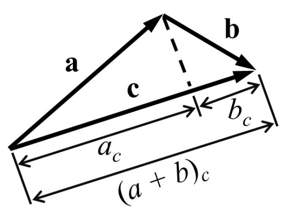

Рис. 1.9. Проєкції векторів a, b та a + b на напрямок вектора c

Із рис. 1.9 очевидно, що

$$|\mathbf{a}|\cos(\mathbf{a}^{\wedge}\mathbf{c}) = a_c, \tag{1.14}$$

де a_c – проєкція вектора $\mathbf{a}$ на напрямок вектора $\mathbf{c}$. Порівнявши вирази (1.13) та (1.14), одержуємо важливу формулу

$$a_c = \mathbf{a}\mathbf{e}_c. \tag{1.15}$$

Оскільки напрямок вектора $\mathbf{c}$ є водночас і напрямком прямої, напрямленої вздовж орта $\mathbf{e}_c$, то формула (1.15) дозволяє сформулювати означення проєкції вектора $\mathbf{a}$ на напрямок прямої у просторі, еквівалентне наданому вище означенню 1.8, але не пов'язане з геометричними побудовами.

Означення 1.12. Проєкцією геометричного вектора на напрямок прямої у просторі є скалярний добуток цього вектора на орт, напрямлений уздовж прямої.

Зауваження 1.9. Згідно з формулою (1.15) проєкція вектора $\mathbf{a}$ на напрямок вектора $\mathbf{c} = |\mathbf{c}|\mathbf{e}_c$ є додатним числом, якщо кут між цими векторами гострий, і від'ємним числом, якщо кут між векторами тупий.

Означення 1.12 дозволяє обчислювати координати вільних векторів у декартовій системі координат, не виконуючи геометричних побудов. Згідно з цим означенням

$$a_x = \mathbf{a}\mathbf{e}_x,\ a_y = \mathbf{a}\mathbf{e}_y,\ a_z = \mathbf{a}\mathbf{e}_z. \tag{1.16}$$

Якщо відомі координати двох векторів у прямокутній системі координат, вирази (1.16) дозволяють легко обчислити скалярний добуток цих векторів, не визначаючи ані їхніх довжин, ані кута між ними. Робочу формулу для такого обчислення легко отримати, розкривши дужки у формулі

$$\mathbf{ab} = (a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z)(b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z)$$

і врахувавши, що скалярні добутки $\mathbf{e}_i\mathbf{e}_k$ дорівнюють одиниці, якщо $i = k$, і нулю, якщо $i \neq k$:

$$\mathbf{ab} = a_x b_x + a_y b_y + a_z b_z. \tag{1.17}$$

Поклавши в цій формулі $\mathbf{a} = \mathbf{b}$ і врахувавши, що

$$|\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}, \tag{1.18}$$

одержуємо формули, які дозволяють поширити поняття "модуль вектора" та "кут між векторами" на різні математичні об'єкти, якщо для них попередньо було означено правило обчислення скалярного добутку:

$$|\mathbf{a}| = \sqrt{\mathbf{aa}}\,, \tag{1.19}$$

$$\cos(\mathbf{a}^\wedge\mathbf{b}) = \frac{\mathbf{ab}}{|\mathbf{a}||\mathbf{b}|}. \tag{1.20}$$

Те, як було отримано формулу (1.17), дає нагоду продемонструвати доцільність введення до розгляду так званого символу Кронекера.

Символ Кронекера є сукупністю чисел, які позначаються літерою з двома індексами, δ_{ik} і *за означенням* дорівнюють одиниці (якщо індекси i та k однакові) або нулю (якщо індекси i та k різні). Індекси символу Кронекера можуть бути або літерами (зазвичай x, y, z) або цілими числами, які в більшості задач набувають додатних значень від 1 до $N \ge 2$. Означення символу Кронекера легко запам'ятати, якщо записати його у формі

$$\delta_{ik} = \begin{cases} 1, & \text{якщо } i = k, \\ 0, & \text{в іншому випадку.} \end{cases} \tag{1.21}$$

За допомогою символу Кронекера скалярні добутки взаємно ортогональних ортів визначаються формулами

$$\mathbf{e}_i\mathbf{e}_k = \delta_{ik}\,,\ i = x, y, z\,, \tag{1.22}$$

а справедливість виразу (1.17) – послідовністю рівностей

$$\mathbf{ab} = \sum_i a_i\mathbf{e}_i \sum_k b_k\mathbf{e}_k = \sum_i\sum_k a_ib_k\mathbf{e}_i\mathbf{e}_k = \sum_i\sum_k a_ib_k\delta_{ik} = \sum_i a_ib_i\,.$$

Важливо, що взаємно ортогональні орти можна позначати літерами з такими індексами, які набувають цілочисельних значень.

1.7. Визначальні властивості скалярного добутку

Скалярному добутку будь-яких геометричних векторів притаманні такі властивості:

1) $\mathbf{ab} = \mathbf{ba}$ (комутативність);
2) $\lambda(\mathbf{ab}) = (\lambda\mathbf{a})\mathbf{b}$ (асоціативність з числом);
3) $(\mathbf{a}+\mathbf{b})\mathbf{c} = \mathbf{ac} + \mathbf{bc}$ (дистрибутивність);

4) $\mathbf{aa} \geq 0$ (скалярний добуток ненульового вектора самого на себе є додатним числом або дорівнює нулю – останнє, якщо $\mathbf{a} = \mathbf{0}$).

Властивості 1), 2) та 4) випливають безпосередньо з формули (1.10), яка є означенням скалярного добутку, а властивість 3) можна довести, виходячи з формул (1.10), (1.14) і того факту, що проєкція суми векторів дорівнює сумі їхніх проєкцій. За означенням скалярного добутку,

$$(\mathbf{a} + \mathbf{b})\mathbf{c} = |\mathbf{a} + \mathbf{b}||\mathbf{c}|\cos(\mathbf{a} + \mathbf{b} \wedge \mathbf{c}). \tag{1.23}$$

Застосовуємо формулу (1.14) до вектора $\mathbf{a} + \mathbf{b}$:

$$|\mathbf{a} + \mathbf{b}|\cos(\mathbf{a} + \mathbf{b} \wedge \mathbf{c}) = (a + b)_c. \tag{1.24}$$

Підставляємо праву частину рівняння (1.24) до правої частини рівняння (1.23):

$$(\mathbf{a} + \mathbf{b})\mathbf{c} = (a + b)_c |\mathbf{c}|.$$

Із рис. 1.9 очевидно, що $(a + b)_c = a_c + b_c$, що дозволяє завершити доведення властивості 3), ще раз взявши до уваги формулу (1.14):

$$(\mathbf{a} + \mathbf{b})\mathbf{c} = (a + b)_c |\mathbf{c}| = a_c |\mathbf{c}| + b_c |\mathbf{c}| = \mathbf{ac} + \mathbf{bc}.$$

Щоб зрозуміти логіку лінійної алгебри, слід зауважити, що властивості скалярного добутку геометричних векторів притаманні і деяким математичним операціям з іншими математичними об'єктами, наприклад функціями.

1.8. Подібність між векторами та функціями

Як відомо, числовою функцією числової змінної t називають правило, за яким кожному значенню змінної ставиться у відповідність одне-єдине число f. Розглянемо сукупність усіх функцій $f_j(t)$, які набувають дійсних значень і неперервні на заданому інтервалі значень *дійсної* змінної t (нижній індекс – додатне ціле число, поставлене, щоб відрізняти функції одна від одної). Зі шкільного курсу математики відомо, що операції додавання функцій притаманні комутативність та асоціативність, а операції множення функції на число – асоціативність, дистрибутивність відносно числа та дистрибутивність відносно функції, тобто:

1) $f_j(t)+f_k(t)=f_k(t)+f_j(t)$;

2) $[f_j(t)+f_k(t)]+f_m(t)=f_j(t)+[f_k(t)+f_m(t)]$;

3) $\mu[\lambda f_j(t)]=(\mu\lambda)f_j(t)$;

4) $(\mu+\lambda)f_j(t)=[\mu f_j(t)+\lambda f_j(t)]$;

5) $\lambda[f_j(t)+f_k(t)]=\lambda f_j(t)+\lambda f_k(t)$,

де λ, μ – дійсні або комплексні числа. Порівнявши ці властивості з властивостями операцій з геометричними векторами (див. властивості 1) – 4) у підрозд. 1.2), бачимо, що можна провести аналогію між функціями та векторами.

Зважаючи на те, що можна додавати функції одна до одної та множити їх на числа, можна утворювати лінійні комбінації функцій. Лінійною комбінацією функцій $f_j(t)$ називають функцію

$$l(t)=\sum_{j=1}^{N} f_j(t). \tag{1.25}$$

За аналогією з геометричними векторами функції називають лінійно незалежними, якщо з них не неможливо утворити *нетривіальну* лінійну комбінацію, що дорівнює нулю, а якщо така лінійна комбінація існує, то функції називають лінійно залежними, тобто, умовою лінійної залежності функцій є рівність

$$\sum_{j=1}^{N}\alpha_j f_j(t)=0,$$

у якій принаймні один з коефіцієнтів α_j не дорівнює нулю.

Якщо в сукупності $\{f_j(t)\}_{j=\overline{1,N}}$ існує n лінійно незалежних функцій, а будь-які $n+1$ функцій лінійно залежні, то будь-яку функцію із сукупності $\{f_j(t)\}_{j=\overline{1,N}}$ можна представити у формі лінійної комбінації n лінійно незалежних функцій. Це твердження легко довести, якщо надати лінійно незалежним функціям номери від одиниці до n і записати умову лінійної залежності n плюс однієї функції:

$$\sum_{j=1}^{n} \alpha_j f_j(t) + \alpha_k f_k(t) = 0\,, \tag{1.26}$$

де $n < k \le N$, $\alpha_k \ne 0$. Поділивши обидві частини рівності (1.26) на α_k і перенісши останній доданок праворуч, можна отримати вираз

$$f_k(t) = \sum_{j=1}^{n} C_j f_j(t)\,, \tag{1.27}$$

який називають розкладом функції $f_k(t)$ по функціях сукупності $\{f_j(t)\}_{j=\overline{1,n}}$. За аналогією з геометричними векторами цю сукупність називають базисом[1].

Аналогія між векторами і функціями підсилюється порівнянням визначеного інтеграла від добутку двох функцій

$$\int_a^b f_j(t) f_k(t)\,dt \tag{1.28}$$

зі скалярним добутком векторів. По-перше, визначений інтеграл обчислюється за правилами, які кожній парі функцій ставлять у відповідність певне число, а скалярний добуток – за правилами, які визначають число, відповідне до кожної пари геометричних векторів. По-друге, визначений інтеграл має властивості, подібні до визначальних властивостей скалярного добутку, сформульованих у підрозд. 1.7, а саме:

1) $\int_a^b f_j(t) f_k(t)\,dt = \int_a^b f_k(t) f_j(t)\,dt$ (комутативність);

2) $\lambda \int_a^b f_j(t) f_k(t)\,dt = \int_a^b [\lambda f_j(t)] f_k(t)\,dt$ (асоціативність з числом);

3) $\int_a^b [f_j(t) + f_k(t)] f_m(t)\,dt = \int_a^b f_j(t) f_m(t)\,dt + \int_a^b f_k(t) f_m(t)\,dt$ (дистрибутивність);

[1] Питання про те, чи є сукупність $\{f_j(t)\}_{j=\overline{1,N}}$ підпростором простору неперервних функцій, тут залишаємо відкритим.

4) $\int_a^b f_j(t) f_j(t) dt \geq 0$ (рівність нулю відповідає тому випадку, коли $f_j(t) = 0$ в усьому інтервалі інтегрування).

Із викладеного в розд. 1 випливає такий висновок: подібність між геометричними векторами та функціями підказує ідею створення математичної дисципліни, яка вивчає вектори і функції із загальних позицій. Ця ідея підсилюється тим, що існують й інші математичні об'єкти, яким притаманні властивості операцій з геометричними векторами і для яких існує правило обчислення скалярного добутку. Серед таких об'єктів дуже важливими, з огляду на їх застосування у багатьох розділах математики, є матриці.

РОЗДІЛ 2
Алгебра матриць

2.1. Означення матриці

У підручниках з вищої алгебри *матрицю* традиційно означають як прямокутну таблицю

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \tag{2.1}$$

де *елементи матриці* a_{ij} є числами з множини всіх дійсних або комплексних чисел, перший і другий індекс кожного елемента є номером рядка та стовпця таблиці, у яких розташований елемент [3]. Слід розуміти, що це означення неповне. Доповнимо його.

Означення 2.1. Матрицями називають *впорядковані* сукупності елементів[1], записані у формі таблиць, для яких означені правила трьох алгебраїчних дій, а саме:
а) додавання матриць;
б) множення матриці на число;
в) множення матриці на матрицю.

Щоб довершити означення матриці, необхідно означити поняття рівності матриць, пояснити поняття впорядкованої сукупності елементів матриці та означити правила алгебраїчних дій з матрицями.

Означення 2.2. Матриця A з елементами a_{ij} дорівнює матриці B з елементами b_{ij}, якщо *дорівнюють один одному відповідні*

---

[1] Елементами матриці у більшості практично важливих задач є числа або функції.

елементи цих матриць, тобто для всіх пар індексів ij справджуються рівності

$$a_{ij} = b_{ij} . \tag{2.2}$$

Із наданого означення випливає, що впорядкованість елементів матриці полягає в тому, що переставлення місцями неоднакових елементів матриці A або B порушує рівність (2.2)[1]. Без означення рівності матриць поняття впорядкованості їхніх елементів не мало б сенсу.

У багатьох випадках матрицю A, що складається з елементів a_{ij}, зручно позначати як $A(a_{ij})$, матрицю B, що складається з елементів b_{ij}, – як $B(b_{ij})$ і т. д.

Означення 2.3. Матрицю, яка складається з m рядків та n стовпців, називають матрицею порядку $m \times n$.

Означення 2.4. Сумою матриць $A(a_{ij})$ та $B(b_{ij})$ називають матрицю $C(c_{ij})$ з елементами

$$c_{ij} = a_{ij} + b_{ij} , \ i = \overline{1,m} , \ \ j = \overline{1,n} . \tag{2.3}$$

Пишуть $C = A + B$.

Означення 2.5. Добутком матриці $A(a_{ij})$ на число λ називають матрицю $B(b_{ij})$ з елементами

$$b_{ij} = \lambda a_{ij} . \tag{2.4}$$

Пишуть $B = \lambda A$.

Означення 2.6. Добутком матриці $A(a_{ij})$ порядку $m \times n$ на матрицю $B(b_{ij})$ порядку $n \times p$ називають матрицю $C(c_{ij})$ порядку $m \times p$ з елементами

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} . \tag{2.5}$$

Пишуть $C = AB$.

---

[1] Отже, декартові координати векторів є впорядкованими трійками чисел (x, y, z). Це робить матриці, що складаються лише з одного рядочка або лише одного стовпчика, подібними до векторів.

Зауваження 2.1. З наданих означень випливає, що:

а) додавати можна лише матриці однакового порядку;

б) помножати на число можна матриці будь-якого порядку;

в) добуток двох матриць існує лише в тому разі, коли кількість стовпців першого співмножника дорівнює кількості рядків другого.

Тепер можна зрозуміти, чому означення матриці було неповним, доти, доки не були означені правила алгебраїчних дій з матрицями. Пояснімо це на простому прикладі. У багатьох спеціалізованих комп'ютерних програмах функції однієї змінної задаються в такий спосіб: спочатку задається масив значень аргументу функції, наприклад, $x := 1, 2 \ldots N$, а потім у певний спосіб задається сама функція. У результаті функція виявляється представленою у вигляді впорядкованої сукупності чисел, що має форму таблиці:

$$f(x) = \begin{pmatrix} 1 & f_1 \\ 2 & f_2 \\ \vdots & \vdots \\ N & f_N \end{pmatrix}. \tag{2.6}$$

Згідно з наданим вище традиційним означенням функція (2.6) є матрицею. Згідно з означенням 2.5 добутку матриці на число має виконуватись така рівність:

$$2f(x) = \begin{pmatrix} 2\cdot 1 & 2\cdot f_1 \\ 2\cdot 2 & 2\cdot f_2 \\ \vdots & \vdots \\ 2\cdot N & 2\cdot f_N \end{pmatrix}, \tag{2.7}$$

тобто $2f(x) = 2f(2x)$. Якщо $f(x) \neq \text{const}\cdot x$, ця рівність помилкова, а отже, таблиця (2.6) не є матрицею, хоча вона і задовольняє традиційне означення цього математичного поняття у вищій алгебрі[1].

---

[1] Зауважимо, що поняття "матриця" використовують не лише у вищій алгебрі, а й в інших розділах математики, і там означення цього поняття може відрізнятися від означення 2.1.

2.2. Властивості дій з матрицями, комутатор матриць

Властивості додавання матриць і множення матриці на число аналогічні з тими, що притаманні цим операціям з геометричними векторами. Перелічимо їх:

1) $A+B=B+A$ (комутативність додавання матриць);
2) $(A+B)+C=A+(B+C)$ (асоціативність додавання матриць);
3) $\mu(\lambda A)=(\mu\lambda)A$ (асоціативність множення матриці на число);
4) $(\mu+\lambda)A=(\mu A+\lambda A)$ (дистрибутивність множення матриці на число відносно числа);
5) $\lambda(A+B)=\lambda A+\lambda B$ (дистрибутивність множення матриці на число відносно матриці).

Операції додавання матриць і множення їх на число зводяться до цих дій з елементами матриць, тобто з числами або функціями, тому їхні властивості такі ж, як операцій з числами, і доводити справедливість рівностей 1) – 5) немає необхідності. На відміну від цього, операція множення матриць специфічна саме для них, і тому її властивості заслуговують на особливу увагу. Перелічимо ці властивості:

1^0. $\lambda(AB)=(\lambda A)B=A(\lambda B)$;

2^0. $(AB)C=A(BC)$;

3^0. $(A_1+A_2)B=A_1B+A_2B$;

4^0. $A(B_1+B_2)=AB_1+AB_2$.

Основною специфічною рисою добутку матриць є відсутність комутативності, але це не значить, що $AB\neq BA$ для *всіх* матриць. Тому важливою для проведення розрахунків величиною є *комутатор двох матриць* $[A,B]$, означений як

$$[A,B]\overset{def}{\equiv}AB-BA. \qquad (2.8)$$

Властивості добутку матриць 1^0 – 4^0 не такі очевидні, як властивості 1) – 5), тому що його обчислення передбачає комбінацію операцій множення та додавання елементів матриць-співмножників (див. вираз (2.5)). Щоб пояснити принцип доведення цих властивостей, але уникнути надто великої кількості формул, доведемо лише найменш очевидну з цих властивостей, а саме, властивість 2^0. Позначимо $\langle AB\rangle_{ik}$ – елемент матриці AB; $\langle (AB)C\rangle_{ij}$ – елемент матриці $(AB)C$. Застосуємо правило обчислення добутку матриць (2.5):

$$\langle AB\rangle_{ik} = \sum_{l} a_{il} b_{lk} ,$$

$$\langle (AB)C\rangle_{ij} = \sum_{k} \langle AB\rangle_{ik} c_{kl} = \sum_{k} \left(\sum_{l} a_{il} b_{lk} \right) c_{kj} =$$

$$= \sum_{l} \left(a_{il} \sum_{k} b_{lk} c_{kj} \right) = \sum_{l} a_{il} \langle BC\rangle_{lj} = \langle A(BC)\rangle_{ij} .$$

2.3. Термінологія алгебри матриць

1. Матрицею, *транспонованою* до матриці $A(a_{ij})$, називають матрицю $A^T(\tilde{a}_{ij})$ з елементами

$$\tilde{a}_{ij} = a_{ji} , \tag{2.9}$$

тобто рядки матриці A^T є стовпцями матриці A.

2. Матрицею, *протилежною* до матриці $A(a_{ij})$, називають матрицю $-A(-a_{ij})$, тобто всі елементи протилежної матриці відрізняються від елементів матриці A лише знаком; суму A та матриці, протилежної до матриці B, називають різницею матриць і пишуть

$$A + (-B) \equiv A - B . \tag{2.10}$$

3. Якщо всі елементи матриці дорівнюють нулю, її називають нульовою матрицею, або просто нулем, і позначають символом 0.

4. Якщо кількість рядків матриці дорівнює кількості стовпців ($m = n$), то матрицю називають *квадратною* матрицею порядку n.

5. Сукупність елементів $\{a_{ii}\}_{i=\overline{1,n}}$ *квадратної* матриці називають її *основною діагоналлю*, а елементи a_{ii} – діагональними елементами.

6. Суму діагональних елементів *квадратної* матриці A називають її *слідом* і позначають $\mathrm{Tr}\,A$, або $\mathrm{Sp}\,A$:

$$\mathrm{Sp}A = \sum_i a_{ii} \; . \qquad (2.11)$$

7. Якщо всі елементи квадратної матриці, окрім діагональних, дорівнюють нулю, матрицю називають діагональною.

8. Якщо всі елементи діагональної матриці дорівнюють одиниці, її називають одиничною матрицею. Сукупність усіх елементів діагональної матриці є символом Кронекера (див. (1.21)), тому позначатимемо таку матрицю $E(\delta_{ij})$, або просто E.

Зауваження 2.2. З означення операцій додавання та множення матриць випливає, що:

а) $A + 0 = A$, якщо ця сума існує;

б) $AE = EA = A$, якщо ці добутки існують.

Перша із цих рівностей очевидна, оскільки додавання матриць виконується шляхом додавання окремих елементів. Друга рівність випливає з того, що

$$\langle AE \rangle_{ik} = \sum_j a_{ij}\delta_{jk} = a_{ik} \, , \; \langle EA \rangle_{ik} = \sum_j \delta_{ij} a_{jk} = a_{ik} \, ,$$

тому що $\delta_{jk} = 0$, якщо $j \neq k$; $\delta_{ij} = 0$, якщо $j \neq i$.

2.4. Приклади дій з матрицями

● **Приклад 2.1.** Додавання матриць та множення матриці на число здійснюються дуже просто, тому надамо приклад цих дій без пояснень:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 2 & 2 \\ 3 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 4 & 5 \\ 6 & 0 & 2 \end{pmatrix},$$

$$5\begin{pmatrix} -1 & 2 & 3 \\ 3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} -5 & 10 & 15 \\ 15 & 0 & 20 \end{pmatrix}. \bullet$$

Операція множення матриці на матрицю заслуговує на пояснення.

● **Приклад 2.2.** Для початку розглянемо добуток матриці порядку (1×2) на матрицю порядку (2×1):

$$AB=\begin{pmatrix}-1 & 3\end{pmatrix},\ B=\begin{pmatrix}4\\2\end{pmatrix}. \tag{2.12}$$

Добуток AB можна обчислити за формулою (2.5). Оскільки у цьому випадку кожен з індексів набуває лише одного значення ($i=1$, $j=1$), то формула (2.5) набуває форми

$$c_{11}=a_{11}b_{12}+a_{12}b_{21}=-1\cdot 4+3\cdot 2=2,$$

а отже, $AB=2$. Однак на практиці для обчислення добутку матриць зручно поставити матриці-співмножники поруч і записати ряд рівностей

$$AB=\begin{pmatrix}-1 & 3\end{pmatrix}\begin{pmatrix}4\\2\end{pmatrix}=-1\cdot 4+3\cdot 2=2.$$

Ці рівності отримано в такий спосіб:

✓ рядок $\begin{pmatrix}-1 & 3\end{pmatrix}$ подумки перетворили на стовпець $\begin{pmatrix}-1\\3\end{pmatrix}$;

✓ уявили, два стовпці $\begin{pmatrix}-1\\3\end{pmatrix}\begin{pmatrix}4\\2\end{pmatrix}$;

✓ перемножили числа, що стоять поруч і додали добутки цих чисел один до одного:

$$-1\cdot 4+3\cdot 2=2.●$$

● **Приклад 2.3.** Обчислимо в описаний вище спосіб добуток матриць другого порядку:

$$\begin{pmatrix}-1 & 2\\3 & -4\end{pmatrix}\begin{pmatrix}2 & 1\\4 & 5\end{pmatrix}=\begin{pmatrix}-1\cdot 2+2\cdot 4 & -1\cdot 1+2\cdot 5\\3\cdot 2-4\cdot 4 & 3\cdot 1-4\cdot 5\end{pmatrix}=\begin{pmatrix}6 & 9\\-10 & -17\end{pmatrix}.●$$

● **Приклад 2.4.** Щоб продемонструвати, що добуток матриць не комутативний, обчислимо добуток переставлених місцями матриць (2.12):

$$BA=\begin{pmatrix}4\\2\end{pmatrix}\begin{pmatrix}-1 & 3\end{pmatrix}=\begin{pmatrix}4\cdot(-1) & 4\cdot 3\\2\cdot(-1) & 2\cdot 3\end{pmatrix}=\begin{pmatrix}-4 & 12\\-2 & 6\end{pmatrix}.$$

Отже, добуток AB є числом, а BA – квадратною матрицею другого порядку.●

● **Приклад 2.5.** Якщо добутки AB та BA є квадратними матрицями однакового порядку, існує комутатор цих матриць. Покажемо це на прикладі матриць

$$A=\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \text{ та } B=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Обчислимо комутатор цих матриць:

$$AB-BA=\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}-\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}=$$

$$=\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}-\begin{pmatrix} 0 & b \\ a & 0 \end{pmatrix}=\begin{pmatrix} 0 & a-b \\ b-a & 0 \end{pmatrix}.$$

У результаті маємо

$$[A,B]=(a-b)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \bullet$$

При першому ознайомленні з алгеброю матриць може виникнути запитання: чому операцію множення матриць означено саме в такий, а не якийсь інший спосіб? Відповідь на це запитання досить проста: таке означення добутку матриць дозволяє застосувати матриці у багатьох розділах математики. Важливим прикладом є використання матриць в одному з важливих правил векторної алгебри, а саме, у правилі перетворення координат радіус-вектора просторової точки при повороті системи координат.

● **Приклад 2.6.** Розглянемо поворот прямокутної системи координат XOY на кут α навколо точки початку координат O. Загальновідомою формою закону перетворення координат є така система двох рівнянь:

$$\begin{cases} x'=x\cos\alpha+y\sin\alpha \\ y'=-x\sin\alpha+y\cos\alpha. \end{cases}$$

Означення 2.6 дає змогу записати цю систему у формі матричного рівняння:

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Це матричне рівняння підказує можливість введення до розгляду *оператора повороту*, $\hat{R}(\alpha)$, який діє на радіус-вектор, перетворюючи його координати за допомогою матриці

$$A=\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Якщо здійснити послідовно два повороти (спочатку на кут α, а потім на кут β), то правило перетворення координат набуде форми матричного рівняння

$$\begin{pmatrix} x" \\ y" \end{pmatrix}=\begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Обчисливши добуток двох матриць у правій частині цього рівняння, легко перевірити, що воно перетворюється на рівняння

$$\begin{pmatrix} x" \\ y" \end{pmatrix}=\begin{pmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ -\sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

яке ілюструє цілком очевидний факт: два послідовних повороти навколо фіксованої осі на кути α та β еквівалентні повороту на кут $\alpha+\beta$.●

Згаданий вище оператор повороту, що діє на геометричний вектор, узагальнюється в лінійній алгебрі, як поняття лінійного оператора, про яке йтиметься далі.

РОЗДІЛ 3
Аксіоматика лінійної алгебри

3.1. Лінійні простори

3.1.1. Означення лінійного простору та його наслідки

Поняття лінійного простору дозволяє поширити алгебру геометричних векторів на інші математичні об'єкти, які традиційно називають *елементами*, або *векторами*. Означення лінійного простору є сукупністю умов, які доцільно розділити на три пункти: а), б), в).

Означення 3.1. Сукупність елементів $\{\mathbf{x}_i\}_{i=1,2,3,...}$ є лінійним простором $\mathcal{L}$, якщо:

а) означені правила обчислення суми цих елементів та множення їх на число;

б) сума будь-яких елементів сукупності та добуток будь-якого елемента сукупності на число λ є елементом цієї ж сукупності;

в) усі елементи сукупності та правила алгебраїчних дій з ними мають 5 властивостей, притаманних діям з геометричними векторами, та 3 додаткові властивості, обумовлення яких необхідне з огляду на те, що уявлення про довжину і напрямок геометричного вектора не можна безпосередньо поширити на елементи не визначеного заздалегідь типу:

1^0. $\mathbf{x}_i + \mathbf{x}_j = \mathbf{x}_j + \mathbf{x}_i$.

2^0. $(\mathbf{x}_i + \mathbf{x}_j) + \mathbf{x}_k = \mathbf{x}_i + (\mathbf{x}_j + \mathbf{x}_k)$.

3^0. $\mu(\lambda \mathbf{x}_i) = (\mu\lambda)\mathbf{x}_i$.

4^0. $(\mu + \lambda)\mathbf{x}_i = \mu \mathbf{x}_i + \lambda \mathbf{x}_i$.

5^0. $\mu(\mathbf{x}_i + \mathbf{x}_j) = \mu \mathbf{x}_i + \mu \mathbf{x}_j$.

6^0. У сукупності $\{\mathbf{x}_i\}$ є елемент $\mathbf{0}$, такий що $\mathbf{x}_i + \mathbf{0} = \mathbf{x}_i$.

7^0. Поряд з кожним елементом $\mathbf{x}_i$ сукупність $\{\mathbf{x}_i\}$ містить елемент $-\mathbf{x}_i$, який задовольняє рівність $\mathbf{x}_i + (-\mathbf{x}_i) = \mathbf{0}$.

8^0. $1 \cdot \mathbf{x}_i = \mathbf{x}_i$.

Властивості 1^0 – 8^0 мають виконуватись для всіх лінійних просторів, тому їх називають *аксіомами лінійної алгебри*.

Означення 3.2. Елемент $\mathbf{0}$ називають нульовим елементом, а елемент $-\mathbf{x}_i$ – елементом, протилежним до вектора $\mathbf{x}_i$.

Означення 3.3. Якщо у властивостях 1^0 – 8^0 числа λ, μ – дійсні, то сукупність елементів $\{\mathbf{x}_i\}$ називають дійсним лінійним простором, а якщо ці числа належать до множини комплексних чисел, то сукупність $\{\mathbf{x}_i\}$ називають комплексним лінійним простором.

Аксіоми лінійної алгебри мають важливі наслідки. Розглянемо два з них.

Наслідок 3.1. Добуток числа нуль на будь-який елемент простору $\mathcal{L}$ дорівнює нульовому елементу.

Наслідок 3.2. Добуток числа (-1) на будь-який елемент $\mathbf{x}$, що належить простору $\mathcal{L}$, дорівнює елементу $-\mathbf{x}$.

Виходячи з принципу побудови аксіоматичної теорії, справедливість наслідків 3.1 та 3.2 слід довести, використовуючи *лише* аксіоми 1^0 – 8^0 та правила алгебраїчних дій з числами.

Наслідок 3.1 можна виразити у формі рівності

$$0 \cdot \mathbf{x} = \mathbf{0}. \tag{3.1}$$

Доведемо її справедливість, указавши ряд послідовних рівностей та номери аксіом, з яких ці рівності випливають:

$$0 \cdot \mathbf{x} \overset{6^0}{=} 0 \cdot \mathbf{x} + \mathbf{0} \overset{7^0}{=} 0 \cdot \mathbf{x} + [\mathbf{x} + (-\mathbf{x})] \overset{2^0}{=} (0 \cdot \mathbf{x} + \mathbf{x}) + (-\mathbf{x}) \overset{8^0}{=}$$

$$= (0 \cdot \mathbf{x} + 1 \cdot \mathbf{x}) + (-\mathbf{x}) \overset{4^0}{=} (0+1)\mathbf{x} + (-\mathbf{x}) = 1 \cdot \mathbf{x} + (-\mathbf{x}) \overset{8^0}{=} [\mathbf{x} + (-\mathbf{x})] \overset{7^0}{=} \mathbf{0}.$$

Виразимо наслідок 3.2 у формі рівності

$$(-1)\mathbf{x} = -\mathbf{x}, \tag{3.2}$$

та доведемо її справедливість:

$$(-1)\mathbf{x}\overset{6^0}{=}(-1)\mathbf{x}+\mathbf{0}\overset{7^0}{=}(-1)\mathbf{x}+[\mathbf{x}+(-\mathbf{x})]\overset{2^0}{=}[(-1)\mathbf{x}+\mathbf{x}]+(-\mathbf{x})\overset{8^0}{=}$$

$$=[(-1)\mathbf{x}+1\mathbf{x}]+(-\mathbf{x})\overset{4^0}{=}(-1+1)\mathbf{x}+(-\mathbf{x})=0\cdot\mathbf{x}+(-\mathbf{x})\overset{(3.1)}{=}$$

$$=[\mathbf{0}+(-\mathbf{x})]\overset{6^0}{=}-\mathbf{x}.$$

Наведемо приклади лінійних просторів, починаючи з найбільш очевидного:

а) сукупність усіх геометричних векторів на заданій площині;

б) сукупність усіх геометричних векторів у тривимірному просторі;

в) сукупність матриць одного (заданого) порядку $m\times n$;

г) сукупність всіх функцій $f_j(x)$, означених і неперервних на відрізку $a\le x\le b$;

д) сукупність поліномів степенів від одиниці до n.

Сукупності елементів, які не є лінійними просторами:

а) сукупність геометричних векторів одиничної довжини не є простором, тому що результатом додавання елементів та множення їх на число може бути вектор неодиничної довжини;

б) сукупність матриць різних порядків не є простором, тому що матриці різних порядків не можна додати одну до одної;

в) сукупність функцій $f_j(x)\equiv|\varphi_j(x)|$, означених і неперервних на відрізку $a\le x\le b$, не є простором, тому що в ній немає протилежних один до одного елементів;

г) сукупність поліномів одного (заданого) порядку n не є простором, тому що до цієї сукупності входить поліном з доданком x^n та поліном з доданком $-x^n$, і сума цих поліномів є поліномом порядку, меншого за n.

3.1.2. Підпростори лінійного простору

Припустимо, що з усіх елементів лінійного простору $\mathcal{L}$ за певною ознакою обрана сукупність елементів $\mathcal{L}'$, до якої входять не всі елементи простору.

Означення 3.4. Сукупність $\mathcal{L}'$ елементів простору $\mathcal{L}$ називають підпростором простору $\mathcal{L}$, якщо ця сукупність є простором відносно означених для простору $\mathcal{L}$ операцій додавання елементів і множення їх на число.

З означення лінійного простору випливає, що для того, щоб сукупність $\mathcal{L}'$ була простором, треба, щоб означені для простору $\mathcal{L}$ операції додавання елементів і множення їх на число, по-перше, не виводили елементи за межі сукупності $\mathcal{L}'$ і, по-друге, задовольняли аксіоми $1^0 - 8^0$. Однак виявляється, що друга з цих вимог завжди випливає з першої завдяки теоремі.

Теорема 3.1. Для того, щоб сукупність $\mathcal{L}'$ елементів простору $\mathcal{L}$ була його підпростором, необхідно й достатньо, щоб означені для простору $\mathcal{L}$ операції не виводили елементи сукупності $\mathcal{L}'$ за її межі[1].

Доведення. Насамперед зауважимо, що всі елементи, що належать $\mathcal{L}'$, є елементами простору $\mathcal{L}$, тому вони задовольняють аксіоми $1^0 - 5^0$ та 8^0. Аксіома 6^0 про наявність у $\mathcal{L}'$ нульового елемента виконана, тому що добуток числа нуль на будь-який елемент сукупності дорівнює нульовому елементу (див. наслідок 3.1) і належить до $\mathcal{L}'$ за умови теореми. Аксіома 7^0 про наявність у $\mathcal{L}'$ протилежних елементів виконана завдяки тому, що добутки числа -1 на всі елементи сукупності дорівнюють протилежним до них елементам (див. наслідок 3.2) і належать до $\mathcal{L}'$ за умови теореми.

Теорема 3.1 спрощує перевірку того, чи є якась сукупність елементів простору його підпростором.

Наведемо приклади підпросторів деяких лінійних просторів:

а) сукупність усіх геометричних векторів на площині є підпростором сукупності всіх геометричних векторів у тривимірному просторі;

б) сукупність поліномів степенів від одиниці до n є підпростором простору функцій, означених та неперервних на всій числовій осі;

---

[1] Тобто, щоб сума будь-яких елементів із сукупності $\mathcal{L}'$ і їхні добутки на будь-які числа належали до сукупності $\mathcal{L}'$.

в) дійсний простір матриць одного (заданого) порядку $m \times n$ є підпростором комплексного простору матриць такого ж порядку.

Сукупності елементів, які не є лінійними просторами:

а) сукупність векторів одиничної довжини не є підпростором простору геометричних векторів;

б) сукупність поліномів степеня n не є підпростором простору функцій, означених та неперервних на всій числовій осі.

3.2. Лінійно незалежні елементи простору. Метод координат у лінійній алгебрі

3.2.1. Умова лінійної незалежності векторів. Вимірність лінійного простору

Означення 3.5. Лінійною комбінацією $\boldsymbol{l}$ елементів $\mathbf{x}_i$, що належать лінійному простору $\mathcal{L}$, називають суму цих елементів, домножених на числові коефіцієнти α_i:

$$\boldsymbol{l} \overset{def}{\equiv} \sum_{i=1}^{p} \alpha_i \mathbf{x}_i \,. \tag{3.3}$$

Означення 3.6. Лінійну комбінацію елементів простору називають тривіальною, якщо всі коефіцієнти α_i дорівнюють нулю і нетривіальною, якщо хоч один з них нулю не дорівнює.

Означення 3.7. Вектори $\{\mathbf{x}_i\}_{i=\overline{1,p}}$ називають лінійно незалежними, якщо з них неможливо утворити нетривіальну лінійну комбінацію рівну нульовому вектору, а якщо така лінійна комбінація існує, то вектори називають лінійно залежними.

З означення 3.7 випливає ***умова лінійної залежності векторів***, яка полягає в тому, що існує така добірка коефіцієнтів α_i, для якої справджується рівність

$$\sum_{i=1}^{p} \alpha_i \mathbf{x}_i = \mathbf{0} \,, \tag{3.4}$$

де *хоча б один з коефіцієнтів не дорівнює нулю.*

Означення 3.8. Якщо в лінійному просторі $\mathcal{L}$ існує n лінійно незалежних елементів, а будь-які елементи, кількість яких перевищує n, лінійно залежні, то ціле число n називають вимірністю цього простору, а простір називають n-вимірним.

Зручно запам'ятати твердження, еквівалентне означенню 3.8: *вимірність лінійного простору дорівнює максимальній кількості лінійно незалежних елементів у ньому*. Це твердження та зауваження 1.7 показують, що означення вимірності лінійного простору є поширенням поняття вимірності простору геометричних векторів на простори, утворені іншими елементами.

Теорема 3.2. Якщо до якоїсь сукупності елементів простору входить нульовий елемент, елементи цієї сукупності є лінійно залежними.

Доведення. З елементів сукупності, яка містить у собі нульовий елемент, можна утворити нетривіальну лінійну комбінацію, що дорівнює нулю: $\boldsymbol{l} = 1 \cdot \mathbf{0} + 0 \cdot \mathbf{x}_1 + 0 \cdot \mathbf{x}_2 + \ldots + 0 \cdot \mathbf{x}_{n-1} + 0 \cdot \mathbf{x}_n$. Це доводить твердження теореми (див. означення 3.7).

3.2.2. Базис лінійного простору. Координати вектора

Згідно із зауваженням 1.8 у підрозд. 1.5 поняття базису в просторі геометричних векторів можна поширити на інші простори.

Означення 3.9. Сукупність лінійно незалежних елементів (векторів), кількість яких дорівнює вимірності простору, називають базисом n-вимірного лінійного простору.

Зауваження 3.1. Оскільки вимірність лінійного простору дорівнює максимальній кількості лінійно незалежних векторів у ньому, то будь-який елемент (вектор) $\mathbf{x}$ цього простору може бути представлений у формі лінійної комбінації базисних векторів $\mathbf{e}_i$:

$$\mathbf{x} = \sum_{i=1}^{n} \xi_i \mathbf{e}_i \,, \tag{3.5}$$

яку називають *розкладом вектора по векторах базису*.

Означення 3.10. Числові коефіцієнти ξ_i розкладу по базисних векторах називають координатами вектора $\mathbf{x}$ у базисі $\{\mathbf{e}_i\}_{i=\overline{1,n}}$.

В основі методу координат лежить теорема, яка уможливлює кількісні розрахунки з елементами різних лінійних просторів.

Теорема 3.3. У заданому базисі $\{\mathbf{e}_i\}_{i=\overline{1,n}}$ кожному вектору $\mathbf{x}$ лінійного простору $\mathcal{L}$ відповідає одна-єдина добірка координат $\{\xi_i\}_{i=\overline{1,n}}$ [1].

Доведення. Припустимо, що існують два різних розклади вектора в одному й тому ж базисі

$$\mathbf{x} = \sum_{i=1}^{n} \xi_i \mathbf{e}_i , \tag{3.6}$$

$$\mathbf{x} = \sum_{i=1}^{n} \eta_i \mathbf{e}_i , \tag{3.7}$$

і доведемо, що з аксіом лінійного простору та їхніх наслідків випливає, що $\xi_i = \eta_i$ для всіх i від одиниці до n. Для цього domножимо праву й ліву частини рівняння (3.7) на число -1, зваживши на аксіоми 3^0 та 5^0:

$$-1 \cdot \mathbf{x} = \sum_{i=1}^{n} \eta_i (-1 \cdot \mathbf{e}_i) . \tag{3.8}$$

Перетворимо рівність (3.8) згідно з наслідком 3.2 аксіом 1^0 – 8^0:

$$-\mathbf{x} = \sum_{i=1}^{n} \eta_i (-\mathbf{e}_i) . \tag{3.9}$$

Додамо одна до одної праві та ліві частини рівнянь (3.6) та (3.9):

$$\mathbf{x} + (-\mathbf{x}) = \sum_{i=1}^{n} [\xi_i \mathbf{e}_i + \eta_i (-\mathbf{e}_i)] .$$

Зваживши на аксіоми 4^0 та 7^0, одержимо рівняння

$$\mathbf{0} = \sum_{i=1}^{n} (\xi_i - \eta_i) \mathbf{e}_i .$$

[1] Зручно вважати, що кількість координат завжди дорівнює вимірності простору, але частина координат деяких векторів дорівнює нулю.

Оскільки базисні вектори лінійно незалежні, то нульовому вектору дорівнює лише тривіальна лінійна комбінація цих векторів, тобто всі коефіцієнти $\xi_i - \eta_i$ дорівнюють нулю.

Зауваження 3.2. З наслідку 3.1 легко зробити висновок, що вектор, усі координати якого дорівнюють нулю, є нульовим елементом простору.

Зауваження 3.3. Цілком очевидно, що в лінійному просторі може існувати безліч базисів, і у різних базисах один і той самий вектор може мати різні координати.

3.2.3. Лінійна оболонка сукупності векторів

Оберемо з лінійного простору $\mathcal{L}$ вектори $\{\mathbf{x}_i\}_{i=\overline{1,p}}$; назвемо їх *обраними векторами.* Згідно з означенням лінійного простору будь-яка лінійна комбінація обраних векторів

$$\boldsymbol{l} = \sum_{i=1}^{p} \alpha_i \mathbf{x}_i \tag{3.10}$$

належить до простору $\mathcal{L}$. Розглянемо сукупність Λ усіх лінійних комбінацій, складених з обраних векторів.

Означення 3.11. Сукупність усіх лінійних комбінацій обраних векторів називають лінійною оболонкою цих векторів.

Сумою будь-яких векторів лінійної оболонки Λ та добутком будь-якого вектора лінійної оболонки на число є якась лінійна комбінація обраних векторів, тобто вектор, що належить Λ. Згідно з теоремою 3.1 *лінійна оболонка* Λ *є підпростором простору* $\mathcal{L}$. Будь-який вектор $\boldsymbol{l}$ лінійної оболонки Λ задовольняє рівняння (3.10).

Розглянемо такі *обрані вектори* $\{\mathbf{x}_i\}_{i=\overline{1,p}}$, серед яких є r лінійно незалежних, а більшої кількості лінійно незалежних векторів серед них не існує.

Теорема 3.4. Вимірність лінійної оболонки обраних векторів дорівнює максимальній кількості *лінійно незалежних обраних векторів* r.

Доведення. Позначимо r лінійно незалежних обраних векторів як $\{\mathbf{x}_i\}_{i=\overline{1,r}}$, а обрані вектори, які є лінійними комбінаціями лінійно незалежних, – як $\{\mathbf{x}_k\}_{k=\overline{r+1,p}}$. Тоді вираз (3.10) набуде форми

$$\boldsymbol{l} = \sum_{i=1}^{r} \alpha_i \mathbf{x}_i + \sum_{k=r+1}^{p} \alpha_k \mathbf{x}_k \,. \tag{3.11}$$

Лінійно залежні вектори є лінійними комбінаціями лінійно незалежних

$$\mathbf{x}_k = \sum_{i=1}^{r} \beta_{ki} \mathbf{x}_i \,,$$

де β_{ki} – числові коефіцієнти. Підставимо ці лінійні комбінації до правої частини рівняння (3.11):

$$\boldsymbol{l} = \sum_{i=1}^{r} \alpha_i \mathbf{x}_i + \sum_{k=r+1}^{p} \alpha_k \sum_{i=1}^{r} \beta_{ki} \mathbf{x}_i \,. \tag{3.12}$$

Рівняння (3.12) тотожне рівнянню

$$\boldsymbol{l} = \sum_{i=1}^{r} \alpha_i \mathbf{x}_i + \sum_{i=1}^{r} \left(\sum_{k=r+1}^{p} \alpha_k \beta_{ki} \right) \mathbf{x}_i \,.$$

Зібравши в останньому рівнянні подібні доданки, доходимо висновку, що будь-який вектор лінійної оболонки є лінійною комбінацією лінійно незалежних векторів обраної сукупності

$$\boldsymbol{l} = \sum_{i=1}^{r} \left(\alpha_i + \sum_{k=r+1}^{p} \alpha_k \beta_{ki} \right) \mathbf{x}_i \,,$$

де числа $\alpha_i + \sum\limits_{k=r+1}^{p} \alpha_k \beta_{ki}$ є коефіцієнтами лінійної комбінації. Звідси випливає, що включення до сукупності лінійно незалежних векторів $\{\mathbf{x}_i\}_{i=\overline{1,r}}$ будь-якого вектора $\boldsymbol{l}$, що належить лінійній оболонці, утворює сукупність лінійно залежних векторів $\left\{\boldsymbol{l}, \{\mathbf{x}_i\}_{i=\overline{1,r}}\right\}$, а отже, кількість лінійно незалежних векторів у лінійній оболонці обраних векторів $\{\mathbf{x}_i\}_{i=\overline{1,p}}$ не перевищує r. Згідно з означенням 3.8 вимірність цієї лінійної оболонки дорівнює r.

Для розв'язання деяких практично важливих задач можуть бути корисними два твердження.

Твердження 3.1. Кожен лінійний простір є лінійною оболонкою свого базису.

Щоб упевнитися у справедливості цього твердження достатньо взяти до уваги, що кожен вектор лінійного простору можна розкласти по векторах базису, тобто представити у вигляді лінійної комбінації цих векторів.

Твердження 3.2. Кожен n-вимірний лінійний простір $\mathcal{L}$ має підпростори всіх вимірностей, від одиниці до n.

Щоб впевнитися у справедливості цього твердження, слід спочатку обрати базисний вектор $\mathbf{e}_1$ і зважити на те, що його лінійна оболонка Λ_1 є одновимірним підпростором простору $\mathcal{L}$. Потім обрати базисні вектори $\mathbf{e}_1, \mathbf{e}_2$, зважити на те, що вони лінійно незалежні, і дійти висновку, що лінійна оболонка цих векторів Λ_2 є двовимірним підпростором простору $\mathcal{L}$. Послідовно збільшуючи кількість обраних базисних векторів на одиницю, розглянути лінійні оболонки Λ_3, Λ_4, … , Λ_{n-1} та взяти до уваги, що вони є підпросторами простору $\mathcal{L}$.

3.3. Скалярний добуток векторів лінійного простору

Як відомо, числовою функцією векторного аргументу $\mathbf{x}$ називають правило, за яким кожному геометричному вектору $\mathbf{x}$ ставиться у відповідність певне число. З цього означення можна зробити висновок про те, що скалярний добуток геометричних векторів є функцією двох векторних аргументів, однак не кожна функція є скалярним добутком. Щоб функція двох векторних аргументів була скалярним добутком, вона має задовольняти означення 1.11, з якого випливають властивості 1) – 4), названі у підрозд. 1.7 "визначальними властивостями" скалярного добутку. Ця назва пояснюється тим, що у лінійній алгебрі властивості 1) – 4) узагальнені і включені до означення скалярного добутку.

Означення 3.12. Скалярним добутком векторів дійсного лінійного простору $\mathcal{L}$ називають таку числову функцію $F(\mathbf{x},\mathbf{y})$ двох векторів, яка має властивості, аналогічні властивостям скалярного добутку геометричних векторів, а саме:

1) $F(\mathbf{x},\mathbf{y}) = F(\mathbf{y},\mathbf{x})$;
2) $F(\mathbf{x}+\mathbf{y},\mathbf{z}) = F(\mathbf{x},\mathbf{z}) + F(\mathbf{y},\mathbf{z})$;
3) $F(\lambda\mathbf{x},\mathbf{y}) = \lambda F(\mathbf{x},\mathbf{y})$;
4) $F(\mathbf{x},\mathbf{x}) \geq 0$, $F(\mathbf{x},\mathbf{x}) = 0$ тоді і лише тоді, коли $\mathbf{x} = \mathbf{0}$, де $\mathbf{x},\mathbf{y}$ та $\mathbf{z}$ – будь-які вектори простору $\mathcal{L}$, λ – дійсне число.

Оскільки функції в різних випадках позначають різними літерами, то літера F у позначенні скалярного добутку не інформативна, тому застосовують скорочені позначення

$F(\mathbf{x},\mathbf{y}) \equiv (\mathbf{x},\mathbf{y})$, $F(\mathbf{x},\mathbf{y}) \equiv \mathbf{x}\cdot\mathbf{y}$, $F(\mathbf{x},\mathbf{y}) \equiv \mathbf{xy}$, $F(\mathbf{x},\mathbf{y}) \equiv \langle x \mid y \rangle$.

У подальшому будемо використовувати одне з двох позначень:

$$F(\mathbf{x},\mathbf{y}) \equiv \mathbf{xy}, \text{ або } F(\mathbf{x},\mathbf{y}) \equiv \langle x \mid y \rangle.$$

Якщо дотримуватись першого з цих позначень, властивості скалярного добутку елементів *дійсного* лінійного простору набудуть форми:

1) $\mathbf{xy} = \mathbf{yx}$;
2) $(\mathbf{x}+\mathbf{y})\mathbf{z} = \mathbf{xz} + \mathbf{yz}$;
3) $(\lambda\mathbf{x})\mathbf{y} = \lambda(\mathbf{xy})$;
4) $\mathbf{xx} \geq 0$, $\mathbf{xx} = 0$ тоді і лише тоді, коли $\mathbf{x} = \mathbf{0}$, де $\mathbf{x},\mathbf{y}$ та $\mathbf{z}$ – будь-які вектори простору $\mathcal{L}$, λ – дійсне число.

Зауваження 3.4. Означення скалярного добутку векторів дійсного простору поширюється на комплексні простори лише з однією, але принципово важливою зміною: властивість 1) змінюється на властивість

$$1_c)\ \mathbf{xy} = (\mathbf{yx})^*,$$

де зіркою позначена операція комплексного спряження. Наслідком цієї властивості та властивості 3) є рівність

$$\mathbf{x}(\lambda\mathbf{y}) = \lambda^*(\mathbf{xy}), \tag{3.13}$$

про яку не можна забувати, виконуючи обчислення. Необхідність заміни властивості 1) на 1_c) випливає з властивості 4), згідно з якою скалярний добуток $\mathbf{xx}$ має бути дійсним числом, тому що комплексне число не можна порівняти з нулем. (Якщо $\mathbf{y} = \mathbf{x}$, то властивість 1_c) дає: $\mathbf{xx} = (\mathbf{xx})^*$ – дійсне число). Принципова важливість умови 4), у свою чергу, пояснюється тим, що в алгебрі геометричних векторів скалярний добуток вектора самого на себе дорівнює квадрату його модуля, а отже, не може бути від'ємним.

Твердження 3.3. Скалярний добуток будь-якого вектора на нульовий вектор дорівнює нулю.

Для геометричних векторів це твердження випливає з означення скалярного добутку (див. (1.10)), а можливість поширити його на будь-який простір Евкліда є наслідком аксіом $1^0 - 8^0$ та наслідку (3.2):

$$\mathbf{x0} \overset{7^0}{=} \mathbf{x}[\mathbf{x} + (-\mathbf{x})] \overset{8^0,(3.2)}{=} \mathbf{x}[1 \cdot \mathbf{x} + (-1) \cdot \mathbf{x}] \overset{4^0}{=} (1-1) \cdot \mathbf{xx} = 0 \cdot \mathbf{xx} = 0 .$$

Означення 3.13. Простір, для елементів якого означено правило обчислення скалярного добутку, називають простором Евкліда[1].

Означення 3.14. Модулем вектора, що належить простору Евкліда, називають квадратний корінь зі скалярного добутку вектора $\mathbf{x}$ самого на себе:

$$|\mathbf{x}| \overset{def}{\equiv} \sqrt{\mathbf{xx}} . \qquad (3.14)$$

Вираз (3.14) є узагальненням виразу (1.19), справедливого для довжини геометричного вектора, на випадок елементу будь-якого іншого лінійного простору. А поняття кута між векторами вдалося узагальнити лише для дійсних лінійних просторів. Для геометричних векторів справедливе рівняння (1.20). Означенням косинуса кута між векторами $\mathbf{x}$ та $\mathbf{y}$, що належать *дійсному* лінійному простору Евкліда, є подібний до цього рівняння вираз

$$\cos(\mathbf{x} \,\hat{}\, \mathbf{y}) \overset{def}{\equiv} \frac{\mathbf{xy}}{|\mathbf{x}||\mathbf{y}|} . \qquad (3.15)$$

Поширити цей вираз на елементи комплексного простору неможливо, оскільки скалярний добуток $\mathbf{xy}$ може бути комплексним

[1] У багатьох джерелах інформації простором Евкліда називають лише дійсний простір, для елементів якого означений скалярний добуток, а комплексний простір, для якого означений скалярний добуток, називають унітарним.

числом, а косинус кута між векторами не може. Більше того, щоб означення (3.15) мало сенс, для будь-яких елементів дійсного простору Евкліда мають справджуватися нерівності

$$-1 \le \frac{\mathbf{xy}}{|\mathbf{x}||\mathbf{y}|} \le 1. \tag{3.16}$$

Доведемо, що властивість 4) скалярного добутку векторів гарантує виконання цих нерівностей.

Доведення. Скалярні добутки $\mathbf{yy}$ та $\mathbf{xy}$ є дійсними числами, тому з будь-яких елементів $\mathbf{x}$ та $\mathbf{y}$ дійсного простору Евкліда можна утворити лінійну комбінацію

$$\mathbf{z} = (\mathbf{yy})\mathbf{x} - (\mathbf{xy})\mathbf{y}, \tag{3.17}$$

яка є елементом того ж простору. Властивість 4) гарантує виконання нерівності $\mathbf{zz} \ge 0$. Підставивши до лівої частини цієї нерівності вираз (3.17), отримуємо такі рівності:

$$\mathbf{zz} = (\mathbf{yy})^2(\mathbf{xx}) - 2(\mathbf{yy})(\mathbf{xy})^2 + (\mathbf{xy})^2(\mathbf{yy}) =$$

$$= (\mathbf{yy})[(\mathbf{yy})(\mathbf{xx}) - (\mathbf{xy})^2] \ge 0.$$

Косинус кута між векторами можна визначити, якщо жоден з векторів не дорівнює нульовому вектору. У цьому випадку скалярний добуток $\mathbf{yy}$ є додатним числом, і тому нерівність $(\mathbf{yy})(\mathbf{xx}) - (\mathbf{xy})^2 \ge 0$, або, що те ж саме, $|\mathbf{y}|^2|\mathbf{x}|^2 - (\mathbf{xy})^2 \ge 0$; звідси з очевидністю випливають нерівності (3.16).

3.4. Ортогональні базиси в просторі Евкліда

3.4.1. Ортогональні базиси в n-вимірному просторі

Неможливість визначення кута між векторами комплексного простору Евкліда не завадила поширити на елементи цього простору критерій ортогональності векторів, який полягає в рівності нулю їхнього скалярного добутку.

Означення 3.15. Належні до дійсного або комплексного простору Евкліда елементи $\mathbf{x}$ та $\mathbf{y}$ називають ортогональними, якщо скалярний добуток цих елементів дорівнює нулю.

Розглянемо таку сукупність ненульових векторів $\{\mathbf{x}_i\}_{i=\overline{1,p}}$, у якій кожен вектор ортогональний до всіх інших. Такі вектори називають *попарно ортогональними*. Попарно ортогональні вектори задовольняють умови $\mathbf{x}_i\mathbf{x}_j = \delta_{ij}$, де δ_{ij} – символ Кронекера.

Теорема 3.5. Якщо вектори $\{\mathbf{x}_i\}_{i=\overline{1,p}}$, що не дорівнюють нулю, попарно ортогональні, то вони лінійно незалежні.

Доведення. Розглянемо лінійну комбінацію попарно ортогональних векторів, яка дорівнює нульовому вектору:

$$\sum_{i=1}^{p} \alpha_i \mathbf{x}_i = \mathbf{0}. \tag{3.18}$$

Утворимо скалярні добутки правої та лівої частин векторної рівності (3.18) з *будь-яким* вектором $\mathbf{x}_j$, що належать сукупності $\{\mathbf{x}_i\}_{i=\overline{1,p}}$:

$$\sum_{i=1}^{p} \alpha_i \mathbf{x}_i \mathbf{x}_j = \mathbf{0}\mathbf{x}_j.$$

Урахуємо рівність $\mathbf{x}_i\mathbf{x}_j = \delta_{ij}$, яка випливає з попарної ортогональності векторів, і твердження 3.3:

$$\sum_{i=1}^{p} \alpha_i \delta_{ij} = 0, \ j = \overline{1,p}.$$

Оскільки символ Кронекера дорівнює нулю (коли $i \neq j$) або одиниці (коли $i = j$), то з останньої рівності випливає, що

$$\alpha_j = 0, \ j = \overline{1,p},$$

тобто всі коефіцієнти лінійної комбінації (3.18) дорівнюють нулю[1]. Це значить, що нульовому елементу дорівнює лише тривіальна лінійна комбінація попарно ортогональних векторів, а отже, вони лінійно незалежні.

Згідно з доведеною теоремою і означенням базису сукупність n попарно ортогональних векторів є базисом n-вимірного простору.

---

[1] Лінійна комбінація (3.18) не зміниться, якщо позначити індекс її коефіцієнтів літерою j.

Означення 3.16. Ортогональним базисом простору Евкліда називають базис, кожен вектор якого ортогональний до всіх інших базисних векторів.

Вектори ортогонального базису задовольняють рівняння

$$\mathbf{e}_i\mathbf{e}_j = |\mathbf{e}_i|^2\,\delta_{ij}\,. \tag{3.19}$$

Означення 3.17. Ортогональний базис $\{\mathbf{e}_i\}_{i=\overline{1,n}}$, утворений з векторів одиничної довжини (ортів), називають ортонормованим.

Орти ортонормованого базису задовольняють умови

$$\mathbf{e}_i\mathbf{e}_j = \delta_{ij}\,, \tag{3.20}$$

врахування яких значно спрощує різноманітні розрахунки та зумовлює широке використання ортонормованих базисів у різних галузях науки та техніки. Серед іншого, значно спрощується процедура визначення координат вектора.

Твердження 3.4. Якщо базис $\{\mathbf{e}_i\}_{i=\overline{1,n}}$ ортонормований, то для визначення координати ξ_j вектора $\mathbf{x}$ достатньо обчислити скалярний добуток $\mathbf{x}\mathbf{e}_j$.

Щоб упевнитись у справедливості цього твердження, слід утворити скалярний добуток правої та лівої частин рівняння (3.6) на орт $\mathbf{e}_j$ та врахувати умови (3.20).

Важливими прикладами ортонормованих базисів є базиси у просторі матриць вимірностей $1\times n$ та $n\times 1$, які називають вектор-рядками та вектор-стовпцями, відповідно. Згідно з означенням добутку матриць добуток двох матриць порядку $1\times n$ не існує, а добуток матриці порядку $1\times n$ на транспоновану до неї матрицю порядку $n\times 1$ є числом і має всі властивості скалярного добутку (див. підрозд. 2.2 та 3.3). Ортонормований базис простору вектор-рядків складається з елементів $\langle i|$, означених як

$$\langle 1| \equiv \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \end{pmatrix},\ \langle 2| \equiv \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \end{pmatrix},\ \ldots,$$

$$\langle n-1| \equiv \begin{pmatrix} 0 & 0 & \cdots & 1 & 0 \end{pmatrix},\ \langle n| \equiv \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Транспоновані до них вектор-стовпці позначають як $|j\rangle$, а добутки рядків на стовпці – як $\langle i\,|\,j\rangle$. Обчислюючи ці добутки за правилом множення матриць, легко отримати співвідношення

$$\langle i \mid j \rangle = \delta_{ij} , \tag{3.21}$$

які доводять, що вектор-рядки $\langle i|$ є ортонормованим базисом у просторі матриць порядку $1 \times n$.

3.4.2. Ортонормований базис простору функцій

Розглянемо неперервні на відрізку $-l/2 \le x \le l/2$ функції $\psi_m(x)$ дійсної змінної x, які можуть набувати як дійсних, так і комплексних значень. Нижній індекс, який набуває додатних цілих значень, нумерує ці функції за певною ознакою. Як було зазначено, ці функції є елементами лінійного простору. Визначені інтеграли

$$\int_{-l/2}^{l/2} \psi_m(x)\psi_k^*(x)dx \equiv \langle \psi_m \mid \psi_k \rangle \tag{3.22}$$

є скалярними добутками в просторі функцій. Якщо функції $\psi_m(x)$ такі, що ці визначені інтеграли дорівнюють нулю, коли $m \neq k$, то вони попарно ортогональні.

Означення 3.18. Функції $\psi_m(x)$ є базисом простору функцій, якщо кожну іншу функцію $f(x)$, неперервну на відрізку $-l/2 \le x \le l/2$, можна представити збіжним рядом[1]:

$$f(x) = \sum_{k=1}^{\infty} C_k \psi_k(x) . \tag{3.23}$$

Згідно з означеннями 3.16 та 3.17 базис простору функцій ортогональний, якщо справджується умова

$$\langle \psi_m \mid \psi_k \rangle = \langle \psi_m \mid \psi_m \rangle \delta_{mk} , \tag{3.24}$$

та ортонормований, за умови

$$\langle \psi_m \mid \psi_k \rangle = \delta_{mk} . \tag{3.25}$$

Щоб розкласти функцію за функціями ортогонального базису, треба знайти коефіцієнти ряду (3.23). Для цього слід помножити

---

[1] За необхідності інформацію про ряди функцій легко знайти у підручнику [4].

обидві частини рівняння (3.23) на функцію $\psi_m^*(x)$ та зінтегрувати результат множення на відрізку неперервності функцій. Результатом цих дій є рівняння

$$\int_{-l/2}^{l/2} f(x)\psi_m^*(x)dx = \sum_{k=1}^{\infty} C_k \int_{-l/2}^{l/2} \psi_k(x)\psi_m^*(x)dx. \qquad (3.26)$$

Очевидним наслідком виразів (3.22) та (3.24) є співвідношення

$$\int_{-l/2}^{l/2} \psi_k(x)\psi_m^*(x)dx = \delta_{mk} \int_{-l/2}^{l/2} \psi_m(x)\psi_m^*(x)dx,$$

завдяки якому в правій частині рівняння (3.26) не дорівнює нулю лише доданок, у якому $k = m$. Зваживши на цей факт і на те, що добуток $\psi_m(x)\psi_m^*(x)$ дорівнює квадрату модуля функції $\psi_m(x)$, знаходимо формулу для обчислення коефіцієнтів нескінченного ряду (3.23):

$$C_m = \frac{\int_{-l/2}^{l/2} f(x)\psi_m^*(x)dx}{\int_{-l/2}^{l/2} |\psi_m(x)|^2\, dx}. \qquad (3.27)$$

Якщо базис ортонормований, то інтеграл, що стоїть у знаменнику, дорівнює одиниці, і вираз (3.27) перетворюється на більш простий:

$$C_m = \int_{-l/2}^{l/2} f(x)\psi_m^*(x)dx. \qquad (3.28)$$

Формули (3.23) та (3.27), або (3.23) та (3.28), дають змогу обчислити коефіцієнти C_m, якщо вдається розрахувати відповідні визначені інтеграли. Якщо ці формули використовують в ході наближених обчислень, то функцію $f(x)$ наближають декількома першими доданками ряду. Кількість доданків, якими наближають функцію, залежить від необхідної точності обчислень. Можливість відкинути всі наступні доданки випливає зі збіжності ряду (3.23).

Важливим прикладом попарно ортогональних функцій є функції

$$\psi_m(x) = \sqrt{\frac{2}{l}} \cos\left(\frac{2\pi m}{l} x\right). \qquad (3.29)$$

Покажемо, що ці функції задовольняють умови (3.25):

$$\left\langle \psi_m \mid \psi_k \right\rangle = \int\limits_{-l/2}^{l/2} \psi_m(x)\psi_k(x)dx = \frac{2}{l}\int\limits_{-l/2}^{l/2} \cos\left(\frac{2\pi m}{l}x\right)\cos\left(\frac{2\pi k}{l}x\right)dx\,;$$

$$2\pi x / l \equiv y\,;$$

$$\left\langle \psi_m \mid \psi_k \right\rangle = \frac{1}{\pi}\int\limits_{-\pi}^{\pi} \cos(my)\cos(ky)dy =$$

$$= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} \cos[(m-k)y] + \cos[(m+k)y]dy =$$

$$= \frac{1}{2\pi}\left\{\frac{\sin[(m-k)y]}{m-k} + \frac{\sin[(m+k)y]}{m+k}\right\}\Bigg|_{-\pi}^{\pi}.$$

Синус кута, кратного π, дорівнює нулю, тому $\left\langle \psi_m \mid \psi_k \right\rangle = 0$, якщо $m \neq k$, а якщо $m = k$, то можна обчислити скалярний добуток $\left\langle \psi_m \mid \psi_m \right\rangle$ у такий спосіб:

$$\left\langle \psi_m \mid \psi_m \right\rangle = \left\{\frac{y}{2\pi}\cdot \lim_{(m-k)y\to 0}\frac{\sin[(m-k)y]}{(m-k)y}\right\}\Bigg|_{-\pi}^{\pi} = \left(\frac{y}{2\pi}\cdot 1\right)\Bigg|_{-\pi}^{\pi} = 1.$$

Будь-яка нетривіальна лінійна комбінація функцій (3.29) є парною функцією, тому непарну функцію $f(x)$ не можна представити нескінченним рядом цих функцій. З цієї причини для отримання ортонормованого базису простору неперервних функцій слід *доповнити* парні функції (3.29) непарними функціями

$$\varphi_m(x) = \sqrt{\frac{2}{l}}\sin\left(\frac{2\pi m}{l}x\right). \tag{3.30}$$

У підручниках із функціонального аналізу можна знайти доведення того, що сукупність функцій $\{\psi_m(x), \varphi_m(x)\}$ утворює *повний* ортонормований базис простору неперервних функцій $f(x)$ (див., напр., [5]).

Із розглянутого прикладу слід зробити такі висновки:

а) простір неперервних функцій нескінченновимірний;

б) перш ніж користуватися якимось базисом простору функцій для проведення розрахунків, слід упевнитися в тому, що цей базис повний.

РОЗДІЛ 4
Визначники квадратних матриць

4.1. Означення та основні властивості визначника

4.1.1. Послідовності додатних цілих чисел

Означення 4.1. *Правильною послідовністю* n чисел називають послідовність додатних цілих чисел, розташованих у порядку зростання,

$$\overline{1,n} \equiv 1,2,...,n\,. \tag{4.1}$$

Означення 4.2. *Довільною послідовністю* n чисел називають послідовність додатних цілих чисел, розташованих у довільному порядку. Позначимо цю послідовність як

$$\overline{j_1, j_n} \equiv j_1, j_2, ..., j_n\,, \tag{4.2}$$

де $1 \le j_p \le n$ є додатними цілими числами, серед яких немає однакових.

Назвемо *елементарним переставленням* чисел j_p таку зміну довільної послідовності цих чисел, за якої два числа міняються місцями, а всі інші залишаються на своїх місцях. Позначимо символом t кількість таких елементарних переставлень, після яких послідовність $\overline{j_1, j_n}$ перетворюється на правильну. Очевидно, що *задану* послідовність можна перетворити на правильну шляхом різної кількості елементарних переставлень. Наприклад, послідовність 321 перетворюється на правильну або одним переставленням чисел 3 та 1, або трьома елементарними переставленнями: $321 \to 312 \to 132 \to 123$. Математики довели твердження, важливе для означення визначника квадратної матриці.

Твердження 4.1. Якщо задану послідовність $\overline{j_1, j_n}$ можна перетворити на правильну шляхом парної кількості елементарних переставлень, то це не можна зробити шляхом непарної кількості переставлень (доведення див., напр., у книзі [3]).

4.1.2. Символ Леві-Чивіти

Означення 4.3. Символом Леві-Чивіти $\varepsilon_{j_1 j_2 j_3 \cdots j_n}$ позначають числа, що дорівнюють нулю, якщо серед додатних цілих чисел $j_1, j_2, ..., j_n$ є хоча б два однакових або $(-1)^t$, якщо всі ці числа різні, тобто утворюють послідовність $\overline{j_1, j_n}$, яку можна перетворити на правильну шляхом t елементарних переставлень.

Зрозуміло, що надане означення не мало б сенсу, якби для заданої послідовності $\overline{j_1, j_n}$ число t могло б бути як парним, так і непарним. Саме в цьому полягає важливість твердження 4.1.

Наслідок 4.1. З наданого означення випливає, що елементарне переставлення двох індексів ненульового символу Леві-Чивіти спричиняє заміну його знака.

Важливим окремим випадком є символ Леві-Чивіти з трьома індексами. Цим символом позначають таку сукупність чисел:

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1,\ \varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1,$$
$$\varepsilon_{j_1 j_1 j_3} = \varepsilon_{j_1 j_2 j_2} = \varepsilon_{j_1 j_2 j_1} = \varepsilon_{j_1 j_1 j_1} = 0. \qquad (4.3)$$

Рівності (4.3) записані з урахуванням того, що переставлення першого індексу на третє місце здійснюється шляхом двох елементарних переставлень, а отже, не змінює знак символу Леві-Чивіти.

4.1.3. Означення визначника

Розглянемо спочатку матрицю 2-го порядку:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

де елементи a_{ij} є числами або функціями, що набувають числових значень.

Означення 4.4. Визначником (детермінантом) матриці 2-го порядку називають таку суму добутків її елементів:

$$\det A \overset{def}{\equiv} a_{11}a_{22} - a_{21}a_{12}. \qquad (4.4)$$

У багатьох випадках для визначника використовують позначення, яке відрізняється від позначення матриці лише формою дужок; у такому разі для визначника матриці 2-го порядку маємо

$$\det A \equiv \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}. \tag{4.5}$$

Означення 4.5. Визначником (детермінантом) матриці порядку $n > 2$ називають таку суму добутків її елементів:

$$\det A \overset{def}{\equiv} \sum_{j_1, j_2, j_3 \cdots j_n} \varepsilon_{j_1 j_2 j_3 \cdots j_n} a_{1 j_1} a_{2 j_2} a_{3 j_3} \ldots a_{n j_n} . \tag{4.6}$$

Широко використовують таке позначення:

$$\det A \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}. \tag{4.7}$$

Означення 4.6. Порядок матриці називають також порядком її визначника; рядки/стовпці матриці називають також рядками/стовпцями її визначника.

Зауваження 4.1. Згідно з означенням 4.5 визначник, елементами якого є числа, – це математичне правило, за яким кожній квадратній матриці ставиться у відповідність певне число.

Розглянемо матриці порядку $n \times 1$ та $1 \times n$, елементами яких є ті ж елементи, з яких утворені стовпці та рядки квадратної матриці A:

$$|a_i\rangle \equiv \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}, \ \langle a_i| \equiv \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}. \tag{4.8}$$

Як зазначено в розд. 3, матриці порядку $n \times 1$ є векторами n-вимірного лінійного простору, і те ж саме стосується матриць порядку $1 \times n$. Якщо якесь твердження однаковою мірою стосується як стовпців, так і рядків визначника, зручно позначати однаково і рядок, і стовпець:

$$\mathbf{a}_i \equiv |a_i\rangle, \langle a_i|. \tag{4.9}$$

Якщо відомо, з яких стовпців/рядків $\mathbf{a}_i$ складається матриця, то відомі всі її елементи, а отже, формула (4.6) дозволяє обчислити її визначник. Звідси випливає, що ***визначник*** n ***-го порядку є функцією від*** n ***векторних аргументів***:

$$\det A = D(\mathbf{a}_i)\,,\ i = \overline{1,n}\,. \qquad (4.10)$$

4.1.4. Основні властивості визначника

Наведемо основні властивості визначника:

1) визначник не змінюється при транспонуванні матриці:

$$\det A = \det A^T\,; \qquad (4.11)$$

2) при переставленні будь-яких двох рядків або стовпців визначник змінює знак:

$$D(\mathbf{a}_i, \mathbf{a}_j) = -D(\mathbf{a}_j, \mathbf{a}_i)\,,\ i \in \overline{1,n},\ j \in \overline{1,n}\,; \qquad (4.12)$$

3) визначник є лінійною функцією своїх рядків і стовпців:

$$D(\mathbf{a}_i, \lambda\mathbf{a}_j + \mu\mathbf{b}_j) = \lambda D(\mathbf{a}_i, \mathbf{a}_j) + \mu D(\mathbf{a}_i, \mathbf{b}_j)\,,\ i \in \overline{1,n},\ j \in \overline{1,n}\,. \qquad (4.13)$$

Доведемо властивість 1). Зважимо на те, що з виразів (2.9) та (4.6) випливає така рівність для визначника матриці $A^T(\tilde{a}_{ij})$:

$$\det A^T \overset{def}{\equiv} \sum_{j_1, j_2, j_3 \ldots j_n} \varepsilon_{j_1 j_2 j_3 \ldots j_n} \tilde{a}_{j_1 1} \tilde{a}_{j_2 2} \tilde{a}_{j_3 3} \ldots \tilde{a}_{j_n n}\,.$$

Розглянемо будь-який доданок суми, що стоїть у правій частині цієї рівності:

$$X \equiv \varepsilon_{j_1 j_2 j_3 \ldots j_n} \tilde{a}_{j_1 1} \tilde{a}_{j_2 2} \tilde{a}_{j_3 3} \ldots \tilde{a}_{j_n n}\,.$$

Шляхом елементарних переставлень *елементів матриці* в добутку елементів, що стоїть у цьому доданку, розташуємо елементи так, щоб їхні перші індекси утворили правильну послідовність:

$$\overline{j_1, j_n} \xrightarrow{t} \overline{1,n}\,, \qquad (4.14)$$

де t – це та кількість переставлень, яка знадобилася для цього. Після цих переставлень елементів матриці другі індекси в доданку X розташуються в якійсь послідовності $\overline{k_1, k_n}$:

$$\overline{1,n} \xrightarrow{t} \overline{k_1, k_n}\,, \qquad (4.15)$$

і тому цей доданок визначника транспонованої матриці виразиться як

$$X = \varepsilon_{j_1 j_2 j_3 \dots j_n} \tilde{a}_{1k_1} \tilde{a}_{2k_2} \tilde{a}_{3k_3} \dots \tilde{a}_{nk_n} .$$

Використавши знову вираз (2.9), одержимо

$$X = \varepsilon_{j_1 j_2 j_3 \dots j_n} a_{k_1 1} a_{k_2 2} a_{k_3 3} \dots a_{k_n n} . \tag{4.16}$$

З означення символу Леві-Чивіти і виразу (4.14) випливає, що $\varepsilon_{j_1 j_2 j_3 \dots j_n} = (-1)^t$, а вираз (4.15) показує, що $\varepsilon_{k_1 k_2 k_3 \dots k_n} = (-1)^t$, і тому справедлива рівність

$$\varepsilon_{j_1 j_2 j_3 \dots j_n} = \varepsilon_{k_1 k_2 k_3 \dots k_n} . \tag{4.17}$$

Підставивши (4.17) до (4.16), знаходимо рівняння

$$X = \varepsilon_{k_1 k_2 k_3 \dots k_n} a_{k_1 1} a_{k_2 2} a_{k_3 3} \dots a_{k_n n} . \tag{4.18}$$

Оскільки немає значення, якою літерою, k чи j, позначені *індекси* символу Леві-Чивіти та елементів матриці, то порівняння виразів (4.18) та (4.6) показує, що будь-який доданок визначника матриці A^T дорівнює одному з доданків визначника матриці A. Як у визначнику матриці A, так і у визначнику матриці A^T, немає двох доданків з однаковою комбінацією нижніх індексів, тому рівність доданків цих матриць взаємно однозначна, а отже, $\det A = \det A^T$.

Доведемо властивість 2). Із цією метою переставимо два рядки у визначнику (4.7), для визначеності – перший та другий. При такому переставленні у формулі (4.6) відбудеться лише одна зміна – нижній індекс 1 зміниться на 2, і навпаки. Визначник з переставленими рядками виразиться як

$$\det A^{1\leftrightarrow 2} = \sum_{j_1, j_2, j_3 \dots j_n} \varepsilon_{j_1 j_2 j_3 \dots j_n} a_{2 j_1} a_{1 j_2} a_{3 j_3} \dots a_{n j_n} .$$

Зі шкільного курсу алгебри відомо, що сума добутків елементів, що стоїть у цьому виразі, не зміниться, якщо у кожному добутку *переставити місцями перший та другий співмножники*, тому

$$\det A^{1\leftrightarrow 2} = \sum_{j_1, j_2, j_3 \dots j_n} \varepsilon_{j_1 j_2 j_3 \dots j_n} a_{1 j_2} a_{2 j_1} a_{3 j_3} \dots a_{n j_n} . \tag{4.19}$$

Тепер використаємо загальну властивість підсумовування за індексами: оскільки обидва індекси j_1 та j_2 набувають цілих значень від одиниці до n, то їх можна *перепозначити* $j_1 \leftrightarrow j_2$ і отримати вираз, рівносильний з (4.19):

$$\det A^{1\leftrightarrow 2} = \sum_{j_1, j_2, j_3 \cdots j_n} \varepsilon_{j_2 j_1 j_3 \dots j_n} a_{1j_1} a_{2j_2} a_{3j_3} \dots a_{nj_n}.$$

Останнє, що залишилося зробити для доведення другої властивості – переставити перший та другий індекси символу Леві-Чивіти, урахувавши, що він при цьому змінить знак:

$$\det A^{1\leftrightarrow 2} = -\sum_{j_1, j_2, j_3 \cdots j_n} \varepsilon_{j_1 j_2 j_3 \dots j_n} a_{1j_1} a_{2j_2} a_{3j_3} \dots a_{nj_n}. \qquad (4.20)$$

Порівняння виразу (4.20) з виразом (4.6) доводить, що $\det A^{1\leftrightarrow 2} = -\det A$.

Оскільки вище було доведено, що транспонування матриці не змінює її визначника, то доходимо висновку, що переставлення двох стовпців визначника також веде до зміни його знака.

Доведемо властивість 3). Для доведення цієї властивості достатньо помітити, що кожен з індексів $1, 2, 3, \dots, n$, який нумерує рядки визначника, зустрічається у формулі (4.6) лише один раз. Звідси випливає, що в кожному з добутків елементів матриці $a_{1j_1} a_{2j_2} a_{3j_3} \dots a_{nj_n}$ є лише один "представник" від кожного рядка. Тому заміна будь-якого рядка визначника, наприклад рядка $\mathbf{a}_2$, на $\lambda \mathbf{b}_2 + \mu \mathbf{c}_2$ веде до заміни

$$a_{1j_1} a_{2j_2} a_{3j_3} \dots a_{nj_n} \to a_{1j_1} (\lambda b_{2j_2} + \mu c_{2j_2}) a_{3j_3} \dots a_{nj_n}. \qquad (4.21)$$

Подальше доведення властивості 3) зводиться до підставлення правої частини виразу (4.21) до правої частини виразу (4.6) та загальновідомих правил алгебраїчних дій з добутками та сумами чисел.

4.1.5. Наслідки основних властивостей визначника

1^0. Визначник, який має однакові рядки або стовпці, дорівнює нулю.

Дійсно, поклавши $\mathbf{a}_i = \mathbf{a}_j$ у рівнянні (4.12), яке виражає властивість 2), одержуємо рівність

$$D(\mathbf{a}_i, \mathbf{a}_i) = -D(\mathbf{a}_i, \mathbf{a}_i).$$

Єдине число, яке дорівнює самому собі після обернення свого знака, – число нуль.

2^0. Якщо відповідні елементи двох рядків/стовпців визначника відрізняються в λ разів, то визначник дорівнює нулю, тобто

$$D(\mathbf{a}_i, \lambda \mathbf{a}_i) = 0 . \tag{4.22}$$

Дійсно, поклавши в рівнянні (4.13), яке виражає властивість 3), $\mu = 0$, одержуємо окремий випадок цього рівняння $D(\mathbf{a}_i, \lambda \mathbf{a}_j) = \lambda D(\mathbf{a}_i, \mathbf{a}_j)$, який показує, що

$$D(\mathbf{a}_i, \lambda \mathbf{a}_i) = \lambda D(\mathbf{a}_i, \mathbf{a}_i) . \tag{4.23}$$

Згідно з наслідком 1^0 маємо $D(\mathbf{a}_i, \mathbf{a}_i) = 0$, а з огляду на (4.23), доходимо висновку, що $D(\mathbf{a}_i, \lambda \mathbf{a}_i) = 0$.

3^0. Якщо до будь-якого рядка/стовпця визначника додати його інший рядок/стовпець, домножений на будь-яке число, то визначник не зміниться.

Дійсно, якщо до рядка/стовпця $\mathbf{a}_j$ визначника $D(\mathbf{a}_i, \mathbf{a}_j)$ додати рядок/стовпець $\mathbf{a}_i$, домножений на число μ, цей визначник перетвориться на $D(\mathbf{a}_i, \mathbf{a}_j + \mu \mathbf{a}_i)$. Поклавши в рівнянні (4.13), яке виражає властивість 3), $\lambda = 1$, $\mathbf{b}_j = \mathbf{a}_i$, отримуємо рівність

$$D(\mathbf{a}_i, \mathbf{a}_j + \mu \mathbf{a}_i) = D(\mathbf{a}_i, \mathbf{a}_j) + \mu D(\mathbf{a}_i, \mathbf{a}_i) .$$

Зваживши на наслідок 1^0, доходимо висновку, що

$$D(\mathbf{a}_i, \mathbf{a}_j + \mu \mathbf{a}_i) = D(\mathbf{a}_i, \mathbf{a}_j) .$$

4^0. Якщо хоча б один рядок/стовпець визначника цілком складається з нулів, цей визначник дорівнює нулю.

Дійсно, якщо до нульового рядка/стовпця визначника $D(\mathbf{a}_i, \mathbf{0})$ додати його рядок $\mathbf{a}_i$, то згідно з наслідком 3^0 визначник від цього не зміниться: $D(\mathbf{a}_i, \mathbf{0}) = D(\mathbf{a}_i, \mathbf{0} + \mathbf{a}_i)$. Зваживши на те, що $\mathbf{0} + \mathbf{a}_i = \mathbf{a}_i$, і враховуючи наслідок 1^0, одержуємо рівності $D(\mathbf{a}_i, \mathbf{0}) = D(\mathbf{a}_i, \mathbf{a}_i) = 0$, які доводять наявність сформульованого наслідку.

4.2. Мінор матриці та алгебраїчне доповнення елемента матриці

Розглянемо матрицю $A(a_{ij})$ порядку $m\times n$. Оберемо у цій матриці рядки з номерами $i_1 < i_2 \dots i_{k-1} < i_k$ і стовпці з номерами $j_1 < j_2 \dots j_{k-1} < j_k$. Утворимо з елементів, що стоять на перехресті цих рядків та стовпців, квадратну матрицю порядку k:

$$S \equiv \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_n j_1} & a_{i_n j_2} & \cdots & a_{i_n j_k} \end{pmatrix}. \qquad (4.24)$$

Матрицю S називають субматрицею матриці A.

Означення 4.7. Визначник квадратної субматриці матриці A називають мінором матриці A[1].

Порядок субматриці S називають порядком мінору матриці A. Матриця порядку $m\times n$ має мінори порядків від $k=2$ до $k=m-1$ (якщо $m<n$) або $k=n-1$ (якщо $n<m$). Надалі використовуватимемо таке позначення:

$$\det S = M^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}. \qquad (4.25)$$

● **Приклад 4.1.** Розглянемо матрицю

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & -2 & -3 & -4 \\ 5 & 6 & 7 & 8 \\ -5 & -6 & -7 & -8 \end{pmatrix}. \qquad (4.26)$$

Обчислимо один з її мінорів:

$$M^{14}_{23} = \begin{vmatrix} 2 & 3 \\ -6 & -7 \end{vmatrix} = 2\cdot(-7) - 3\cdot(-6) = 4,$$

(див. формулу (4.4)).●

---

[1] Надалі будемо називати мінор квадратної матриці також мінором її визначника.

Розглянемо тепер *квадратну* матрицю порядку n. Якщо видалити з цієї матриці k рядків та k стовпців, де $k < n-1$, то ті елементи, що залишаться, утворять квадратну матрицю $\bar{S}$.

Означення 4.8. Визначник матриці $\bar{S}$ називають *доповняльним мінором* до мінору $M_{j_1 j_2 \cdots j_k}^{i_1 i_2 \dots i_k}$:

$$\det \bar{S} = \bar{M}_{j_1 j_2 \cdots j_k}^{i_1 i_2 \dots i_k} . \tag{4.27}$$

Зауваження 4.2. У позначенні доповняльного мінору вказують ті рядки та стовпці матриці A, які були видалені, а не ті, до яких належать елементи доповняльного мінору.

● **Приклад 4.2.** Доповняльним до мінору (4.27) є мінор

$$\bar{M}_{23}^{14} = \begin{vmatrix} -1 & -4 \\ 5 & 8 \end{vmatrix} = -8 + 20 = 12 . \bullet$$

Якщо видалити з матриці рядок і стовпчик, до яких належить елемент a_{ij}, залишиться матриця $\bar{S}$ порядку $n-1$. Її визначник $\bar{M}_j^i$ є доповняльним мінором елемента a_{ij}.

Означення 4.9. Алгебраїчним доповненням елемента a_{ij} квадратної матриці A називають визначник, пов'язаний з доповняльним мінором $\bar{M}_j^i$ формулою

$$A_j^i \overset{def}{\equiv} (-1)^{i+j} \bar{M}_j^i . \tag{4.28}$$

● **Приклад 4.3.** Алгебраїчним доповненням елемента a_{11} матриці (4.26) є визначник

$$A_1^1 = \begin{vmatrix} -2 & -3 & -4 \\ 6 & 7 & 8 \\ -6 & -7 & -8 \end{vmatrix},$$

а алгебраїчним доповненням елемента a_{12} – визначник

$$A_2^1 = -\begin{vmatrix} -1 & -3 & -4 \\ 5 & 7 & 8 \\ -5 & -7 & -8 \end{vmatrix} . \bullet$$

Поняття алгебраїчного доповнення елемента матриці дозволяє вивести зручну аналітичну формулу для обчислення визначників порядку $n > 2$, яку називають розкладом визначника за елементами рядка або стовпця.

4.3. Розклад визначника за елементами рядка або стовпця. Формули Лапласа

Розглянемо визначник квадратної матриці порядку n:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{4.29}$$

Як було зазначено, рядки та стовпці визначника є векторами лінійного простору, а визначник – лінійною функцією цих векторів (див. основну властивість 3)). Представимо останній вектор-рядок визначника (4.29) у формі такої лінійної комбінації:

$$\mathbf{a}_n = \begin{pmatrix} a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = a_{n1}\begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} +$$
$$+ a_{n2}\begin{pmatrix} 0 & 1 & \ldots & 0 \end{pmatrix} + \ldots + a_{nn}\begin{pmatrix} 0 & 0 & \ldots & 1 \end{pmatrix}$$

Застосувавши до визначника (4.29) формулу (4.13), представимо його у формі суми визначників:

$$\det A = a_{n1}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{vmatrix} + a_{n2}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 1 & \cdots & 0 \end{vmatrix} + \ldots$$
$$+ a_{nn}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \equiv a_{n1}\Delta_1 + a_{n2}\Delta_2 + \ldots + a_{nn}\Delta_n. \tag{4.30}$$

Доведемо, що останній визначник дорівнює доповняльному мінору елемента a_{nn}:[1]

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = \sum_{j_1, j_2 \cdots j_{n-1}, j_n} \varepsilon_{j_1 j_2 \cdots j_{n-1} j_n} a_{1 j_1} a_{2 j_2} \ldots a_{n-1\, j_{n-1}} a_{n j_n} =$$

$$= \sum_{j_1, j_2 \cdots j_{n-1}} \varepsilon_{j_1 j_2 \cdots j_{n-1} n} a_{1 j_1} a_{2 j_2} \ldots a_{n-1\, j_{n-1}} \cdot 1.$$

При написанні останньої із цих рівностей було враховано, що індекс j_n є номером елемента з останнього рядка визначника, *єдиний* ненульовий елемент цього рядка стоїть на останньому місті і дорівнює одиниці. З огляду на це, у сумі за індексом j_n є лише один ненульовий доданок, пропорційний символу $\varepsilon_{j_1 j_2 \cdots j_{n-1} n}$. Далі слід зважити, що, якщо певною кількістю переставлень якась послідовність індексів $\overline{j_1, j_{n-1}}$ перетворюється на правильну послідовність $\overline{1, n-1}$, то послідовність індексів символу $\varepsilon_{j_1 j_2 \cdots j_{n-1} n}$ обов'язково стає правильною послідовністю $\overline{1, n}$, і тому справджується рівність

$$\varepsilon_{j_1 j_2 \cdots j_{n-1} n} = \varepsilon_{j_1 j_2 \cdots j_{n-1}}.$$

Підставивши цю рівність в останній з наведених вище виразів для визначника Δ_n, знаходимо формулу

$$\Delta_n = \sum_{j_1, j_2 \cdots j_{n-1}} \varepsilon_{j_1 j_2 \cdots j_{n-1} n} a_{1 j_1} a_{2 j_2} \ldots a_{n-1\, j_{n-1}},$$

яка є означенням визначника порядку $n-1$, утвореного з тих елементів визначника (4.29), які розташовані в рядках і стовпцях з номерами від одиниці, до $n-1$. Отже, визначник Δ_n дорівнює доповняльному мінору $\bar{M}_n^n$. За означенням алгебраїчного допов-

[1] Автори деяких підручників це твердження не доводять, але воно не очевидне з тієї причини, що порядок визначника Δ_n перевищує на одиницю порядок доповняльного мінору.

нення елемента a_{nn} маємо $A_n^n = (-1)^{n+n}\bar{M}_n^n = \bar{M}_n^n$. З цього випливає рівність, яку треба було довести:

$$\Delta_n = \bar{M}_n^n .$$

Розглянемо тепер визначник Δ_k, у якому одиниця, що розташована в останньому рядку, належить до k-го стовпця. Зробивши $n-k$ переставлень цього стовпця з сусідніми, можна розташувати одиницю на останньому місці останнього рядку, повторити все те, що було зроблено для визначника Δ_n, і в такий спосіб отримати рівність $\Delta_k = (-1)^{n-k}\bar{M}_k^n$. Домноживши праву частину цієї рівності на $(-1)^{2k}$, одержимо вираз $\Delta_k = (-1)^{n+k}\bar{M}_k^n$ і врахуємо означення алгебраїчного доповнення (4.28). У результаті одержимо рівність

$$\Delta_k = A_k^n . \tag{4.31}$$

Записавши тотожну рівність з виразу (4.30) у формі

$$\det A \equiv \sum_{k=1}^{n} a_{nk}\Delta_k ,$$

і врахувавши рівність (4.31), одержуємо формулу розкладу визначника за елементами останнього рядка:

$$\det A = \sum_{k=1}^{n} a_{nk} A_k^n .$$

Неважко впевнитись, що визначник можна розкласти за елементами будь-якого рядка, скориставшись формулою

$$\det A = \sum_{k=1}^{n} a_{ik} A_k^i , \quad i = \overline{1,n} . \tag{4.32}$$

Оскільки транспонування матриці не змінює її визначника, то його можна розкласти і за елементами будь-якого стовпця згідно з формулою

$$\det A = \sum_{i=1}^{n} a_{ik} A_k^i , \quad k = \overline{1,n} . \tag{4.33}$$

Формули (4.32), (4.33) дозволяють перетворити визначник n-го порядку на суму визначників порядку $n-1$, кожен з яких можна перетворити на суму визначників порядку $n-2$ і так далі, доки визначник не перетвориться на суму визначників 2-го порядку, які можна буде обчислити за формулою (4.4).

Насамкінець зауважимо, що формули (4.32) і (4.33) дозволяють обчислити будь-який визначник, тому кожну з них можна вважати означенням визначника [1]. Узагальненням формул (4.32), (4.33) є формули Лапласа

$$\sum_{i=1}^{n} a_{ij} A_k^i = \delta_{kj} \det A, \tag{4.34}$$

$$\sum_{k=1}^{n} a_{ik} A_k^j = \delta_{ij} \det A, \tag{4.35}$$

де δ_{kj} та δ_{ij} – символи Кронекера. Якщо у формулі Лапласа (4.34) покласти $j = k$, то вона перетвориться на розклад визначника за елементами стовпця (див. (4.33)). Якщо ж у формулі Лапласа (4.35) покласти $j = i$, то вона стане розкладом визначника за елементами рядка (див. (4.32)). Суть формул Лапласа частково формулюють у формі *теореми Лапласа*: сума добутків елементів i-го рядка/стовпця визначника на алгебраїчні доповнення відповідних елементів іншого рядка/стовпця дорівнює нулю. Доведемо цю теорему. Для цього розглянемо розклад визначника за елементами першого стовпця:

$$\sum_{i=1}^{n} a_{i1} A_1^i = \det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{4.36}$$

Далі розглянемо ліву частину рівняння (4.34), поклавши в ній $j = 2$, $k = 1$: $\sum\limits_{j=1}^{n} a_{i2} A_1^i$. Вона відрізняється від лівої частини рівняння (4.36) *лише* заміною елементів першого стовпця на елементи другого стовпця визначника, і тому дорівнює визначнику, у якому перший та другий стовпці однакові:

$$\sum_{i=1}^{n} a_{i2} A_1^i = \begin{vmatrix} a_{12} & a_{12} & \cdots & a_{1n} \\ a_{22} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Згідно з наслідком 1^0 основних властивостей визначника визначник з двома однаковими стовпцями дорівнює нулю, що і треба було довести. Оскільки визначник матриці A можна розкласти по елементах будь-якого стовпця, то доведення теореми не зміниться, якщо розглянути формулу (4.34) з будь-якими індексами $j \neq k$. Для доведення формули (4.35) слід розглянути розклад визначника за елементами певного рядка, а потім замінити елементи цього рядка на елементи іншого.

4.4. Обернена матриця

4.4.1. Означення оберненої матриці

Означення 4.10. Матрицею, що обернена матриці A, називають матрицю A^{-1}, добуток якої на матрицю A дорівнює одиничній матриці, тобто

$$AA^{-1} = A^{-1}A = E . \tag{4.37}$$

Першу з рівностей (4.37) формулюють у формі твердження, яке має ознаки сленгу, але зручне для запам'ятовування.

Твердження 4.2. Права обернена матриця дорівнює лівій.

Щоб довести справедливість цього твердження слід показати, що з рівності

$$AA^{-1} = E \tag{4.38}$$

випливає рівність

$$A^{-1}A = E .$$

Зробимо це, домноживши обидві частини рівності (4.38) на матрицю A з правого боку

$$(AA^{-1})A = EA .$$

Візьмемо до уваги асоціативність добутку матриць і пункт б) зауваження 2.2:

$$A(A^{-1}A) = AE \Rightarrow (A^{-1}A) = E .$$

4.4.2. Розрахунок елементів оберненої матриці

Позначимо елемент оберненої матриці, що належить i-му рядку та j-му стовпцю, як $\left\langle A^{-1} \right\rangle_{ij}$. Доведемо, що елементи матриці, що обернена матриці $A(a_{ij})$, можна визначати за формулою

$$\left\langle A^{-1} \right\rangle_{ij} = \frac{A_i^j}{\det A}, \tag{4.39}$$

де A_i^j є алгебраїчним доповненням елемента a_{ji} матриці A.

Доведення. Для доведення формули (4.39) достатньо довести, що, якщо обчислити елементи матриці A^{-1} за цією формулою, то матриця A^{-1} задовольнить рівняння $AA^{-1} = E$, яке є формальним означенням оберненої матриці. Щоб зробити це, зауважимо, що згідно з означенням добутку матриць

$$\left\langle AA^{-1} \right\rangle_{ij} = \sum_{k=1}^{n} a_{ik} \left\langle A^{-1} \right\rangle_{kj},$$

тому рівняння (4.39) можна виразити в такій формі:

$$\left\langle AA^{-1} \right\rangle_{ij} = \sum_{k=1}^{n} a_{ik} \frac{A_k^j}{\det A}. \tag{4.40}$$

Спростимо праву частину рівняння за допомогою формули Лапласа (4.35):

$$\sum_{k=1}^{n} a_{ik} \frac{A_k^j}{\det A} = \frac{\sum\limits_{k=1}^{n} a_{ik} A_k^j}{\det A} = \frac{\delta_{ij} \det A}{\det A} = \delta_{ij}.$$

Спрощене рівняння (4.40) має форму

$$\left\langle AA^{-1} \right\rangle_{ij} = \delta_{ij} \tag{4.41}$$

і справедливе для всіх елементів матриці AA^{-1}. Оскільки сукупність елементів одиничної матриці є символом Кронекера, то сукупність рівнянь (4.41) еквівалентна матричному рівнянню $AA^{-1} = E$.

Означення 4.11. Матрицю, визначник якої дорівнює нулю, називають *особливою*.

Зауваження 4.3. Із формули (4.39) випливає, що матриці, що обернена особливій, не існує.

4.5. Лінійна залежність рядків або стовпців визначника. Критерій рівності визначника нулю

4.5.1. Умова лінійної залежності рядків або стовпців матриці

Як було зазначено, рядки/стовпці визначника $\mathbf{a}_i$ є векторами лінійного простору. З огляду на це, на них поширюється поняття лінійної комбінації, означене як

$$\boldsymbol{l} \overset{def}{\equiv} \sum_{i=1}^{p} \alpha_i \mathbf{a}_i \,, \tag{4.42}$$

де α_i – числові коефіцієнти, та умова лінійної залежності векторів, яка полягає в тому, що існує така добірка коефіцієнтів, для якої справджується рівність

$$\sum_{i=1}^{p} \alpha_i \mathbf{a}_i = \mathbf{0}, \tag{4.43}$$

де хоча б один коефіцієнт не дорівнює нулю.

4.5.2. Базисний мінор. Ранг матриці

Розглянемо матрицю $A(a_{ij})$ порядку $m \times n$, вважаючи для визначеності, що $m < n$. Ця матриця має мінори порядків $2, 3 \ldots m$. У загальному випадку серед цих мінорів є такі, що дорівнюють нулю, і такі, що нулю не дорівнюють. Позначимо числами $r_1 < r_2 < \ldots < r$ порядки тих мінорів, що не дорівнюють нулю. За таких позначень r – це *найбільший з порядків ненульових мінорів.*

Означення 4.12. Найбільший з порядків ненульових мінорів називають *рангом матриці*, пишуть $\operatorname{rang} A = r$.

Для матриці $A(a_{ij})$ може існувати декілька *не рівних нулю мінорів порядку* r. Кожен з таких мінорів називають *базисним мінором* матриці $A(a_{ij})$ [1].

[1] Базисний мінор матриці називають також базисним мінором її визначника.

Означення 4.13. Рядки та стовпці матриці, з елементів яких складається базисний мінор, називають базисними рядками та стовпцями цієї матриці[1].

Чому ненульовий мінор порядку r і його рядки/стовпці називають базисними пояснюється двома теоремами.

Теорема 4.1. Базисні рядки/стовпці матриці лінійно незалежні.

Доведення. Припустимо, що базисні рядки/стовпці матриці лінійно залежні, і впевнимось в тому, що таке припущення хибне. За такого припущення має справджуватися умова лінійної залежності рядків/стовпців базисного мінору:

$$\sum_{i=1}^{r} \alpha_i \mathbf{a}_i = \mathbf{0},$$

де хоча б один коефіцієнт (для визначеності – α_k) не дорівнює нулю[2]. Поділивши обидві частини рівняння (4.43) на цей коефіцієнт, виразимо зроблене припущення у формі рівняння:

$$\mathbf{a}_k + \sum_{i \neq k}^{r} (\alpha_i / \alpha_k) \mathbf{a}_i = \mathbf{0}. \qquad (4.44)$$

Позначимо базисний мінор як $M(\mathbf{a}_i)$, де $i = \overline{1, r}$. За означенням 4.12, базисний мінор не може дорівнювати нулю: $M(\mathbf{a}_i) \neq 0$.

Згідно з доведеним у п. 4.1.5 наслідком 3^0 основних властивостей визначника *базисний мінор не зміниться*, якщо спочатку замінити вектор $\mathbf{a}_k$ на $\mathbf{a}_k + (\alpha_1 / \alpha_k)\mathbf{a}_1$, потім вектор $\mathbf{a}_k + (\alpha_1 / \alpha_k)\mathbf{a}_1$ замінити на $\mathbf{a}_k + (\alpha_1 / \alpha_k)\mathbf{a}_1 + (\alpha_2 / \alpha_k)\mathbf{a}_2$ і так далі, аж доки не отримаємо визначник, у якому під номером k стоятиме рядок/стовпець $\mathbf{a}_k + \sum_{i \neq k}^{r} (\alpha_i / \alpha_k)\mathbf{a}_i$, усі елементи якого за припущенням (4.44) дорівнюють нулю. Згідно з наслідком 4^0 основних властивостей визначника і формулою розкладу за елементами

---

[1] Базисні рядки та стовпці матриці називають також рядками та стовпцями її визначника.

[2] Оскільки переставлення рядків/стовпців базисного мінору веде лише до зміни його знака, вважаємо, що ці рядки/стовпці мають номери від 1 до r.

рядка/стовпця цей визначник дорівнює нулю, чого не може бути, оскільки він дорівнює базисному мінору. Звідси випливає, що припущення про лінійну залежність базисних рядків/стовпців матриці хибне.

Теорема 4.2. Будь-який рядок/стовпець матриці можна представити у формі лінійної комбінації базисних рядків/стовпців.

Доведення. Достатньо довести, що будь-який з небазисних рядків/стовпців може бути представлений у формі лінійної комбінації базисних. Для цього слід довести, що для кожного небазисного рядка/стовпця $\mathbf{a}_p$ існують такі числові коефіцієнти C_k, які задовольняють рівність

$$\mathbf{a}_p = \sum_{k=1}^{r} C_k \mathbf{a}_k \ . \qquad (4.45)$$

Щоб зробити це, додамо до базисного мінору небазисний рядок $\mathbf{a}_p$ і небазисний стовпець $\mathbf{a}_q$ та розглянемо визначник порядку $r+1$, який згідно з означенням базисного мінору має дорівнювати нулю:

$$\Delta \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} & a_q \\ a_{12} & a_{22} & \cdots & a_{2r} & a_{2q} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} & a_{rq} \\ a_{p1} & a_{p2} & \cdots & a_{pr} & a_{pq} \end{vmatrix} = 0 \ .$$

Розкладемо цей визначник за елементами останнього рядка:

$$\Delta = \sum_{k=1}^{r} a_{pk} A_k^p + a_{pq} M(\mathbf{a}_i) \ , \qquad (4.46)$$

де A_k^p – алгебраїчні доповнення, що складаються з елементів базисного мінору. Із розкладу (4.46) випливають рівняння

$$a_{pq} = -\sum_{k=1}^{r} a_{pk} \frac{A_k^p}{M(\mathbf{a}_i)} \ , \ \ p = \overline{r+1, n} \ .$$

Зваживши на те, що алгебраїчні доповнення базисного мінору не залежать від того, яким є останній рядок визначника Δ, введемо до розгляду коефіцієнти $C_k \equiv -A_k^p / M(\mathbf{a}_i)$ й отримаємо рівняння

$$a_{pq} = \sum_{k=1}^{r} C_k a_{pk} + \sum_{k=r+1}^{n} 0 \cdot a_{pk}\,,\ \ p = \overline{r+1,n}\,,$$

еквівалентні векторному рівнянню

$$\mathbf{a}_p = \sum_{\substack{k=1\\k\neq p}}^{n} C_k \mathbf{a}_k\,, \tag{4.47}$$

яке і доводить твердження теореми.

Доведені теореми 4.1 та 4.2 показують, що рядки/стовпці матриці є елементами векторного простору вимірності r, а сукупність рядків/стовпці матриці, належних базисному мінору, є базисом цього простору, що і пояснює термін "базисний мінор".

Наслідок 4.2. Якщо справедлива нерівність $n > r$, то стовпці $\{\mathbf{a}_k\}_{k=\overline{1,n}}$ лінійно залежні.

Дійсно, рівняння (4.47) еквівалентне рівнянню

$$\sum_{k=1}^{n} C_k \mathbf{a}_k = \mathbf{0}\,,$$

яке є умовою лінійної залежності векторів $\{\mathbf{a}_k\}_{k=\overline{1,n}}$, тому що принаймні один коефіцієнт ($C_p = -1$) не дорівнює нулю.

Наслідок 4.3. Ранг матриці дорівнює максимальній кількості її лінійно незалежних рядків/стовпців.

Наслідок 4.3 іноді вважають означенням рангу матриці, і з цього означення роблять висновок, що ранг матриці дорівнює найбільшому порядку ненульових мінорів її визначника.

4.5.3. Критерій рівності визначника нулю

Сформулюємо критерій рівності визначника нулю: *визначник дорівняє нулю тоді і лише тоді, коли його рядки/стовпці лінійно залежні.* Цей критерій містить пряму та обернену теореми.

Пряма теорема: відомо, що вектори $\mathbf{a}_i$ у визначнику $D(\mathbf{a}_i)$ лінійно залежні. Треба довести, що визначник $D(\mathbf{a}_i)$ дорівнює нулю.

Доведення. Доведення цієї теореми подібне доведенню теореми 4.1. За умовою прямої теореми справджується рівняння (4.43), у якому $p = n$ і коефіцієнт $\alpha_1 \neq 0$ [1]. Виразимо з нього вектор $\mathbf{a}_1$:

$$\mathbf{a}_1 = -\sum_{i=2}^{n} (\alpha_i / \alpha_1)\mathbf{a}_i \ . \tag{4.48}$$

Згідно з доведеним у п. 4.1.5 наслідком 3^0 основних властивостей визначника визначник $D(\mathbf{a}_i)$ не зміниться, якщо замінити вектор $\mathbf{a}_1$ на $\mathbf{a}_1 + (\alpha_2 / \alpha_1)\mathbf{a}_2$, тому має виконуватись рівність

$$D(\mathbf{a}_i) = D_1\left(\mathbf{a}_1 + \frac{\alpha_2}{\alpha_1}\mathbf{a}_2, \mathbf{a}_k\right),$$

де $i = \overline{1,n}$, $k = \overline{2,n}$. Додавши до першого рядка/стовпця визначника D_1 доданок $(\alpha_3 / \alpha_1)\mathbf{a}_3$, отримаємо рівність

$$D(\mathbf{a}_i) = D_2\left(\mathbf{a}_1 + \frac{\alpha_2}{\alpha_1}\mathbf{a}_2 + \frac{\alpha_3}{\alpha_1}\mathbf{a}_3, \mathbf{a}_k\right).$$

Діючи надалі у такий спосіб, дійдемо висновку, що має справджуватись рівність

$$D(\mathbf{a}_i) = D_{n-1}\left(\mathbf{a}_1 + \sum_{i=2}^{n}\frac{\alpha_i}{\alpha_1}\mathbf{a}_i, \mathbf{a}_k\right).$$

Із виразу (4.48) випливає, що перший рядок/стовпець визначника D_{n-1} дорівнює нулю, а отже, дорівнює нулю і сам визначник D_{n-1}, і визначник D.

Обернена теорема: відомо, що визначник $D(\mathbf{a}_i)$ дорівнює нулю. Треба довести, що вектори $\mathbf{a}_i$ лінійно залежні.

Доведення. Оскільки визначник $D(\mathbf{a}_i)$ дорівнює нулю, то його порядок перевищує порядок базисного мінору. Це означає, що в

---

[1] Переставлення місцями рядків/стовпців визначника змінює його знак, не змінюючи його абсолютної величини, тому можна вважати, що рядок/стовпець, коефіцієнт якого не дорівнює нулю, поставлений на перше місце.

цьому визначнику є рядки $\mathbf{a}_p$, які не входять до базисного мінору. Згідно з теоремою 4.2 кожен із цих рядків можна представити у формі лінійної комбінації базисних рядків, а потім утворити лінійну комбінацію

$$\mathbf{a}_p - \sum_{k=1}^{n} C_k \mathbf{a}_k = \mathbf{0}$$

(див. (4.47)). Це доводить твердження теореми, оскільки утворена лінійна комбінація не тривіальна, якщо $\mathbf{a}_p \neq \mathbf{0}$, а якщо $\mathbf{a}_p = \mathbf{0}$, то можна утворити нетривіальну комбінацію, що дорівнює нулю:

$$1 \cdot \mathbf{0} + \sum_{k \neq p}^{n} 0 \cdot \mathbf{a}_k = \mathbf{0}.$$

Зауваження 4.4. Усі інші випадки, у яких визначник дорівнює нулю, є окремими випадками лінійної залежності його рядків/стовпців.

Наприклад, щойно було показано, що рядки визначника лінійно залежні, якщо серед них є рядок, усі елементи якого дорівнюють нулю, і це показує, що сформульований у п. 4.1.5 наслідок 4^0 основних властивостей визначника відповідає доведеному критерію. Ще більш очевидно, що йому відповідають наслідки 1^0 та 2^0.

РОЗДІЛ 5
Системи лінійних алгебраїчних рівнянь

5.1. Означення та різні форми запису системи лінійних алгебраїчних рівнянь

Означення 5.1. Системою лінійних алгебраїчних рівнянь для невідомих x_j $(j=\overline{1,n})$ називають таку систему:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n = b_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{5.1}$$

Системи лінійних алгебраїчних рівнянь називають також лінійними алгебраїчними системами. Числа a_{ij} $(i=\overline{1,m})$ називають коефіцієнтами лінійної алгебраїчної системи, а числа b_i – правими частинами рівнянь.

Означення 5.2. Систему лінійних алгебраїчних рівнянь називають однорідною, якщо праві частини всіх рівнянь дорівнюють нулю, і неоднорідною, якщо хоча б одна з правих частин не дорівнює нулю.

Система рівнянь (5.1) може бути виражена у формі одного матричного рівняння:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix}. \tag{5.2}$$

З огляду на громіздкість виразів (5.1), (5.2), введемо позначення

$$\mathbf{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},\ \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},\ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\ A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (5.3)$$

і запишемо перший із цих виразів у скороченій формі

$$\sum_{j=1}^{n} a_{ij} x_j = b_i\ , \quad (5.4)$$

а другий – у формі

$$A\mathbf{x} = \mathbf{b}\ . \quad (5.5)$$

У подальшому будемо користуватися також "векторною" формою запису системи (5.2):

$$\sum_{j=1}^{n} x_j \mathbf{a}_j = \mathbf{b}\ . \quad (5.6)$$

Означення 5.3. Матрицю A називають основною, а матрицю

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \quad (5.7)$$

– розширеною матрицею лінійної алгебраїчної системи.

Означення 5.4. Якщо кількість рівнянь системи дорівнює кількості невідомих, тобто, якщо $m = n$, систему називають квадратною.

Означення 5.5. Розв'язком системи лінійних алгебраїчних рівнянь називають такі значення невідомих, підставлення яких до системи перетворює всі її рівняння на тотожності; позначатимемо ці значення як $x_j^{(\mathrm{s})}$ ($j = \overline{1,n}$), а утворений з них вектор-стовпець – як $\mathbf{x}^{(\mathrm{s})}$, де "s" – перша літера слова "solution".

Означення 5.6. Тривіальним розв'язком системи лінійних рівнянь називають розв'язок $\mathbf{x}^{(\mathrm{s})} = \mathbf{0}$, або, що те саме, $x_j^{(\mathrm{s})} = 0$ для всіх значень $j = \overline{1,n}$.

Означення 5.7. Лінійну алгебраїчну систему називають сумісною, якщо вона має хоча б один розв'язок[1], а в оберненому випадку її називають несумісною.

Далі наведемо стислий виклад лише тих відомостей, які знадобляться в наступних розділах підручника. Більш повний виклад теорії лінійних алгебраїчних систем та методів їхнього розв'язання можна знайти в підручниках [6, 7].

5.2. Однорідні лінійні системи

Розглянемо однорідну систему лінійних алгебраїчних рівнянь:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = 0, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n = 0. \end{cases} \tag{5.8}$$

Цілком очевидно, що однорідна система лінійних алгебраїчних рівнянь має тривіальний розв'язок $x_j^{(s)} = 0$ для всіх $j = \overline{1,n}$. Не таким очевидним, але надзвичайно важливим, є *питання про наявність нетривіальних розв'язків однорідної системи лінійних алгебраїчних рівнянь*. Надати повний перелік математичних, фізичних та інженерних задач, для розв'язання яких це питання має вирішальне значення, неможливо, тому вкажемо лише на ті, що пов'язані зі змістом наступних розділів підручника, а саме:

- ✓ відшукання власних значень і власних векторів лінійних операторів (підрозд. 6.3);
- ✓ зведення квадратичної форми до суми квадратів та визначення канонічного базису (підрозд. 7.3);
- ✓ інтегрування систем лінійних диференціальних рівнянь першого порядку (розд. 9);

---

[1] Навіть, якщо єдиний розв'язок – тривіальний. Більш правильно було б казати, що сумісними є рівняння, що входять до системи, але надане означення загальновживане.

✓ дослідження стійкості стаціонарного стану фізичних та інженерних систем при виникненні малого збурення цього стану (розд. 10);

✓ визначення спектрів вільних коливань динамічних систем (розд. 7, п. 7.3.3).

Для з'ясування умови, за якої система рівнянь (5.8) має нетривіальний розв'язок, виразимо її у векторній формі:

$$\sum_{j=1}^{n} x_j \mathbf{a}_j = \mathbf{0}. \qquad (5.9)$$

Теорема 5.1. Система однорідних лінійних алгебраїчних рівнянь має нетривіальний розв'язок тоді і лише тоді, коли кількість невідомих перевищує кількість лінійно незалежних рівнянь системи.

Теорема 5.1 містить пряме й обернене твердження:

а) якщо кількість невідомих n перевищує кількість лінійно незалежних рівнянь r, то система (5.9) має нетривіальний розв'язок;

б) якщо система (5.9) має нетривіальний розв'язок $x_j^{(s)}$, то кількість невідомих перевищує кількість лінійно незалежних рівнянь.

Доведемо твердження а). Відомо, що $n > r$. Треба довести наявність нетривіального розв'язку. Для доведення зазначимо, що кількість лінійно незалежних рівнянь системи дорівнює кількості базисних, тобто лінійно незалежних, рядків та стовпців її основної матриці A (див. означення 4.13 та теорему 4.1). Стовпці матриці A – це вектори $\mathbf{a}_j$. Згідно з твердженням а) кількість базисних векторів дорівнює r та є меншою від їхньої повної кількості n. Згідно з теоремою 4.2 кожний небазисний вектор $\mathbf{a}_k$ є лінійною комбінацією базисних, тобто справджується рівність

$$\mathbf{a}_k = \sum_{j=1}^{r} \alpha_j \mathbf{a}_j, \qquad (5.10)$$

у якій *усі або деякі коефіцієнти* α_j *не дорівнюють нулю*[1]. Із виразу (5.10) випливає, що є справедливою рівність

$$\sum_{j=1}^{r}\alpha_j\mathbf{a}_j - \mathbf{a}_k + \sum_{j=r+1}^{k-1} 0\cdot\mathbf{a}_j + \sum_{j=k+1}^{n} 0\cdot\mathbf{a}_j = \mathbf{0}.$$

Надамо невідомим x_j значення $x_j = 1$, якщо $j = k$, $x_j = -\alpha_j$, якщо $1 \le j \le r$, $x_j = 0$, якщо $r < j < k$, $k+1 \le j \le n$. Ці значення задовольняють рівність

$$\sum_{j=1}^{r} x_j\mathbf{a}_j + \sum_{j=r+1}^{k-1} x_j\mathbf{a}_j + x_k\mathbf{a}_k + \sum_{j=k+1}^{n} x_j\mathbf{a}_j = \mathbf{0}.$$

Урахувавши, що

$$\sum_{j=1}^{r} x_j\mathbf{a}_j + \sum_{j=r+1}^{k-1} x_j\mathbf{a}_j + x_k\mathbf{a}_k + \sum_{j=k+1}^{n} x_j\mathbf{a}_j = \sum_{j=1}^{n} x_j\mathbf{a}_j,$$

дійдемо висновку, що значення, надані невідомим, задовольняють векторне рівняння (5.9), а отже, вони є нетривіальним розв'язком однорідної системи лінійних алгебраїчних рівнянь.

Доведемо твердження б). Відомо, що система має нетривіальний розв'язок. Треба довести, що $n > r$. Для цього достатньо взяти до уваги, що, якщо однорідна лінійна система рівнянь має нетривіальний розв'язок $x_j = x_j^{(s)}$, то є виконаною рівність

$$\sum_{j=1}^{n} x_j^{(s)}\mathbf{a}_j = \mathbf{0}, \tag{5.11}$$

у якій ті коефіцієнти $x_j^{(s)}$, що стоять біля базисних векторів, не дорівнюють нулю. Ця рівність є умовою лінійної залежності векторів $\left\{\mathbf{a}_j\right\}_{j=\overline{1,n}}$ (див. означення 3.7 та рівність (3.4)).

Наслідок 5.1. Квадратна система однорідних лінійних алгебраїчних рівнянь має нетривіальний розв'язок тоді і лише тоді, коли її визначник дорівнює нулю.

---

[1] У виразі (5.10) враховано, що сума у системі рівнянь (5.9), а отже, і система рівнянь, не змінюється при переставленні доданків, тому можна вважати, що базисні вектори мають номери від одиниці до r.

Наслідок 5.1 випливає з того, що кількість невідомих у квадратній системі рівнянь дорівнює порядку визначника основної матриці системи n, а кількість лінійно незалежних рівнянь системи – кількості лінійно незалежних рядків/стовпців цього визначника r, тому справедлива нерівність $n > r$. Згідно з наслідком 4.2, у такому разі стовпці визначника лінійно залежні. Лінійна залежність стовпців є критерієм рівності визначника нулю.

Зауважимо, що для фізико-математичних наук, інженерії та змісту наступних розділів підручника наслідок 5.1 навіть важливіший, ніж сама теорема 5.1. Саме з умови рівності нулю визначника системи лінійних алгебраїчних рівнянь визначаються спектри механічних та електромагнітних коливань, власні числа лінійних операторів та багато інших фізичних і математичних величин.

5.3. Неоднорідні лінійні системи. Формули Крамера

Умова наявності розв'язку системи неоднорідних лінійних алгебраїчних рівнянь сформульована в теоремі, відомій як теорема Кронекера–Капелі.

Теорема 5.2. Неоднорідна система лінійних алгебраїчних рівнянь сумісна тоді і лише тоді, коли ранг її основної матриці дорівнює рангу розширеної матриці.

Теорема 5.2, як і попередня, містить пряме й обернене твердження:

а) якщо $\operatorname{rang} A = \operatorname{rang} B$, то система

$$\sum_{j=1}^{n} x_j \mathbf{a}_j = \mathbf{b} \tag{5.12}$$

сумісна[1];

б) якщо система (5.12) сумісна, то $\operatorname{rang} A = \operatorname{rang} B$.

[1] Вираз (5.12) є векторною формою неоднорідної системи лінійних алгебраїчних рівнянь.

Доведення твердження а). Відомо, що $\operatorname{rang} A = \operatorname{rang} B$. Треба довести, що система (5.12) сумісна. Для цього позначимо $\operatorname{rang} A \equiv r$ і згадаємо, що за означенням 4.12 ранг матриці дорівнює порядку її базисного мінору. Усі стовпці $\left\{\mathbf{a}_j\right\}_{j=\overline{1,r}}$ базисного мінору матриці A входять до складу матриці B, а оскільки ранги цих матриць однакові, то базисний мінор матриці A водночас є базисним мінором матриці B. Згідно з теоремою 4.2 стовпець $\mathbf{b}$ матриці B може бути представлений у формі лінійної комбінації стовпців, належних її базисному мінору:

$$\mathbf{b} = \sum_{j=1}^{r} \alpha_j \mathbf{a}_j \,. \tag{5.13}$$

Надамо невідомим, що входять до системи (5.12), значення $x_j = \alpha_j$, $1 \le j \le r$, та $x_j = 0$, $r < j \le n$. Із виразу (5.13) випливає, що ці значення задовольняють векторне рівняння

$$\mathbf{b} = \sum_{j=1}^{n} x_j \mathbf{a}_j \,,$$

еквівалентне неоднорідній системі алгебраїчних рівнянь (5.12).

Доведення твердження б). Відомо, що система рівнянь (5.12) сумісна, тобто існують такі значення невідомих, що справджується рівність

$$\sum_{j=1}^{n} x_j^{(s)} \mathbf{a}_j = \mathbf{b} \,. \tag{5.14}$$

Треба довести, що $\operatorname{rang} A = \operatorname{rang} B$. Для цього слід взяти до уваги, що ліва частина векторної рівності (5.14) є лінійною комбінацією векторів, що належать сукупності $\left\{\mathbf{a}_j\right\}_{j=\overline{1,n}}$. З огляду на це, вектор $\mathbf{b}$ є *одним із векторів лінійної оболонки* цієї сукупності. Вимірність лінійної оболонки сукупності векторів дорівнює максимальній кількості лінійно незалежних векторів цієї сукупності, тобто рангу матриці A (див. теорему 3.4 і наслідок 4.3). Оскільки в *усій лінійній оболонці* стовпців $\left\{\mathbf{a}_j\right\}_{j=\overline{1,n}}$ основної матриці існує лише r лінійно незалежних векторів, то в сукупності вектор-стовпців $\{\mathbf{a}_j, \mathbf{b}\}_{j=\overline{1,n}}$ розширеної матриці їх не може бути більше.

Теорема Кронекера–Капелі стосується неоднорідної лінійної системи, що складається з m рівнянь для n невідомих, тобто основна матриця такої системи не квадратна. Але саме квадратні системи найчастіше зустрічаються у підручниках з фізико-математичних та фізико-технічних наук. Існують декілька методів розв'язання таких систем, але найбільш важливим є метод, у якому розв'язок неоднорідної лінійної системи виражається *формулами Крамера*, тому що ці формули необхідні для розгляду важливих питань, що стосуються лінійних операторів, білінійних форм, лінійних диференціальних рівнянь тощо. Ці формули містять визначник основної матриці неоднорідної лінійної системи алгебраїчних рівнянь $\det A$ та *доповняльні визначники* $\Delta^{(k)}$, які відрізняються від $\det A$ лише тим, що стовпець $\mathbf{a}_k$ основної матриці замінений у них на стовпець $\mathbf{b}$, утворений з правих частин рівнянь. Формули Крамера мають таку форму:

$$x_k = \frac{\Delta^{(k)}}{\det A}, \quad k = \overline{1,n}. \tag{5.15}$$

Доведемо справедливість формул Крамера. Для цього домножимо праві та ліві частини рівнянь (5.4) на алгебраїчні доповнення елементів a_{ik} матриці A:

$$\sum_{j=1}^{n} a_{ij} x_j A_k^i = b_i A_k^i, \; i = \overline{1,n}. \tag{5.16}$$

Сума правих частин рівнянь (5.16) дорівнює сумі лівих частин:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_j A_k^i = \sum_{i=1}^{n} b_i A_k^i. \tag{5.17}$$

Розглянувши розклад визначника матриці A за елементами стовпця $\mathbf{a}_k$ (рівняння (4.33)), помічаємо, що права частина рівняння (5.17) відрізняється від цього розкладу лише заміною стовпця $\mathbf{a}_k$ на стовпець $\mathbf{b}$, утворений з правих частин. Щодо лівої частини рівняння (5.17), то в ній можна змінити порядок обчислення сум і отримати рівняння

$$\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij} x_j A_k^i = \Delta^{(k)},$$

де $\Delta^{(k)}$ – це визначник, який відрізняється від визначника матриці A лише заміною стовпця $\mathbf{a}_k$ на стовпець $\mathbf{b}$. Множник x_j однаковий для всіх доданків суми за індексом i, тому його можна винести за знак суми й отримати рівняння

$$\sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ij} A_k^i = \Delta^{(k)}.$$

Формула Лапласа (4.34) дозволяє позбутися суми за індексом i та отримати рівність

$$\sum_{j=1}^{n} x_j \delta_{kj} \det A = \Delta^{(k)},$$

у якій під знаком суми стоїть символ Кронекера. Завдяки цьому не дорівнює нулю лише той доданок суми, у якому $j = k$, а отже, справджуються рівності

$$x_k \det A = \Delta^{(k)},$$

з яких безпосередньо випливають формули Крамера.

Цілком очевидно, що формули Крамера не дозволяють розв'язати квадратну систему лінійних алгебраїчних рівнянь у тому випадку, коли визначник її основної матриці дорівнює нулю. Розглянемо цей випадок у такий спосіб, який дозволяє зрозуміти властивості лінійних операторів, квадратичних форм та систем лінійних диференціальних рівнянь першого порядку. Для цього візьмемо до уваги, що рядки визначника, що дорівнює нулю, лінійно залежні (див. п. 4.5.3), а отже, лінійно залежними є рівняння лінійної алгебраїчної системи (5.4) з визначником основної матриці, що дорівнює нулю.

Першим кроком при розв'язуванні системи з нульовим визначником основної матриці є визначення базисного мінору цієї матриці. Ті рівняння, коефіцієнти яких входять до рядків базисного мінору, лінійно незалежні, їхня кількість дорівнює порядку базисного мінору r. Назвемо ці рівняння базисними і надамо їм номери від одиниці до r[1]. Такі ж номери надамо тим невідомим, біля

[1] Зміна нумерації рівнянь не змінює їхнього розв'язку.

яких стоять коефіцієнти, які входять до базисного мінору. На другому кроці, ті доданки рівнянь, коефіцієнти яких не входять до базисного мінору, перенесемо до правих частин рівнянь. У результаті отримаємо систему, що складається з r рівнянь

$$\sum_{j=1}^{r} a_{ij}x_j = b_i - \sum_{j=r+1}^{n} a_{ij}x_j\ ,\quad i=\overline{1,n}\ .$$

Небазисні рівняння можна не розглядати, тому що вони є лінійними комбінаціями базисних (див. теорему 4.2) і не накладають додаткових умов на величини x_j. Позначивши

$$b_i - \sum_{j=r+1}^{n} a_{ij}x_j \equiv \tilde{b}_i\ ,$$

отримаємо квадратну систему рівнянь:

$$\sum_{j=1}^{r} a_{ij}x_j = \tilde{b}_i\ ,\quad i=\overline{1,r}\ . \tag{5.18}$$

Третім кроком є розв'язання системи рівнянь (5.18). Визначник основної матриці цієї системи є базисним мінором M матриці A, тому він не дорівнює нулю, а отже, для відшукання розв'язку придатні формули Крамера, які в цьому випадку мають форму

$$x_k = \frac{\tilde{\Delta}^{(k)}}{M}\ ,\quad k=\overline{1,r}\ , \tag{5.19}$$

де $\tilde{\Delta}^{(k)}$ є допоміжними визначниками, у яких стовпці з номером k складаються з правих частин $\tilde{b}_i$ рівнянь (5.18).

● **Приклад 5.1.** Для пояснення описаної вище процедури розв'яжемо систему рівнянь

$$\begin{cases} x+y+z=3, \\ x-y+z=1, \\ 3x+y+3z=7. \end{cases} \tag{5.20}$$

Визначником основної матриці цієї системи є

$$\det A = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 3 & 1 & 3 \end{vmatrix} = 0\ .$$

Бачимо, що третій рядок цього визначника дорівнює сумі другого рядка та першого рядка, помноженого на 2.

На першому кроці обчислюємо мінор

$$M_{12}^{12} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$$

і бачимо, що він не дорівнює нулю, а отже, є базисним мінором. Тому перше та друге рівняння системи (5.20) є базисними рівняннями, а невідомі x та y – базисними невідомими.

На другому кроці відкидаємо із системи небазисне рівняння та переносимо небазисну змінну до правих частин базисних рівнянь:

$$\begin{cases} x + y = 3 - z \\ x - y = 1 - z. \end{cases}$$

На третьому кроці розв'язуємо систему базисних рівнянь за допомогою формул (5.19)[1]:

$$\tilde{\Delta}^{(x)} = \begin{vmatrix} 3-z & 1 \\ 1-z & -1 \end{vmatrix} = 2z - 4\,, \quad \tilde{\Delta}^{(y)} = \begin{vmatrix} 1 & 3-z \\ 1 & 1-z \end{vmatrix} = -2\,,$$

$$x = \frac{\tilde{\Delta}^{(x)}}{M_{12}^{12}} = 2 - z\,, \quad y = \frac{\tilde{\Delta}^{(y)}}{M_{12}^{12}} = 1\,.$$

Бачимо, що система рівнянь (5.20) визначає лише зв'язок між двома з трьох невідомих і не накладає обмежень на величину невідомої z. Тому цю невідому слід вважати довільною сталою та виразити розв'язок системи рівнянь (5.20) у формі

$$z = C\,, \quad y = 1\,, \quad x = 2 - C\,.\bullet$$

Додаткові відомості щодо розв'язання лінійних систем алгебраїчних рівнянь можна знайти в підручнику [6].

[1] Цю систему легше розв'язати, додавши рівняння одне до одного, але застосування формул Крамера ілюструє загальний підхід до проблеми.

РОЗДІЛ 6
Лінійні оператори

6.1. Означення лінійного оператора. Матриця лінійного оператора в ортонормованому базисі

Як відомо, "оператор" – це той, хто виконує певні (зазвичай передбачені інструкціями або правилами) *дії*. Щоб зрозуміти, чому слово "оператор" застосовується в лінійній алгебрі, наведемо два приклади – один з оптики і один з електроніки.

На рис. 6.1 показано зображення радіус-вектора $\mathbf{r}$ точки $P(x, y, z)$ на листі паперу, або екрані комп'ютера. Око людини, яка дивиться на зображення вектора $\mathbf{r}$, проєктує його з листа паперу на сітківку ока, зменшуючи модуль вектора $\mathbf{r}$ та змінюючи напрямок на протилежний. Можна сказати, що око є приладом, *дія* якого полягає в перетворенні геометричного вектора $\mathbf{r}$ з координатами x, y, z на геометричний вектор $\mathbf{r}'$ з координатами $x' = -kx$, $y' = -ky$, $z' = z = 0$ (рис. 6.1).

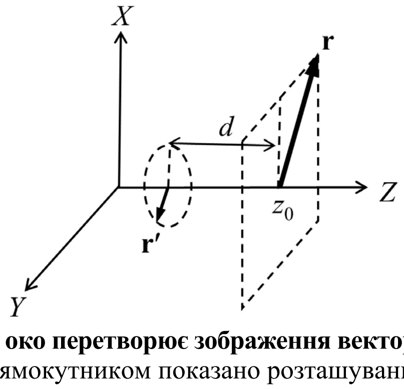

Рис. 6.1. Людське око перетворює зображення вектора r на вектор r′.
(Пунктирним прямокутником показано розташування листа паперу або екрана комп'ютера, а пунктирним колом – положення сітківки ока, z_0 – відстань від початку координат до листа паперу, d – відстань від листа паперу або екрана до сітківки ока)

Важливими приладами електроніки є підсилювачі електричних сигналів. Підсилювач перетворює "вхідний" сигнал

$$u(t) = u_A \cos \omega t$$

на "вихідний"

$$U(t) = k(\omega) u_A \cos \omega t,$$

де t – час, u_A та ω – відповідно амплітуда і частота вхідного сигналу, $k(\omega)$ – залежний від частоти коефіцієнт підсилення сигналу. З погляду математики підсилювач є приладом, *дія* якого полягає в перетворенні функції $u(t)$ на функцію $U(t)$.

Означення 6.1. Оператором $\hat{L}$, що *діє* в лінійному просторі $\mathcal{L}$, називають правило, за яким кожний елемент $\mathbf{x}$, що належить цьому простору, перетворюється на елемент $\mathbf{y}$, що належить цьому ж простору.

Пишуть $\hat{L}\mathbf{x} = \mathbf{y}$; кажуть, що оператор $\hat{L}$ перетворює елемент $\mathbf{x}$ на елемент $\mathbf{y}$.

Зауваження 6.1. Якщо оператор перетворює кожен елемент простору $\mathcal{L}_1$ на елемент простору $\mathcal{L}_2$, то кажуть, що оператор діє з простору $\mathcal{L}_1$ у простір $\mathcal{L}_2$.

Зауваження 6.2. Якщо згадати, що функцією числової змінної називають правило, за яким кожному значенню змінної ставиться у відповідність певне число, стає зрозуміло, що поняття "оператор" є поширенням поняття "функція" із чисел на елементи лінійних просторів.

Означення 6.2. Оператор $\hat{\mathbf{I}}$, який не перетворює (не змінює) елементи простору, називають *тотожним оператором*, його дія визначається рівністю $\hat{\mathbf{I}}\mathbf{x} = \mathbf{x}$, для будь-якого елементу $\mathbf{x}$, що належить простору $\mathcal{L}$. Оператор, який перетворює будь-який елемент простору на нульовий елемент, називають нульовим оператором; оскільки результат його дії тотожний результату множення елементу простору на число нуль, то нульовий оператор позначають символом 0 і кажуть, що цей оператор дорівнює нулю.

Означення 6.3. Оператор $\hat{L}$ називають лінійним, якщо його дія на елементи лінійного простору має такі властивості:

I) $\hat{L}(\mathbf{x}+\mathbf{y}) = \hat{L}\mathbf{x} + \hat{L}\mathbf{y}$;

II) $\hat{L}\lambda\mathbf{x} = \lambda\hat{L}\mathbf{x}$,

де λ є дійсним чи комплексним числом залежно від того, у якому просторі (або з якого у який простір) діє оператор.

Для розв'язання інженерних задач і вивчення природних явищ необхідно *кількісно* характеризувати результат дії лінійного оператора на елементи простору, тобто поширити на оператори метод координат. Зрозуміти, як це зробили математики, можна на простому прикладі.

● **Приклад 6.1.** Розкладемо радіус-вектор $\mathbf{r}$ точки простору по ортах декартової системи координат XYZ:

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z . \tag{6.1}$$

Подіємо на вектор $\mathbf{r}$ оператором повороту $\hat{R}_\alpha$ радіус-векторів усіх точок простору на кут α навколо осі OZ, зваживши на те, що цей оператор лінійний, тобто, урахувавши властивості I), II) з означення 6.3:

$$\hat{R}_\alpha\mathbf{r} = x\hat{R}_\alpha\mathbf{e}_x + y\hat{R}_\alpha\mathbf{e}_y + z\hat{R}_\alpha\mathbf{e}_z .$$

Введемо такі позначення для векторів, перетворених оператором повороту: $\hat{R}_\alpha\mathbf{r} \equiv \mathbf{r}'$, $\hat{R}_\alpha\mathbf{e}_x \equiv \mathbf{e}'_x$, $\hat{R}_\alpha\mathbf{e}_y \equiv \mathbf{e}'_y$. Врахувавши, що поворот навколо осі OZ не змінює напрямок орта $\mathbf{e}_z$, отримаємо рівняння

$$\mathbf{r}' = x\mathbf{e}'_x + y\mathbf{e}'_y + z\mathbf{e}_z . \tag{6.2}$$

Проєкція вектора $\mathbf{r}'$ на вісь OX дорівнює скалярному добутку цього вектора на орт $\mathbf{e}_x$; утворивши скалярний добуток обох частин рівності (6.2) на цей орт, знаходимо рівняння:

$$x' = x(\mathbf{e}'_x\mathbf{e}_x) + y(\mathbf{e}'_y\mathbf{e}_x) + z(\mathbf{e}_z\mathbf{e}_x)$$

Із цього рівняння, виразу (1.20) і рис. 6.2 випливає відома формула

$$x' = x\cos\alpha - y\sin\alpha . \tag{6.3}$$

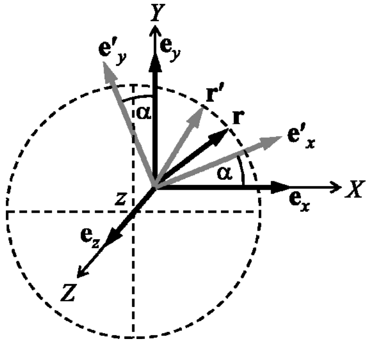

Рис. 6.2. Перетворення радіус-вектора r довільної точки простору та ортів декартової системи координат під дією оператора повороту навколо осі *OZ*

(пунктиром показано, що кінець радіус-вектора рухається під дією цього оператора по колу з центром у точці з координатами (0, 0, z))

Утворивши скалярний добуток обох частин рівності (6.2) на орт $\mathbf{e}_y$, знаходимо рівняння

$$y' = x(\mathbf{e}'_x\mathbf{e}_y) + y(\mathbf{e}'_y\mathbf{e}_y) + z(\mathbf{e}_z\mathbf{e}_y) =$$
$$= x\cos(\pi/2 - \alpha) + y\cos\alpha,$$

та, як наслідок, ще одну відому формулу

$$y' = x\sin\alpha + y\cos\alpha\,. \tag{6.4}$$

Утворивши скалярний добуток обох частин рівності (6.2) на орт $\mathbf{e}_z$, дістаємо очевидну рівність

$$z' = z\,. \tag{6.5}$$

Усі три формули перетворення координат радіус-вектора під дією оператора повороту навколо осі *OZ* можна записати у формі одного матричного рівняння:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{6.6}$$

У правій частині цього рівняння стоїть матриця

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.7}$$

яку називають матрицею повороту на кут α навколо осі OZ. Якщо $\alpha = \pi/2$, матриця повороту набуває форми

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.8}$$

Підставивши цю матрицю до рівняння (6.6), отримуємо

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y \\ x \\ z \end{pmatrix}.$$

Із рис. 6.2 легко побачити, що в тому разі, коли кут повороту α дорівнює $\pi/2$, орти $\mathbf{e}'_x$, $\mathbf{e}'_y$, $\mathbf{e}'_z$ пов'язані з ортами $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ рівняннями

$$\mathbf{e}'_x = \mathbf{e}_y,\ \mathbf{e}'_y = -\mathbf{e}_x,\ \mathbf{e}'_z = \mathbf{e}_z. \tag{6.9}$$

Ці три рівняння еквівалентні матричному рівнянню

$$\begin{pmatrix} \mathbf{e}'_x \\ \mathbf{e}'_y \\ \mathbf{e}'_z \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix}, \tag{6.10}$$

тому що з нього та з правила обчислення добутку матриць випливає рівність

$$\begin{pmatrix} \mathbf{e}'_x \\ \mathbf{e}'_y \\ \mathbf{e}'_z \end{pmatrix} = \begin{pmatrix} \mathbf{e}_y \\ -\mathbf{e}_x \\ \mathbf{e}_z \end{pmatrix},$$

еквівалентна рівностям (6.3) – (6.5). Залучена до правої частини рівняння (6.10) матриця третього порядку

$$A^T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.11}$$

транспонована до матриці оператора повороту (6.8).●

Узагальнюючи наведений приклад, покажемо, що *в ортонормованому* базисі $\{\mathbf{e}_i\}_{i=\overline{1,n}}$ лінійному оператору $\hat{L}$ відповідає матриця $A(a_{ik})$ з *матричними елементами*, що дорівнюють скалярним добуткам ортів $\mathbf{e}_i$ на вектори $\mathbf{e}'_k \equiv \hat{L}\mathbf{e}_k$:

$$a_{ik} = \mathbf{e}_i \hat{L} \mathbf{e}_k . \tag{6.12}$$

З цією метою розглянемо дію оператора $\hat{L}$ на вектор $\mathbf{x}$, що належить лінійному простору $\mathcal{L}$: $\hat{L}\mathbf{x} = \mathbf{y}$. Знайдемо координату y_i вектора $\mathbf{y}$, урахувавши твердження 3.4, сформульоване для координат вектора в ортонормованому базисі:

$$y_i = \mathbf{e}_i \mathbf{y} = \mathbf{e}_i \hat{L} \mathbf{x} . \tag{6.13}$$

Розкладемо вектор $\mathbf{x}$ у базисі $\{\mathbf{e}_i\}_{i=\overline{1,n}}$:

$$\mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{e}_k . \tag{6.14}$$

Підставимо розклад (6.14) до виразу (6.13):

$$y_i = \mathbf{e}_i \hat{L} \sum_{k=1}^{n} x_k \mathbf{e}_k .$$

Візьмемо до уваги властивості лінійного оператора I) та II) з означення 6.3:

$$y_i = \mathbf{e}_i \sum_{k=1}^{n} \hat{L} x_k \mathbf{e}_k = \mathbf{e}_i \sum_{k=1}^{n} x_k \hat{L} \mathbf{e}_k .$$

Зважимо на дистрибутивність скалярного добутку та його асоціативність із числом:

$$y_i = \sum_{k=1}^{n} \mathbf{e}_i x_k \hat{L} \mathbf{e}_k = \sum_{k=1}^{n} x_k \mathbf{e}_i \hat{L} \mathbf{e}_k .$$

Зваживши на позначення (6.12), виразимо правило перетворення координат вектора під дією лінійного оператора у такій формі:

$$y_i = \sum_{k=1}^{n} a_{ik} x_k . \tag{6.15}$$

Оскільки кожне фіксоване значення *першого* індексу матричного елементу відповідає одному з рядків матриці $A(a_{ik})$, то правило (6.15) зручно запам'ятати у формі простого твердження.

Твердження 6.1. Координати вектора перетворюються за допомогою рядків матриці лінійного оператора.

Легко впевнитися у справедливості ще одного важливого твердження.

Твердження 6.2. Базисні орти перетворюються за допомогою стовпців матриці лінійного оператора.

Щоб упевнитися у справедливості цього твердження, розкладемо по базисних ортах вектор $\mathbf{e}_i' \equiv \hat{L}\mathbf{e}_i$:

$$\mathbf{e}_i' = \sum_{k=1}^{n} \zeta_k \mathbf{e}_k \ . \tag{6.16}$$

Узявши до уваги твердження 3.4 та позначення (6.12), виразимо координати ζ_k вектора $\mathbf{e}_i'$ через матричні елементи оператора $\hat{L}$:

$$\zeta_k = \mathbf{e}_k \mathbf{e}_i' = \mathbf{e}_k \hat{L} \mathbf{e}_i = a_{ki} \ .$$

Підставивши знайдений вираз до розкладу (6.16), одержимо формулу

$$\mathbf{e}_i' = \sum_{k=1}^{n} a_{ki} \mathbf{e}_k \ , \tag{6.17}$$

яка доводить твердження 6.2, тому що фіксоване значення *другого* індексу матричного елементу відповідає одному стовпцю матриці $A(a_{ki})$.

Зауваження 6.3. Елементи матриці $A(a_{ik})$ називають матричними елементами лінійного оператора в заданому базисі. Якщо базис ортонормований, для відшукання матричних елементів оператора достатньо обчислити скалярні добутки (6.12). У багатьох розділах фізико-математичних і фізико-технічних наук визначну роль відіграють оператори, що діють у просторі неперервних на відрізку $a \le t \le b$ функцій змінної t, які можуть набувати комплексних значень. У такому разі скалярні добутки є визначеними інтегралами, а матричні елементи лінійного оператора слід обчислювати за формулою

$$a_{jk} = \int_a^b \psi_j(t) \hat{L} \psi_k^*(t) dt \ , \tag{6.18}$$

де $\psi_j(t)$ та $\psi_k(t)$ – функції ортонормованого базису.

● **Приклад 6.2.** Знайдемо матрицю оператора диференціювання в просторі функцій

$$x(t) = x_1\psi_1(t) + x_2\psi_2(t),$$

де $\psi_1(t) = \dfrac{1}{\sqrt{\pi}}\sin t$, $\psi_2(t) = \dfrac{1}{\sqrt{\pi}}\cos t$, $0 \le x \le 2\pi$, x_1 та x_2 – числові коефіцієнти. Для цього зауважимо, що сукупність функцій $x(t)$ є лінійною оболонкою двох лінійно незалежних функцій $\psi_1(t)$ та $\psi_2(t)$, тому функції $x(t)$ є елементами двовимірного простору (див. теорему 3.4). Функції $\psi_1(t)$ та $\psi_2(t)$ є ортонормованим базисом у цьому просторі, тому що їхні скалярні добутки задовольняють умови (3.25):

$$\langle\psi_1|\psi_1\rangle = \frac{1}{\pi}\int_0^{2\pi}\sin^2 t dt = 1,\quad \langle\psi_2|\psi_2\rangle = \frac{1}{\pi}\int_0^{2\pi}\cos^2 t dt = 1,$$

$$\langle\psi_1|\psi_2\rangle = \frac{1}{\pi}\int_0^{2\pi}\sin t\cos t dt = 0.$$

Обчислимо матричні елементи оператора $\dfrac{d}{dt}$ у цьому базисі:

$$a_{11} = \int_0^{2\pi}\psi_1(t)\frac{d}{dt}\psi_1(t)dt = \frac{1}{\pi}\int_0^{2\pi}\sin t\cos t dt = 0,$$

$$a_{22} = \int_0^{2\pi}\psi_2(t)\frac{d}{dt}\psi_2(t)dt = -\frac{1}{\pi}\int_0^{2\pi}\cos t\sin t dt = 0,$$

$$a_{12} = \int_0^{2\pi}\psi_1(t)\frac{d}{dt}\psi_2(t)dt = -\frac{1}{\pi}\int_0^{2\pi}\sin^2 t dt = -1,$$

$$a_{21} = \int_0^{2\pi}\psi_2(t)\frac{d}{dt}\psi_1(t)dt = \frac{1}{\pi}\int_0^{2\pi}\cos^2 t dt = 1.$$

Обчислення показали, що оператор диференціювання в базисі $\{\psi_1, \psi_2\}$ має матрицю

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Коефіцієнти x_1 та x_2 є координатами елементу $f(t)$ у базисі $\{\psi_1, \psi_1\}$. Застосуємо до них твердження 6.1 щодо перетворення координат під дією лінійного оператора. З цією метою виразимо рівняння (6.15) у матричній формі

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Виконавши операцію множення матриць, що стоять у правій частині цього рівняння, знаходимо рівняння

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix},$$

яке показує, що під дією оператора диференціювання функція

$$x(t) = \frac{1}{\sqrt{\pi}}(x_1 \sin t + x_2 \cos t)$$

перетворюється на функцію

$$y(t) = \frac{1}{\sqrt{\pi}}(-x_2 \sin t + x_1 \cos t).$$

Цей результат підтверджується загальновідомими правилами диференціювання тригонометричних функцій $\sin t$ та $\cos t$.●

6.2. Сума, добуток та комутатор операторів

Розглянемо оператори $\hat{L}_1$ та $\hat{L}_2$, що діють у лінійному просторі, і *будь-який* вектор цього простору $\mathbf{x}$. Оператор $\hat{L}_1$ перетворює вектор $\mathbf{x}$ на певний вектор $\mathbf{y}$, тобто $\hat{L}_1\mathbf{x} = \mathbf{y}$. Під дією оператора $\hat{L}_2$ вектор $\mathbf{y}$ перетворюється на певний вектор $\mathbf{z}$, тобто $\hat{L}_2\mathbf{y} = \hat{L}_2\hat{L}_1\mathbf{x} = \mathbf{z}$. Це означає, що в тому разі, якщо в лінійному просторі означені оператори $\hat{L}_1$ та $\hat{L}_2$, то існує правило, за яким кожен вектор $\mathbf{x}$, що належить лінійному простору, перетворюється на певний вектор $\mathbf{z}$ того ж простору. Це правило є оператором $\hat{L}$, який називають добутком операторів $\hat{L}_1$ та $\hat{L}_2$. Пишуть

$$\hat{L}\mathbf{x} = \hat{L}_2\hat{L}_1\mathbf{x}\,,\ \hat{L} = \hat{L}_2\hat{L}_1\,.$$

Якщо подіяти операторами $\hat{L}_1$ та $\hat{L}_2$ на *будь-який* вектор $\mathbf{x}$, що належить лінійному простору, він перетвориться на вектори $\mathbf{y}_1$ та $\mathbf{y}_2$, сума яких також буде вектором цього простору $L_1\mathbf{x} + L_2\mathbf{x} = \mathbf{y}_1 + \mathbf{y}_2 = \mathbf{z}$. Тому існує оператор $\hat{L}$, який перетворює вектор $\mathbf{x}$ на вектор $\mathbf{z}$. Цей оператор називають сумою операторів $\hat{L}_1$ та $\hat{L}_2$. Пишуть

$$\hat{L}\mathbf{x} = \hat{L}_1\mathbf{x} + \hat{L}_2\mathbf{x}\,,\ \hat{L} = \hat{L}_1 + \hat{L}_2\,.$$

Аналогічно означають різницю операторів та пишуть:

$$\hat{L}\mathbf{x} = \hat{L}_1\mathbf{x} - \hat{L}_2\mathbf{x}\,,\ \hat{L} = \hat{L}_1 - \hat{L}_2\,.$$

Згідно з твердженням 6.1, кожному оператору відповідає квадратна матриця, якою визначається зміна координат вектора в заданому базисі. У підрозд. 2.2 було зазначено, що операція множення матриць не комутативна, і означено поняття комутатора двох матриць (див. рівняння (2.8)). Звідси випливає, що результат дії на вектор $\mathbf{x}$ оператора $\hat{L}_2\hat{L}_1$ може відрізнятися від результату дії на цей самий вектор оператора $\hat{L}_1\hat{L}_2$, і тоді $\hat{L}_1\hat{L}_2\mathbf{x} - \hat{L}_2\hat{L}_1\mathbf{x} \neq \mathbf{0}$. Тому для правильного проведення обчислень важливе поняття комутатора двох операторів[1].

Означення 6.4. Комутатором операторів $\hat{L}_1$ та $\hat{L}_2$ називають оператор, який є різницею добутків $\hat{L}_1\hat{L}_2$ та $\hat{L}_2\hat{L}_1$:

$$[\hat{L}_1, \hat{L}_2] \overset{def}{\equiv} \hat{L}_1\hat{L}_2 - \hat{L}_2\hat{L}_1\,. \tag{6.19}$$

Залежно від того, якими є оператори $\hat{L}_1$ та $\hat{L}_2$, комутатор $[\hat{L}_1, \hat{L}_2]$ може або дорівнювати, або не дорівнювати нулю

---

[1] Слово *commutator* перекладається з англійської або як *те, що з'єднує,* або як *те, що переставляє.* У лінійній алгебрі використовується друге з цих двох значень слова.

(див. означення 6.2). Якщо комутатор операторів $\hat{L}_1$ та $\hat{L}_2$ дорівнює нулю, то кажуть, що оператор $\hat{L}_1$ комутує з оператором $\hat{L}_2$, і пишуть $[\hat{L}_1, \hat{L}_2] = 0$. В оберненому випадку кажуть, що $\hat{L}_1$ не комутує з $\hat{L}_2$, а при виконанні обчислень користуються операторним рівнянням

$$\hat{L}_1 \hat{L}_2 = \hat{L}_2 \hat{L}_1 + [\hat{L}_1, \hat{L}_2]. \tag{6.20}$$

Зауваження 6.4. Добутку, сумі та комутатору операторів відповідають сума, добуток та комутатор матриць цих операторів.

Зауваження 6.5. Якщо оператор $\hat{L}_1$ комутує з оператором $\hat{L}_2$, то матриця оператора $\hat{L}_1$ комутує з матрицею оператора $\hat{L}_2$.

Слід зауважити, що поняття лінійного оператора та комутатора двох операторів відіграє визначальну роль в аксіоматичній будові квантової механіки. У цій фундаментальній галузі сучасної фізики кожній фізичній величині відповідає лінійний оператор. Якщо якась фізична величина f, що характеризує рухому фізичну систему (імпульс, орбітальний момент тощо) не змінюється в часі (зберігається у процесі руху системи), то оператор цієї величини $\hat{f}$ комутує з оператором енергії $\hat{\mathcal{H}}$[1]. Тому рівняння $[\hat{f}, \hat{\mathcal{H}}] = 0$ є математичним виразом закону збереження фізичної величини f. Принципи застосування лінійної алгебри у квантовій механіці стисло викладено в загальновідомому підручнику [8].

Означення 6.5. Оператор $\hat{L}^{-1}$ називають оберненим до оператора $\hat{L}$, якщо для будь-якого вектора $\mathbf{x}$, що належить простору $\mathcal{L}$, виконуються умови $\hat{L}^{-1}\hat{L}\mathbf{x} = \hat{L}\hat{L}^{-1}\mathbf{x} = \mathbf{x}$, або, що те ж саме, якщо добуток $\hat{L}^{-1}\hat{L}$ є тотожним оператором (див. означення 6.2).

Зауваження 6.6. Оскільки вектор $\mathbf{x}$ не змінюється під дією оператора $\hat{L}^{-1}\hat{L}$, то цьому оператору відповідає одинична матри-

[1] Оператор енергії називають гамільтоніаном механічної системи, тому його позначають першою літерою прізвища видатного ірландського науковця Вільяма Роуена Гамільтона (*William Rowan Hamilton*).

ця. Згідно з означенням 4.10 оператору $\hat{L}^{-1}$ відповідає матриця A^{-1}, обернена до матриці A. Тому матричні елементи оператора $\hat{L}^{-1}$ можна обчислити за допомогою формули (4.39), якщо відомі матричні елементи a_{ij}.

6.3. Власні числа та власні значення лінійного оператора

У ході розв'язання великої кількості інженерних та наукових задач виникає потреба у визначенні такого базису, у якому матриця лінійного оператора має діагональну форму. Щоб з'ясувати, коли і як можна знайти такий базис, розглянемо оператор $\hat{L}$, що діє в *комплексному* просторі Евкліда $\mathcal{L}$.

Означення 6.6. Якщо *ненульовий* вектор $\mathbf{x}$ задовольняє рівняння

$$\hat{L}\mathbf{x} = \lambda\mathbf{x}, \tag{6.21}$$

то цей вектор називають власним вектором оператора $\hat{L}$, а число λ – власним числом цього оператора.

Зауваження 6.7. Оскільки оператор $\hat{L}$ – лінійний, то поряд з кожним власним вектором $\mathbf{x}$ він має безліч власних векторів $C\mathbf{x}$, де C – довільний числовий множник. Це твердження є наслідком властивостей I), II), вказаних в означенні 6.3, і доводиться такими співвідношеннями:

$$\hat{L}C\mathbf{x} = C\hat{L}\mathbf{x} = C\lambda\mathbf{x} = \lambda(C\mathbf{x}).$$

Слід також зауважити, що згідно з теоремою 3.1 сукупність векторів $C\mathbf{x}$ є одновимірним лінійним підпростором простору $\mathcal{L}$, тому що сума таких векторів і добуток такого вектора на число належить до цієї сукупності.

Визначення напрямків всіх власних векторів лінійного оператора є важливим завданням багатьох розділів вищої математики, теоретичної фізики та інших фізико-математичних наук. Щоб виконати це завдання, достатньо знайти координати власних векторів в ортонормованому базисі. Згідно з рівняннями (6.21) координати вектора $\hat{L}\mathbf{x}$ мають дорівнювати відповідним координатам вектора $\lambda\mathbf{x}$, тобто мають справджуватися рівняння

$$\sum_{k=1}^{n} a_{jk} x_k = \lambda x_j \,,\ \ j = \overline{1,n}\,,$$

еквівалентні однорідній системі лінійних алгебраїчних рівнянь

$$\sum_{k=1}^{n} \left(a_{jk} - \lambda \delta_{jk} \right) x_k = 0 \,,\ \ j = \overline{1,n}\,. \tag{6.22}$$

Згідно з наслідком теореми 5.1 така система рівнянь має *нетривіальний* розв'язок тоді і лише тоді, коли визначник її основної матриці дорівнює нулю (див. підрозд. 5.2), тобто, якщо справджується таке алгебраїчне рівняння порядку n для невідомої величини λ:

$$\det(a_{jk} - \lambda \delta_{jk}) = 0\,. \tag{6.23}$$

Означення 6.7. Рівняння (6.23) називають характеристичним рівнянням лінійного оператора.

Згідно з так званою "основною теоремою алгебри" рівняння (6.23) має принаймні один корінь (дійсний або комплексний). Унаслідок цього, система (6.22) має принаймні один розв'язок, а кожен лінійний оператор, що діє в комплексному просторі Евкліда, – принаймні один власний вектор. Якщо кратність усіх коренів рівняння (6.23) дорівнює одиниці, то воно має n *різних* коренів. Можна довести (див. підручник [9], § 10), що в такому разі власні вектори $\left\{ \mathbf{x}^{(p)} \right\}_{p=\overline{1,n}}$ лінійно незалежні. Оскільки їхня кількість дорівнює вимірності того простору, у якому діє оператор, то вони є базисом цього простору. Тоді з рівностей

$$\hat{L} x_j^{(p)} = \lambda_p x_j^{(p)} = \sum_{k=1}^{n} \lambda_p \delta_{jk} x_k^{(p)}$$

і твердження 6.2 випливає, що матричні елементи a_{jk} оператора $\hat{L}$ у базисі $\left\{ \mathbf{x}^{(p)} \right\}_{p=\overline{1,n}}$ дорівнюють числам $\lambda_p \delta_{jk}$.

Висновок 6.1. Одним із шляхів діагоналізації матриці лінійного оператора є визначення його власних чисел і власних векторів.

Власні числа оператора є розв'язками характеристичного рівняння (6.23), тому процедуру діагоналізації слід починати з його розв'язання. Якщо кратність усіх коренів характеристичного рівняння дорівнює одиниці, то *матриця оператора має діагональну форму в базисі, утвореному з власних векторів, і власні числа оператора є діагональними елементами його матриці.* У такому разі для діагоналізації матриці достатньо визначити власні числа оператора з характеристичного рівняння, а якщо треба, визначені з характеристичного рівняння власні числа слід почергово підставляти до лінійної однорідної системи (6.22) та визначати координати власних векторів.

Якщо оператор має менше ніж n лінійно незалежних власних векторів, діагоналізувати його матрицю неможливо. Питання про спрощення форми матриці такого оператора складне, його розгляд можна знайти у розд. III підручника [9].

● **Приклад 6.3.** Нехай матриця

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$$

є матрицею лінійного оператора, що діє у двовимірному дійсному просторі Евкліда. Знайдемо власні числа та власні вектори цього оператора. Для цього розв'яжемо характеристичне рівняння

$$\begin{vmatrix} 3-\lambda & 2 \\ 1 & 2-\lambda \end{vmatrix} = 0 \quad \Rightarrow \quad \lambda^2 - 5\lambda + 4 = 0 \quad \Rightarrow \quad \lambda_1 = 1, \quad \lambda_2 = 4.$$

Розглянемо систему (6.22) з $a_{11} = 3$, $a_{12} = 2$, $a_{21} = 1$, $a_{22} = 2$:

$$\begin{cases} (3-\lambda)x_1 + 2x_2 = 0, \\ x_1 + (2-\lambda)x_2 = 0. \end{cases} \tag{6.24}$$

Підставимо до неї власне число λ_1:

$$\begin{cases} 2x_1^{(1)} + 2x_2^{(1)} = 0, \\ x_1^{(1)} + x_2^{(1)} = 0. \end{cases} \tag{6.25}$$

Рівняння системи (6.25) виявилися лінійно залежними, тому що рівність нулю визначника основної матриці лінійної системи рівнянь є критерієм лінійної залежності цих рівнянь. Розв'язком системи (6.24) є власний вектор $\mathbf{x}^{(1)}$, координати якого пов'язані між собою співвідношенням $x_1^{(1)} = -x_2^{(1)}$, яке фіксує напрямок цього вектора на площині змінних x_1, x_2, але залишає невизначеною його довжину. Тому власному значенню λ_1 відповідають вектори

$$\mathbf{x}^{(1)} = (C_1, -C_1),$$

де C_1 – довільна стала. Підставивши до системи (6.24) значення $\lambda = \lambda_2 = 4$, легко знайти власний вектор

$$\mathbf{x}^{(2)} = (C_2, 2C_2).$$

Скалярний добуток цих векторів дорівнює $-C_1 C_2$, а отже, знайдені власні вектори не ортогональні.●

6.4. Спряжені оператори. Самоспряжений оператор

Існують декілька специфічних різновидів лінійних операторів, які особливо важливі з огляду на їхнє застосування у багатьох розділів математики, фізико-математичних та фізико-технічних наук, і насамперед у квантовій механіці. Одним із таких різновидів є самоспряжені оператори, загальновідомі також як *оператори Ерміта*. Розглянемо ті властивості операторів Ерміта, які зумовлюють їхню виняткову важливість. З цією метою розглянемо матричні елементи $a_{jk} = \mathbf{e}_j \hat{L} \mathbf{e}_k$ оператора $\hat{L}$ та матричні елементи $\tilde{a}_{jk} = \mathbf{e}_j \hat{L}^{+} \mathbf{e}_k$ іншого оператора $\hat{L}^{+}$ в ортонормованому базисі $\{\mathbf{e}_j\}_{j=\overline{1,n}}$.

Означення 6.6. Оператор $\hat{L}^{+}$ називають спряженим з оператором $\hat{L}$, якщо його матричні елементи задовольняють умову

$$\tilde{a}_{jk} = a_{kj}^{*}, \tag{6.26}$$

де зірка позначає комплексне спряження елементу a_{kj}.

Матрицю спряженого оператора позначають A^{+}. Умова (6.24) показує, що для того, щоб з матриці A оператора $\hat{L}$ отримати матрицю A^{+} оператора $\hat{L}^{+}$, треба транспонувати матрицю A та замінити всі її елементи на комплексно-спряжені. Оскільки це можна зробити з будь-якою матрицею, то для кожного оператора існує спряжений з ним оператор.

Зауваження 6.7. Означенням 6.6 зручно користуватися в тому випадку, коли розрахунки виконуються з використанням ортонормованого базису. У багатьох посібниках використовується більш загальне, але менш зручне для розрахунків, означення: оператори $\hat{L}$ та $\hat{L}^{+}$ спряжені один з одним, якщо вони задовольняють рівняння

$$\mathbf{x}\hat{L}^{+}\mathbf{y} = (\mathbf{y}\hat{L}\mathbf{x})^{*}. \tag{6.27}$$

Покажемо, що умова (6.26) випливає з рівняння (6.27) у випадку, коли вектори $\mathbf{x}$ та $\mathbf{y}$ розкладені в ортонормованому базисі $\{\mathbf{e}_j\}_{j=\overline{1,n}}$. У цьому випадку ліва частина рівняння (6.27) набуває форми[1]

$$\mathbf{x}\hat{L}^{+}\mathbf{y} = \sum_{j=1}^{n} x_j \mathbf{e}_j \hat{L}^{+} \sum_{k=1}^{n} y_k^{*} \mathbf{e}_k \ .$$

Урахуємо, що оператор $\hat{L}^{+}$ лінійний, а скалярний добуток векторів – дистрибутивний та асоціативний із числом

$$\mathbf{x}\hat{L}^{+}\mathbf{y} = \sum_{j=1}^{n} x_j \mathbf{e}_j \sum_{k=1}^{n} y_k^{*} \hat{L}^{+} \mathbf{e}_k = \\ = \sum_{j=1}^{n} x_j \sum_{k=1}^{n} y_k^{*} \mathbf{e}_j \hat{L}^{+} \mathbf{e}_k \ .$$

Узявши до уваги означення матричного елементу лінійного оператора (див. (6.12)), виразимо ліву частину рівняння (6.27) через матричні елементи оператора $\hat{L}^{+}$:

$$\mathbf{x}\hat{L}^{+}\mathbf{y} = \sum_{j=1}^{n} x_j \sum_{k=1}^{n} y_k^{*} \tilde{a}_{jk} \ . \tag{6.28}$$

[1] Те, чому в розкладі вектора **y** по ортах базису його координати комплексно спряжені, пояснює зауваження 3.4 у підрозд. 3.3.

Здійснимо тотожні перетворення правої частини рівняння (6.27)[1]:

$$(\mathbf{y}\hat{L}\mathbf{x})^* = \left(\sum_{k=1}^{n} y_k \mathbf{e}_k \hat{L} \sum_{j=1}^{n} x_j^* \mathbf{e}_j\right)^* = \sum_{k=1}^{n} y_k^* \mathbf{e}_k \left(\hat{L} \sum_{j=1}^{n} x_j^* \mathbf{e}_j\right)^* =$$

$$= \sum_{k=1}^{n} y_k^* \sum_{j=1}^{n} x_j (\mathbf{e}_k \hat{L} \mathbf{e}_j)^*$$

та виразимо її через матричні елементи оператора $\hat{L}$:

$$(\mathbf{y}\hat{L}\mathbf{x})^* = \sum_{k=1}^{n} y_k^* \sum_{j=1}^{n} x_j a_{kj}^* . \qquad (6.29)$$

Із порівняння виразів (6.28) та (6.29) випливає, що рівняння (6.27) справджується лише за умови (6.26).

Означення 6.7. Оператор $\hat{L}$ називають самоспряженим оператором, або оператором Ерміта, якщо

$$L^+ = L . \qquad (6.30)$$

Як показують вирази (6.27) та (6.30), оператор Ерміта задовольняє рівняння

$$\mathbf{x}\hat{L}\mathbf{y} = (\mathbf{y}\hat{L}\mathbf{x})^* . \qquad (6.31)$$

Якщо оператори $\hat{L}$ та L^+ дорівнюють один одному, то в заданому базисі їм відповідають однакові матриці. Тому "ознакою ермітовості" лінійного оператора є рівність

$$A = A^+ . \qquad (6.32)$$

Із виразів (6.26) та (6.32) випливає, що матричні елементи оператора Ерміта в ортонормованому базисі задовольняють умови

$$a_{jk} = a_{kj}^* . \qquad (6.33)$$

Зауваження 6.8. Якщо оператор $\hat{L}$ діє не в комплексному, а в дійсному просторі, то його матричні елементи є дійсними числами й умови (6.32) та (6.33) перетворюються на умови

$$A = A^T , \quad a_{jk} = a_{kj} . \qquad (6.34)$$

Матриці, що задовольняють умову (6.32), називають симетричними відносно головної діагоналі. Множина дійсних

---

[1] Результат додавання векторів не залежить від того, якою літерою позначений індекс підсумовування.

чисел є підмножиною комплексних, а отже, якщо оператор діє в дійсному просторі, і його матриця симетрична, то цей оператор є окремим випадком оператора Ерміта.

Зазначимо ті властивості операторів Ерміта, які зумовлюють їхнє застосування у фізико-математичних науках.

I. Діагональні матричні елементи оператора Ерміта є дійсними числами.

II. Власні вектори оператора Ерміта є ортогональним базисом того простору, у якому діє оператор.

III. Усі власні числа оператора Ерміта є дійсними числами.

Перша з цих властивостей безпосередньо випливає з умов (6.33); доведення другої властивості можна знайти в § 16 підручника [9], а третя властивість є наслідком першої, тому що в базисі з власних векторів матриця оператора Ерміта має діагональну форму, її діагональні елементи є власними числами оператора (див. висновок 6.1). Проілюструємо властивості I – III конкретним прикладом.

● **Приклад 6.4.** Знайдемо власні числа та власні вектори оператора Ерміта, який діє у двовимірному комплексному просторі Евкліда. У загальному випадку такому оператору відповідає матриця

$$A = \begin{pmatrix} a_{11} & \alpha + i\beta \\ \alpha - i\beta & a_{22} \end{pmatrix},$$

де a_{11}, a_{22}, α та β є *дійсними* числами. Система рівнянь для визначення координат власних векторів має форму

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 = 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 = 0. \end{cases} \tag{6.35}$$

Із цієї системи знаходимо характеристичне рівняння

$$\begin{vmatrix} a_{11} - \lambda & \alpha + i\beta \\ \alpha - i\beta & a_{22} - \lambda \end{vmatrix} = 0,$$

яке є алгебраїчним рівнянням другого порядку

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - \alpha^2 - \beta^2 = 0\,.$$

За допомогою загальновідомої формули для розв'язків такого рівняння знаходимо загальну формулу для власних чисел оператора Ерміта, що діє у двовимірному просторі:

$$\lambda_{1,2} = \frac{1}{2}\left[(a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + 4(\alpha^2 + \beta^2)} \right]. \tag{6.36}$$

Оскільки підрадикальний вираз у формулі (6.36) не може бути від'ємним ні за яких значень чисел a_{11}, a_{22}, α та β, то формула (6.36) підтверджує властивість III, згідно з якою власні числа оператора Ерміта не можуть бути комплексними.

Щоб уникнути громіздких формул, знайдемо координати власних векторів оператора для конкретних значень параметрів, що входять до матриці A:

$$a_{11} = 1,\ a_{22} = 3,\ \alpha = \sqrt{2},\ \beta = 1.$$

Підставивши ці значення до формули (6.36), знаходимо відповідні значення власних чисел:

$$\lambda_1 = 0,\ \lambda_2 = 4.$$

Для першого власного числа $\lambda = \lambda_1$ система рівнянь (6.35) перетворюється на еквівалентні одне одному рівняння

$$\begin{cases} x_1 + (\sqrt{2} + i)x_2 = 0, \\ (\sqrt{2} - i)x_1 + 3x_2 = 0. \end{cases} \tag{6.37}$$

Якщо домножити перше з рівнянь системи (6.37) на $(\sqrt{2} - i)$, воно перетвориться на друге, а отже, рівняння цієї системи лінійно залежні. Лінійна залежність рівнянь є наслідком того, що головний визначник системи дорівнює нулю. Нагадаємо, що лінійна залежність рядків є критерієм рівності визначника нулю (див. п. 4.5.3). Оскільки рівняння системи (6.37) еквівалентні, то вони встановлюють лише співвідношення координат першого власного вектора $x_1^{(1)} = -(\sqrt{2} + i)x_2^{(1)}$, тому одна з координат залишається невизначеною і може набувати будь-яких значень. Поклавши $x_2^{(1)} = -C_1$, де C_1 – довільна стала, знаходимо перший власний вектор оператора Ерміта:

$$\mathbf{x}^{(1)} = (x_1^{(1)},\ x_2^{(1)}) = C_1(\sqrt{2} + i,\ -1). \tag{6.38}$$

Для другого власного числа $\lambda = \lambda_2$ система рівнянь (6.35) перетворюється на рівняння

$$\begin{cases} -3x_1 + (\sqrt{2} + i)x_2 = 0, \\ (\sqrt{2} - i)x_1 - x_2 = 0. \end{cases}$$

Із цих рівнянь легко знайти другий власний вектор

$$\mathbf{x}^{(2)} = (x_1^{(2)}\,, x_2^{(2)}) = C_2(1\,, \sqrt{2} - i)\,, \tag{6.39}$$

де C_2 є довільною сталою. З виразів (6.38) та (6.39) випливає, що власні вектори утворюють ортогональний базис двовимірного простору, тому що скалярний добуток цих власних векторів дорівнює нулю:

$$\mathbf{x}^{(1)}\mathbf{x}^{(2)} = x_1^{(1)}x_1^{(2)*} + x_2^{(1)}x_2^{(2)*} = 0\,.$$

Обчислимо модулі цих векторів:

$$\sqrt{\mathbf{x}^{(1)}\mathbf{x}^{(1)}} = \sqrt{x_1^{(1)}x_1^{(1)*} + x_2^{(1)}x_2^{(1)*}} = 2\,|\,C_1\,|\,,$$

$$\sqrt{\mathbf{x}^{(2)}\mathbf{x}^{(2)}} = \sqrt{x_1^{(2)}x_1^{(2)*} + x_2^{(2)}x_2^{(2)*}} = 2\,|\,C_2\,|\,.$$

Зафіксувавши значення довільних сталих $C_1 = C_2 = 1/2$, знаходимо ортонормований базис того двовимірного простору, у якому діє оператор Ерміта:

$$\mathbf{e}^{(1)} = \frac{1}{2}(\sqrt{2} + i\,, -1)\,,\ \mathbf{e}^{(2)} = \frac{1}{2}(1\,, \sqrt{2} - i)\,.\bullet$$

Слід зауважити, що саме оператори Ерміта відіграють визначальну роль в аксіоматичній будові квантової механіки, тому що, як було вже зазначено, у цій галузі сучасної фізики кожній фізичній величині, що характеризує рухому механічну систему (енергії, імпульсу, орбітальному моменту тощо) відповідає лінійний оператор. Тепер можна додати, що оператори фізичних величин є операторами Ерміта. Справа в тому, що власні числа лінійного оператора квантово-механічної величини є тими значеннями цієї величини, які можуть бути виміряні експериментально. Оскільки результати фізичних вимірювань механічних величин описуються дійсними числами, то власні числа квантово-механічних операторів обов'язково мають бути дійсними, а самі оператори – операторами Ерміта. Власні вектори операторів фізичних величин також мають суттєве значення для квантової механіки, але пояснення цього факту досить складне і викладено навіть не в усіх підручниках з квантової механіки. Одне з найкращих пояснень цього факту можна знайти в підручнику [8].

6.5. Унітарний оператор. Ортогональний оператор

Продовжимо розгляд операторів, що діють у комплексних просторах Евкліда. При розв'язуванні великої кількості задач і вивченні важливих природних явищ застосовуються оператори, які не змінюють величину скалярного добутку векторів цього простору. Прикладом такого оператора є розглянутий у підрозд. 6.1 оператор повороту геометричних векторів навколо фіксованої осі.

Означення 6.8. Оператор $\hat{L}$ називають унітарним оператором, якщо він не змінює величину скалярного добутку будь-яких двох елементів, що належать комплексному простору Евкліда.

Доведемо, що матриця унітарного оператора $\hat{L}$ задовольняє умови

$$A^{+}A = E\,, \tag{6.40}$$

де E – одинична матриця. Для цього розглянемо скалярний добуток $\mathbf{xy}$ двох векторів, що належать комплексному простору Евкліда. Подіємо на ці вектори унітарним оператором $\hat{L}$:

$$\mathbf{x}' = \hat{L}\mathbf{x}\,, \quad \mathbf{y}' = \hat{L}\mathbf{y}\,.$$

Оскільки унітарний оператор не змінює величину скалярного добутку, то справджується "умова унітарності" оператора $\hat{L}$:

$$\mathbf{x}'\mathbf{y}' = \mathbf{xy}\,. \tag{6.41}$$

Якщо виразити скалярні добутки через координати векторів, ця умова набуде форми

$$\sum_{i=1}^{n} x_i' y_i'^{*} = \sum_{k=1}^{n} x_k y_k^{*}\,. \tag{6.42}$$

Тут знову враховано, що у скалярному добутку $\mathbf{xy}$ векторів комплексного простору координати другого співмножника комплексно-спряжені з координатами вектора $\mathbf{y}$. Урахувавши формули перетворення координат векторів $\mathbf{x}$ та $\mathbf{y}$ під дією лінійного оператора

$$y_i'^{*} = \sum_{j=1}^{n} a_{ij}^{*} y_j^{*}\,, \quad x_i' = \sum_{k=1}^{n} a_{ik} x_k\,,$$

виразимо рівняння (6.42) у формі

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} a_{ik} a_{ij}^{*} x_k y_j^{*} = \sum_{k=1}^{n} x_k y_k^{*} . \tag{6.43}$$

Узявши до уваги тотожність

$$\sum_{k=1}^{n} x_k y_k^{*} = \sum_{k=1}^{n}\sum_{j=1}^{n} x_k y_j^{*} \delta_{jk} \tag{6.44}$$

і змінивши послідовність обчислення сум, із виразів (6.43) та (6.44) легко отримати рівність

$$\sum_{k=1}^{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ik} a_{ij}^{*} - \delta_{jk}\right) x_k y_j^{*} = 0 ,$$

яка має виконуватись для координат x_k , y_j будь-яких двох векторів простору. Звідси випливає, що має виконуватись рівність

$$\sum_{i=1}^{n} a_{ik} a_{ij}^{*} - \delta_{jk} = 0 . \tag{6.45}$$

Узявши до уваги правило обчислення добутку матриць (2.5), доходимо висновку, що сума, яка стоїть у лівій частині рівності (6.45), дорівнює елементу $\left\langle A^{+} A \right\rangle_{jk}$ добутку матриці A^{+} на матрицю A, тобто

$$\sum_{i=1}^{n} a_{ij}^{*} a_{ik} = \left\langle A^{+} A \right\rangle_{jk} . \tag{6.46}$$

Оскільки символ Кронекера δ_{jk} позначає елементи одиничної матриці $\left\langle E \right\rangle_{jk}$, то з рівностей (6.45) та (6.46) випливає матричне рівняння

$$A^{+} A - E = 0 ,$$

яке є наслідком умови унітарності (6.41). Це рівняння тотожне з рівнянням (6.40), яке слід було довести і яке є корисною "ознакою унітарності" оператора $\hat{L}$ для проведення багатьох розрахунків.

З означення оберненої матриці (4.37) випливає, що рівняння (6.40) тотожне рівнянню

$$A^{+} = A^{-1} . \tag{6.47}$$

Означення 6.9. Оператор $\hat{L}$ називають ортогональним оператором, якщо він не змінює величину скалярного добутку будь-яких двох елементів, що належать дійсному простору Евкліда.

Із рівнянь (6.40), (6.47) безпосередньо випливає, що матриця ортогонального оператора задовольняє умови

$$A^T A = E\,,\ A^T = A^{-1}\,. \tag{6.48}$$

Найвідомішим прикладом ортогонального оператора є оператор повороту навколо фіксованої осі в тривимірному просторі. Матриця (6.7) оператора повороту навколо осі OZ задовольняє умови (6.48). Щоб впевнитися у цьому, достатньо взяти до уваги, що умова $A^T A = E$ тотожна умові $A^T = A^{-1}$; оператором, оберненим до оператора повороту на кут α є оператор повороту на кут $-\alpha$; матриця (6.7) перетворюється на транспоновану матрицю при заміні α на $-\alpha$.

6.6. Перетворення матриці оператора в результаті лінійного перетворення базису

Як вже було зазначено в підрозд. 6.1, при розв'язуванні великої кількості інженерних і наукових задач виникає потреба у визначенні такого базису, у якому матриця лінійного оператора $\hat{L}$ має найпростішу з усіх можливих форму. У багатьох випадках визначення оптимального для розв'язання задачі базису здійснюється шляхом лінійного перетворення "вихідного" базису $\{\mathbf{e}_j\}_{j=\overline{1,n}}$ на "оптимальний" базис $\{\tilde{\mathbf{e}}_j\}_{j=\overline{1,n}}$ за формулою

$$\tilde{\mathbf{e}}_j = \sum_{k=1}^{n} c_{kj}\mathbf{e}_k\,. \tag{6.49}$$

Ця формула аналогічна формулі (6.17), за якою перетворюються базисні орти під дією лінійного оператора, тому числові коефіцієнти c_{kj} можна вважати матричними елементами лінійного оператора $\hat{T}$, який перетворює базис $\{\mathbf{e}_j\}_{j=\overline{1,n}}$ на базис $\{\tilde{\mathbf{e}}_j\}_{j=\overline{1,n}}$. Дія цього оператора на орт $\mathbf{e}_j$ означається виразом

$$\tilde{\mathbf{e}}_j = \hat{T}\mathbf{e}_j . \tag{6.50}$$

З'ясуємо, на яку матрицю перетворюється матриця A лінійного оператора $\hat{L}$, визначена в базисі $\{\mathbf{e}_j\}_{j=\overline{1,n}}$, у результаті лінійного перетворення цього базису. З цією метою подіємо на орт $\mathbf{e}_j$ добутком трьох лінійних операторів $\hat{T}^{-1}\hat{L}\hat{T}$. У результаті отримаємо вектор

$$\mathbf{x}_j \equiv \hat{T}^{-1}\hat{L}\hat{T}\mathbf{e}_j . \tag{6.51}$$

Зваживши на рівняння (6.50), отримаємо вираз

$$\mathbf{x}_j = \hat{T}^{-1}\hat{L}\tilde{\mathbf{e}}_j . \tag{6.52}$$

Оператор $\hat{L}$ діє на орти $\tilde{\mathbf{e}}_j$ за допомогою своїх матричних елементів у цьому базисі (див. рівняння (6.15)). Позначивши ці елементи як $\tilde{a}_{kj}$, виразимо дію оператора на орти $\tilde{\mathbf{e}}_j$ рівнянням

$$\hat{L}\tilde{\mathbf{e}}_j \equiv \sum_{k=1}^{n} \tilde{a}_{kj}\tilde{\mathbf{e}}_k .$$

Підставимо праву частину цього рівняння до правої частини рівняння (6.52), врахуємо, що оператор $\hat{T}^{-1}$ є лінійним оператором, і в результаті отримаємо вираз

$$\mathbf{x}_j = \sum_{k=1}^{n} \tilde{a}_{kj}\hat{T}^{-1}\tilde{\mathbf{e}}_k . \tag{6.53}$$

Оскільки оператор $\hat{T}$ перетворює $\mathbf{e}_k$ на $\tilde{\mathbf{e}}_k$, то обернений оператор $\hat{T}^{-1}$ перетворює $\tilde{\mathbf{e}}_k$ на $\mathbf{e}_k$, тому з виразу (6.53) випливає подібне до формули (6.17) рівняння

$$\mathbf{x}_j = \sum_{k=1}^{n} \tilde{a}_{kj}\mathbf{e}_k . \tag{6.54}$$

Вирази (6.51) та (6.54) показують, що матричні елементи оператора $\hat{L}$ в оптимальному базисі водночас є матричними елементами оператора $\hat{T}^{-1}\hat{L}\hat{T}$ у вихідному базисі, а отже, матриця

$\tilde{A}(\tilde{a}_{jk})$ оператора $\hat{L}$ в оптимальному базисі $\{\tilde{\mathbf{e}}_j\}_{j=\overline{1,n}}$ пов'язана з матрицею A цього ж оператора у вихідному базисі $\{\mathbf{e}_j\}_{j=\overline{1,n}}$ співвідношенням

$$\tilde{A} = C^{-1}AC\,, \tag{6.55}$$

де $C(c_{kj})$ є матрицею оператора $\hat{T}$. Ця матриця складається з коефіцієнтів лінійного перетворення (6.49) і визначає взаємозв'язок між матрицями оператора $\hat{L}$ у двох лінійно пов'язаних між собою базисах.

РОЗДІЛ 7
Лінійні, білінійні та квадратичні форми

7.1. Лінійні функції векторного аргументу

Розглянемо лінійний простір вимірності n та базис цього простору $\{\mathbf{e}_j\}_{j=\overline{1,n}}$.

Означення 7.1. Правило, за яким кожному вектору $\mathbf{x}$, що належить лінійному простору, поставлено у відповідність число $f(\mathbf{x})$, називають функцією векторного аргументу.

Означення 7.2. Функцію векторного аргументу називають лінійною, якщо вона має дві властивості, подібні до властивостей лінійного оператора, а саме:

III) $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$;

IV) $f(\lambda\mathbf{x}) = \lambda f(\mathbf{x})$.

Щоб поширити на лінійні функції метод координат, розкладемо вектор $\mathbf{x}$ по векторах базису:

$$\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j .$$

Виразимо лінійну функцію через координати її векторного аргументу, узявши до уваги спочатку властивість I), а потім властивість II):

$$f(\mathbf{x}) = f\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{n} f(x_j \mathbf{e}_j) = \sum_{j=1}^{n} x_j f(\mathbf{e}_j) .$$

Значення функції, що відповідають базисним векторам, є числами, які зручно позначити літерами, пронумерованими відповідно до нумерації ортів

$$a_j \overset{def}{\equiv} f(\mathbf{e}_j) , \tag{7.1}$$

та отримати *лінійну форму* функції $f(\mathbf{x})$:

$$f(\mathbf{x}) = \sum_{j=1}^{n} a_j x_j \ . \tag{7.2}$$

Оскільки коефіцієнти a_j пов'язані з базисними векторами, то вони певним чином змінюються при лінійному перетворенні базису, означеного формулою (6.49). Щоб з'ясувати, як саме, розкладемо вектор $\mathbf{x}$ по векторах перетвореного базису $\{\tilde{\mathbf{e}}_j\}_{j=\overline{1,n}}$:

$$\mathbf{x} = \sum_{j=1}^{n} \tilde{x}_j \tilde{\mathbf{e}}_j$$

та отримаємо вираз

$$f(\mathbf{x}) = \sum_{j=1}^{n} \tilde{x}_j f(\tilde{\mathbf{e}}_j)\,,$$

де $f(\tilde{\mathbf{e}}_j) \equiv \tilde{a}_j$ є коефіцієнтами лінійної форми

$$f(\mathbf{x}) = \sum_{j=1}^{n} \tilde{a}_j \tilde{x}_j$$

у перетвореному базисі. Узявши до уваги формулу лінійного перетворення (6.49), знаходимо співвідношення

$$\tilde{a}_j = f(\tilde{\mathbf{e}}_j) = f\left(\sum_{k=1}^{n} c_{kj} \mathbf{e}_k \right) = \sum_{k=1}^{n} c_{kj} f(\mathbf{e}_k)\,,$$

з яких випливає, що внаслідок лінійного перетворення базису коефіцієнти лінійної форми перетворюються так, як базисні орти, тобто

$$\tilde{a}_j = \sum_{k=1}^{n} c_{kj} a_k \ . \tag{7.3}$$

7.2. Білінійні функції векторного аргументу

Означення 7.3. Правило, за яким кожній парі векторів $\mathbf{x}, \mathbf{y}$, що належить лінійному простору $\mathcal{L}$, поставлено у відповідність певне число $f(\mathbf{x}, \mathbf{y})$, називають функцією двох векторних аргументів.

Означення 7.4. Функцію двох векторних аргументів називають білінійною, якщо вона лінійна за кожним зі своїх аргументів, тобто має такі властивості:

I) $f(\mathbf{x}_1+\mathbf{x}_2,\mathbf{y})=f(\mathbf{x}_1,\mathbf{y})+f(\mathbf{x}_2,\mathbf{y})$;

II) $f(\lambda\mathbf{x},\mathbf{y})=\lambda f(\mathbf{x},\mathbf{y})$;

III) $f(\mathbf{x},\mathbf{y}_1+\mathbf{y}_2)=f(\mathbf{x},\mathbf{y}_1)+f(\mathbf{x},\mathbf{y}_2)$;

IV) $f(\mathbf{x},\lambda\mathbf{y})=\lambda f(\mathbf{x},\mathbf{y})$.

Щоб виразити білінійну функцію $f(\mathbf{x},\mathbf{y})$ через координати векторів $\mathbf{x}$ та $\mathbf{y}$, розкладемо ці вектори в базисі $\{\mathbf{e}_j\}_{j=\overline{1,n}}$:

$$\mathbf{x}=\sum_{j=1}^{n}x_j\mathbf{e}_j\,,\quad \mathbf{y}=\sum_{k=1}^{n}y_k\mathbf{e}_k\,.$$

Узявши до уваги спочатку властивості I), II), а потім властивості III), IV), знаходимо

$$f(\mathbf{x},\mathbf{y})=f\left(\sum_{j=1}^{n}x_j\mathbf{e}_j\sum_{k=1}^{n}y_k\mathbf{e}_k\right)=\sum_{j=1}^{n}x_j\sum_{k=1}^{n}y_kf\left(\mathbf{e}_j,\mathbf{e}_k\right).$$

Урахувавши, що множник x_j однаковий для всіх доданків суми, що стоїть праворуч від нього, внесемо його під знак суми

$$f(\mathbf{x},\mathbf{y})=\sum_{j=1}^{n}\sum_{k=1}^{n}x_jy_kf\left(\mathbf{e}_j,\mathbf{e}_k\right).$$

Увівши позначення

$$a_{jk}\overset{def}{\equiv}f(\mathbf{e}_j,\mathbf{e}_k)\,, \tag{7.4}$$

отримаємо *білінійну форму* функції $f(\mathbf{x},\mathbf{y})$:

$$f(\mathbf{x},\mathbf{y})=\sum_{j=1}^{n}\sum_{k=1}^{n}a_{jk}x_jy_k\,. \tag{7.5}$$

Коефіцієнти a_{jk} утворюють квадратну матрицю.

Означення 7.5. Квадратну матрицю

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix},$$

називають матрицею білінійної форми.

Елементи матриці білінійної форми є тими значеннями функції, які відповідають базисним ортам, а отже, вони змінюються при лінійному перетворенні базису $\{\mathbf{e}_j\}_{j=\overline{1,n}}$ на базис $\{\tilde{\mathbf{e}}_j\}_{j=\overline{1,n}}$ (див. рівняння (6.49)). Унаслідок такого перетворення базису матриця A перетворюється на матрицю $\tilde{A}$ за формулою

$$\tilde{A} = C^T A C\,, \tag{7.6}$$

де C – матриця, елементами якої є коефіцієнти лінійного перетворення (6.49). Співвідношення (7.6) називають перетворенням подібності для матриці A. Воно доводиться виконанням алгебраїчних дій, подібних до тих, що були здійснені для отримання виразу (7.3):

$$\tilde{a}_{jk} \equiv f(\tilde{\mathbf{e}}_j, \tilde{\mathbf{e}}_k) = f\left(\sum_{i=1}^{n} c_{ij}\mathbf{e}_i, \sum_{m=1}^{n} c_{mk}\mathbf{e}_m\right) = \sum_{i=1}^{n}\sum_{m=1}^{n} c_{ij}c_{mk} f(\mathbf{e}_i, \mathbf{e}_m) =$$

$$= \sum_{i=1}^{n}\sum_{m=1}^{n} c_{ij}c_{mk}a_{im} = \sum_{i=1}^{n} c_{ij}\sum_{m=1}^{n} a_{im}c_{mk} = \sum_{i=1}^{n} c_{ij}\left\langle AC\right\rangle_{ik} =$$

$$= \left\langle C^T\right\rangle_{ji}\left\langle AC\right\rangle_{ik} = \left\langle C^T A C\right\rangle_{jk}\,,$$

де використані такі ж позначення для елементів добутків матриць, як у підрозд. 2.2.

Зауваження 7.1. Подальший розгляд білінійних форм буде звужений до випадку дійсних просторів Евкліда. Якщо $\mathbf{x}$ та $\mathbf{y}$ є елементами простору Евкліда, виникає можливість вважати матрицю білінійної форми матрицею лінійного оператора $\hat{L}$, що діє в цьому просторі. Ця можливість основана на тому, що для простору Евкліда означено правило обчислення скалярного добутку векторів, згідно з яким можна обчислити скалярний добуток вектора $\mathbf{x}$ та вектора $\hat{L}\mathbf{y}$, на який перетворюється вектор $\mathbf{y}$ під дією оператора $\hat{L}$. Розклавши вектор $\mathbf{y}$ по векторах ортонормованого базису $\{\mathbf{e}_j\}_{j=\overline{1,n}}$ і зваживши на властивості лінійного оператора (див. означення 6.3), виразимо скалярний добуток $\mathbf{x}\hat{L}\mathbf{y}$ через координати векторів:

$$\mathbf{x}\hat{L}\mathbf{y} = \sum_{j=1}^{n} x_j\mathbf{e}_j \hat{L}\sum_{k=1}^{n} y_k\mathbf{e}_k = \sum_{j=1}^{n} x_j\mathbf{e}_j \sum_{k=1}^{n} y_k\hat{L}\mathbf{e}_k = \sum_{j=1}^{n}\sum_{k=1}^{n} y_k x_j \mathbf{e}_j\hat{L}\mathbf{e}_k\ .$$

Праворуч від останнього знака рівності під знаками сум стоять скалярні добутки $\mathbf{e}_j \hat{L} \mathbf{e}_k$. Згідно з виразом (6.12) ці скалярні добутки є матричними елементами оператора $\hat{L}$ у базисі $\{\mathbf{e}_j\}_{j=\overline{1,n}}$.

З цього випливає білінійна форма скалярного добутку $\mathbf{x}\hat{L}\mathbf{y}$:

$$\mathbf{x}\hat{L}\mathbf{y} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j y_k \,. \tag{7.7}$$

Аналогічність виразів (7.5) та (7.7) дозволяє вважати матрицю, що входить до білінійної форми функції $f(\mathbf{x},\mathbf{y})$*, матрицею лінійного оператора.*

7.3. Квадратичні форми

7.3.1. Квадратична форма та її канонічний базис

Розглянемо *симетричну* білінійну функцію $f(\mathbf{x},\mathbf{y})$, яка задовольняє умову $f(\mathbf{x},\mathbf{y}) = f(\mathbf{y},\mathbf{x})$. Матриця білінійної форми такої функції задовольняє умову $A = A^T$, а її коефіцієнти – умови

$$a_{jk} = a_{kj} \,. \tag{7.8}$$

Якщо у симетричній білінійній формі (7.5) покласти $\mathbf{x} = \mathbf{y}$, то отримаємо вираз

$$f(\mathbf{x},\mathbf{x}) = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j x_k \,, \tag{7.9}$$

який називають *квадратичною формою* функції $f(\mathbf{x},\mathbf{x})$.

Зауваження 7.2. Згідно із загально визнаною домовленістю кожна квадратична форма має відповідати одній-єдиній функції $f(\mathbf{x},\mathbf{x})$. З цієї причини подвійну суму

$$\sum_{j=1}^{n} \sum_{k=1}^{n} b_{jk} x_j x_k \,,$$

у якій $b_{jk} \neq b_{kj}$, квадратичною формою не вважають, тому що будь-яка матриця $B(b_{ik})$ тотожно дорівнює сумі симетричної матриці $A = (B + B^T)/2$ та антисиметричної матриці

$D = (B - B^T)/2$, і антисиметрична матриця не дає внеску до квадратичної форми. Цей факт випливає з того, що елементи матриці D задовольняють умови $d_{jk} = -d_{kj}$. Доведемо, що за цих умов справджується рівність

$$\sum_{j=1}^{n}\sum_{k=1}^{n} d_{jk} x_j x_k = 0\,. \qquad (7.10)$$

Для доведення слід взяти до уваги, що взаємозаміна індексів $i \leftrightarrow k$ не змінює величину подвійної суми, тому що ці індекси набувають значень від одиниці до n незалежно один від одного, тому справджується рівність

$$\sum_{j=1}^{n}\sum_{k=1}^{n} d_{jk} x_j x_k = \sum_{k=1}^{n}\sum_{j=1}^{n} d_{kj} x_k x_j\,.$$

Зважимо, що права частина цієї рівності не зміниться, якщо переставити співмножники в добутках $x_k x_j$ і змінити порядок сум:

$$\sum_{j=1}^{n}\sum_{k=1}^{n} d_{jk} x_j x_k = \sum_{j=1}^{n}\sum_{k=1}^{n} d_{kj} x_j x_k\,.$$

Урахувавши, що $d_{ik} = -d_{ki}$, одержимо рівність

$$\sum_{j=1}^{n}\sum_{k=1}^{n} d_{jk} x_j x_k = -\sum_{j=1}^{n}\sum_{k=1}^{n} d_{jk} x_j x_k\,,$$

яка доводить, що ліва частина рівняння (7.10) не змінюється при зміні її знака, а отже, вона дорівнює нулю.

● **Приклад 7.1.** Розглянемо квадратичну форму в тривимірному лінійному просторі:

$$f(\mathbf{x},\mathbf{x}) = x_1^2 - 3x_2^2 + 5x_3^2 + 2x_1x_2 + 4x_1x_3\,. \qquad (7.11)$$

Їй відповідає матриця

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -3 & 0 \\ 2 & 0 & 5 \end{pmatrix}.$$

Слід звернути увагу на те, що недіагональні елементи матриці вдвічі менші від коефіцієнтів, з якими входять до виразу (7.11) добутки координат x_1x_2 та x_1x_3, тому що $2x_1x_2 = x_1x_2 + x_2x_1$ та $4x_1x_3 = 2x_1x_3 + 2x_3x_1$.●

При розгляді багатьох фундаментальних питань математики, фізики та деяких інших галузей науки, а також при розв'язуванні багатьох практично важливих задач, виникає потреба у визначенні такого базису, у якому матриця $A(a_{jk})$ є діагональною матрицею, тобто $a_{jk} = \lambda_k \delta_{jk}$, де λ_k – дійсні числа (див. зауваження 7.1). Такий базис називають *канонічним*. У канонічному базисі квадратична форма є такою сумою:

$$f(\mathbf{x}, \mathbf{x}) = \sum_{k=1}^{n} \lambda_k \tilde{x}_k^2 , \tag{7.12}$$

де $\tilde{x}_k$ є координатами вектора $\mathbf{x}$ у канонічному базисі. З огляду на це, процедуру діагоналізації матриці квадратичної форми називають *зведенням квадратичної форми до суми квадратів*, маючи на увазі квадрати координат вектора $\mathbf{x}$.

Завдяки тому, що матрицю квадратичної форми можна вважати матрицею лінійного оператора (див. зауваження 7.1), коефіцієнти λ_k квадратичної форми в канонічному базисі та координати векторів цього базису можна знаходити так, як власні числа та координати власних векторів лінійного оператора, тобто з характеристичного рівняння (6.23) та системи рівнянь (6.22). Пояснімо це простими прикладами.

● **Приклад 7.2.** Зведемо до суми квадратів квадратичну форму

$$f(\mathbf{x}, \mathbf{x}) = x_1^2 + 4x_2^2 + 4x_1x_2 ,$$

задану у двовимірному лінійному просторі, і знайдемо вектори канонічного базису. Цій квадратичній формі відповідає матриця

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}. \tag{7.13}$$

З виразу (6.23) знаходимо рівняння для коефіцієнтів квадратичної форми в канонічному базисі

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} = 0 ,$$

яке має корені $\lambda_1 = 0$, $\lambda_2 = 5$. Із виразу (6.22) знаходимо систему рівнянь для координат базисних векторів

$$\begin{cases} (1-\lambda)x_1 + 2x_2 = 0, \\ 2x_1 + (4-\lambda)x_2 = 0, \end{cases}$$

з якої знаходимо базисні вектори $\mathbf{x}^{(1)} = C_1(2,-1)$ та $\mathbf{x}^{(2)} = C_2(1,2)$ для $\lambda_1 = 0$ та $\lambda_2 = 5$, відповідно. Легко помітити, що скалярний добуток $\mathbf{x}^{(1)}\mathbf{x}^{(2)}$ дорівнює нулю, тому за фіксованих значень довільних сталих C_1 та C_2 вектори $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$ є векторами *ортогонального* канонічного базису, у якому квадратична форма зводиться до суми квадратів $f(\mathbf{x},\mathbf{x}) = 5\tilde{x}_2^2$.●

Наведений приклад цікавий тим, що визначник матриці A дорівнює нулю (див. вираз (7.13)). Саме тому характеристичне рівняння має нульовий корінь, а це, у свою чергу, спричиняє відсутність доданка з $\tilde{x}_1^2$ у квадратичній формі, представленій у канонічному базисі.

● **Приклад 7.3.** Зведемо до суми квадратів квадратичну форму

$$f(\mathbf{x},\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 + 2x_2x_3\,, \tag{7.14}$$

задану в тривимірному лінійному просторі, і знайдемо її канонічний базис. Оскільки поставлене завдання не відрізняється принципово від того, що було поставлене в попередньому прикладі, то слід пояснити шлях його виконання конспективно.

Квадратичній формі (7.14) відповідає матриця

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Рівняння для коефіцієнтів квадратичної форми

$$\begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = 0$$

є кубічним алгебраїчним рівнянням

$$(1-\lambda)[(1-\lambda)^2 - 2] = 0\,,$$

яке має корені

$$\lambda_1 = 1\,,\ \lambda_2 = 1+\sqrt{2}\,,\ \lambda_3 = 1-\sqrt{2}\,.$$

Ці корені є коефіцієнтами, що входять до виразу квадратичної форми в канонічному базисі

$$f(\mathbf{x},\mathbf{x}) = \tilde{x}_1^2 + (1+\sqrt{2})\tilde{x}_2^2 + (1-\sqrt{2})\tilde{x}_3^2 .$$

З виразу (6.22) знаходимо систему рівнянь

$$\begin{cases} (1-\lambda)x_1 + x_3 = 0, \\ (1-\lambda)x_2 + x_3 = 0, \\ x_1 + x_2 + (1-\lambda)x_3 = 0. \end{cases}$$

Підставивши до цієї системи значення $\lambda = \lambda_1 = 1$, знаходимо координати першого базисного вектора $x_3 = 0$, $x_2 = -x_1$, $x_1 = C_1$ і сам вектор $\mathbf{x}^{(1)} = C_1(1,-1,0)$. Таким самим чином знаходимо два інші базисні вектори: $\mathbf{x}^{(2)} = C_2(1,1,\sqrt{2})$, $\mathbf{x}^{(3)} = C_3(1,1,-\sqrt{2})$. Числа C_1, C_2, C_3 є довільними сталими. Упевнюємося, що скалярні добутки базисних векторів задовольняють умову

$$\mathbf{x}^{(1)}\mathbf{x}^{(2)} = \mathbf{x}^{(1)}\mathbf{x}^{(3)} = \mathbf{x}^{(2)}\mathbf{x}^{(3)} = 0 ,$$

і доходимо висновку, що канонічний базис – ортогональний.●

В обох наведених вище прикладах *не випадково канонічні базиси виявилися ортогональними*. Справа в тому, що матриця квадратичної форми відповідає лінійному оператору та задовольняє умову $A = A^T$. У підрозд. 6.4 було зазначено, що таку умову задовольняє матриця оператора Ерміта, якщо цей оператор діє в дійсному просторі. Також було зазначено, що власні вектори оператора Ерміта утворюють ортонормований базис. Саме цим і пояснюється загальна риса канонічних базисів квадратичних форм, розглянутих у прикладах 7.2 та 7.3. Більше того, оскільки матриця оператора Ерміта стає діагональною при переході до базису, утвореному власними векторами, то канонічний базис існує для кожної квадратичної форми.

7.3.2. Додатна визначеність квадратичної форми

Наведені вище приклади показують, що вирази для базисних векторів містять довільні сталі, тому кожна квадратична форма має безліч канонічних базисів, які відрізняються один від одного довжиною базисних векторів та/або зміною напрямків деяких з них на протилежні. Ці приклади показують також, що коефіцієнти λ_k зведеної до суми квадратів квадратичної форми можуть бути додатними, від'ємними або дорівнювати нулю. Якщо всі коефіцієнти зведеної до суми квадратів квадратичної форми додатні, то цілком очевидно, що вона набуває мінімальної величини, коли всі змінні x_k обертаються на нуль. Тобто вектор $\mathbf{x}=\mathbf{0}$ "надає мінімальне значення" функції $f(\mathbf{x},\mathbf{x})$, причому $f(\mathbf{0},\mathbf{0})=0$, та $f(\mathbf{x},\mathbf{x})>0$, якщо $\mathbf{x}\neq\mathbf{0}$. Цілком очевидно, що якщо хоча б один з коефіцієнтів λ_k зведеної до суми квадратів квадратичної форми від'ємний, то поява ненульової величини відповідної змінної x_k зменшує величину квадратичної форми, а отже, вектор $\mathbf{x}=\mathbf{0}$ не надає мінімального значення функції $f(\mathbf{x},\mathbf{x})$. Факт наявності мінімуму функції та її мінімальне значення не можуть змінитися при заміні базису, тому, якщо в якомусь з канонічних базисів усі коефіцієнти квадратичної форми додатні, то і в будь-якому іншому канонічному базисі від'ємних коефіцієнтів в неї немає.

Означення 7.6. Якщо всі коефіцієнти зведеної до суми квадратів квадратичної форми додатні, її називають додатно визначеною.

Щоб квадратична форма мала мінімум, вона має бути додатно визначеною. Якщо функція $f(\mathbf{x},\mathbf{x})$ задана в багатовимірному просторі, то зведення квадратичної форми до суми квадратів може бути важким завданням. З огляду на це, бажано з'ясовувати, чи є квадратична форма додатно визначеною, не переходячи до канонічного базису. Таку можливість надає критерій Сильвестра.

Критерій Сильвестра: квадратична форма (7.9) є додатно визначеною тоді і лише тоді, коли визначник матриці $A(a_{jk})$ та всі її головні мінори додатні.

Зауваження 7.3. Головним мінором першого порядку є елемент a_{11}, а головними мінорами вищих порядків – визначники

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},\ \dots,\ \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1(n-1)} \\ a_{21} & a_{22} & \cdots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ a_{(n-1)1} & a_{(n-1)2} & \cdots & a_{(n-1)(n-1)} \end{vmatrix}.$$

Згідно з критерієм Сильвестра, якщо хоч один із цих визначників від'ємний або дорівнює нулю, квадратична форма не є додатно визначеною.

Доведення критерію Сильвестра досить складне (див., напр., § 6 підручника [9]). Щоб зрозуміти, як застосовувати цей критерій на практиці, достатньо розглянути прості приклади.

● **Приклад 7.4.** З'ясуємо, чи є додатно визначеною квадратична форма (7.14), обчисливши головні мінори та визначник її матриці:

$$a_{11} = 1\,, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1\,, \ \det A = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -1\,.$$

Висновок: квадратична форма (7.14) не є додатно визначеною. Цей висновок узгоджується з тим фактом, що знайдений при розгляді прикладу 7.3 коефіцієнт λ_3 від'ємний.●

● **Приклад 7.5.** Визначимо, чи додатно визначена квадратична форма

$$f(\mathbf{x},\mathbf{x}) = 5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3\,. \tag{7.15}$$

Обчислимо її головні мінори та визначник:

$$a_{11} = 5\,, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1\,, \ \det A = \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1$$

Висновок: квадратична форма (7.15) є додатно визначеною.●

7.3.3. Застосування квадратичних форм для вивчення коливальних систем з декількома ступенями вільності

Для багатьох сфер людської діяльності (медицини, інженерії, наукових досліджень) вирішально важливим є явище резонансу (магнітного, електричного, акустичного тощо). Як буде показано далі, резонансні частоти близькі за величиною до частот вільних коливань. Значна частка широко відомих резонансних методик використовує коливання одразу декількох складових векторної фізичної величини. З огляду на це, у базових курсах теоретичної, експериментальної та прикладної фізики розглянуто вільні коливання систем з декількома ступенями вільності. Щоб показати, як теорія таких коливань використовує основи лінійної алгебри, розглянемо в простішій спосіб коливання фізичної системи, яка має n ступенів вільності[1]. За певних припущень та фізичних умов вираз для енергії такої системи має форму

$$E = \frac{1}{2}\sum_{j=1}^{n} g\dot{x}_j^2 + U(x_j, x_k), \tag{7.16}$$

де залежні від часу величини x_j, x_k описують малі відхилення системи від рівноваги; $\dot{x}_j$ – швидкості зміни цих величин; сума, що стоїть у правій частині рівняння (7.16), є кінетичною енергією системи; функція $U(x_j, x_k)$ описує потенціальну енергію і має локальний мінімум у положенні рівноваги; стала g характеризує інерційність фізичної системи[2].

У багатьох практично важливих випадках величини x_j, x_k є координатами фізичного вектора $\mathbf{x}$ в *ортогональному* базисі $\left\{\mathbf{e}_j\right\}_{j=\overline{1,n}}$. У стані рівноваги координати x_j, x_k та перші похідні функції $U(x_j, x_k)$ дорівнюють нулю, тому цю функцію можна наблизити такою формулою:

$$U(x_j, x_k) \approx U(0,0) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{\partial^2 U}{\partial x_j \partial x_k}\right)_{\mathbf{x}=\mathbf{0}} x_j x_k .$$

Похідні потенціальної енергії набувають у положенні рівноваги певних сталих значень, тому зручно ввести позначення

$$\left(\frac{\partial^2 U}{\partial x_j \partial x_k}\right)_{\mathbf{x}=\mathbf{0}} \equiv a_{jk}$$

і представити потенціальну енергію як суму її величини в положенні рівноваги $U(0,0)$ і квадратичної форми, яка описує

[1] Більш докладний розгляд цього явища представлений у 1-му томі всесвітньо відомого курсу теоретичної фізики [10].

[2] У випадку механічних коливань матеріальної точки ця стала дорівнює масі точки.

зміну потенціальної енергії при відхиленні системи від цього положення:

$$U(x_j, x_k) \approx U(0,0) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk} x_j x_k . \qquad (7.17)$$

Зауважимо, що згідно із загальними властивостями частинних похідних другого порядку $a_{jk} = a_{kj}$.

Розглядаючи кінетичну енергію коливальної системи в простішій спосіб, вважатимемо залежні від часу координати x_j та x_k пропорційними функції $\exp(i\omega t)$, де ω є кутовою частотою коливань. Диференціювання цієї функції спричиняє лише її множення на сталу $i\omega$. Завдяки цьому вираз (7.16) для суми кінетичної та потенціальної енергії набуває форми

$$E \approx \frac{1}{2}\sum_{j=1}^{n}(-g\omega^2)x_j^2 + U(0,0) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk} x_j x_k .$$

Спектр частот вільних коливань підлягає визначенню. Щоб отримати потрібне для цього рівняння, слід здійснити тотожне алгебраїчне перетворення на зразок того, що було виконано для виведення рівняння (6.45), і виразити енергію формулою

$$E \approx U(0,0) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}(a_{jk} - g\omega^2\delta_{jk})x_j x_k .$$

Згідно із загальними принципами теоретичної фізики частоти вільних коливань визначають з необхідних умов мінімуму енергії: $\partial E / \partial x_j = 0$. Із цих умов одержують систему однорідних алгебраїчних рівнянь

$$\sum_{k=1}^{n}(a_{jk} - g\omega^2\delta_{jk})x_k = 0 , \qquad (7.18)$$

яка має нетривіальний розв'язок, лише в тому разі, якщо визначник її матриці дорівнює нулю:

$$\det(a_{jk} - g\omega^2\delta_{jk}) = 0 . \qquad (7.19)$$

Вираз (7.19) і є тим рівнянням, з якого визначають спектр частот вільних коливань.

Якщо ввести позначення $g\omega^2 \equiv \lambda$, вирази (7.18) та (7.19) перетворяться на систему рівнянь (6.22) і характеристичне рівняння (6.23), відповідно, за допомогою яких зводять до суми квадратів квадратичну форму, що входить до формули потенціальної енергії (7.17). Із рівняння (7.19) визначають резонансні частоти ω_α, по черзі підставляють їх до рівнянь (7.18) та знаходять вектори $\tilde{\mathbf{e}}_\alpha$, які утворюють ортогональний базис квадратичної форми. У цьому базисі вираз для енергії коливальної системи перетворюється на суму

$$E = \sum_{j=1}^{n} E^{(\alpha)} ,$$

доданки якої

$$E^{(\alpha)} = \frac{1}{2} g\dot{\tilde{x}}_\alpha^2 + \lambda_\alpha \tilde{x}_\alpha^2 \tag{7.20}$$

є енергією одновимірних коливань. Ці одновимірні коливання називають *нормальними коливаннями* фізичної системи[1]. Збудження нормальних коливань зовнішніми періодичними силами лежить в основі багатьох широко застосовних резонансних методик.

---

[1] У виразі (7.16) кінетична енергія пропорційна до суми квадратів координат, тобто, до скалярного добутку **xx**. Тому і в канонічному базисі вона пропорційна до цього добутку. У формулі (7.20) враховано, що канонічний базис є ортогональним, тому і в ньому скалярний добуток **xx** дорівнює сумі квадратів координат.

РОЗДІЛ 8
Застосування лінійної алгебри для інтегрування лінійних диференціальних рівнянь

8.1. Вступні зауваження та означення основних понять

Необхідність вивчення різних методів розв'язання диференціальних рівнянь зумовлена тим, що такі рівняння описують важливі процеси, які відбуваються в природі та суспільстві. Серед процесів, які підлягають вивченню, є багато таких, що описуються *лінійними* диференціальними рівняннями. Докладне пояснення основних методів розв'язання таких рівнянь можна знайти в підручниках [11–13]. Ці методи значною мірою базуються на аксіомах та теоремах лінійної алгебри. Цей і наступний розділи підручника мають на меті:

а) вказати на доцільність застосування лінійної алгебри для пояснення методів розв'язання лінійних диференціальних рівнянь та систем лінійних диференціальних рівнянь;

б) продемонструвати використання цих методів на прикладі рівняння вимушених коливань та рівняння, що описує поширення гармонічних хвиль у просторі;

в) пояснити суть дослідження в лінійному наближенні стійкості стаціонарних станів динамічних систем (фізичних, хімічних, технічних тощо).

Перш ніж почати виконувати поставлені завдання, нагадаємо термінологію, що стосується диференціальних рівнянь.

Означення 8.1. Звичайним диференціальним рівнянням називають таке рівняння, яке пов'язує невідому функцію однієї змінної $y(x)$ з її похідними і самою змінною x. Рівняння, яке пов'язує функцію декількох змінних з її частинними похідними, називають рівнянням у частинних похідних.

Наведемо позначення та означення основних понять, що зустрічаються при вивченні та розв'язуванні *звичайних диференціальних рівнянь*, називаючи їх просто *диференціальними рівняннями* (ДР).

У багатьох випадках функцію, яку треба визначити з диференціального рівняння, буває зручно зобразити у вигляді графіка на площині, тому її доцільно позначити як $y(x)$ і вважати x та y декартовими координатами точок цієї площини. Похідні першого, другого та третього порядків від функції $y(x)$ позначають символами y', y'' та y''', відповідно, а символом $y^{(n)}$ позначають похідну порядку n.

Означення 8.2. Якщо рівняння містить похідну $y^{(n)}$, а похідних більш високого порядку в ньому немає, то число n називають *порядком диференціального рівняння.*

Будь-яке диференціальне рівняння n-го порядку можна записати у формі

$$F(x, y, y', y'', y'''...y^{(n)}) = 0, \tag{8.1}$$

тому що праву частину рівняння можна перенести ліворуч, змінивши її знак. В окремих випадках у диференціальному рівнянні n-го порядку можуть бути відсутніми незалежна змінна та/або невідома функція, та/або деякі похідні порядку, меншого від n. Наприклад, якщо $F = y' + 2y$, то в рівнянні відсутня незалежна змінна, а якщо $F = 2y'' + \sin 2y$, то в рівнянні немає першої похідної від невідомої функції.

Означення 8.3. *Розв'язком диференціального рівняння* називають таку функцію, підстановка якої до рівняння перетворює його на тотожність.

Із цього означення випливає можливість знайти розв'язок рівняння шляхом підстановки до нього пробної функції; ця можливість використовується при розв'язуванні лінійних ДР.

Рівняння типу

$$y' = f(x) \tag{8.2}$$

є диференціальним рівнянням 1-го порядку, розв'язаним відносно похідної y'. Його розв'язок виражається через невизначений інтеграл формулою

$$y = \int f(x)dx + C\,. \tag{8.3}$$

Права частина рівності (8.3) містить одну довільну сталу C, яка виникла внаслідок інтегрування правої частини рівняння (8.2). Щоб розв'язати диференціальне рівняння 2-го порядку, треба виконати процедуру інтегрування двічі, після чого в розв'язку рівняння виникнуть дві довільні сталі. Розв'язання рівняння n-го порядку передбачає n-кратне інтегрування, унаслідок чого до розв'язку такого рівняння входять n довільних сталих.

Означення 8.4. Розв'язок диференціального рівняння n-го порядку, який містить n довільних сталих, називають *загальним інтегралом* цього рівняння, а процедуру відшукання загального інтеграла називають *інтегруванням диференціального рівняння.*

Означення 8.5. *Частинним розв'язком* диференціального рівняння називають функцію, яку отримують із загального інтеграла, зафіксувавши значення наявних у ньому довільних сталих.

Означення 8.6. Графік частинного розв'язку диференціального рівняння називають *інтегральною кривою* цього рівняння.

Загальний інтеграл диференціального рівняння зображується сім'єю інтегральних кривих. Розглядаючи окремі інтегральні криві, часто вдається дійти певних висновків про загальний інтеграл.

8.2. Операторна форма лінійного диференціального рівняння та алгебраїчний підхід до інтегрування лінійних диференціальних рівнянь

Означення 8.7. Лінійним ДР порядку n називають рівняння

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \ldots + a_1(x)y' + a_0(x)y = b(x). \tag{8.4}$$

Функції $a_j(x)$ $(j=0,1,2,...n)$ називають коефіцієнтами лінійного ДР, а функцію $b(x)$ – його правою частиною.

Тією рисою лінійного ДР, що пов'язує його з основами лінійної алгебри, є те, що його можна записати в операторній формі

$$\hat{L}(x)y(x)=b(x)\,, \tag{8.5}$$

де

$$\hat{L}(x)=a_n(x)\frac{d^n}{dx^n}+a_{n-1}(x)\frac{d^{n-1}}{dx^{n-1}}+...+a_1(x)\frac{d}{dx}+a_0(x) \tag{8.6}$$

є лінійним оператором, тому що похідна від суми функцій дорівнює сумі похідних і, якщо перед функцією $y(x)$ стоїть числовий коефіцієнт C, його можна винести за знак похідної:

$$\frac{d^j}{dx^j}[y_1(x)+y_2(x)]=\frac{d^j}{dx^j}y_1(x)+\frac{d^j}{dx^j}y_2(x)\,,$$

$$\frac{d^j}{dx^j}Cy(x)=C\frac{d^j}{dx^j}y(x)\,.$$

Завдяки цьому виконуються умови:

$$\begin{aligned}&1)\quad \hat{L}(x)\big(y_1(x)+y_2(x)\big)=\hat{L}(x)y_1(x)+\hat{L}(x)y_2(x),\\ &2)\quad \hat{L}(x)Cy(x)=C\hat{L}y(x),\end{aligned} \tag{8.7}$$

тобто оператор $\hat{L}(x)$ підпадає під означення 6.3 лінійного оператора, що діє в просторі функцій.

Зауваження 8.1. З того, що перша умова лінійності виконується для двох функцій, як це зазначено у виразах (8.7), випливає, що вона справедлива і для більшої кількості функцій.

Визначаючи, чи є лінійним те рівняння, яке треба зінтегрувати для розв'язання тієї чи іншої задачі, доцільно пам'ятати, що до лінійного ДР входять лише перші степені функції $y(x)$ та її похідних і не входять $y^2, y'^2, y''^2,\ldots$ та інші степені цих величин.

Означення 8.8. Однорідним називають лінійне диференціальне рівняння, права частина якого дорівнює нулю.

Запишемо однорідне лінійне ДР у двох еквівалентних формах:

$$a_n(x)y^{(n)}+a_{n-1}(x)y^{(n-1)}+\ldots+a_1(x)y'+a_0(x)y=0 \qquad (8.8)$$

та

$$\hat{L}(x)y(x)=0\,. \qquad (8.9)$$

З того, що означений формулою (8.6) оператор $\hat{L}$ задовольняє вимоги лінійності (8.7), випливає основна властивість однорідних лінійних ДР – принцип суперпозиції розв'язків.

Принцип суперпозиції розв'язків *полягає в тому, що будь-яка лінійна комбінація розв'язків однорідного лінійного ДР є його розв'язком.*

Доведемо справедливість наведеного принципу. Подіємо оператором $\hat{L}$ на довільну лінійну комбінацію розв'язків

$$y(x)=\sum_{\alpha=1}^{p}C_\alpha y_\alpha(x)\,, \qquad (8.10)$$

урахувавши умови лінійності цього оператора:

$$\hat{L}\sum_{\alpha=1}^{p}C_\alpha y_\alpha(x)=\sum_{\alpha=1}^{p}\hat{L}C_\alpha y_\alpha(x)=\sum_{\alpha=1}^{p}C_\alpha \hat{L}y_\alpha(x).$$

Якщо функції $y_\alpha(x)$ є розв'язками однорідного лінійного ДР, то підстановка кожного з них до рівняння (8.10) перетворює рівняння на тотожність $\hat{L}y_\alpha(x)\equiv 0$, завдяки чому

$$\hat{L}\sum_{\alpha=1}^{p}C_\alpha y_\alpha(x)\equiv 0\,,$$

тобто функція (8.10) є розв'язком рівняння (8.9), а отже, і еквівалентного йому рівняння (8.8).

Простота доведення принципу суперпозиції розв'язків лінійного ДР, вираженого в операторній формі, демонструє доцільність застосування поняття лінійного оператора до однорідних лінійних ДР. Подальший розгляд таких ДР базується на означенні 3.5 лінійної комбінації елементів лінійного простору та означенні 3.7 лінійно незалежних елементів.

Використовуючи операторну форму неоднорідного лінійного ДР, легко довести теорему, яку використовують для інтегрування багатьох рівнянь такого типу.

Теорема 8.1. Загальний інтеграл неоднорідного лінійного рівняння є сумою загального інтеграла однорідного рівняння та частинного розв'язку неоднорідного:

$$y_{\text{з.н.}} = y_{\text{з.о.}} + y_{\text{ч.н.}}\,. \tag{8.11}$$

Доведення. Згідно з означенням 8.3 підставлення функції $y_{\text{з.н}}$ $y_{\text{з.н.}}$ до неоднорідного лінійного рівняння (8.5) і функції $y_{\text{ч.н}}$ $y_{\text{ч.н.}}$ до однорідного лінійного рівняння (8.9) перетворює це рівняння на тотожності

$$\begin{aligned} &\hat{L}y_{\text{з.н.}} \equiv b(x),\\ &\hat{L}y_{\text{з.о.}} \equiv 0. \end{aligned} \tag{8.12}$$

Для доведення теореми достатньо розглянути різницю загальних інтегралів однорідного та неоднорідного рівнянь

$$y_{\text{з.н.}} - y_{\text{з.о.}} \overset{def}{\equiv} \psi(x)\,, \tag{8.13}$$

і довести, що ця різниця є частинним розв'язком неоднорідного рівняння. З цією метою подіємо оператором $\hat{L}$ на праву та ліву частини тотожності (8.13); оскільки права та ліва частини тотожності (8.13) – це два різних позначення різниці функцій $y_{\text{з.н.}}$ та $y_{\text{з.о.}}$, то під дією оператора вони перетворюються на одну й ту саму функцію, тобто справджується тотожність

$$\hat{L}(y_{\text{з.н.}} - y_{\text{з.о.}}) \equiv \hat{L}\psi(x)\,.$$

Ураховуючи лінійність оператора, отримуємо тотожність

$$\hat{L}y_{\text{з.н.}} - \hat{L}y_{\text{з.о.}} \equiv \hat{L}\psi(x)\,. \tag{8.14}$$

Згідно з тотожностями (8.12) обидва доданки в лівій частині тотожності (8.14) дорівнюють нулю, тобто справджується тотожність $\hat{L}\psi(x) \equiv 0$, яка доводить, що різниця загальних інтегралів неоднорідного та однорідного рівнянь є розв'язком однорідного рівняння.

8.3. Загальний інтеграл однорідного лінійного рівняння

8.3.1. Суперпозиція лінійно незалежних частинних розв'язків

З огляду на принцип суперпозиції розв'язків і на той факт, що загальний інтеграл диференціального рівняння порядку n містить n довільних сталих, у математиків виникла ідея утворити загальний інтеграл $y_{\text{з.о.}}(x)$ однорідного лінійного ДР у вигляді лінійної комбінації частинних розв'язків, кількістю n:

$$y_{\text{з.о.}}(x) = \sum_{\alpha=1}^{n} C_\alpha y_\alpha(x). \qquad (8.15)$$

При цьому виникло запитання: чи з будь-яких частинних розв'язків кількістю n можна утворити загальний інтеграл? Відповідь на поставлене запитання майже очевидна. Щоб дати її, розглянемо однорідне лінійне ДР 3-го порядку і його частинні розв'язки $y_1(x)$, $y_2(x)$, $y_3(x)$. Утворимо лінійну комбінацію цих розв'язків:

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + C_3 y_3(x). \qquad (8.16)$$

Припустимо, що частинні розв'язки $y_1(x)$, $y_2(x)$, $y_3(x)$ лінійно залежні. Тоді один із них, для визначеності $y_3(x)$, є лінійною комбінацією двох інших:

$$y_3(x) = K_1 y_1(x) + K_2 y_2(x), \qquad (8.17)$$

де K_1, K_2 – сталі коефіцієнти. Підставивши вираз (8.17) до виразу (8.16), одержуємо функцію

$$y(x) = (C_1 + K_1) y_1(x) + (C_2 + K_2) y_2(x),$$

яка містить лише дві довільні сталі, а саме: $\tilde{C}_1 \equiv C_1 + K_1$ та $\tilde{C}_2 \equiv C_2 + K_2$. Отже, ця функція не може бути загальним інтегралом ДР 3-го порядку, до якого обов'язково входять три довільні сталі. Зменшення кількості довільних сталих від трьох, які мають бути наявними в загальному інтегралі ДР 3-го порядку, до двох трапилося внаслідок того, що функція $y_3(x)$ є, за припущенням, лінійно залежною від $y_1(x)$ та $y_2(x)$. Звідси випливає надважливе твердження, яке є відповіддю на поставлене запитання.

Твердження: загальний інтеграл однорідного лінійного ДР порядку n є лінійною комбінацією n *лінійно незалежних* частинних розв'язків.

Справедливість цього твердження була щойно продемонстрована для ДР 3-го порядку, а для ДР порядку n це твердження доводять шляхом розгляду задачі Коші, яка полягає у відшуканні такого розв'язку однорідного лінійного ДР, що задовольняє початкові умови

$$y(x_0) = y_0,\ y'(x_0) = y_0{}',\ldots,\ y^{(n-1)}(x_0) = y_0{}^{(n-1)}, \qquad (8.18)$$

де величини y_0', y_0'', ... $y_0^{(n-1)}$ є не похідними від сталої y_0, а зручними позначеннями сталих значень похідних порядку 1, 2, ... $n-1$ від функції $y(x)$. Розглядаючи задачу Коші, беруть до уваги, що для кожного фіксованого значення x_0 її розв'язок відповідає певним значенням довільних сталих у загальному інтегралі (8.15), а змінюючи значення x_0, можна отримати всі розв'язки рівняння[1].

Функція (8.15) задовольняє однорідне лінійне ДР за своїм означенням, а для того, щоб бути розв'язком задачі Коші, вона має задовольнити початкові умови (8.18). Після підстановки функції (8.15) до початкових умов ці умови стають системою алгебраїчних рівнянь

$$\begin{cases} C_1 y_1(x_0) + C_2 y_2(x_0) + \ldots + C_n y_n(x_0) = y_0 \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) + \ldots + C_n y_n'(x_0) = y_0' \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \qquad\qquad \vdots \\ C_1 y_1^{(n-1)}(x_0) + C_2 y_2^{(n-1)}(x_0) + \ldots. + C_n y_n^{(n-1)}(x_0) = y_0^{(n-1)} \end{cases} \qquad (8.19)$$

для невідомих C_α. Із цієї системи лише тоді можна знайти такі значення невідомих C_α, за яких лінійна комбінація частинних

[1] Умови існування розв'язку задачі Коші вважають виконаними.

розв'язків $y_\alpha(x)$ є розв'язком задачі Коші, коли визначник системи не дорівнює нулю. Це випливає, серед іншого, із формул Крамера для розв'язку неоднорідної лінійної системи алгебраїчних рівнянь. Для системи рівнянь (8.19) це формули

$$C_\alpha = \frac{\Delta^{(\alpha)}}{W}, \ \ \alpha = 1, 2, \ldots n-1, n,$$

де

$$W = \begin{vmatrix} y_1(x_0) & y_2(x_0) & \ldots & y_n(x_0) \\ y_1'(x_0) & y_2'(x_0) & \ldots & y_n'(x_0) \\ \vdots & \vdots & \vdots & \vdots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \ldots & y_n^{(n-1)}(x_0) \end{vmatrix}$$

є визначником основної матриці цієї системи рівнянь, а $\Delta^{(\alpha)}$ – визначниками, у яких стовпчик визначника W з номером α замінений на стовпчик, елементами якого є праві частини рівнянь (8.19).

Тепер зважимо на те, що задачу Коші можна поставити для будь-якого значення змінної x, а отже, *лінійна комбінація частинних розв'язків є загальним інтегралом однорідного лінійного ДР лише за умови*

$$W \equiv \begin{vmatrix} y_1(x) & y_2(x) & \ldots & y_n(x) \\ y_1'(x) & y_2'(x) & \ldots & y_n'(x) \\ \vdots & \vdots & \vdots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \ldots & y_n^{(n-1)}(x) \end{vmatrix} \neq 0. \qquad (8.20)$$

Визначник W названо *визначником Вронського* на честь відомого польського математика.

Щоб завершити доведення сформульованого вище твердження, зауважимо, що умова $W \neq 0$ є умовою лінійної незалежності стовпчиків визначника Вронського, а отже, і функцій $y_\alpha(x)$ (див. п. 4.5.3).

Висновок 8.1. Алгебраїчний підхід до розв'язання лінійного однорідного диференціального рівняння полягає в пошуку сукупності n лінійно незалежних частинних розв'язків рівняння і представленні загального інтеграла рівняння у формі лінійної комбінації цих частинних розв'язків.

Зауваження 8.2. Математики назвали сукупність n лінійно незалежних частинних розв'язків однорідного лінійного ДР *фундаментальною системою розв'язків* і довели, що вона існує для будь-якого однорідного лінійного ДР.

8.3.2. Загальний інтеграл однорідного рівняння зі сталими коефіцієнтами

При розгляді багатьох практично важливих питань фізико-технічних наук виникає необхідність у розв'язанні однорідних лінійних ДР зі сталими коефіцієнтами[1]. Розглянемо однорідне лінійне ДР:

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0, \qquad (8.21)$$

у якому a_j – дійсні константи, $j = \overline{1,n}$. Оскільки експоненціальні функції з різними показниками $y_\alpha = e^{k_\alpha x}$ лінійно незалежні (жодну з них не можна представити у вигляді лінійної комбінації експонент $e^{k_\beta x}$ з показниками $k_\beta x \neq k_\alpha x$), то зі сформульованого вище твердження випливає ідея розглянути пробну функцію

$$y(x) = e^{kx} \qquad (8.22)$$

і з'ясувати, яким має бути коефіцієнт k, щоб ця функція була частинним розв'язком рівняння (8.21). Для цього слід підставити функцію (8.22) до цього рівняння. Кожне диференціювання експоненти e^{kx} помножує її на k, тобто

$$\frac{d^m}{dx^m} e^{kx} = k^m e^{kx}. \qquad (8.23)$$

Підставивши функцію (8.22) до рівняння (8.21), зваживши на вираз (8.23) і скоротивши ліву частину рівняння на e^{kx}, доходимо висновку, що функція (8.22) є частинним розв'язком

[1] Деякі з цих питань будуть проаналізовані далі.

рівняння (8.21) лише в тому разі, якщо коефіцієнт k задовольняє таке алгебраїчне рівняння порядку n:

$$a_n k^n + a_{n-1} k^{n-1} + \ldots + a_1 k + a_0 = 0 . \qquad (8.24)$$

Означення 8.9. Рівняння (8.24) називають характеристичним рівнянням лінійного ДР зі сталими коефіцієнтами, а його корені – характеристичними числами.

Яким саме є розв'язок однорідного лінійного ДР, залежить від того, які корені має характеристичне рівняння. Саме тому це рівняння називають характеристичним.

1. Якщо характеристичне рівняння має лише *прості* (не кратні) корені $k = k_\alpha$ і всі ці корені *дійсні*, то кількість коренів дорівнює порядку рівняння. Тоді слід розглянути добірку з n лінійно незалежних частинних розв'язків $y_\alpha(x) = e^{k_\alpha x}$ та утворити з них загальний інтеграл

$$y(x) = C_1 e^{k_1 x} + C_2 e^{k_2 x} + \ldots + C_{n-1} e^{k_{n-1} x} + C_n e^{k_n x} . \qquad (8.25)$$

2. Якщо серед *простих* коренів характеристичного рівняння є *комплексні*, то поряд з кожним комплексним коренем

$$k_\alpha = \gamma_\alpha + i\omega_\alpha$$

є комплексно-спряжений корінь[1] $k_\alpha = \gamma_\alpha - i\omega_\alpha$, де γ_α та ω_α – дійсна та уявна частини кореня k_α. Тоді диференціальне рівнян-

---

[1] Із курсу алгебри відомо, що алгебраїчне рівняння можна записати у формі $a_n \prod\limits_\alpha (k - k_\alpha)^{m_\alpha} = 0$, де m_α – кратність кореня k_α. Добуток двох комплексно-спряжених коренів є дійсним числом:

$$[k - (\gamma_\alpha + i\omega_\alpha)][k - (\gamma_\alpha - i\omega_\alpha)] = (k - \gamma_\alpha)^2 + \omega_\alpha^2 .$$

Якщо ж корінь $k_\alpha = \gamma_\alpha + i\omega_\alpha$ "не має пари", то до рівняння ввійде доданок з уявним коефіцієнтом, пропорційним числу ω_α.

ня (8.21) задовольняють комплексні функції $y_{\alpha}^{(c)} = e^{(\gamma_{\alpha}+i\omega_{\alpha})x}$ та $y_{\alpha+1}^{(c)} = e^{(\gamma_{\alpha}-\omega_{\alpha})x}$, з яких слід утворити дійсні частинні розв'язки

$$y_{\alpha} = \frac{1}{2}(y_{\alpha}^{(c)} + y_{\alpha+1}^{(c)}) = e^{\gamma_{\alpha}x}\cos\omega_{\alpha}x,$$

$$y_{\alpha+1} = \frac{1}{2i}(y_{\alpha}^{(c)} - y_{\alpha+1}^{(c)}) = e^{\gamma_{\alpha}x}\sin\omega_{\alpha}x$$

і включити до загального інтеграла доданки

$$C_{\alpha}y_{\alpha} + C_{\alpha+1}y_{\alpha+1} = e^{\gamma_{\alpha}x}(C_{\alpha}\cos\omega_{\alpha}x + C_{\alpha+1}\sin\omega_{\alpha}x). \qquad (8.26)$$

3. Якщо серед коренів характеристичного рівняння є *дійсний корінь кратності* m, то можна довести, що до загального інтеграла слід включити доданки

$$\begin{aligned} C_{\alpha}y_{j} + C_{\alpha}y_{\alpha+1} + \ldots + C_{\alpha}y_{\alpha+m-1} = e^{\gamma_{\alpha}x}(C_{\alpha} + C_{\alpha+1}x + \\ + C_{\alpha+2}x^{2} + \ldots + C_{\alpha+m-1}x^{\alpha+m-1}). \end{aligned} \qquad (8.27)$$

4. Також можна довести, що за наявності *комплексного кореня кратності* m до загального інтеграла слід включити доданки

$$\begin{aligned} (A_{\alpha} + A_{\alpha+1}x + A_{\alpha+2}x^{2} + \ldots + A_{\alpha+m-1}x^{\alpha+m-1})e^{\gamma_{\alpha}x}\cos\omega_{\alpha}x + \\ +(B_{\alpha} + B_{\alpha+1}x + B_{\alpha+2}x^{2} + \ldots + B_{\alpha+m-1}x^{\alpha+m-1})e^{\gamma_{\alpha}x}\sin\omega_{\alpha}x. \end{aligned} \qquad (8.28)$$

Необхідність включення до загального інтеграла доданків (8.27), (8.28) можна пояснити так: якщо б кожному значенню кратного кореня відповідав лише один доданок до загального інтеграла, то це призвело б до того, що кількість частинних розв'язків $y_{\alpha}(x)$, з яких утворений загальний інтеграл, була б меншою за порядок рівняння, а отже, і кількість довільних сталих була б меншою за порядок рівняння n, а це не відповідає означенню загального інтеграла.

Висновок 8.2. Щоб зінтегрувати однорідне лінійне ДР зі сталими коефіцієнтами слід розв'язати характеристичне рівняння і знайти загальний інтеграл за такими правилами:

1) кожному простому дійсному кореню характеристичного рівняння відповідає доданок $C_{\alpha}y^{k_{\alpha}x}$, який слід включити до формули загального інтеграла;

2) кожній парі комплексно-спряжених коренів характеристичного рівняння відповідають два доданки (8.26), які слід включити до формули загального інтеграла;

3) кожному кратному дійсному кореню характеристичного рівняння відповідає сукупність доданків (8.27), яку слід включити до формули загального інтеграла;

4) кожному кратному комплексному кореню характеристичного рівняння відповідає сукупність доданків (3.38), яку слід включити до формули загального інтеграла.

Загальний інтеграл лінійного однорідного ДР є сумою доданків, що відповідають всім кореням характеристичного рівняння.

Приклади того, як застосовуються правила 1) – 4) для розв'язання однорідних лінійних ДР із *заданими* числовими коефіцієнтами, наведено в посібнику [13]. Але при викладенні багатьох важливих питань фізико-математичних і технічних дисциплін виникає потреба у визначенні та поясненні розв'язку рівнянь, до яких входять такі параметри, які можуть набувати *різних* числових значень. Одним із таких рівнянь є рівняння вільних коливань

8.3.3. Загальний інтеграл рівняння вільних коливань

Рівняння вільних коливань

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0 \qquad (8.29)$$

є однорідним лінійним рівнянням 2-го порядку для змінної величина x, що характеризує відхилення динамічної системи (фізичної, хімічної, біологічної тощо) від рівноважного стану. У цьому рівнянні $\dot{x} = dx/dt$, $\ddot{x} = d^2x/dt^2$, t – час, γ – коефіцієнт загасання або наростання коливань (якщо він додатний або від'ємний, відповідно), ω_0 − кутова частота коливань. Якщо параметри системи не змінюються під час її коливань, то коефіцієнти рівняння сталі. При записі рівняння коливань було використано те, що всі його доданки можна поділити на коефіцієнт біля другої похідної.

Для визначеності *вважатимемо коефіцієнт* γ *додатним*. Перше, що слід зауважити, це те, що коливання можуть виявитися неможливими в системі з великим коефіцієнтом γ, який у фізиці зазвичай характеризує дисипацію енергії системи, тобто перетворення енергії на тепло. Цей висновок випливає з аналізу характеристичного рівняння

$$k^2 + 2\gamma k + \omega_0^2 = 0, \tag{8.30}$$

розв'язки якого

$$k_{\pm} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2} \tag{8.31}$$

дійсні, якщо $\gamma > \omega_0$. У такому разі, згідно з правилом 1) загальний інтеграл рівняння є суперпозицією двох експонент з від'ємними показниками:

$$x = C_1 e^{k_+ t} + C_2 e^{k_- t}, \quad k_- < k_+ < 0. \tag{8.32}$$

Час наближення системи до рівноваги визначається коренем k_+, оскільки внаслідок нерівності $|k_+| < |k_-|$ величина першого доданка у формулі (8.32) зменшується із часом повільніше, ніж величина другого.

Слід зауважити, що рівняння (8.26) перетворюється на ДР першого порядку $2\gamma\dot{x} = -\omega_0^2 x$, якщо його першим доданком можна знехтувати порівняно з другим. Це можливо, коли сили інерції системи слабкі, а формально, коли $\omega_0^2 / 2\gamma << k_+$, тобто $\omega_0 << 2\gamma$.

Вільні коливання динамічної системи можливі, якщо

$$\gamma < \omega_0. \tag{8.33}$$

У цьому разі корені характеристичного рівняння є комплексними, а отже, його загальним інтегралом є функція

$$x_{\text{з.о.}} = e^{-\gamma t}(C_1 \cos\omega t + C_2 \sin\omega t), \tag{8.34}$$

де $\omega \equiv \sqrt{\omega_0^2 - \gamma^2}$ (див. вираз (8.31)). Замість довільних сталих C_1 та C_2 введемо до розгляду сталі величини a та α за формулами

$$C_1 = a\cos\alpha, \quad C_2 = a\sin\alpha. \tag{8.35}$$

Легко впевнитися, що це можна зробити за будь-яких величин C_1 та C_2.

Загальний інтеграл рівняння вільних коливань виражається через введені величини формулою

$$x_{\text{з.о.}} = ae^{-\gamma t}\cos(\omega t + \alpha)\,, \qquad (8.36)$$

яка показує, що під час вільних коливань динамічної системи їхня амплітуда експоненціально зменшується.

На практиці стала величина a визначається за амплітудою вільних коливань, яка не може набувати будь-яких значень. Обмеження на її величину визначаються індивідуальними особливостями динамічної системи. Фаза коливань α визначається тим, чому дорівнювало відхилення системи від рівноваги на момент початку спостереження коливань.

Для природничих наук та інженерії велике значення мають не лише вільні коливання, але й вільне поширення хвиль у просторі, яке описується хвильовим рівнянням. Хвильове рівняння є рівнянням у частинних похідних. Оскільки вивчення рівнянь у частинних похідних виходить за межі курсу звичайних ДР, то стислий розгляд хвильового рівняння та його розв'язків винесений до підрозд. 8.5.

8.4. Загальний інтеграл неоднорідного лінійного рівняння

8.4.1. Відшукання загального інтеграла методом варіації довільних сталих

Розглянемо *неоднорідне* лінійне ДР

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y' + a_0 y = b(x)\,, \qquad (8.37)$$

у якому $a_j = \text{const}$, $j = \overline{1,n}$. Як було доведено, загальний інтеграл *однорідного* лінійного диференціального рівняння є лінійною комбінацією лінійно незалежних частинних розв'язків y_α цього рівняння (див. (8.15)). Спробуємо знайти розв'язок *неоднорідного* рівняння у формі

$$y(x)=\sum_{\alpha=1}^{n}u_\alpha(x)y_\alpha(x). \tag{8.38}$$

Цей розв'язок відрізняється від загального інтеграла однорідного рівняння заміною довільних сталих на невідомі функції $u_\alpha(x)$. Щоб визначити невідомі функції, підставимо пробний розв'язок (8.38) до рівняння (8.37). Оскільки це рівняння містить похідні $y', y'', \ldots, y^{(n-1)}, y^{(n)}$, то для відшукання функцій $u_\alpha(x)$ треба обчислити ці похідні.

Надважливий момент: для відшукання функцій $u_1(x)$, $u_2(x)$, ... $u_n(x)$ є лише одне рівняння (8.37), тому при обчисленні похідних можна накласти на ці функції $n-1$ таких додаткових умов, які спростять подальші розрахунки

$$y'=\sum_\alpha u_\alpha y'_\alpha+\sum_\alpha u'_\alpha y_\alpha\ ;$$

умова 1: $\sum\limits_\alpha u'_\alpha y_\alpha=0$,

$$y''=\sum_\alpha u_\alpha y''_\alpha+\sum_\alpha u'_\alpha y'_\alpha\ ;$$

умова 2: $\sum\limits_\alpha u'_\alpha y'_\alpha=0$,

$$y^{(n-1)}=\sum_\alpha u_\alpha y_\alpha{}^{(n-1)}+\sum_\alpha u'_\alpha y_\alpha{}^{(n-2)}\ ;$$

умова $n-1$: $\sum\limits_\alpha u'_\alpha y_\alpha{}^{(n-2)}=0$,

$$y^{(n)}=\sum_\alpha u_\alpha y_\alpha{}^{(n)}+\sum_\alpha u'_\alpha y_\alpha{}^{(n-1)}.$$

При обчисленні останньої похідної умова не накладається, тому що вже є n рівнянь (рівняння (8.37) та рівняння накладених умов) для функцій $u_\alpha(x)$.

Підставимо вирази для всіх похідних до рівняння (8.37):

$$a_n(x)\sum_{\alpha} u_{\alpha} y_{\alpha}^{(n)} + a_{(n-1)}(x)\sum_{\alpha} u_{\alpha} y_{\alpha}^{(n-1)} \ldots$$
$$\ldots + a_1(x)\sum_{\alpha} u_{\alpha} y'_a + a_0(x)\sum_{\alpha} u_{\alpha} y_a + a_n(x)\sum_{\alpha} u'_{\alpha} y_{\alpha}^{(n-1)} = b(x). \quad (8.39)$$

В усіх доданках рівняння (8.39), окрім останнього, є співмножник u_{α}, який можна винести за дужки

$$\sum_{\alpha} u_{\alpha}[a_n(x) y_{\alpha}^{(n)} + a_{n-1}(x) y_{\alpha}^{(n-1)} + \ldots + a_0 y_{\alpha}]$$
$$+ a_n(x)\sum_{\alpha} u'_{\alpha} y_{\alpha}^{(n-1)} = b(x). \quad (8.40)$$

Доданки у квадратних дужках є лівою частиною однорідного лінійного ДР, а y_{α} − його частинними розв'язками, тому сума доданків у квадратних дужках дорівнює нулю. Завдяки цьому рівняння (8.40) перетворюється на *умову* n :

$$a_n(x)\sum_{\alpha} u'_{\alpha} y_{\alpha}^{(n-1)} = b(x),$$

якою слід доповнити умови, накладені на невідомі функції. Доповнена сукупність умов є системою n диференціальних рівнянь 1-го порядку для n функцій $u_{\alpha}(x)$. Важливо, що ці рівняння зводяться до квадратур, якщо розв'язки $y_{\alpha}(x)$ знайдені в ході інтегрування однорідного лінійного ДР.

Зауваження 8.3. Тут *не* використано умову про те, що коефіцієнти a_m сталі, але у випадку коефіцієнтів, що залежні від координат, серйозна проблема полягає у відшуканні розв'язків однорідного рівняння.

● ***Приклад 8.1.*** Зінтегруємо рівняння

$$y'' + y = 1 + e^x . \quad (8.41)$$

Характеристичним рівнянням однорідного ДР $y'' + y = 0$ є рівняння $k^2 + 1 = 0$, яке має пару комплексно-спряжених уявних коренів $k = \pm i$, тому загальний інтеграл однорідного ДР має форму

$$y_{\text{з.о.}} = C_1 \cos x + C_2 \sin x . \quad (8.42)$$

Підставимо до рівняння (8.41) пробну функцію

$$y = u_1(x)\cos x + u_2(x)\sin x\ . \qquad (8.43)$$

Диференціюючи цю функцію, накладемо *умову* 1, яка в даному випадку має форму

$$u_1'\cos x + u_2'\sin x = 0\ . \qquad (8.44)$$

За цієї умови похідні функції y дорівнюють

$$y' = -u_1\sin x + u_2\cos x\ ,$$

$$y'' = -u_1\cos x - u_2\sin x - u_1'\sin x + u_2'\cos x.$$

Підставивши їх до рівняння (8.41), одержуємо *умову* 2:

$$-u_1'\sin x + u_2'\cos x = 1 + e^x \qquad (8.45)$$

Із системи рівнянь (8.44), (8.45) знаходимо вирази для похідних

$$u_2' = (1+e^x)\cos x,$$

$$u_1' = -(1+e^x)\sin x.$$

Виконуємо інтегрування:

$$u_2 = \int (1+e^x)\cos x dx = \sin x + \frac{1}{2}e^x(\sin x + \cos x) + C_2,$$

$$u_1 = -\int (1+e^x)\sin x dx = \cos x - \frac{1}{2}e^x(\sin x - \cos x) + C_2.$$

Підставивши ці вирази для u_2 та u_1 до пробного розв'язку (8.43), знаходимо загальний інтеграл неоднорідного лінійного ДР (8.41):

$$y(x) = C_1\cos x + C_2\sin x + \left(1 + \frac{1}{2}e^x\right). \qquad (8.46)$$

Зауважимо, що цей загальний інтеграл виявився сумою загального інтеграл однорідного ДР (див. (8.42)) та частинного розв'язку неоднорідного ДР. Те, що функція $y(x) = 1 + e^x/2$ є розв'язком рівняння (8.41), легко перевірити, підставивши її до рівняння.●

8.4.2. Визначення загального інтеграла неоднорідного лінійного рівняння методом подібності розв'язку до правої частини рівняння

Запишемо неоднорідне лінійне диференціальне рівняння у двох еквівалентних формах

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y' + a_0 y = b(x), \qquad (8.47)$$

де $a_j = \text{const}$, та

$$\hat{L}y(x) = b(x). \qquad (8.48)$$

Метод подібності застосовують, якщо права частина рівняння $b(x)$ є такою функцією, що похідні $b('x)$, $b''(x)$, $b'''(x), \ldots b^{(n)}(x)$ є функціями такого ж типу, як і $b(x)$, тобто, якщо права частина рівняння є або експонентою $b(x) = \eta e^{\beta x}$, або лінійною комбінацією синусів та косинусів $b(x) = \eta_1 \sin(\beta x) + \eta_2 \cos(\beta x)$, або поліномом $b(x) = \eta_m x^m + \eta_{m-1} x^{m-1} + \ldots + \eta_1 x + \eta_0$, або сумою таких функцій, або добутком полінома на експоненту чи полінома на лінійну комбінацію синусів та косинусів.

Метод подібності спирається на теорему 8.1, згідно з якою загальний інтеграл неоднорідного лінійного ДР є сумою загального інтеграла однорідного лінійного ДР та частинного розв'язку неоднорідного.

Оператор $\hat{L}$ перетворює експоненту $e^{\beta x}$ на добуток $(a_n \beta^n + a_{n-1} \beta^{n-1} + \ldots + a_1 \beta + a_0) e^{\beta x}$, лінійну комбінацію синуса та косинуса на іншу лінійну комбінацію синуса та косинуса, поліном на інший поліном. З огляду на це, якщо права частина рівняння є:

- експонентою;
- лінійною комбінацією синуса та косинуса;
- поліномом,

то слід взяти *відповідну* пробну функцію

- $y_1(x) = A e^{\beta x}$;
- $y_2(x) = A \sin(\beta x) + B \cos(\beta x)$;
- $y_3 = A_m x^m + A_{m-1} x^{m-1} + \ldots + A_1 x + A_0$

та знайти такі коефіцієнти A, B, $A_0, \dots A_m$, щоб ця функція задовольнила рівняння (8.47), тобто була його частинним розв'язком. Завдяки цьому буде знайдений частинний розв'язок неоднорідного ДР. Якщо права частина ДР є сумою або добутком перелічених вище функцій, то слід шукати частинний розв'язок у формі суми або добутку функцій y_1, y_2, y_3.

Щоб знайти загальний інтеграл неоднорідного лінійного ДР з правою частиною одного із зазначених вище типів, необхідно:

а) відкинути його праву частину і розв'язати характеристичне рівняння однорідного лінійного ДР;

б) за правилами 1) – 4) скласти загальний інтеграл однорідного ДР;

в) додати до загального інтеграла однорідного ДР знайдений частинний розв'язок неоднорідного ДР.

У багатьох випадках такий метод інтегрування ДР виявляється набагато простішим, ніж метод варіації довільних сталих. Продемонструємо це на прикладі рівняння (8.41). Його права частина $b(x) = 1 + e^x$ вказує на те, що частинний розв'язок слід шукати у формі пробної функції $y = A + Be^x$. Оскільки $y'' = Be^x$, $y'' + y = A + 2Be^x$, то підстановка цієї функції до рівняння (8.41) дає умову $A + 2Be^x = 1 + e^x$. А оскільки ця умова має бути виконана для всіх значень змінної x, то з неї випливає, що $A = 1$, $B = 1/2$, а отже, розв'язок рівняння (8.41) є сумою загального інтеграла (8.42) і частинного розв'язку

$$y_{\text{ч.н}} = 1 + \frac{1}{2}e^x,$$

у повній відповідності до результату (8.46), отриманому методом варіації довільних сталих, що потребував більш складних обчислень.

8.4.3. Вимушені коливання під дією періодичної сили

Одним із найбільш важливих лінійних неоднорідних ДР є рівняння вимушених коливань. Розглянемо коливання під дією періодичної сили

$$f(t) = f_0 \cos \Omega t, \tag{8.49}$$

де f_0 − максимальна величина сили, Ω − кутова частота зміни сили. Такі коливання описуються рівнянням

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f_0 \cos \Omega t \ . \tag{8.50}$$

За формулами (8.11) та (8.36) знаходимо такий вираз для загального інтеграла цього неоднорідного лінійного рівняння:

$$x_{\text{з.н.}} = ae^{-\gamma t} \cos(\omega t + \alpha) + x_{\text{ч.н.}} . \tag{8.51}$$

У більшості випадків для сучасної інженерії дуже важливі коливання звукових, ультразвукових високих та надвисоких частот. У разі таких коливань перший доданок у загальному інтегралі (8.51) зазвичай швидко зменшується за величиною, тому зосередимось на розгляді частинного розв'язку рівняння, який відповідає вимушеним коливанням.

Застосуємо до рівняння (8.50) метод подібності розв'язку до правої частини рівняння. За правилами, сформульованими в п. 8.4.2, частинний розв'язок рівняння (8.50) слід шукати у формі лінійної комбінації синуса та косинуса. Однак до рівняння (8.50) входять декілька параметрів, тому такий спосіб відшукання частинного розв'язку веде до громіздких обчислень. Виявляється, легше діяти в інший спосіб, а саме, розглянути два рівняння:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f_0 \cos \Omega t,$$

$$i\ddot{y} + 2i\gamma\dot{y} + i\omega_0^2 y = if_0 \sin \Omega t,$$

додати їх одне до одного та отримати рівняння для комплексної змінної $z \equiv x + iy$:

$$\ddot{z} + 2\gamma\dot{z} + \omega_0^2 z = f_0 e^{i\Omega t}, \tag{8.52}$$

яке не еквівалентне рівнянню (8.50), але має такий розв'язок, дійсна частина якого є розв'язком цього рівняння. Спрощення обчислень відбувається тому, що частинний розв'язок рівняння (8.52) слід шукати у формі експоненти

$$z_{\text{ч.н.}} = Ae^{i\Omega t} \tag{8.53}$$

і знаходити сталу A, підставляючи його до рівняння (8.52). Зробимо це:

$$A(-\Omega^2 + 2i\gamma\Omega + \omega_0^2)e^{i\Omega t} = f_0 e^{i\Omega t}.$$

Звідси одержуємо простий вираз

$$A = \frac{f_0}{\omega_0^2 - \Omega^2 + 2i\gamma\Omega}.$$

для сталої A, яку називають *комплексною амплітудою* коливань. Далі слід обчислити дійсну частину x комплексної величини z з очевидних співвідношень

$$x = \mathrm{Re}\, z_1 \times \mathrm{Re}\, z_2 - \mathrm{Im}\, z_1 \times \mathrm{Im}\, z_2,$$

де $\mathrm{Re}\, z_{1,2}$ та $\mathrm{Im}\, z_{1,2}$ – дійсні й уявні частини комплексних величин $z_1 = A$ та $z_2 = e^{i\Omega t}$, тобто

$$\mathrm{Re}\, z_2 = \cos\Omega t,$$

$$\mathrm{Im}\, z_2 = \sin\Omega t,$$

$$\mathrm{Re}\, A = \frac{1}{2}(A + A^*) = \frac{1}{2}\left(\frac{f_0}{\omega_0^2 - \Omega^2 + 2i\gamma\Omega} + \frac{f_0}{\omega_0^2 - \Omega^2 - 2i\gamma\Omega}\right),$$

$$\mathrm{Im}\, A = \frac{1}{2i}(A - A^*) = \frac{1}{2i}\left(\frac{f_0}{\omega_0^2 - \Omega^2 + 2i\gamma\Omega} - \frac{f_0}{\omega_0^2 - \Omega^2 - 2i\gamma\Omega}\right).$$

Звівши дроби до спільного знаменника та підставивши знайдені $\mathrm{Re}\, A$ та $\mathrm{Im}\, A$ до формули

$$x = \mathrm{Re}\, A \times \cos\Omega t - \mathrm{Im}\, A \times \sin\Omega t,$$

одержуємо вираз для змінної $x(t)$:

$$x(t) = f_0 \frac{\omega_0^2 - \Omega^2}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}\cos\Omega t + f_0 \frac{2\gamma\Omega}{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}\sin\Omega t. \tag{8.54}$$

Позначимо

$$\begin{aligned} \frac{\omega_0^2 - \Omega^2}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}} &\equiv \cos\delta, \\ \frac{2\gamma\Omega}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}} &\equiv \sin\delta. \end{aligned} \tag{8.55}$$

Це можна зробити, оскільки введені позначення забезпечують виконання тригонометричної рівності $\sin^2\delta + \cos^2\delta = 1$. Підставимо $\sin\delta$ та $\cos\delta$ до рівняння (8.54):

$$x(t) = f_0 \frac{\cos\delta\cos\Omega t}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}} + f_0 \frac{\sin\delta\sin\Omega t}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}}.$$

Введемо до розгляду безрозмірну функцію, означенням якої є формула

$$\chi(\Omega) \overset{def}{\equiv} \frac{\omega_0^2}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\gamma^2\Omega^2}}, \tag{8.56}$$

Означення 8.10. Функцію $\chi(\Omega)$ називають *динамічною сприйнятливістю* коливальної системи.

Величина відхилення системи від положення рівноваги пов'язана з динамічною сприйнятливістю співвідношенням

$$x(t) = \frac{f_0}{\omega_0^2}\chi(\Omega)\cos(\Omega t - \delta). \tag{8.57}$$

Проаналізуємо це співвідношення. *По-перше*, воно показує, що δ є різницею фаз періодичного відхилення коливальної системи від рівноваги, та тієї сили, що збуджує коливання. *По-друге*, амплітуда вимушених коливань пропорційна не лише амплітудному значенню f_0 сили, що спричиняє коливання, але й величині динамічної сприйнятливості коливальної системи.

8.4.4. Явище резонансу

Стала сила f_0 спричиняє незалежне від часу відхилення динамічної системи від рівноваги, тому $\ddot{x} = \dot{x} = 0$ і рівняння (8.50) набуває форми $\omega_0^2 x = f_0$, тобто

$$x = f_0 \big/ \omega_0^2 \qquad (8.58)$$

– це величина відхилення системи від рівноваги під дією сталої сили.

Досвід показує, що відхилення системи від положення рівноваги під дією періодичної сили може значно перевищувати відхилення під дією сталої сили, що дорівнює за величиною амплітудному значенню періодичної сили. Це може здатися дивним, оскільки в такому разі середня величина модуля періодичної сили в $\sqrt{2}$ разів менша за сталу силу, але це можливо, якщо за певних умов виконується сильна нерівність

$$\chi(\Omega) >> 1, \qquad (8.59)$$

(див. (8.57) та (8.58)). Ці умови називають *умовами резонансу*, а те, що відбувається за таких умов, – явищем резонансу, або *резонансом*. Дослідимо це явище у випадку, коли коефіцієнт загасання коливань малий, тобто справджується сильна нерівність

$$\gamma << \omega_0 . \qquad (8.60)$$

Для цього проаналізуємо функцію $\chi(\Omega)$ і різницю фаз δ, виходячи з означення (8.56) та виразу, який випливає з виразів (8.55):

$$\operatorname{tg}\delta = \frac{2\gamma\Omega}{\omega_0^2 - \Omega^2} . \qquad (8.61)$$

Розглянемо три випадки, а саме:

а) граничний випадок сталої сили $f(t)$, тобто $\Omega \to 0$; у цьому випадку $\operatorname{tg}\delta \to +0 \Rightarrow \delta \to 0;\ \chi \to 1,\ x \to f_0 \big/ \omega_0^2$, тобто маємо статичне відхилення від рівноваги (8.58);

б) граничний випадок сили високої частоти $\Omega \to \infty$; у цьому випадку $\operatorname{tg}\delta \to -0 \Rightarrow \delta \to \pi;\ \chi \to 0$, тобто система не встигає відхилитися від рівноваги за половину періоду дії сили;

в) випадок коливань під дією сили, частота якої відповідає максимальній величині динамічної сприйнятливості; у цьому випадку

$$\frac{\partial\chi}{\partial\Omega}=0, \quad \Rightarrow \frac{d}{d\Omega}\Big[(\omega_0^2-\Omega^2)^2+4\gamma^2\Omega^2\Big]^{-\frac{1}{2}}=0\,.$$

Обчисливши похідну, бачимо, що динамічна сприйнятливість набуває максимального значення, якщо частота зовнішньої сили дорівнює *резонансній частоті*

$$\Omega_{\text{рез}}=\sqrt{\omega_0^2-2\gamma^2}\,, \tag{8.62}$$

близькій за величиною до частоти вільних коливань.

Залежність динамічної сприйнятливості від частоти сили представлено на рис. 8.1 для двох малих величин відношення γ/ω_0.

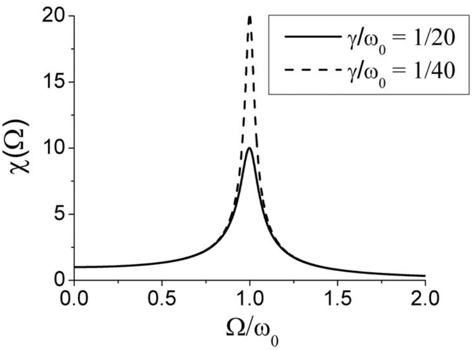

Рис. 8.1. Залежність динамічної сприйнятливості коливальної системи від частоти періодичної сили

Зваживши на сильну нерівність (8.60), бачимо, що на резонансній частоті (8.62) виконуються співвідношення

$$\chi(\Omega_{\text{рез}})\approx\frac{\omega_0}{2\gamma}>>1\,,\ \operatorname{tg}\delta=\frac{2\gamma\sqrt{\omega_0^2-2\gamma^2}}{2\gamma^2}\approx\frac{\omega_0}{\gamma}>>1\,.$$

Ці співвідношення показують, що явище резонансу відбувається за виконання умов

$$\gamma<<\omega_0\,,\ \Omega\approx\Omega_{\text{рез}}\,,\ \delta\approx\pi/2\,, \tag{8.63}$$

тобто, якщо коефіцієнт загасання коливань малий порівняно із частотою вільних коливань, частота сили близька до резонансної частоти, а різниця фаз між періодичною силою і спричиненим нею відхиленням від рівноваги приблизно дорівнює $\pi/2$.

Означення 8.11. Умови (8.63) називають умовами резонансу або, що те саме, резонансними умовами.

Вчені досліджують величезну кількість коливальних систем різних типів, і для кожного типу систем розроблені певні методи спостереження явища резонансу.[1] Значна частина цих методів базується на утворенні *стоячих хвиль* у спеціальних приладах, які називають *резонаторами*. Щоб пояснити суть того, що називають стоячими хвилями, слід проаналізувати розв'язки *хвильового рівняння.*

8.5. Хвильове рівняння, поширення хвиль, стоячі хвилі

Хвильове рівняння математично описує хвилі, що можуть поширюватись у просторі на великі відстані або існувати в обмежених ділянках простору. Такі хвилі характеризуються змінними величинами, які є функціями часу t і просторових координат x, y, z. З цієї причини хвильове рівняння є лінійним рівнянням у частинних похідних. Для визначеності розглянемо змінне електричне поле $\mathbf{E}(x, y, z; t)$ електромагнітної хвилі, що поширюється у вакуумі, і виразимо хвильове рівняння в операторній формі

$$\Box\mathbf{E}(x, y, z, t) = 0 \tag{8.64}$$

[1] Винайдення та дослідження нових типів резонансних явищ актуальне й дотепер. Зокрема, нещодавно був винайдений експериментально й досліджений так, як описано в пп. 8.4.3 та 8.4.4, індукований змінним магнітним полем резонанс пружних коливань у п'єзоелектричній плівці з введеними до неї мікрочастинками феромагнітного металу [14, 15].

із застосуванням лінійного диференціального оператора

$$\Box \overset{def}{\equiv} \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}, \tag{8.65}$$

де x – це просторова координата, t – час, c – сталий параметр, що має фізичну розмірність швидкості. Оператор (8.65) називають оператором д'Аламбера.

На великій відстані від випромінювача електромагнітних хвиль можна вважати хвилі плоскими. Якщо обрати систему координат так, щоб вісь OX була зорієнтована вздовж напрямку поширення хвилі, то вектор електричного поля плоскої хвилі залежатиме лише від змінної x і матиме дві складові E_y та E_z, кожна з яких задовольнятиме хвильове рівняння

$$\frac{\partial^2 E(x,t)}{\partial x^2} - \frac{1}{c^2}\cdot\frac{\partial^2 E(x,t)}{\partial t^2} = 0, \tag{8.66}$$

яке є окремим випадком рівняння (8.66), і тому має таку ж назву. Проведемо заміну змінних $\xi = x + ct$, $\eta = x - ct$, поступово виразивши через ці змінні ліву частину рівняння (8.66):

$$\frac{\partial E}{\partial x} = \frac{\partial E}{\partial \xi}\cdot\frac{\partial \xi}{\partial x} + \frac{\partial E}{\partial \eta}\cdot\frac{\partial \eta}{\partial x} = \frac{\partial E}{\partial \xi} + \frac{\partial E}{\partial \eta} = \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)E,$$

$$\frac{\partial E}{\partial t} = \frac{\partial E}{\partial \xi}\cdot\frac{\partial \xi}{\partial t} + \frac{\partial E}{\partial \eta}\cdot\frac{\partial \eta}{\partial t} = c\left(\frac{\partial E}{\partial \xi} - \frac{\partial E}{\partial \eta}\right) = c\left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta}\right)E,$$

$$\frac{\partial^2 E}{\partial x^2} = \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)^2 E;$$

$$\frac{\partial^2 E}{\partial t^2} = c^2\left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta}\right)^2 E,$$

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2}\cdot\frac{\partial^2 E}{\partial t^2} = \left[\left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)^2 - \left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta}\right)^2\right]E.$$

Розкривши дужки у цих формулах, легко переконатись, що два доданки з $\partial^2 E / \partial \xi^2$ та два доданки з $\partial^2 E / \partial \eta^2$ входять з різними знаками, і тому взаємно знищуються. У результаті рівняння (8.66) набуває форми

$$4\frac{\partial^2 E}{\partial\xi\partial\eta} = 0 .$$

Це рівняння еквівалентне рівнянню

$$\frac{\partial}{\partial\xi}\left(\frac{\partial E}{\partial\eta}\right) = 0 ,$$

яке показує, що похідна $\partial E / \partial\eta$ не залежить від змінної ξ, тобто є функцією лише змінної η:

$$\frac{\partial E}{\partial\eta} = f(\eta) .$$

У такому разі виконується співвідношення

$$E(\eta,\xi) = \int f(\eta)d\eta + V(\xi) , \qquad (8.67)$$

де $V(\xi)$ є довільною функцією змінної ξ. У справедливості співвідношення (8.67) легко переконатися, узявши частинну похідну за змінною η від його правої частини і зваживши на те, що

$$\frac{\partial}{\partial\eta}\int f(\eta)d\eta = f(\eta) \quad \text{та} \quad \partial V(\xi)/\partial\eta = 0 .$$

Невизначений інтеграл у правій частині співвідношення (8.67) є функцією змінної η:

$$\int f(\eta)d\eta \equiv U(\eta) .$$

Узявши до уваги означення змінних η та ξ, співвідношення (8.67) можна виразити у такій формі:

$$E(x,t) = U(x-ct) + V(x+ct) . \qquad (8.68)$$

Дуже важливо, що співвідношення (8.68) було виведене з рівняння (8.66) без накладання умов або обмежень на функцію $E(x,t)$, а отже, ця функція задовольняє хвильове рівняння, якими б не були функції $U(x-ct)$ та $V(x+ct)$ (треба лише щоб існували похідні $\partial^2 E / \partial\xi^2$, $\partial^2 E / \partial\eta^2$ та $\partial^2 E / \partial\xi\partial\eta$, а ця властивість притаманна фізичним полям, оскільки швидкість їхньої зміни у часі та просторі є неперервною функцією координат та часу). Це означає, що сукупність усіх функцій $E(x,t)$, які виражаються співвідношеннями (8.68), є загальним інтегралом рівняння (8.66).

У випадку, коли функція $V(x+ct)$ тотожно дорівнює нулю, поле $E(x,t)$ є сталою величиною, якщо є сталою різниця $x-ct$. Як відомо, $x-ct=\text{const}$ є рівнянням площини, перпендикулярної до координатної осі *OX*. Із плином часу ця площина рухається від менших до більших значень просторової координати $x=\text{const}+ct$. Отже, якщо поле певним чином розподілене в просторі, цей розподіл рухається від менших до більших значень координат. Кажуть, що

$$E(x,t)=U(x-ct) \qquad (8.69)$$

– це хвиля, що рухається в напрямку збільшення просторової координати x. Її називають *прямою хвилею*. Аналогічно доходимо висновку, що

$$E(x,t)=V(x+ct) \qquad (8.70)$$

– це хвиля, що рухається в протилежному напрямку. Її називають *оберненою хвилею*.

Висновок 8.3. Загальний інтеграл *хвильового рівняння* (8.66) описує суперпозицією прямої та оберненої хвилі і при цьому, просторовий розподіл поля хвилі може бути дуже різним (математично – будь-якою функцією, змінних η та ξ, для якої існують другі похідні за цими змінними).

Зазвичай хвильові процеси асоціюються з *гармонічними* коливаннями (пружними, електромагнітними тощо), які поширюються в просторі. Але в природі існує безліч різноманітних звуків і відтінків кольорів. Зроблений висновок пояснює цей факт і ще багато практично важливих фактів, але не зменшує важливості вивчення гармонічних коливань.

Із теорії рядів Фур'є відомо, що навіть складний розподіл поля у хвилі можна наближено представити суперпозицією гармонічних функцій. Більше того, утворення та поширення гармонічних коливань є процесами, що важливі для функціювання багатьох інженерних пристроїв. Тому розглянемо суперпозицію прямих та обернених гармонічних коливань

$$E(x,t) = a\cos(kx - \omega t) + b\cos(kx + \omega t)\,, \tag{8.71}$$

де $k = \text{const}$. Згідно з наведеним висновком гармонічні коливання зможуть поширюватися в просторі лише в тому випадку, якщо величини $kx \pm \omega t$ будуть пропорційними $x \pm ct$, тобто, якщо $\omega / k = c$, тобто, якщо справджується рівняння

$$\omega = ck\,, \tag{8.72}$$

яке за традицією називають *дисперсійним рівнянням* хвиль типу (8.71). За такою термінологією це рівняння виражає собою *лінійний за параметром* k *закон дисперсії* хвиль, зазвичай електромагнітних або пружних.

Параметр k *називають хвильовим числом*. Щоб зрозуміти, як пов'язаний цей параметр із загальновідомими характеристиками хвиль, достатньо зазначити, що поле $E(x,t)$ не зміниться, якщо додати до x величину $\Delta x = 2\pi / k$, тому Δx – це довжина хвилі λ. Звідси випливає просте, але важливе співвідношення

$$k = \frac{2\pi}{\lambda}\,. \tag{8.73}$$

Підставивши цей вираз хвильового числа до дисперсійного рівняння, знаходимо співвідношення між довжиною хвилі, кутовою частотою ω та частотою коливань ν:

$$\lambda = \frac{2\pi c}{\omega} = \frac{c}{\nu}\,. \tag{8.74}$$

Частота коливань пов'язана з кутовою частотою співвідношенням $\omega = 2\pi\nu$.

Висновок 8.4. Дисперсійне рівняння показує, що, *чим більшою буде задана конструкцією інженерного пристрою частота хвилі, тим меншою буде її довжина, або, чим більшою буде задана довжина хвилі, тим меншою буде її частота.*

Дійти висновку 8.4 легко і в інший спосіб, а саме, підставивши функцію (8.71) до хвильового рівняння (8.66). Після такої підстановки це рівняння набуває форми

$$[a\cos(kx-\omega t)+b\cos(kx-\omega t)](k^2-\omega^2/c^2)=0$$

та перетворюється на тотожність, якщо справджується закон дисперсії (8.72).

Поряд із хвилями, що поширюються в просторі, на практиці часто утворюються та використовуються так звані *стоячі хвилі.* Щоб описати стоячу хвилю математично, розглянемо поле, яке утворюється, коли одна й та сама хвиля періодично відбивається від перешкод, розташованих на відстані $\Delta x=\lambda/2$ одна від одної:

$$E(x,t)=a\left[\cos(kx-\omega t)+\cos(kx+\frac{k\lambda}{2}+\omega t)\right].$$

Із рівняння (8.73) випливає, що $k\lambda/2=\pi$, а тому

$$E(x,t)=a\left[\cos(kx-\omega t)-\cos(kx+\omega t)\right]=2a\sin kx\sin\omega t\,.$$

Увівши позначення

$$A(x)\equiv 2a\sin kx\,, \tag{8.75}$$

виражаємо поле хвилі у формі

$$E(x,t)=A(x)\sin\omega t\,. \tag{8.76}$$

Із рівнянь (8.75), (8.76) випливає, що в точках відбиття хвилі (точках із координатами $x=0$ та $x=\lambda/2$) поле $E(x,t)$ дорівнює нулю, а при віддаленні від цих точок воно зростає і сягає найбільшого значення посередині відрізка між точками відбиття. Якщо така хвиля відображується на екрані осцилографа, а її період менший за час реакції сітківки ока людини на кванти світла, спостерігач побачить на дисплеї нерухомий образ, який цілком природно назвати стоячою хвилею. Саме стоячі хвилі утворюються в приладах, які називають *резонаторами*. Вище розглянуто найдовшу зі стоячих хвиль, довжина якої λ дорівнює подвоєній довжині резонатора. Таку хвилю називають *першою резонансною гармонікою*. Хвилі довжиною $\lambda/2$, $\lambda/3$,… також можуть збуджуватися в резонаторах різних типів. Такі хвилі називають *вищими резонансними гармоніками*.

РОЗДІЛ 9
Застосування лінійної алгебри для інтегрування систем лінійних диференціальних рівнянь першого порядку

9.1. Нормальні та автономні системи рівнянь

Системи диференціальних рівнянь 1-го порядку вперше виникли при розв'язуванні задач механіки, насамперед задачі про визначення траєкторій рухомих фізичних тіл (планет, гарматних ядер тощо). Функціями, що підлягають визначенню, були координати тіл, а змінною, від якої залежать ці функції, був час t. Тому в більшості підручників для функцій, що визначаються із систем ДР, використовується позначення $x_i(t)$, де індекс $i=1,2,...n$ є номером функції. Диференціювання функції за змінною t зазвичай позначають крапкою над функцією. У цьому розділі для невідомих функцій використано позначення $y_i(t)$, яке у випадку $n=1$ перетворюється на прийняте в попередньому розділі позначення невідомої функції $y(t)$ і природним чином пов'язує цей розділ з попереднім.

Означення 9.1. Нормальною системою диференціальних рівнянь 1-го порядку називають систему, усі рівняння якої розв'язані відносно похідної:

$$\begin{cases} \dot{y}_1 = F_1(y_1, y_2, \ldots, y_n; t) \\ \dot{y}_2 = F_2(y_1, y_2, \ldots, y_n; t) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dot{y}_n = F_n(y_1, y_2, \ldots, y_n; t). \end{cases}$$

Для зручності використовують скорочені форми запису нормальної системи

$$\dot{y}_i = F_i(y_j;t), \;\; i,j = 1,2,...n\,, \tag{9.1}$$

або

$$\dot{\mathbf{y}} = \mathbf{F}(\mathbf{y};t)\,, \tag{9.2}$$

де $\mathbf{y}$ та $\mathbf{F}$ – це вектор-стовпці:

$$\mathbf{y} = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}; \;\; \mathbf{F} = \begin{pmatrix} F_1(\mathbf{y};t) \\ F_2(\mathbf{y};t) \\ \vdots \\ F_n(\mathbf{y};t) \end{pmatrix}. \tag{9.3}$$

Означення 9.2. Розв'язком системи диференціальних рівнянь (9.1) називають сукупність n функцій, підстановка яких до рівнянь перетворює їх на тотожності; сукупність усіх розв'язків називають загальним інтегралом системи диференціальних рівнянь.

Кількість похідних від функцій $y_i(t)$, що входять до системи рівнянь (9.1), дорівнює кількості рівнянь n. Щоб знайти з рівнянь ці функції, треба обчислити n невизначених інтегралів, тому загальний інтеграл системи містить n довільних сталих. Зафіксувавши величини довільних сталих, отримаємо *частинний розв'язок* системи ДР.

Означення 9.3. Простір векторів $\mathbf{y}$ називають фазовим простором системи диференціальних рівнянь 1-го порядку.

Зауваження 9.1. Класична механіка вивчає рухи сукупності матеріальних точок у тривимірному просторі. Якщо кількість точок дорівнює N, фазовим простором називають $6N$-вимірний простір, координатами елементів якого є координати радіус-векторів та швидкостей (або імпульсів) цих матеріальних точок. Сенс такої назви можна пояснити простим прикладом.

● ***Приклад 9.1.*** Розглянемо систему рівнянь:

$$\begin{cases} \dot{y}_1 = -2\gamma y_1 + y_2 \\ \dot{y}_2 = -\omega_0^2 y_1. \end{cases} \tag{9.4}$$

Продиференціювавши обидві частини 1-го рівняння за змінною t, одержуємо лінійне ДР другого порядку $\ddot{y}_1 = -2\gamma\dot{y}_1 + \dot{y}_2$. Підставивши до цього рівняння похідну $\dot{y}_2$ із 2-го рівняння системи (9.4), доходимо висновку, що ця система еквівалентна рівнянню

$$\ddot{y}_1 + 2\gamma\dot{y}_1 + \omega_0^2 y_1 = 0.$$

Якщо змінна t є часом, а функція $y_1(t)$ – координатою матеріальної точки, то це рівняння описує вільні одновимірні коливання (див. п. 8.3.3). З погляду механіки, вектори з координатами (y_1, y_2) є елементами двовимірного фазового простору матеріальної точки, яка здійснює одновимірні коливання, а згідно з означенням 9.3, ці вектори є елементами двовимірного фазового простору системи рівнянь (9.4).●

Зауваження 9.2. Можливість описати динаміку механічної системи або диференціальними рівняннями 2-го порядку, або системами диференціальних рівнянь 1-го порядку, породила два підходи до класичної механіки, які називають формалізмом Лагранжа та формалізмом Гамільтона. Формалізм Лагранжа базується на диференціальних рівняннях 2-го порядку, а формалізм Гамільтона – на системах ДР 1-го порядку.

Означення 9.4. Систему диференціальних рівнянь

$$\dot{y}_i = F_i(y_j) \tag{9.5}$$

називають *автономною*.

Ознакою автономної системи є те, що всі функції $F_i(y_j)$ залежать від часу t лише через функції $y_j(t)$, тобто змінна t не входить до правих частин рівнянь явно. Назва "автономна система" походить з механіки: автономними системами рівнянь описується рух замкненої системи фізичних тіл, що взаємодіють лише між собою, тобто їхній рух відбувається автономно (незалежно від інших систем фізичних тіл).

Розглянемо простір, координатами елементів якого є значення змінної t та значення функцій $y_i(t)$, $i = 1, 2, ...n$. Вимірність цього простору перевищує кількість рівнянь нормальної системи на одиницю. Інтегральні криві системи рівнянь є кривими в цьому $n+1$-вимірному просторі. Проєкції цих кривих на фазовий простір називають *фазовими траєкторіями*.

● ***Приклад 9.2.*** Розглянемо поняття фазової траєкторії на прикладі системи рівнянь (9.4), для простоти поклавши в ній $\gamma = 0$. Оскільки ця система еквівалентна рівнянню вільних коливань матеріальної точки, то з рівняння (8.36) випливає, що її загальним інтегралом є функції

$$y_1 = a\cos(\omega t + \alpha),$$
$$y_2 = -a\omega\sin(\omega t + \alpha),$$

а фазові траєкторії є еліпсами

$$\frac{y_1^2}{a^2} + \frac{y_2^2}{a^2\omega^2} = 1, \tag{9.6}$$

побудованими на площині, координатами точок якої є значення функцій $y_1(t)$ та $y_2(t)$ у фіксовані моменти часу. У тому, що функції $y_1(t), y_2(t)$ задовольняють рівняння (9.6), легко впевнитися, підставивши функції y_1 та y_2 до цього рівняння і врахувавши тригонометричну тотожність

$$\sin^2(\omega t + \alpha) + \cos^2(\omega t + \alpha) = 1.$$

Рівняння (9.6) встановлює залежність між функціями $y_1(t)$ та $y_2(t)$. З одного боку, ці функції є координатами вектора, що належить фазовому простору, а з іншого, – координатами точки, яка рухається на площині (y_1, y_2) уздовж еліпса (9.6). Саме цей еліпс і є фазовою траєкторію рухомої точки.●

Означення 9.5. Задачею Коші для нормальної системи диференціальних рівнянь називають задачу відшукання такого розв'язку системи, який задовольняє задані початкові умови. Формальним означенням задачі Коші є система рівнянь

$$\begin{cases} \dot{y}_i = f_i(y_i;t) \\ y_i(t_0) = y_{i0}, \quad i = 1,2...n, \end{cases} \tag{9.7}$$

де y_{i0} – числа, задані в умові задачі.

Задачу Коші можна задати також і у векторній формі

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{F}(\mathbf{y}; t) \\ \mathbf{y}(t_0) = \mathbf{y}_0, \end{cases} \tag{9.8}$$

де

$$\mathbf{y}_0 = \begin{pmatrix} y_{10} \\ y_{20} \\ \vdots \\ y_{n0} \end{pmatrix}, \tag{9.9}$$

вектори $\mathbf{y}$ та $\mathbf{F}$ означені виразами (9.3).

9.2. Системи лінійних диференціальних рівнянь першого порядку

9.2.1.Основні означення

Означення 9.6. Лінійною системою диференціальних рівнянь першого порядку називають систему[1]

$$\begin{cases} \dot{y}_1 = a_{11}(t)y_1 + a_{12}(t)y_2 + \ldots + a_{1n}(t) + f_1(t), \\ \dot{y}_2 = a_{21}(t)y_1 + a_{22}(t)y_2 + \ldots + a_{2n}(t) + f_2(t), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dot{y}_n = a_{n1}(t)y_1 + a_{n2}(t)y_2 + \ldots + a_{nn}(t) + f_n(t). \end{cases} \tag{9.10}$$

Систему (9.10) називають лінійною, тому що її можна записати в матричній формі:

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}, \tag{9.11}$$

ввести до розгляду *лінійний* оператор

[1] Тут і далі розглянуто системи, у яких кількість рівнянь дорівнює кількості невідомих функцій.

$$\hat{L} \equiv \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \frac{d}{dt} - \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix} \tag{9.12}$$

та тотожно перетворити матричне рівняння (9.10) на операторне

$$\hat{L}\mathbf{y}(t) = \mathbf{f}(t), \tag{9.13}$$

де

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \ \mathbf{f}(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}. \tag{9.14}$$

Означення 9.7. Систему лінійних диференціальних рівнянь

$$\hat{L}\mathbf{y}(t) = \mathbf{0} \tag{9.15}$$

називають *однорідною*.

Якщо замінити вектор-стовпець $\mathbf{y}(t)$ на функцію $y(t)$, операторне рівняння (9.15) перетвориться на однорідне лінійне рівняння (8.9). Цей факт зумовлює можливість поширення алгебраїчного підходу до розв'язання однорідних лінійних ДР на системи однорідних лінійних рівнянь.

9.2.2. Алгебраїчний підхід до розв'язання систем однорідних лінійних диференціальних рівнянь

Оскільки вектор-стовпці (9.14) є елементами n-вимірного лінійного простору, а оператор $\hat{L}$ – лінійним оператором, то для однорідних лінійних систем справедливий принцип суперпозиції розв'язків. Згідно з цим принципом будь-яка лінійна комбінація частинних розв'язків однорідної лінійної системи є її розв'язком.

Доведення справедливості принципу суперпозиції розв'язків для систем лінійних ДР подібне доведенню, яке описане в підрозд. 8.2 для частинних розв'язків одного лінійного рівняння. Достатньо виконати це доведення для двох розв'язків системи двох однорідних лінійних рівнянь. Щоб зробити це, зауважимо, що згідно з означенням частинного розв'язку підстановка частинних розв'язків $\mathbf{y}_1$ та $\mathbf{y}_2$ до системи (9.15) перетворює її на тотожності

$$\hat{L}\mathbf{y}_1 \equiv 0 \text{ та } \hat{L}\mathbf{y}_2 \equiv 0.$$

Подіємо оператором $\hat{L}$ на лінійну комбінацію розв'язків:

$$\hat{L}(C_1\mathbf{y}_1 + C_2\mathbf{y}_2) = \hat{L}C_1\mathbf{y}_1 + \hat{L}C_2\mathbf{y}_2 = C_1\hat{L}\mathbf{y}_1 + C_2\hat{L}\mathbf{y}_2.$$

Звідси випливає тотожність

$$\hat{L}(C_1\mathbf{y}_1 + C_2\mathbf{y}_2) \equiv 0, \tag{9.16}$$

яка доводить справедливість принципу суперпозиції розв'язків для систем однорідних лінійних ДР. Узагальнення на випадок більшої кількості розв'язків виконується збільшенням кількості доданків у їхній лінійній комбінації.

Так само, як і у випадку одного рівняння, із принципу суперпозиції виникає можливість утворити загальний інтеграл однорідної лінійної системи у вигляді лінійної комбінації частинних розв'язків:

$$\mathbf{y}(t) = \sum_{\alpha=1}^{p} C_\alpha \mathbf{y}_\alpha(t), \tag{9.17}$$

де C_α – довільні сталі, а $\mathbf{y}(t)$ та $\mathbf{y}_\alpha(t)$ є вектор-стовпцями, кожен з яких складається з n функцій:

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \ \mathbf{y}_\alpha(t) = \begin{pmatrix} y_{1\alpha}(t) \\ \vdots \\ y_{n\alpha}(t) \end{pmatrix}. \tag{9.18}$$

Функції $y_{i\alpha}(t)$ позначені парою індексів $i\alpha$; перший індекс – це номер функції, яка входить до системи рівнянь (9.15) і підлягає визначенню, а другий індекс – це номер частинного розв'язку в лінійній комбінації вектор-стовпців (9.17). Лінійну комбінацію вектор-стовпців (9.17) можна представити у формі лінійних комбінацій функцій $y_{i\alpha}(t)$:

$$y_i(t) = \sum_{\alpha=1}^{p} C_\alpha y_{i\alpha}(t). \tag{9.19}$$

Доведемо, що лінійна комбінація частинних розв'язків є загальним інтегралом лише тоді, коли ці частинні розв'язки лінійно незалежні.

Доведення. Виходимо з того, що загальний інтеграл має містити розв'язки задачі Коші для системи лінійних однорідних рівнянь (9.10). Ця задача полягає у відшуканні такого розв'язку системи, який задовольняє початкові умови

$$y_1(t_0) = y_{10},\, y_2(t_0) = y_{20}, \ldots, y_n(t_0) = y_{n0}\,;\; y_{\alpha 0} = \text{const}\,. \quad (9.20)$$

Принцип суперпозиції гарантує, що функції (9.19) задовольняють систему рівнянь, і залишається лише з'ясувати, за яких значень сталих C_α вони задовольняють початкові умови. Підставивши функції (9.19) до початкових умов (9.20), одержуємо систему неоднорідних алгебраїчних рівнянь для невідомих C_α:

$$\begin{cases} y_{11}(t_0)C_1 + y_{12}(t_0)C_2 + \ldots + y_{1n}(t_0)C_2 = y_{10}, \\ y_{21}(t_0)C_1 + y_{22}(t_0)C_2 + \ldots + y_{2n}(t_0)C_n = y_{20}, \\ \\ y_{n1}(t_0)C_1 + y_{n2}(t_0)C_2 + \ldots + y_{nn}(t_0)C_n = y_{n0}. \end{cases} \quad (9.21)$$

Розв'язок цієї системи можна знайти, скориставшись формулами Крамера, справедливість яких доведена в підрозд. 5.3:

$$C_\alpha = \frac{\tilde{\Delta}^{(\alpha)}}{W}\,, \quad (9.22)$$

де

$$W \equiv \begin{vmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{vmatrix}, \quad (9.23)$$

$\tilde{\Delta}^{(\alpha)}$ є визначниками, у кожному з яких стовпець визначника W з номером α замінений на стовпець, елементами якого є праві частини рівнянь (9.21).

Визначник W називають визначником Вронського системи лінійних ДР.

Формули (9.22) показують, що функція (9.19) може задовольнити початкові умови задачі Коші лише за умови, що визначник Вронського не дорівнює нулю. Оскільки загальний інтеграл має містити всі частинні розв'язки, то умова $W \neq 0$ має виконуватись за різних значень змінної t, тобто у формулі (9.23) замість фіксованих величин $y_{i\alpha}(t_0)$ містяться функції $y_{i\alpha}(t)$. Визначник Вронського не дорівнює нулю лише тоді, коли вектор-стовпці $\mathbf{y}_\alpha$ – лінійно незалежні[1]. Саме це і треба було довести.

Висновок 9.1. Щоб зінтегрувати лінійну однорідну систему ДР, що складається з n незалежних рівнянь, достатньо знайти n лінійно незалежних частинних розв'язків[2]. Їх називають *фундаментальними розв'язками* системи однорідних лінійних диференціальних рівнянь.

Для простого пояснення алгебраїчного методу розв'язання однорідних лінійних систем доцільно розглянути системи двох ДР зі сталими коефіцієнтами.

9.2.3. Алгебраїчний метод інтегрування систем однорідних лінійних рівнянь зі сталими коефіцієнтами

Розглянемо систему двох однорідних лінійних ДР зі сталими коефіцієнтами

$$\begin{cases} \dot{y}_1 = a_{11}y_1 + a_{12}y_2, \\ \dot{y}_2 = a_{21}y_1 + a_{22}y_2. \end{cases} \tag{9.24}$$

Запишемо її у матричній формі

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{9.25}$$

Підставивши до рівняння (9.25) пробні функції

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} e^{kt} \tag{9.26}$$

---

[1] Як було доведено в п. 4.5.3, рівність визначника нулю є критерієм лінійної залежності його стовпців.

[2] Той факт, що загальний інтеграл є лінійною комбінацією n лінійно незалежних частинних розв'язків, вказує на те, що частинні розв'язки є елементами n-вимірного лінійного простору.

зі сталими коефіцієнтами λ_1, λ_2, одержимо матричне рівняння

$$k\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix},$$

яке тотожно перетворюється на рівняння

$$\left[\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}-k\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right]\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=0$$

і на систему однорідних алгебраїчних рівнянь для невідомих λ_1, λ_2:

$$\begin{cases}(a_{11}-k)\lambda_1+a_{12}\lambda_2=0,\\ a_{21}\lambda_1+(a_{22}-k)\lambda_2=0.\end{cases}\tag{9.27}$$

Як уже було зазначено, така система має нетривіальний розв'язок тоді і лише тоді, коли її визначник дорівнює нулю, тобто за умови, що

$$\begin{vmatrix}a_{11}-k & a_{12}\\ a_{21} & a_{22}-k\end{vmatrix}=0.\tag{9.28}$$

Рівняння (9.28) є квадратним алгебраїчним рівнянням для відшукання таких чисел $k=k_\alpha$ ($\alpha=1,2$), за яких функції

$$\begin{pmatrix}y_{1\alpha}\\ y_{2\alpha}\end{pmatrix}=\begin{pmatrix}\lambda_{1\alpha}\\ \lambda_{2\alpha}\end{pmatrix}e^{k_\alpha t},\ \ (\alpha=1,2)\tag{9.29}$$

будуть задовольняти задану систему однорідних лінійних ДР зі сталими коефіцієнтами.

Функції (9.29) є фундаментальними розв'язками системи однорідних лінійних диференціальних рівнянь (9.24). Загальний інтеграл цієї системи є лінійною комбінацією фундаментальних розв'язків

$$\begin{pmatrix}y_1\\ y_2\end{pmatrix}_{\text{з.о.}}=\begin{pmatrix}\lambda_{11}\\ \lambda_{21}\end{pmatrix}e^{k_1 t}+\begin{pmatrix}\lambda_{12}\\ \lambda_{22}\end{pmatrix}e^{k_2 t}.\tag{9.30}$$

Формула (9.30) показує, що загальний інтеграл однорідної системи визначається (характеризується) тим, які корені має рівняння (9.28), тому його називають *характеристичним* *рівнянням*. Щоб знайти загальний інтеграл (9.30), треба виконати такі дії:

а) знайти корені характеристичного рівняння $k_{1,2}$;

б) знайти функції (9.29), почергово підставивши до системи (9.27) знайдені k_1 та k_2.

При цьому можуть виникнути три різні випадки:

1) $k_1 \neq k_2$, дійсні корені рівняння (9.28);

2) $k_1 = k_2$, кратний дійсний корінь;

3) $k_1 = \gamma + i\omega$, $k_2 = \gamma - i\omega$, комплексно-спряжені корені.

Зауваження 9.3. У підручниках з диференціальних рівнянь доводиться, що у випадку кратного кореня характеристичного рівняння загальний інтеграл однорідної системи має форму

$$x = (C_1 + C_2 t)e^{k_1 t},$$

$$y = (l_1 + l_2 t)e^{k_1 t}.$$

Цей факт легко зрозуміти виходячи з того, що загальний інтеграл має бути лінійною комбінацією двох лінійно незалежних частинних розв'язків, а у випадку кратного кореня характеристичному рівнянню задовольняє лише одна експоненціальна функція. Тому в підручниках доводиться, що існують такі лінійні комбінації l_1, l_2 довільних сталих C_1, C_2, що вектор-стовпчик, пропорційний добутку $te^{k_1 t}$, задовольняє матричне рівняння (9.25).

Зауваження 9.4. У випадку комплексних коренів характеристичного рівняння із частинних розв'язків (9.29) слід скласти дійсний загальний інтеграл. Найлегше зробити це, скориставшись формулою

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\text{з.о.}} = \operatorname{Re}\begin{pmatrix} y_{11} \\ y_{21} \end{pmatrix} + \operatorname{Im}\begin{pmatrix} y_{12} \\ y_{22} \end{pmatrix}. \tag{9.31}$$

Зауваження 9.5. Оскільки визначник (9.28) алгебраїчної системи (9.27) дорівнює нулю, то після підстановки до цієї системи характеристичного числа k_1 або k_2 її рівняння виявляються лінійно залежними (еквівалентними), і з них вдається знайти лише відношення $\lambda_{21}/\lambda_{11}$ (підстановкою до них $k = k_1$) та $\lambda_{22}/\lambda_{12}$ (підстановкою $k = k_2$). При цьому величина одного коефіцієнта з кожної пари чисел ($(\lambda_{11}, \lambda_{21})$ та $(\lambda_{22}, \lambda_{12})$) може набувати довільних значень, тобто є довільною сталою C_1 та C_2 (відповідно), що входить до загального інтеграла системи ДР.

Пояснімо процедуру розв'язання системи однорідних лінійних ДР прикладами, узявши до уваги, що автори більшості задачників із диференціальних рівнянь використовують позначення $y_1 \equiv x$, $y_2 \equiv y$.

● ***Приклад 9.3.*** Зінтегруємо однорідну лінійну систему

$$\begin{cases} \dot{x} = x + 2y, \\ \dot{y} = 2x - 2y. \end{cases} \tag{9.32}$$

Для цього розглянемо матрицю коефіцієнтів цієї системи

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$$

і розв'яжемо характеристичне рівняння

$$\begin{vmatrix} 1-k & 2 \\ 2 & -2-k \end{vmatrix} = 0,$$

яке є алгебраїчним рівнянням 2-го порядку $k^2 + k - 6 = 0$ і має дійсні корені $k_1 = 2$ та $k_2 = -3$.

Знайдемо розв'язок, що відповідає кореню $k = 2$ з матричного рівняння

$$\begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \end{pmatrix} = 0,$$

яке тотожне системі алгебраїчних рівнянь

$$\begin{cases} -\lambda_{11} + 2\lambda_{21} = 0, \\ 2\lambda_{11} - 4\lambda_{21} = 0. \end{cases}$$

Ця система дозволяє знайти зв'язок між коефіцієнтами $\lambda_{11} = 2\lambda_{21} \equiv 2C_1$ і перший фундаментальний розв'язок

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 2C_1 \\ C_1 \end{pmatrix} e^{2t}.$$

Знайдемо розв'язок, що відповідає кореню $k = -3$ із матричного рівняння

$$\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \lambda_{12} \\ \lambda_{22} \end{pmatrix} = 0 ,$$

тотожного системі рівнянь

$$\begin{cases} 4\lambda_{12} + 2\lambda_{22} = 0, \\ 2\lambda_{12} + \lambda_{22} = 0. \end{cases}$$

Ця система дає змогу знайти співвідношення $\lambda_{22} = -2\lambda_{12} \equiv -2C_2$ та другий фундаментальний розв'язок

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} C_2 \\ -2C_2 \end{pmatrix} e^{-3t} .$$

Знайдемо суму фундаментальних розв'язків:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2C_1 \\ C_1 \end{pmatrix} e^{2t} + \begin{pmatrix} C_2 \\ -2C_2 \end{pmatrix} e^{-3t} .$$

Виразимо з неї загальний інтеграл системи (9.32):

$$x = 2C_1 e^{2t} + C_2 e^{-3t},$$
$$y = C_1 e^{2t} - 2C_2 e^{-3t}. \bullet$$

● ***Приклад 9.4.*** Зінтегруємо систему однорідних лінійних рівнянь

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -x, \end{cases} \tag{9.33}$$

діючи так само, як у попередньому прикладі. Спочатку розглянемо матрицю коефіцієнтів цієї системи

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

сформулюємо характеристичне рівняння

$$\begin{vmatrix} -k & 1 \\ -1 & -k \end{vmatrix} = 0$$

та рівняння для визначення коефіцієнтів $\lambda_{i\alpha}$:

$$\begin{cases} -k\lambda_1 + \lambda_2 = 0, \\ -\lambda_1 - k\lambda_2 = 0. \end{cases}$$

Потім обчислимо корені характеристичного рівняння $k_1 = i$, $k_2 = -i$ та співвідношення між коефіцієнтами $\lambda_{21} = i\lambda_{11} \equiv iC_1$, $\lambda_{22} = -i\lambda_{12} \equiv iC_2$. Визначимо комплексні розв'язки

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \end{pmatrix} e^{k_1 t} = \begin{pmatrix} C_1 \\ iC_1 \end{pmatrix} e^{it},$$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \lambda_{12} \\ \lambda_{22} \end{pmatrix} e^{-it} = \begin{pmatrix} -C_2 \\ iC_2 \end{pmatrix} e^{-it},$$

та утворимо з них дійсний загальний інтеграл, використавши для цього формулу (9.31), яка в цьому випадку дає:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathrm{Re}\begin{pmatrix} C_1 \\ +iC_1 \end{pmatrix} e^{it} + \mathrm{Im}\begin{pmatrix} -C_2 \\ iC_2 \end{pmatrix} e^{-it} =$$

$$= \mathrm{Re}\begin{pmatrix} C_1 \\ iC_1 \end{pmatrix} (\cos t + i \sin t) + \mathrm{Im}\begin{pmatrix} -C_2 \\ iC_2 \end{pmatrix} (\cos t - i \sin t) =$$

$$= \begin{pmatrix} C_1 \cos t \\ -C_1 \sin t \end{pmatrix} + \begin{pmatrix} C_2 \sin t \\ C_2 \cos t \end{pmatrix}.$$

Ці співвідношення показують, що загальний інтеграл системи рівнянь (9.33) виражається рівняннями

$$x = C_1 \cos t + C_2 \sin t,$$
$$y = -C_1 \sin t + C_2 \cos t. \quad \bullet$$

Зауваження 9.6. Наведені приклади показують, що алгебраїчний метод інтегрування систем лінійних однорідних ДР зі сталими коефіцієнтами застосовний також і до систем, що містять більше двох рівнянь, але він потребує розв'язання характеристичного рівняння порядку $n > 2$, що може призвести до серйозних ускладнень.

9.2.4. Інтегрування системи неоднорідних лінійних рівнянь зі сталими коефіцієнтами методом подібності розв'язку до правих частин рівнянь

Метод подібності застосовують, якщо похідні $df_i(t)/dt$ правих частин обох рівнянь системи є функціями такого ж типу, як і $f_i(t)$, тобто, якщо праві частини рівняння є або експонентами, або лінійними комбінаціями синусів та косинусів, або поліномами. Вони можуть бути також або сумою таких функцій, або добутком полінома на експоненту чи полінома на лінійну комбінацію синусів та косинусів.

Метод подібності оснований на теоремі, згідно з якою *загальний інтеграл системи неоднорідних лінійних ДР є сумою загального інтеграла системи однорідних ДР та частинного розв'язку неоднорідної системи.* Для системи двох рівнянь твердження теореми виражається у формі рівняння для вектор-стовпців $\mathbf{y}(t)$, елементами яких є функції $y_1(t)$ та $y_2(t)$:

$$\mathbf{y}_{\text{з.н.}}(t) = \mathbf{y}_{\text{з.о.}}(t) + \mathbf{y}_{\text{ч.н.}}(t)\,,$$

тобто

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\text{з.н.}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\text{з.о.}} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\text{ч.н.}}. \tag{9.34}$$

У підрозд. 8.2 цю теорему було доведено для неоднорідного лінійного ДР, виходячи з операторної форми цього рівняння (8.5). Оскільки операторне рівняння (9.13) відрізняється від неоднорідного лінійного ДР 1-го порядку (8.5) лише заміною функцій $y(x)$ та $b(x)$ на стовпці $\mathbf{y}(t)$ та $\mathbf{f}(t)$, то ця теорема доводиться для системи рівнянь так само, як вона була доведена в підрозд. 8.2.

Зважаючи на вказану теорему, щоб знайти загальний інтеграл неоднорідної лінійної системи ДР необхідно:

а) відкинути праві частини рівнянь і розв'язати характеристичне рівняння лінійної системи однорідних ДР;
б) знайти загальний інтеграл системи однорідних ДР;
в) знайти частинний розв'язок системи неоднорідних ДР;
г) додати до загального інтеграла системи однорідних ДР знайдений частинний розв'язок системи неоднорідних ДР.

Проілюструємо описану процедуру інтегрування системи неоднорідних лінійних ДР простим прикладом, позначивши, як і в попередніх прикладах $y_1 \equiv x$, $y_2 \equiv y$.

● ***Приклад 9.5.*** Зінтегруємо систему рівнянь:

$$\begin{cases} \dot{x} = y, \\ \dot{y} = x + \cos t. \end{cases} \tag{9.35}$$

Для цього виконаємо такі кроки:

а) відкинемо праву частину другого рівняння та розв'яжемо характеристичне рівняння

$$\begin{vmatrix} -k & 1 \\ 1 & -k \end{vmatrix} = 0,$$

$$k^2 - 1 = 0, \ k_1 = 1, \ k_2 = -1;$$

б) знайдемо загальний інтеграл системи однорідних ДР:

$$\begin{pmatrix} -k & 1 \\ 1 & -k \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0,$$

$$\begin{cases} -k\lambda_1 + \lambda_2 = 0, \\ \lambda_1 - k\lambda_2 = 0; \end{cases}$$

$k_1 = 1$:

$$-\lambda_{11} + \lambda_{21} = 0, \quad \lambda_{11} = \lambda_{21} \equiv C_1,$$

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} C_1 \\ C_1 \end{pmatrix} e^t;$$

$k_2 = -1$:

$$\lambda_{12} + \lambda_{22} = 0, \quad \lambda_{12} = -\lambda_{22} \equiv C_2,$$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} C_2 \\ -C_2 \end{pmatrix} e^{-t},$$

$$x_{\text{з.о.}} = C_1 e^t + C_2 e^{-t}, \quad y_{\text{з.о.}} = C_1 e^t - C_2 e^{-t};$$

в) знайдемо частинний розв'язок неоднорідної системи ДР, підставивши до неї пробні функції, подібні до правої частини другого рівняння[1]:

$$x = A_1 \sin t + B_1 \cos t, \ \ y = A_2 \sin t + B_2 \cos t.$$

Знаходимо умови

$$A_1 \cos t - B_1 \sin t = A_2 \sin t + B_2 \cos t,$$

$$A_2 \cos t - B_2 \sin t = A_1 \sin t + (B_1 + 1) \cos t.$$

[1] Правою частиною другого рівняння системи є косинус змінної t, але до пробних функцій додано ще й синус, оскільки похідна від синуса дорівнює косинусу, а отже, вона дорівнює правій частині рівняння.

Ці умови мають виконуватись для всіх значень змінної t, а отже, мають дорівнювати один одному коефіцієнти як біля косинусів, так і біля синусів цієї змінної. Коефіцієнти біля косинусів дорівнюють один одному, якщо

$$A_1 = B_2 \text{ та } A_2 = B_1 + 1,$$

а біля синусів, якщо

$$-B_1 = A_2 \text{ та } -B_2 = A_1.$$

Отже, маємо систему чотирьох рівнянь для чотирьох невідомих, розв'язком якої є коефіцієнти $A_1 = B_2 = 0$, $A_2 = -B_1 = 1/2$. Частинним розв'язком системи неоднорідних ДР є функції

$$x_{\text{ч.н.}} = -\frac{1}{2}\cos t, \quad y_{\text{ч.н.}} = \frac{1}{2}\sin t.$$

Додамо до знайденого частинного розв'язку загальний інтеграл системи однорідних рівнянь, і в такий спосіб визначимо загальний інтеграл системи рівнянь (9.35):

$$x_{\text{з.н.}} = C_1 e^t + C_2 e^{-t} - \frac{1}{2}\cos t,$$

$$y_{\text{з.н.}} = C_1 e^t - C_2 e^{-t} + \frac{1}{2}\sin t.$$

Бачимо, що до знайденого загального інтеграла входить не лише косинус, але й синус змінної t, а отже, обмежуватися в пробних функціях лише косинусом було б помилкою.●

РОЗДІЛ 10
Дослідження стійкості стаціонарних станів динамічної системи в лінійному наближенні

10.1. Постановка задачі

Домовимось, що *динамічна система* (ДС) – це об'єкт або сукупність об'єктів фізичної, хімічної, соціальної або іншої природи, а *система рівнянь* – автономна система двох диференціальних рівнянь 1-го порядку для функцій $y_1(t)$ та $y_2(t)$:

$$\begin{cases} \dot{y}_1(t) = F_1(y_1, y_2), \\ \dot{y}_2(t) = F_2(y_1, y_2). \end{cases} \tag{10.1}$$

Розглянемо динамічну систему, зміна стану якої описується функціями $y_1(t)$ та $y_2(t)$. Якщо ДС може перебувати в незмінному стані, то система рівнянь має *стаціонарний* (тобто незалежний від часу) розв'язок

$$y_1 = y_{10} = \text{const}, \quad y_2 = y_{20} = \text{const}. \tag{10.2}$$

Припустимо, що ДС, яка перебуває в незмінному стані, у момент $t = 0$ потрапляє під раптовий короткотривалий вплив зовнішнього чинника (для ДС різної природи це може бути механічний поштовх, короткотривала зміна електричної напруги, спалах світла, втручання, що веде до раптової зміни інформаційного трафіку тощо). Під впливом такого чинника стан спокою системи порушується, вона починає рухатися, і сталі величини (10.2), що описують стан системи, замінюються на залежні від часу величини

$$y_1^* = y_{10} + \tilde{y}_1(t), \quad y_2^* = y_{20} + \tilde{y}_2(t). \tag{10.3}$$

Дуже часто короткотривалий вплив на ДС буває непередбачуваним, тому зумовлені цим впливом доданки $\tilde{y}_1(t)$ та $\tilde{y}_2(t)$ до стаціонарного розв'язку називають *випадковими збуреннями* величин y_{10} та y_{20}. У багатьох реальних випадках навіть невеликий зовнішній вплив може призвести до великих (і зазвичай небезпечних) наслідків. Це трапляється, якщо внаслідок певних властивостей ДС, а формально – для певних функцій $F_1(y_1, y_2)$ та $F_2(y_1, y_2)$, що описують ці властивості, малі за початковою величиною збурення $|\tilde{y}_1(0)| << |y_{10}|$ та/або $|\tilde{y}_2(0)| << |y_{20}|$ зростають із часом до величин, порівняних із y_{10} та/або y_{20}. У такому разі кажуть, що *стаціонарний стан* динамічної системи *нестійкий* до зовнішнього впливу, а стаціонарний розв'язок системи рівнянь нестійкий до малих випадкових збурень.

Підвищення стійкості ДС до зовнішніх впливів є одним із найважливіших завдань інженерії, а дослідження стійкості стаціонарних розв'язків системи рівнянь до малих збурень – однією з важливих задач математики.

10.2. Лінеаризація системи диференціальних рівнянь в околі точок спокою на її фазовій площині

Фазова площина системи диференціальних рівнянь (10.1) – це площина, координатами точок якої є значення функцій y_1, y_2. Якщо стан динамічної системи змінюється, розв'язок системи рівнянь зображується певною кривою $\Phi(y_1, y_2) = 0$ на фазовій площині. Цю криву називають *фазовою траєкторією* динамічної системи, тому що кожному "миттєвому" стану системи відповідає певна точка на кривій. Зручно вважати, що ця точка рухається вздовж фазової траєкторії і похідні $\dot{y}_1, \dot{y}_2$ є декартовими координатами вектора швидкості.

Точку, яка відповідає стаціонарному стану ДС, називають *точкою спокою*. Її координатами є стаціонарні розв'язки $y_1 = y_{10}$, $y_2 = y_{20}$, а її швидкість дорівнює нулю. Стаціонарні розв'язки слід знаходити із системи рівнянь (10.1), у якій покладено $\dot{y}_1 = \dot{y}_2 = 0$, тобто із системи рівнянь

$$\begin{cases} F_1(y_{10}, y_{20}) = 0, \\ F_2(y_{10}, y_{20}) = 0. \end{cases} \qquad (10.4)$$

У лінійному за малими збуреннями наближенні для дослідження стійкості точок спокою (а на практиці – стаціонарних станів ДС) використовують *лінеаризовану систему* диференціальних рівнянь. Точки спокою класифікують відповідно до того, які корені має характеристичне рівняння лінеаризованої системи рівнянь, і додатково розподіляють точки спокою кожного типу на стійкі та нестійкі до малих збурень. Щоб отримати лінеаризовану систему рівнянь, розглянемо функції $F_1(y_1, y_2)$ та $F_2(y_1, y_2)$, що стоять у системі рівнянь (9.36), вважаючи, що y_1, y_2 – це "збурені" функції (9.38), у яких збурення малі за величиною настільки, що $\left|\tilde{y}_{1,2}\right|^2 << \left|\tilde{y}_{1,2}\right|$. Наближено представимо $F_{1,2}(y_1, y_2)$ формулою Тейлора, урахувавши поряд з доданками, що не залежать від збурень, лише лінійні за збуреннями доданки[1]:

$$F_1(y_1^*, y_2^*) \approx F_1(y_{10}, y_{20}) + \left(\frac{\partial F_1}{\partial y_1^*}\right)_{\tilde{y}_{1,2}=0} \cdot \tilde{y}_1 + \left(\frac{\partial F_1}{\partial y_2^*}\right)_{\tilde{y}_{1,2}=0} \cdot \tilde{y}_2,$$
$$F_2(y_1^*, y_2^*) \approx F_2(y_{10}, y_{20}) + \left(\frac{\partial F_2}{\partial y_1^*}\right)_{\tilde{y}_{1,2}=0} \cdot \tilde{y}_1 + \left(\frac{\partial F_2}{\partial y_2^*}\right)_{\tilde{y}_{1,2}=0} \cdot \tilde{y}_2. \qquad (10.5)$$

Оскільки y_{10} та y_{20} сталі, то похідні у формулах (10.5) є константами, які доцільно позначити літерами з двома індексами:

$$a_{ij} \equiv \left(\frac{\partial F_i}{\partial y_j^*}\right)_{\tilde{y}_{1,2}=0}. \qquad (10.6)$$

Окрім того, слід зважити на рівняння (10.4). Після цього вирази (10.5) набудуть спрощеної форми

$$F_1(y_1^*, y_2^*) = a_{11}\tilde{y}_1 + a_{12}\tilde{y}_2, \quad F_2(y_1^*, y_2^*) = a_{21}\tilde{y}_1 + a_{22}\tilde{y}_2,$$

а система рівнянь (10.1) перетвориться на однорідну лінійну систему диференціальних рівнянь

$$\begin{cases} \dot{\tilde{y}}_1 = a_{11}\tilde{y}_1 + a_{12}\tilde{y}_2, \\ \dot{\tilde{y}}_2 = a_{21}\tilde{y}_1 + a_{22}\tilde{y}_2. \end{cases} \qquad (10.7)$$

---

[1] Саме в цьому полягає суть лінійного наближення.

(Оскільки $\dot{y}_{10} = \dot{y}_{20} = 0$, то $\dot{y}_1^* = \dot{\tilde{y}}_1$, $\dot{y}_2^* = \dot{\tilde{y}}_2$). Систему рівнянь (10.7) називають *лінеаризованою за малими збуреннями* $\tilde{y}_1$ та $\tilde{y}_2$, а дослідження стійкості ДС, основане на визначенні з цієї системи часової залежності малих збурень, – дослідженням стійкості ДС у лінійному наближенні.

10.3. Ознака нестійкості стаціонарного стану динамічної системи

Щоб спростити форму запису лінеаризованої системи рівнянь, перепозначимо малі збурення

$$\tilde{y}_1 \equiv x, \ \tilde{y}_2 \equiv y \tag{10.8}$$

і перепишемо вирази (10.7) у формі

$$\begin{cases} \dot{x} = a_{11}x + a_{12}y, \\ \dot{y} = a_{21}x + a_{22}y. \end{cases} \tag{10.9}$$

Згідно з цими позначеннями площина (x, y) є фазовою площиною лінеаризованої системи рівнянь (10.9).

У розд. 9 було показано, що загальний інтеграл системи однорідних лінійних ДР є сумою двох фундаментальних розв'язків[1]

$$\begin{pmatrix} x \\ y \end{pmatrix}_1 = \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \end{pmatrix} e^{k_1 t}, \quad \begin{pmatrix} x \\ y \end{pmatrix}_2 = \begin{pmatrix} \lambda_{12} \\ \lambda_{22} \end{pmatrix} e^{k_2 t}, \tag{10.10}$$

де k_1 та k_2 є коренями характеристичного рівняння (9.28), а коефіцієнти $\lambda_{i\alpha}$ із точністю до двох довільних сталих визначаються із системи алгебраїчних рівнянь (9.27). Загальний інтеграл лінеаризованої системи зручно виразити у формі двох рівнянь:

$$\begin{aligned} x &= \lambda_{11}e^{k_1 t} + \lambda_{12}e^{k_2 t}, \\ y &= \lambda_{21}e^{k_1 t} + \lambda_{22}e^{k_2 t}. \end{aligned} \tag{10.11}$$

---

[1] Нагадаємо, що *фундаментальними* називають лінійно незалежні частинні розв'язки системи однорідних лінійних ДР, кількість яких дорівнює кількості рівнянь у системі. Ці розв'язки є параметричним завданням функцій $y_1(x_1)$ та $y_2(x_2)$ через параметр t.

Із виразів (10.11) можна зробити важливі висновки.

Висновок 10.1. Якщо корені характеристичного рівняння дійсні, то фундаментальні розв'язки зображуються на фазовій площині прямими лініями, які називають *сепаратрисами*.

І справді, якщо корені характеристичного рівняння є дійсними числами, то і коефіцієнти $\lambda_{1\alpha}$ та $\lambda_{2\alpha}$ – також дійсні числа. Із виразів (10.11) випливають співвідношення $(y/x)_1 = \lambda_{21}/\lambda_{11}$ та $(y/x)_2 = \lambda_{22}/\lambda_{12}$, тому рівняння двох фундаментальних розв'язків є рівняннями двох прямих

$$y = h_1 x \text{ та } y = h_2 x \tag{10.12}$$

із кутовими коефіцієнтами $h_1 = \lambda_{21}/\lambda_{11}$, $h_2 = \lambda_{22}/\lambda_{12}$. Рівняння (10.12) є рівняннями сепаратрис.

Висновок 10.2. Загальний інтеграл лінеаризованої системи зображується сім'єю фазових траєкторій $y(x, C_1, C_2)$. Із системи алгебраїчних рівнянь (9.27) визначаються відношення коефіцієнтів $\lambda_{21}/\lambda_{11}$ та $\lambda_{22}/\lambda_{12}$, а два із чотирьох коефіцієнтів залишаються невизначеними[1]. Невизначені коефіцієнти є довільними сталими в загальному інтегралі лінеаризованої системи рівнянь: $C_1 \equiv \lambda_{11}$, $C_2 \equiv \lambda_{12}$.

У виразах (10.11) змінна t є часом, а x та y – малими збуреннями функцій $y_1(t)$ та $y_2(t)$. Якщо хоча б один дійсний корінь характеристичного рівняння є додатним числом або хоча б один комплексний корінь має додатну дійсну частину, то існують такі збурення, які експоненціально зростають із часом. На практиці, якщо зростає хоча б одне випадкове збурення, то науковці та інженери доходять висновку про *нестійкість* стаціонарного стану динамічної системи. (Оскільки випадкові збурення виникають непередбачувано, то може виникнути саме таке збурення, яке потім зростатиме з часом і спричинить певні наслідки).

[1] Корені характеристичного рівняння занулюють визначник системи рівнянь (9.27). Якщо визначник системи однорідних лінійних алгебраїчних рівнянь дорівнює нулю, то рівняння системи лінійно залежні, а отже, рівносильні.

Висновок 10.3. Наявність хоча б одного додатного розв'язку характеристичного рівняння або додатної дійсної частини хоча б в одному з комплексних розв'язків вважається *достатньою ознакою нестійкості* стаціонарного стану динамічної системи[1].

Зауваження 10.1. Якщо серед коренів характеристичного рівняння немає додатних, але є хоча б один такий, що дорівнює нулю, дослідження стійкості стаціонарного стану системи в лінійному наближенні не є достатнім і треба проводити додаткове, більш точне дослідження.

10.4. Класифікація точок спокою

При розв'язуванні деяких важливих задач недостатньо просто зробити висновок про стійкість або нестійкість стаціонарного стану динамічної системи, а треба ще й визначити, як проходять фазові траєкторії в околі різних точок спокою. Щоб пояснити, як це можна зробити, зауважимо, що точка спокою має координати $x = y = 0$, тому що вона відповідає стаціонарному стану, а в ньому збурення дорівнюють нулю (див. (10.2), (10.3)). Класифікація точок спокою базується на класифікації коренів характеристичного рівняння лінеаризованої системи за їхнім типом та величиною. Розглянемо почергово різні випадки.

1. Характеристичне рівняння має *два дійсних корені* $k_1 = \gamma_1$, $k_2 = \gamma_2$, і ці корені є числами *одного знака*. У цьому випадку загальний інтеграл визначається формулою

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \end{pmatrix} e^{\gamma_1 t} + \begin{pmatrix} \lambda_{12} \\ \lambda_{22} \end{pmatrix} e^{\gamma_2 t}, \qquad (10.13)$$

де відношення $h_1 = \lambda_{21}/\lambda_{11}$ та $h_2 = \lambda_{22}/\lambda_{12}$ визначаються із характеристичного рівняння. Корені характеристичного

[1] Цей висновок справедливий і в тих випадках, коли стан динамічної системи описується більш ніж двома диференціальними рівняннями, а характеристичне рівняння має більш ніж два корені.

рівняння завжди можна пронумерувати так, щоб виконувалась нерівність

$$\gamma_1 > \gamma_2 . \qquad (10.14)$$

Розглянемо рух точки з координатами $x(t)$, $y(t)$ уздовж фазових траєкторій лінеаризованої системи рівнянь на площині XOY, вважаючи, що *обидва корені характеристичного рівняння додатні*, а час t зростає від від'ємних значень до додатних. У далекому минулому показники експонент у рівності (10.13) набували великих за абсолютною величиною від'ємних значень, тому обидва доданки правої частини цієї рівності були малими за величиною, і перший доданок був менше другого. Унаслідок цього точка, що зображує збурення, була розташована біля точки спокою, і траєкторія її руху пролягала ближче до сепаратриси $y = h_2 x$, ніж до $y = h_1 x$, як показано на рис. 10.1. Напрямок руху точки, координати якої дорівнюють величинам випадкових збурень, зображено на рисунку стрілочками.

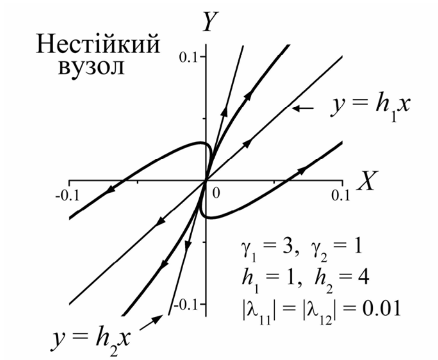

Рис. 10.1. Фазові траєкторії лінеаризованої системи рівнянь, розраховані для випадку двох додатних коренів характеристичного рівняння

Оскільки корені характеристичного рівняння додатні, то з плином часу обидва доданки в правій частині рівності (10.13) збільшуються і рухома точка віддаляється від точки спокою, яка лежить на перетині координатних осей. Унаслідок нерівності (10.14) перший доданок збільшується швидше, ніж другий, тому фазові траєкторії відхиляються від сепаратриси $y = h_2 x$, а їхні напрямки наближаються до напрямку сепаратриси $y = h_1 x$. Із рис. 10.1 очевидно, що випадкові збурення зростають із часом, тому стаціонарний стан динамічної системи нестійкий. Зазначимо, що фазові траєкторії, розташовані в різних квадрантах системи координат XOY, відповідають різним знакам довільних сталих λ_{11} та λ_{12}, а абсолютні величини цих сталих обрані так, щоб для зображених на рис. 10.1 фазових траєкторій виконувалися умови $|x|^2 << |x|$, $|y|^2 << |y|$, за яких була здійснена лінеаризація системи рівнянь.

Якщо *обидва корені характеристичного рівняння від'ємні*, то фазові траєкторії відрізняються від зображених на рис. 10.1 лише зміною напрямків усіх стрілочок (рис. 10.2). У цьому разі випадкові збурення експоненціально згасають із часом, тому стаціонарний стан динамічної системи стійкий.

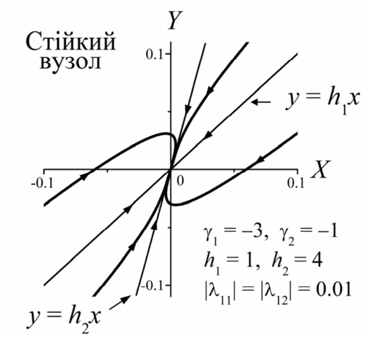

Рис. 10.2. Фазові траєкторії лінеаризованої системи рівнянь, розраховані для випадку двох від'ємних коренів характеристичного рівняння

Означення 10.1. Якщо характеристичне рівняння лінеаризованої системи має *два дійсних корені одного знака*, точку спокою називають *вузлом.* Вузол називають нестійким, якщо корені додатні, і стійким, коли вони від'ємні. Додатні корені характеристичного рівняння називають інкрементами зростання збурень, а від'ємні – декрементами загасання.

2. Характеристичне рівняння має *два дійсних корені* $k_1 = \gamma_1$, $k_2 = \gamma_2$, і вони є числами *різного знака*. Головне в цьому випадку те, що стаціонарний стан ДС нестійкий, оскільки один із коренів характеристичного рівняння додатний. Щоб з'ясувати, як у цьому випадку проходять фазові траєкторії, слід знову розглянути загальний інтеграл (10.13). З огляду на нерівність (10.14), корінь γ_1 додатний, а корінь γ_2 від'ємний. За від'ємних значень змінної t показники експонент у правій частині рівняння (10.13) задовольняють нерівності $\gamma_2 t > 0$, $\gamma_1 t < 0$, тому другий доданок у правій частині рівняння (10.13) набагато перевищує перший. Унаслідок цього фазова траєкторія рухомої точки проходить поблизу сепаратриси $y = h_2 x$ (рис. 10.3).

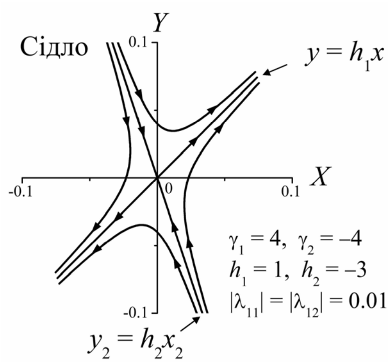

Рис. 10.3. Фазові траєкторії лінеаризованої системи рівнянь, розраховані для випадку різних за знаком дійсних коренів характеристичного рівняння

Із плином часу показник експоненти в другому доданку правої частини рівняння (10.13) зменшується і стає від'ємним, а у першому – збільшується і стає додатним. Завдяки цьому траєкторія рухомої точки відхиляється від сепаратриси $y = h_2 x$ і наближається до сепаратриси $y = h_1 x$, як показано стрілочками на рис. 10.3. Кожна із фазових траєкторій, показаних на рис. 10.3, відповідає одній із чотирьох можливих комбінацій знаків сталих

λ_{11} та λ_{12}. Форма зображених траєкторій асоціюється із формою кінського сідла, тому точку спокою називають *сідлом*.

Означення 10.2. Якщо характеристичне рівняння лінеаризованої системи має *два дійсних корені різного знака*, точку спокою називають *сідловою точкою*, *або сідлом.*

3. Характеристичне рівняння має *два уявні корені* $k_1 = i\omega$, $k_2 = -i\omega$, які комплексно-спряжені з огляду на те, що коефіцієнти лінеаризованої системи рівнянь є дійсними числами. Як було показано в розд. 9, у такому випадку загальний інтеграл лінеаризованої системи має форму

$$\begin{aligned} x &= C_1 \cos\omega t + C_2 \sin\omega t, \\ y &= l_1 \cos\omega t + l_2 \sin\omega t, \end{aligned} \tag{10.15}$$

де l_1 та l_2 є лінійними комбінаціями довільних сталих C_1 та C_2. Щоб з'ясувати, яку форму в цьому випадку мають фазові траєкторії, доцільно виразити загальний інтеграл (10.15) через інші довільні сталі, означені співвідношеннями

$$C_1 = A\cos\alpha, \quad C_2 = -A\sin\alpha,$$
$$l_1 = B\sin\beta, \quad l_2 = B\cos\beta,$$

тому що завдяки цьому загальний інтеграл лінеаризованої системи виражається формулами

$$x = A\cos(\omega t + \alpha), \; y = B\sin(\omega t + \beta), \tag{10.16}$$

з яких випливає, що випадкові збурення періодично змінюються із часом. Якщо $\alpha = \beta$, то рівняннями фазових траєкторій є рівняння еліпсів з осями довжиною в $2A$ та $2B$:

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1. \tag{10.17}$$

Якщо ж $\alpha \neq \beta$, то фазові траєкторії мають форму еліпсів, що повернуті відносно осей координат (рис. 10.4).

На рис. 10.4 видно, що фазові траєкторії замкнені і цілком лежать в обмеженій області фазової площини, а отже, випадкові збурення ніколи не перевищують певних значень. Як вже було зазначено, у цьому разі лінійне за збуреннями наближення недостатнє для висновку про стійкість або нестійкість стаціонарного стану ДС до впливу зовнішніх чинників, і тому треба проводити більш точне дослідження цього питання.

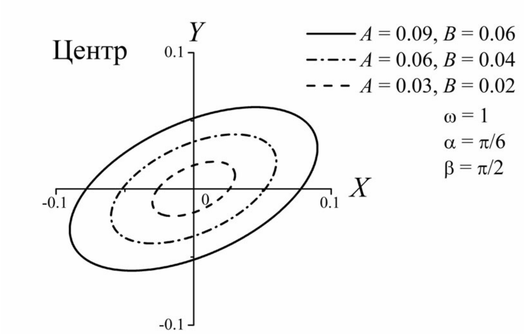

Рис. 10.4. Фазові траєкторії лінеаризованої системи рівнянь, розраховані для випадку уявних коренів характеристичного рівняння

Означення 10.3. Якщо характеристичне рівняння лінеаризованої системи рівнянь має *два уявних корені*, то точку спокою називають *центром.*

Назва цієї точки спокою пояснюється тим, що вона розташована всередині замкнених фазових траєкторій.

4. Характеристичне рівняння має *два комплексно-спряжених корені* $k_1 = \gamma + i\omega$, $k_2 = \gamma - i\omega$, які мають ненульову дійсну частину γ. У такому разі загальний інтеграл лінеаризованої системи рівнянь має форму

$$x = Ae^{\gamma t}\cos(\omega t + \alpha),$$

$$y = Be^{\gamma t}\sin(\omega t + \beta).$$

Можна легко уявити, яку форму матимуть фазові траєкторії, якщо вважати, що $\gamma << \omega$. Тоді точка з координатами (x, y) рухається по фазовій площині за майже еліптичною траєкторією, поступово віддаляючись від початку координат, коли $\gamma > 0$, або наближаючись до нього, коли $\gamma < 0$. Отже, фазова траєкторія має форму спіралі (рис. 10.5).

Означення 10.4. Якщо характеристичне рівняння лінеаризованої системи має *два комплексно-спряжених корені з ненульовою дійсною частиною*, то точку спокою називають *фокусом.*

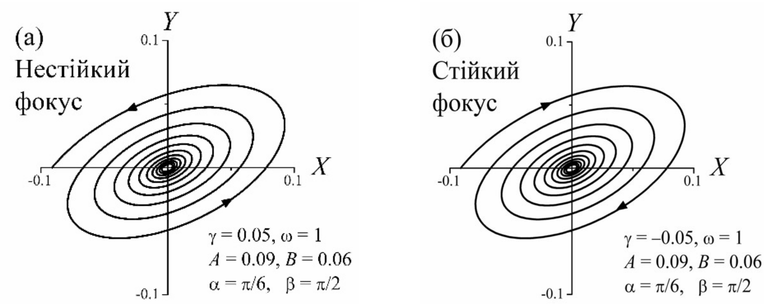

Рис. 10.5. Фазові траєкторії лінеаризованої системи рівнянь, розраховані для випадку комплексних коренів характеристичного рівняння

Точка спокою "фокус" буде нестійкою, якщо дійсна частина кореня характеристичного рівняння додатна, і стійкою, якщо дійсна частина кореня характеристичного рівняння від'ємна.

5. Характеристичне рівняння має *один кратний корінь* $k=\gamma$. Кратний корінь є дійсним, оскільки дійсними є коефіцієнти характеристичного рівняння. Як пояснено в зауваженні 9.3, у такому разі загальний інтеграл лінеаризованої системи набуває форми

$$\begin{aligned} x &= (C_1 + C_2 t)e^{\gamma t}, \\ y &= (l_1 + l_2 t)e^{\gamma t}, \end{aligned} \tag{10.18}$$

де l_1 та l_2 – лінійні комбінації довільних сталих C_1 та C_2. Оскільки збурення x та y пропорційні експоненті, то вони зростають із часом, якщо $\gamma > 0$, або зменшуються, якщо $\gamma < 0$.

Означення 10.5. Якщо характеристичне рівняння лінеаризованої системи має один кратний корінь, точку спокою називають *виродженим вузлом.*

Вироджений вузол нестійкий, якщо $\gamma > 0$, і стійкий, якщо $\gamma < 0$. Дослідження фазових траєкторій в околі виродженого вузла не має великого практичного значення, оскільки на практиці коефіцієнти системи ДР, а отже, і коефіцієнти характеристичного рівняння, визначаються з певною точністю, а більш точне визначення цих коефіцієнтів зазвичай перетворює кратний корінь на два близьких за величиною дійсних корені, або на два комплексно-спряжених корені з малою уявною частиною. Відповідно до цього вироджений вузол перетворюється на невироджений або на фокус. (Зображення фазової траєкторії в околі виродженого вузла можна знайти, напр., у підручнику [11]).

10.5. Приклад дослідження стійкості та визначення типу точок спокою

Щоб краще пояснити процедуру дослідження стійкості стаціонарних станів динамічної системи, розглянемо таку нелінійну систему диференціальних рівнянь:

$$\begin{cases} \dot{y}_1 = 2y_1 + y_2^2 - 1, \\ \dot{y}_2 = 6y_1 - y_2^2 + 1. \end{cases} \tag{10.19}$$

Координати точок спокою, тобто точок, у яких $\dot{x} = \dot{y} = 0$, знайдемо з рівнянь

$$\begin{cases} 2y_1 + y_2^2 - 1 = 0, \\ 6y_1 - y_2^2 + 1 = 0. \end{cases} \tag{10.20}$$

Система алгебраїчних рівнянь (10.20) має розв'язки $y_{10} = 0$, $y_{20} = \pm 1$, звідки випливає, що система (10.19) має дві точки спокою на площині Y_1OY_2 змінних y_1, y_2, а саме:

$$P_1(0,1), \quad P_2(0,-1). \tag{10.21}$$

Визначимо тип та дослідимо стійкість цих точок спокою.

Спочатку лінеаризуємо систему рівнянь (10.19). Для цього додамо до сталих величин малі збурення x та y:

$$y_1 = y_{10} + x = x, \quad y_2 = y_{20} + y, \quad y_2^2 = 1 + 2y_{20}y + y^2,$$

знехтуємо величиною y^2, підставимо збурені величини y_1, y_2 та $y_2^2 \approx 1 + 2y_{20}y$ до системи (10.19) та одержимо лінеаризовану систему рівнянь[1]

$$\begin{cases} \dot{x} = 2x + 2y_{20}y, \\ \dot{y} = 6x - 2y_{20}y. \end{cases} \qquad (10.22)$$

Тепер дослідимо стійкість та визначимо тип точки P_1. Поклавши в лінеаризованій системі $y_{20} = 1$, одержуємо характеристичне рівняння

$$\begin{vmatrix} 2-k & 2 \\ 6 & -2-k \end{vmatrix} = 0$$

та систему рівнянь

$$\begin{cases} (2-k)\lambda_{11} + 2\lambda_{12} = 0, \\ 6\lambda_{21} - (2+k)\lambda_{22} = 0 \end{cases} \qquad (10.23)$$

для визначення коефіцієнтів загального інтеграла лінеаризованої системи рівнянь. Знаходимо корені характеристичного рівняння

$$k \equiv \gamma_1 = 4,\ k \equiv \gamma_2 = -4.$$

Висновок 10.1. *Точка* P_1 *є сідловою точкою, тому вона нестійка.*

По черзі підставивши знайдені корені до системи рівнянь (10.23), визначимо, що для кореня γ_1 виконується співвідношення $\lambda_{21}/\lambda_{11} = 1$, а для кореня γ_2 – співвідношення $\lambda_{22}/\lambda_{12} = -3$. Після цього знаходимо загальний інтеграл лінеаризованої системи

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_{11} \\ \lambda_{11} \end{pmatrix} e^{4t} + \begin{pmatrix} \lambda_{12} \\ -3\lambda_{12} \end{pmatrix} e^{-4t},$$

де λ_{11} та λ_{12} можуть набувати будь-яких дійсних значень. Вводимо до розгляду величини $h_1 \equiv \lambda_{21}/\lambda_{11} = 1$, $h_2 \equiv \lambda_{22}/\lambda_{12} = -3$, фіксуємо малі значення довільних сталих $\lambda_{11} = \pm 0,1$ та $\lambda_{12} = \pm 0,1$, і повертаємось до рис. 10.3.

---

[1] У більш складних випадках слід розкладати праві частини нелінійної системи рівнянь за малими збуреннями, користуючись формулою Тейлора (10.5).

Висновок 10.2. Із рис. 10.3 стає очевидно, що в околі точки P_1 фазові траєкторії мають такий вигляд, як показано у верхній частині рис. 10.6. *На завершення цього прикладу дослідимо стійкість і визначимо тип точки* P_2. Поклавши в лінеаризованій системі $y_{20} = -1$, одержимо характеристичне рівняння

$$\begin{vmatrix} 2-k & -2 \\ 6 & 2-k \end{vmatrix} = 0$$

та лінеаризовану систему рівнянь

$$\begin{cases} \dot{x} = 2x - 2y, \\ \dot{y} = 6x + 2y. \end{cases} \tag{10.24}$$

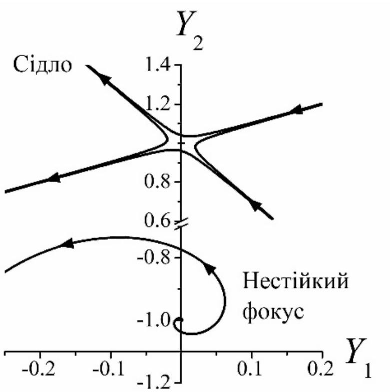

Рис. 10.6. Фазові траєкторії нелінійної системи ДР (6.20) в околі точок спокою $P_1(0,1)$ та $P_2(0,-1)$

Характеристичне рівняння має комплексно-спряжені корені $k_{1,2} = 2(1 \pm \sqrt{3}i)$ з додатною дійсною частиною $\gamma = 2$.

Висновок 10.3. Точка P_2 є нестійким фокусом.

В околі цієї точки спокою форма фазових траєкторій системи (10.19) на площині змінних y_1, y_2 визначається з рівнянь для загального інтеграла

$$\begin{aligned} x &= e^{2t}(\lambda_{11}\cos 2\sqrt{3}t - C_{12}\sin 2\sqrt{3}t), \\ y = x &= e^{2t}(\lambda_{11}\sqrt{3}\sin 2\sqrt{3}t + \lambda_{12}\sqrt{3}\cos 2\sqrt{3}t). \end{aligned} \tag{10.25}$$

Одну із цих фазових траєкторій зображено в нижній частині рис. 10.6. Вона значно відрізняється за формою від фазових траєкторій, накреслених на рис. 10.5, оскільки відповідає значенню $\gamma = 2$, а на рис. 10.5 накреслено траєкторії, розраховані для малої величини γ.

Висновок 10.4. Рухома точка обертається навколо точки спокою P_2 у напрямку, протилежному до напрямку обертання годинникової стрілки, швидко віддаляючись від цієї точки.

Зазначимо, що як і має бути, напрямок руху навколо точки P_2 узгоджується з напрямком руху повз точку P_1. (Це відображає напрямок стрілочок на фазових траєкторіях, зображених на рис. 10.6).

10.6. Стійкість стаціонарних розв'язків диференціального рівняння другого порядку

Стаціонарні стани і динаміка різних динамічних систем (фізичних, хімічних, соціальних тощо) у багатьох випадках описується не системою двох ДР 1-го порядку, а одним ДР 2-го порядку. Як досліджується стійкість стаціонарних станів ДС у таких випадках, пояснімо на прикладі нелінійного диференціального рівняння 2-го порядку:

$$\ddot{y} + 2\gamma\dot{y} + a^2 y^2 - b^2 = 0 . \qquad (10.26)$$

Для визначеності вважатимемо, що параметри, якими описується ДС, задовольняють нерівності

$$\gamma > 0 , \quad ab > 0 , \; \gamma^2 < 2ab . \qquad (10.27)$$

Розв'язки, що відповідають стаціонарним станам, знаходимо з умови $\ddot{y} = \dot{y} = 0$, яка перетворює рівняння (10.26) на рівняння стаціонарних станів $a^2 y^2 - b^2 = 0$. Це рівняння має два розв'язки. Дослідимо стійкість відповідних до них стаціонарних станів.

1. Перший розв'язок $y = y_1 = -b/a$. Додамо до цього розв'язку випадкове збурення і підставимо збурену функцію $y = y_1 + \tilde{y}$ до рівняння (10.26):

$$\ddot{\tilde{y}} + 2\gamma\dot{\tilde{y}} + a^2\left(-\frac{b}{a} + \tilde{y}\right)^2 - b^2 = 0.$$

Знехтувавши малим доданком $a^2\tilde{y}^2$, знайдемо рівняння, лінеаризоване за малим збуренням

$$\ddot{\tilde{y}} + 2\gamma\dot{\tilde{y}} - 2ab\tilde{y} = 0\,.$$

Характеристичне рівняння $k^2 + 2\gamma k - 2ab = 0$ має два дійсних корені

$$k_{\pm} = -\gamma \pm \sqrt{\gamma^2 + 2ab}\,. \qquad (10.28)$$

Із нерівності $ab > 0$ випливає, що корінь k_+ додатний, а отже, *стаціонарний стан, що відповідає розв'язку* $y = -b/a$ *нестійкий.*

2. Другий розв'язок $y = y_2 = b/a$. Для дослідження стійкості відповідного до цього розв'язку стаціонарного стану достатньо замінити $b \to -b$ у формулі (10.28) для коренів характеристичного рівняння. Унаслідок цієї заміни ця формула набуває форми

$$k_{\pm} = -\gamma \pm \sqrt{\gamma^2 - 2ab}.$$

З умови $\gamma^2 < 2ab$ випливає, що корені характеристичного рівняння, які відповідають другому розв'язку рівняння стаціонарних станів, є двома комплексними числами

$$k_{\pm} = -\gamma \pm i\sqrt{2ab - \gamma^2}\,,$$

що мають від'ємну дійсну частину $-\gamma$, а отже, *стаціонарний стан, що відповідає розв'язку* $y = b/a$, *стійкий.*

Цікаво зазначити, що рівняння (10.26) не змінюється при зміні знака параметра b, і тому висновок про те, що стаціонарний стан, який відповідає розв'язку $y_1 = -b/a$, нестійкий, а розв'язку $y_2 = b/a$, – стійкий, може викликати здивування. Здивування зникне, якщо помітити, що при заміні b на $-b$ розв'язок y_1 перетворюється на розв'язок y_2, а розв'язок y_2 стає розв'язком y_1; завдяки цьому заміна знака параметра b не впливає на результат дослідження стійкості стаціонарних станів рівняння (10.26).

РОЗДІЛ 11
Застосування лінійної алгебри для розкладу функції в ряд Фур'є. інтегральне перетворення Фур'є

11.1. Постановка задачі

Ще у стародавні часи люди відчули, що кілька струн на музичних інструментах можуть утворювати дуже приємний звук, коли коливаються одночасно. Згодом стало зрозуміло, що коливання струн описуються функцією (8.76), яка є розв'язком хвильового рівняння, розглянутого в підрозд. 8.5, а різноманітні звуки люди чують завдяки коливанням повітря, що поширюються в атмосфері Землі. Можливо, що саме завдяки цьому розумінню і тому факту, що розв'язки рівняння коливань та хвильового рівняння є тригонометричними функціями (див. (8.42), (8.68) (8.71)) виникла ідея про представлення різних функцій змінної x суперпозицією синусів та косинусів цієї змінної. Із погляду лінійної алгебри, доцільно використовувати для цього функції

$$\psi_m(x)=\sqrt{\frac{2}{l}}\cos\left(\frac{2\pi m}{l}x\right) \text{ та } \varphi_m(x)=\sqrt{\frac{2}{l}}\sin\left(\frac{2\pi m}{l}x\right), \quad (11.1)$$

де $m=1,2,3...$, $l=\text{const}$, тому що вони, як було зазначено в п. 3.4.2, утворюють ортонормований базис у просторі функцій, неперервних в інтервалі значень змінної $-l/2\le x\le l/2$.

Розглянемо функцію $f(x)$, неперервну в інтервалі $-l/2\le x\le l/2$, і розкладемо її по базисних функціях. Розклад функції $f(x)$ по функціях (11.1) є нескінченним рядом, подібним до ряду (3.23):

$$f(x)=\frac{a_0}{\sqrt{2l}}+\sum_{m=1}^{\infty}[a_m\varphi_m(x)+b_m\psi_m(x)] \quad (11.2)$$

Означення 11.1. Нескінченний ряд (11.2) називають тригонометричним рядом Фур'є.

Практично важлива задача полягає у визначенні таких коефіцієнтів a_m *та* b_m, *за яких справджується рівняння* (11.2). Причину, з якої доданок із $m = 0$, відокремлений від інших доданків, буде пояснено далі.

Коли коефіцієнти, які задовольняють рівняння (11.2), знайдені, кажуть, що *функція* $f(x)$ *розкладена в ряд Фур'є в інтервалі значень* $-l/2 \le x \le l/2$ *змінної* x. Далі буде показано, що результат обчислення коефіцієнтів тригонометричного ряду Фур'є залежить від величини параметра l, чим і пояснюється така термінологія. Математики довели, що функція може бути розкладена в ряд Фур'є (11.1) на інтервалі $-l/2 \le x \le l/2$ за умови, що на цьому інтервалі вона обмежена, має скінченну кількість екстремумів та скінченну кількість точок розриву першого роду[1].

11.2. Обчислення коефіцієнтів тригонометричного ряду Фур'є

Щоб знайти формулу для обчислення коефіцієнта a_0 ряду Фур'є, зінтегруємо ліву та праву частини рівності (11.2) у симетричних межах, від $-l/2$ до $l/2$, та зважимо на рівності

$$\int_{-l/2}^{l/2} \varphi_m(x)dx = 0 \quad \text{та} \quad \int_{-l/2}^{l/2} \psi_m(x)dx = 0\,,$$

які є очевидними з того, що інтеграли від $-l/2$ до $l/2$ є інтегралами за цілим числом періодів тих синусів та косинусів, до яких пропорційні функції $\varphi_m(x)$ та $\psi_m(x)$. У такий спосіб отримаємо з (11.2) рівняння

$$\int_{-l/2}^{l/2} f(x)dx = \frac{a_0}{\sqrt{2l}} \int_{-l/2}^{l/2} dx\,,$$

з якого випливає, що

$$a_0 = \sqrt{\frac{2}{l}} \int_{-l/2}^{l/2} f(x)dx\,. \qquad (11.3)$$

[1] Ця умова називається умовою Діріхле.

Щоб знайти формулу для обчислення інших коефіцієнтів тригонометричного ряду Фур'є, помножимо обидві частини рівності (11.2) на функцію $\varphi_n(x)$ та зінтегруємо їх у межах від $-l/2$ до $l/2$:

$$\int_{-l/2}^{l/2} f(x)\varphi_n(x)dx = \sum_{m=1}^{\infty}\left[a_m \int_{-l/2}^{l/2} \varphi_m(x)\varphi_n(x)dx + b_m \int_{-l/2}^{l/2} \psi_m(x)\varphi_n(x)dx\right].$$

З огляду на умову ортонормованості базисних функцій (див. (3.22), (3.25)), уся сума, що стоїть у правій частині рівняння, а отже, і інтеграл, що стоїть у правій частині, дорівнює коефіцієнту a_n, тобто

$$a_n = \int_{-l/2}^{l/2} f(x)\varphi_n(x)dx . \tag{11.4}$$

Оскільки рівняння (11.1) можна помножити на будь-яку із функцій $\varphi_n(x)$, то вираз (11.4) є чинним для всіх коефіцієнтів a_n.

Помноживши обидві частини рівності (11.2) на функцію $\psi_n(x)$, і зінтегрувавши їх у межах від $-l/2$ до $l/2$, легко отримати вираз для коефіцієнтів b_n:

$$b_n = \int_{-l/2}^{l/2} f(x)\psi_n(x)dx . \tag{11.5}$$

Підставивши функції (11.1) до виразів (11.2), (11.4) та (11.5), знаходимо явний вираз тригонометричного ряду Фур'є

$$f(x) = \frac{a_0}{\sqrt{2l}} + \sqrt{\frac{2}{l}}\sum_{n=1}^{\infty}\left[a_n \cos\left(\frac{2\pi n}{l}x\right) + b_n \sin\left(\frac{2\pi n}{l}x\right)\right] \tag{11.6}$$

та формули, за якими слід обчислювати його коефіцієнти:

$$a_n = \sqrt{\frac{2}{l}}\int_{-l/2}^{l/2} f(x)\cos\left(\frac{2\pi n}{l}x\right)dx , \tag{11.7}$$

$$b_n = \sqrt{\frac{2}{l}}\int_{-l/2}^{l/2} f(x)\sin\left(\frac{2\pi n}{l}x\right)dx . \tag{11.8}$$

Формули (11.6) – (11.8) вирішують поставлену задачу про розклад функції $f(x)$ у ряд Фур'є в інтервалі її неперервності $-l/2 \le x \le l/2$.

Ряд Фур'є широко застосовується для опису динамічних явищ, у тому числі звуків, створених музичними інструментами. Це зумовило застосування до рядів Фур'є певної термінології.

Означення 11.2. Доданки суми у правій частині рівняння (11.6) називають *гармоніками функції* $f(x)$ (доданок з $n=1$ – першою гармонікою, з $n=2$ – другою, з $n=3$ – третьою, …).

Зауваження 11.1. За формулами (11.6) – (11.8) можна розкласти функцію $f(x)$ у ряд Фур'є на інтервалі неперервності цієї функції $x_1 \le x \le x_2$, межі якого розташовані несиметрично до початку координат. Для цього слід ввести позначення $x_2 - x_1 \equiv l$, замінити у формулах (11.6) – (11.8) змінну x на нову змінну $\tilde{x} = x - (x_1 + l/2)$ та обчислити коефіцієнти a_n, b_n, виконавши інтегрування за новою змінною в межах від $-l/2$ до $l/2$.

Зауваження 11.2. Перша гармоніка розкладу функції в ряд Фур'є (11.6) є періодичною функцією з періодом l, а вищі гармоніки – періодичні з періодом l/n. Унаслідок цього сума гармонік, а отже, і права частина рівняння (11.6), є періодичною функцією з періодом l. Із цієї причини функція, яка не є періодичною, не може бути представлена рядом (11.6) поза межами інтервалу $x_1 \le x \le x_2$, ширина якого дорівнює l.

Зауваження 11.3. Якщо функція $f(x)$ парна, тобто $f(-x) = f(x)$, то в правій частині рівняння (11.6) не можуть бути присутні непарні функції, тому всі коефіцієнти b_n дорівнюють нулю, а отже, обчислювати їх не треба. Якщо ж функція $f(x)$ непарна, тобто $f(-x) = -f(x)$, то в правій частині рівняння (11.6) не можуть бути присутні парні функції, тому дорівнюють нулю коефіцієнти a_n та a_0 [1].

---

[1] Це підтверджується виразами (11.7), (11.8) і тим фактом, що інтеграл із симетричними межами від будь-якої непарної функції дорівнює нулю.

Зауваження 11.4. Завдяки тому, що перший доданок у правій частині рівняння (11.2) поділений на $\sqrt{2l}$, формули (11.7), (11.8) придатні для обчислення всіх коефіцієнтів ряду, включно з коефіцієнтом a_0. Саме тому цей доданок був відокремлений від інших доданків суми в правій частині рівняння (11.2).

Зауваження 11.5. Вирази (11.6) – (11.8) були отримані з розкладу функції $f(x)$ в ортонормованому базисі, що складається із функцій (11.1). Це було зроблено для того, щоб зауважити, що розклад функції в ряд Фур'є є окремим випадком розкладу елемента лінійного простору по елементах базису. У багатьох підручниках наведено спрощені вирази

$$f(x) = \frac{1}{2}\overline{a}_0 + \sum_{n=1}^{\infty}\left[\overline{a}_n \cos(nx) + \overline{b}_n \sin(nx)\right],$$

$$\overline{a}_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx,$$

$$\overline{b}_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx,$$

за якими здійснюється розклад функції $f(x)$ у тригонометричний ряд Фур'є в інтервалі $-\pi \le x \le \pi$. Поклавши $l = 2\pi$ у формулах (11.7), (11.8), легко знайти рівності $a_n = \overline{a}_n\sqrt{\pi}$, $n = 0,1,2...$.

11.3. Наближене представлення функції сумою гармонік

Тригонометричний ряд Фур'є застосовується для виконання багатьох практично важливих завдань. Одним із таких завдань є наближене представлення різних (не лише періодичних) функцій на інтервалі $-l/2 \le x \le l/2$ сумою певної кількості доданків цього ряду за наближеною формулою

$$f(x) \approx \frac{a_0}{\sqrt{2l}} + \sqrt{\frac{2}{l}}\sum_{n=1}^{N}\left[a_n \cos\left(\frac{2\pi n}{l}x\right) + b_n \sin\left(\frac{2\pi n}{l}x\right)\right], \quad (11.9)$$

тобто розкладення функції $f(x)$ на певну кількість гармонік. Кількість гармонік N, які слід врахувати у формулі (11.9), залежить від того, з якою точністю необхідно обчислити значення функції $f(x)$ для розв'язання тієї чи іншої задачі. Заміна точного рівняння (11.6) наближеною формулою (11.9) можлива тоді, коли тригонометричний ряд Фур'є є збіжним рядом. Ретельне формулювання та доведення умов, яким має задовольняти функція, щоб до неї можна було застосовувати формулу (11.9), є досить складним завданням. Це завдання не є критично важливим з погляду лінійної алгебри, але з методичного погляду корисно розглянути приклад, який покаже, як точність наближеного представлення функції сумою гармонік залежить від їхньої кількості.

● **Приклад 11.1.** Представимо сумою гармонік функцію $f(x)=|x|$ на інтервалі $-1 \le x \le 1$ та подивимось, як точність такого представлення залежить від кількості врахованих у рівнянні (11.9) гармонік.

Передусім візьмемо до уваги, що в цьому випадку, функція $f(x)$ парна, і внаслідок цього $b_n = 0$. Зваживши на це і поклавши $n=0$ у формулі (11.7), знайдемо перший доданок суми в рівнянні (11.9):

$$a_0 = 2\int_0^1 x dx = 1 \quad \Rightarrow \quad \frac{a_0}{\sqrt{2l}} = \frac{1}{2}.$$

Далі знайдемо коефіцієнти розкладу функції по гармоніках:

$$a_n = 2\int_0^1 x\cos(\pi nx)dx = \frac{2}{\pi n}\left[x\sin(\pi nx)\,|_0^1 - \int_0^1 \sin(\pi nx)dx \right]$$

$$= \frac{2}{\pi^2 n^2}\cos(\pi nx)\,|_0^1 = \frac{2}{\pi^2 n^2}[\cos(\pi n) - 1].$$

Урахувавши, що $[\cos(\pi n) - 1]$ дорівнює нулю для парних значень цілого числа n, і мінус двом – для непарних (тобто для $n = 2m+1$, де $m = 0,1,2,...$), знаходимо вираз

$$a_m = -\frac{4}{\pi^2 (2m+1)^2}, \; m = 0,1,2,...,$$

і розклад функції по гармоніках

$$|x| \approx \frac{1}{2} + \frac{4}{\pi^2}\sum_{m=0}^{N} \frac{\cos[\pi(2m+1)x]}{(2m+1)^2}. \tag{11.10}$$

Щоб візуалізувати точність представлення функції $f(x)=|x|$ сумою гармонік, порівняємо її графік з графіками функцій

$$S_N(x)=\frac{1}{2}+\frac{4}{\pi^2}\sum_{m=0}^{N}\frac{\cos[\pi(2m+1)x]}{(2m+1)^2},$$

показаних на рис. 11.1.

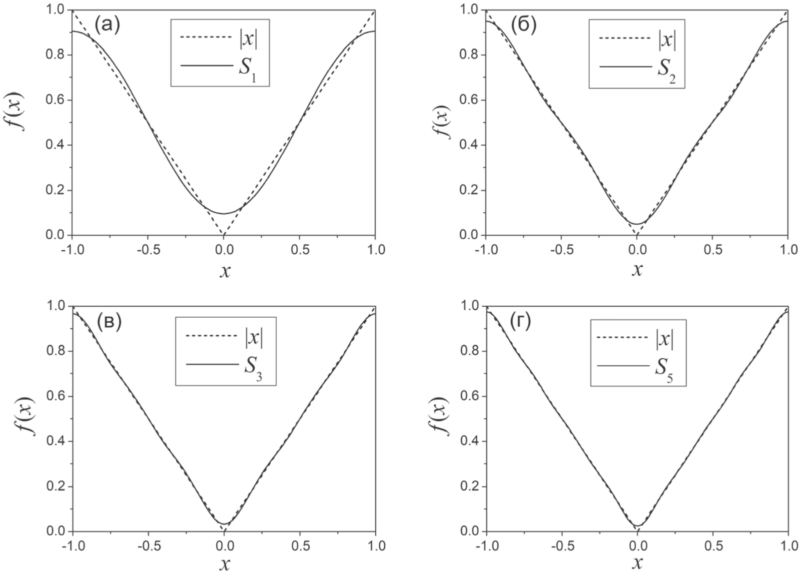

Рис. 11.1. Функція $f(x)=|x|$ та її представлення

а) першою гармонікою, б) сумою двох, в) трьох, г) п'ятьох гармонік

Рис. 11.1, *а* наочно доводить, що точність представлення функції першою гармонікою низька. Рис. 11.1, *б* показує, що точність представлення сумою двох гармонік є набагато вищою, але різниця між графіками функції та її представлення $S_2(x)$ ще помітна. Графіки представлення функції трьома та п'ятьма гармоніками майже зливаються з графіком функції всюди, за винятком вузького околу початку координат і точок з координатами (–1, 1) та (1,1) (рис. 11.1 *в*, *г*), тому точність представлення функції такою кількістю гармонік слід проаналізувати додатково.

Абсолютна похибка представлення функції $f(x) = |x|$ сумою гармонік дорівнює різниці $|x| - S_N(x)$. Величина абсолютної похибки представлення цієї функції сумою трьох гармонік не перевищує числа 0,033 (рис. 11.2, *а*), а сумою п'яти гармонік – числа 0,025 (рис. 11.2, *б*).

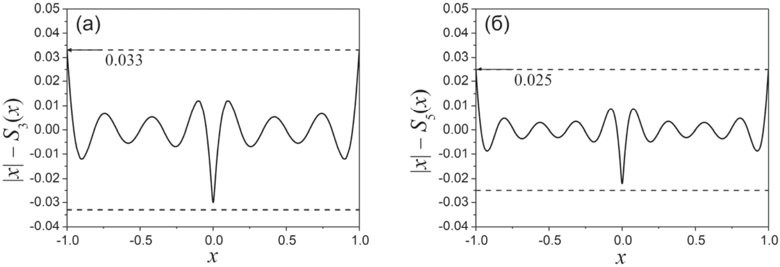

Рис. 11.2. Абсолютна похибка представлення функції $f(x) = |x|$ сумою

а) трьох, б) п'ятьох гармонік

Відносну похибку представлення функції $f(x) = |x|$ сумою гармонік можна охарактеризувати відношенням[1]

$$[|x| - S_N(x)] / S_N(x) \equiv \Delta_N(x).$$

Графіки функції $\Delta_N(x)$ показано на рис. 11.3. Всюди, за винятком відносно вузького інтервалу $|x| < 0,012$, величина абсолютної похибки представлення цієї функції сумою трьох гармонік не перевищує числа 0,03 (рис. 11.3, *а*), а сумою п'яти гармонік – числа 0,025 (рис. 11.3, *б*). Великі значення відносної похибки в інтервалі $|x| < 0,012$ зумовлені тим, що похідна функції $f(x) = |x|$ має розрив у точці $x = 0$, а похідні всіх доданків ряду Фур'є є неперервними функціями. Слід зауважити, що в інтервалі $|x| < 0,012$ функція набуває значень набагато менших від одиниці.

[1] Таке відношення, у якому $|x|$ стоїть у знаменнику дробу, використовувати не зручно, тому що $|x|$ дорівнює нулю напочатку координат.

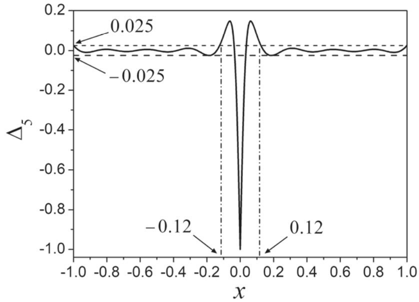

Рис. 11.3. Відносна похибка представлення функції $f(x) = |x|$ сумою п'ятьох гармонік

Узагальнюючи цей приклад, зазначимо, що якщо представлення функції певною кількістю гармонік призводить до значної похибки там, де функція набуває малих значень, така похибка може виявитися припустимою для розв'язання певних задач. Наприклад, якщо на тіло діє змінна сила, то те, як саме вона змінювалася в той проміжок часу, коли її величина була малою, практично не впливає на результат її дії. Тоді може бути прийнятним, щоб відносна похибка вимірювання сили у цей часовий проміжок була порядку одиниці.●

11.4. Комплексна форма ряду Фур'є

Введемо позначення

$$\frac{2\pi n}{l} \equiv k_n \tag{11.11}$$

та виразимо тригонометричний ряд Фур'є (11.6) у формі

$$f(x) = \frac{a_0}{\sqrt{2l}} + \sqrt{\frac{2}{l}} \sum_n^{\infty} [a_n \cos(k_n x) + b_n \sin(k_n x)], \tag{11.12}$$

де

$$a_n = \sqrt{\frac{2}{l}} \int_{-l/2}^{l/2} f(x) \cos\left(k_n x\right) dx, \quad b_n = \sqrt{\frac{2}{l}} \int_{-l/2}^{l/2} f(x) \sin\left(k_n x\right) dx. \tag{11.13}$$

Розглянемо ряд

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{ik_n x} \ . \qquad (11.14)$$

Узявши до уваги, що

$$k_n = -k_{-n}, \qquad (11.15)$$

з'ясуємо шляхом тотожних математичних перетворень, якими мають бути коефіцієнти C_n, щоб права частина рівняння (11.4) дорівнювала правій частині рівняння (11.2):

$$\sum_{n=-\infty}^{\infty} C_n e^{iknx} = C_0 + \sum_{n=-\infty}^{-1} C_n e^{iknx} + \sum_{n=1}^{\infty} C_n e^{iknx} \ ;$$

$$\sum_{n=-\infty}^{-1} C_n e^{ik_n x} = \sum_{n=-1}^{-\infty} C_n e^{ik_n x} = C_{-1} e^{ik_{-1} x} + C_{-2} e^{ik_{-2} x} + C_{-3} e^{ik_{-3} x} + \ldots$$

$$= C_{-1} e^{-ik_1 x} + C_{-2} e^{-ik_2 x} + C_{-3} e^{-ik_3 x} + \ldots = \sum_{n=1}^{\infty} C_{-n} e^{-ik_n x};$$

$$\sum_{n=-\infty}^{\infty} C_n e^{ik_n x} = C_0 + \sum_{n=1}^{\infty} (C_{-n} e^{-ik_n x} + C_n e^{ik_n x})$$

$$= C_0 + \sum_{n=1}^{\infty} \{C_{-n}[\cos(k_n x) - i\sin(k_n x)] + C_n[\cos(k_n x) + i\sin(k_n x)]\}.$$

Перегрупувавши доданки, бачимо, що ряд (11.14) перетворився на тригонометричний ряд

$$f(x) = C_0 + \sum_{n=1}^{\infty} [(C_n + C_{-n})\cos(k_n x) + i(C_n - C_{-n})\sin(k_n x)] \ .$$

Порівнявши цей ряд із рядом (11.12), знаходимо систему рівнянь

$$\begin{cases} C_n + C_{-n} = a_n, \\ i(C_n - C_{-n}) = b_n, \end{cases}$$

з якої випливає, що ці ряди тотожні, якщо справджується рівність

$$C_n = \frac{1}{\sqrt{2l}} (a_n - ib_n) \ .$$

Підставивши до правої частини цієї рівності коефіцієнти (11.3), одержуємо такий вираз:

$$C_n = \frac{1}{l} \int_{-l/2}^{l/2} f(x) e^{-ik_n x} dx \ . \qquad (11.16)$$

Висновок 11.1. Ряд (11.4) з коефіцієнтами (11.16) є вираженим у комплексній формі рядом Фур'є, який скорочено називають комплексним рядом Фур'є[1].

Зауваження 11.6. Із рівності (11.15) та виразу (11.16) випливає важлива властивість коефіцієнтів комплексного ряду Фур'є:

$$C_{-n} = C_n^*, \tag{11.17}$$

де C_n^* – число, комплексно спряжене з числом C_n. Завдяки цій властивості сума ряду (11.14) є функцією, яка набуває лише дійсних значень, як це і має бути, оскільки ця сума дорівнює функції $f(x)$.

11.5. Представлення функції інтегралом Фур'є, перетворення Фур'є функції однієї змінної

Розглянемо функцію $f(x)$, неперервну за всіх значень змінної x. З'ясуємо, чим треба замінити комплексний ряд Фур'є у граничному випадку $l \to \infty$. Із цієї метою розглянемо рівняння (11.14) та візьмемо до уваги, що згідно з позначенням (11.11) справджуються такі рівності:

$$\Delta k_n \equiv k_{n+1} - k_n = \frac{2\pi}{l}, \quad \Delta k_n \frac{l}{2\pi} = 1. \tag{11.18}$$

Сума ряду (11.14) не зміниться, якщо всі його доданки помножити на одиницю, тому комплексний ряд Фур'є (11.14) тотожно перетворюється на ряд

$$f(x) = \sum_{n=-\infty}^{\infty} \Delta k_n \frac{l}{2\pi} C_n e^{ik_n x}.$$

Із рівнянням (11.11) випливає, що $n \equiv k_n l / 2\pi$, тому коефіцієнти C_n залежать від параметра k_n. Введемо позначення

$$\Phi(k_n) \equiv lC_n = \int_{-l/2}^{l/2} f(x) e^{-ik_n x} dx \tag{11.19}$$

[1] Якщо додержуватись більш коректної термінології, слід казати: "Цей ряд є сумою функцій комплексної змінної".

та виразимо комплексний ряд Фур'є для функції $f(x)$ у такій формі:

$$f(x)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\Phi(k_n)e^{ik_n x}\Delta k_n\,. \qquad (11.20)$$

Здійснимо в рівнянні (11.20) граничний перехід $l\to\infty$. Коли l прямує до нескінченності, величина Δk_n прямує до нуля (див. (11.18)), а отже, сума, що стоїть у правій частині рівняння (11.20), перетворюється на інтеграл:

$$\lim_{l\to\infty}\left(\sum_{n=-\infty}^{\infty}\Phi(k_n)e^{ik_n x}\Delta k_n\right)=\lim_{\Delta k_n\to 0}\left(\sum_{n=-\infty}^{\infty}\Phi(k_n)e^{ik_n x}\Delta k_n\right)=\int_{-\infty}^{\infty}\Phi(k)e^{ikx}dk\,.$$

Висновок 11.2. Неперервну за всіх значень змінної x функцію $f(x)$ можна представити у формі інтеграла, який називають інтегралом Фур'є:

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi(k)e^{ikx}dk\,, \qquad (11.21)$$

де

$$\Phi(k)=\int_{-\infty}^{\infty}f(x)e^{-ikx}dx\,, \qquad (11.22)$$

(див. (11.19)).

В алгебраїчному аспекті формула (11.22) є означенням лінійного оператора, який діє на функцію $f(x)$, значення якої є дійсними числами, *перетворюючи* її на функцію $\Phi(k)$, яка може набувати як дійсних, так і комплексних значень.

Означення 11.3. Інтегральне співвідношення (11.22) називають *перетворенням Фур'є* функції $f(x)$, або *прямим перетворенням Фур'є*; співвідношення (11.21) називають *оберненим перетворенням Фур'є.* Функцію $\Phi(k)$ називають фур'є-образом функції $f(x)$.

Зауваження 11.7. Пояснити різницю між поняттями "перетворення Фур'є" та "інтеграл Фур'є" можна одним реченням: інтеграл Фур'є це той інтеграл, який треба обчислити, щоб здійснити пряме перетворення Фур'є функції, означеної умовами завдання, яке треба вирішити.

11.6. Застосування перетворення Фур'є, спектральна щільність сигналу

Ряд і перетворення Фур'є застосовуються у природничих та комп'ютерних науках, у різних галузях інженерії, економіки та військової справи. Такі застосування передбачають можливість представлення складних функцій відповідними до них сумами гармонік. Для багатьох важливих застосувань буває необхідно представити сумою гармонік а) функцію координат; б) функцію часу; в) функцію координат та часу, а потім, докладно аналізувати це представлення, тобто провести *гармонічний аналіз* функції координат та/або часу.

Необхідність гармонічного аналізу функції координат $f(x)$ виникає, наприклад, при проєктуванні, створенні та експлуатації транспортних засобів, мостів, будівель різного призначення тощо. У такому разі використовують залежно від специфіки поставленого завдання ряд або перетворення Фур'є, виражені у формі (11.12), (11.14) або (11.21), де параметри k_n або k (відповідно) мають фізичну вимірність оберненої довжини, а фур'є-гармоніки $\Phi(k)$ зручно вважати нерухомими хвилями із хвильовими числами k_n або k та пов'язаними з ними довжинами (див. (8.73)).

Гармонічний аналіз функцій часу виконується передавачем та одержувачем сигналів, що містять практично важливу інформацію[1]. З одного боку, гармонічний аналіз є великим та непростим розділом математики, а з іншого, – складним інженерно-технічним завданням.

Опишемо дуже коротко ті властивості та особливості перетворення Фур'є, про які варто знати перед вивченням спеціальної

---

[1] *У такому контексті* сигнал – це зміна в часі певної фізичної величини (тиску – в акустиці, електромагнітного поля – в радіотехніці та оптиці тощо), яка використовується для передачі інформації (повідомлення) від передавача до одержувача або від одного до іншого елемента складного технічного пристрою.

літератури з гармонічного аналізу[1]. З цією метою розглянемо пряме та обернене перетворення Фур'є функції часу $f(t)$, значення якої є дійсними в усі моменти часу:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \qquad (11.23)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega, \qquad (11.24)$$

відповідно.

Щоб пояснити суть перетворення Фур'є, зручно вважати функцію $f(t)$ *безрозмірною фізичною величиною*. Тоді із загальновідомого співвідношення

$$e^{\pm i\omega t} = \cos(\omega t) \pm i \sin(\omega t) \qquad (11.25)$$

випливає, що згідно з оберненим перетворенням Фур'є (11.24) функція часу є суперпозицією коливань з різними кутовими частотами ω (див. рівняння (8.74) та його пояснення). Оскільки в цьому випадку ω є неперервною змінною, то існує неперервний *спектр коливань*, а добуток $F(\omega)d\omega$ показує, яка частка від повної кількості коливань має частоту, що лежить у вузькому інтервалі частот від ω до $\omega + d\omega$.

Означення 11.4. Функцію $F(\omega)$ називають спектральною щільністю коливань, суперпозицією яких є сигнал $f(t)$.

Зауваження 11.9. У підручниках зазвичай не акцентують на важливості з погляду фізики множника $1/2\pi$ у рівнянні (11.24), яка полягає в тому, що кутова частота має вимірюється в радіанах за секунду, тому відношення $\omega/2\pi$ вимірюється в обернених секундах; оскільки спектральна щільність (11.23) безрозмірної функції $f(t)$ вимірюється в секундах, то наявність множника $d\omega/2\pi$ забезпечує узгодження фізичних розмірностей правої та лівої частин рівняння (11.24).

---

[1] Ці базові знання необхідні також для вивчення багатьох природничих та технічних наук.

Зауваження 11.10. У літературі з гармонічного аналізу застосовується як означене вище поняття спектральної щільності коливань $F(\omega)$, так і поняття спектральної густини потужності сигналу. Математичним означенням спектральної густини потужності сигналу $S(\omega)$ вважають співвідношення

$$S(\omega) \overset{def}{\equiv} \lim_{T\to\infty} \frac{|F_T(\omega)|^2}{T},$$

де

$$F_T(\omega) = \int_{-T/2}^{T/2} f(t)e^{-i\omega t}dt .$$

(З огляду на зауваження 11.9 функція $F_T(\omega)$, означена в деяких джерелах інформації, може відрізнятися від наведеної вище множником $1/2\pi$ або $1/\sqrt{2\pi}$).

Розглянемо властивості спектральної щільності коливань у їхньому зв'язку з властивостями функції $f(t)$. З цією метою підставимо експоненту, виражену у формі (11.25), до правих частин рівнянь (11.23), (11.24) і в такий спосіб виразимо *означення спектральної щільності та оберненого перетворення функції* $f(t)$ у тригонометричній формі:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)[\cos(\omega t) - i\sin(\omega t)]dt, \qquad (11.26)$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)[\cos(\omega t) + i\sin(\omega t)]d\omega. \qquad (11.27)$$

Покажемо, які властивості спектральної щільності випливають з рівнянь (11.26), (11.27). Ці властивості часто називають *принципами*, справедливими для спектральної щільності коливань, або, у більш широкому розумінні, фур'є-образу функції $f(t)$.

1. *Принцип (властивість) дуалізму*: якщо функція $F(\omega)$ є спектральною щільністю функції $f(t)$, то функція $f(-\omega)$ є спектральною щільністю функції $F(t)/2\pi$.

Доведення. Оскільки косинус є парною функцією свого аргументу, а синус – непарною, то зі співвідношення (11.27) безпосередньо випливає вираз

$$f(-t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)[\cos(\omega t) - i\sin(\omega t)]d\omega .$$

Оскільки в цьому виразі змінну інтегрування та аргумент функції $f(-t)$ можна позначити будь-якою літерою, то зробимо перепозначення $t \leftrightarrow \omega$ й отримаємо рівняння

$$f(-\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(t)[\cos(\omega t) - i\sin(\omega t)]dt ,$$

або, що теж саме,

$$f(-\omega) = \int_{-\infty}^{\infty} \frac{F(t)}{2\pi}[\cos(\omega t) - i\sin(\omega t)]dt . \qquad (11.28)$$

Права частина рівняння (11.28) відрізняється від правої частини рівняння (11.26) лише заміною функції $f(t)$ на функцію $F(t)/2\pi$. Оскільки рівняння (11.26) є означенням спектральної щільності тієї функції, що стоїть під знаком інтеграла в правій частині цього рівняння, то функція $f(-\omega)$ є спектральною щільністю функції $F(t)/2\pi$, яка стоїть у правій частині рівняння (11.28).

2. *Принцип (властивість) комплексної спряженості*:

$$F^*(\omega) = F(-\omega) , \qquad (11.29)$$

де значення функції $F^*(\omega)$ дорівнюють комплексно-спряженим значенням функції $F(\omega)$.

Доведення. Здійснивши комплексне спряження лівої та правої частини рівняння (11.26), перетворюємо його на рівняння

$$F^*(\omega) = \int_{-\infty}^{\infty} f^*(t)[\cos(\omega t) + i\sin(\omega t)]dt .$$

Оскільки значення функції $f(t)$ дійсні, то $f^*(t) = f(t)$. Ураховуючи, що косинус – парна, а синус – непарна функція свого аргументу, одержуємо з цього рівняння вираз

$$F^*(\omega) = \int_{-\infty}^{\infty} f(t)[\cos(-\omega t) - i\sin(-\omega t)]dt ,$$

а з означення спектральної щільності (11.26) безпосередньо випливає рівність

$$F(-\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)[\cos(-\omega t) - i\sin(-\omega t)]dt ,$$

що і доводить принцип комплексної спряженості (11.29).

3. *Принцип (властивість) дійсної парності та уявної непарності*: дійсна частина спектральної функції є парною функцією, а уявна – непарною, тобто:

$$\operatorname{Re} F(-\omega) = \operatorname{Re} F(\omega) , \tag{11.30}$$

$$\operatorname{Im} F(-\omega) = -\operatorname{Im} F(\omega) . \tag{11.31}$$

Доведення. Будь-яка функція, що набуває комплексних значень, є сумою дійсної частини та уявної частини, помноженої на уявну одиницю, тобто

$$F(\omega) = \operatorname{Re} F(\omega) + i \operatorname{Im} F(\omega) , \tag{11.32}$$

де $\operatorname{Re} F(\omega)$ та $\operatorname{Im} F(\omega)$ – функції, значення яких є дійсними числами. Завдяки цьому принцип комплексної спряженості (11.29) виражається у формі рівняння

$$[\operatorname{Re} F(\omega) + i \operatorname{Im} F(\omega)]^* = \operatorname{Re} F(-\omega) + iF(-\omega) ,$$

з якого випливає така рівність:

$$\operatorname{Re} F(\omega) - i \operatorname{Im} F(\omega) = \operatorname{Re} F(-\omega) + iF(-\omega) .$$

За кожного значення параметра ω ліва та права частини рівності є комплексними числами. Два комплексних числа дорівнюють одне одному тоді і лише тоді, коли дійсна частина першого числа дорівнює дійсній частині другого, а уявна частина першого – уявній частині другого, тобто тоді і лише тоді, коли справджуються рівності (11.30) та (11.31).

Зауваження 11.11. Важливим наслідком вищезазначених властивостей спектральної щільності є косинус-перетворення парної функції $f_S(t)$

$$F_S(\omega) = \int_{-\infty}^{\infty} f_S(t)\cos(\omega t)dt \tag{11.33}$$

та синус-перетворення непарної функції $f_A(t)$

$$F_A(\omega) = \int_{-\infty}^{\infty} f_A(t)\sin(\omega t)dt\,, \qquad (11.34)$$

де $f_A(t) = -f_A(-t)$, $f_S(t) = f_S(-t)$. Перетворення, обернені до (11.33) та (11.34), мають форму

$$f_S(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_S(\omega)\cos(\omega t)d\omega \qquad (11.35)$$

та

$$f_A(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_A(\omega)\sin(\omega t)d\omega\,. \qquad (11.36)$$

(Нижній індекс у позначенні кожної із функцій вказує на те, що парну функцію називають симетричною (*Symmetric*), а непарну – антисиметричною (*Antisymmetric*) до зміни знака її аргументу). Вирази (11.35) та (11.36) легко вивести із (11.27), якщо взяти до уваги, що інтеграли від непарних функції, обчислені в межах від $-\infty$ до ∞, дорівнюють нулю.

11.7. Важливі приклади перетворення Фур'є

1. *Спектральна щільність прямокутного імпульсу електричної напруги*:

$$U(t) = \begin{cases} 1\text{ В}, & if \quad -T/2 \le t \le T/2, \\ 0\text{ В}, & otherwise, \end{cases} \qquad (11.37)$$

зображеного на рис. 11.4, *а*.

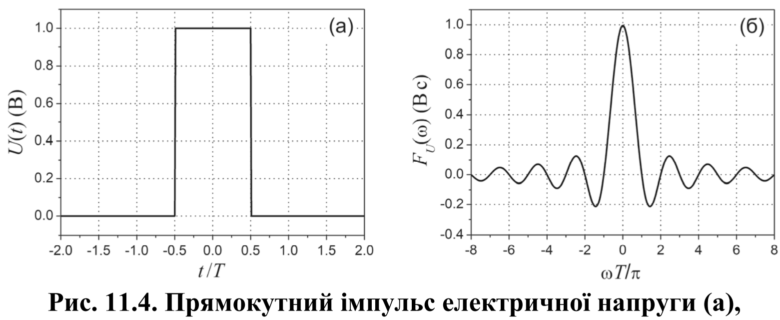

Рис. 11.4. Прямокутний імпульс електричної напруги (а), та його спектральна щільність (б)

Знайдемо спектральну щільність прямокутного імпульсу, виходячи з означення спектральної щільності, вираженого формулою (11.23):

$$F_U(\omega) = \int_{-\infty}^{\infty} U(t) e^{-i\omega t} dt \,.$$

Узявши до уваги вирази (11.37) та (11.25), отримаємо рівності:

$$F_U(\omega) = \int_{-T/2}^{T/2} e^{-i\omega t} dt = \int_{-T/2}^{T/2} \cos(\omega T) dt - \int_{-T/2}^{T/2} \sin(\omega T) dt \,.$$

Виконавши інтегрування, знайдемо остаточну формулу для спектральної щільності прямокутного імпульсу електричної напруги (рис. 11.4, *б*):

$$F_U(\omega) = \frac{\sin(\omega T)}{\omega} \text{ В} \cdot \text{с} \,. \tag{11.38}$$

2. *Фур'є-образ функції Гаусса, що описує "нормальний" розподіл імовірностей:*

$$P(t) = \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{t^2}{2\tau^2}} \,. \tag{11.39}$$

Сталу величину $\tau = \text{const}$ називають параметром нормального розподілу ймовірностей $P(t)$. У теорії імовірностей показано, що цей розподіл імовірностей притаманний великій кількості випадкових процесів[1]. Графіки функції (11.39) наведено на рис. 11.5, *а* для різних значень параметра. Ці графіки показують, що зменшення величини параметра τ веде до зменшення "півширини гаусівського піку"[2].

Знайдемо спектральну щільність функції Гаусса, виходячи з означення спектральної щільності:

$$F_P(\omega) = \int_{-\infty}^{\infty} P(t) e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi\tau^2}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2\tau^2} - i\omega t} dt \,.$$

---

[1] Суть змінної t залежить від того, який саме випадковий процес описується функцією Гаусса, а тут цю змінну позначено так, щоб зберегти позначення, введене в підрозд. 11.6.

[2] Тобто до зменшення різниці $t_2 - t_1$ між тими значеннями змінної t, за яких значення функції Гаусса є меншим від максимального в $2^{1/2}$ разів.

Використавши тотожність

$$-\frac{t^2}{2\tau^2} - i\omega t \equiv -\frac{1}{2}\left(\frac{t}{\tau} + i\omega t\right)^2 - \frac{\omega^2\tau^2}{2},$$

одержимо вираз

$$F_P(\omega) = \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{\omega^2\tau^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{t}{\tau}+i\omega\tau\right)^2} dt.$$

Зробивши заміну змінних

$$\frac{1}{\sqrt{2}}\left(\frac{t}{\tau} + i\omega t\right) \equiv u, \quad dt = \sqrt{2}\tau du,$$

виразимо спектральну щільність функції Гаусса через інтеграл Пуассона

$$\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$$

та отримаємо остаточну формулу:

$$F_P(\omega) = e^{-\frac{\omega^2\tau^2}{2}}. \qquad (11.40)$$

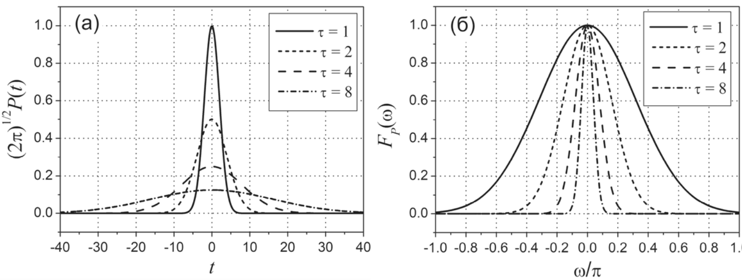

Рис. 11.5. Графіки функції Гаусса (а) та її спектральної щільності (б), побудовані для різних значень параметра τ

Графіки функції (11.40) наведено на рис. 11.5, *б* для різних значень параметра τ. Порівняння цих графіків з графіками, наведеними на рис. 11.5, *а*, показує, що зменшення півширини гауссівського піку веде до збільшення півширини піку спектральної щільності.

3. *Фур'є-образ δ-функції Дірака*. Зробивши у функції Гаусса (11.39) заміну $x = t + a$, $a = \text{const}$, отримаємо функцію

$$P(x) = \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(x-a)^2}{2\tau^2}},$$

яку також називають функцією Гаусса. Неважко помітити, що значення цієї функції в точці $x = a$ обернено пропорційне параметру τ, тому величина $P(a)$ необмежено зростає, якщо величина τ прямує до нуля. Якщо ж змінна x не дорівнює a, то коли величина τ прямує до нуля функція, $P(x)$ експоненціально зменшується. Ці властивості функції Гаусса виражають умовною рівністю

$$\lim_{\tau \to 0} P(a) = \begin{cases} \infty, & if\ x = a, \\ 0, & otherwise. \end{cases} \qquad (11.41)$$

Окрім того, легко упевнитись, що

$$\int_{-\infty}^{\infty} P(x) dx = 1, \qquad (11.42)$$

а отже, виконується умова

$$\lim_{\tau \to 0} \int_{-\infty}^{\infty} P(x) dx = 1. \qquad (11.43)$$

Умовні рівності на зразок рівностей (11.41), (11.42) справджуються і для багатьох інших функцій, тому використовується поняття *узагальненої функції* $\delta(x - a)$, означенням якої є рівності

$$\delta(x - a) \overset{def}{\equiv} \begin{cases} \infty, & if\ x = a, \\ 0, & otherwise, \end{cases} \qquad (11.44)$$

$$\int_{-\infty}^{\infty} \delta(x - a) dx = 1. \qquad (11.45)$$

Функцію $\delta(x - a)$ називають "δ-функція Дірака", або "δ-функція".

Зауваження 11.10. Співвідношення (11.44), (11.45) широко використовуються в літературі з фізико-технічних наукових дисциплін, але вважаються некоректними з погляду математичного аналізу. Тому математики вважають означенням δ-функції співвідношення

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) dx = f(a). \qquad (11.46)$$

Прихильність фізиків та інженерів до співвідношень (11.44), (11.45) пояснюється зручністю використання таких виразів для проведення та пояснення математичних розрахунків. Продемонструємо цю зручність, вивівши формулу (11.46) зі співвідношень (11.44) та (11.45). Співвідношення (11.44) показує, що підінтегральна функція того інтеграла, що стоїть у лівій частині рівності (11.46), дорівнює нулю в усіх точках, окрім точки $x = a$, а отже, те, яких значень набуває функція $f(x)$ у точках $x \neq a$, не впливає на величину інтеграла і можна замінити функцію $f(x)$ на число $f(a)$. Тому справедлива рівність

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = \int_{-\infty}^{\infty} f(a)\delta(x-a)dx .$$

Винісши число $f(a)$ за знак інтеграла, отримаємо рівність

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a)\int_{-\infty}^{\infty} \delta(x-a)dx .$$

Завдяки властивості (11.44) ця рівність з очевидністю перетворюється на (11.45).

Зауваження 11.11. Оскільки в граничному випадку $\tau \to 0$ функція Гаусса перетворюється на δ-функцію Дірака[1], то кажуть, що *функція Гаусса здійснює інтегральне представлення δ-функції.*

Знайдемо спектральну щільність F_δ нескінченно короткого в часі сигналу $\delta(t)$, тобто здійснімо перетворення Фур'є δ-функції в окремому випадку $a = 0$:

$$F_\delta(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t}\delta(t)dt .$$

Згідно із формулою (11.46) маємо

$$\int_{-\infty}^{\infty} e^{-i\omega t}\delta(t)dx = e^{-i\omega t}\,\big|_{t=0} = 1 ,$$

а отже,

$$F_\delta(\omega) = 1 .$$

---

[1] Порівняйте праву частину рівнянь (11.41) із правою частиною рівняння (11.44), а рівняння (11.43) з рівнянням (11.45).

Цей результат випливає також із формули (11.40) для спектральної щільності функції Гаусса (11.39), яка перетворюється на $\delta(t)$ у граничному випадку $\tau \to 0$. Згідно із цією формулою

$$\delta(t) = \lim_{\tau \to 0} e^{-\frac{\omega^2 \tau^2}{2}} = 1 .$$

Той факт, що спектральна щільність δ-функції є сталою величиною, можна було передбачити з рис. 11.5, який вказує на те, що зменшення параметра τ зменшує півширину гауссівського піку, уповільнюючи водночас залежність спектральної щільності від частоти, тобто поступово перетворює функцію $F_\delta(\omega)$ на константу.

ЗАВДАННЯ ТА ЗАПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

До розділу 1

1. Зробіть простий рисунок, який покаже, що операції додавання векторів притаманна асоціативність. (Подібно до того, як рис. 1.2 вказує на комутативність цієї операції).

2. Розгляньте розклади радіус-векторів точок *A* та *B* по ортах прямокутної декартової системи координат. Застосувавши правило трикутника та властивості операцій додавання векторів і множення вектора на число, доведіть, що декартові координати геометричного вектора, початок якого лежить у точці *A*, а кінець – у точці *B*, дорівнюють різницям координат цих точок.

3. Кут між векторами $\mathbf{a} = \mathbf{e}_x + \mathbf{e}_y - \mathbf{e}_z$ та $\mathbf{b} = b_x\mathbf{e}_x - 2\mathbf{e}_y + \mathbf{e}_z$ дорівнює 90°. Чому дорівнює координата b_x? Чому дорівнює кут між сумою цих векторів та віссю OZ?

До розділу 2

1. Чи існує добуток AB матриці другого порядку A та матриці B, порядку 2×3?

2. Чому дорівнює елемент a_{zz} матриці $A(a_{ik})$, що описує поворот системи координат на кут α навколо осі OZ?

3. Обчисліть добуток матриці повороту на 180° навколо осі OZ та матриці повороту на 180° навколо осі OY. Який поворот описує ця матриця?

До розділу 3

1. Чи є лінійним простором сукупність компланарних геометричних векторів?

2. Доведіть справедливість співвідношення $\mathbf{x}(\lambda\mathbf{y}) = \lambda^*(\mathbf{xy})$, виходячи з означення скалярного добутку елементів комплексного лінійного простору.

3. Зваживши на співвідношення $\mathbf{x}(\lambda\mathbf{y}) = \lambda^*(\mathbf{xy})$, доведіть, що в тому разі, якщо $\mathbf{e}_j\mathbf{e}_k = \delta_{jk}$, вектори комплексного простору Евкліда $\mathbf{x} = \mathbf{e}_1 + i\mathbf{e}_2$ та $\mathbf{y} = \mathbf{e}_1 - i\mathbf{e}_2$ – ортогональні.

До розділу 4

1. Виходячи з означення символу Леві-Чивіти ε_{jkl} та символу Кронекера δ_{kl}, доведіть, що для будь-яких значень індексів j,k,l справджуються рівності $\varepsilon_{jkl}\delta_{kl} = 0$.

2. У скільки разів збільшиться визначник матриці третього порядку, якщо помножити цю матрицю на число 2?

3. Обчисліть за допомогою формули (4.39) елементи матриці A^{-1}, оберненої до матриці $A = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}$. Перевірте результат обчислення, обрахувавши добуток $A^{-1}A$. Чи можливо було б виконати це обчислення, якщо елемент a_{21} дорівнював би одиниці?

До розділу 5

1. Відомо, що квадратна матриця другого порядку A складена з ненульових елементів a_{jk}, E – одинична матриця, $\mathbf{x}$ та $\mathbf{b}$ – вектор-стовпці, кожен з яких складений із двох елементів. Виразіть матричне рівняння $(A - \lambda E)\mathbf{x} = \mathbf{b}$ у формі неоднорідної лінійної системи двох алгебраїчних рівнянь.

2. Відомо, що третє рівняння системи чотирьох неоднорідних лінійних алгебраїчних рівнянь є наслідком трьох інших. Яким може бути ранг основної матриці системи?

До розділу 6

1. Чи можна вважати, що матриця повороту декартової системи координат навколо заданого напрямку відповідає оператору, що діє на радіус-вектори точок простору? Якщо "так", то чи є цей оператор а) лінійним, б) ортогональним?

2. Трьом лінійним операторам відповідають матриці:

$$A_1 = \begin{pmatrix} 1 & 1+i \\ 1-i & i \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1+i \\ 1-i & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 1-i \\ 1-i & 1 \end{pmatrix}.$$

Який (які) із цих операторів є оператором (операторами) Ерміта?

До розділу 7

1. Перетворіть матрицю A білінійної форми

$$f(\mathbf{x},\mathbf{x}) = x_1y_1 + 3x_2y_2 - x_3y_3 + 2x_1y_2 + 6x_2y_1 - 4x_2y_3 + 4y_2x_3$$

на суму симетричної та антисиметричної матриці, A_S та A_A, відповідно. Чому дорівнює визначник антисиметричної матриці?

2. Запишіть вираз для квадратичної форми, матриця якої дорівнює симетричній матриці A_S. Чи є ця квадратична форма додатно визначеною?

До розділу 8

1. Якою має бути функція $z[y(x)]$, щоб після заміни функції $y(x)$ на функцію $z[y(x)]$ рівняння $xy'+1 = e^{x-y}$ перетворилося на неоднорідне лінійне рівняння для $z(x)$?

2. Перетворіть операторне рівняння $\hat{H}\psi(x) = E\psi(x)$, у якому $\hat{H} = -(\hbar^2/2m)d^2/dx^2$, $\hbar$, m та E – сталі параметри, на однорідне лінійне рівняння другого порядку зі сталими коефіцієнтами. Як має бути пов'язаний з параметрами рівняння коефіцієнт k, щоб функція $\psi(x) = e^{ikx}$ була розв'язком цього рівняння, а E – власним числом оператора $\hat{H}$?

До розділу 9

1. Коефіцієнти трьох різних систем однорідних лінійних ДР утворюють матриці

$$A_1=\begin{pmatrix}1 & 4\\ -3 & 0\end{pmatrix},\ A_2=\begin{pmatrix}2 & 2\\ 2 & -2\end{pmatrix},\ A_3=\begin{pmatrix}-2 & -2\\ 6 & -6\end{pmatrix}.$$

Зі скількох рівнянь складається кожна з цих систем? Чому дорівнюють корені характеристичного рівняння кожної системи?

2. Зінтегруйте систему, коефіцієнти якої утворюють матрицю A_2.

До розділу 10

1. Доведіть, що точки фазової площини $P_1=(4,1)$, $P_2=(2,2)$, $P_3=(-2,-2)$ є точками спокою динамічної системи, рух якої описується рівняннями $\begin{cases}\dot{x}=xy-4\\ \dot{y}=(x-4)(y-x)\end{cases}$, та лінеаризуйте ці рівняння в околі кожної з точок спокою.

2. Дослідіть стійкість стаціонарних станів, яким відповідають точки спокою $P_1=(4,1)$, $P_2=(2,2)$, $P_3=(-2,-2)$.

До розділу 11

1. Що треба зробити, щоб розкласти в ряд Фур'є або представити у формі інтеграла Фур'є функцію $f(x)$, задану на несиметричному інтервалі $x_1\le x\le x_2$?

2. Якщо випадковій величині t притаманний нормальний розподіл імовірностей $P(t)$, то середнє значення функції $f(t)$ дорівнює інтегралу $\int_{-\infty}^{\infty} f(t)P(t)dt$. Чому дорівнює середнє значення квадрата випадкової величини, якій притаманний нормальний розподіл імовірностей?

ЛІТЕРАТУРА

1. Слепушенко В. Н. Аналитическая геометрия / В. Н. Слепушенко, И. А. Яковлев. Изд. 2-е. – Харьков : ХВВКИУ, 1964.

2. Придатченко Ю. В. Векторна алгебра та аналітична геометрія / Ю. В. Придатченко, В. А. Львов, С. В. Ефименко. – Київ : ВПЦ "Київський університет", 2010.

3. Lancaster P. Theory of Matrices / P. Lancaster. – N.-Y. ; London : Academic Press, 1969.

4. Fichtenholz G. M. Funktional Series / G. M. Fichtenholz. – London : Gordon&Breach, 1970. – 372 p.

5. Read M. Functional Analysis / M. Read, B. Simon. – N.-Y. ; London : Academic Press, 1980. – 416 p.

6. Гриньов Б. В. Вища алгебра / Б. В. Гриньов, І. К. Кириченко. – Харків : Гімназія, 2008. – 182 с.

7. Anton H. Elementary Linear Algebra / H. Anton, C. Rorres. 9th ed. N.-Y. : Wiley, 2005. – 416 p.

8. Messiah A. Quantum Mechanics / A. Messiah. Amsterdam : North-Holland Publishing Company, 1961. – Vol. I, Ch. VII. P. 243–293.

9. Gel'fand I. M. Lectures on Linear Algebra / I. M. Gel'fand. – N.-Y. : Dover Publications, 1989. – 208 p.

10. Landau, L. D. Course of theoretical physics, Vol. 1, Mechanics / L. D. Landau, E. M. Lifshitz. – Oxford, UK : Pergamon Press, 1969. – 167 p.

11. Elsgolts L. Differential Equations and the Calculus of Variations / L. Elsgolts. – Forest Grove, Oregon, USA : University Press of the Pacific, 2003. – 444 p.

12. Кривошея С. А. Диференціальні та інтегральні рівняння / С. А. Кривошея, М. О. Перестюк, В. М. Бурим. – Київ : Либідь, 2004. – 407 с.

13. Львов В. А. Просто про складне: звичайні диференціальні рівняння : навч. посіб. / В. А. Львов, А. О. Косогор, Д. Л. Попадюк. – К. : ВПЦ "Київський університет", 2021.

14. In a search for effective giant magnetoelectric coupling: Magnetically induced elastic resonance in Ni-Mn-Ga/P(VDF-TrFE) composites / P. Martins, A. C. Lima, V. A. L'vov et al. // Applied Materials Today, 2022. – Vol. 29. – 101682.

15. Giant magnetoelectric effect of Ni–Mn-Ga/piezopolymer composites tailored by a martensitic transformation / V. A. L'vov, P. Martins, N. Pereira et al. // Composite Science and Technology, 2023. – Vol. 241. – 110101.

ЗМІСТ

Навчальне видання

ЛЬВОВ Віктор Анатолійович
КОСОГОР Анна Олексіївна
БАРАБАНОВ Олександр Валерійович

ПРОСТО ПРО СКЛАДНЕ
ОСНОВИ ТА ЗАСТОСУВАННЯ ЛІНІЙНОЇ АЛГЕБРИ

Підручник

Редактор *Л. Львова*

Оригінал-макет виготовлено ВПЦ "Київський університет"

Формат 60x84$^{1/16}$. Ум. друк. арк. 13,02. Наклад 100. Зам. № 223-10760.
Гарнітура Times New Roman. Папір офсетний. Друк офсетний. Вид. № Рф8.
Підписано до друку 07.02.24

Видавець і виготовлювач
ВПЦ "Київський університет"
Б-р Тараса Шевченка, 14, м. Київ, 01601, Україна
☎ (38044) 239 32 22; (38044) 239 31 58; (38044) 239 31 28
e-mail: vpc@knu.ua; vpc_div.chief@univ.net.ua; redaktor@univ.net.ua
http: vpc.knu.kiev.ua
Свідоцтво суб'єкта видавничої справи ДК № 1103 від 31.10.02