Київський національний університет імені Тараса Шевченка
Механіко-математичний факультет

**Євген Бондаренко**
**Олександра Десятерик**
**Андрій Олійник**
**Андрій Руссєв**

# Конкретна математика

Навчальний посібник

для студентів механіко-математичного факультету

Київ–2026

**Бондаренко Є. В., Десятерик О. О., Олійник А. С., Руссєв А. В.** Конкретна математика. Навчальний посібник для студентів механіко-математичного факультету. — К., 2026. — 230 с.

**Рецензенти**: канд. фіз.-мат. наук, доц. Касімова Н. В.
канд. фіз.-мат. наук, доц. Ловейкін А. В.

Рекомендовано до друку вченою радою механіко-математичного факультету (протокол № 10 від 6 березня 2026 року)

Ухвалено науково-методичною радою Київського національного університету імені Тараса Шевченка (протокол № 08-26 від 18 червня 2026 року)

# Зміст

# Вступ

Курс конкретної математики читається для студентів першого курсу механіко-математичного факультету спеціальностей «Математика» та «Статистика». В цьому курсі розглядаються методи доведень, основні поняття математичної логіки, елементи теорії чисел, дискретної математики та алгебри.

Метою посібника є ознайомити читача з основними поняттями різних розділів математики, які складають основу для подальшого, поглибленого вивчення цих розділів та використання отриманих знань для дослідження різних математичних об'єктів.

Розділ 1 присвячений вивченню елементів математичної логіки та основним методам доведень. Розглянуто метод математичної індукції, який широко використовується в усіх розділах математики і є одним з основних методів доведень. Як основа математичної логіки розглядається логіка висловлювань. У даному розділі введені основні логічні зв'язки, їх задання за допомогою таблиць істинності. Ознайомлення з основними поняттями математичної логіки є дуже важливим як для абстрактої математики (наприклад, логіка дає основу для побудови формальної теорії множин), так і для практичних застовувань (наприклад, у роботі з теорією алгоритмів).

Розділ 2 містить основні поняття теорії чисел. У ньому доведено теорему про ділення з остачею для цілих чисел. Розглянуто властивості подільності цілих чисел. Досліджено такі важливі поняття, як НСД та НСК для цілих чисел. Розглянуто розширений алгоритм Евкліда для цілих чисел. Також приділено увагу простим числам, розкладу цілого числа на прості множники та використанню даного розкладу.

Розділ 3 розглядає комплексні числа. У розділі вводиться поняття комплексного числа, операції над комплексними числами. Розглянуто алгебраїчні та геометричні властивості комплексних чисел, їх коренів і,

зокрема, коренів з одиниці.

У розділі 4 наводиться та обґрунтовується метод розв'язання лінійних діофантових рівнянь, отримані явні формули для коренів. Також вводяться поняття модулярної арифметики, оскільки теорія чисел, інші розділи алгебри (наприклад, теорія груп та теорія полів), широко оперують ними.

Розділи 5–7 розповідають про матриці, підстановки та визначники. Введено поняття матриці та визначено арифметичні операції над матрицями. Окрему увагу приділено оберненим матрицям та їх властивостям. Розглянуто елементарні перетворення матриць та їх використання для перевірки оборотності матриць. Показано можливість використання оберненої матриці для розв'язання систем лінійних рівнянь. Далі розглядаються поняття підстановки, множення підстановок та їх властивості. Далі розглядаються визначники матриць і способи їх обчислення.

Розділи 8–10 присвячено многочленам та їх властивостям. Розглянуто алгоритм Евкліда для многочленів та його зв'язок з НСД двох многочленів. Для читача має бути цікавим прослідкувати аналогію з відповідними поняттями для цілих чисел, які розглядалися у попередніх розділах. Значну увагу приділено властивостям подільності та властивостям незвідності многочленів. Дослідження властивостей подільності многочленів є важливими, зокрема, у теорії кілець при побудові фактор-кілець. Введено загальні поняття результанта та дискримінанта, наведено властивості які, їх пов'язують. На завершення наводяться різні ознаки незвідності многочленів, які можна використовувати за різних умов, що накладені на многочлен.

У навчальному посібнику наведені вправи для самостійної роботи, які мають допомогти читачу перевірити засвоєння нового матеріалу та покращити його розуміння. Деякі вправи спонукають читача глибше зануритися у матеріал відповідної теми. У списку літератури наведено джерела як базового рівня, так і призначені для глибшого вивчення понять, які розглянуто у посібнику.

# Розділ 1

# Логіка висловлювань та доведення

## 1.1 Висловлювання

Поняття висловлювання є основним в логіці висловлювань. Висловлювання бувають прості і складені. Поняття простого висловлювання є первісним і в межах логіки висловлювань не визначається. Змістовно, хоча й дещо спрощено, можна вважати, що прості висловлювання — це прості речення, в яких щось стверджується або заперечується. Прикладами висловлювань є такі речення:

а) «Діагоналі довільного паралелограма рівні між собою»

б) «Дія додавання раціональних чисел є асоціативною»

в) «Множина простих чисел є скінченною»

г) «Кожне натуральне число розкладається в добуток простих множників»

В твердженнях б) і г) виражено правильну думку, а в твердженнях а) і в) — неправильну. Висловлюваннями не будуть запитальні та окличні речення, а також означення, адже в них лише про щось домовляються.

Кожне просте висловлювання має певне логічне значення: «істина» чи «хиба», які далі позначатимемо символами «1» і «0» відповідно. На рівні алгебри висловлювань ніякі інші властивості простих висловлювань, а також їх внутрішня будова не враховуються.

З простих висловлювань можна конструювати складені, застосовуючи дії над висловлюваннями. Вони є уточненням найбільш уживаних

конструкцій для побудови складних речень за допомогою сполучників чи мовних зворотів. Основними діями над висловлюваннями є:

1) кон'юнкція (логічний добуток) — утворена на основі сполучника «і»;

2) диз'юнкція (логічна сума) — утворена на основі сполучника «або», вжитого у з'єднувальному значенні;

3) імплікація — утворена на основі сполучення «якщо ..., то ...», причому в разі хибності умови результат застосування імплікації вважається істинним (така імплікація називається матеріальною);

4) еквіваленція — утворена на основі звороту «... тоді й тільки тоді, коли ...»;

5) заперечення — утворене на основі звороту «неправда, що ...».

Дії 1)–4) є бінарними, дія 5) — унарна. Позначаються вони символами $\wedge$ (кон'юнкція), $\vee$ (диз'юнкція), $\to$ (імплікація), $\leftrightarrow$ (еквіваленція), $\neg$ (заперечення). В імплікації $x \to y$ висловлювання $x$ називається антецедентом (умовою), а висловлювання $y$ — консеквентом (наслідком). Дії однозначно характеризуються своїми таблицями логічних значень (табл. 1.1, 1.2).

| $x$ | $y$ | $x \wedge y$ | $x \vee y$ | $x \to y$ | $x \leftrightarrow y$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 |

Табл. 1.1

| $x$ | $\neg x$ |
|---|---|
| 1 | 0 |
| 0 | 1 |

Табл. 1.2

Складені висловлювання, які конструюються з простих за допомогою дій 1)–5), можуть мати довільну довжину. Крім того, поняття змісту речень в логіці висловлювань не розглядається; кожне висловлювання характеризується лише своїм логічним значенням. А тому висловлюваннями можуть бути конструкції, які як речення є беззмістовними.

Наприклад, застосувавши імплікацію до висловлювань

$$a = \text{«Париж — столиця Франції»}$$

та

$$b = \text{«Дніпро — найдовша річка України»},$$

дістанемо складене висловлювання $a \to b$, яке природно прочитати, як речення «Якщо Париж — столиця Франції, то Дніпро — найдовша річка України», яке позбавлене змісту. Проте це речення є висловлюванням і його логічне значення дорівнює 1, адже воно утворене з двох істинних простих висловлювань (див. табл. 1.1).

Якщо прості висловлювання позначено літерами якогось алфавіту, то будь-якому складеному висловлюванню, що з них побудоване, відповідає певна формула. Наприклад, висловлювання «Якщо добуток $xy$ натуральних чисел $x$ та $y$ ділиться на просте число $p$, то принаймні одне з них також ділиться на $p$» записується у вигляді формули таким чином. Введемо позначення

$$\begin{aligned} a &= \text{«добуток } xy \text{ натуральних чисел } x \text{ та } y \\ &\qquad \text{ділиться на просте число } p\text{»}, \\ b &= \text{«число } x \text{ ділиться на } p\text{»}, \\ c &= \text{«число } y \text{ ділиться на } p\text{»}. \end{aligned}$$

Скороченим записом початкового висловлювання тоді буде формула

$$a \to (b \vee c).$$

Так побудовані формули називаються формулами логіки висловлювань або формулами алгебри логіки. Строге визначення формул дається таким чином.

Алфавітом алгебри логіки називається множина символів чотирьох типів:

i) нескінченна послідовність $x_1$, $x_2$, $x_3$, ... символів для позначення простих висловлювань (символи змінних висловлювань);

ii) символи 1 (істинне висловлювання) і 0 (хибне висловлювання) — сталі логіки висловлювань;

iii) символи $\wedge, \vee, \to, \leftrightarrow, \neg$ для позначення основних логічних дій;

iv) $(,)$ (ліва і права дужки) — технічні символи.

**Означення 1.1.** а) кожен символ послідовності i) і символи з ii) є формулами логіки висловлювань;

б) якщо $F_1, F_2$ — формули, то $(F_1 \wedge F_2)$, $(F_1 \vee F_2)$, $(F_1 \to F_2)$, $(F_1 \to F_2)$, $(F_1 \leftrightarrow F_2)$, $(\neg F_1)$ є формулами логіки висловлювань;

в) послідовність символів алфавіту логіки висловлювань буде формулою логіки висловлювань тоді й лише тоді, коли вона є формулою або внаслідок пункту а) цього означення або внаслідок пункту б) цього означення.

Наприклад, послідовності символів

$$(x_1 \to (x_2 \to (\neg x_3))),$$

$$((\neg(x_1 \to x_2)) \wedge (x_2 \leftrightarrow x_3)),$$

$$(((\neg x_1) \leftrightarrow (\neg x_2)) \to x_3)$$

є формулами, а послідовності

$$\neg x_1) \to x_2(x_3 \vee x_4))),$$

$$\neg(\to x_1 \wedge x_2)$$

формулами логіки висловлювань не будуть.

Кажуть, що символ $x$ послідовності i) входить у формулу $F$, якщо він зустрічається в записі цієї формули. В такому випадку формула $F$ залежить явно від $x$. В деяких випадках зручно вважати, що формула $F$ також залежить, але вже неявно, від $x$, якщо цей символ не входить у формулу $F$. Таким чином, можна вважати, що довільна формула $F$ залежить від довільної скінченної підмножини символів послідовності i), яка містить усі ті символи цієї послідовності, що входять у $F$. Зокрема, для усіх достатньо великих $n$ формула $F$ залежить від $x_1$, $x_2$, $\ldots$, $x_n$. Позначатимемо це так: $F(x_1, x_2, \ldots, x_n)$. Зауважимо, що для довільного скінченного числа формул можна вважати, що вони залежать від одних і тих же символів послідовності i).

Щоб встановити, чи буде дана послідовність символів алфавіту логіки висловлювань формулою, чи ні, можна скористатись методом послідовного аналізу будови формули. Для цього виділяємо в формулі всі символи простих висловлювань, а потім послідовно розглядаємо підформули даної формули все більшої довжини, які будуються з раніше виділених на основі пункту b) означення формули. Процес продовжуємо, поки не дістанемо початкову формулу. В цьому випадку аналізована послідовність символів алфавіту є формулою; в іншому разі (коли процес аналізу обірветься раніше) вона формулою не буде.

## 1.2 Тавтології

Якщо у формулу логіки висловлювань

$$F(x_1, x_2, \ldots, x_n)$$

замість змінних $x_1$, $x_2$, …, $x_n$ підставити, відповідно, висловлювання

$$a_1, a_2, \ldots, a_n,$$

то дістанемо деяке висловлювання $F(a_1, a_2, \ldots, a_n)$. Це висловлювання має певне логічне значення, яке однозначно обчислюється за логічними значеннями висловлювань $a_1$, $a_2$, …, $a_n$. При обчисленні логічного значення висловлювання

$$F(a_1, a_2, \ldots, a_n)$$

досить знати лише логічні значення складових частин

$$a_1, a_2, \ldots, a_n.$$

Підставляючи у формулу $F(x_1, x_2, \ldots, x_n)$ замість $x_1$, $x_2$, …, $x_n$ логічні константи 0, 1, дістанемо вираз, складений з символів 0, 1 і символів логічних дій. Його значення можна обчислити, користуючись таблицями 1.1 і 1.2. У такий спосіб приходимо до поняття інтерпретації формули логіки висловлювань.

**Означення 1.2.** Інтерпретацією формули $F(x_1, x_2, \ldots, x_n)$ алгебри логіки називається довільна заміна

$$x_1 \rightsquigarrow \alpha_1, x_2 \rightsquigarrow \alpha_2, \ldots, x_n \rightsquigarrow \alpha_n \tag{1.1}$$

змінних цієї формули на логічні константи. Результатом інтерпретації (1.1) є логічна константа $F(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Кожна формула, яка залежить від $n$ змінних, може бути проінтерпретована $2^n$ різними способами, адже саме стільки послідовностей вигляду $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ можна утворити з логічних констант 0, 1. Список найможливіших інтерпретацій формули $F(x_1, x_2, \ldots, x_n)$ і їх результатів записують у вигляді таблиці, яка називається таблицею значень або таблицею істинності формули $F$ (табл. 1.3).

| $x_1$ | $x_2$ | $\ldots$ | $x_n$ | $F(x_1, x_2, \ldots, x_n)$ |
|---|---|---|---|---|
| 1 | 1 | $\ldots$ | 1 | $F(1, 1, \ldots, 1)$ |
| 1 | 1 | $\ldots$ | 0 | $F(1, 1, \ldots, 0)$ |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots\ldots\ldots\ldots$ |
| 0 | 0 | $\ldots$ | 0 | $F(0, 0, \ldots, 0)$ |

Табл. 1.3

Впорядкований набір $\bar{a} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ логічних сталих 0, 1 називається булевим вектором, а число $n$ називається довжиною булевого вектора $\bar{a}$. Кожна інтерпретація

$$x_1 \rightsquigarrow \alpha_1, x_2 \rightsquigarrow \alpha_2, \ldots, x_n \rightsquigarrow \alpha_n$$

формули $F(x_1, x_2, \ldots, x_n)$ однозначно визначається булевим вектором

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

і навпаки.

За допомогою поняття інтерпретації всі формули логіки висловлювань поділяються на три класи: тотожно істинні, виконливі і тотожно хибні.

**Означення 1.3.** Формула $F$ називається:

а) тотожно істинною (або тавтологією), якщо вона набуває значення 1 при кожній інтерпретації своїх змінних;

б) виконливою, якщо вона набуває значення 1 принаймні при одній інтерпретації змінних;

в) тотожно хибною, якщо вона при жодній інтерпретації змінних не має значення 1.

Кожна тотожно істинна формула є виконливою, але не навпаки. Заперечення тотожно хибної формули є тотожно істинною, а заперечення тотожно хибної формули є тотожно істинною формулою.

Конструювати нові тавтології з відомих можна, застосовуючи до них різноманітні правила побудови тавтологій. Основними з них є правила підстановки.

Нехай $F$ — деяка формула логіки висловлювань, $A$ — підформула формули $F$, тобто така її частина, яка сама є формулою в розумінні означення 1.1. У формулі $F$ підформула $A$ може зустрічатися декілька разів. Кожну з таких підформул називатимемо входженням $A$ в $F$. Наприклад, формула

$$A = x_1 \to \neg x_2$$

має три входження у формулу

$$F = ((\underline{x_1 \to \neg x_2}) \to x_3) \vee ((\underline{x_1 \to \neg x_2}) \wedge (x_3 \to (\underline{x_1 \to \neg x_2}))),$$

які виділено, як підкреслені фрагменти $F$.

**Означення 1.4.** Підстановкою формули $B$ в формулу $F$ замість формули $A$ називається послідовна заміна усіх входжень формули $A$ в $F$ на формулу $B$.

Результатом підстановки є нова формула, яка позначається $\Pi_A^B F$. Зокрема, якщо здійснюється підстановка замість окремої змінної $x$, то її результатом буде формула $\Pi_x^B F$.

Наприклад, розглянемо формулу

$$F = (x_1 \to x_2) \vee (x_3 \to \neg x_2).$$

Нехай $A = x_1 \to x_2$, $B = (x_1 \wedge x_2) \to x_3$. Тоді

$$\Pi_A^B F = ((x_1 \wedge x_2) \to x_3) \vee (x_3 \to \neg x_2),$$
$$\Pi_{x_2}^A F = (x_1 \to (x_1 \to x_2)) \vee (x_3 \to \neg(x_1 \to x_2)),$$
$$\Pi_{x_2}^B = (x_1 \to ((x_1 \wedge x_2) \to x_3)) \vee (x_3 \to \neg((x_1 \wedge x_2) \to x_3)).$$

**Теорема 1.1** (Перший принцип підстановки для тавтологій)**.** *Нехай $F$ — тавтологія, $x$ — деяка змінна формули $F$ і $A$ — довільна формула. Тоді формула $\Pi_x^A F$ також є тавтологією.*

*Доведення.* Нехай $x_1, x_2, \ldots, x_n$ — змінні, від яких залежить формула $F$. Не обмежуючи загальності будемо вважати, що підстановка формули $A$ в $F$ здійснюється замість змінної $x_1$. Для довільного булевого вектора $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in L^n$ значення $\Pi_x^A F(\alpha_1, \alpha_2, \ldots, \alpha_n)$ можна знайти, обчисливши спочатку значення $A(\alpha_1, \alpha_2, \ldots, \alpha_n) = \beta$, а потім значення

$$F(\beta, \alpha_2, \ldots, \alpha_n).$$

Оскільки $F$ — тавтологія, то воно дорівнює 1. Звідси випливає, що $\Pi_x^A F$ — тавтологія. □

## 1.3 Рівносильність

**Означення 1.5.** Областю істинності формули $F(x_1, \ldots, x_n)$ логіки висловлювань називається множина всіх булевих векторів довжини $n$, при підстановці яких замість змінних у формулу $F$ дістаємо логічне значення 1. Областю хибності формули $F(x_1, \ldots, x_n)$ називається множина тих векторів $(\alpha_1, \ldots, \alpha_n)$, для яких $F(\alpha_1, \ldots, \alpha_n) = 0$.

Область істинності формули $F$ позначається символом $I(F)$, а її область хибності — символом $O(F)$. Для довільної формули $F(x_1, \ldots, x_n)$ маємо

$$O(F) \cup I(F) = L^n, \quad O(F) \cap I(F) = \varnothing.$$

**Означення 1.6.** Формула $F(x_1, \ldots, x_n)$ є логічним наслідком формули $G(x_1, \ldots, x_n)$, якщо $I(G) \subset I(F)$. Формули $F$ і $G$ називаються рівносильними, якщо кожна з них є логічним наслідком іншої, тобто $I(F) = I(G)$.

Таким чином, формули $F$ і $G$ рівносильні тоді й лише тоді, коли вони мають одинакові таблиці логічних значень, тобто визначають одну й ту ж булеву функцію. Рівносильність формул $F$ і $G$ будемо позначати символом $F \equiv G$. Зауважимо, що всі тотожно істинні формули і, відповідно, всі тотожно хибні формули, будуть попарно рівносильними.

**Теорема 1.2.** *Рівносильність $F_1 \equiv F_2$ має місце тоді й лише тоді, коли формула $F_1 \leftrightarrow F_2$ є тавтологією.*

*Доведення.* Не порушуючи загальності, можна вважати, що формули $F_1$ і $F_2$ залежать від тих самих змінних $x_1$, $x_2$, …, $x_n$ (можливо, від деяких неявно).

Нехай має місце рівносильність

$$F_1(x_1, x_2, \ldots, x_n) \equiv F_2(x_1, x_2, \ldots, x_n).$$

Тоді для довільного булевого вектора $(\alpha_1, \ldots, \alpha_n)$ виконується рівність

$$F_1(\alpha_1, \ldots, \alpha_n) = F_2(\alpha_1, \ldots, \alpha_n).$$

Це означає, що

$$(F_1 \leftrightarrow F_2)(\alpha_1, \ldots, \alpha_n) = F_1(\alpha_1, \ldots, \alpha_n) \leftrightarrow F_2(\alpha_1, \ldots, \alpha_n) = 1.$$

Звідси отримуємо, що формула $F_1 \leftrightarrow F_2$ є тавтологією.

Припустимо тепер, що

$$F_1(x_1, x_2, \ldots, x_n) \not\equiv F_2(x_1, x_2, \ldots, x_n).$$

Тоді знайдеться вектор $(\alpha_1, \ldots, \alpha_n)$, для якого

$$F_1(\alpha_1, \ldots, \alpha_n) \neq F_2(\alpha_1, \ldots, \alpha_n).$$

Але тоді

$$(F_1 \leftrightarrow F_2)(\alpha_1, \ldots, \alpha_n) = F_1(\alpha_1, \ldots, \alpha_n) \leftrightarrow F_2(\alpha_1, \ldots, \alpha_n) = 0,$$

тобто формула $F_1 \leftrightarrow F_2$ не тавтологія. □

Кожну формулу $F$ логіки висловлювань можна подати у вигляді $F = F_1 \circ F_2$, $\circ \in \{\vee, \wedge, \rightarrow, \leftrightarrow\}$ або $F = \neg F_1$, де $F_1, F_2$ — деякі формули. Якщо формула $F$ має вигляд $F_1 \circ F_2$, то логічна операція $\circ$, яка застосовується при побудові формули останньою, називається головною зв'язкою формули $F$. З теореми 1.2 випливає, що проблема характеризації рівносильності формул логіки висловлювань еквівалентна задачі характеризації тих тотожно істинних формул, головною зв'язкою яких є еквівалентність.

**Теорема 1.3** (Другий принцип підстановки для тавтологій)**.** *Нехай $F$ — тавтологія, $A$ — деяка підформула $F$, і $A \equiv B$. Тоді формула $\Pi_A^B F$ також є тавтологією.*

*Доведення.* Нехай $x_1, x_2, \ldots, x_n$ — всі змінні, від яких залежать формули $F$, $A$, $B$. Значення формули $\Pi_A^B F(x_1, x_2, \ldots, x_n)$ на булевому векторі

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \in L^n$$

обчислюється послідовно, починаючи від змінних, і переходячи до підформул більшої довжини. На певному кроці доведеться обчислювати логічні значення підформул $B(x_1, x_2, \ldots, x_n)$. Оскільки $A \equiv B$, то всі ці значення дорівнюють $A(\alpha_1, \alpha_2, \ldots, \alpha_n)$. А тому всі наступні кроки обчислення логічного значення

$$\Pi_A^B F(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

будуть такими ж, як і при обчисленні значення $F(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Отже, для довільного вектора $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in L^n$ має місце рівність

$$\Pi_A^B F(\alpha_1, \alpha_2, \ldots, \alpha_n) = F(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1.$$

Це й означає, що формула $\Pi_A^B F$ є тавтологією. □

## 1.4 Логічний наслідок

Простий аналіз показує, що на практиці використовуються міркування, умова яких складається з кількох, по суті незалежних висловлювань. Ця ситуація легко формалізується за допомогою поняття логічного наслідку для формул логіки висловлювань.

Нехай $F_1, F_2, \ldots, F_s$ — довільним чином вибрані формули логіки висловлювань, які залежать від змінних $x_1$, $x_2$, …, $x_n$.

**Означення 1.7.** Формула $F$ називається логічним наслідком формул

$$F_1, F_2, \ldots, F_s,$$

якщо для довільного набору висловлювань $a_1$, $a_2$, …, $a_n$ з того, що всі висловлювання

$$F_1(a_1, a_2, \ldots, a_n), F_2(a_1, a_2, \ldots, a_n), \ldots, F_s(a_1, a_2, \ldots, a_n)$$

є істинними, випливає істинність висловлювання $F(a_1, a_2, \ldots, a_n)$.

Іншими словами, це означає, що кожен булевий вектор, який входить до областей істинності усіх формул $F_1$, $F_2$, …, $F_s$, входитиме також до області істинності формули $F$, тобто

$$\bigcap_{i=1}^{s} I(F_i) \subset I(F). \tag{1.2}$$

Той факт, що формула $F$ є логічним наслідком формул

$$F_1, F_2, \ldots, F_s,$$

позначатимемо так:

$$F_1, F_2, \ldots, F_s \models F.$$

**Теорема 1.4.** *Співвідношення $F_1, F_2, \ldots, F_s \models F$ для формул логіки висловлювань має місце тоді і тільки тоді, коли тавтологією буде кожна з наступних формул:*

$$(F_1 \wedge F_2 \wedge \ldots \wedge F_s) \to F, \tag{1.3}$$

$$F_1 \to (F_2 \to \ldots \to (F_s \to F)\ldots). \tag{1.4}$$

*Доведення.* Нехай $n$ — число змінних у формулах $F_1$, $F_2$, …, $F_s$, $F$.

1) Якщо формула (1.3) є тавтологією, то для довільного булевого вектора

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \bigcap_{i=1}^{s} I(F_i)$$

маємо

$$(F_1 \wedge F_2 \wedge \ldots \wedge F_s)(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1,$$

а тому $F(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1$. Отже, має місце включення (1.2).

З іншого боку, нехай виконується включення (1.2) і $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ — довільний булевий вектор. Якщо

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \notin \bigcap_{i=1}^{s} I(F_i),$$

то

$$(F_1 \wedge F_2 \wedge \ldots \wedge F_s)(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0,$$

а тому вся формула (1.3) на цьому векторі набуває значення 1. У тому випадку, коли

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \bigcap_{i=1}^{s} I(F_i),$$

отримуємо

$$(F_1 \wedge F_2 \wedge \ldots \wedge F_s)(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1, \quad \text{і} \quad F(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1,$$

тобто формула (1.3) також набуває значення 1. Отже, (1.3) — це тавтологія.

2) Нехай формула (1.4) є тавтологією. Припустимо від супротивного, що співвідношення $F_1, F_2, \ldots, F_s \models F$ не виконується. Тоді існує булевий вектор $\bar{a} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, для якого

$$F_1(\bar{a}) = F_2(\bar{a}) = \ldots = F_s(\bar{a}) = 1, \quad \text{але} \quad F(\bar{a}) = 0.$$

Звідси дістаємо

$$(F_s \to F)(\bar{a}) = 1 \to 0 = 0, \quad (F_{s-1} \to (F_s \to F))(\bar{a}) = 1 \to 0 = 0.$$

Продовжуючи рухатись ліворуч до початку формули (1.4) і враховуючи на кожному кроці результат попередніх обчислень, на останньому кроці дістанемо, що значення формули (1.4) на векторі $\bar{a}$ дорівнює

$$F_1(\bar{a}) \to 0 = 1 \to 0 = 0.$$

Отже, формула (1.4) не є тавтологією, що суперечить умові теореми.

З іншого боку, нехай має місце включення (1.2). Якщо булевий вектор $\bar{a} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ міститься в

$$\bigcap_{i=1}^{s} I(F_i),$$

то

$$F_1(\bar{a}) = F_2(\bar{a}) = \ldots = F_s(\bar{a}) = 1,$$

а тому й $F(\bar{a}) = 1$. Отже, в такому разі формула (1.4) на векторі $\bar{a}$ набуває значення

$$1 \to (1 \to \ldots \to (1 \to 1)\ldots) = 1.$$

Якщо ж

$$\bar{a} \notin \bigcap_{i=1}^{s} I(F_i),$$

то існує індекс $k$, $1 \le k \le s$, такий, що $\bar{a} \notin I(F_k)$. Виберемо найменше число $k$ з цією властивістю. Тоді

$$F_k(\bar{a}) \to (F_{k+1}(\bar{a}) \to \ldots \to (F_s(\bar{a}) \to F(\bar{a}))\ldots) = 0 \to (\ldots) = 1$$

незалежно від того, яке логічне значення набуває консеквент цієї формули. Звідси випливає, що на векторі $\bar{a}$ формула (1.4) набуває значення

$$F_1(\bar{a}) \to (F_2(\bar{a}) \to \ldots \to (F_{k-1}(\bar{a}) \to 1)\ldots).$$

З вибору числа $k$ випливає, що $F_1(\bar{a}) = F_2(\bar{a}) = \ldots = F_{k-1}(\bar{a}) = 1$. Тому останній вираз має логічне значення 1. Отже, (1.4) є тавтологією. □

## 1.5 Основні методи доведень

При доведенні теорем в тому чи іншому розділі математики окрім фактів теорії, для потреб якої формулюються і доводяться теореми, використовуться певні логічні міркування. Метод доведення теореми залежить від того, які саме логічні міркування застосовуються, та в якій послідовності вони проводяться. Загальні схеми, які при цьому виникають, є схемами логічного наслідку.

### 1.5.1 Доведення від супротивного

Доведення твердження $A$ цим методом здійснюється таким чином. Припускається, що $A$ є хибним твердженням, тобто істинним твердженням є $\neg A$. Після цього доводиться, що з припущення випливає деяке хибне твердження $B$, тобто таке, що $\neg B$ є істинним. На підставі цього робиться висновок про істинність твердження $A$. Це висновок правильний, оскільки він опирається на схему логічного міркування

$$\neg A \to B, \neg B \models A.$$

Проаналізуємо доведення твердження «Число $\sqrt{2}$ є ірраціональним». Позначимо це висловлювання символом $A$. Тоді $\neg A$ є висловлюванням «Число $\sqrt{2}$ є раціональним». Якщо $\neg A$ є істинним, то існують натуральні числа $m, n$ такі, що $\sqrt{2} = m/n$, звідки $2 = m^2/n^2$, тобто $2n^2 = m^2$. Остання рівність означає, що істинним є твердження «Існує натуральне число, яке при діленні на степені числа 2 різної парності дає в частці обидва рази непарне число». Позначимо це твердження символом $B$. Оскільки $B$ — хибне твкрдження, тобто $\neg B$ — істинне, то маємо умову $\neg A \to B, \neg B$. З цієї умови, використовуючи правило логічного наслідку modus tollens, дістаємо твердження $A$. Таким чином, наведене доведення твердження «Число $\sqrt{2}$ є ірраціональним» здійснюється за наведеною вище логічною схемою.

### 1.5.2 Метод зведення до абсурду

Це різновидність методу доведення від супротивного. Нехай потрібно довести деяке твердження $A$. Припускаючи, що воно хибне, встановлюється, що для деякого твердження $B$ істинними є імплікації $\neg A \to B$ і $\neg A \to \neg B$. Використовуючи схему логічного наслідку

$$\neg A \to B, \neg A \to \neg B \models A,$$

звідси отримуємо істинність твердження $A$.

### 1.5.3 Вичерпування можливостей для умови

Нехай теорема, яку потрібно довести, має вигляд $A \to B$, де $A$ — умова теореми, а $B$ — її висновок. Припустимо, що твердження $A$ можна подати, як диз'юнкцію певних тверджень:

$$A = A_1 \vee A_2 \vee \ldots \vee A_s.$$

Тоді для доведення теореми досить переконатися, що кожна з імплікацій

$$A_1 \to B, A_2 \to B, \ldots, A_s \to B$$

є істинним твердженням. Справді, легко пересвідчитися, що має місце таке відношення логічного наслідку

$$A_1 \vee A_2 \vee \ldots \vee A_s, A_1 \to B, A_2 \to B, \ldots, A_s \to B \models A_1 \vee A_2 \vee \ldots \vee A_s \to B.$$

Застосовуючи цю схему логічного наслідку в нашій ситуації, дістаємо потрібне.

З'ясуємо, за якою схемою логічних міркувань здійснюється доведення теореми «Кут, що вписаний у коло, вимірюється половиною дуги, на яку він спирається».

Умовою цієї теореми є висловлювання $A$ = «В коло вписано деякий кут $\alpha$», а її наслідком — висловлювання $B$ = «Кут $\alpha$ дорівнює половині дуги, на яку він спирається». Умова $A$ є диз'юнкцією трьох тверджень, які характеризують розташування кута в колі:

$A_1$ = «Центр кола лежить на одній зі сторін кута»,

$A_2$ = «Центр кола лежить всередині кута»,

$A_3$ = «Центр кола лежить поза кутом».

Зрозуміло, що $A = A_1 \vee A_2 \vee A_3$. Тому для доведення теореми досить довести істинність імплікацій $A_1 \to B$, $A_2 \to B$, $A_3 \to B$. Отже, твердження теореми отримуємо на основі схеми логічного наслідку

$$A_1 \vee A_2 \vee A_3, A_1 \to B, A_2 \to B, A_3 \to B \models A_1 \vee A_2 \vee A_3 \to B.$$

### 1.5.4 Вичерпування можливостей для висновку

Нехай потрібно довести деяке твердження $A \to B$, а $B_1, \ldots, B_s$ — інші можливості для висновку $B$, причому одна з них обов'язково правильна, тобто $B \vee B_1 \vee \ldots \vee B_s = 1$. Тоді досить довести хибність імплікацій $A \to B_1, \ldots, A \to B_s$. Звідси отримуємо істинність імплікації $A \to B$. Це випливає з такого співвідношення логічного наслідку:

$$B \vee B_1 \vee \ldots \vee B_s, \neg(A \to B_1), \ldots, \neg(A \to B_s) \models A \to B.$$

### 1.5.5 Метод математичної індукції

Нехай потрібно довести твердження про те, що кожен елемент деякої множини $\Omega$ має певну властивість $P$, причому множина $\Omega$ є зліченною, тобто кожен її елемент має унікальний натуральний номер:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \ldots\}$$

Доводимо такі твердження:

(a) елемент $\omega_1$ має властивість $P$, тобто $P(\omega_1) = 1$;

(b) для довільного натурального $n$ із припущення, що $\omega_n$ має властивість $P$, випливає, що цю властивість має і елемент $\omega_{n+1}$, тобто $P(\omega_n) \to P(\omega_{n+1}) = 1$.

Тоді кожен елемент множини $\Omega$ має властивість $P$.

Цей метод доведення називається методом математичної індукції. Він базується на тому, що для кожного $n \geq 1$ має місце співвідношення

$$P(\omega_1), P(\omega_1) \to P(\omega_2), \ldots, P(\omega_{n-1}) \to P(\omega_n) \models P(\omega_n).$$

Доведення частина (a) називають базою індукції, а доведення частини (b) — індуктивним кроком.

### 1.5.6 Пряма і обернена теореми

Якщо теорема записана у вигляді імплікації $A \to B$, то поряд із нею можна розглядати такі імплікації:

$$B \to A, \quad \neg A \to \neg B, \quad \neg B \to \neg A.$$

Перша з них називається оберненою теоремою, друга — протилежною, а третя — протилежною до оберненої теореми. Саму ж імплікацію $A \to B$ при цьому звуть прямою теоремою. Оскільки мають місце рівносильності

$$A \to B \equiv \neg B \to \neg A, \quad B \to A \equiv \neg A \to \neg B,$$

то пряма теорема рівносильна з протилежною до оберненої, а обернена теорема рівносильна з протилежною. Пряма і обернена теорема, взагалі кажучи, не рівносильні.

Таким чином, замість того, щоб доводити пряму теорему, можна доводити протилежну до оберненої, а замість того, щоб будувати доведення оберненої, можна доводити протилежну теорему.

Рівносильність прямого і оберненого тверджень теореми означає, що вона має вигляд $A \leftrightarrow B$, тобто формулюється у вигляді критерія. Доведення такої теореми відбувається на основі схеми логічного наслідку

$$A \to B, B \to A \models A \leftrightarrow B.$$

## 1.6 Вправи

1. Побудуйте таблицю істинності формули і перевірте, чи буде вона тавтологією:

   а) $x_1 \wedge (\neg x_2 \leftrightarrow x_3)$;

   б) $(\neg x_2 \vee x_3) \to x_1$.

2. На основі поняття логічного наслідку встановіть, чи є логічними такі міркування:

   *«Студент С не складе екзамен, якщо погано до нього підготується. Якщо ж він не складе екзамен, то не буде отримувати стипендію. Отже, якщо студент С погано підготується до екзамену, то він не буде отримувати стипендію.»*

3. Для заданої теореми сформулюйте обернене твердження. Чи буде воно правильним? Чому?

   а) Діагоналі паралелограма в точці перетину діляться навпіл.

   б) В прямокутному трикутнику квадрат гіпотенузи дорівнює сумі квадратів катетів.

4. Для кожного з наведених тверджень сформулюйте протилежне, обернене і протилежне до оберненого твердження. Які з них правильні і чому?

   а) Навколо прямокутника можна описати коло.

   б) Якщо в чотирикутник можна вписати коло, то він є ромбом.

5. Доведіть методом математичної індукції, що має місце рівність

   а) $1^2 + 2^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$, $n \geq 1$;

   б) $\dfrac{1}{1 \cdot 2} + \dfrac{1}{2 \cdot 3} + \cdots + \dfrac{1}{(n-1) \cdot n} = \dfrac{n-1}{n}$, $n \geq 2$;

   в) $\left(1 - \dfrac{1}{2}\right)\left(1 - \dfrac{1}{3}\right) \cdots \left(1 - \dfrac{1}{n}\right) = \dfrac{1}{n}$, $n \geq 2$;

   г) $\dfrac{1}{1 + \sqrt{2}} + \dfrac{1}{\sqrt{2} + \sqrt{3}} + \cdots + \dfrac{1}{\sqrt{n-1} + \sqrt{n}} = \sqrt{n} - 1$, $n \geq 2$.

# Розділ 2

# Основи теорії чисел

## 2.1 Подільність цілих чисел. Прості числа

При обчисленнях з цілими числами ми часто стикаємося із ситуацією, коли після ділення залишається певний «надлишок». Наступна теорема, яку називають теоремою про ділення з остачею, формалізує процедуру ділення, враховуючи таку ситуацію, та визначає однозначний її результат.

**Теорема 2.1.** *Для довільної пари цілих чисел $a$ і $b$, $b \neq 0$, існують такі цілі числа $q$ та $r$, що*

$$a = qb + r \quad і \quad 0 \leq r < |b|. \tag{2.1}$$

*Числа $q$ і $r$ цими умовами визначаються однозначно.*

Числа $q$ і $r$ з рівності (2.1) називаються відповідно (неповною) *часткою* та *остачею* від ділення $a$ на $b$.

*Доведення.* Доведемо існування чисел $q$ і $r$. Маємо наступні випадки.

Випадок $a \geq 0$ і $b > 0$. Розглянемо множину

$$A = \{a - nb \mid n \in \mathbb{Z}\},$$

тобто множину, що містить елементи

$$\ldots, a + 3b, a + 2b, a + b, a, a - b, a - 2b, a - 3b, \ldots$$

Серед елементів множини $A$ є невід'ємні числа. Наприклад, таким буде число $a$. Тому серед невід'ємних чисел із множини $A$ є найменше. Позначимо його $r$. Нехай $r = a - qb$ — це задання $r$ як елемента множини $A$. Тоді $r < b$. Справді, з припущення $r \geq b$ відразу випливало б, що множина $A$ містить невід'ємне число $r - b = a - (q+1)b$, яке менше за $r$. Отже, $0 \leq r < b$ і $a = qb + r$, тобто числа $q$ і $r$ — шукані.

Випадок $a \geq 0$ і $b < 0$. За доведеним у попередньому випадку, існують числа $m$ і $r$ такі, що

$$a = m \cdot |b| + r = (-m) \cdot b + r \quad \text{і} \quad 0 \leq r < |b|.$$

Тобто числа $q = -m$ і $r$ — шукані.

Випадок $a < 0$. За доведеним у попередніх двох випадках, існують числа $m$ і $k$ такі, що

$$|a| = m \cdot b + k \quad \text{і} \quad 0 \leq k < |b|.$$

Якщо $k = 0$, то $q = -m$ та $r = k = 0$. Якщо ж $k > 0$, то

$$a = -|a| = (-m) \cdot b + (-k) = (-m - \operatorname{sign} b) \cdot b + (|b| - k),$$

де $\operatorname{sign} b = 1$ при $b > 0$ та $\operatorname{sign} b = -1$ при $b < 0$, і умови теореми задовольняють $q = -m - \operatorname{sign} b$ та $r = |b| - k$, оскільки $0 < |b| - k < |b|$.

Існування чисел $q$ і $r$ доведено.

Припустимо, що окрім пари чисел $q$ та $r$, умова (2.1) виконується для пари $q_1$ та $r_1$, тобто

$$a = q_1 b + r_1 \quad \text{і} \quad 0 \leq r_1 < |b|.$$

Тоді з рівності $q_1 b + r_1 = qb + r$ випливає $|q_1 - q| \cdot |b| = |r - r_1| < |b|$, оскільки числа $r$ і $r_1$ належать напівінтервалу $[0, |b|)$. Тому $|q_1 - q| < 1$. Єдине ціле число, яке задовольняє останню нерівність, це $|q_1 - q| = 0$. Звідки $q_1 = q$ та $r_1 = (q - q_1)b + r = r$, що й треба було довести. □

**Приклад 2.1.** При діленні 17 на 7 одержуємо

$$17 = 2 \cdot 7 + 3,$$

тобто частка 2 і остача 3, а при діленні $-17$ на 7 одержуємо

$$-17 = (-3) \cdot 7 + 4,$$

тобто частка 3 і остача 4.

**Вправа 2.1.** Знайдіть критерій, коли остача від ділення $a$ на $b$ рівна $a$.

**Вправа 2.2.** Вкажіть правило, як обчислити знак частки від ділення $a$ на $b$.

**Означення 2.1.** Кажуть, що ціле число $a$ *ділиться* на ціле число $b$, якщо існує ціле число $c$ таке, що $a = b \cdot c$. Число $a$ називають *кратним* числа $b$, а число $b$ — *дільником* числа $a$.

Для позначення того, що ціле число $a$ ділиться на ціле число $b$ використовують запис $b \mid a$ (читається «$b$ є дільником $a$») або $a \vdots b$ (читається «$a$ ділиться на $b$»). Якщо ж потрібно вказати, що ціле число $a$ *не* ділиться на ціле число $b$, то використовують перекреслені символи: $b \nmid a$ (читається «$b$ не є дільником $a$») або $a \not\vdots b$ (читається «$a$ не ділиться на $b$»).

Безпосередньо з означення випливає, що ціле число $a$ ділиться на ціле число $b \neq 0$ тоді й лише тоді, коли остача від ділення $a$ на $b$ рівна нулю. Також, відповідно до означення, можна говорити про подільність на 0 і число 0 є єдиним числом, що ділиться на 0.

Зв'язок «бути дільником» чи «ділитися» між парою чисел природним чином визначає бінарне відношення на множині цілих чисел. Основні властивості цього відношення вказує

**Теорема 2.2.** *Відношення подільності має наступні властивості:*

*а)* $a \mid a$ *(рефлексивність);*

*б) якщо* $a \mid b$ *і* $b \mid c$*, то* $a \mid c$ *(транзитивність);*

*в) якщо* $a \mid b$*, то* $a \mid (b \cdot c)$*;*

*г) якщо* $a \mid b$ *і* $a \mid c$*, то* $a \mid (b \pm c)$*;*

*д)* 0 *ділиться на кожне ціле число, і лише* 0 *має таку властивість;*

*е) кожне число ділиться на* $\pm 1$*, і лише* $\pm 1$ *мають таку властивість;*

*ж) якщо* $a \mid b$ *і* $c \mid d$*, то* $(ac) \mid (bd)$.

Зазначені властивості відношення подільності безпосередньо випливають з означення 2.1. Проілюструємо це на прикладі транзитивності. З умов $a \mid b$ і $b \mid c$ випливає існування таких цілих чисел $a_1$ і $b_1$, що $b = a \cdot a_1$ і $c = b \cdot b_1$. Звідси $c = b \cdot b_1 = (a \cdot a_1) \cdot b_1 = a \cdot (a_1 \cdot b_1)$, тобто $a \mid c$.

**Вправа 2.3.** Доведіть всі інші властивості.

Нехай $a \neq 0$ — ціле число та $b$ його дільник. За означенням існує ціле число $c$, причому $c \neq 0$ таке, що $a = b \cdot c$. Оскільки $|c| \geq 1$, то виконується нерівність $|a| = |b| \cdot |c| \geq |b|$, яку можна переписати у вигляді $-|a| \leq b \leq |a|$. Іншими словами, всі дільники числа $a$ знаходяться між $\pm a$. З рівностей $a = a \cdot 1 = (-a) \cdot (-1)$ випливає, що числа $\pm 1$ і $\pm a$ завжди є дільниками числа $a$. Ці дільники називають *невласними*. Будь-який інший дільник $b$ числа $a$, тобто такий, що задовольняє умову $1 < |b| < |a|$, називається *власним*.

**Означення 2.2.** Ціле число називається *складеним*, якщо воно має власні дільники. Ціле число називається *простим*, якщо воно відмінне від $0, \pm 1$ і не має власних дільників.

Таким чином, кожне ціле число, окрім 0 і $\pm 1$, або є простим, або є складеним. При цьому числа 0 і $\pm 1$ не відносять ні до простих, ні до складених. Легко бачити, що ціле (натуральне) число $a$ є простим тоді й тільки тоді, коли воно має рівно 4 цілі дільники: $\pm 1$ і $\pm a$ (рівно 2 натуральні дільники: 1 і $a$).

Безпосередньо з означення простого числа випливає

**Твердження 2.1.** *Ціле число $a$, відмінне від $0, \pm 1$, є простим тоді й лише тоді, коли для будь-якого розкладу $a$ у добуток двох множників $a = b \cdot c$ один із них дорівнює $\pm 1$.*

Доведення простоти цілого числа за означенням вимагає перевірки великої кількості чисел, що можуть виявитись власним дільником. Цю кількість можна значно скоротити, якщо шукати мінімальний власний дільник.

**Твердження 2.2.** *Натуральне число $a > 1$ є простим тоді й лише тоді, коли $a$ не ділиться на жодне натуральне число $b$ таке, що $1 < b \leq \sqrt{a}$.*

*Доведення.* Необхідність є очевидною, якщо врахувати, що, за умови $a > 1$, виконується нерівність $\sqrt{a} < a$.

Доведемо достатність. Нехай $a = b \cdot c$ деякий розклад натурального числа $a$ у добуток двох множників. Не обмежуючи загальності, можна вважати, що $|b| \leq |c|$. Тоді $b^2 \leq |b| \cdot |c| = |b \cdot c| = a$. Звідки одержуємо, що дільник $|b|$ числа $a$ задовольняє нерівність $1 \leq |b| \leq \sqrt{a}$. Оскільки $a$ не ділиться на натуральне число $|b|$, що задовольняє умові $1 < |b| \leq \sqrt{a}$, то залишається лише випадок $|b| = 1$. Отже, за твердженням 2.1 число $a$ є простим. □

Прості числа виконують роль «будівельних блоків» для натуральних чисел, оскільки кожне таке число можна подати як добуток простих. Пізніше ми переконаємось, що за додаткових умов, такий розклад єдиний. Поки ми обмежимось доведенням існування такого розкладу. Для цього нам буде потрібна

**Лема 2.1.** *Кожне відмінне від* $\pm 1$ *ціле число ділиться на деяке просте число.*

*Доведення.* Число 0 ділиться на просте число 2. Для ненульових чисел доводимо індукцією наступне твердження, з якого випливає твердження леми. Кожне ціле число $k$, яке задовольняє нерівність $2 \leq |k| \leq n$, ділиться на деяке просте число.

База індукції $n = 2$. Числа $\pm 2$ діляться на просте число 2.

Крок індукції. Припустимо, що твердження правильне для $n - 1$. Щоб довести для $n$ потрібно розглянути лише «нові» числа $\pm n$. Якщо число $n$ просте, то простим дільником чисел $\pm n$ буде число $n$. Якщо ж $n$ складене, то воно має власний натуральний дільник $m$. Оскільки $1 < m < |n|$, то, за припущенням індукції, $m$ має простий дільник $p$. З транзитивності відношення подільності випливає $p \mid (\pm n)$. □

В подальшому, коли ми говоримо про добуток $m_1 \cdot m_2 \cdot \ldots \cdot m_k$ деяких $k$ множників, будемо використовувати наступну домовленість. Єдине число у випадку $k = 1$ також будемо називати добутком, що містить один множник. Крім того, число 1 будемо вважати добутком нульової кількості множників (випадок $k = 0$).

**Теорема 2.3.** *Кожне натуральне число розкладається у добуток простих чисел.*

*Доведення.* Доводимо індукцією за $n \in \mathbb{N}$.

База індукції. Для числа $n = 1$ твердження теореми є правильним згідно з домовленістю.

Крок індукції. Розглянемо довільне $n > 1$ і припустимо, що для всіх чисел, менших від $n$, теорему вже доведено. За попередньою лемою $n$ можна розкласти у добуток $n = p_1 m$, де $p_1$ — просте і $1 \leq m < n$. За припущенням індукції для $m$ існує розклад $m = p_2 p_3 \ldots p_k$ у добуток простих чисел. Тоді $n = p_1 p_2 \ldots p_k$, що й буде шуканим розкладом для числа $n$. □

Класичним результатом, що встановлює нескінченність множини простих чисел є

**Теорема 2.4** (Евклід)**.** *Множина всіх простих чисел є нескінченною.*

*Доведення.* Доводимо методом від супротивного. Припустимо, що множина $\mathbb{P}$ усіх простих чисел скінченна. Нехай

$$\mathbb{P} = \{p_1, p_2, \ldots, p_k\}.$$

Розглянемо натуральне число $n = |p_1 p_2 \ldots p_k| + 1$. За доведеною лемою число $n$ має простий дільник $p$. Цей дільник не може збігатися з жодним із чисел $p_1, \ldots, p_k$, бо при діленні $n$ з остачею на кожне з них отримуємо остачу 1. Отже, просте число $p$ не зустрічається в списку $p_1, \ldots, p_k$. Це суперечить припущенню про те, що $\mathbb{P}$ містить усі прості числа. Теорему доведено. □

## 2.2 НСД і НСК

**Означення 2.3.** Число $d$ називають *найбільшим спільним дільником* чисел $a$ і $b$, якщо виконуються умови:

1) $d \mid a$ і $d \mid b$;

2) якщо $c \mid a$ і $c \mid b$, то $c \mid d$.

Слово «найбільший» означає не найбільший за величиною, а те, що він є кратним будь-якого іншого спільного дільника чисел $a$ і $b$. Наприклад, спільними дільниками чисел 12 і 18 є числа $\pm 1$, $\pm 2$, $\pm 3$ і $\pm 6$. Лише числа $\pm 6$ задовольняють другу умову означення, тому найбільшим спільним дільником чисел 12 і 18 є число 6 або $-6$.

Означення 2.3 накладає певні вимоги на число $d$, щоб можна було його назвати найбільшим спільним дільником заданих чисел. Але існування такого числа $d$ потребує доведення. Це встановлює

**Теорема 2.5.** *Для будь-яких цілих чисел $a$ і $b$ існує їх найбільший спільний дільник.*

*Доведення.* Безпосередньо з означення маємо, що найбільший спільний дільник чисел 0 і 0 є число 0. Тому далі вважаємо, що серед чисел $a$ і $b$ є ненульове.

Розглянемо множину

$$I = \{ma + nb \mid m, n \in \mathbb{Z}\}.$$

Вона, зокрема, містить числа

$$\pm a = \pm 1 \cdot a + 0 \cdot b \quad \text{та} \quad \pm b = 0 \cdot a + (\pm 1) \cdot b.$$

Тому серед елементів множини $I$ є натуральні числа.

Позначимо через $d$ найменше з них і доведемо, що $d$ є найбільшим спільним дільником чисел $a$ і $b$. Справді, нехай $d = m_0 a + n_0 b$. Припустимо, що $a$ не ділиться на $d$. Тоді розділимо $a$ на $d$ з остачею:

$$a = qd + r, \quad 0 < r < d,$$

і матимемо

$$r = a - qd = (1 - qm_0)a + (-qn_0)b \in I,$$

що суперечить вибору числа $d$. Аналогічно доводиться, що $d \mid b$. Отже, першу умову з означення найбільшого спільного дільника число $d$ задовольняє.

Припустимо тепер, що $c \mid a$ і $c \mid b$. Тоді існують розклади $a = ca_1$ і $b = cb_1$ та

$$d = m_0 \cdot ca_1 + n_0 \cdot cb_1 = c \cdot (m_0 a_1 + n_0 b_1).$$

Звідки одержуємо $c \mid d$, тобто виконується і друга умова означення. □

Як показав приклад вище, може існувати кілька чисел, кожне з яких є найбільшим спільним дільником. Наступне твердження встановлює зв'язок між всіма такими числами.

**Твердження 2.3.** *Найбільший спільний дільник визначений однозначно з точністю до знака.*

*Доведення.* З означення 2.3 випливає, що разом із числом $d$ найбільшим спільним дільником чисел $a$ і $b$ буде також і число $-d$.

Нехай тепер $d_1$ і $d_2$ — найбільші спільні дільники чисел $a$ і $b$. З означення маємо $d_1 \mid d_2$ і $d_2 \mid d_1$, тобто $d_2 = d_1 \cdot q_1$ і $d_1 = d_2 \cdot q_2$. Якщо $d_2 = 0$, то $d_1 = 0$. У випадку $d_2 \neq 0$ одержуємо $d_2 = d_2 \cdot q_1 q_2$. Звідки $q_1 q_2 = 1$, що можливо лише при $q_1 = \pm 1$. Отже, $d_1 = \pm d_2$. □

Доведене твердження дозволяє досягнути однозначності, якщо домовитись розглядати лише невід'ємне значення найбільшого спільного дільника. Такий найбільший спільний дільник чисел $a$ і $b$ позначають $\mathsf{НСД}(a, b)$ або просто $(a, b)$.

Безпосередньо з доведення теореми 2.5 випливає

**Твердження 2.4** (Тотожність Безу)**.** *Якщо* $d = \mathsf{НСД}(a, b)$*, то існують цілі числа* $m$ *і* $n$ *такі, що* $ma + nb = d$*.*

**Твердження 2.5.** *Якщо* $a = qb + r$*,* $b \neq 0$*, то* $\mathsf{НСД}(a, b) = \mathsf{НСД}(b, r)$*.*

*Доведення.* Нехай $\mathsf{НСД}(a, b) = d_1$, $\mathsf{НСД}(b, r) = d_2$. Оскільки $b \neq 0$, то числа $d_1$ і $d_2$ не можуть бути нульовими, тобто є додатними. З рівності $a = qb + r$ випливає, що $d_2 \mid a$, бо $d_2$ ділить праву частину. Крім того, $d_2 \mid b$. Тому за означенням найбільшого спільного дільника $d_2 \mid d_1$. Аналогічно з рівності $r = a - qb$ одержуємо $d_1 \mid r$ та $d_1 \mid d_2$. Отже, числа $d_1$ і $d_2$ діляться одне на одне, тобто існують цілі числа $q_1$ і $q_2$ такі, що $d_2 = d_1 \cdot q_1$ і $d_1 = d_2 \cdot q_2$. Звідси $d_2 = d_2 \cdot q_1 q_2$, що еквівалентно $q_1 q_2 = 1$. Остання рівність можлива лише при $q_1 = \pm 1$, тобто $d_1 = \pm d_2$. Оскільки $d_1$ і $d_2$ додатні, то $d_1 = d_2$. □

**Вхід:** $a$, $b$ — цілі числа, $a > b \geq 0$
**Вихід:** $\mathsf{НСД}(a, b)$
**while** $b > 0$ **do**
  обчислюємо остачу $r$ від ділення $a$ на $b$;
  $a = b$;
  $b = r$;
**end**
**return** $a$;

**Алгоритм 1.** Алгоритм Евкліда

**Теорема 2.6** (про коректність і обчислювальну складність)**.** *За допомогою алгоритму Евкліда обчислюється найбільший спільний дільник цілих чисел* $a$ *і* $b$*,* $a > b \geq 0$*. При цьому вимагається менше, ніж* $2 \log_2 a$ *ділень з остачею.*

*Доведення.* Коректність алгоритму Евкліда. У випадку $b = 0$ маємо $\mathsf{НСД}(a, b) = a$. Алгоритм Евкліда також повертає це значення. Далі розглядаємо випадок $b \neq 0$.

Позначимо $r_{-1} = a$, $r_0 = b$ та розглянемо послідовність чисел $\{r_i\}$, які обчислюються алгоритмом послідовним діленням з остачею

$$r_{i-1} = q_{i+1} r_i + r_{i+1}, \quad 0 < r_{i+1} < r_i, \qquad i \geq 0. \tag{2.2}$$

Ці числа утворюють монотонно спадну послідовність

$$r_{-1} > r_0 > r_1 > r_2 > r_3 > \dots,$$

натуральних чисел, яка не може бути нескінченною. Тому на якомусь кроці відбудеться ділення націло $r_{k-1} = q_{k+1}r_k$, та алгоритм закінчить обчислення і поверне остачу $r_k$, яка є найбільшим спільним дільником чисел $a$ і $b$.

Справді, з твердження 2.5 і правила побудови послідовності (2.2) маємо

$$\mathsf{НСД}(a,b) = \mathsf{НСД}(b,r_1) = \mathsf{НСД}(r_1,r_2) = \ldots = \mathsf{НСД}(r_{k-1},r_k) = r_k.$$

Обчислювальна складність алгоритму Евкліда. Покажемо, що в послідовності остач

$$r_{-1}, r_0, r_1, r_2, \ldots, r_{k-1}, r_k$$

за будь-які два кроки остача зменшується принаймні вдвічі, тобто, що для всіх $i$, $1 \le i \le k$, виконується нерівність $r_i < r_{i-2}/2$. Справді, якщо

$$r_{i-1} \le r_{i-2}/2,$$

то

$$r_i < r_{i-1} \le r_{i-2}/2.$$

Якщо ж

$$r_{i-1} > r_{i-2}/2,$$

то

$$0 < r_{i-2} - r_{i-1} < r_{i-2} - r_{i-2}/2 = r_{i-2}/2 < r_{i-1}.$$

Але з рівності $r_{i-2} = r_{i-1} + (r_{i-2} - r_{i-1})$ тепер випливає, що при діленні $r_{i-2}$ на $r_{i-1}$ остача дорівнює

$$r_i = r_{i-2} - r_{i-1} < r_{i-2}/2.$$

Нехай тепер $2m-1$ — найбільше непарне число, яке не перевищує $k$. Тоді $k \le 2m$. Крім того, з нерівностей $r_i < r_{i-2}/2$ випливає, що

$$r_{2m-1} < a/2^m.$$

Оскільки $r_{2m-1} \ge 1$, то $2^m < a$, звідки $m < \log_2 a$. Отже,

$$k \le 2m < 2\log_2 a,$$

що й треба було довести. □

**Твердження 2.6.** *Алгоритм Евкліда можна розширити для знаходження лінійного зображення $d = ma+nb$ найбільшого спільного дільника $d$ чисел $a$ і $b$.*

*Доведення.* Справді, з рівності

$$r_{k-2} = q_k r_{k-1} + r_k$$

можна одержати зображення

$$d = r_k = r_{k-2} - q_k r_{k-1}$$

найбільшого спільного дільника як лінійної комбінації остач $r_{k-1}$ та $r_{k-2}$. Використовуючи попередню рівність

$$r_{k-3} = q_{k-1} r_{k-2} + r_{k-1},$$

одержимо

$$d = r_{k-2} - q_k(r_{k-3} - q_{k-1} r_{k-2}) = (1 + q_k q_{k+1}) r_{k-2} - q_k r_{k-3}$$

як лінійної комбінації остач $r_{k-2}$ та $r_{k-3}$. Піднімаючись ланцюжком рівностей з алгоритму Евкліда вгору, на наступному кроці одержимо зображення $d$ як лінійної комбінації остач $r_{k-3}$ і $r_{k-4}$, потім — як лінійної комбінації остач $r_{k-4}$ та $r_{k-5}$, і т.д., доки за допомогою першої рівності $a = q_1 b + r_1$ не одержимо зображення $d$ як лінійної комбінації $a$ і $b$. □

Альтернативний варіант розширеного алгоритму Евкліда, який не вимагає збереження всієї послідовності часток та остач в процесі своєї роботи, представлено нижче. Цей алгоритм в порівнянні зі звичайним алгоритмом Евкліда виконує додаткові обчислення, які не впливають на кількість ітерацій. Його коректність випливає з коректності звичайного алгоритму Евкліда та того факту, що рівності $x = m_x a + n_x b$ та $y = m_y a + n_y b$ виконуються до початку виконання циклу та після кожної ітерації циклу.

**Вхід:** $a$, $b$ — цілі числа, $a > b \geq 0$
**Вихід:** $d$, $m$, $n$ — цілі числа; $d = \mathsf{НСД}(a, b) = ma + nb$
$x = a$; $m_x = 1$; $n_x = 0$; $\quad \{\, x = m_x a + n_x b \,\}$
$y = b$; $m_y = 0$; $n_y = 1$; $\quad \{\, y = m_y a + n_y b \,\}$
**while** $y > 0$ **do**
  обчислюємо частку $q$ та остачу $r$ від ділення $x$ на $y$;
  $\{\, r = x - qy = (m_x a + n_x b) - q(m_y a + n_y b)$
  $\quad = (m_x - qm_y)a + (n_x - qn_y)b \,\}$
  $m_r = m_x - qm_y$; $n_r = n_x - qn_y$; $\quad \{\, r = m_r a + n_r b \,\}$
  $x = y$; $m_x = m_y$; $n_x = n_y$;
  $y = r$; $m_y = m_r$; $n_y = n_r$;
**end**
**return** $(x, m_x, n_x)$;

**Алгоритм 2.** Розширений алгоритм Евкліда

**Вправа 2.4.** Доведіть наступні твердження про розширений алгоритм Евкліда.

- На кожній ітерації циклу виконуються рівності $|m_r| = |m_x| + q|m_y|$ та $|n_r| = |n_x| + q|n_y|$. Як наслідок $|m_r| \geq |m_y|$ та $|n_r| \geq |n_y|$.
- На кожній ітерації циклу виконується рівність $|m_r n_y - n_r m_y| = |m_y n_x - n_y m_x|$. Як наслідок $|m_r n_y - n_r m_y| = 1$.
- На останній ітерації циклу $|m_r| = b/d$ та $|n_r| = a/d$.
- Виконуються нерівності $|m| \leq b/d$ та $|n| \leq a/d$.
- Якщо $b > 0$, то виконуються більш сильні обмеження

$$|m| \leq \frac{b}{2d} \quad \text{та} \quad |n| \leq \frac{a}{2d}.$$

Поняття найбільшого спільного дільника можна узагальнити на довільну скінченну кількість чисел наступним чином.

**Означення 2.4.** Число $d$ називається *найбільшим спільним дільником* чисел

$$a_1, a_2, \ldots, a_k, \quad k \geq 1,$$

якщо $d$ задовольняє такі умови:

1) $d \mid a_1$, $d \mid a_2$, …, $d \mid a_k$;

2) якщо $c \mid a_1$, $c \mid a_2$, …, $c \mid a_k$, то $c \mid d$.

Нескладно переконатись аналогічно до випадку двох чисел, що найбільший спільний дільник чисел $a_1$, $a_2$, …, $a_k$ визначений з точністю до знака. Це дозволяє досягти однозначності, обмежившись лише невід'ємним значенням, яке позначають $\mathsf{НСД}(a_1, a_2, \ldots, a_k)$.

Для обчислення найбільшого спільного дільника довільної скінченної кількості чисел можна використовувати алгоритм Евкліда для двох чисел, що є наслідком

**Твердження 2.7.** $\mathsf{НСД}(a_1, a_2, \ldots, a_k) = \mathsf{НСД}(\mathsf{НСД}(a_1, a_2), a_3, \ldots, a_k)$.

*Доведення.* Позначимо $d_{12} = \mathsf{НСД}(a_1, a_2)$ та $d = \mathsf{НСД}(d_{12}, a_3, \ldots, a_k)$. За означенням $d \mid d_{12}$ та $d \mid a_3$, …, $d \mid a_k$. Крім того, $d_{12} \mid a_1$ і $d_{12} \mid a_2$, звідки за транзитивністю відношення подільності $d \mid a_1$ і $d \mid a_2$. Отже, перша умова означення виконується.

Нехай $c$ таке, що $c \mid a_1$, $c \mid a_2$, …, $c \mid a_k$. Тоді $c \mid d_{12}$ за означенням найбільшого спільного дільника для двох чисел, а за означенням для $k-1$ числа $d_{12}$, $a_3$, …, $a_k$ маємо $c \mid d$. Отже, виконується і друга умова означення та

$$\mathsf{НСД}(a_1, a_2, \ldots, a_k) = d.$$

Твердження доведено. □

Нерідко при обчисленнях з цілими числами чи проведенні певних міркувань зручно мати справу з числами, що не мають спільних простих дільників, або, що те саме, не мають спільних дільників, відмінних від $\pm 1$. Такі числа отримали окрему назву, на що вказує

**Означення 2.5.** Цілі числа $a$ і $b$ називаються *взаємно простими*, якщо їх найбільший спільний дільник дорівнює 1.

У випадку більшої кількості чисел говорять про взаємно простоту (у сукупності), коли найбільший спільний дільник всіх чисел дорівнює 1, та про попарну взаємну простоту, коли всі пари чисел є взаємно простими.

**Теорема 2.7** (Критерій взаємної простоти)**.** *Цілі числа $a$ і $b$ будуть взаємно простими тоді й лише тоді, коли існують цілі числа $c$ і $d$ такі, що $ac + bd = 1$.*

*Доведення.* Необхідність прямо випливає з доведення теореми 2.5 (або з тотожності Безу).

Достатність маємо з того, що довільний спільний дільник чисел $a$ і $b$ буде також дільником числа $ac + bd = 1$, тобто має дорівнювати $\pm 1$. □

З критерію взаємної простоти випливає ряд властивостей взаємно простих чисел, які сформульовані в наступних твердженнях.

**Твердження 2.8.** *Якщо* $\mathsf{НСД}(a, b) = 1$ *і* $\mathsf{НСД}(a, c) = 1$, *то* $\mathsf{НСД}(a, bc) = 1$.

*Доведення.* За критерієм взаємної простоти чисел умови $\mathsf{НСД}(a, b) = 1$ і $\mathsf{НСД}(a, c) = 1$ рівносильні існуванню таких цілих чисел $m$, $n$, $k$, $l$, що

$$ma + nb = 1, \quad ka + lc = 1.$$

Перемноживши ці рівності, одержимо:

$$(mka + mlc + nkb) \cdot a + nl \cdot bc = 1.$$

За тим же критерієм звідси випливає, що $\mathsf{НСД}(a, bc) = 1$. □

**Твердження 2.9.** *Якщо* $a \mid bc$ *і* $\mathsf{НСД}(a, b) = 1$, *то* $a \mid c$.

*Доведення.* Помножимо обидві частини рівності $ma + nb = 1$, яка виконується за критерієм взаємної простоти, на $c$:

$$mac + nbc = c.$$

Оскільки $a \mid mac$ і $a \mid nbc$, то $a \mid c$. □

**Твердження 2.10.** *Якщо* $a \mid c$, $b \mid c$ *і* $\mathsf{НСД}(a, b) = 1$, *то* $ab \mid c$.

*Доведення.* За умовою існують такі $a_0$, $b_0$, $m$ і $n$, що $c = aa_0 = bb_0$ і $ma + nb = 1$. Тоді

$$c = mac + nbc = ma \cdot bb_0 + nb \cdot aa_0 = (mb_0 + na_0)ab.$$

Отже, $ab \mid c$. □

**Твердження 2.11.** *Якщо* $\mathsf{НСД}(a, b) = d$ *і* $a = a_0 d$, $b = b_0 d$, *то числа* $a_0$ *і* $b_0$ *взаємно прості.*

*Доведення.* Зобразимо число $d$ у вигляді $d = ma + nb = ma_0 d + nb_0 d$. Тоді після скорочення на $d$ матимемо: $1 = ma_0 + nb_0$. Отже, за критерієм взаємної простоти чисел, $\mathsf{НСД}(a_0, b_0) = 1$. □

**Вправа 2.5.** Доведіть, що для довільного цілого $a$ та простого $p$ числа $a$ і $p$ будуть взаємно простими тоді й лише тоді, коли $a$ не ділиться на $p$.

**Означення 2.6.** Число $m$ називається *найменшим спільним кратним* чисел $a$ і $b$, якщо воно задовольняє дві умови:

1) $a \mid m$ і $b \mid m$;

2) якщо $a \mid n$ і $b \mid n$, то $m \mid n$.

Слово «найменше» у означенні розуміється в сенсі подільності, а не величини. Це спільне кратне, кратним якого буде будь-яке спільне кратне чисел $a$ і $b$. Найменше спільне кратне, як і найбільший спільний дільник, визначене з точністю до знака. Це дозволяє досягнути однозначності, якщо домовитись розглядати лише невід'ємне значення. Таке найменше спільне кратне чисел $a$ і $b$ позначають $\mathsf{НСК}(a,b)$ або $[a,b]$.

**Теорема 2.8** (Теорема про існування НСК і зв'язок з НСД)**.** *Для довільних натуральних чисел $a$ і $b$ має місце рівність*

$$\mathsf{НСК}(a,b) = \frac{ab}{\mathsf{НСД}(a,b)}.$$

*Доведення.* Нехай $\mathsf{НСД}(a,b) = d$. Тоді $a = a_0 d$, $b = b_0 d$ і

$$\frac{ab}{\mathsf{НСД}(a,b)} = a_0 b_0 d.$$

Число $m = a_0 b_0 d = b_0 a = a_0 b$ є спільним кратним чисел $a$ і $b$, тобто перша умова означення виконується.

Нехай тепер $M$ — якесь спільне кратне чисел $a$ і $b$. Тоді $M = aa_1 = bb_1$. З твердження 2.11 маємо $\mathsf{НСД}(a_0, b_0) = 1$, а тому існують такі цілі числа $k$ і $l$, що $1 = ka_0 + lb_0$. Помножимо обидві частини цієї рівності на $M$:

$$M = ka_0 M + lb_0 M = ka_0 bb_1 + lb_0 aa_1 = kmb_1 + lma_1 = m(kb_1 + la_1).$$

Звідки випливає $m \mid M$, тобто виконується і друга умова означення найменшого спільного кратного для числа $m$. □

Поняття найменшого спільного кратного також можна узагальнити на довільну скінченну кількість чисел наступним чином.

**Означення 2.7.** Число $m$ називається *найменшим спільним кратним* чисел

$$a_1, a_2, \ldots, a_k, \quad k \geq 1,$$

якщо $m$ задовольняє такі умови:

1) $a_1 \mid m$, $a_2 \mid m$, $\ldots$, $a_k \mid m$;

2) якщо $a_1 \mid n$, $a_2 \mid n$, $\ldots$, $a_k \mid n$, то $m \mid n$.

Для обчислення найменшого спільного кратного довільної скінченної кількості чисел достатньо вміти знаходити найменше спільне кратне двох чисел, що встановлює

**Твердження 2.12.** $\mathsf{HCK}(a_1, a_2, \ldots, a_k) = \mathsf{HCK}(\mathsf{HCK}(a_1, a_2), a_3, \ldots, a_k)$.

*Доведення.* Позначимо $m_{12} = \mathsf{HCK}(a_1, a_2)$ та $m = \mathsf{HCK}(m_{12}, a_3, \ldots, a_k)$. За означенням $m_{12} \mid m$ та $a_3 \mid m, \ldots, a_k \mid m$. Крім того, $a_1 \mid m_{12}$ і $a_2 \mid m_{12}$, звідки за транзитивністю відношення подільності $a_1 \mid m$ і $a_2 \mid m$. Отже, перша умова означення виконується.

Нехай $n$ таке, що $a_1 \mid n$, $a_2 \mid n$, $\ldots$, $a_k \mid n$. Тоді $m_{12} \mid n$ за означенням найменшого спільного кратного для двох чисел, а за означенням для $k-1$ числа $m_{12}$, $a_3$, $\ldots$, $a_k$ маємо $m \mid n$. Отже, виконується і друга умова означення та

$$\mathsf{HCK}(a_1, a_2, \ldots, a_k) = m.$$

Твердження доведено. □

## 2.3 Розклад на прості множники

**Теорема 2.9** (Основна теорема арифметики)**.** *Кожне натуральне число розкладається в добуток простих чисел, причому такий розклад є однозначним з точністю до порядку і знаків множників.*

*Доведення.* Існування розкладу доводить теорема 2.3.

Доведемо однозначність розкладу індукцією за довжиною $k$ найкоротшого розкладу (тобто розкладу, який містить найменшу кількість простих множників) натурального числа на прості множники.

База індукції. При $k = 0$ (число 1) однозначність розкладу має місце.

Крок індукції. Припустимо, що для всіх чисел, для яких існує розклад на менше ніж $k$ простих множників, однозначність розкладу вже

доведена. Нехай $n$ є числом, довжина деякого найкоротшого розкладу якого дорівнює $k$ та маємо два розклади числа $n$ на прості множники:

$$n = p_1 p_2 \dots p_k = q_1 q_2 \dots q_m, \quad m \geq k. \tag{2.3}$$

Якщо $\mathsf{НСД}(p_1, q_1) = 1$, $\mathsf{НСД}(p_1, q_2) = 1$, …, $\mathsf{НСД}(p_1, q_m) = 1$, то з твердження 2.8 випливає, що $\mathsf{НСД}(p_1, q_1 q_2 \dots q_m) = 1$. Але це суперечить рівності (2.3). Тому існує таке $i$, що $\mathsf{НСД}(p_1, q_i) \neq 1$. Оскільки $p_1$ і $q_i$ — прості числа, то вони або рівні, або протилежні. Змінивши, у разі потреби, в добутку $q_1 q_2 \dots q_m$ порядок множників і знаки двох множників, можемо вважати, що $i = 1$ та $p_1 = q_1$. Тоді рівність (2.3) набуває вигляду

$$p_1 p_2 \dots p_k = p_1 q_2 \dots q_m.$$

Після скорочення на $p_1$ отримаємо:

$$p_2 \dots p_k = q_2 \dots q_m.$$

Але добуток у лівій частині цієї рівності містить уже менше, ніж $k$ множників. Тому, за припущенням індукції, кількість множників в обох частинах рівності однакова (тобто $k = m$), а добутки $p_2 \dots p_k$ і $q_2 \dots q_m$ відрізняються щонайбільше порядком і знаками множників. Враховуючи $p_1 = q_1$, однозначність розкладу для числа $n$ доведена. □

Оскільки від'ємне число $n$ можна записати у вигляді $n = -|n|$, то з теореми 2.9 одразу випливає існування і однозначність розкладу в добуток простих множників і для від'ємних цілих чисел за виключенням числа $-1$. Це виключення пов'язано з тим, що число $-1$ є дільником числа 1, яке подається як добуток нульової кількості простих множників.

Природність виокремлення чисел $\pm 1$ при визначенні простих та складених чисел підкреслюється формулюванням теореми 2.9. Якби числа $\pm 1$ відносили до простих чисел, то зникла б однозначність розкладу в добуток простих, бо до кожного розкладу можна було б приєднати довільну кількість множників 1 та довільну парну кількість множників $-1$.

Легко бачити, що у розкладі натурального числа $n$ в добуток простих усі множники можна взяти додатними і впорядкувати їх за зростанням. Цим пояснюється коректність

**Означення 2.8.** *Канонічним* називається розклад натурального числа $n$ на прості множники вигляду

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m},$$

де $p_1$, $p_2$, …, $p_m$ — прості числа і $0 < p_1 < p_2 < \dots < p_m$.

Безпосередньо з основної теореми арифметики випливає

**Наслідок 2.1.** *Для кожного натурального числа існує єдиний канонічний розклад.*

Інколи корисно вважати, що деякі прості числа можуть зустрічатися в канонічних розкладах із нульовими показниками. Таким чином, наприклад, для числа 20 розклад

$$20 = 2^2 \cdot 5^1$$

є канонічним за означенням, а розклади

$$20 = 2^2 \cdot 3^0 \cdot 5^1 = 2^2 \cdot 3^0 \cdot 5^1 \cdot 7^0 = \dots$$

канонічні за домовленістю. Зручність такої домовленості можна побачити вже в наступній лемі.

**Лема 2.2.** *Якщо канонічний розклад натурального числа $n$ має вигляд*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m},$$

*то кожний натуральний дільник $d$ числа $n$ має вигляд*

$$d = p_1^{\beta_1} p_2^{\beta_2} \dots p_m^{\beta_m},$$

*де* $0 \le \beta_1 \le \alpha_1$, $0 \le \beta_2 \le \alpha_2$, $\dots$, $0 \le \beta_m \le \alpha_m$.

*Доведення.* Нехай $n = d \cdot f$ і канонічні розклади множників $d$ і $f$ мають вигляд

$$d = p_1^{\beta_1} p_2^{\beta_2} \dots p_m^{\beta_m}, \qquad f = p_1^{\gamma_1} p_2^{\gamma_2} \dots p_m^{\gamma_m},$$

де $\beta_1, \beta_2, \dots, \beta_m$ і $\gamma_1, \gamma_2, \dots, \gamma_m$ — цілі невід'ємні числа. Тоді канонічний розклад числа $n$ має вигляд

$$n = p_1^{\beta_1+\gamma_1} p_2^{\beta_2+\gamma_2} \dots p_m^{\beta_m+\gamma_m}.$$

Порівнюючи показники в канонічних розкладах числа $n$ та використовуючи єдиність розкладу, одержуємо $\alpha_i = \beta_i + \gamma_i$, звідки випливає $\alpha_i \ge \beta_i$ для всіх $i$, $1 \le i \le m$. □

Використовуючи лему 2.2 легко довести наступну теорему, яка встановлює зв'язок між канонічними розкладами натуральних чисел та їх найбільшим спільним дільником та найменшим спільним кратним.

**Теорема 2.10.** *Якщо для натуральних чисел $a$ і $b$ відомі їх канонічні розклади*

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m} \quad і \quad b = p_1^{\beta_1} p_2^{\beta_2} \dots p_m^{\beta_m},$$

*то*

$$\text{НСД}(a,b) = p_1^{\min\{\alpha_1,\beta_1\}} p_2^{\min\{\alpha_2,\beta_2\}} \dots p_m^{\min\{\alpha_m,\beta_m\}},$$
$$\text{НСК}(a,b) = p_1^{\max\{\alpha_1,\beta_1\}} p_2^{\max\{\alpha_2,\beta_2\}} \dots p_m^{\max\{\alpha_m,\beta_m\}}.$$

**Наслідок 2.2.** *Числа $a$ і $b$ є взаємно простими тоді і лише тоді, коли кожне просте число входить із додатним показником у канонічний розклад не більше ніж одного з чисел $a$ і $b$.*

Зауважимо, що задача знаходження розкладу натурального числа в добуток простих чисел є обчислювально складною, тобто для її розв'язання в загальному випадку не відомий швидкий алгоритм. Тому використання теореми 2.10 для знаходження найбільшого спільного дільника чи найменшого спільного кратного на практиці є неефективним.

## 2.4 Функція Ойлера

**Означення 2.9.** *Функцією Ойлера* називається функція

$$\varphi : \mathbb{N} \to \mathbb{N},$$

значення $\varphi(n)$ якої від $n \in \mathbb{N}$ дорівнює кількості натуральних чисел, які не більші за $n$ і взаємно прості з $n$.

Безпосередньою перевіркою легко переконатись, що

$$\varphi(1) = 1, \qquad \varphi(3) = 2, \qquad \varphi(5) = 4, \qquad \varphi(7) = 6,$$
$$\varphi(2) = 1, \qquad \varphi(4) = 2, \qquad \varphi(6) = 2, \qquad \varphi(8) = 4.$$

**Означення 2.10.** Функція $f : \mathbb{N} \to \mathbb{R}$ називається *мультиплікативною*, якщо функція $f$ не є тотожно рівною нулю та для довільних взаємно простих натуральних чисел $m$ і $n$ виконується рівність $f(mn) = f(m)f(n)$.

**Теорема 2.11.** *Функція Ойлера є мультиплікативною.*

*Доведення.* Оскільки $\varphi(1) = 1$, то функція $\varphi$ не є тотожно рівною нулю.

Покажемо, що для довільних взаємно простих натуральних чисел $m$ і $n$ виконується рівність $\varphi(mn) = \varphi(m)\varphi(n)$. Для цього всі натуральні числа від 1 до $mn$ випишемо у вигляді такої таблиці:

$$\begin{array}{cccccc} 1 & 2 & 3 & \cdots & m \\ m+1 & m+2 & m+3 & \cdots & 2m \\ 2m+1 & 2m+2 & 2m+3 & \cdots & 3m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (n-1)m+1 & (n-1)m+2 & (n-1)m+3 & \cdots & nm \end{array}$$

і знайдемо кількість тих чисел у цій таблиці, які взаємно прості з $mn$.

За твердженням 2.5 для довільних натуральних $k$, $m$ і $r$ маємо

$$\mathsf{НСД}(km + r, m) = \mathsf{НСД}(r, m),$$

оскільки остача від ділення чисел $km + r$ і $r$ на число $m$ однакова. Тому числа, які взаємно прості з $m$, а відтак і з $mn$, можуть бути лише в стовпчиках, номери $r$ яких взаємно прості з числом $m$. За означенням функції $\varphi$ кількість таких стовпчиків дорівнює $\varphi(m)$.

Стовпчик із номером $r$ складається з чисел

$$r, m + r, 2m + r, \ldots, (n-1)m + r.$$

Всі ці числа при діленні на $n$ дають попарно різні остачі. Справді, припустимо, що числа $k_1m + r$ і $k_2m + r$ при діленні на $n$ дають однакові остачі. Тоді

$$(k_2m + r) - (k_1m + r) = (k_2 - k_1)m$$

ділиться на $n$. Але $\mathsf{НСД}(m, n) = 1$, тому $k_2 - k_1$ ділиться на $n$, і з умови $0 \leq k_1, k_2 \leq n - 1$ випливає $k_1 = k_2$. Отже, всі числа стовпчика з номером $r$ при діленні на $n$ дають попарно різні остачі. Але кожен стовпчик містить рівно $n$ чисел, тому цими остачами будуть, у певному порядку, числа 0, 1, …, $n - 1$. За твердженням 2.5 число буде взаємно простим з $n$ тоді й лише тоді, коли взаємно простою з $n$ буде його остача від ділення на $n$. Тому кожний стовпчик містить рівно $\varphi(n)$ чисел, які взаємно прості з $n$.

Таким чином, в таблиці числа, взаємно прості з $m$, знаходяться в одному з $\varphi(m)$ стовпчиків, а кожний такий стовпчик містить $\varphi(n)$ чисел, взаємно простих з $n$. За твердженням 2.8 число буде взаємно простим з $mn$ тоді і тільки тоді, коли воно взаємно просте з $m$ та $n$. Тому

таблиця містить рівно $\varphi(m)\varphi(n)$ чисел, взаємно простих з $mn$. За означенням функції Ойлера кількість таких чисел дорівнює $\varphi(mn)$. Отже, $\varphi(mn) = \varphi(m)\varphi(n)$. □

**Теорема 2.12.** *Нехай $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ — канонічний розклад числа $n$. Тоді*

$$\varphi(n) = n\Big(1 - \frac{1}{p_1}\Big)\Big(1 - \frac{1}{p_2}\Big) \cdots \Big(1 - \frac{1}{p_k}\Big).$$

*Зокрема, для простого числа $p$*

$$\varphi(p) = p - 1 \quad і \quad \varphi(p^\alpha) = p^\alpha - p^{\alpha-1}.$$

*Доведення.* Розглянемо спочатку випадок $n = p^\alpha$, де $p$ — просте число. Число $m$ буде взаємно простим із числом $p^\alpha$ тоді й лише тоді, коли $m$ не ділиться на $p$. Серед чисел 1, 2, …, $p^\alpha$ на $p$ буде ділитися кожне $p$-те число. Тому таких чисел буде $p^\alpha/p = p^{\alpha-1}$. Решта $p^\alpha - p^{\alpha-1}$ чисел взаємно прості з $p$, а отже і з $p^\alpha$. Тому $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$. Зокрема, $\varphi(p) = p^1 - p^0 = p - 1$.

Тепер з мультиплікативності функції $\varphi$ у загальному випадку маємо:

$$\begin{aligned}\varphi(p_1^{\alpha_1} \dots p_k^{\alpha_k}) = \varphi(p_1^{\alpha_1}) \dots \varphi(p_k^{\alpha_k}) = (p_1^{\alpha_1} - p_1^{\alpha_1-1}) \dots (p_k^{\alpha_k} - p_k^{\alpha_k-1}) = \\ = p_1^{\alpha_1}\Big(1 - \frac{1}{p_1}\Big) \dots p_k^{\alpha_k}\Big(1 - \frac{1}{p_k}\Big) = n\Big(1 - \frac{1}{p_1}\Big) \dots \Big(1 - \frac{1}{p_k}\Big).\end{aligned}$$

Теорему доведено. □

**Теорема 2.13** (Ойлер)**.** *Для взаємно простих чисел $a$ та $n$ число $a^{\varphi(n)} - 1$ ділиться на $n$.*

*Доведення.* Нехай $k_1, k_2, \dots, k_{\varphi(n)}$ — всі натуральні числа, які не більші за $n$ і взаємно прості з числом $n$. Розглянемо числа

$$ak_1, \quad ak_2, \quad \dots, \quad ak_{\varphi(n)}. \tag{2.4}$$

За твердженням 2.8 всі вони є взаємно простими з числом $n$. Крім того, числа (2.4) дають попарно різні остачі при діленні на $n$. Дійсно, якщо $ak_i$ і $ak_j$ дають однакові остачі, то їх різниця $ak_i - ak_j = a(k_i - k_j)$ ділиться на $n$. Тоді за твердженням 2.9 на $n$ ділиться число $k_i - k_j$, що можливо лише у випадку $k_i = k_j$. Отже, множина остач від ділення

чисел (2.4) на $n$ співпадає з множиною чисел $k_1, k_2, \ldots, k_{\varphi(n)}$ та добуток всіх чисел (2.4) можна записати двома способами:

$$ak_1 \cdot ak_2 \cdot \ldots \cdot ak_{\varphi(n)} = (q_1 n + k_1)(q_2 n + k_2) \ldots (q_{\varphi(n)} n + k_{\varphi(n)}), \quad (2.5)$$

де множники зліва та справа знаходяться у взаємно однозначній відповідності та множнику $ak_i$ зліва, що дає частку $q_r$ і остачу $k_r$ при діленні на $n$, відповідає множник $q_r n + k_r$ справа. При розкритті дужок в правій частині рівності всі доданки, за виключенням добутку $k_1 \cdot k_2 \cdot \ldots \cdot k_{\varphi(n)}$ будуть ділитися на $n$, оскільки будуть містити множник $n$. Якщо їх суму позначити через $A \cdot n$, то рівність (2.5) можна переписати у вигляді

$$ak_1 \cdot ak_2 \cdot \ldots \cdot ak_{\varphi(n)} = A \cdot n + k_1 \cdot k_2 \cdot \ldots \cdot k_{\varphi(n)}.$$

Звідси випливає, що різниця

$$\begin{aligned} &ak_1 \cdot ak_2 \cdot \ldots \cdot ak_{\varphi(n)} - k_1 \cdot k_2 \cdot \ldots \cdot k_{\varphi(n)} = \\ &\quad = a^{\varphi(n)} k_1 k_2 \ldots k_{\varphi(n)} - k_1 k_2 \ldots k_{\varphi(n)} = (a^{\varphi(n)} - 1) k_1 k_2 \ldots k_{\varphi(n)} \end{aligned}$$

ділиться на $n$. За твердженням 2.8 добуток $k_1 k_2 \ldots k_{\varphi(n)}$ і число $n$ взаємно прості, а тому за твердженням 2.9 число $a^{\varphi(n)} - 1$ ділиться на $n$. □

## 2.5 Вправи

1. Обчисліть натуральне число $b$ і остачу $r$ від ділення $a = 19178$ на $b$, якщо частка від ділення $a$ на $b$ дорівнює 129.

2. Доведіть, що для кожного невід'ємного цілого числа $n$

$$7 \mid 2^{2^{2n}} + 5.$$

3. Доведіть, що існує нескінченно багато простих чисел, які мають вигляд $4k + 3$.

4. Знайдіть всі такі натуральні прості числа $p$, для яких кожне із чисел $p+4$ і $p+14$ також буде простим. Якою буде відповідь, якщо не вимагати, щоб число $p$ було натуральним?

5. Чи існують такі натуральні числа $a$, $b$, $c$, що $a \mid bc$ і $a \nmid b$, $a \nmid c$?

6. Користуючись розширеним алгоритмом Евкліда знайдіть найбільший спільний дільник і найменше спільне кратне чисел 9367 і 4318 та подайте найбільший спільний дільник у вигляді лінійної комбінації заданих чисел.

7. Доведіть, що

$$\text{НСК}(a,b,c) = \frac{abc \cdot \text{НСД}(a,b,c)}{\text{НСД}(a,b) \cdot \text{НСД}(a,c) \cdot \text{НСД}(b,c)}.$$

8. Обчисліть $\varphi(n)$ для

   а) $n = 2310$;

   б) $n = 16632$;

   в) $n = 43560$.

9. Знайдіть всі такі натуральні числа $n$, що $n$ є добутком двох різних простих чисел, а $\varphi(n)$ дорівнює:

   а) 24;

   б) 100.

# Розділ 3

# Комплексні числа

## 3.1 Поняття кільця та поля

Нехай $M$ — непорожня множина. *Бінарною операцією* на множині $M$ називається відображення, яке кожній парі елементів $a, b \in M$ ставить у відповідність елемент з множини $M$. Для позначення бінарних операцій використовують різні математичні символи, серед яких $+$, $\cdot$, $*$, при цьому застосування відповідної бінарної операції до елементів $a, b \in M$ записують у вигляді $a + b$, $a \cdot b$, $a * b$.

Бінарна операція $*$ на множині $M$ називається *асоціативною*, якщо для довільних $a, b, c \in M$ виконується рівність

$$(a * b) * c = a * (b * c).$$

Операція $*$ називається *комутативною*, якщо для довільних $a, b \in M$ виконується рівність

$$a * b = b * a.$$

Елемент $e \in M$ називається *нейтральним* відносно операції $*$, якщо для всіх $a \in M$ виконується

$$a * e = e * a = a.$$

При цьому елемент $b \in M$ називають *оберненим* до $a \in M$ відносно операції $*$, якщо

$$a * b = b * a = e.$$

**Означення 3.1.** Множина $G$ з визначеною на ній бінарною операцією $*$ називається *групою*, якщо операція $*$ є асоціативною, відносно цієї операції існує нейтральний елемент та кожний елемент з множини $G$ має

обернений відносно $*$. Група $G$ називається *абелевою* (або *комутативною*), якщо операція $*$ є комутативною.

Прикладами груп є множина цілих чисел відносно додавання, множина додатних раціональних чисел відносно множення, множина всіх підмножин фіксованої множини відносно симетричної різниці. Всі ці групи є прикладами абелевих груп. Прикладом неабелевої групи є множина всіх взаємно однозначних перетворень фіксованої множини відносно композиції.

Про дві бінарні операції $*$, $+$ на множині $M$ кажуть, що операція $*$ є *дистрибутивною* відносно операції $+$, якщо для довільних $a, b, c \in M$ виконуються рівності

$$a * (b + c) = (a * b) + (a * c),$$
$$(b + c) * a = (b * a) + (c * a).$$

**Означення 3.2.** Множина $K$ з визначеними на ній бінарними операціями $+$ та $\cdot$, які прийнято називати додаванням та множенням, називається *кільцем*, якщо $K$ відносно додавання є абелевою групою, множення є асоціативним та дистрибутивним відносно додавання. Якщо множення є комутативним, то кільце $K$ називається *комутативним*. Якщо відносно множення існує нейтральний елемент, то кільце $K$ називається *кільцем з одиницею*.

**Означення 3.3.** Множина $P$ з визначеними на ній бінарними операціями $+$ та $\cdot$, які прийнято називати додаванням та множенням, називається *полем*, якщо множина $P$ відносно додавання є абелевою групою, нейтральний елемент якої прийнято називати *нулем*, множина ненульових елементів з множини $P$ відносно множення є абелевою групою та множення є дистрибутивним відносно додавання.

Прикладами кілець є цілі числа та парні цілі числа відносно звичайних операцій додавання і множення. Ці кільця є комутативними, перше з яких є кільцем з одиницею. Множини раціональних та дійсних чисел відносно звичайних операцій додавання і множення утворюють поле, при цьому вони також є прикладами комутативних кілець з одиницею, оскільки кожне поле є комутативним кільцем з одиницею. Зауважимо, що множина цілих чисел $\mathbb{Z}$ з діями додавання і множення, які на ній визначені, полем не є (чому?).

## 3.2 Побудова поля комплексних чисел

Нагадаємо, що непорожню множину $P$ називають полем, якщо:

1) на ній визначено дві дії, одна з яких називається додаванням, а друга множенням, причому кожна з цих дій є асоціативною і комутативною, а множення є дистрибутивним відносно додавання;

2) вона містить такі елементи 0 і 1 такі, що додавання 0 і множення на 1 не змінює жоден елемент з $P$;

3) кожен елемент з $P$ має протилежний елемент і кожен ненульовий елемент з $P$ має обернений елемент (це, відповідно, дає змогу віднімати від довільного елемента з $P$ будь-який інший елемент і ділити довільний елемент на довільний ненульовий елемент, тобто вводити в $P$ дії віднімання і ділення).

Проблема побудови поля комплексних чисел включає в себе завдання побудувати поле, яке

1) містить поле дійсних чисел;

2) містить розв'язок рівняння $x^2 = -1$;

3) результати виконання дій додавання і множення для комплексних чисел, які є дійсними, збігається з раніше визначеними результатами для них як для дійсних чисел.

Тобто вимагається фактично так розширити поле дійсних чисел, щоб нова множина знову була полем, але вже містила деякий новий об'єкт — розв'язок рівняння $x^2 = -1$. Розширення здійснюється таким чином.

**Означення 3.4.** Комплексним числом називається вираз вигляду $a+bi$, в якому $a$ та $b$ — деякі дійсні числа, а символ $i$ позначає розв'язок рівняння $x^2 = -1$, тобто $i^2 = -1$.

Запис $a + bi$ називають ще алгебраїчною формою комплексного числа. Для комплексного числа $z = a + bi$ дійсне число $a$ називають його дійсною частиною і позначають $\operatorname{Re} z$, а дійсне число $b$ називають його уявною частиною і позначають $\operatorname{Im} z$. Комплексне число, у якого уявна частина рівна 0, ототожнюють з його дійсною частиною (тобто воно є дійсним, а відтак виконана перша умова поставленої проблеми). Якщо у комплексного числа дійсна частина рівна 0, то таке число називають чисто уявним. Таким буде зокрема число $i$ (отже, виконується і друга умова поставленої проблеми).

Два комплексних числа $z_1 = a+bi$ і $z_2 = c+di$ називаються рівними, якщо у них рівні як дійсні, так і уявні частини, тобто $z_1 = z_2$ в тому

і лише тому випадку, коли $a = c$ і $b = d$. Це означає, що для кожного комплексного числа його дійсна та уявна частини визначені однозначно.

Для числа $z = a + bi$ спряженим до нього числом назвемо число $\bar{z} = a - bi$. Зрозуміло, що $\bar{\bar{z}} = z$.

Символом $\mathbb{C}$ позначимо множину всіх комплексних чисел. Визначимо на цій множині арифметичні дії. Нехай $z_1 = a + bi$ і $z_2 = c + di$ — деякі комплексні числа.

**Означення 3.5.** Сумою комплексних чисел $z_1, z_2$ називається число

$$z_1 + z_2 = (a + c) + (b + d)i.$$

Протилежним до числа $z_1$ буде число $-a + (-b)i$. Тоді різниця чисел $z_1$ і $z_2$ визначається рівністю

$$z_1 - z_2 = z_1 + (-z_2) = (a - c) + (b - d)i.$$

**Означення 3.6.** Добутком комплексних чисел $z_1, z_2$ називається число, яке обчислюється за правилом

$$z_1 z_2 = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Звідси одразу одержуємо, що

$$z_1 \bar{z}_1 = a^2 + b^2,$$

тобто добуток комплексного числа на спряжене до нього є числом дійсним.

Якщо число $z_1$ є ненульовим, то обернене до нього число можна обчислити так:

$$(z_1)^{-1} = \frac{1}{z_1} = \frac{\bar{z}_1}{z_1 \bar{z}_1} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} i.$$

Тепер можемо записати правило обчислення частки, яка визначена лише при $z_2 \neq 0$

$$\frac{z_1}{z_2} = z_1 z_2^{-1} = (a + bi)\left(\frac{c}{c^2 + d^2} + \frac{-d}{c^2 + d^2} i\right) = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2} i.$$

Таким чином, сума, різниця і частка (якщо вона визначена) комплексних чисел є комплексними числами.

Оскільки додавання комплексних чисел зводиться до додавання їх дійсних і уявних частин, то дія додавання в $\mathbb{C}$ є асоціативною і комутативною. Комплексне число $0 = 0 + 0i$ є нейтральним елементом для додавання. Можна перевірити, що множення комплексних чисел також є асоціативною і комутативною операцією, а також дистрибутивною відносно додавання. Крім того, множення на число $1 = 1 + 0i$ не змінює комплексні числа. Це означає, що множина $\mathbb{C}$ справді є полем, причому третя умова задачі про побудову поля комплексних чисел виконується. (Перевірте всі твердження цього абзацу.)

Задачу побудови поля комплексних чисел можна розв'язати більш строго і показати, що в певному розумінні вона має єдиний розв'язок.

## 3.3 Геометрична інтерпретація

Розглянемо на площині прямокутну декартову систему координат. Комплексному числу $z = a + bi$ поставимо у відповідність точку на цій площині, яка має координати $(a, b)$, тобто абсцисою цієї точки є дійсна частина числа $z$, а ординатою — його уявна частина. Така відповідність між комплексними числами і точками площини буде взаємно однозначною, тобто кожному числу відповідає єдина точка і кожній точці — єдине число. Площину, точки якої взаємно однозначно відповідають комплексним числам, будемо називати комплексною площиною. При цьому вісь абсцис називають дійсною віссю, а вісь ординат — уявною віссю. Зрозуміло, що дійсним числам відповідають точки дійсної осі, а чисто уявним — точки уявної осі (Рис. 3.1).

Іноді зручно співставляти числу $z = a + bi$ не лише точку з координатами $(a, b)$, а й радіус-вектор цієї точки, тобто вектор, початком якого є початок координат, а кінцем — точка $(a, b)$. Тоді радіус-вектор, що відповідає сумі двох комплексних чисел є сумою радіус-векторів, які відповідають доданкам (Рис. 3.2).

Введемо на комплексній площині полярну систему координат, вибравши її полюсом початок прямокутної декартової системи координат, а полярною віссю — додатну дійсну піввісь. В полярній системі координат положення точки визначається відстанню від неї до полюса (тобто довжиною радіус-вектора цієї точки) і кутом між полярною віссю і радіус-вектором точки, причому додатним відліком будемо вважати відлік проти руху стрілки годинника (від додатної дійсної до додатної уявної півосі). Для комплексного числа $z$ довжину радіус-вектора точки

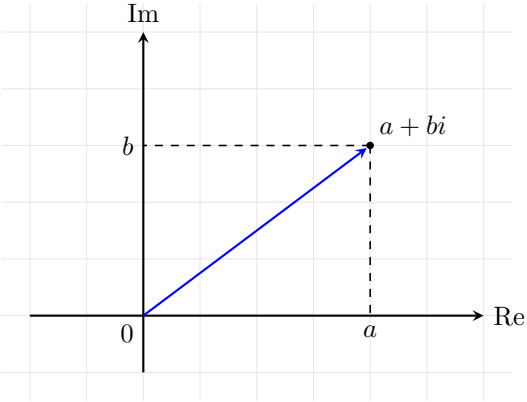



Рис. 3.1: Комплексна площина

на комплексній площині, яка йому відповідає, будемо називати модулем цього комплексного числа і позначати $|z|$. Кут же між цим радіус-вектором і полярною віссю назвемо аргументом числа $z$ і позначимо $\arg z$ (Рис. 3.3).

Для кожного комплексного числа його модуль є невід'ємним дійсним числом і визначається однозначно. Поняття модуля комплексного числа узагальнює добре відоме поняття модуля дійсного числа. Аргумент є числом дійсним і визначається неоднозначно. Точніше, для ненульового комплексного числа дійсні числа, які є його аргументами — це множина всіх тих дійсних чисел, з яких будь-які два відрізняються на доданок, цілочисельно кратний $2\pi$. При цьому для числа 0 аргумент вважається невизначеним.

## 3.4 Тригонометрична форма

Для комплексного числа $z = a + bi$ позначимо його модуль символом $r$, а аргумент — символом $\varphi$. Враховуючи зв'язок між координатами точки у прямокутній декартовій та полярній системах координат, маємо рівності $a = r\cos\varphi$, $b = r\sin\varphi$, звідки випливає рівність

$$z = r(\cos\varphi + i\sin\varphi).$$

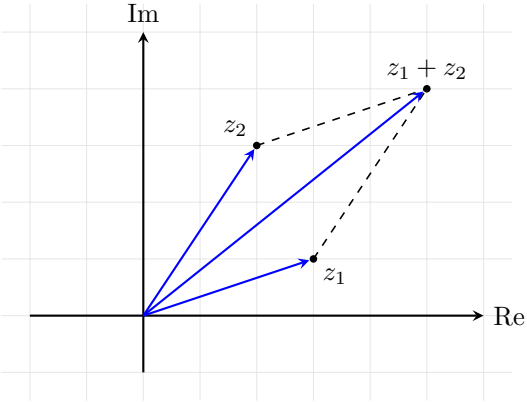



Рис. 3.2: Додавання комплексних чисел

Праву частину цієї рівності називають тригонометричною формою комплексного числа $z$. Якщо модуль комплексного числа дорівнює 1, то в його тригонометричній формі його, як правило, не пишуть. Для переходу до тригонометричної форми від алгебраїчної маємо рівності $r = \sqrt{a^2 + b^2}$ і $\cos\varphi = a/r$, $\sin\varphi = b/r$ $(z \neq 0)$.

За означенням рівності двох комплексних чисел, рівність

$$r_1(\cos\varphi_1 + i\sin\varphi_1) = r_2(\cos\varphi_2 + i\sin\varphi_2)$$

має місце тоді й лише тоді, коли $r_1 = r_2$ і $\varphi_1 - \varphi_2 = 2\pi k$ для деякого цілого $k$.

Добуток двох комплексних чисел

$$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1) \quad \text{і} \quad z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$$

обчислюється таким чином:

$$\begin{aligned} z_1 z_2 = (r_1\cos\varphi_1 + ir_1\sin\varphi_1)(r_2\cos\varphi_2 + ir_2\sin\varphi_2) = \\ = (r_1 r_2\cos\varphi_1\cos\varphi_2 - r_1 r_2\sin\varphi_1\sin\varphi_2) + \\ + i(r_1 r_2\cos\varphi_1\sin\varphi_2 + r_1 r_2\cos\varphi_2\sin\varphi_1) = \\ = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)). \end{aligned}$$

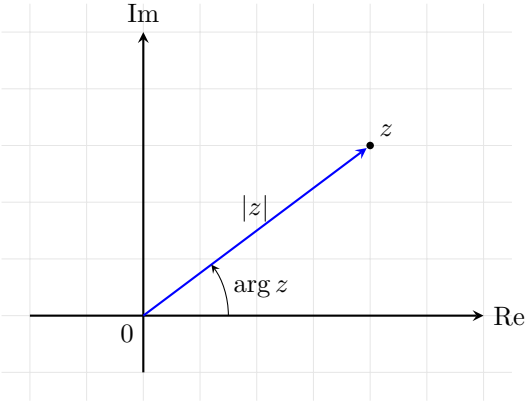



Рис. 3.3: Модуль і аргумент комплексних чисел

Якщо ж число $z_2 \neq 0$, то обернене до нього число дорівнює

$$\frac{1}{z_2} = \frac{\cos\varphi_2 - i\sin\varphi_2}{r_2(\cos^2\varphi_2 + \sin^2\varphi_2)} = r_2^{-1}(\cos(-\varphi_2) + i\sin(-\varphi_2)).$$

Звідси для обчислення частки одержуємо

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)).$$

Таким чином, дістаємо таке правило для множення і ділення комплексних чисел в тригонометричній формі. Модуль добутку (частки) комплексних чисел дорівнює добутку (частці) модулів множників, а аргумент добутку (частки) — сумі (різниці) аргументів множників:

$$|z_1 z_2| = |z_1| \cdot |z_2|, \quad \arg(z_1 z_2) = \arg(z_1) + \arg(z_2),$$

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \quad \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2).$$

Це означає, що радіус-вектор точки, яка відповідає добутку $z_1 z_2$ отримується з радіус-вектора точки, що відповідає числу $z_1$, поворотом на кут $\arg z_2$ в додатному напрямку і розтягненням у $|z_2|$ раз.

Як наслідок з цього спостереження одержуємо зручну формулу для піднесення комплексного числа до цілого степеня.

### 3.4.1 Формула Муавра

Для комплексного числа $z = r(\cos\varphi + i\sin\varphi)$ і цілого числа $n$ має місце рівність

$$z^n = r^n(\cos n\varphi + i\sin n\varphi).$$

*Доведення.* Перевіримо методом математичної індукції, що рівність виконується для невід'ємних степенів $n$.

База індукції: $n = 0$. Рівність виконується очевидним чином.

Припустимо, що для деякого невід'ємного $n$ твердження про рівність доведено і зробимо індуктивний крок. Маємо

$$\begin{aligned} z^{n+1} = z^n z = \ &\text{(за припущенням індукції)} \ = \\ &= r^n(\cos n\varphi + i\sin n\varphi)\cdot r(\cos\varphi + i\sin\varphi) = \\ = \ &\text{(за правилом множення)} \ = r^{n+1}(\cos(n+1)\varphi + i\sin(n+1)\varphi). \end{aligned}$$

Отже, для невід'ємних цілих степенів рівність має місце.

Нехай тепер $n$ — від'ємне ціле число. Тоді

$$\begin{aligned} z^n = (z^{-n})^{-1} = \frac{1}{z^{-n}} = \ &\text{(за щойно доведеним)} \ = \\ = \frac{1}{r^{-n}(\cos(-n)\varphi + i\sin(-n)\varphi)} = \ &\text{(за правилом ділення)} \ = \\ &= r^n(\cos n\varphi + i\sin n\varphi). \end{aligned}$$

Отже, рівність перевірено для всіх цілих степенів. □

## 3.5 Корені з комплексних чисел

Зафіксуємо деяке натуральне число $n$.

**Означення 3.7.** Коренем $n$-го степеня з комплексного числа $z$ назвемо таке комплексне число $\omega$, що $\omega^n = z$.

З означення добутку комплексних чисел маємо, що коренем $n$-го степеня з числа 0 буде саме тільки число 0. Тому далі розглядаємо лише ненульове число $z$. Запишемо його в тригонометричній формі:

$$z = r(\cos\varphi + i\sin\varphi).$$

Потрібно встановити, чи існують корені $n$-го степеня з числа $z$, а якщо так, то знайти їх кількість і загальний вигляд. Будемо шукати корінь $n$-го степеня з цього числа у тригонометричній формі:

$$\omega = \rho(\cos\xi + i\sin\xi).$$

Рівність $\omega^n = z$ тоді перепишеться у вигляді

$$\rho^n(\cos n\xi + i\sin n\xi) = r(\cos\varphi + i\sin\varphi).$$

Вона рівносильна рівностям

$$\rho^n = r, \quad n\xi = \varphi + 2\pi k \text{ для деякого } k \in \mathbb{Z}.$$

Тобто

$$\rho = \sqrt[n]{r}, \quad \xi = \frac{\varphi + 2\pi k}{n} \text{ для деякого } k \in \mathbb{Z}.$$

Позначимо

$$\omega_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right), \quad k \in \mathbb{Z}$$

і покажемо, що числа $\omega_k$ і $\omega_l$ рівні тоді й лише тоді, коли різниця $k - l$ ділиться на $n$. Справді, рівність $\omega_k = \omega_l$ рівносильна рівності

$$\frac{\varphi + 2\pi k}{n} = \frac{\varphi + 2\pi l}{n} + 2\pi m, \quad m \in \mathbb{Z},$$

яка, в свою чергу, рівносильна такій рівності

$$k - l = mn, \quad m \in \mathbb{Z}.$$

Це й означає, що $k - l$ ділиться на $n$. Таким чином, кожне з чисел $\omega_k$, $k \in \mathbb{Z}$, дорівнює якомусь з чисел $\omega_l$, $l = 0, 1, \ldots, n-1$, які є попарно різними. Підсумовуючи все сказане, одержуємо

**Твердження 3.1.** *Для ненульового комплексного числа*

$$z = r(\cos\varphi + i\sin\varphi)$$

*і натурального $n$ існує рівно $n$ коренів $n$-го степеня з числа $z$ і вони обчислюються за формулою*

$$\omega_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right), \quad k = 0, 1, \ldots, n-1.$$

**Приклад 3.1.** Нехай $z = 2 + 2i$, $n = 3$. Тоді $z = \sqrt{8}(\cos \pi/4 + i \sin \pi/4)$ і корені третього степеня з $z$ мають вигляд (див. рис. 3.4)

$$\omega_k = \sqrt[6]{8}\left(\cos\frac{\pi/4 + 2\pi k}{3} + i\sin\frac{\pi/4 + 2\pi k}{3}\right), \quad k = 0, 1, 2.$$

Звідси

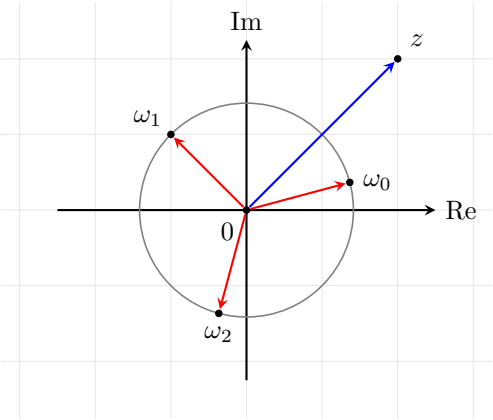



Рис. 3.4: Корені степеня 3 з числа $2 + 2i$

$$\omega_0 = \sqrt{2}\Big(\cos\frac{\pi}{12} + i\sin\frac{\pi}{12}\Big),$$
$$\omega_1 = \sqrt{2}\Big(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}\Big) = -1 + i,$$
$$\omega_2 = \sqrt{2}\Big(\cos\frac{17\pi}{12} + i\sin\frac{17\pi}{12}\Big).$$

**Вправа 3.1.** Знайдіть алгебраїчну форму $\omega_0$ і $\omega_2$.

На комплексній площині кореням $n$-го степеня з числа $z$ відповідають вершини правильного $n$-кутника, вписаного в коло радіуса $\sqrt[n]{|z|}$ з центром у початку координат, причому одна з вершин лежить на перетині з колом променя, який проходить через початок координат і нахилений до осі $Ox$ під кутом $\varphi/n$.

## 3.6 Корені з одиниці

Тригонометрична форма числа 1 має вигляд

$$\cos 0 + i \sin 0.$$

Тому для обчислення коренів $n$-го степеня з одиниці отримуємо таку просту формулу

$$\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, \quad k = 0, 1, \ldots, n-1.$$

Одним з цих коренів буде саме число 1. Крім того,

$$\varepsilon_1^k = \varepsilon_k, \quad k = 0, 1, \ldots, n-1.$$

На комплексній площині кореням $n$-го степеня з одиниці відповідають вершини правильного $n$-кутника (див. рис. 3.5), вписаного в одиничне коло з центром у початку координат, причому одна з вершин відповідає числу 1.

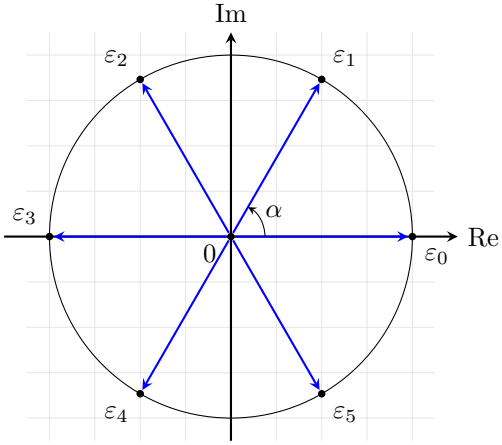



Рис. 3.5: Корені з одиниці степеня 6

Множину коренів $n$-го степеня з одиниці позначимо символом $\mathbb{C}_n$. Зауважимо, що ця множина збігається з множиною розв'язків рівняння $x^n - 1 = 0$. Сформулюємо основні властивості цієї множини.

**Твердження 3.2.** *1) Добуток коренів $n$-го степеня з одиниці є коренем $n$-го степеня з одиниці.*
*2) Число, обернене до кореня $n$-го степеня з одиниці, є коренем $n$-го степеня з одиниці.*
*3) Число, спряжене до кореня $n$-го степеня з одиниці, є коренем $n$-го степеня з одиниці.*

*Доведення.* 1) Нехай $z_1, z_2 \in \mathbb{C}_n$. Тоді $(z_1 z_2)^n = z_1^n z_2^n = 1$, звідки $z_1 z_2 \in \mathbb{C}_n$.

2) Нехай $z \in \mathbb{C}_n$. Тоді $(z^{-1})^n = (z^n)^{-1} = 1$, звідки $z^{-1} \in \mathbb{C}_n$.

3) Це твердження випливає з того, що для числа $z = \cos\varphi + i\sin\varphi$ з модулем одиниця спряжене до нього число $\bar{z} = \cos\varphi - i\sin\varphi$ дорівнює оберненому до нього $z^{-1} = \cos(-\varphi) + i\sin(-\varphi)$. Після цього залишається лише застосувати попередню властивість. □

Серед коренів степеня $n$ з одиниці важливу роль відіграють так звані первісні корені. А саме, корінь $\varepsilon$ $n$-го степеня з одиниці називається первісним коренем $n$-го степеня з одиниці, якщо він не є коренем з одиниці ніякого степеня, меншого за $n$, тобто $\varepsilon^m \neq 1$ для всіх натуральних $m < n$. З формули Муавра одержуємо, що число

$$\varepsilon_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$$

є первісним коренем $n$-го степеня з 1. Також легко зрозуміти, що кожен корінь $n$-го степеня з одиниці буде первісним коренем з одиниці деякого степеня. А саме, має місце

**Твердження 3.3.** *Нехай $d$ — найбільший спільний дільник цілого числа $k$ і числа $n$. Тоді число $\varepsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}$ є первісним коренем степеня $n/d$ з одиниці.*

*Доведення.* Маємо $k = dk_1$, $n = dn_1$ для деяких цілого $k_1$ і натурального $n_1$, причому числа $k_1$ і $n_1$ є взаємно простими. Тому аргумент числа $\varepsilon_k$ дорівнює $\frac{2\pi k_1}{n_1}$. Звідси за формулою Муавра $\varepsilon_k^{n_1} = 1$. Тобто число $\varepsilon_k$ є коренем степеня $n_1$ з одиниці.

Покажемо, що $n_1$ є найменшим степенем, у якому $\varepsilon_k$ дорівнює одиниці. Якщо $\varepsilon_k^m = 1$, то $\frac{2\pi k_1 m}{n_1} = 2\pi l$ для деякого цілого $l$. Звідси $k_1 m = n_1 l$. Враховуючи взаємну простоту чисел $k_1$ і $n_1$ звідси одразу одержуємо, що $m$ ділиться на $n_1$, тобто $m$ не може бути натуральним числом, меншим за $n_1$. □

Звідси отримуємо таку характеризацію первісних коренів $n$-го степеня з одиниці.

**Наслідок 3.1.** *Число* $\varepsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}$ $(k \in \mathbb{Z})$ *є первісним коренем* $n$*-го степеня з одиниці тоді й лише тоді, коли числа* $k$ *і* $n$ *є взаємно простими.*

*Доведення.* Нехай $d$ — найбільший спільний дільник $k$ і $n$. За попереднім твердженням $\varepsilon_k$ є первісним коренем степеня $n/d$ з одиниці. Але $n = n/d$ тоді й лише тоді, коли $k$ і $n$ є взаємно простими числами. □

Для натурального числа $n$ визначимо число $\varphi(n)$ як кількість натуральних чисел, які не перевищують $n$ і є взаємно простими з $n$. Тоді отримаємо натуральнозначну функцію натурального аргумента $\varphi : \mathbb{N} \to \mathbb{N}$, яку називають функцією Ойлера. З попереднього твердження тепер маємо

**Наслідок 3.2.** *Кількість первісних коренів степеня* $n$ *з одиниці дорівнює* $\varphi(n)$.

Наведемо тепер ще одне твердження, котре характеризує первісні корені $n$-го степеня з одиниці.

**Теорема 3.1.** *Корінь* $\varepsilon$ $n$*-го степеня з одиниці є первісним коренем* $n$*-го степеня з одиниці тоді й лише тоді, коли кожен корінь* $n$*-го степеня з одиниці можна отримати, піднісши* $\varepsilon$ *до деякого натурального степеня.*

*Доведення.* Нехай $\varepsilon$ є первісним коренем $n$-го степеня з одиниці. Розглянемо числа $\varepsilon$, $\varepsilon^2$, …, $\varepsilon^n$. Всі вони є коренями $n$-го степеня з одиниці і досить переконатися, що вони є попарно різними. Припустимо, що це не так. Нехай $\varepsilon^l = \varepsilon^m$, де $1 \le l < m \le n$. Це рівносильно тому, що для натурального числа $m - l$, яке менше за $n$, має місце рівність $\varepsilon^{m-l} = 1$, тобто $\varepsilon$ не є первісним коренем $n$-го степеня з одиниці. Суперечність. Отже, всі $n$ коренів степеня $n$ з одиниці містяться серед $\varepsilon$, $\varepsilon^2$, …, $\varepsilon^n$.

Навпаки, припустимо, що число $\varepsilon$ не є первісним коренем $n$-го степеня з одиниці, тобто $\varepsilon^k = 1$ для деякого $k < n$. Виберемо довільне натуральне $m$ і поділимо його з остачею на $k$: $m = kq + r$ і $0 \le r \le k - 1$. Тоді $\varepsilon^m = (\varepsilon^k)^q\varepsilon^r = \varepsilon^r$. Тому, підносячи $\varepsilon$ до натурального степеня, можна отримати не більше $k$ різних чисел. Оскільки ж коренів степеня $n$ з одиниці є $n$, то всіх їх так отримати неможливо. □

## 3.7 Вправи

1. Обчисліть в алгебраїчній формі:

   а) $(2+i)(3-i)+(2+3i)(3+4i)$;

   б) $\dfrac{(5+i)(7-6i)}{3+i}$.

2. Розв'яжіть рівняння
$$z^2 = 3-4i.$$

3. Знайдіть тригонометричну форму числа:

   а) $1-i$;

   б) $-1+i\sqrt{3}$.

4. Який геометричний зміст модуля різниці двох комплексних чисел?

5. Зобразіть на комплексній площині множину точок, що відповідають заданій множині комплексних чисел:

   а) $|z|=1$;

   б) $\arg z = \pi/3$;

   в) $|z-1-i|<1$.

6. Обчисліть корені $n$-го степеня з числа $z$:

   а) $z=-4$, $n=4$;

   б) $z=8(\sqrt{3}-i)$, $n=4$;

   в) $z=1+i$, $n=3$.

# Розділ 4

# Модулярна арифметика

## 4.1 Лінійні діофантові рівняння

*Лінійним діофантовим* рівнянням від $m$ змінних, $m \geq 2$, називається рівняння вигляду

$$a_1x_1 + \cdots + a_mx_m = b, \tag{4.1}$$

в якому $a_1, \ldots, a_m, b \in \mathbb{Z}$, а *розв'язком* є впорядкований набір $m$ цілих чисел, при підстановці якого в рівняння замість змінних одержується правильна рівність. Якщо всі коефіцієнти дорівнюють нулю, то або рівняння не має розв'язків у випадку $b \neq 0$, або будь-який набір цілих чисел є розв'язком у випадку $b = 0$. Тому надалі вважаємо, що серед коефіцієнтів рівняння є ненульові.

Обмежимось спочатку випадком лінійного діофантового рівняння з двома невідомими. Критерій існування розв'язку такого рівняння дає

**Теорема 4.1.** *Лінійне діофантове рівняння $ax + by = c$ має розв'язки тоді й лише тоді, коли його права частина $c$ ділиться на найбільший спільний дільник чисел $a$ і $b$.*

*Доведення.* Необхідність. Нехай розв'язок заданого рівняння існує. Сума $am+bn$ завжди ділиться на найбільший спільний дільник чисел $a$ і $b$.

Достатність. Позначимо символом $d$ найбільший спільний дільник чисел $a$ і $b$. Розглянемо для нього тотожність Безу $ma + nb = d$, де $m$ і $n$ — цілі. Нехай $c = c_0d$. Тоді з рівності $amc_0 + bnc_0 = dc_0$ випливає, що пара цілих чисел $(mc_0, nc_0)$ є розв'язком рівняння $ax + by = c$. □

**Теорема 4.2.** *Якщо лінійне діофантове рівняння*

$$ax + by = c \tag{4.2}$$

*має розв'язок* $(x_0, y_0)$*, то воно має нескінченно багато розв'язків, і кожен з них можна знайти за формулами*

$$x = x_0 - \frac{b}{d} \cdot k, \quad y = y_0 + \frac{a}{d} \cdot k, \tag{4.3}$$

*де* $d = \mathsf{НСД}(a, b)$*, а* $k$ *— довільне ціле число.*

*Доведення.* Позначимо $a/d = a_0$, $b/d = b_0$. За умовою теореми

$$ax_0 + by_0 = c. \tag{4.4}$$

Тому для цілих $x_0 - b_0 k$ і $y_0 + a_0 k$ виконується

$$a(x_0 - b_0 k) + b(y_0 + a_0 k) = ax_0 - \frac{ab}{d} \cdot k + by_0 + \frac{ba}{d} \cdot k = ax_0 + by_0 = c.$$

Отже, формули (4.3) справді визначають розв'язок рівняння (4.2).

Нехай тепер $(x, y)$ — довільний розв'язок рівняння (4.2). Віднімаючи (4.2) від рівності (4.4), одержимо

$$a(x_0 - x) + b(y_0 - y) = 0,$$

після ділення на $d$,

$$a_0(x_0 - x) + b_0(y_0 - y) = 0. \tag{4.5}$$

Звідси випливає, що $a_0 \mid b_0(y - y_0)$. Але числа $a_0$ і $b_0$ взаємно прості за твердженням 2.11. Тому $a_0 \mid y - y_0$. Отже, існує таке ціле число $k$, що $y - y_0 = a_0 k$ і $y = y_0 + a_0 k$. Рівність (4.5) тепер набуває вигляду

$$a_0(x_0 - x) = b_0 a_0 k,$$

звідки $x_0 - x = b_0 k$ і $x = x_0 - b_0 k$. Таким чином, кожен розв'язок рівняння (4.2) має вигляд (4.3), що й треба було довести. □

Таким чином, щоб розв'язати лінійне діофантове рівняння $ax+by = c$, потрібно виконати такі дії.

1. За допомогою алгоритму Евкліда знайти найбільший спільний дільник $d$ чисел $a$, $b$ і перевірити, чи $d \mid c$. Якщо $d \nmid c$, то рівняння розв'язків не має.

2. Якщо $d \mid c$, то за допомогою розширеного алгоритму Евкліда знайти лінійне зображення $d = ma + nb$ і частковий розв'язок
$$x_0 = m \cdot \frac{c}{d}, \quad y_0 = n \cdot \frac{c}{d}.$$

3. Знайти загальний розв'язок
$$x = x_0 - \frac{b}{d} \cdot k, \quad y = y_0 + \frac{a}{d} \cdot k, \quad k \in \mathbb{Z}.$$

Коректність цього алгоритму випливає з теорем 4.1 та 4.2.

Перейдемо тепер до загального випадку. Позначимо через $d$ найбільший спільний дільник коефіцієнтів рівняння (4.1). Якщо це рівняння має розв'язки, і впорядкований набір $(k_1, k_2, \ldots, k_m)$ є одним з них, то сума
$$a_1k_1 + a_2k_2 + \cdots + a_mk_m$$
ділиться на $d$, а тому і число $b$ ділиться на $d$. Отже, якщо число $b$ не ділиться на $d$, то рівняння (4.1) не має розв'язків. Надалі розглядаємо лише випадок, коли число $b$ ділиться на $d$.

Будемо казати, що множина розв'язків рівняння (4.1) має цілочисельну параметризацію, якщо вона збігається з множиною впорядкованих наборів $(k_1, k_2, \ldots, k_m)$, які задаються рівностями
$$\begin{aligned}
k_1 &= c_{10} + c_{11}t_1 + c_{12}t_2 + \cdots + c_{1,m-1}t_{m-1},\\
k_2 &= c_{20} + c_{21}t_1 + c_{22}t_2 + \cdots + c_{2,m-1}t_{m-1},\\
&\vdots\\
k_m &= c_{m0} + c_{m1}t_1 + c_{m2}t_2 + \cdots + c_{m,m-1}t_{m-1},
\end{aligned}$$
де $c_{ij}$ — фіксовані цілі числа, $1 \le i \le m$, $0 \le j \le m-1$, а $t_j$ — параметри, $1 \le j \le m-1$, які пробігають множину цілих чисел.

Серед коефіцієнтів рівняння (4.1) оберемо ненульовий, який є мінімальним за модулем. Не обмежуючи загальності, можна вважати, що таким коефіцієнтом є $a_1$. Позначимо через $a'_2$, $a'_3$, $\ldots$, $a'_m$ остачі від ділення чисел $a_2$, $a_3$, $\ldots$, $a_m$ на число $a_1$ та розглянемо лінійне діофантове рівняння
$$a_1x'_1 + a'_2x'_2 + \cdots + a'_mx'_m = b. \tag{4.6}$$

Зв'язок між множинами розв'язків рівнянь (4.1) і (4.6) встановлює

**Лема 4.1.** *Множина розв'язків рівняння* (4.1) *має цілочисельну параметризацію тоді й лише тоді, коли множина розв'язків рівняння* (4.6) *має цілочисельну параметризацію.*

*Доведення.* Коефіцієнти рівнянь (4.1) і (4.6) пов'язані рівностями

$$a_i = q_i a_1 + a'_i, \quad 2 \le i \le m,$$

в яких числа $q_i$ позначають частки при діленні з остачею числа $a_i$ на число $a_1$.

Необхідність. Нехай множина розв'язків рівняння (4.1) має цілочисельну параметризацію, тобто множина розв'язків має вигляд

$$\{(f_1(t_1,\dots,t_{m-1}),\dots,f_m(t_1,\dots,t_{m-1})) \mid t_1,\dots,t_{m-1} \in \mathbb{Z}\},$$

де функції $f_i$, $1 \le i \le m$, є лінійними з цілими коефіцієнтами.

Розглянемо цілочисельний набір $(k'_1, k'_2, \dots, k'_m)$, який є розв'язком рівняння (4.6), тобто виконується рівність

$$\begin{aligned} b &= a_1k'_1 + a'_2k'_2 + \dots + a'_mk'_m = \\ &= a_1k'_1 + (a_2 - q_2a_1)k'_2 + \dots + (a_m - q_ma_1)k'_m = \\ &= a_1(k'_1 - q_2k'_2 - \dots - q_mk'_m) + a_2k'_2 + \dots + a_mk'_m. \end{aligned}$$

Звідки випливає, що $(k'_1 - q_2k'_2 - \dots - q_mk'_m, k'_2, \dots, k'_m)$ є розв'язком рівняння (4.1), а тому для деяких $t_1,\dots,t_{m-1} \in \mathbb{Z}$ виконуються рівності

$$\begin{aligned} k'_1 - q_2k'_2 - \dots - q_mk'_m &= f_1(t_1,\dots,t_{m-1}),\\ k'_2 &= f_2(t_1,\dots,t_{m-1}),\\ &\vdots\\ k'_m &= f_m(t_1,\dots,t_{m-1}), \end{aligned}$$

з яких маємо

$$k'_1 = f_1(t_1,\dots,t_{m-1}) + q_2f_2(t_1,\dots,t_{m-1}) + \dots + q_mf_m(t_1,\dots,t_{m-1}).$$

Отже, множина розв'язків рівняння (4.6) є підмножиною множини

$$\{(g(t_1,\dots,t_{m-1}), f_2(t_1,\dots,t_{m-1}),\dots,f_m(t_1,\dots,t_{m-1})) \mid t_1,\dots,t_{m-1} \in \mathbb{Z}\}, \quad (4.7)$$

де функція $g$ визначається рівністю

$$g(t_1,\dots,t_{m-1}) = f_1(t_1,\dots,t_{m-1}) + \sum_{i=2}^{m} q_i f_i(t_1,\dots,t_{m-1}),$$

яка очевидно є лінійною з цілими коефіцієнтами.

З іншого боку, для будь-яких $t_1,\dots,t_{m-1} \in \mathbb{Z}$ впорядкований набір

$$(g(t_1,\dots,t_{m-1}), f_2(t_1,\dots,t_{m-1}),\dots, f_m(t_1,\dots,t_{m-1}))$$

є розв'язком рівняння (4.6), оскільки

$$\begin{aligned} a_1 g(t_1,\dots,t_{m-1}) + \sum_{i=2}^{m} a_i' f_i(t_1,\dots,t_{m-1}) = \\ = a_1 g(t_1,\dots,t_{m-1}) + \sum_{i=2}^{m} (a_i - q_i a_1) f_i(t_1,\dots,t_{m-1}) = \\ = \sum_{i=1}^{m} a_i f_i(t_1,\dots,t_{m-1}) = b. \end{aligned}$$

Отже, множина розв'язків рівняння (4.6) збігається з множиною (4.7), а тому множина розв'язків рівняння (4.6) має цілочисельну параметризацію.

Достатність доводиться аналогічно. □

Зауважимо, що доведення леми вказує, як цілочисельну параметризацію множини розв'язків одного з рівнянь (4.1) чи (4.6) перетворити у цілочисельну параметризацію множини розв'язків іншого.

При переході від рівняння (4.1) до рівняння (4.6) сума модулів коефіцієнтів зменшується, якщо хоча б два коефіцієнти рівняння (4.1) є ненульовими. Дійсно, оскільки $a_1$ є мінімальним за модулем серед ненульових, то при $2 \le i \le m$ маємо $a_i' = a_i = 0$ або $|a_i| \ge |a_1| > |a_i'|$, де остання нерівність виконується за теоремою про ділення з остачею. Отже, застосовуючи описану процедуру, можна побудувати послідовність рівнянь, зв'язки між цілочисельними параметризаціями множин розв'язків яких відомі, причому сума модулів коефіцієнтів цих рівнянь спадає. Спадна послідовність цілих чисел, яка обмежена знизу числом 0, є скінченною, тому і побудована послідовність рівнянь теж буде скінченною. В останньому рівнянні побудованої послідовності рівно один із коефіцієнтів буде ненульовим. Не обмежуючи загальності, можна вважати,

що це рівняння буде мати вигляд

$$a_1'' x_1'' + 0x_2'' + \cdots + 0x_m'' = b. \tag{4.8}$$

**Вправа 4.1.** Доведіть, що $|a_1''| = d$.

Множина розв'язків рівняння (4.8) має очевидну цілочисельну параметризацію

$$\begin{aligned} x_1'' &= b/d, \\ x_2'' &= t_1, \\ &\vdots \\ x_m'' &= t_{m-1}, \end{aligned}$$

а тому за лемою 4.1 множина розв'язків рівняння (4.1) теж має цілочисельну параметризацію, яку можна одержати, використовуючи зв'язок між параметризаціями з доведення цієї леми.

**Приклад 4.1.** Розв'яжемо лінійне діофантове рівняння

$$6x + 10y + 15z = 1.$$

Мінімальним за модулем (серед ненульових) є коефіцієнт при $x$. Решту коефіцієнтів ділимо на 6 з остачею. Числа 10 і 15 при діленні на 6 дають частки 1 і 2 та остачі 4 і 3. Частки від ділень задають коефіцієнти заміни

$$x' = x + y + 2z,$$

після якої рівняння набуває вигляду

$$6x' + 4y + 3z = 1.$$

Наступним мінімальним за модулем є коефіцієнт при $z$. Тепер ділимо числа 6 і 4 на число 3. В результаті заміни

$$z' = 2x' + y + z$$

одержуємо рівняння

$$0x' + y + 3z' = 1,$$

в якому мінімальним за модулем серед ненульових є коефіцієнт при $y$. Після останньої заміни

$$y' = 0x' + y + 3z'$$

рівняння набуває вигляду

$$0x' + y' + 0z' = 1.$$

Оскільки єдиний ненульовий коефіцієнт ділить праву частину, число 1, то початкове рівняння має розв'язки, та для того, щоб їх знайти, починаємо з очевидної параметризації

$$x' = t_1, \quad y' = 1, \quad z' = t_2.$$

Для того, щоб знайти цілочисельну параметризацію для множини розв'язків початкового рівняння, використовуємо заміни у зворотньому порядку:

$$\begin{aligned} y &= y' - 0x' - 3z' = 1 - 3t_2; \\ z &= z' - 2x' - y = -1 - 2t_1 + 4t_2; \\ x &= x' - y - 2z = 1 + 5t_1 - 5t_2. \end{aligned}$$

Отже, загальний розв'язок початкового рівняння задається цілочисельною параметризацією

$$x = 1 + 5t_1 - 5t_2, \quad y = 1 - 3t_2, \quad z = -1 - 2t_1 + 4t_2, \qquad t_1, t_2 \in \mathbb{Z}.$$

З вправи 4.1 випливає альтернативний алгоритм обчислення найбільшого спільного дільника $m$ чисел, $m > 2$. Для цього необхідно виписати всі числа в рядок та поки не залишиться єдине ненульове число виконувати такі дії: обирати мінімальне за модулем ненульове число $a$, а решту замінювати на остачі від ділення на $a$. Єдине ненульове число, яке залишиться в кінці, є найбільшим спільним дільником початкових чисел.

## 4.2 Конгруенції

Зафіксуємо натуральне число $n$.

Кажуть, що цілі числа $a$ і $b$ *конгруентні за модулем* $n$, якщо різниця $a - b$ ділиться на $n$, та для позначення цього використовують запис $a \equiv b \pmod{n}$. Такий запис читається «$a$ конгруентне $b$ за модулем $n$» та називається *конгруенцією*.

**Теорема 4.3.** *$a \equiv b \pmod{n}$ тоді й лише тоді, коли числа $a$ і $b$ при діленні на $n$ дають однакові остачі.*

*Доведення.* Розділимо $a$ і $b$ на $n$ з остачею:

$$a = q_1 n + r_1, \quad b = q_2 n + r_2.$$

Тоді

$$a - b = (q_1 - q_2)n + (r_1 - r_2).$$

Перший доданок у правій частині ділиться на $n$. Тому подільність $a - b$ на $n$ рівносильна подільності на $n$ другого доданку $r_1 - r_2$. Але

$$-n < -r_2 \leq r_1 - r_2 \leq r_1 < n.$$

Єдине число з проміжку $(-n, n)$, яке ділиться на $n$, це 0. Отже, подільність $a - b$ на $n$ рівносильна умові $r_1 - r_2 = 0$, тобто $r_1 = r_2$. □

**Наслідок 4.1.** *Якщо остача від ділення цілого числа $a$ на $n$ дорівнює $r$, то $a \equiv r \pmod{n}$.*

Між конгруенціями та рівностями можна помітити певну аналогію. Так конгруенції з однаковим модулем можна додавати, що стверджує

**Теорема 4.4.** *Якщо $a_1 \equiv b_1 \pmod{n}$ і $a_2 \equiv b_2 \pmod{n}$, то*

$$a_1 + a_2 \equiv b_1 + b_2 \pmod{n}.$$

*Доведення.* $(a_1 + a_2) - (b_1 + b_2) = (a_1 - b_1) + (a_2 - b_2)$. Але за умовою кожний з двох доданків правої частини ділиться на $n$. Тому й ліва частина ділиться на $n$. □

**Наслідок 4.2.** *До обох частин конгруенції можна додати одне і те саме ціле число.*

*Доведення.* Додати до обох частин конгруенції $a \equiv b \pmod{n}$ число $c$ — це те саме, що додати конгруенцію $c \equiv c \pmod{n}$. □

**Наслідок 4.3.** *Будь-який з доданків можна перенести на інший бік конгруенції, змінивши знак доданку на протилежний.*

*Доведення.* Перенести в конгруенції $a + b \equiv c \pmod{n}$ на інший бік доданок $b$ — це все одно, що додати конгруенцію $-b \equiv -b \pmod{n}$. □

Конгруенції з однаковим модулем також можна множити.

**Теорема 4.5.** *Якщо* $a_1 \equiv b_1 \pmod{n}$ *і* $a_2 \equiv b_2 \pmod{n}$, *то*

$$a_1 a_2 \equiv b_1 b_2 \pmod{n}.$$

*Доведення.* Додамо і віднімемо одне і те саме число до різниці $a_1a_2 - b_1b_2$:

$$a_1a_2 - b_1b_2 = a_1a_2 - a_1b_2 + a_1b_2 - b_1b_2 = a_1(a_2 - b_2) + (a_1 - b_1)b_2.$$

Права частина ділиться на $n$, бо, за умовою, множники $a_2 - b_2$ і $a_1 - b_1$ кратні $n$. Тому й ліва частина ділиться на $n$. □

**Наслідок 4.4.** *Обидві частини конгруенції можна помножити на одне і те ж число.*

*Доведення.* Помножити обидві частини на число $c$ — це все одно, що помножити на конгруенцію $c \equiv c \pmod{n}$. □

**Наслідок 4.5.** *Обидві частини конгруенції можна піднести до одного й того ж натурального степеня* $k$.

*Доведення.* Піднести обидві частини конгруенції $a \equiv b \pmod{n}$ до степеня $k$ — це все одно, що перемножити $k$ однакових конгруенцій $a \equiv b \pmod{n}$. □

На відміну від рівностей, конгруенцію не завжди можна скорочувати на ненульове ціле число, а саме, має місце

**Теорема 4.6.** *Якщо* $ac \equiv bc \pmod{n}$ *і число* $c$ *взаємно просте з* $n$, *то* $a \equiv b \pmod{n}$.

*Доведення.* За умовою $ac - bc = (a - b)c$ ділиться на $n$. Із взаємної простоти $n$ і $c$ та твердження 2.9 випливає, що $n \mid a - b$. □

Вимога взаємної простоти чисел $n$ і $c$ є суттєвою. Наприклад, має місце конгруенція $11 \cdot 2 \equiv 5 \cdot 2 \pmod{12}$, але $11 \not\equiv 5 \pmod{12}$. Різниця чисел 11 і 5 ділиться лише на 6. Зауважимо, що рівність $6 = 12/2$ не є випадковістю. Між конгруенціями з різними модулями є зв'язки, що вказані в наступній теоремі.

**Теорема 4.7.** *Мають місце такі твердження:*

*а) Для кожного натурального числа* $k$ *конгруенції* $a \equiv b \pmod{n}$ *і* $ak \equiv bk \pmod{nk}$ *рівносильні.*

*б) Якщо* $a \equiv b \pmod{n}$ *і* $d$ *ділить* $n$*, то* $a \equiv b \pmod{d}$.

*в) Якщо* $a \equiv b \pmod{n_1}$, $\ldots$, $a \equiv b \pmod{n_k}$*, то для найменшого спільного кратного* $n$ *чисел* $n_1, \ldots, n_k$ *виконується конгруенція* $a \equiv b \pmod{n}$.

*Доведення.* Перше твердження можна переформулювати наступним чином. Число $ak - bk = (a-b)k$ ділиться на $nk$ тоді і тільки тоді, коли $a-b$ ділиться на $n$. В такому формулюванні воно стає очевидним.

В другому твердженні за умовою число $a - b$ ділиться на $n$. Тоді це число ділиться і на дільник $d$ числа $n$, що і означає, що числа $a$ і $b$ конгруентні за модулем $d$.

За умовою третього твердження число $a-b$ ділиться на кожне з чисел $n_1, \ldots, n_k$. Тому $a-b$ є їх спільним кратним та ділиться на їх найменше спільне кратне $n$, тобто числа $a$ і $b$ конгруентні за модулем $n$. □

## 4.3 Кільце класів лишків

Нехай $n$ — натуральне число. Множину усіх цілих чисел, які попарно конгруентні між собою за модулем $n$, називають *класом лишків* за модулем $n$. Іншими словами, множина $S \subset \mathbb{Z}$ є класом лишків за модулем $n$, якщо виконуються такі дві умови:

1) для довільних цілих $a, b \in S$ виконується $a \equiv b \pmod{n}$;

2) для довільних цілих $a \in S$, $b \notin S$ виконується $a \not\equiv b \pmod{n}$.

Клас лишків, який містить елемент $a$ позначають символом $\overline{a}$. Елементи, що належать класу лишків, називають *представниками* цього класу лишків. Множину всіх класів лишків за модулем $n$ позначають символом $\mathbb{Z}_n$.

**Теорема 4.8.** *Має місце рівність* $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$.

На множині $\mathbb{Z}_n$ можна визначити операції додавання та множення. А саме, *сумою* класів $\overline{a}$ і $\overline{b}$ назвемо клас $\overline{a+b}$, а їх *добутком* — клас $\overline{ab}$. Оскільки можна обирати різні представники класів, що додаються чи множаться, виникає питання коректності означення цих операцій. Потрібно переконатись, що результат операції не змінюється, якщо вибрати інші представники.

Результат додавання $\overline{a+b}$ не залежить від вибору представників $a$ і $b$ класів $\overline{a}$ і $\overline{b}$. Справді, якщо $a'$ і $b'$ — інші представники, то

$$a \equiv a' \pmod{n}, \quad b \equiv b' \pmod{n}.$$

Тоді за теоремою про додавання конгруенцій

$$a + b \equiv a' + b' \pmod{n}$$

і тому $\overline{a+b} = \overline{a'+b'}$.

Коректність означення дії множення доводиться аналогічно.

**Теорема 4.9.** *Множина $\mathbb{Z}_n$ відносно введених дій додавання і множення утворює комутативне кільце з одиницею.*

Клас лишків $\overline{a} \in \mathbb{Z}_n$ називається *оборотним*, якщо існує такий клас $\overline{b} \in \mathbb{Z}_n$, що $\overline{a} \cdot \overline{b} = \overline{1}$. Такий клас $\overline{b}$ називають *оберненим до* $\overline{a}$.

**Вправа 4.2.** Обернений до $\overline{a}$ клас лишків $\overline{b}$, якщо він існує, визначений однозначно.

Обернений до $\overline{a}$ клас лишків будемо позначати $\overline{a}^{-1}$. Наприклад, оберненим до $\overline{2}$ в кільці $\mathbb{Z}_{15}$ є клас $\overline{2}^{-1} = \overline{8}$, оскільки $\overline{2} \cdot \overline{8} = \overline{16} = \overline{1}$.

**Теорема 4.10** (Критерій оборотності в кільці класів лишків)**.** *Клас лишків $\overline{a} \in \mathbb{Z}_n$ оборотний тоді й лише тоді, коли числа $a$ і $n$ взаємно прості.*

*Доведення.* Необхідність. Нехай клас лишків $\overline{a}$ оборотний і $\overline{a} \cdot \overline{b} = \overline{1}$. Тоді $\overline{ab} = \overline{1}$ і $ab \equiv 1 \pmod{n}$, тобто $ab - 1 = kn$ для деякого цілого числа $k$. Із рівності $1 = ab - kn$ і критерію взаємної простоти випливає, що числа $a$ і $n$ взаємно прості.

Достатність. Якщо $a$ і $n$ взаємно прості, то за критерієм взаємної простоти існують такі числа $b$ і $m$, що $1 = ab + mn$. Але тоді

$$\overline{1} = \overline{ab + mn} = \overline{ab} = \overline{a} \cdot \overline{b}.$$

Отже, клас лишків $\overline{a}$ є оборотним. □

Безпосередньо з теореми випливає, що кільце $\mathbb{Z}_n$ містить рівно $\varphi(n)$ оборотних елементів. Якщо число $n$ є простим, то всі елементи, крім $\overline{0}$, є оборотними. В такому випадку кільце $\mathbb{Z}_n$ є полем.

**Твердження 4.1.** *У кільці $\mathbb{Z}_n$ добуток оборотних елементів також є оборотним елементом.*

*Доведення.* Нехай класи $\overline{a}_1, \overline{a}_2 \in \mathbb{Z}_n$ є оборотними та $\overline{a} = \overline{a}_1 \cdot \overline{a}_2$. Тоді існують такі класи $\overline{b}_1, \overline{b}_2 \in \mathbb{Z}_n$, що $\overline{a}_1 \cdot \overline{b}_1 = \overline{a}_2 \cdot \overline{b}_2 = \overline{1}$. Розглянемо клас $\overline{b} = \overline{b}_1 \cdot \overline{b}_2$. Оскільки

$$\overline{a} \cdot \overline{b} = (\overline{a}_1 \cdot \overline{a}_2) \cdot (\overline{b}_1 \cdot \overline{b}_2) = (\overline{a}_1 \cdot \overline{b}_1) \cdot (\overline{a}_2 \cdot \overline{b}_2) = \overline{1} \cdot \overline{1} = \overline{1},$$

клас $\overline{b}$ є оберненим до $\overline{a}$ та клас $\overline{a}$ є оборотним. □

Набір цілих чисел називається *повною системою лишків за модулем $n$*, якщо він містить рівно по одному представнику з кожного класу лишків за модулем $n$.

**Твердження 4.2.** *Числа $a_1$, $a_2$, …, $a_n$ утворюють повну систему лишків за модулем $n$ тоді й лише тоді, коли остачі $r_1$, $r_2$, …, $r_n$ від ділення цих чисел на $n$ попарно різні.*

**Приклад 4.2.** Набір $0, 1, 2, \ldots, n-1$ є повною системою лишків за модулем $n$. Це *система найменших невід'ємних лишків.*

**Приклад 4.3.** Набір усіх цілих чисел з $(-n/2, n/2]$ також є повною системою лишків за модулем $n$. Це *система абсолютно найменших лишків.*

Набір цілих чисел називається *зведеною системою лишків за модулем $n$*, якщо всі числа з набору взаємно прості з $n$ і він містить рівно по одному представнику з кожного оборотного класу лишків за модулем $n$.

**Твердження 4.3.** *Числа $a_1$, $a_2$, …, $a_{\varphi(n)}$ утворюють зведену систему лишків за модулем $n$ тоді й лише тоді, коли остачі $r_1$, $r_2$, …, $r_{\varphi(n)}$ від ділення цих чисел на $n$ попарно різні і взаємно прості з $n$.*

**Приклад 4.4.** Набір $1, 2, \ldots, p-1$ є зведеною системою лишків за простим модулем $p$.

## 4.4 Конгруенції з невідомими

Конгруенція вигляду

$$a_0 x^m + a_1 x^{m-1} + \cdots + a_{m-1} x + a_m \equiv 0 \pmod{n}, \tag{4.9}$$

де $a_0, a_1, \ldots, a_m \in \mathbb{Z}$, називається *конгруенцією з невідомою $x$ степеня $m$ за модулем $n$*.

Ціле число $b$ називається *частковим розв'язком конгруенції* (4.9), якщо правильною є числова конгруенція

$$a_0 b^m + a_1 b^{m-1} + \cdots + a_{m-1} b + a_m \equiv 0 \pmod{n}.$$

Конгруенція з невідомою називається *сумісною*, якщо вона має хоча б один частковий розв'язок. Інакше вона називається *несумісною*.

**Твердження 4.4.** *Ціле число $b'$, конгруентне за модулем $n$ частковому розв'язку $b$ конгруенції* (4.9)*, також буде частковим розв'язком цієї конгруенції.*

Отже, якщо число $b$ є частковим розв'язком, то всі елементи класу лишків $\bar{b}$ за модулем $n$ також є частковими розв'язками. Лишок $\bar{b}$ за модулем $n$ називається *розв'язком конгруенції* (4.9), якщо число $b$ є її частковим розв'язком. Дві конгруенції з невідомою $x$ називаються *еквівалентними*, якщо множини їх часткових розв'язків рівні. Дві конгруенції з невідомою $x$ за модулем $n$ називаються *рівносильними*, якщо множини їх розв'язків рівні.

**Твердження 4.5.** *Нехай в* (4.9) *коефіцієнти $a_0$, $a_1$, …, $a_m$ замінено конгруентними їм за модулем $n$ коефіцієнтами $a'_0$, $a'_1$, …, $a'_m$, тобто $a_0 \equiv a'_0 \pmod{n}$, $a_1 \equiv a'_1 \pmod{n}$, …, $a_m \equiv a'_m \pmod{n}$. Тоді ціле число $b$ буде розв'язком конгруенції*

$$a'_0 x^m + a'_1 x^{m-1} + \cdots + a'_{m-1} x + a'_m \equiv 0 \pmod{n}$$

*тоді й лише тоді, коли $b$ є розв'язком* (4.9).

З попереднього твердження випливає, що знайти всі розв'язки (4.9) — це те ж саме, що розв'язати над множиною $\mathbb{Z}_n$ рівняння

$$\bar{a}_0 x^m + \bar{a}_1 x^{m-1} + \cdots + \bar{a}_{m-1} x + \bar{a}_m = \bar{0}.$$

В загальній ситуації знаходження розв'язку конгруенції з невідомою може бути непростою задачею, особливо для великих модулів. На складності такої задачі, наприклад, ґрунтуються криптосистеми RSA та Рабіна. Тому ми обмежимось лише випадком лінійної конгруенції.

**Теорема 4.11.** *Лінійна конгруенція*

$$ax \equiv b \pmod{n} \tag{4.10}$$

*сумісна тоді й лише тоді, коли $b$ ділиться на $d = \mathsf{НСД}(a, n)$. Якщо конгруенція* (4.10) *сумісна, то вона має рівно $d$ розв'язків.*

*Доведення.* Нехай $x_0$ — частковий розв'язок конгруенції (4.10). Тоді число $ax_0 - b$ ділиться на $n$, тобто $ax_0 - b = kn$ для деякого цілого числа $k$. У правій частині рівності $b = ax_0 - kn$ доданки діляться на $d$, тому $b$ ділиться на $d$.

Навпаки, нехай $b$ ділиться на $d$. Тоді $b$ має вигляд $b = b_0 d$. Число $d = \mathsf{НСД}(a, n)$ можна записати у вигляді $d = ka + mn$. Тому

$$b = b_0 d = ab_0 k + b_0 mn.$$

Оскільки число $ab_0 k - b = -b_0 mn$ ділиться на $n$, то

$$ab_0 k \equiv b \pmod{n}$$

і число $b_0 k$ є частковим розв'язком (4.10). Отже, конгруенція (4.10) є сумісною.

Нехай конгруенція (4.10) сумісна. За доведеним вище число $b$ ділиться на $d$, і можна записати: $a = a_0 d$, $b = b_0 d$, $n = n_0 d$. Конгруенція (4.10) набуває вигляду

$$a_0 dx \equiv b_0 d \pmod{n_0 d}.$$

Вона еквівалентна конгруенції

$$a_0 x \equiv b_0 \pmod{n_0}. \tag{4.11}$$

Числа $a_0$ і $n_0$ взаємно прості. Тому клас лишків $\bar{a}_0 \in \mathbb{Z}_{n_0}$ є оборотним і існує таке число $c_0$, що $a_0 c_0 \equiv 1 \pmod{n_0}$. Але тоді (4.11) рівносильна конгруенції

$$x \equiv c_0 b_0 \pmod{n_0}. \tag{4.12}$$

Справді, конгруенція (4.12) одержується з (4.11) множенням обох частин на $c_0$, а конгруенція (4.11) одержується з (4.12) множенням обох частин на $a_0$. Залишилось показати, що клас лишків $\overline{c_0 b_0} \in \mathbb{Z}_{n_0}$ при переході до модуля $n$ розпадається на $d$ класів.

Позначимо $r$ остачу від ділення $c_0 b_0$ на $n_0$. Тоді $0 \leq r < n_0$. Клас лишків $\bar{t} \in \mathbb{Z}_n$ буде розв'язком конгруенції (4.10) тоді й лише тоді, коли $t$

буде частковим розв'язком конгруенції (4.11), або, що те саме, частковим розв'язком рівносильної конгруенції (4.12). Останнє еквівалентне тому, що $t$ при діленні на $n_0$ дає в остачі $r$. Число $t$ можна вибрати з множини

$$\{0, 1, 2, \ldots, n-1 = n_0 d - 1\}.$$

Але з цієї множини остачу $r$ при діленні на $n_0$ дають лише числа

$$r, n_0 + r, 2n_0 + r, \ldots, (d-1)n_0 + r.$$

Отже, конгруенція (4.10) має рівно $d$ розв'язків:

$$\overline{r}, \overline{n_0 + r}, \overline{2n_0 + r}, \ldots, \overline{(d-1)n_0 + r} \in \mathbb{Z}_n. \qquad \square$$

Правильність конгруенції $ax_0 \equiv b \pmod{n}$ рівносильна існуванню такого цілого числа $k$, що $ax_0 - kn = b$. На пару $(x_0, -k)$ можна дивитись як на розв'язок лінійного діофантового рівняння

$$ax + ny = b. \tag{4.13}$$

Навпаки, кожен розв'язок $(x_0, y_0)$ діофантового рівняння (4.13) дає частковий розв'язок $x_0$ конгруенції (4.10), бо $ax_0 - b = -ny_0$ ділиться на $n$. Тому конгруенцію (4.10) можна розглядати як діофантове рівняння (4.13), в якому нас цікавить лише перша компонента розв'язку. Цим пояснюється той факт, що критерій сумісності лінійної конгруенції (4.10) збігається з критерієм сумісності лінійного діофантового рівняння (4.13). Для розв'язування конгруенції (4.10) переходимо до лінійного діофантового рівняння (4.13) і розв'язуємо його.

В прикладних задачах виникає потреба, щоб невідоме ціле число задовольняло кілька конгруенцій з невідомою одночасно. Таким чином виникають системи конгруенцій. Оскільки вже наявність однієї конгруенції з невідомою, степінь якої більше 1, суттєво ускладнює задачу, будемо розглядати лише системи лінійних конгруенцій від однієї невідомої, тобто системи вигляду

$$\begin{cases} a_1 x \equiv b_1 \pmod{n_1}; \\ a_2 x \equiv b_2 \pmod{n_2}; \\ \quad \vdots \\ a_m x \equiv b_m \pmod{n_m}. \end{cases} \tag{4.14}$$

Якщо якась із конгруенцій

$$a_k x \equiv b_k \pmod{n_k} \tag{4.15}$$

цієї системи несумісна, то і вся система несумісна. Якщо ж конгруенція (4.15) сумісна, то $b_k$ ділиться на найбільший спільний дільник $d_k$ чисел $a_k$ і $n_k$, і можна її замінити рівносильною їй конгруенцією

$$a'_k x \equiv b'_k \pmod{n'_k},$$

де $a'_k = a_k/d_k$, $b'_k = b_k/d_k$, $n'_k = n_k/d_k$. Зауважимо, що числа $a'_k$ і $n'_k$ взаємно прості.

Таким чином, систему (4.14) можна замінити рівносильною їй системою

$$\begin{cases} a'_1 x \equiv b'_1 \pmod{n'_1}; \\ a'_2 x \equiv b'_2 \pmod{n'_2}; \\ \quad\vdots \\ a'_m x \equiv b'_m \pmod{n'_m}. \end{cases} \tag{4.16}$$

Кожний коефіцієнт $a'_k$ взаємно простий з відповідним модулем $n'_k$. Тому для кожного $k$ існує таке число $c_k$, що $a'_k \cdot c_k \equiv 1 \pmod{n'_k}$. Домноживши кожну з конгруенцій системи (4.16) на відповідний множник $c_k$, одержимо рівносильну систему

$$\begin{cases} x \equiv b'_1 c_1 \pmod{n'_1}; \\ x \equiv b'_2 c_2 \pmod{n'_2}; \\ \quad\vdots \\ x \equiv b'_m c_m \pmod{n'_m}. \end{cases}$$

Тому в подальшому обмежимось лише системами вигляду

$$\begin{cases} x \equiv b_1 \pmod{n_1}; \\ x \equiv b_2 \pmod{n_2}; \\ \quad\vdots \\ x \equiv b_m \pmod{n_m}. \end{cases} \tag{4.17}$$

**Теорема 4.12.** *Якщо система* (4.17) *сумісна, то її розв'язки утворюють клас лишків за модулем числа* $\mathsf{HCK}(n_1, n_2, \ldots, n_m)$.

*Доведення.* Розглянемо спочатку випадок $m = 2$. Система (4.17) набуває вигляду

$$\begin{cases} x \equiv b_1 \pmod{n_1}; \\ x \equiv b_2 \pmod{n_2}. \end{cases}$$

Припустимо, що вона є сумісною, а $x_0$ — її частковий розв'язок. Позначимо $n = \mathsf{HCK}(n_1, n_2)$.

Кожен елемент $x_1$ із класу лишків $\bar{x}_0 \in \mathbb{Z}_n$ також буде її розв'язком. Справді, $x_1$ має вигляд $x_1 = x_0 + kn$. Тому число

$$x_1 - b_1 = (x_0 + kn) - b_1 = (x_0 - b_1) + kn$$

ділиться на $n_1$, бо кожен із двох доданків правої частини ділиться на $n_1$. Отже, $x_1$ задовольняє першу конгруенцію. Аналогічно доводиться, що $x_1$ задовольняє і другу конгруенцію.

Навпаки, якщо $x_2$ — інший частковий розв'язок системи, то число $x_2 - x_0 = (x_2 - b_1) - (x_0 - b_1)$ ділиться на $n_1$, бо кожний з двох доданків правої частини ділиться на $n_1$. Аналогічно доводиться, що $x_2 - x_0$ ділиться на $n_2$. Тому $x_2 - x_0$ є спільним кратним чисел $n_1$ і $n_2$ і тому ділиться на їх найменше спільне кратне $n$. Отже, виконується $x_2 \equiv x_0 \pmod{n}$, тобто $x_2 \in \bar{x}_0$.

Таким чином, розв'язки системи (4.17) у випадку $m = 2$ утворюють клас лишків $\bar{x}_0 \in \mathbb{Z}_n$. Для $m = 2$ твердження теореми доведено.

Розглянемо загальний випадок. Скористаємось методом математичної індукції за $m$. Для $m = 1$ і $m = 2$ твердження правильне. Розглянемо $m \geq 3$ і припустимо, що для всіх систем меншого розміру теорема доведена. Нехай $n' = \mathsf{HCK}(n_1, n_2, \ldots, n_{m-1})$. Тоді, за припущенням індукції, система

$$\begin{cases} x \equiv b_1 \pmod{n_1}; \\ x \equiv b_2 \pmod{n_2}; \\ \quad \vdots \\ x \equiv b_{m-1} \pmod{n_{m-1}} \end{cases}$$

рівносильна одній конгруенції вигляду

$$x \equiv b' \pmod{n'}.$$

Тоді початкова система (4.17) рівносильна системі

$$\begin{cases} x \equiv b' \pmod{n'}; \\ x \equiv b_m \pmod{n_m}. \end{cases}$$

Як уже доведено у випадку $m = 2$, ця система рівносильна конгруенції вигляду

$$x \equiv b \pmod{n},$$

де

$$n = \mathsf{НСК}(n', n_m) = \mathsf{НСК}(\mathsf{НСК}(n_1, n_2, \ldots, n_{m-1}), n_m) = \\ = \mathsf{НСК}(n_1, n_2, \ldots, n_{m-1}, n_m).$$

Отже, розв'язки системи (4.17) утворюють один клас лишків за модулем числа $\mathsf{НСК}(n_1, n_2, \ldots, n_m)$, що й потрібно було довести. □

Вкажемо процедуру, за допомогою якої можна знайти розв'язок системи конгруенцій (4.17). Для цього знадобиться

**Лема 4.2.** *Нехай $n_1$ і $n_2$ взаємно прості і $1 = k_1 n_1 + k_2 n_2$. Тоді число $k_2 n_2$ буде частковим розв'язком системи*

$$\begin{cases} x \equiv 1 \pmod{n_1}; \\ x \equiv 0 \pmod{n_2}. \end{cases}$$

*Доведення.* Із рівності $k_2 n_2 - 1 = -k_1 n_1$ випливає, що число $k_2 n_2$ задовольняє першій конгруенції. Перевірка другої конгруенції тривіальна. □

Нехай числа $n_1$, $n_2$, …, $n_m$ попарно взаємно прості. Тоді для кожного $k$, $1 \le k \le m$, числа $n_k$ і $t_k = n_1 n_2 \ldots n_m / n_k$ взаємно прості. Позначимо через $B_k$ розв'язок системи конгруенцій

$$\begin{cases} x \equiv 1 \pmod{n_k}; \\ x \equiv 0 \pmod{t_k}; \end{cases}$$

який будується за допомогою попередньої леми.

**Теорема 4.13** (Китайська теорема про остачі)**.** *Якщо модулі $n_1$, $n_2$, …, $n_m$ попарно взаємно прості, то множина розв'язків системи конгруенцій* (4.17) *збігається з класом лишків*

$$\overline{b_1 B_1 + b_2 B_2 + \cdots + b_m B_m} \in \mathbb{Z}_{n_1 n_2 \ldots n_m}.$$

*Зокрема, система вигляду* (4.17) *із попарно взаємно простими модулями завжди сумісна.*

*Доведення.* Зауважимо, що $t_k = n_1 n_2 \dots n_m / n_k$ ділиться на кожне з чисел $n_1, \dots, n_{k-1}, n_{k+1}, \dots, n_m$. Тому з конгруенції $B_k \equiv 0 \pmod{t_k}$ випливає, що для кожного $i \neq k$ виконується конгруенція $B_k \equiv 0 \pmod{n_i}$. Крім того, $B_k \equiv 1 \pmod{n_k}$. Але тоді для кожного $k$, $1 \leq k \leq m$,

$$b_1 B_1 + \dots + b_{k-1} B_{k-1} + b_k B_k + b_{k+1} B_{k+1} + \dots + b_m B_m \equiv$$
$$\equiv b_1 \cdot 0 + \dots + b_{k-1} \cdot 0 + b_k \cdot 1 + b_{k+1} \cdot 0 + \dots + b_m \cdot 0 \equiv b_k \pmod{n_k}.$$

Отже, система (4.17) сумісна і число

$$b_1 B_1 + b_2 B_2 + \dots + b_m B_m$$

є її розв'язком. Але множина розв'язків системи (4.17) утворює клас лишків за модулем числа $\mathsf{HCK}(n_1, n_2, \dots, n_m)$. Оскільки числа $n_1$, $n_2$, $\dots$, $n_m$ попарно взаємно прості, то $\mathsf{HCK}(n_1, n_2, \dots, n_m) = n_1 n_2 \dots n_m$. Таким чином, множиною розв'язків системи (4.17) є клас лишків

$$\overline{b_1 B_1 + b_2 B_2 + \dots + b_m B_m} \in \mathbb{Z}_{n_1 n_2 \dots n_m}.$$

Теорему доведено. □

Якщо числа $n_1, n_2, \dots, n_m$ не є попарно взаємно простими, то можна перейти до системи конгруенцій еквівалентної системі (4.17), за умови її сумісності, в якій модулі вже будуть попарно взаємно простими. Досягти цього можна, наприклад, наступним чином. Спочатку кожну конгруенцію системи виду

$$x \equiv c \pmod{n}$$

замінити на кілька конгруенцій

$$\begin{aligned} x &\equiv c \pmod{p_1^{\alpha_1}}, \\ x &\equiv c \pmod{p_2^{\alpha_2}}, \\ &\vdots \\ x &\equiv c \pmod{p_m^{\alpha_m}}, \end{aligned}$$

модулі яких визначаються канонічним розкладом $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m}$. Така заміна є еквівалентним перетворенням за теоремою 4.7. Після чого, для кожного простого числа $p$ розглянути всі конгруенції виду

$$\begin{aligned} x &\equiv c_1 \pmod{p^{\beta_1}}, \\ x &\equiv c_2 \pmod{p^{\beta_2}}, \\ &\vdots \\ x &\equiv c_k \pmod{p^{\beta_k}}, \end{aligned}$$

де $\beta_1 \geq \beta_2 \geq \ldots \geq \beta_k$, які будуть сумісними тоді й лише тоді, коли

$$c_1 \equiv c_i \pmod{p^{\beta_i}}, \quad 2 \leq i \leq k. \tag{4.18}$$

Якщо умови (4.18) не виконуються, то початкова система конгруенцій несумісна, тобто не має розв'язків, якщо ж умови (4.18) виконуються, то можна залишити лише одну конгруенцію

$$x \equiv c_1 \pmod{p^{\beta_1}}.$$

Нова система конгруенцій за теоремою 4.7 буде еквівалентна початковій. В результаті таких перетворень система буде містити конгруенції за модулями, що є степенями попарно різних простих чисел, а отже, модулі будуть попарно взаємно простими.

## 4.5 Вправи

1. Розв'яжіть лінійне діофантове рівняння

   а) $15x - 39y = 55$;
   б) $187x - 143y = 77$;
   в) $35x + 55y + 77z = 1$.

2. Знайдіть остачу від ділення числа $5 \cdot 11 \cdot 19 \cdot 29 \cdot 101 \cdot 197$ на 13.

3. Знайдіть дві останні цифри числа $3^{100}$.

4. Розв'яжіть лінійні конгруенції з невідомою:

   а) $128x \equiv 238 \pmod{94}$;
   б) $59x \equiv 17 \pmod{71}$.

5. Розв'яжіть систему лінійних конгруенцій з невідомою

$$\begin{cases} x \equiv 13 \pmod{21}; \\ x \equiv 7 \pmod{22}; \\ x \equiv 9 \pmod{23}. \end{cases}$$

6. Розв'яжіть систему лінійних конгруенцій з невідомою

$$\begin{cases} 15x \equiv 24 \pmod{51}; \\ 28x \equiv 24 \pmod{72}; \\ 30x \equiv 24 \pmod{46}. \end{cases}$$

7. Для яких значень параметра $a$ система конгруенцій з невідомою

$$\begin{cases} x \equiv 5 \pmod 6; \\ x \equiv 8 \pmod{15}; \\ x \equiv a \pmod{10} \end{cases}$$

буде сумісною?

8. Знайдіть найменше натуральне число, яке при діленні на $m$, $n$ і $k$ дає відповідно остачі $a$, $b$ і $c$, якщо

$$m = 12, \quad n = 13, \quad k = 14, \quad a = 5, \quad b = 6, \quad c = 7.$$

# Розділ 5

# Алгебра матриць

Алгебра матриць вивчає операції над прямокутними таблицями чисел. Як і в арифметиці, матриці можна додавати та множити, однак ці операції утворюють значно складнішу алгебраїчну структуру. Матрична арифметика має суттєві особливості: результат множення матриць залежить від порядку множників, а добуток двох ненульових матриць може дорівнювати нулю. Це потребувало розробки специфічного математичного апарату, який ми будемо вивчати в цьому розділі.

## 5.1 Поняття матриці

Матриці є одним з базових інструментів лінійної алгебри. Вони використовуються для компактного запису систем лінійних рівнянь, опису лінійних перетворень, реалізації чисельних алгоритмів тощо.

**Означення 5.1.** *Матрицею розміру* $m \times n$ називається прямокутна таблиця чисел, що складається з $m$ рядків та $n$ стовпців і має вигляд:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Числа $a_{ij}$, що входять до складу матриці, називаються її *елементами*. Індекси $i$ та $j$ вказують на розташування елемента: число $a_{ij}$ знаходиться у $i$-му рядку та $j$-му стовпці. Для скороченого позначення матриці

використовують символ $A = (a_{ij})_{i=1,j=1}^{m,n}$ або просто $A = (a_{ij})$, якщо розміри матриці є відомими.

Матриці $A = (a_{ij})$ і $B = (b_{ij})$ називаються *рівними*, якщо вони мають однакову розмірність і $a_{ij} = b_{ij}$ для всіх $i, j$.

Множина всіх матриць розміру $m \times n$, коефіцієнти яких належить множині $F$, позначається через $M_{m\times n}(F)$. У випадку, коли кількість рядків матриці збігається з кількістю її стовпців ($m = n$), матрицю називають *квадратною*, а число $n$ — її *порядком*. Множина всіх квадратних матриць порядку $n$, коефіцієнти яких належить множині $F$, позначається через $M_n(F)$.

**Приклад 5.1.** Розглянемо матриці

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 7 \end{pmatrix} \quad \text{та} \quad B = \begin{pmatrix} -4 & 5 \\ 7 & -3 \end{pmatrix}.$$

Матриця $A$ має розмір $2 \times 3$ (два рядки та три стовпці). Її елементами, зокрема, є $a_{12} = 2$ та $a_{23} = 7$. Матриця $B$ є квадратною порядку 2; її елементами, наприклад, є $b_{11} = -4$ та $b_{21} = 7$.

Серед усього різноманіття матриць виділяють певні типи, що мають спеціальну структуру та властивості. Елементи $a_{11}$, $a_{22}$, …, $a_{nn}$ квадратної матриці порядку $n$ утворюють її *головну діагональ*. Квадратна матриця називається *діагональною*, якщо всі її елементи, що лежать поза головною діагоналлю, є нульовими (тобто $a_{ij} = 0$ при $i \neq j$):

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Діагональна матриця, у якої всі елементи на головній діагоналі однакові, називається *скалярною*:

$$C = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & a & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a \end{pmatrix}.$$

Якщо ж усі діагональні елементи такої матриці дорівнюють одиниці, то вона називається *одиничною* і позначається $E$ або $E_n$, якщо потрібно вказати порядок матриці. Квадратна матриця порядку $n$, у якої всі елементи дорівнюють нулю, називається *нульовою* і позначається $O_n$ або $O$, якщо розмір матриці є зрозумілим із контексту:

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \qquad O = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Квадратна матриця називається *верхньою трикутною*, якщо $a_{ij} = 0$ при $i > j$, та *нижньою трикутною*, якщо $a_{ij} = 0$ при $i < j$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

**Приклад 5.2.** Розглянемо матриці

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix},$$
$$B = \begin{pmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 1 & -3 \\ 0 & -7 & 2 \\ 0 & 0 & 5 \end{pmatrix}.$$

Матриця $E_3$ є одиничною порядку 3, матриця $A$ є діагональною, матриця $B$ є скалярною, а матриця $C$ є верхньою трикутною.

## 5.2 Операції над матрицями

У цьому розділі будемо розглядати матриці з елементами з деякого поля $F$ (наприклад, $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ або скінченне поле $\mathbb{Z}_p$, де $p$ — просте число).

**Означення 5.2.** *Сумою матриць* $A = (a_{ij})$ і $B = (b_{ij})$ однакового розміру $m \times n$ називається матриця $C = (c_{ij})$ того самого розміру, елементи якої

обчислюються за правилом:

$$c_{ij} = a_{ij} + b_{ij}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

*Різницею матриць* $A = (a_{ij})$ і $B = (b_{ij})$ називається матриця $D = (d_{ij})$, елементи якої обчислюються за правилом:

$$d_{ij} = a_{ij} - b_{ij}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Операції додавання та віднімання не визначаються для матриць різних розмірів.

**Означення 5.3.** *Добутком матриці* $A = (a_{ij})$ *на число* $\lambda \in F$ називається матриця

$$\lambda A = (\lambda a_{ij}), \quad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Матриця $(-1)A = (-a_{ij})$ називається *протилежною* до матриці $A$ і позначається $-A$. Це дозволяє записати різницю матриць як суму: $A - B = A + (-B)$.

**Приклад 5.3.** Розглянемо матриці

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 1 & -3 \\ 5 & -2 & 4 \end{pmatrix}.$$

Знайдемо суму та різницю цих матриць:

$$A + B = \begin{pmatrix} 1+2 & -2+1 & 3+(-3) \\ 0+5 & 4+(-2) & -1+4 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ 5 & 2 & 3 \end{pmatrix},$$

$$A - B = \begin{pmatrix} 1-2 & -2-1 & 3-(-3) \\ 0-5 & 4-(-2) & -1-4 \end{pmatrix} = \begin{pmatrix} -1 & -3 & 6 \\ -5 & 6 & -5 \end{pmatrix}.$$

Помножимо матрицю $A$ на скаляр $\lambda = 2$:

$$2A = \begin{pmatrix} 2 & -4 & 6 \\ 0 & 8 & -2 \end{pmatrix}.$$

Обчислимо лінійну комбінацію матриць:

$$2A - 3B = \begin{pmatrix} 2 & -4 & 6 \\ 0 & 8 & -2 \end{pmatrix} - \begin{pmatrix} 6 & 3 & -9 \\ 15 & -6 & 12 \end{pmatrix} = \begin{pmatrix} -4 & -7 & 15 \\ -15 & 14 & -14 \end{pmatrix}.$$

Операції додавання матриць і множення матриці на число мають властивості, аналогічні відповідним властивостям для чисел.

**Твердження 5.1** (властивості додавання матриць і множення матриць на скаляри)**.** *Нехай $A, B, C \in M_{m\times n}(F)$ і $\lambda, \mu \in F$. Тоді виконуються такі властивості:*

1. $A + B = B + A$ *(комутативність додавання);*
2. $(A + B) + C = A + (B + C)$ *(асоціативність додавання);*
3. $A + O = A$;
4. $A + (-A) = O$;
5. $\lambda(A + B) = \lambda A + \lambda B$;
6. $(\lambda + \mu)A = \lambda A + \mu A$;
7. $\lambda(\mu A) = (\lambda\mu)A$;
8. $1 \cdot A = A$.

*Доведення.* Оскільки операції додавання матриць та множення матриці на скаляр визначені поелементно, усі перелічені властивості зводяться до відповідних властивостей операцій у полі $F$. □

Добуток двох матриць $A$ і $B$ визначається тільки тоді, коли кількість стовпців матриці $A$ дорівнює кількості рядків матриці $B$.

**Означення 5.4.** *Добутком матриць* $A = (a_{ij}) \in M_{m\times k}(F)$ та $B = (b_{ij}) \in M_{k\times n}(F)$ називається матриця $C = (c_{ij}) \in M_{m\times n}(F)$, де

$$c_{ij} = \sum_{r=1}^{k} a_{ir}b_{rj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj}.$$

Матриця $C = AB$ має розмір $m \times n$, де $m$ — це кількість рядків матриці $A$, а $n$ — кількість стовпців матриці $B$. Елемент $c_{ij}$, що стоїть на перетині $i$-го рядка та $j$-го стовпця, утворюється як скалярний добуток $i$-го рядка матриці $A$ та $j$-го стовпця матриці $B$.

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ik} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mk} \end{pmatrix} \begin{pmatrix} b_{11} & \ldots & b_{1j} & \ldots & b_{1n} \\ b_{21} & \ldots & b_{2j} & \ldots & b_{2n} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & \ldots & b_{kj} & \ldots & b_{kn} \end{pmatrix} = \begin{pmatrix} c_{11} & \ldots & c_{1j} & \ldots & c_{1n} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \ldots & c_{ij} & \ldots & c_{in} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \ldots & c_{mi} & \ldots & c_{mn} \end{pmatrix}$$

**Приклад 5.4.** Розглянемо матриці

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & -5 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 3 \\ -1 & 2 \\ 5 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}.$$

Обчислимо добутки:

$$AB = \begin{pmatrix} 1\cdot 2+(-1)\cdot(-1)+2\cdot 5 & 1\cdot 3+(-1)\cdot 2+2\cdot(-4) \\ 3\cdot 2+4\cdot(-1)+(-5)\cdot 5 & 3\cdot 3+4\cdot 2+(-5)\cdot(-4) \end{pmatrix} = \begin{pmatrix} 13 & -7 \\ -23 & 37 \end{pmatrix},$$

$$BC = \begin{pmatrix} 2\cdot 1+3\cdot 3 & 2\cdot 2+3\cdot(-4) \\ -1\cdot 1+2\cdot 3 & -1\cdot 2+2\cdot(-4) \\ 5\cdot 1+(-4)\cdot 3 & 5\cdot 2+(-4)\cdot(-4) \end{pmatrix} = \begin{pmatrix} 11 & -8 \\ 5 & -10 \\ -7 & 26 \end{pmatrix},$$

$$C^2 = \begin{pmatrix} 1\cdot 1+2\cdot 3 & 1\cdot 2+2\cdot(-4) \\ 3\cdot 1+(-4)\cdot 3 & 3\cdot 2+(-4)\cdot(-4) \end{pmatrix} = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix}.$$

У той же час деякі операції над вказаними матрицями неможливо виконати через невідповідність їхніх розмірів. Зокрема, добуток $AC$ не визначений, оскільки кількість стовпців матриці $A$ не збігається з кількістю рядків матриці $C$. Так само не визначено добуток $B^2$, оскільки матриця $B$ не є квадратною.

**Приклад 5.5.** Розглянемо добутки матриці-рядка та матриці-стовпця:

$$\begin{pmatrix} 1 & -2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1\cdot 4+(-2)\cdot 5+3\cdot 6 \end{pmatrix} = \begin{pmatrix} 12 \end{pmatrix},$$

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \begin{pmatrix} 1 & -2 & 3 \end{pmatrix} = \begin{pmatrix} 4\cdot 1 & 4\cdot(-2) & 4\cdot 3 \\ 5\cdot 1 & 5\cdot(-2) & 5\cdot 3 \\ 6\cdot 1 & 6\cdot(-2) & 6\cdot 3 \end{pmatrix} = \begin{pmatrix} 4 & -8 & 12 \\ 5 & -10 & 15 \\ 6 & -12 & 18 \end{pmatrix}.$$

**Твердження 5.2** (властивості добутку матриць)**.** *Нехай $A$, $B$, $C$ — матриці над полем $F$, розміри яких дозволяють виконувати вказані операції, та $\lambda \in F$. Тоді:*

1. $(AB)C = A(BC)$ *(асоціативність);*
2. $A(B+C) = AB + AC$ *(ліва дистрибутивність);*

3. $(A+B)C = AC + BC$ *(права дистрибутивність);*

4. $\lambda(AB) = (\lambda A)B = A(\lambda B)$*;*

5. $E_m A = AE_n = A$ *для будь-якої матриці* $A \in M_{m\times n}(F)$.

*Доведення.* Доведемо найбільш трудомістку властивість — асоціативність. Решта властивостей доводяться аналогічно через розкриття сум за визначенням добутку матриць.

Нехай $A \in M_{m\times k}(F)$, $B \in M_{k\times p}(F)$, $C \in M_{p\times n}(F)$. Розглянемо елемент на позиції $(i,j)$ матриці $(AB)C$. За визначенням добутку:

$$((AB)C)_{ij} = \sum_{s=1}^{p} (AB)_{is} c_{sj} = \sum_{s=1}^{p} \left( \sum_{r=1}^{k} a_{ir} b_{rs} \right) c_{sj}.$$

Використовуючи дистрибутивність та асоціативність множення у полі $F$, ми можемо змінити порядок підсумовування:

$$((AB)C)_{ij} = \sum_{s=1}^{p} \sum_{r=1}^{k} (a_{ir} b_{rs}) c_{sj} = \sum_{r=1}^{k} a_{ir} \left( \sum_{s=1}^{p} b_{rs} c_{sj} \right).$$

Вираз у дужках є елементом $(BC)_{rj}$ матриці $BC$. Отже:

$$((AB)C)_{ij} = \sum_{r=1}^{k} a_{ir} (BC)_{rj} = (A(BC))_{ij}.$$

Оскільки всі відповідні елементи матриць рівні, то $(AB)C = A(BC)$. □

**Приклад 5.6.** Добуток матриць не є комутативним, тобто для деяких матриць $A$ і $B$ виконується $AB \neq BA$. Наприклад,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Кажуть, що матриці $A$ і $B$ *комутують*, якщо $AB = BA$.

**Зауваження 5.1.** Матриці можна розглядати не лише як таблиці чисел, а й як відображення між просторами. Кожна матриця $A \in M_{m\times n}(F)$ визначає відображення $f_A \colon F^n \to F^m$, яке діє на вектор-стовпець $\boldsymbol{x} \in F^n$ за правилом

$$f_A(\boldsymbol{x}) = A\boldsymbol{x}.$$

З властивостей множення матриці на вектор випливає, що це відображення є лінійним, тобто

$$A(\boldsymbol{x}+\boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y}, \qquad A(\lambda\boldsymbol{x}) = \lambda(A\boldsymbol{x}).$$

З цієї точки зору множення матриць відповідає композиції відповідних лінійних відображень. Нехай $A \in M_{m\times k}(F)$ і $B \in M_{k\times n}(F)$. Тоді $f_A\colon F^k \to F^m$ та $f_B\colon F^n \to F^k$, і для будь-якого $\boldsymbol{x} \in F^n$ маємо

$$f_A \circ f_B(\boldsymbol{x}) = f_A(f_B(\boldsymbol{x})) = A(B\boldsymbol{x}) = (AB)\boldsymbol{x}.$$

Справді, якщо $\boldsymbol{x} = (x_1, \ldots, x_n)$, то

$$B\boldsymbol{x} = \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{k1} & \ldots & b_{kn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n b_{1j}x_j \\ \vdots \\ \sum_{j=1}^n b_{kj}x_j \end{pmatrix},$$

$$(A(B\boldsymbol{x}))_i = \sum_{r=1}^k a_{ir}(B\boldsymbol{x})_r = \sum_{r=1}^k a_{ir}\left(\sum_{j=1}^n b_{rj}x_j\right) = \sum_{j=1}^n \left(\sum_{r=1}^k a_{ir}b_{rj}\right) x_j.$$

Отже, композиція відображень $f_A \circ f_B$ задається матрицею $C$, елементи якої визначаються формулою, такою самою, як і при множенні матриць $AB$:

$$c_{ij} = \sum_{r=1}^k a_{ir}b_{rj}.$$

**Приклад 5.7.** Розглянемо добуток матриці на матрицю-стовпець:

$$\begin{pmatrix} 1 & 2 & -3 \\ -2 & 0 & 4 \\ 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1\cdot 4 + 2\cdot 3 + (-3)\cdot 2 \\ (-2)\cdot 4 + 0\cdot 3 + 4\cdot 2 \\ 3\cdot 4 + 1\cdot 3 + 5\cdot 2 \end{pmatrix} =$$
$$= 4\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + 2\begin{pmatrix} -3 \\ 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 25 \end{pmatrix}.$$

Зверніть увагу, що результат множення матриці на стовпець є лінійною комбінацією стовпців першої матриці з коефіцієнтами з другого

стовпця. У загальному випадку:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = b_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + b_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Звідси випливає наступна властивість добутку матриць.

**Твердження 5.3.** *Нехай $A \in M_{m\times k}(F)$ і $B \in M_{k\times n}(F)$, причому*

$$B = [\boldsymbol{b}_1\ \boldsymbol{b}_2\ \dots\ \boldsymbol{b}_n],$$

*де $\boldsymbol{b}_i$ — стовпці матриці $B$. Тоді добуток $AB$ можна записати через стовпці матриці $B$ як*

$$AB = [A\boldsymbol{b}_1\ A\boldsymbol{b}_2\ \dots\ A\boldsymbol{b}_n].$$

*Зокрема, кожний стовпець матриці $AB$ є лінійною комбінацією стовпців матриці $A$.*

Квадратну матрицю можна множити на себе.

**Означення 5.5.** Нехай $A \in M_n(F)$ — квадратна матриця. *Степенем* матриці $A$ з цілим невід'ємним показником $k$ називають добуток $k$ копій матриці $A$:

$$A^0 := E_n, \qquad A^k := \underbrace{A \cdot A \cdot \ldots \cdot A}_{k \text{ разів}}, \quad k \geq 1.$$

Для степенів квадратних матриць виконуються стандартні властивості:

$$A^k \cdot A^l = A^{k+l}, \quad (A^k)^l = A^{kl}, \qquad k, l \geq 0.$$

**Означення 5.6.** Нехай $A \in M_n(F)$ і

$$p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m \in F[x]$$

— многочлен над полем $F$. Тоді *многочлен від матриці $A$* визначається як матриця

$$p(A) = a_0 E_n + a_1 A + a_2 A^2 + \dots + a_m A^m.$$

Для многочленів від матриці виконуються стандартні властивості: для $p(x), q(x) \in F[x]$ і $A \in M_n(F)$ маємо

$$(p+q)(A) = p(A) + q(A) \qquad \text{та} \qquad (p \cdot q)(A) = p(A)\, q(A).$$

**Приклад 5.8.** Нехай

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}, \quad p(x) = 3 + 5x - 2x^2.$$

Тоді:

$$p(A) = 3E_2 + 5A - 2A^2 = 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 5\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} - 2\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}^2 =$$
$$= 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 5\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} - 2\begin{pmatrix} 3 & -5 \\ 5 & 8 \end{pmatrix} = \begin{pmatrix} 7 & 5 \\ -5 & 2 \end{pmatrix}.$$

**Означення 5.7.** *Транспонованою* до матриці $A = (a_{ij}) \in M_{m\times n}(F)$ називається матриця $A^T = (b_{ij}) \in M_{n\times m}(F)$, елементи якої визначаються так:

$$b_{ij} = a_{ji}, \quad i = 1, \ldots, n, \ j = 1, \ldots, m.$$

Тобто рядки матриці $A$ стають стовпцями матриці $A^T$.

**Приклад 5.9.** Розглянемо матриці та їх транспоновані:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \ A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -2 \\ -5 & 4 \end{pmatrix}, \ B^T = \begin{pmatrix} 3 & -5 \\ -2 & 4 \end{pmatrix}.$$

**Твердження 5.4** (властивості транспонування матриць)**.** *Нехай $A, B \in M_{m\times n}(F)$ та $\lambda \in F$. Тоді виконуються такі властивості:*

1. $(A^T)^T = A$;
2. $(A + B)^T = A^T + B^T$;
3. $(\lambda A)^T = \lambda A^T$;
4. *Якщо $A \in M_{m\times k}(F)$ та $B \in M_{k\times n}(F)$, то $(AB)^T = B^T A^T$.*

*Доведення.* Властивості 1–3 виконуються поелементно за означенням транспонування. Доведемо пункт 4. Маємо:

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{r=1}^{k} a_{jr}b_{ri} = \sum_{r=1}^{k} (B^T)_{ir}(A^T)_{rj} = (B^T A^T)_{ij}.$$

Отже, $(AB)^T = B^T A^T$. □

**Приклад 5.10.** Зверніть увагу, що $(AB)^T$ може не дорівнювати $A^T B^T$:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix},$$
$$A^T B^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 11 \\ 4 & 16 \end{pmatrix},$$
$$(AB)^T = \begin{pmatrix} 4 & 7 \\ 8 & 15 \end{pmatrix}^T = \begin{pmatrix} 4 & 8 \\ 7 & 15 \end{pmatrix}.$$

**Означення 5.8.** Квадратна матриця $A$ називається *симетричною*, якщо $A^T = A$, тобто $a_{ij} = a_{ji}$ для всіх $i, j$. Інакше кажучи, елементи симетричної матриці симетричні відносно головної діагоналі.

Квадратна матриця $A$ називається *кососиметричною*, якщо $A^T = -A$, тобто $a_{ij} = -a_{ji}$ для всіх $i, j$.

**Приклад 5.11.** Розглянемо кілька матриць:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 4 \\ -1 & 4 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 2 & -3 \\ -2 & 0 & 1 \\ 3 & -1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Матриця $A$ є симетричною, матриця $B$ є кососиметричною, а матриця $C$ не є ні симетричною, ні кососиметричною.

**Твердження 5.5** (властивості симетричних матриць)**.** *Нехай $A, B \in M_n(F)$ — симетричні матриці та $\lambda \in F$. Тоді:*

1. *Сума $A + B$ та добуток на скаляр $\lambda A$ є симетричними матрицями.*
2. *Добуток $AB$ є симетричною матрицею тоді і тільки тоді, коли $A$ та $B$ комутують.*

3. *Для довільної матриці $M \in M_{m\times n}(F)$ добутки $MM^T$ та $M^TM$ є симетричними матрицями (розмірів $m\times m$ та $n\times n$ відповідно).*

*Доведення.* За властивостями транспонування матриць перевіряємо симетричність матриць:

$$(A+B)^T = A^T + B^T = A + B, \qquad (\lambda A)^T = \lambda A^T = \lambda A,$$

$$(AB)^T = B^T A^T = BA, \qquad (MM^T)^T = MM^T, \quad (M^TM)^T = M^TM. \ \square$$

**Приклад 5.12.** Нехай

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 1 \end{pmatrix}.$$

Тоді:

$$AA^T = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -1 & 4 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 14 & -1 \\ -1 & 17 \end{pmatrix},$$

$$A^TA = \begin{pmatrix} 2 & 0 \\ -1 & 4 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 6 \\ -2 & 17 & 1 \\ 6 & 1 & 10 \end{pmatrix}.$$

Обидві матриці $AA^T$ та $A^TA$ є симетричними.

## 5.3 Обернена матриця

**Означення 5.9.** Квадратна матриця $A \in M_n(F)$ називається *оборотною*, якщо існує така матриця $B \in M_n(F)$, що $AB = BA = E$, де $E$ — одинична матриця. В цьому випадку, матриця $B$ називається *оберненою* до $A$ і позначається $A^{-1}$.

**Приклад 5.13.** Розглянемо матриці:

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Матриця $B$ є оберненою до матриці $A$, а матриця $A$ є оберненою до $B$:

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad BA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Матриця $C$ не є оборотною, оскільки добуток $C \cdot X$ для довільної матриці $X$ не може дорівнювати одиничній матриці:

$$C \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ a+c & b+d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{cases} a+c=1, \\ b+d=0, \\ a+c=0, \\ b+d=1, \end{cases}$$

що неможливо. Отже, $C$ не має оберненої матриці.

**Твердження 5.6.** *Кожна оборотна матриця $A$ має єдину обернену матрицю.*

*Доведення.* Нехай $B$ і $C$ — обернені до матриці $A$. Тоді:

$$B = BE = B(AC) = (BA)C = EC = C. \qquad \square$$

**Приклад 5.14.** Дослідимо умови оборотності матриці другого порядку та виведемо формулу для оберненої матриці. Нехай

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}.$$

За означенням оберненої матриці маємо:

$$A \cdot A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Перемноживши матриці та прирівнявши відповідні елементи, отримуємо дві системи лінійних рівнянь:

$$\begin{cases} ax + bz = 1, \\ cx + dz = 0 \end{cases} \qquad \text{та} \qquad \begin{cases} ay + bt = 0, \\ cy + dt = 1. \end{cases}$$

Розв'язавши ці системи, дістаємо, що розв'язок існує тоді і тільки тоді, коли величина $\Delta = ad - bc \neq 0$. У цьому випадку значення невідомих дорівнюють:

$$x = \frac{d}{ad-bc}, \quad z = \frac{-c}{ad-bc}, \quad y = \frac{-b}{ad-bc}, \quad t = \frac{a}{ad-bc}.$$

Таким чином, формула для оберненої матриці порядку два має вигляд:

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

**Теорема 5.1** (властивості оборотних матриць)**.** *Нехай $A, B \in M_n(F)$ — оборотні матриці.*

1. *Матриця $A^{-1}$ також є оборотною і*

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

2. *Для будь-якого $\lambda \in F$, $\lambda \neq 0$, матриця $\lambda A$ є оборотною і*

$$(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}.$$

3. *Добуток матриць $AB$ є оборотною матрицею і*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

4. *Транспонована матриця $A^T$ є оборотною і*

$$(A^T)^{-1} = (A^{-1})^T.$$

*Доведення.* 1) За означенням оберненої матриці маємо $AA^{-1} = A^{-1}A = E$. Ця ж рівність свідчить про те, що для матриці $A^{-1}$ оберненою є матриця $A$. Отже, $(A^{-1})^{-1} = A$.

2) Використовуючи властивості множення матриці на скаляр та асоціативність, отримуємо:

$$(\lambda A) \cdot \left(\frac{1}{\lambda} A^{-1}\right) = \left(\lambda \cdot \frac{1}{\lambda}\right)(AA^{-1}) = 1 \cdot E = E.$$

Аналогічно, $\left(\frac{1}{\lambda} A^{-1}\right) \cdot (\lambda A) = E$, що доводить рівність $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$.

3) Використовуючи асоціативність множення, отримуємо:

$$(AB) \cdot (B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E,$$

$$(B^{-1}A^{-1}) \cdot (AB) = B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E.$$

Отже, $(AB)^{-1} = B^{-1}A^{-1}$. Зверніть увагу на зміну порядку множників.

4) Скористаємося властивістю $(AB)^T = B^T A^T$:

$$A^T \cdot (A^{-1})^T = (A^{-1}A)^T = E^T = E, \quad (A^{-1})^T \cdot A^T = (AA^{-1})^T = E^T = E.$$

Це підтверджує, що $(A^T)^{-1} = (A^{-1})^T$. □

Алгебра матриць дозволяє компактно записувати та аналізувати системи лінійних рівнянь. Розглянемо систему $m$ лінійних рівнянь з $n$ невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

Введемо матрицю коефіцієнтів, вектор невідомих та вектор правих частин:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Помноживши матрицю на вектор невідомих, отримаємо:

$$A\boldsymbol{x} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Отже, систему лінійних рівнянь можна записати у матричному вигляді як $A\boldsymbol{x} = \boldsymbol{b}$.

**Твердження 5.7.** *Нехай $A \in M_n(F)$ — оборотна матриця, а $\boldsymbol{b} \in F^n$. Тоді система $A\boldsymbol{x} = \boldsymbol{b}$ має єдиний розв'язок, який задається формулою $\boldsymbol{x} = A^{-1}\boldsymbol{b}$.*

*Доведення.* Підставимо $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ в систему:

$$A\boldsymbol{x} = A(A^{-1}\boldsymbol{b}) = (AA^{-1})\boldsymbol{b} = E\boldsymbol{b} = \boldsymbol{b}.$$

Отже, $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ дійсно є розв'язком. Помноживши обидві частини системи на обернену до матриці $A$, отримуємо єдиність розв'язку:

$$A^{-1}(A\boldsymbol{x}) = A^{-1}\boldsymbol{b} \quad \Rightarrow \quad (A^{-1}A)\boldsymbol{x} = E\boldsymbol{x} = \boldsymbol{x} = A^{-1}\boldsymbol{b}. \qquad \square$$

**Приклад 5.15.** Розглянемо систему лінійних рівнянь:

$$\begin{cases} 3x + 2y = 2, \\ x + 2y = 3, \end{cases} \qquad \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Матриця коефіцієнтів

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$$

є оборотною, її обернена

$$A^{-1} = \frac{1}{4} \begin{pmatrix} 2 & -2 \\ -1 & 3 \end{pmatrix}.$$

Тоді розв'язком системи є:

$$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 2 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -2 \\ 7 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{7}{4} \end{pmatrix}.$$

**Зауваження 5.2.** Квадратна матриця $A \in M_n(F)$ визначає лінійне відображення $f_A : F^n \to F^n$ за правилом $f_A(\boldsymbol{x}) = A\boldsymbol{x}$. Якщо матриця $A$ є оборотною, то для будь-якого $\boldsymbol{y} \in F^n$ рівняння $A\boldsymbol{x} = \boldsymbol{y}$ має єдиний розв'язок $\boldsymbol{x} = A^{-1}\boldsymbol{y}$. Це означає, що кожному образу відповідає єдиний прообраз, тобто відображення $f_A$ — бієктивне. Навпаки, якщо $f_A$ є бієктивним, то для кожного вектора стандартного базису $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ існує рівно один такий вектор $\boldsymbol{b}_i$, що $A\boldsymbol{b}_i = \boldsymbol{e}_i$. Якщо скласти з цих стовпців $\boldsymbol{b}_i$ матрицю $B = [\boldsymbol{b}_1\ \boldsymbol{b}_2\ \ldots\ \boldsymbol{b}_n]$, то за правилом множення матриць отримаємо $AB = E$. Аналогічно доводиться, що $BA = E$. Отже, матриця $A$ є оборотною, а матриця $B$ — її оберненою. Таким чином, матриця $A$ є оборотною тоді і тільки тоді, коли відображення $f_A$ є бієктивним. У цьому випадку обернене відображення задається оберненою матрицею: $f_A^{-1}(\boldsymbol{y}) = A^{-1}\boldsymbol{y}$.

## 5.4 Елементарні перетворення та ступінчасті форми

У багатьох задачах лінійної алгебри виникає потреба виконувати перетворення рядків або стовпців матриці. За допомогою таких перетворень матрицю можна поступово привести до більш простого вигляду, не змінюючи її суттєвих властивостей.

**Означення 5.10.** *Елементарними перетвореннями над рядками* (або стовпцями) матриці називаються такі операції:

1. Перестановка двох рядків (стовпців) місцями. Позначення: $R_i \leftrightarrow R_j$ (для рядків) або $C_i \leftrightarrow C_j$ (для стовпців).

2. Множення рядка (стовпця) на ненульове число $\lambda \in F \setminus \{0\}$. Позначення: $R_i \to \lambda R_i$ або $C_i \to \lambda C_i$.

3. Додавання до одного рядка (стовпця) іншого рядка (стовпця), помноженого на число $\lambda \in F$. Позначення: $R_i \to R_i + \lambda R_j$ або $C_i \to C_i + \lambda C_j$.

**Приклад 5.16.** Зведемо матрицю

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 5 \\ 3 & 8 & 4 \end{pmatrix}$$

до діагонального вигляду елементарними перетвореннями над рядками:

1. Використаємо перший рядок, щоб обнулити елементи у першому стовпці під головною діагоналлю. До другого рядка додамо перший, помножений на $-2$, а до третього — перший, помножений на $-3$:
$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 5 \\ 3 & 8 & 4 \end{pmatrix} \xrightarrow[\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 3R_1}]{} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 3 \\ 0 & 2 & 1 \end{pmatrix}.$$

2. Оскільки на позиції $(2, 2)$ утворився нуль, переставимо другий та третій рядки місцями:
$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 3 \\ 0 & 2 & 1 \end{pmatrix} \xrightarrow[R_2 \leftrightarrow R_3]{} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

3. Обнулимо елементи над головною діагоналлю. До першого та другого рядків додамо третій, помножений на $-\frac{1}{3}$:
$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \xrightarrow[\substack{R_1 \to R_1 - \frac{1}{3}R_3 \\ R_2 \to R_2 - \frac{1}{3}R_3}]{} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

4. Обнулимо елемент над двійкою у другому стовпці:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \xrightarrow[R_1 \to R_1 - R_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

**Зауваження 5.3.** Будь-яка система лінійних рівнянь (СЛР) однозначно задає свою розширену матрицю, де рядки відповідають рівнянням, а стовпці — коефіцієнтам при невідомих та вільним членам:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \Leftrightarrow \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right).$$

Елементарні перетворення рядків такої матриці не змінюють множину розв'язків системи, тому вони перетворюють СЛР на еквівалентну їй систему. Якщо за допомогою таких перетворень звести матрицю до простої форми (наприклад, діагональної), це дозволяє одразу виписати розв'язки. Наприклад, розглянемо систему лінійних рівнянь, запишемо розширену матрицю та зведемо її до простого вигляду:

$$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 2x_1 + 5x_2 + 3x_3 = 9 \\ x_1 + 3x_2 + 4x_3 = 6 \end{cases} \Leftrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 2 & 5 & 3 & 9 \\ 1 & 3 & 4 & 6 \end{array}\right),$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 2 & 5 & 3 & 9 \\ 1 & 3 & 4 & 6 \end{array}\right) \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 3 & 2 \end{array}\right) \xrightarrow[R_3 \to R_3 - R_2]{}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right) \xrightarrow[R_3 \to \frac{1}{2}R_3]{} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & \frac{1}{2} \end{array}\right) \xrightarrow[R_1 \to R_1 - R_3]{R_2 \to R_2 - R_3}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & \frac{7}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \end{array}\right) \xrightarrow[R_1 \to R_1 - 2R_2]{} \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{5}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \end{array}\right) \Leftrightarrow \begin{cases} x_1 = \frac{5}{2}, \\ x_2 = \frac{1}{2}, \\ x_3 = \frac{1}{2}. \end{cases}$$

Отже, початкова СЛР рівносильна останній, яка має розв'язок $x_1 = \frac{5}{2}$, $x_2 = \frac{1}{2}$, $x_3 = \frac{1}{2}$.

**Означення 5.11.** Матриці $A$ та $B$ називаються *рядково* (відповідно *стовпцево*) *еквівалентними*, якщо одну з них можна отримати з іншої за допомогою скінченної кількості елементарних перетворень над рядками (відповідно стовпцями). Позначається $A \sim B$ (або $A \overset{r}{\sim} B$ та $A \overset{c}{\sim} B$ для уточнення).

**Зауваження 5.4.** Важливо зауважити, що кожне елементарне перетворення має зворотне. Тому, якщо $A \sim B$, то і $B \sim A$. Також $A \sim A$, а з умов $A \sim B$ та $B \sim C$ випливає, що $A \sim C$. Це означає, що рядкова (стовпцева) еквівалентність є відношенням еквівалентності на множині матриць розміру $m \times n$. У кожному класі еквівалентності природно виділити матрицю найпростішого вигляду. Хоча не кожну матрицю можна звести до діагонального вигляду, будь-яку матрицю можна звести до ступінчастої форми.

**Означення 5.12.** Матриця має *рядково-ступінчасту форму*, якщо виконуються такі умови:

1. Усі нульові рядки (якщо вони є) розташовані нижче за будь-який ненульовий рядок.
2. Перший ненульовий елемент кожного наступного рядка (так званий *ведучий елемент*) розташований праворуч від ведучого елемента попереднього рядка.

**Приклад 5.17.** Приклади матриць у рядково-ступінчастій формі:

$$\begin{pmatrix} \textcircled{1} & 2 & -1 & 0 & 4 & 5 \\ 0 & 0 & \textcircled{3} & 1 & -2 & 0 \\ 0 & 0 & 0 & \textcircled{7} & 6 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} \textcircled{1} & 2 & 0 & -1 & 3 & 4 & 5 \\ 0 & \textcircled{3} & 1 & 0 & 2 & -2 & 6 \\ 0 & 0 & 0 & 0 & \textcircled{2} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{5} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

де кружечком обведені ведучі елементи рядків. Приклади матриць, які не мають рядково-ступінчастої форми:

$$\begin{pmatrix} -2 & 1 & 2 \\ 1 & 3 & 0 \\ 0 & 5 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & -3 & 4 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 3 & 0 & 5 \\ 0 & 0 & 1 & -4 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & -3 & 2 \\ 0 & 0 & 2 & 1 & 0 & 2 \\ 0 & 1 & -4 & 3 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

У першій матриці ведучий елемент другого рядка стоїть не правіше за ведучий елемент першого рядка. У другій матриці нульовий рядок розташований вище за ненульовий, а в третій матриці ведучий елемент третього рядка стоїть не правіше за ведучий елемент другого рядка.

**Теорема 5.2** (Гаусса)**.** *Будь-яка матриця за допомогою скінченної кількості елементарних перетворень рядків може бути зведена до рядково-ступінчастої форми.*

*Доведення.* Доведемо твердження індукцією за розмірами матриці.

*База індукції.* Якщо матриця має один рядок, то вона вже має рядково-ступінчасту форму.

*Крок індукції.* Припускаємо, що твердження виконується для всіх матриць розміру менше ніж $m \times n$. Розглянемо матрицю $A$ розміру $m \times n$. Якщо всі елементи першого стовпця дорівнюють нулю, то відкинемо цей стовпець і застосуємо індукційне припущення до підматриці з $m$ рядків і $n-1$ стовпців. Нехай у першому стовпці є ненульовий елемент. Перестановкою рядків перенесемо один з таких елементів у перший рядок. Нехай він дорівнює $a_{11} \neq 0$. Далі за допомогою елементарних перетворень виду $R_i \to R_i - \frac{a_{i1}}{a_{11}} R_1$ для $i = 2, \ldots, m$, занулюємо всі елементи під $a_{11}$ у першому стовпці. У результаті матриця зводиться до вигляду:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

За припущенням індукції підматрицю, утворену рядками $2, \ldots, m$, можна звести до рядково-ступінчастої форми. Оскільки ці перетворення не зачіпають перший рядок і не змінюють нулі у першому стовпці, вся матриця в результаті набуде рядково-ступінчастої форми. □

Алгоритмічна процедура, описана в доведенні теореми Гаусса, називається *прямим ходом методу Гаусса*. Він полягає у послідовному обнуленні елементів під ведучими елементами при русі зліва направо та зверху вниз. Результатом прямого ходу завжди є матриця у рядково-ступінчастій формі.

**Приклад 5.18.** Зведемо матрицю

$$A = \begin{pmatrix} 1 & -2 & 3 & 1 \\ 2 & -4 & 7 & 5 \\ -1 & 3 & -1 & 0 \end{pmatrix}$$

до рядково-ступінчастої форми прямим ходом методу Гаусса:

1. Обнулимо елементи у першому стовпці під ведучою одиницею. До другого рядка додамо перший, помножений на $-2$, а до третього рядка додамо перший:

$$\begin{pmatrix} 1 & -2 & 3 & 1 \\ 2 & -4 & 7 & 5 \\ -1 & 3 & -1 & 0 \end{pmatrix} \xrightarrow[R_3 \to R_3 + R_1]{R_2 \to R_2 - 2R_1} \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & 2 & 1 \end{pmatrix}.$$

2. Зауважимо, що на позиції $(2, 2)$ стоїть нуль, а під ним у третьому рядку — ненульовий елемент. Поміняємо другий та третій рядки місцями:

$$\begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & 2 & 1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

Отримана матриця має ступінчасту форму.

**Зауваження 5.5.** Матриця може бути зведена до різних ступінчастих форм. Наприклад, розглянемо матрицю $A$ і зведемо її до ступінчастої форми різними способами:

$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 5 \\ 1 & 2 & 3 \end{pmatrix} \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - 2R_1} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 5 \\ 1 & 2 & 3 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 5 \\ 1 & 1 & 2 \end{pmatrix} \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - 2R_1}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -1 \\ 0 & -1 & -1 \end{pmatrix} \xrightarrow{R_3 \to R_3 - \frac{1}{2} R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -1 \\ 0 & 0 & -1/2 \end{pmatrix}.$$

Розглянемо форму матриць, яка визначається однозначно.

**Означення 5.13.** Матриця має *редуковану рядково-ступінчасту форму*, якщо виконуються такі умови:

1. Матриця має рядково-ступінчасту форму.
2. Кожен перший ненульовий елемент (ведучий елемент) будь-якого ненульового рядка дорівнює одиниці.
3. У стовпці, де стоїть ведуча одиниця, всі інші елементи дорівнюють нулю.

**Приклад 5.19.** Приклади редукованих рядково-ступінчастих форм:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 0 & 2 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 0 & -2 & 4 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Наступні матриці мають ступінчасту форму, але не редуковану:

$$\begin{pmatrix} 2 & 1 & 4 & 3 & 0 & 5 \\ 0 & 0 & 1 & -2 & 5 & 1 \\ 0 & 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 1 & 4 & 5 \\ 0 & 1 & 0 & 2 & 3 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

У першої матриці ведучий елемент першого рядка не є одиничним (він дорівнює 2), а у другої матриці не всі елементи над ведучими одиницями дорівнюють нулю.

**Теорема 5.3** (Гаусс)**.** *Кожна матриця за допомогою елементарних перетворень рядків зводиться до редукованої рядково-ступінчастої форми.*

*Доведення.* Ми вже знаємо, що будь-яку матрицю можна звести до рядково-ступінчастої форми прямим ходом методу Гаусса. Опишемо *зворотній хід методу Гауссу*, який зводить ступінчасту матрицю до редукованої форми. Починаючи з останнього ненульового рядка і рухаючись угору, для кожного рядка виконуємо такі дії:

1. Ділимо рядок на значення його ведучого елемента, щоб перетворити цей елемент на одиницю.
2. Для кожного рядка, що розташований вище поточного, віднімаємо від нього поточний рядок, помножений на елемент цього рядка у стовпці ведучого елемента. У результаті всі елементи над ведучою одиницею стають рівними нулю.

Після виконання цієї процедури для всіх ненульових рядків матриця переходить у редуковану рядково-ступінчасту форму. □

**Зауваження 5.6.** Можна показати, що редукована рядково-ступінчаста форма матриці визначається однозначно. Тому дві матриці $A$ та $B$ є рядково-еквівалентними тоді і тільки тоді, коли вони мають однакову редуковану рядково-ступінчасту форму.

**Приклад 5.20.** Зведемо матрицю

$$A = \begin{pmatrix} 2 & 4 & -2 & 2 \\ 1 & 2 & 1 & 3 \\ -1 & -2 & 3 & 3 \end{pmatrix}$$

до редукованої рядково-ступінчастої форми методом Гаусса. Прямий хід:

1. На позиції $(1,1)$ стоїть елемент 2. Поміняємо перший та другий рядки місцями, щоб отримати ведучу одиницю:

$$\begin{pmatrix} 2 & 4 & -2 & 2 \\ 1 & 2 & 1 & 3 \\ -1 & -2 & 3 & 3 \end{pmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{} \begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 4 & -2 & 2 \\ -1 & -2 & 3 & 3 \end{pmatrix}.$$

2. Обнулимо елементи у першому стовпці під ведучою одиницею:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 4 & -2 & 2 \\ -1 & -2 & 3 & 3 \end{pmatrix} \xrightarrow[\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 + R_1}]{} \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 4 & 6 \end{pmatrix}.$$

3. Зауважимо, що другий стовпець став нульовим нижче першого рядка, тому ведучий елемент другого рядка знаходиться у третьому стовпці. Обнулимо елемент під ним:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 4 & 6 \end{pmatrix} \xrightarrow[R_3 \to R_3 + R_2]{} \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Отримана матриця має ступінчасту форму. Зворотній хід:

4. Нормалізуємо третій рядок, поділивши його на 2:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & 2 \end{pmatrix} \xrightarrow[R_3 \to \frac{1}{2} R_3]{} \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

5. Обнулимо елементи у четвертому стовпці над ведучим елементом третього рядка:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\substack{R_2 \to R_2 + 4R_3 \\ R_1 \to R_1 - 3R_3}]{} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

6. Нормалізуємо другий рядок, поділивши його на $-4$:

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_2 \to -\frac{1}{4}R_2]{} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

7. Обнулимо елемент у третьому стовпці над ведучим елементом другого рядка:

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_1 \to R_1 - R_2]{} \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Отримана матриця є редукованою рядково-ступінчастою формою матриці $A$.

## 5.5 Елементарні матриці та матричні рівняння

**Означення 5.14.** Матриця називається *елементарною*, якщо вона отримується з одиничної матриці шляхом виконання одного з елементарних перетворень.

Існує три типи елементарних матриць, які відповідають трьом типам елементарних перетворень:

1. Перестановка $i$-го та $j$-го рядків:

$$P_{ij} = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}.$$

2. Множення $i$-го рядка на ненульовий скаляр $\lambda \neq 0$:

$$E_i(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}.$$

3. Додавання до $i$-го рядка $j$-го рядка, помноженого на скаляр $\lambda$:

$$E_{ij}(\lambda) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}.$$

**Приклад 5.21.** Кілька прикладів елементарних матриць порядку три:

$$P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad E_2(-5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$E_{12}(7) = \begin{pmatrix} 1 & 7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{23}(-2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Твердження 5.8.** *1. Елементарні перетворення над рядками матриці евівалентні множенню матриці зліва на відповідну елементарну матрицю.*

*2. Елементарні перетворення над стовпцями матриці евівалентні множенню матриці справа на відповідну елементарну матрицю.*

*Доведення.* Твердження перевіряється множенням матриці на відповідну елементарну матрицю. □

**Приклад 5.22.** Виконаємо кілька елементарних перетворень над рядками матриці та перевіримо множення на елементарну матрицю. Переставимо перший і третій рядки матриці

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 4 \\ 1 & 2 & 3 \end{pmatrix}.$$

Це відповідає множенню зліва на елементарну матрицю:

$$P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad P_{13}A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 4 \\ 1 & 2 & 3 \end{pmatrix}.$$

Помножимо другий рядок на 5 за допомогою відповідної елементарної матриці:

$$E_2(5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_2(5)A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 20 \\ 2 & 0 & 1 \end{pmatrix}.$$

Додамо до третього рядка перший рядок:

$$E_{31}(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \qquad E_{31}(1)A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 3 & 2 & 4 \end{pmatrix}.$$

Кожне елементарне перетворення рядків або стовпців має зворотнє перетворення. У термінах алгебри матриць це означає, що елементарні матриці є оборотними, а їх обернені матриці є елементарними.

**Твердження 5.9.** *Кожна елементарна матриця є оборотною, а її обернена матриця також є елементарною матрицею того ж типу:*

$$P_{ij}^{-1} = P_{ij}; \qquad E_i(\lambda)^{-1} = E_i(\lambda^{-1}); \qquad E_{ij}(\lambda)^{-1} = E_{ij}(-\lambda).$$

Послідовність елементарних перетворень відповідає множенню на добуток елементарних матриць. Тому, зводячи матрицю до ступінчастої форми методом Гауса, можна визначити, чи є вона оборотною.

**Теорема 5.4.** *Нехай $A \in M_n(F)$. Наступні умови рівносильні:*

1. *Матриця $A$ є оборотною.*

2. *Матриця $A$ зводиться до одиничної матриці $E$ за допомогою елементарних перетворень над рядками (стовпцями).*

3. *Матриця $A$ розкладається у добуток елементарних матриць.*

*Доведення.* (1) $\Rightarrow$ (2). Нехай $A$ — оборотна матриця. За допомогою елементарних перетворень рядків (прямий хід методу Гаусса) зводимо матрицю $A$ до ступінчатої форми $R$:

$$R = \begin{pmatrix} r_{11} & * & * & \dots & * \\ 0 & r_{22} & * & \dots & * \\ 0 & 0 & r_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ 0 & 0 & 0 & \dots & r_{nn} \end{pmatrix}.$$

Оскільки кожне елементарне перетворення відповідає множенню на елементарну матрицю, то існують такі елементарні матриці $E_1, \dots, E_k$, що $E_k \cdots E_1 A = R$. Враховуючи, що матриця $A$ та всі елементарні матриці $E_i$ є оборотними, їхній добуток $R$ також має бути оборотною матрицею. Тому $R$ не може містити нульового рядка, і всі діагональні елементи $r_{11}, r_{22}, \dots, r_{nn}$ є ненульовим. Поділивши рядки на відповідні діагональні елементи та застосовуючи зворотній хід методу Гаусса, матриця зводиться до одиничної $E$.

(2) $\Rightarrow$ (3). Нехай матриця $A$ зводиться до одиничної $E$ за допомогою елементарних перетворень рядків. Тоді існують елементарні матриці $E_1, \dots, E_k$, такі що $E_k \cdots E_1 A = E$. Помноживши на обернені матриці, одержуємо $A = E_1^{-1} \cdots E_k^{-1}$. Оскільки обернені до елементарних матриць також є елементарними, матриця $A$ розкладається у добуток елементарних матриць.

(3) $\Rightarrow$ (1). Кожна елементарна матриця є оборотною, а добуток оборотних матриць є оборотним.Тому добуток елементарних матриць є оборотною матрицею. $\square$

Критерій, наведений у теоремі, дозволяє використовувати метод Гаусса для перевірки оборотності матриці. Для цього задана матриця зводиться до ступінчатої форми: матриця є оборотною тоді і лише тоді, коли у її ступінчастій формі немає нульових рядків.

**Приклад 5.23.** Перевіримо матриці на оборотність, де

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 1 & -1 & 10 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -3 & 4 \\ 3 & -1 & 3 \end{pmatrix}.$$

Для цього зведемо їх до ступінчатої форми за допомогою методу Гаусса:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 1 & -1 & 10 \end{pmatrix} \xrightarrow[R_3-R_1]{R_2-2R_1} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 0 & -3 & 7 \end{pmatrix} \xrightarrow{R_3+3R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$B = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -3 & 4 \\ 3 & -1 & 3 \end{pmatrix} \xrightarrow[R_3-3R_1]{R_2-2R_1} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -7 & 6 \\ 0 & -7 & 6 \end{pmatrix} \xrightarrow{R_3-R_2} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -7 & 6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ступінчаста форма матриці $A$ не має нульових рядків, тому вона є оборотною. А у ступінчастій формі матриці $B$ з'явився нульовий рядок, тому вона не є оборотною.

**Приклад 5.24.** Розкладемо матрицю у добуток елементарних:

$$A = \begin{pmatrix} 2 & -1 & -2 \\ -3 & 2 & 1 \\ 1 & 0 & 5 \end{pmatrix}.$$

Зводимо матрицю до одиничної матриці $E$ методом Гауса, запам'ятовуючи перетворення, які ми робимо над рядками матриці:

$$\begin{pmatrix} 2 & -1 & -2 \\ -3 & 2 & 1 \\ 1 & 0 & 5 \end{pmatrix} \xrightarrow{P_{13}} \begin{pmatrix} 1 & 0 & 5 \\ -3 & 2 & 1 \\ 2 & -1 & -2 \end{pmatrix} \xrightarrow[E_{31}(-2)]{E_{21}(3)} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 2 & 16 \\ 0 & -1 & -12 \end{pmatrix}$$

$$\xrightarrow{P_{23}} \begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & -12 \\ 0 & 2 & 16 \end{pmatrix} \xrightarrow{E_{32}(2)} \begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & -12 \\ 0 & 0 & -8 \end{pmatrix} \xrightarrow{E_3(-\frac{1}{8})}$$

$$\begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & -12 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[E_{13}(-5)]{E_{23}(12)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{E_2(-1)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Оскільки кожне елементарне перетворення над рядками відповідає множенню матриці зліва на елементарну матрицю, ми отримаємо:

$$E_2(-1)E_{13}(-5)E_{23}(12)E_3(-\tfrac{1}{8})E_{32}(2)P_{23}E_{31}(-2)E_{21}(3)P_{13}A = E.$$

Множимо на обернені до елементарних матриць, які в свою чергу є елементарними, і отримаємо розклад матриці $A$ у добуток елементарних:

$$A = P_{13}E_{21}(-3)E_{31}(2)P_{23}E_{32}(-2)E_3(-8)E_{23}(-12)E_{13}(5)E_2(-1)$$
$$= \begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\-3&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&-2&1\end{pmatrix}$$
$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&-8\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&-12\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&5\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}.$$

Теорема 5.4 дозволяє не тільки перевіряти, чи є матриця оборотною, а й знаходити обернену матрицю. *Метод Гаусса* для знаходження оберненої до матриці $A$ полягає в наступному:

1. Записуємо розширену матрицю $(A \mid E)$.

2. Виконуючи елементарні перетворення над рядками матриці $(A \mid E)$, зводимо її ліву частину $A$ до одиничної матриці $E$.

3. Якщо в процесі перетворень отримати одиничну матрицю неможливо (у ступінчастій формі з'являється нульовий рядок), то обернена матриця до $A$ не існує.

4. Якщо в результаті перетворень отримано матрицю $(E \mid B)$, то матриця $A$ є оборотною, а матриця $B$ є оберненою до $A$, тобто $B = A^{-1}$.

**Приклад 5.25.** Знайдемо обернену до матриці

$$A = \begin{pmatrix}2&1&1\\1&1&0\\1&2&1\end{pmatrix}.$$

Застосуємо метод Гаусса:

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[P_{12}]{} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R_3-R_1]{R_2-2R_1}$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & -2 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array}\right) \xrightarrow[P_{23}]{} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \\ 0 & -1 & 1 & 1 & -2 & 0 \end{array}\right) \xrightarrow{R_3+R_2}$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \\ 0 & 0 & 2 & 1 & -3 & 1 \end{array}\right) \xrightarrow{R_3\cdot\frac{1}{2}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{R_2-R_3}$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{R_1-R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right).$$

Отже,

$$A^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix}.$$

Зробимо перевірку:

$$A \cdot A^{-1} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Теорема 5.5.** *Нехай $A, B \in M_n(F)$ і матриця $A$ є оборотною. Тоді кожне з матричних рівнянь $AX = B$ та $XA = B$ має єдиний розв'язок, який знаходиться як $X = A^{-1}B$ та $X = BA^{-1}$ відповідно.*

*Доведення.* Розглянемо рівняння $AX = B$. Перевіряємо, що $X = A^{-1}B$ дійсно є розв'язком:

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

Для доведення єдиності розв'язку, помножимо рівняння $AX = B$ зліва на $A^{-1}$:

$$A^{-1}(AX) = A^{-1}B \quad \Rightarrow \quad A^{-1}AX = (E)X = A^{-1}B \quad \Rightarrow \quad X = A^{-1}B.$$

Аналогічно розв'язується рівняння $XA = B$. □

Матричні рівняння вигляду $AX = B$ та $XA = B$ можна розв'язувати методом Гаусса, аналогічно до процедури знаходження оберненої матриці. Цей підхід часто є більш ефективним, оскільки дозволяє знайти невідому матрицю без явного обчислення $A^{-1}$. Метод Гаусса для знаходження розв'язків рівняння $AX = B$:

1. Записуємо розширену матрицю $(A \mid B)$, де ліворуч стоїть матриця коефіцієнтів, а праворуч — матриця вільних членів.

2. Виконуючи елементарні перетворення над рядками матриці $(A \mid B)$, зводимо її ліву частину $A$ до одиничної матриці $E$.

3. Якщо в процесі зведення у лівій частині утворюється нульовий рядок, це свідчить про те, що матриця $A$ не є оборотною.

4. Якщо в результаті отримано матрицю виду $(E \mid C)$, то матриця $C$ є шуканим розв'язком: $X = C = A^{-1}B$.

Аналогічно розв'язується рівняння $XA = B$, яке можна звести до попереднього виду за допомогою транспонування: $(XA)^T = A^T X^T = B^T$.

Обґрунтування цього методу ґрунтується на зв'язку між елементарними перетвореннями та множенням на елементарні матриці. Нехай послідовність елементарних перетворень рядків, яка зводить матрицю $A$ до одиничної $E$, задається елементарними матрицями $E_1, E_2, \ldots, E_k$. Тоді $(E_k \cdots E_2 E_1)A = E$ і добуток $E_k \cdots E_2 E_1$ є оберненою матрицею $A^{-1}$. Коли ми виконуємо ці ж самі перетворення над розширеною матрицею $(A \mid B)$, ми множимо її зліва на ту саму матрицю $A^{-1} = E_k \cdots E_2 E_1$ і права частина набуває вигляду $X = A^{-1}B$, що є шуканим розв'язком рівняння.

**Приклад 5.26.** Розв'яжемо матричне рівняння

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 2 & -1 & 0 \end{pmatrix} \cdot X = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & 2 \\ 0 & 1 & 4 \end{pmatrix}.$$

Записуємо розширену матрицю та застосовуємо метод Гаусса:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & 3 & 2 \\ 2 & -1 & 0 & 0 & 1 & 4 \end{array}\right) \xrightarrow[R_3-2R_1]{} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & 3 & 2 \\ 0 & -5 & -2 & -4 & 1 & 2 \end{array}\right) \xrightarrow[R_3+5R_2]{}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & 3 & 2 \\ 0 & 0 & 13 & 1 & 16 & 12 \end{array}\right) \xrightarrow[R_3\cdot\frac{1}{13}]{} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & 3 & 2 \\ 0 & 0 & 1 & \frac{1}{13} & \frac{16}{13} & \frac{12}{13} \end{array}\right) \xrightarrow[R_1-R_3]{R_2-3R_3}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{25}{13} & -\frac{16}{13} & \frac{1}{13} \\ 0 & 1 & 0 & \frac{10}{13} & -\frac{9}{13} & -\frac{10}{13} \\ 0 & 0 & 1 & \frac{1}{13} & \frac{16}{13} & \frac{12}{13} \end{array}\right) \xrightarrow[R_1-2R_2]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{13} & \frac{2}{13} & \frac{21}{13} \\ 0 & 1 & 0 & \frac{10}{13} & -\frac{9}{13} & -\frac{10}{13} \\ 0 & 0 & 1 & \frac{1}{13} & \frac{16}{13} & \frac{12}{13} \end{array}\right).$$

Отже, шуканий розв'язок:

$$X = \begin{pmatrix} 5/13 & 2/13 & 21/13 \\ 10/13 & -9/13 & -10/13 \\ 1/13 & 16/13 & 12/13 \end{pmatrix} = \frac{1}{13}\begin{pmatrix} 5 & 2 & 21 \\ 10 & -9 & -10 \\ 1 & 16 & 12 \end{pmatrix}.$$

## 5.6 Вправи

1. Виконайте операції над матрицями:

   а) $3\begin{pmatrix} 2 & 3 \\ 1 & 5 \end{pmatrix} - 2\begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix}^T$; б) $3\begin{pmatrix} -2 & 1 \\ 4 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & -2 \end{pmatrix} \cdot \begin{pmatrix} -3 & 2 \\ 5 & 1 \end{pmatrix}$;

   в) $\begin{pmatrix} 2 & 5 \\ -1 & 3 \\ 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix}$; г) $\begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & -3 \\ 5 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 & 5 \\ 1 & -3 & 2 \\ 3 & 7 & -1 \end{pmatrix}$;

   д) $\begin{pmatrix} 4 & 2 & 3 \\ 1 & -3 & 5 \\ 2 & 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 & 2 \\ 2 & 4 & 5 \\ 1 & 3 & 2 \end{pmatrix} + \begin{pmatrix} 5 & -1 & 2 \\ 3 & 2 & 4 \\ 4 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 4 & 1 \\ 3 & 5 & 2 \\ 1 & 2 & 4 \end{pmatrix}$.

2. Обчисліть значення многочлена $f(x)$ від матриці $A$:

   а) $f(x) = x^3 - 2x^2 + 4x + 5$, $A = \begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix}$;

   б) $f(x) = 4x^2 - 3x + 2$, $A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & 1 \\ -2 & 0 & 4 \end{pmatrix}$.

3. Зведіть матрицю до редукованої рядково-ступінчастої форми:

а) $\begin{pmatrix} 1 & -2 & 3 & -4 \\ 2 & -4 & 7 & -9 \\ -1 & 2 & 0 & 1 \end{pmatrix}$; б) $\begin{pmatrix} 2 & 4 & -2 & 2 & 4 \\ 1 & 2 & -1 & 2 & 0 \\ 3 & 6 & -3 & 1 & 10 \end{pmatrix}$;

в) $\begin{pmatrix} 1 & 2 & -1 & 2 & 1 & 3 \\ 2 & 4 & -1 & 5 & 0 & 8 \\ -1 & -2 & 3 & 0 & 5 & 1 \\ 1 & 2 & 1 & 4 & -1 & 7 \end{pmatrix}$.

4. Знайдіть обернену матрицю, якщо вона існує:

а) $\begin{pmatrix} 5 & 8 \\ 3 & 5 \end{pmatrix}$; б) $\begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}$; в) $\begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix}$.

5. Розв'яжіть матричне рівняння:

а) $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 3 \\ -1 & 3 & 1 \end{pmatrix} X = \begin{pmatrix} 5 & 0 \\ 11 & 3 \\ 10 & -5 \end{pmatrix}$;

б) $X \begin{pmatrix} 3 & -1 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 4 & -3 \\ 12 & -3 & 11 \end{pmatrix}$.

6. Розкладіть матрицю у добуток елементарних:

а) $\begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}$; б) $\begin{pmatrix} 2 & 1 & 3 \\ -1 & 3 & 1 \\ 3 & 5 & -2 \end{pmatrix}$; в) $\begin{pmatrix} 4 & 5 & 3 \\ 2 & -3 & 2 \\ -2 & 2 & 3 \end{pmatrix}$.

7. Нехай $A, B, C$ — квадратні матриці порядку $n$. Доведіть, що з умов $AB = E$ та $CA = E$ випливає рівність $B = C$.

8. Доведіть, що для симетричних матриць $A, B \in M_n(\mathbb{R})$ добуток $AB$ є симетричною матрицею тоді і тільки тоді, коли $AB = BA$.

9. Доведіть, що будь-яку квадратну матрицю $A \in M_n(\mathbb{R})$ можна єдиним чином подати у вигляді суми симетричної та кососиметричної матриць.

10. Опишіть усі матриці $A \in M_n(\mathbb{R})$ такі, що $AA^T = O$.

11. Опишіть усі матриці $A \in M_2(\mathbb{R})$ такі, що $A^2 = O$.

12. Опишіть усі матриці $A \in M_2(\mathbb{R})$ такі, що $A^2 = A$.

13. Доведіть, що не існує таких матриць $A, B \in M_n(\mathbb{R})$, для яких виконується рівність $AB - BA = E$.

14. Доведіть, що якщо матриця $E + AB$ є оборотною, то матриця $E + BA$ також є оборотною.

# Розділ 6

# Підстановки

## 6.1 Поняття підстановки

**Означення 6.1.** *Підстановкою* на множині $X = \{1, 2, \ldots, n\}$ називається взаємно однозначне відображення множини $X$ на саму себе.

Іншими словами, підстановка — це бієкція $\sigma\colon X \to X$. Підстановку зручно записувати у вигляді таблиці:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix},$$

де у верхньому рядку стоять елементи множини $X$, а у нижньому — їхні образи. Оскільки $\sigma$ є взаємно однозначним відображенням, кожне з чисел $1, 2, \ldots, n$ зустрічається у нижньому рядку рівно один раз. Отже, нижній рядок є перестановкою елементів верхнього. Зауважимо, що порядок елементів у верхньому рядку не є суттєвим: його можна довільно змінювати, якщо відповідно переставити елементи в нижньому рядку, тобто

$$\sigma = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \sigma(a_1) & \sigma(a_2) & \cdots & \sigma(a_n) \end{pmatrix},$$

де $a_1, a_2, \ldots, a_n$ — довільна перестановка чисел $1, 2, \ldots, n$. Тому підстановка визначається відображенням $\sigma$, а не конкретним виглядом таблиці.

**Приклад 6.1.** Підстановка на множині $\{1, 2, 3\}$ та її табличне зобра-

ження:

$$\sigma: \begin{array}{l} 1 \mapsto 3, \\ 2 \mapsto 2, \\ 3 \mapsto 1 \end{array} \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 1 & 2 \end{pmatrix}.$$

Підстановка на множині $\{1, 2, 3, 4, 5, 6, 7\}$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 1 & 6 & 7 & 4 & 5 \end{pmatrix}.$$

Наступні відображення не є підстановками:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 1 & 6 & 7 & 4 & 4 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 1 & 6 & 7 & 4 & 8 \end{pmatrix}.$$

Перше відображення не є бієкцією, оскільки значення 4 зустрічається двічі, а значення 5 не зустрічається жодного разу. Друге відображення не є відображенням множини $X$ на саму себе, оскільки $8 \notin X$.

Множину всіх підстановок на множині $\{1, 2, \ldots, n\}$ позначають $S_n$:

$$S_n = \{\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\} \mid \sigma \text{ — бієкція}\} =$$
$$= \left\{ \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \,\middle|\, a_1, a_2, \ldots, a_n \text{ — перестановка чисел } 1, 2, \ldots, n \right\}.$$

**Твердження 6.1.** *Множина $S_n$ містить рівно $n!$ елементів.*

*Доведення.* Оскільки підстановка є бієкцією, всі образи $\sigma(1)$, $\sigma(2)$, …, $\sigma(n)$ мають бути різними. Образ $\sigma(1)$ можна обрати $n$ способами, після чого образ $\sigma(2)$ — $(n-1)$ способами, і так далі. За правилом добутку загальна кількість підстановок дорівнює

$$n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n!. \qquad \square$$

Випишемо всі підстановки на множині $\{1, 2, 3\}$. Оскільки $|S_3| = 3! = 6$, маємо шість підстановок:

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \right.$$
$$\left. \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}.$$

## 6.2 Множення підстановок

Оскільки підстановки є відображеннями, можна розглянути їх композицію.

**Означення 6.2.** *Добутком* підстановок $\sigma, \tau \in S_n$ називається їх композиція $\sigma \cdot \tau$, тобто підстановка, що діє за правилом $(\sigma \cdot \tau)(i) = \sigma(\tau(i))$.

Іншими словами, щоб знайти образ елемента $i$ під дією добутку $\sigma \cdot \tau$, спочатку застосовуємо підстановку $\tau$ до $i$, а потім застосовуємо підстановку $\sigma$ до отриманого результату.

Якщо підстановки $\sigma$ і $\tau$ задані у вигляді таблиць

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix},$$

то їх добуток обчислюється за правилом

$$\sigma \cdot \tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(\tau(1)) & \sigma(\tau(2)) & \cdots & \sigma(\tau(n)) \end{pmatrix}.$$

Практично це означає таке: щоб знайти елемент у нижньому рядку добутку під номером $i$, спочатку знаходимо $\tau(i)$ (у таблиці $\tau$), а потім підставляємо цей результат у таблицю $\sigma$.

**Приклад 6.2.** Розглянемо підстановки на множині $\{1, 2, 3\}$ та обчислимо їх добуток:

$$\sigma: \begin{array}{l} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 1 \end{array}, \quad \tau: \begin{array}{l} 1 \mapsto 3 \\ 2 \mapsto 1 \\ 3 \mapsto 2 \end{array}, \quad \sigma \cdot \tau: \begin{array}{l} 1 \mapsto 3 \mapsto 1 \\ 2 \mapsto 1 \mapsto 2 \\ 3 \mapsto 2 \mapsto 3 \end{array}.$$

Обчислимо добуток підстановок, заданих таблицями:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 6 & 1 & 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 6 & 5 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 2 & 3 & 5 & 1 \end{pmatrix}.$$

У цьому добутку, наприклад, ми отримали $1 \mapsto 6$, оскільки $1 \mapsto 3$ другою підстановкою, а потім $3 \mapsto 6$ першою. Аналогічно, оскільки $2 \mapsto 1$ другою підстановкою, а потім $1 \mapsto 4$ першою, маємо $2 \mapsto 4$.

**Твердження 6.2** (властивості множення підстановок)**.** *Множина $S_n$ відносно множення підстановок є групою. Це означає, що виконуються наступні властивості:*

1. *Замкненість: для всіх $\sigma, \tau \in S_n$ добуток $\sigma \cdot \tau$ належить $S_n$.*

2. *Асоціативність: для всіх $\sigma, \tau, \rho \in S_n$ виконується*

$$(\sigma \cdot \tau) \cdot \rho = \sigma \cdot (\tau \cdot \rho).$$

3. *Нейтральний елемент: існує тотожна підстановка*

$$\varepsilon = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} \in S_n,$$

*така, що для всіх $\sigma \in S_n$ виконується $\varepsilon \cdot \sigma = \sigma \cdot \varepsilon = \sigma$.*

4. *Оберненість: для кожної $\sigma \in S_n$ існує $\tau \in S_n$, обернена до $\sigma$, така що $\sigma \cdot \tau = \tau \cdot \sigma = \varepsilon$.*

*Доведення.* 1. Композиція двох бієкцій є бієкцією, отже $\sigma \cdot \tau \in S_n$.

2. Асоціативність випливає з асоціативності композиції відображень: для кожного $i$ маємо $((\sigma \cdot \tau) \cdot \rho)(i) = \sigma(\tau(\rho(i))) = (\sigma \cdot (\tau \cdot \rho))(i)$. Отже, $(\sigma \cdot \tau) \cdot \rho = \sigma \cdot (\tau \cdot \rho)$.

3. Тотожна підстановка $\varepsilon$ не змінює жодного елемента, тому $(\varepsilon \cdot \sigma)(i) = (\sigma \cdot \varepsilon)(i) = \sigma(i)$ для всіх $i$.

4. Оскільки $\sigma$ є бієкцією на множині $\{1, 2, \ldots, n\}$, існує обернене відображення $\sigma^{-1}$ на цій множині, яке також є бієкцією, а отже, $\sigma^{-1} \in S_n$. За означенням оберненого відображення

$$\sigma \cdot \sigma^{-1} = \sigma^{-1} \cdot \sigma = \varepsilon. \qquad \square$$

**Означення 6.3.** Група $S_n$ називається *симетричною групою степеня $n$*.

**Зауваження 6.1.** З асоціативності випливає, що добуток підстановок $\sigma_1 \cdot \sigma_2 \cdot \ldots \cdot \sigma_k$ не залежить від того, в якому порядку вони обчислюються, тобто результат від розстановки дужок однаковий. Тому дужки можна не писати, вираз $\sigma_1 \cdot \sigma_2 \cdot \ldots \cdot \sigma_k$ є однозначно визначеним. У той же час множення не є комутативним, тобто загалом $\sigma \cdot \tau \neq \tau \cdot \sigma$. Наприклад,

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Кожна підстановка має єдину обернену підстановку, яка є оберненим відображенням. Обернену підстановку легко знайти з таблиці: достатньо поміняти місцями верхній і нижній рядки (а потім можна впорядкувати стовпці так, щоб верхній рядок набув натурального порядку):

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \quad \Rightarrow \quad \sigma^{-1} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

**Приклад 6.3.** Розглянемо підстановку та її обернену:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}, \qquad \sigma^{-1} = \begin{pmatrix} 3 & 1 & 4 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

**Твердження 6.3** (властивості оберненої підстановки)**.** *Для всіх підстановок $\sigma, \tau, \sigma_i \in S_n$ виконуються наступні властивості:*

1. $(\sigma^{-1})^{-1} = \sigma$.
2. $(\sigma \cdot \tau)^{-1} = \tau^{-1} \cdot \sigma^{-1}$.
3. $(\sigma_1 \cdot \sigma_2 \cdot \ldots \cdot \sigma_k)^{-1} = \sigma_k^{-1} \cdot \ldots \cdot \sigma_2^{-1} \cdot \sigma_1^{-1}$.

*Доведення.* 1. За означенням, $\sigma^{-1}$ є оберненою до $\sigma$, тобто $\sigma \cdot \sigma^{-1} = \varepsilon$. Це означає, що також $\sigma$ є оберненою до $\sigma^{-1}$, тобто $(\sigma^{-1})^{-1} = \sigma$.

2. Перевіримо, що $\tau^{-1} \cdot \sigma^{-1}$ є оберненою до $\sigma \cdot \tau$:

$$(\sigma \cdot \tau) \cdot (\tau^{-1} \cdot \sigma^{-1}) = \sigma \cdot (\tau \cdot \tau^{-1}) \cdot \sigma^{-1} = \sigma \cdot \varepsilon \cdot \sigma^{-1} = \sigma \cdot \sigma^{-1} = \varepsilon.$$

Аналогічно перевіряється, що $(\tau^{-1} \cdot \sigma^{-1}) \cdot (\sigma \cdot \tau) = \varepsilon$. Оскільки обернена підстановка єдина, маємо $(\sigma \cdot \tau)^{-1} = \tau^{-1} \cdot \sigma^{-1}$.

3. Випливає з пункту 2 індукцією за $k$. База $k = 2$ доведена вище. Якщо твердження виконується для добутку $k-1$ підстановок, то

$$(\sigma_1 \cdot \ldots \cdot \sigma_{k-1} \cdot \sigma_k)^{-1} = \sigma_k^{-1} \cdot (\sigma_1 \cdot \ldots \cdot \sigma_{k-1})^{-1} = \sigma_k^{-1} \cdot \sigma_{k-1}^{-1} \cdot \ldots \cdot \sigma_1^{-1}. \ \square$$

Для підстановки $\sigma \in S_n$ і цілого числа $k$ визначимо її $k$-й степінь:

$$\sigma^0 = \varepsilon, \qquad \sigma^k = \underbrace{\sigma \cdot \sigma \cdot \ldots \cdot \sigma}_{k}, \quad k \in \mathbb{N}, \qquad \sigma^{-k} = (\sigma^{-1})^k, \quad k \in \mathbb{N}.$$

Для степенів виконуються звичайні правила:

$$\sigma^m \cdot \sigma^k = \sigma^{m+k} \quad \text{та} \quad (\sigma^m)^k = \sigma^{mk}.$$

**Приклад 6.4.** Обчислимо степені підстановки $\sigma$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}, \quad \sigma^4 = \varepsilon.$$

## 6.3 Орбіти та цикли

Розглянемо, як підстановки діють на елементи множини при повторному застосуванні.

**Означення 6.4.** *Орбітою* елемента $i \in \{1, 2, \ldots, n\}$ відносно підстановки $\sigma \in S_n$ називається множина

$$\mathcal{O}_\sigma(i) = \{i, \sigma(i), \sigma^2(i), \sigma^3(i), \ldots\}.$$

**Приклад 6.5.** Розглянемо підстановку

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 1 & 7 & 9 & 8 & 4 & 2 & 10 & 6 & 5 \end{pmatrix}.$$

Простежимо орбіти:

$$1 \mapsto 3 \mapsto 7 \mapsto 2 \mapsto 1, \quad 4 \mapsto 9 \mapsto 6 \mapsto 4, \quad 5 \mapsto 8 \mapsto 10 \mapsto 5.$$

Отже,

$$\mathcal{O}_\sigma(1) = \{1, 2, 3, 7\}, \quad \mathcal{O}_\sigma(4) = \{4, 6, 9\}, \quad \mathcal{O}_\sigma(5) = \{5, 8, 10\}.$$

**Лема 6.1.** *Нехай $\sigma \in S_n$ та $i \in \{1, 2, \ldots, n\}$. Тоді існує найменше натуральне число $k \geq 1$ таке, що $\sigma^k(i) = i$. При цьому орбіта $\mathcal{O}_\sigma(i)$ складається рівно з $k$ різних елементів:*

$$\mathcal{O}_\sigma(i) = \{i, \sigma(i), \sigma^2(i), \ldots, \sigma^{k-1}(i)\}.$$

*Доведення.* Розглянемо послідовність

$$i, \sigma(i), \sigma^2(i), \ldots$$

Оскільки множина $\{1, 2, \ldots, n\}$ скінченна, у цій послідовності деякі елементи повторюються: існують $r < s$ такі, що $\sigma^r(i) = \sigma^s(i)$. Застосувавши $\sigma^{-r}$, дістаємо $\sigma^{s-r}(i) = i$. Нехай $k \geq 1$ — найменше число з властивістю $\sigma^k(i) = i$.

Покажемо, що елементи $i$, $\sigma(i)$, $\ldots$, $\sigma^{k-1}(i)$ попарно різні. Якщо $\sigma^r(i) = \sigma^s(i)$ для $0 \leq r < s < k$, то $\sigma^{s-r}(i) = i$, що суперечить мінімальності $k$. Отже, орбіта $\mathcal{O}_\sigma(i)$ складається з $k$ різних елементів $i$, $\sigma(i)$, $\ldots$, $\sigma^{k-1}(i)$. □

**Лема 6.2.** *Будь-які дві орбіти підстановки або не перетинаються, або збігаються. Зокрема, орбіти підстановки $\sigma \in S_n$ утворюють розбиття множини $\{1, 2, \ldots, n\}$.*

*Доведення.* Нехай орбіти $\mathcal{O}_\sigma(i)$ та $\mathcal{O}_\sigma(j)$ містять спільний елемент $l \in \mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j)$. Тоді $l = \sigma^r(i)$ та $l = \sigma^s(j)$ для деяких $r, s$, звідки $j = \sigma^{r-s}(i)$. Для довільного $\sigma^m(j) \in \mathcal{O}(j)$ маємо $\sigma^m(j) = \sigma^{m+r-s}(i) \in \mathcal{O}_\sigma(i)$, тому $\mathcal{O}_\sigma(j) \subseteq \mathcal{O}_\sigma(i)$. Аналогічно $\mathcal{O}_\sigma(i) \subseteq \mathcal{O}_\sigma(j)$, отже $\mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j)$.

Оскільки кожен елемент $i$ належить своїй орбіті $\mathcal{O}_\sigma(i)$, орбіти утворюють розбиття множини $\{1, 2, \ldots, n\}$. □

**Зауваження 6.2.** Орбіти можна також визначити через поняття відношення еквівалентності. Нехай $\sigma \in S_n$. Визначимо відношення на множині $\{1, 2, \ldots, n\}$ за правилом: $i \sim j$, якщо $\sigma^k(i) = j$ для деякого $k \in \mathbb{Z}$. Перевіримо, що це відношення еквівалентності:

- Рефлексивність: $\sigma^0(i) = i$, тому $i \sim i$.
- Симетричність: якщо $j = \sigma^k(i)$, то $i = \sigma^{-k}(j)$, отже з $i \sim j$ випливає $j \sim i$.
- Транзитивність: якщо $j = \sigma^k(i)$ та $l = \sigma^m(j)$, то $l = \sigma^{m+k}(i)$, отже з $i \sim j$ та $j \sim l$ випливає $i \sim l$.

Класи еквівалентності цього відношення збігаються з орбітами підстановки $\sigma$, а тому вони утворюють розбиття множини $\{1, 2, \ldots, n\}$.

Підстановка діє циклічно на кожній орбіті, а сукупність цих циклів однозначно визначає підстановку. Це дозволяє записувати підстановку у вигляді набору циклів.

**Приклад 6.6.** Підстановка

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 1 & 2 & 7 & 6 & 4 \end{pmatrix}$$

має орбіти $\{1, 3\}$, $\{2, 5, 7, 4\}$, $\{6\}$, на кожній з яких діє циклічно:

$$1 \mapsto 3 \mapsto 1, \quad 2 \mapsto 5 \mapsto 7 \mapsto 4 \mapsto 2, \quad 6 \mapsto 6.$$

Тому $\sigma$ однозначно задається набором циклів (1 3)(2 5 7 4)(6). Кожен з таких циклів можна розглядати як окрему підстановку, а весь запис — як добуток таких підстановок.

**Означення 6.5.** Підстановка називається *циклом*, якщо вона має рівно одну орбіту потужності більше за 1. Кількість елементів у цій орбіті називається *довжиною* циклу. Якщо довжина дорівнює $k$, то цикл називають *$k$-циклом*.

Іншими словами, підстановка $\sigma \in S_n$ є циклом довжини $k \geq 2$, якщо існують попарно різні елементи $a_1, a_2, \ldots, a_k \in \{1, 2, \ldots, n\}$ такі, що

$$\sigma(a_1) = a_2, \quad \sigma(a_2) = a_3, \quad \ldots, \quad \sigma(a_{k-1}) = a_k, \quad \sigma(a_k) = a_1,$$

а всі інші елементи залишаються нерухомими: $\sigma(i) = i$ виконується для $i \notin \{a_1, \ldots, a_k\}$. Такий цикл позначається $(a_1\ a_2\ \ldots\ a_k)$.

Цикл довжини 2, тобто вигляду $(i\ j)$, називається *транспозицією*. Цикл вигляду $(i\ \ i{+}1)$ називається *транспозицією сусідніх*.

**Приклад 6.7.** Розглянемо підстановки:

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 5 & 4 & 1 & 6 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 1 & 6 & 2 & 4 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 3 & 5 & 6 \end{pmatrix}.$$

Підстановка $\sigma_1$ має орбіти $\{1, 3, 5\}$ та $\{2\}$, $\{4\}$, $\{6\}$, тому вона є циклом $(1\ 3\ 5)$ довжини 3. Підстановка $\sigma_2$ має три орбіти $\{1, 3\}$, $\{2, 5\}$, $\{4, 6\}$ потужності більше за 1, тому вона не є циклом. Підстановка $\sigma_3$ має дві орбіти $\{1, 2\}$ та $\{3, 4\}$ потужності більше за 1, тому вона не є циклом.

**Приклад 6.8.** Цикл $(2\ 7\ 3\ 6\ 4) \in S_8$ представляє підстановку:

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

Оскільки цикли є підстановками, ми можемо їх множити; при цьому добуток циклів не обов'язково є циклом.

**Приклад 6.9.** Обчислимо добутки циклів на множині $\{1, 2, 3, 4, 5, 6\}$:

$$\sigma_1 = (1\ 3\ 5)(2\ 4) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 1 & 6 \end{pmatrix} = (1\ 3\ 5)(2\ 4),$$

$$\sigma_2 = (1\ 2\ 5)(2\ 3\ 6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 6 & 4 & 1 & 5 \end{pmatrix} = (1\ 2\ 3\ 6\ 5),$$

$$\sigma_3 = (1\ 2)(2\ 4)(3\ 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 1 & 3 & 6 \end{pmatrix} = (1\ 2\ 4)(3\ 5).$$

Підстановки $\sigma_1$ та $\sigma_3$ не є циклами, натомість $\sigma_2$ є циклом.

**Зауваження 6.3.** Запис циклу не є єдиним: циклічна перестановка елементів задає той самий цикл, тобто

$$(a_1\ a_2\ \dots\ a_k) = (a_2\ a_3\ \dots\ a_k\ a_1) = \dots = (a_k\ a_1\ \dots\ a_{k-1}).$$

**Означення 6.6.** Два цикли $(a_1\ a_2\ \dots\ a_k)$ і $(b_1\ b_2\ \dots\ b_m)$ називаються *незалежними*, якщо серед елементів $a_1, a_2, \dots, a_k$ та $b_1, b_2, \dots, b_m$ немає однакових.

**Зауваження 6.4.** Незалежні цикли комутують, тобто $\sigma\cdot\tau = \tau\cdot\sigma$. Справді, для елемента з орбіти $\sigma$ підстановка $\tau$ діє тривіально, тому результат визначається лише $\sigma$, і навпаки. Зокрема, порядок множників у добутку незалежних циклів не має значення.

**Твердження 6.4.** *Кожна нетотожна підстановка розкладається у добуток незалежних циклів, і цей розклад є єдиним з точністю до порядку множників.*

*Доведення. Існування.* Нехай $\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_s$ — орбіти підстановки $\sigma$ потужності більше за 1. Для кожної орбіти $\mathcal{O}_i$ визначимо цикл $\tau_i$, який діє на $\mathcal{O}_i$ так само, як $\sigma$, а всі інші елементи залишає нерухомими. Оскільки орбіти попарно не перетинаються, цикли $\tau_1, \dots, \tau_s$ є незалежними, і $\sigma = \tau_1 \cdot \tau_2 \cdots \tau_s$.

*Єдиність.* Орбіти підстановки $\sigma$ визначені однозначно, і кожен цикл у розкладі повністю визначається своєю орбітою та дією $\sigma$ на ній. Отже, розклад єдиний з точністю до порядку множників. □

**Приклад 6.10.** Розкладемо підстановку

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 7 & 1 & 9 & 5 & 2 & 6 & 10 & 4 & 8 \end{pmatrix}$$

у добуток незалежних циклів. Прослідкувавши орбіти, отримуємо:

$$\sigma = (1\ 3)(2\ 7\ 6)(4\ 9)(8\ 10).$$

Запис підстановок у добуток незалежних циклів дозволяє компактніше їх записувати. Наприклад, випишемо всі 24 елементи групи $S_4$ у циклічному вигляді:

$$\varepsilon,\ (1\,2),\ (1\,3),\ (1\,4),\ (2\,3),\ (2\,4),\ (3\,4), (1\,2\,3),\ (1\,3\,2),\ (1\,2\,4),$$
$$(1\,4\,2),\ (1\,3\,4),\ (1\,4\,3),\ (2\,3\,4),\ (2\,4\,3),\ (1\,2)(3\,4),\ (1\,3)(2\,4),\ (1\,4)(2\,3),$$
$$(1\,2\,3\,4),\ (1\,2\,4\,3),\ (1\,3\,2\,4),\ (1\,3\,4\,2),\ (1\,4\,2\,3),\ (1\,4\,3\,2).$$

**Теорема 6.1.** *Кожна підстановка розкладається у добуток транспозицій.*

*Доведення.* Оскільки будь-яка підстановка є добутком незалежних циклів, твердження достатньо довести для довільного циклу

$$\sigma = (a_1\ a_2\ \dots\ a_k)$$

. Безпосередньо перевіряється наступний розклад:

$$(a_1\ a_2\ \dots\ a_k) = (a_1\ a_2)(a_2\ a_3)\cdots(a_{k-1}\ a_k).$$

Справді, елемент $a_i$ при $i \neq k$ не змінюється транспозиціями $(a_{i+2}\ a_{i+3})$, $\dots$, $(a_{k-1}\ a_k)$, потім переходить у $a_{i+1}$ під дією транспозиції $(a_i\ a_{i+1})$, а всі наступні транспозиції залишають $a_{i+1}$ нерухомим. Отже, добуток транспозицій переводить $a_i$ у $a_{i+1}$. Елемент $a_k$ послідовно переходить у $a_{k-1}$, $a_{k-2}$, $\dots$, $a_1$ під дією транспозицій $(a_{k-1}\ a_k)$, $\dots$, $(a_1\ a_2)$, тобто в результаті $a_k$ переходить у $a_1$. Всі елементи поза орбітою залишаються нерухомими. □

**Приклад 6.11.** Розкладемо підстановку

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 7 & 1 & 10 & 9 & 3 & 6 & 5 & 8 & 4 \end{pmatrix} = (1\ 2\ 7\ 6\ 3)(4\ 10)(5\ 9\ 8)$$

у добуток транспозицій:

$$\sigma = (1\ 2)(2\ 7)(7\ 6)(6\ 3)(4\ 10)(5\ 9)(9\ 8).$$

Розклад у добуток транспозицій не є єдиним. Наприклад,

$$\sigma = (1\ 3)(1\ 6)(1\ 7)(1\ 2)(4\ 10)(5\ 8)(5\ 9).$$

**Теорема 6.2.** *Кожна підстановка розкладається у добуток транспозицій сусідніх.*

*Доведення.* За попередньою теоремою, достатньо показати, що кожна транспозиція розкладається у добуток транспозицій сусідніх. Для транспозиції $(i\ j)$ з $i < j$ маємо:

$$(i\ j) = (i\ i{+}1)(i{+}1\ i{+}2)\cdots(j{-}1\ j)\cdots(i{+}1\ i{+}2)(i\ i{+}1).$$

Справді, ця тотожність послідовно переміщує елемент $i$ вправо до позиції $j$, потім повертає елемент $j$ вліво до позиції $i$. □

**Приклад 6.12.** Розкладемо підстановку $(1\ 3\ 5)(2\ 4)$ у добуток транспозицій сусідніх. Спочатку розкладемо у добуток транспозицій:

$$(1\ 3\ 5)(2\ 4) = (1\ 3)(3\ 5)(2\ 4).$$

Далі розкладемо кожну транспозицію у добуток транспозицій сусідніх:

$$(1\ 3) = (1\ 2)(2\ 3)(1\ 2), \quad (3\ 5) = (3\ 4)(4\ 5)(3\ 4), \quad (2\ 4) = (2\ 3)(3\ 4)(2\ 3).$$

Отже,

$$(1\ 3\ 5)(2\ 4) = (1\ 2)(2\ 3)(1\ 2)(3\ 4)(4\ 5)(3\ 4)(2\ 3)(3\ 4)(2\ 3).$$

## 6.4 Парні та непарні підстановки

Хоча розклад підстановки у добуток транспозицій не є єдиним, парність кількості транспозицій у такому розкладі завжди залишається незмінною. Тому всі підстановки можна розділити на два класи: парні та непарні. Для обґрунтування цієї властивості введемо поняття інверсії підстановки.

**Означення 6.7.** *Інверсією* підстановки $\sigma$ називається впорядкована пара $(i, j)$, де $i, j \in \{1, 2, \ldots, n\}$ і виконуються умови: $i < j$ та $\sigma(i) > \sigma(j)$. Кількість інверсій підстановки $\sigma$ позначається через $\mathsf{inv}(\sigma)$.

**Приклад 6.13.** Знайдемо інверсії підстановки

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 1 & 8 & 2 & 5 & 4 & 7 \end{pmatrix}.$$

Перебираємо всі пари $i < j$ та перевіряємо, чи $\sigma(i) > \sigma(j)$:

$$\begin{array}{llll} (1,3):\ 3>1, & (1,5):\ 3>2, & (2,3):\ 6>1, & (2,5):\ 6>2, \\ (2,6):\ 6>5, & (2,7):\ 6>4, & (4,5):\ 8>2, & (4,6):\ 8>5, \\ (4,7):\ 8>4, & (4,8):\ 8>7, & (6,7):\ 5>4. & \end{array}$$

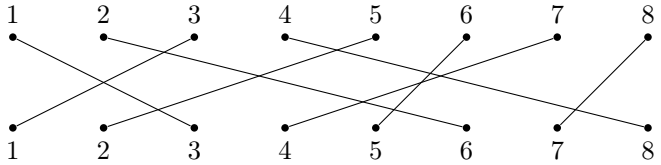
Отже, $\mathsf{inv}(\sigma) = 11$.

**Зауваження 6.5.** Кількість інверсій підстановки має наочну інтерпретацію через *діаграму підстановки*. Для підстановки $\sigma \in S_n$ розглянемо дві горизонтальні прямі, на кожній з яких позначено точки $1, 2, \ldots, n$. Для

кожного $i$ з'єднуємо точку $i$ у верхньому ряду з точкою $\sigma(i)$ у нижньому. Тоді кожне перетинання відрізків у цій діаграмі відповідає одній інверсії підстановки. Отже, кількість інверсій підстановки $\sigma$ дорівнює кількості перетинів у цій діаграмі. Наприклад, для підстановки

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 1 & 8 & 2 & 5 & 4 & 7 \end{pmatrix}$$

маємо таку діаграму:




Діаграма має 11 перетинів, що відповідає $\mathsf{inv}(\sigma) = 11$.

**Означення 6.8.** Підстановка $\sigma$ називається *парною* (відповідно *непарною*), якщо кількість її інверсій є парною (відповідно непарною), тобто $\mathsf{inv}(\sigma)$ є парним (відповідно непарним) числом.

*Знаком підстановки* $\sigma$ називається число

$$\mathsf{sgn}(\sigma) = (-1)^{\mathsf{inv}(\sigma)} = \begin{cases} +1, & \text{якщо } \sigma \text{ парна}, \\ -1, & \text{якщо } \sigma \text{ непарна}. \end{cases}$$

**Приклад 6.14.** Розглянемо підстановки:

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 5 & 1 & 6 & 2 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 1 & 6 & 3 & 5 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 3 & 2 & 1 \end{pmatrix}.$$

Оскільки $\mathsf{inv}(\sigma_1) = 8$, $\mathsf{inv}(\sigma_2) = 5$, $\mathsf{inv}(\sigma_3) = 12$, то $\sigma_2$ є непарною, а $\sigma_1$, $\sigma_3$ є парними.

**Лема 6.3.** *1. Кожна транспозиція є непарною підстановкою.*

*2. Множення підстановки на транспозицію змінює її парність на протилежну, тобто для всіх $\sigma \in S_n$ і будь-якої транспозиції $\tau \in S_n$ виконується*

$$\mathit{sgn}(\sigma \cdot \tau) = -\mathit{sgn}(\sigma).$$

*Доведення.* 1. Розглянемо транспозицію $(i\ j)$ з $i < j$:

$$(i\ j) = \begin{pmatrix} 1 & 2 & \cdots & i & i{+}1 & \cdots & j{-}1 & j & \cdots & n \\ 1 & 2 & \cdots & j & i{+}1 & \cdots & j{-}1 & i & \cdots & n \end{pmatrix}.$$

Інверсіями $(i\ j)$ є такі пари:

- пара $(i, j)$, оскільки $i < j$, але $j > i$;
- пари $(i, k)$ для кожного $k$ з $i < k < j$, оскільки $i < k$, але $j > k$;
- пари $(k, j)$ для кожного $k$ з $i < k < j$, оскільки $k < j$, але $k > i$.

Таким чином, кількість інверсій дорівнює $1+(j-i-1)+(j-i-1) = 2(j-i)-1$, що є непарним числом.

2. Нехай $\tau = (i\ j)$ з $i < j$. Тоді підстановка $\sigma \cdot \tau$ отримується переставлянням місцями позицій $i$ та $j$ у нижньому рядку підстановки $\sigma$:

$$\sigma = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix},$$
$$\sigma \cdot \tau = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix}.$$

Розглянемо, як при цьому змінюється кількість інверсій:

- Пара $(i, j)$ змінює свій статус: якщо $a_i > a_j$, то після перестановки маємо $a_j < a_i$, і навпаки. Тобто ця пара або додає одну інверсію, або прибирає одну.
- Для кожного $k \notin \{i, j\}$ розглянемо пари, що містять $i$ або $j$. Якщо $k < i$ або $k > j$, то пари $(k, i)$ і $(k, j)$ або $(i, k)$ і $(j, k)$ просто міняються місцями, і кількість інверсій серед них не змінюється.

  Якщо $i < k < j$, то пари $(i, k)$ і $(k, j)$ змінюються так, що рівно одна з них була інверсією до перестановки тоді і тільки тоді, коли рівно одна є інверсією після перестановки. Тому загальна кількість інверсій серед цих пар також не змінюється.
- Усі інші пари не змінюються.

Отже, кількість інверсій змінюється на одиницю і $\mathsf{sgn}(\sigma \cdot \tau) = -\mathsf{sgn}(\sigma)$.

□

**Наслідок 6.1.** *Якщо підстановка $\sigma$ розкладається у добуток $k$ транспозицій, то $\mathsf{sgn}(\sigma) = (-1)^k$. Зокрема, парність кількості транспозицій у будь-якому розкладі підстановки є однаковою.*

*Доведення.* Нехай $\sigma = \tau_1 \cdot \tau_2 \cdot \ldots \cdot \tau_k$ — розклад у добуток транспозицій. Застосовуємо лему послідовно $k$ разів:

$$\begin{aligned}\mathsf{sgn}(\sigma) &= \mathsf{sgn}(\tau_1 \cdot \tau_2 \cdot \ldots \cdot \tau_k) = (-1) \cdot \mathsf{sgn}(\tau_1 \cdot \tau_2 \cdot \ldots \cdot \tau_{k-1}) = \ldots = \\ &= (-1)^{k-1} \cdot \mathsf{sgn}(\tau_1) = (-1)^k.\end{aligned}$$

□

**Наслідок 6.2.** *$k$-цикл є парною підстановкою тоді і тільки тоді, коли $k$ є непарним.*

*Доведення.* Будь-який $k$-цикл розкладається у добуток $k-1$ транспозицій:

$$\sigma = (a_1\ a_2\ \ldots\ a_k) = (a_1\ a_2)(a_2\ a_3)\cdots(a_{k-1}\ a_k).$$

Отже, $\mathsf{sgn}(\sigma) = (-1)^{k-1}$, що дорівнює $+1$ тоді і тільки тоді, коли $k-1$ є парним, тобто $k$ є непарним. □

**Твердження 6.5** (властивості знаку підстановки)**.** *Для всіх підстановок $\sigma, \tau \in S_n$ справедливі такі властивості:*

1. $\mathsf{sgn}(\sigma \cdot \tau) = \mathsf{sgn}(\sigma) \cdot \mathsf{sgn}(\tau)$;
2. $\mathsf{sgn}(\sigma^{-1}) = \mathsf{sgn}(\sigma)$.

*Доведення.* 1. Нехай $\tau = \tau_1 \cdot \tau_2 \cdot \ldots \cdot \tau_k$ — розклад у добуток транспозицій, тоді $\mathsf{sgn}(\tau) = (-1)^k$. За лемою, кожне множення на транспозицію змінює знак, тому

$$\mathsf{sgn}(\sigma \cdot \tau) = \mathsf{sgn}(\sigma \cdot \tau_1 \cdot \tau_2 \cdot \ldots \cdot \tau_k) = \mathsf{sgn}(\sigma) \cdot (-1)^k = \mathsf{sgn}(\sigma) \cdot \mathsf{sgn}(\tau).$$

2. З пункту 1 маємо $\mathsf{sgn}(\sigma) \cdot \mathsf{sgn}(\sigma^{-1}) = \mathsf{sgn}(\sigma \cdot \sigma^{-1}) = \mathsf{sgn}(\varepsilon) = 1$, оскільки тотожна підстановка має нуль інверсій. Тому $\mathsf{sgn}(\sigma^{-1}) = \mathsf{sgn}(\sigma)$. □

Для встановлення парності підстановки не обов'язково підраховувати інверсії чи розкладати її у добуток транспозицій — достатньо знайти її циклічний розклад і використати наступний критерій.

**Наслідок 6.3** (Критерій парності підстановки)**.** *Підстановка $\sigma \in S_n$ є парною тоді і лише тоді, коли у її розкладі в добуток незалежних циклів міститься парна кількість циклів парної довжини.*

*Доведення.* Нехай $\sigma = \sigma_1 \cdot \sigma_2 \cdot \ldots \cdot \sigma_k$ — розклад у добуток незалежних циклів. За наслідком 6.2, кожний цикл парної довжини має знак $-1$, а непарної — $+1$. Тоді

$$\mathsf{sgn}(\sigma) = \mathsf{sgn}(\sigma_1) \cdot \mathsf{sgn}(\sigma_2) \cdot \ldots \cdot \mathsf{sgn}(\sigma_k) = (-1)^l,$$

де $l$ — кількість циклів парної довжини. Отже, $\sigma$ є парною тоді і тільки тоді, коли $l$ є парним. □

**Приклад 6.15.** Визначимо парність підстановки

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 7 & 1 & 9 & 5 & 2 & 6 & 10 & 4 & 8 \end{pmatrix} = (1\ 3)(2\ 7\ 6)(4\ 9)(8\ 10).$$

Цикли мають довжини $2, 3, 2, 2$, серед яких три цикли парної довжини. Оскільки 3 є непарним числом, підстановка $\sigma$ є непарною. Підстановка

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 7 & 8 & 2 & 10 & 5 & 4 & 1 & 9 & 6 \end{pmatrix} = (1\ 3\ 8)(2\ 7\ 4)(5\ 10\ 6)$$

має нуль циклів парної довжини. Оскільки нуль є числом парним, підстановка $\tau$ є парною.

Множина всіх парних підстановок на множині з $n$ елементів позначається $A_n$:

$$A_n = \{\sigma \in S_n : \sigma \text{ є парною}\} = \{\sigma \in S_n : \mathsf{sgn}(\sigma) = 1\}.$$

**Твердження 6.6** (властивості парних підстановок)**.** *Множина $A_n$ є підгрупою групи $S_n$, яка називається* знакозмінною групою степеня $n$. *Це означає, що виконуються такі властивості:*

1. *Замкненість: якщо $\sigma, \tau \in A_n$, то $\sigma \cdot \tau \in A_n$.*
2. *Оберненість: якщо $\sigma \in A_n$, то $\sigma^{-1} \in A_n$.*

*Крім того, при $n \geq 2$ група $A_n$ містить рівно $n!/2$ елементів.*

*Доведення.* Властивості 1 та 2 випливають з властивостей знаку:

$$\mathsf{sgn}(\sigma \cdot \tau) = \mathsf{sgn}(\sigma) \cdot \mathsf{sgn}(\tau) = 1 \cdot 1 = 1, \qquad \mathsf{sgn}(\sigma^{-1}) = \mathsf{sgn}(\sigma) = 1.$$

Для підрахунку $|A_n|$ розглянемо відображення

$$\varphi\colon S_n \mapsto S_n, \quad \varphi(\sigma) = (1\ 2) \cdot \sigma.$$

Відображення $\varphi$ є оберненим до самого себе:

$$\varphi(\varphi(\sigma)) = (1\ 2)\cdot((1\ 2)\cdot\sigma) = ((1\ 2)\cdot(1\ 2))\cdot\sigma = \varepsilon\cdot\sigma = \sigma.$$

Тому $\varphi$ є бієкцією. За лемою 6.3, $\varphi$ переводить парні підстановки в непарні і навпаки. Отже, кількість парних і непарних підстановок однакова і дорівнює $n!/2$. □

Знакозмінна група $A_4$ містить $4!/2 = 12$ елементів:

$$A_4 = \{\varepsilon,\ (1\,2\,3),\ (1\,3\,2),\ (1\,2\,4),\ (1\,4\,2),\ (1\,3\,4),\ (1\,4\,3),\\ (2\,3\,4),\ (2\,4\,3),\ (1\,2)(3\,4),\ (1\,3)(2\,4),\ (1\,4)(2\,3)\}.$$

## 6.5 Вправи

1. Знайдіть добуток підстановок:

   а) $\begin{pmatrix}1&2&3&4&5&6\\5&3&6&1&2&4\end{pmatrix}\cdot\begin{pmatrix}1&2&3&4&5&6\\2&4&1&6&5&3\end{pmatrix}$;

   б) $\begin{pmatrix}1&2&3&4&5\\4&3&5&1&2\end{pmatrix}\cdot\begin{pmatrix}1&2&3&4&5\\2&4&1&5&3\end{pmatrix}\cdot\begin{pmatrix}1&2&3&4&5\\3&5&2&4&1\end{pmatrix}$;

   в) $\begin{pmatrix}1&2&3&4&5&6&7&8\\2&5&1&8&3&7&4&6\end{pmatrix}\cdot\begin{pmatrix}1&2&3&4&5&6&7&8\\4&3&2&7&1&8&5&6\end{pmatrix}$.

2. Опишіть орбіти підстановки та розкладіть її у добуток незалежних циклів:

   а) $\begin{pmatrix}1&2&3&4&5&6&7&8\\3&6&1&8&2&5&4&7\end{pmatrix}$; в) $\begin{pmatrix}1&2&3&4&5&6&7&8\\2&1&4&3&6&5&8&7\end{pmatrix}$;

   б) $\begin{pmatrix}1&2&3&4&5&6&7&8\\5&8&4&3&7&1&6&2\end{pmatrix}$; г) $\begin{pmatrix}1&2&3&4&5&6&7&8\\4&3&2&5&1&8&7&6\end{pmatrix}$.

3. Випишіть всі інверсії підстановки та визначіть її парність:

   а) $\begin{pmatrix}1&2&3&4&5&6\\4&1&5&2&6&3\end{pmatrix}$;  в) $\begin{pmatrix}1&2&3&4&5&6&7&8\\5&2&7&1&8&3&6&4\end{pmatrix}$;

   б) $\begin{pmatrix}1&2&3&4&5&6\\3&6&2&5&1&4\end{pmatrix}$;  г) $\begin{pmatrix}1&2&3&4&5&6&7&8\\7&5&3&1&2&4&6&8\end{pmatrix}$.

4. Знайдіть парність підстановки, розклавши її у добуток циклів:

   а) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 6 & 4 & 1 & 2 & 7 & 3 & 5 & 9 & 8 \end{pmatrix}$; б) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 1 & 4 & 5 & 3 & 8 & 9 & 7 & 6 \end{pmatrix}$.

5. Доведіть, що будь-яка підстановка $\sigma \in S_n$ може бути подана як добуток не більше ніж $n-1$ транспозицій.

6. Доведіть, що кожна підстановка з $S_n$ розкладається у добуток транспозицій вигляду $(1\,2), (1\,3), \ldots, (1\,n)$.

7. Доведіть, що кожна парна підстановка розкладається у добуток циклів довжини три.

8. Доведіть, що для будь-яких двох підстановок $\sigma, \tau \in S_n$ підстановки $\tau$ та $\sigma\tau\sigma^{-1}$ завжди мають однакову парність.

9. Доведіть, що для будь-яких двох підстановок $\sigma, \tau \in S_n$ їхній комутатор $[\sigma, \tau] = \sigma\tau\sigma^{-1}\tau^{-1}$ завжди є парною підстановкою.

10. Доведіть, що кожна підстановка з $S_n$ може бути подана як добуток лише двох підстановок: транспозиції $(1\,2)$ та циклу $(1\,2\,3\,\ldots\,n)$.

11. Доведіть, що при $n \geq 3$ єдиною підстановкою, яка комутує з усіма іншими підстановками (тобто $\sigma\tau = \tau\sigma$ для будь-якої $\tau \in S_n$), є лише тотожна підстановка $\varepsilon$.

# Розділ 7

# Визначники

## 7.1 Означення визначника

Кожній квадратній матриці можна поставити у відповідність число, яке несе важливу геометричну та алгебраїчну інформацію про матрицю.

**Означення 7.1.** *Визначником* квадратної матриці $A = (a_{ij})$ порядку $n$ називається число

$$\det A = |A| = \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)} =$$

$$= \sum_{\left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{smallmatrix}\right) \in S_n} \mathsf{sgn} \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} a_{1i_1}\, a_{2i_2} \cdots a_{ni_n}.$$

У цій сумі $n!$ доданків, кожен з яких є добутком $n$ елементів матриці, по одному з кожного рядка і кожного стовпця, взятим зі знаком $\mathsf{sgn}(\sigma)$.

**Приклад 7.1.** Розглянемо означення визначника для матриць порядку 2. Група $S_2$ містить дві підстановки:

$$\varepsilon = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad \mathsf{sgn}(\varepsilon) = 1 \quad \text{та} \quad \sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \mathsf{sgn}(\sigma) = -1.$$

Тоді формула для визначника набуває вигляду:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \mathsf{sgn}(\varepsilon)\, a_{1\varepsilon(1)} a_{2\varepsilon(2)} + \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} = a_{11}a_{22} - a_{12}a_{21}.$$

Наприклад,

$$\begin{vmatrix} 3 & 5 \\ -2 & 4 \end{vmatrix} = 3 \cdot 4 - 5 \cdot (-2) = 12 + 10 = 22.$$

**Приклад 7.2.** Розглянемо визначник матриць порядку 3. Група $S_3$ містить шість підстановок:

$$\begin{aligned} \mathsf{sgn} = +1 \qquad & \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \\ \mathsf{sgn} = -1 \qquad & \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}. \end{aligned}$$

Тому формула для визначника матриці $3 \times 3$ має вигляд:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Наприклад,

$$\begin{vmatrix} 1 & -2 & 3 \\ 4 & -2 & 5 \\ -3 & 7 & 1 \end{vmatrix} = 1 \cdot (-2) \cdot 1 + (-2) \cdot 5 \cdot (-3) + 3 \cdot 4 \cdot 7 - 3 \cdot (-2) \cdot (-3) - 1 \cdot 5 \cdot 7 - (-2) \cdot 4 \cdot 1 = = -2 + 30 + 84 - 18 - 35 + 8 = 67.$$

## 7.2 Властивості визначників

Обчислення визначника за означенням для великих $n$ є нереалістичним, оскільки кількість доданків $n!$ швидко зростає. Наприклад, для матриці $10 \times 10$ їх $10! \approx 3{,}6$ мільйона. У цьому розділі ми встановимо основні властивості визначника, які дозволяють обчислювати його значно ефективніше. Ключова ідея полягає у зведенні матриці до трикутного вигляду за допомогою елементарних перетворень, після чого визначник обчислюється тривіально.

**Твердження 7.1.** *Визначник верхньої (або нижньої) трикутної матри-*

*ці дорівнює добутку її діагональних елементів:*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

*Доведення.* Розглянемо означення визначника:

$$\det A = \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Для верхньої трикутної матриці елементи $a_{ij} = 0$ при $i > j$. Тому доданок, що відповідає підстановці $\sigma$, є ненульовим лише тоді, коли $\sigma(i) \geq i$ для всіх $i$. Єдина підстановка з цією властивістю — тотожна $\varepsilon$. Отже,

$$\det A = \mathsf{sgn}(\varepsilon)\, a_{11}\, a_{22} \cdots a_{nn} = a_{11}a_{22}\cdots a_{nn}. \qquad \square$$

**Твердження 7.2.** *Нехай квадратна матриця $A$ має блочно-трикутний вигляд*

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

*де $B$ і $C$ — квадратні матриці. Тоді* $\det A = \det B \cdot \det C$.

*Доведення.* Нехай матриці $B$ і $C$ мають розміри $k \times k$ та $l \times l$ відповідно, де $l = n - k$. Оскільки $a_{ij} = 0$ при $i > k$ та $j \leq k$, то кожен доданок у сумі в означенні визначника, що відповідає підстановці $\sigma$, є нульовим, якщо $\sigma(i) > k$ для деякого $i \leq k$, або $\sigma(i) \leq k$ для деякого $i > k$. Тому достатньо розглядати $\sigma$ вигляду:

$$\sigma = \begin{pmatrix} 1 & \cdots & k & k+1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(k) & \sigma(k+1) & \cdots & \sigma(n) \end{pmatrix},$$

де $\sigma(1), \ldots, \sigma(k)$ є перестановкою чисел $1, 2, \ldots, k$, а $\sigma(k+1), \ldots, \sigma(n)$ є перестановкою чисел $k+1, k+2, \ldots, n$. Це дозволяє записати $\sigma$ у вигляді добутку $\sigma = \sigma_B \cdot \sigma_C$, де $\sigma_B$ діє тотожно на $\{k+1, \ldots, n\}$ і $\sigma_C$ діє тотожно

на $\{1,\dots,k\}$. При цьому $\mathsf{sgn}(\sigma) = \mathsf{sgn}(\sigma_B)\cdot\mathsf{sgn}(\sigma_C)$. Тоді

$$\begin{aligned}\det A &= \sum_{\substack{\sigma_B\in S_k,\\ \sigma_C\in S_l}} \mathsf{sgn}(\sigma_B)\,\mathsf{sgn}(\sigma_C)\,a_{1\sigma_B(1)}\cdots a_{k\sigma_B(k)}\,a_{k+1,\sigma_C(k+1)}\cdots a_{n\sigma_C(n)} = \\ &= \left(\sum_{\sigma_B\in S_k} \mathsf{sgn}(\sigma_B)\,a_{1\sigma_B(1)}\cdots a_{k\sigma_B(k)}\right)\times \\ &\qquad \times\left(\sum_{\sigma_C\in S_l} \mathsf{sgn}(\sigma_C)\,a_{k+1,\sigma_C(k+1)}\cdots a_{n\sigma_C(n)}\right) = \\ &= \det B\cdot\det C. \end{aligned}$$

□

**Наслідок 7.1.** *Нехай квадратна матриця $A$ має блочно-трикутний вигляд*

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ 0 & A_{22} & \cdots & A_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{mm} \end{pmatrix},$$

*де $A_{11}, A_{22}, \dots, A_{mm}$ — квадратні матриці. Тоді*

$$\det A = \det A_{11}\cdot\det A_{22}\cdot\ldots\cdot\det A_{mm}.$$

Перш ніж розглянути, як змінюється визначник при елементарних перетвореннях, встановимо, що всі властивості, сформульовані для рядків, автоматично переносяться на стовпці. Перехід від рядків до стовпців відповідає транспонуванню матриці, тому достатньо довести таке твердження.

**Твердження 7.3.** *Для будь-якої квадратної матриці $A$ виконується*

$$\det A^T = \det A.$$

*Доведення.* Нехай $A = (a_{ij})$ та $A^T = (b_{ij})$, де за означенням транспонування $b_{ij} = a_{ji}$. Запишемо визначник матриці $A^T$:

$$\begin{aligned}\det A^T &= \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)}\cdots b_{n\sigma(n)} = \\ &= \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2}\cdots a_{\sigma(n)n}.\end{aligned}$$

Зробимо заміну $\tau = \sigma^{-1}$. Оскільки $\sigma$ пробігає всю групу $S_n$, то і $\tau$ пробігає всю групу $S_n$. Крім того, $\mathsf{sgn}(\tau) = \mathsf{sgn}(\sigma)$. Оскільки $\sigma(i) = j \Leftrightarrow i = \tau(j)$, маємо:

$$a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n} = a_{1\tau(1)}a_{2\tau(2)}\cdots a_{n\tau(n)}.$$

Отже,

$$\det A^T = \sum_{\tau\in S_n} \mathrm{sgn}(\tau)a_{1\tau(1)}a_{2\tau(2)}\cdots a_{n\tau(n)} = \det A. \qquad \square$$

При розв'язанні систем лінійних рівнянь метод Гаусса дозволяв зводити матрицю до ступінчастої форми без зміни множини розв'язків. На відміну від розв'язків системи, визначник матриці може змінюватися при таких перетвореннях. Проте ці зміни відбуваються за чіткими правилами, які ми далі встановимо.

**Твердження 7.4.** *При перестановці двох рядків квадратної матриці визначник змінює знак.*

*Доведення.* Нехай $B$ отримана з $A$ перестановкою $i$-го та $j$-го рядків. Тоді $b_{ik} = a_{jk}$ та $b_{jk} = a_{ik}$ для всіх $k$, а всі інші елементи незмінні. Запишемо визначник $B$:

$$\begin{aligned}\det B &= \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma)\, b_{1\sigma(1)}\cdots b_{i\sigma(i)}\cdots b_{j\sigma(j)}\cdots b_{n\sigma(n)} = \\ &= \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)}\cdots a_{j\sigma(i)}\cdots a_{i\sigma(j)}\cdots a_{n\sigma(n)}.\end{aligned}$$

Зробимо заміну $\tau = \sigma\cdot(i\ j)$. Оскільки $\sigma(i) = \tau(j)$ та $\sigma(j) = \tau(i)$, а для $k \notin \{i, j\}$ маємо $\sigma(k) = \tau(k)$, то:

$$\begin{aligned}\det B &= \sum_{\tau\in S_n} \mathsf{sgn}(\tau\cdot(i\ j))\, a_{1\tau(1)}\cdots a_{i\tau(i)}\cdots a_{j\tau(j)}\cdots a_{n\tau(n)} = \\ &= -\sum_{\tau\in S_n} \mathsf{sgn}(\tau)\, a_{1\tau(1)}\cdots a_{n\tau(n)} = -\det A.\end{aligned}$$

де ми використали $\mathsf{sgn}(\sigma) = \mathsf{sgn}(\tau\cdot(i\ j)) = -\mathsf{sgn}(\tau)$. $\square$

**Наслідок 7.2.** *Якщо квадратна матриця має два однакових рядки, то її визначник дорівнює нулю.*

*Доведення.* Нехай $i$-й та $j$-й рядки матриці $A$ однакові. Переставимо ці два рядки місцями — отримаємо ту саму матрицю $A$. За твердженням, $\det A = -\det A$, звідси $\det A = 0$. □

**Твердження 7.5.** *При множенні одного рядка квадратної матриці на число $\lambda$ визначник множиться на $\lambda$:*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & \cdots & \lambda a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

*Доведення.* Нехай $B$ отримана з $A$ множенням $i$-го рядка на $\lambda$. Тоді $b_{ik} = \lambda a_{ik}$ для всіх $k$, а решта елементів незмінні. Кожен доданок у визначнику містить рівно один елемент $i$-го рядка, тому

$$\det B = \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} \cdots \lambda a_{i\sigma(i)} \cdots a_{n\sigma(n)} = \lambda \det A. \qquad \square$$

**Твердження 7.6.** *При додаванні до одного рядка квадратної матриці іншого рядка, помноженого на число, визначник не змінюється.*

*Доведення.* Нехай $B$ отримана з $A$ додаванням до $i$-го рядка $j$-го рядка, помноженого на число $\lambda$. Тоді $b_{ik} = a_{ik} + \lambda a_{jk}$ для всіх $k$, а решта елементів незмінні. Запишемо визначник:

$$\begin{aligned} \det B &= \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} \cdots (a_{i\sigma(i)} + \lambda a_{j\sigma(i)}) \cdots a_{n\sigma(n)} = \\ &= \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} + \\ &\qquad + \lambda \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(i)} \cdots a_{n\sigma(n)}. \end{aligned}$$

Перша сума дорівнює $\det A$. У другій сумі $i$-й та $j$-й елементи однакові, тому вона є визначником матриці з двома однаковими рядками. За наслідком 7.2 цей визначник дорівнює нулю. Отже, $\det B = \det A$. □

Ефективний спосіб обчислення визначника полягає у застосуванні *методу Гаусса*: зводимо матрицю до верхньої трикутної форми елементарними перетвореннями над рядками і стовпцями, відстежуючи зміни визначника:

- перестановка двох рядків або стовпців змінює його знак;
- множення рядка або стовпця на $\lambda$ множить визначник на $\lambda$;
- додавання до одного рядка іншого, помноженого на число, не змінює визначника.

Після зведення визначник трикутної матриці обчислюється як добуток діагональних елементів.

**Приклад 7.3.** Обчислимо визначник порядку 4 методом Гаусса:

$$\begin{vmatrix} 2 & 1 & -1 & 3 \\ 4 & 2 & 1 & 5 \\ -2 & 3 & 2 & 1 \\ 6 & 1 & -2 & 4 \end{vmatrix} \underset{\substack{R_3 \to R_3 + R_1 \\ R_4 \to R_4 - 3R_1}}{\overset{R_2 \to R_2 - 2R_1}{=}} \begin{vmatrix} 2 & 1 & -1 & 3 \\ 0 & 0 & 3 & -1 \\ 0 & 4 & 1 & 4 \\ 0 & -2 & 1 & -5 \end{vmatrix} \underset{R_2 \leftrightarrow R_3}{=}$$

$$= -\begin{vmatrix} 2 & 1 & -1 & 3 \\ 0 & 4 & 1 & 4 \\ 0 & 0 & 3 & -1 \\ 0 & -2 & 1 & -5 \end{vmatrix} \underset{R_4 \to R_4 + \frac{1}{2}R_2}{=} -\begin{vmatrix} 2 & 1 & -1 & 3 \\ 0 & 4 & 1 & 4 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & \frac{3}{2} & -3 \end{vmatrix} \underset{R_4 \to R_4 - \frac{1}{2}R_3}{=}$$

$$= -\begin{vmatrix} 2 & 1 & -1 & 3 \\ 0 & 4 & 1 & 4 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -\frac{5}{2} \end{vmatrix} = -2 \cdot 4 \cdot 3 \cdot \left(-\tfrac{5}{2}\right) = 60.$$

Встановлені властивості визначника при елементарних перетвореннях є окремими випадками більш загальної властивості — полілінійності.

**Твердження 7.7** (полілінійність визначника)**.** *Визначник матриці є лінійною функцією від кожного рядка (та стовпця). Це означає, що якщо $i$-й рядок $R_i$ квадратної матриці $A$ є лінійною комбінацією векторів:*

$$R_i = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k,$$

*то визначник матриці $A$ дорівнює лінійній комбінації відповідних визначників:*

$$\det A = \lambda_1 \det B_1 + \lambda_2 \det B_2 + \cdots + \lambda_k \det B_k,$$

*де кожна матриця $B_j$ отримана з матриці $A$ шляхом заміни її $i$-го рядка на вектор $v_j$ при збереженні всіх інших рядків незмінними.*

*Доведення.* За означенням визначника,

$$\det A = \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} \cdots \left( \sum_{j=1}^{k} \lambda_j (v_j)_{\sigma(i)} \right) \cdots a_{n\sigma(n)} =$$

$$= \sum_{j=1}^{k} \lambda_j \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma)\, a_{1\sigma(1)} \cdots (v_j)_{\sigma(i)} \cdots a_{n\sigma(n)} = \sum_{j=1}^{k} \lambda_j \det B_j. \ \square$$

**Зауваження 7.1.** Можна показати, що існує єдина функція від рядків квадратної матриці, яка є полілінійною, кососиметричною та набуває значення 1 на одиничній матриці. Цією функцією є визначник. Тому в деяких підручниках визначник вводять саме через ці три властивості, а формулу через підстановки отримують як наслідок.

**Приклад 7.4.** Проілюструємо полілінійність визначника. Нехай

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Запишемо перший рядок як лінійну комбінацію:

$$(1,2) = 3 \cdot (1,0) + 1 \cdot (-2,2).$$

Тоді за полілінійністю:

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 3 \begin{vmatrix} 1 & 0 \\ 3 & 4 \end{vmatrix} + 1 \begin{vmatrix} -2 & 2 \\ 3 & 4 \end{vmatrix} = 3 \cdot 4 + 1 \cdot (-8 - 6) = 12 - 14 = -2.$$

Це збігається з безпосереднім обчисленням: $\det A = 1 \cdot 4 - 2 \cdot 3 = -2$.

**Приклад 7.5.** Обчислимо визначник

$$\Delta_n = \begin{vmatrix} x_1 + y_1 & x_1 + y_2 & \cdots & x_1 + y_n \\ x_2 + y_1 & x_2 + y_2 & \cdots & x_2 + y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n + y_1 & x_n + y_2 & \cdots & x_n + y_n \end{vmatrix}.$$

Запишемо $i$-й рядок як суму $R_i = x_i \cdot (1,1,\ldots,1) + (y_1, y_2, \ldots, y_n)$. За полілінійністю визначник розкладається у суму $2^n$ визначників, у кожному з яких $i$-й рядок є або $x_i \cdot (1,1,\ldots,1)$, або $(y_1, y_2, \ldots, y_n)$. Кожний

визначник має два однакових рядки при $n \geq 3$. Оскільки визначник з двома однаковими рядками дорівнює нулю, то $\Delta_n = 0$ при $n \geq 3$. Окремо обчислюємо

$$\Delta_2 = \begin{vmatrix} x_1 + y_1 & x_1 + y_2 \\ x_2 + y_1 & x_2 + y_2 \end{vmatrix} = \\ = (x_1 + y_1)(x_2 + y_2) - (x_1 + y_2)(x_2 + y_1) = (x_1 - x_2)(y_2 - y_1).$$

Встановимо зв'язок між визначником матриці та основними алгебраїчними операціями над матрицями.

**Твердження 7.8.** *Елементарні матриці мають такі визначники:*

$$\det P_{ij} = -1; \qquad \det E_i(\lambda) = \lambda; \qquad \det E_{ij}(\lambda) = 1.$$

*Доведення.* Кожна елементарна матриця отримується з одиничної матриці $E$ одним елементарним перетворенням. Оскільки $\det E = 1$, то твердження випливає безпосередньо з властивостей зміни визначника при елементарних перетвореннях. □

**Теорема 7.1** (визначник добутку матриць)**.** *Для будь-яких квадратних матриць $A$ і $B$ розміру $n \times n$ виконується*

$$\det(AB) = \det A \cdot \det B.$$

*Доведення.* Спочатку доведемо твердження для випадку, коли $A$ є елементарною матрицею. Множення зліва (справа) на елементарну матрицю рівносильне виконанню відповідного елементарного перетворення рядків (стовпців) матриці $B$. За властивостями визначника при елементарних перетвореннях, маємо $\det(AB) = \det A \cdot \det B$.

Розглянемо два випадки залежно від того, чи матриця $A$ є оборотною. Якщо $A$ є оборотною, то за теоремою 5.4 вона розкладається у добуток елементарних матриць $A = E_1 E_2 \cdots E_k$. Послідовно застосовуючи доведений частковий випадок, отримуємо:

$$\det(A) = \det(E_1 \cdots E_k) = \det E_1 \cdots \det E_k,$$
$$\det(AB) = \det(E_1 \cdots E_k B) = \det E_1 \cdots \det E_k \cdot \det B = \\ = \det(E_1 \cdots E_k) \cdot \det B = \det A \cdot \det B.$$

Якщо $A$ не є оборотною, то її ступінчаста форма має нульовий рядок. Тому існують елементарні матриці $E_1, \ldots, E_k$ такі, що матриця

$\tilde{A} = E_k \cdots E_1 A$ має нульовий рядок. Тоді матриця $\tilde{A}B = E_k \cdots E_1(AB)$ також має нульовий рядок за правилом множення матриць. Тому матриці $\tilde{A}$ та $\tilde{A}B$ мають нульові визначники. Послідовно застосовуючи доведений частковий випадок, отримуємо

$$\det \tilde{A} = \det E_k \cdots \det E_1 \cdot \det A = 0,$$
$$\det(\tilde{A}B) = \det E_k \cdots \det E_1 \cdot \det(AB) = 0.$$

Оскільки $\det E_i \neq 0$ для кожного $i$, маємо $\det(A) = 0$ і $\det(AB) = 0 = \det A \cdot \det B$. □

**Приклад 7.6.** Обчислимо визначник матриці $A$ порядку $n$, де $a_{ij} = ij+1$:

$$A = \begin{pmatrix} 1\cdot 1+1 & 1\cdot 2+1 & \dots & 1\cdot n+1 \\ 2\cdot 1+1 & 2\cdot 2+1 & \dots & 2\cdot n+1 \\ \vdots & \vdots & \ddots & \vdots \\ n\cdot 1+1 & n\cdot 2+1 & \dots & n\cdot n+1 \end{pmatrix}.$$

Матрицю $A$ можна подати у вигляді добутку двох матриць порядку $n$:

$$A = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 2 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & 1 & 0 & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

При $n \geq 3$ кожна з цих матриць має нульовий стовпець (рядок), тому обидва визначники рівні нулю і $\det A = 0$. Для $n = 1$ маємо $\det A = 2$, а для $n = 2$ обчислюємо:

$$\det \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = 10 - 9 = 1.$$

**Зауваження 7.2.** Зазначимо, що визначник суми матриць, взагалі кажучи, не дорівнює сумі визначників. Наприклад, для матриць

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

маємо $\det A = 0$, $\det B = 0$, тому $\det A + \det B = 0$, проте

$$\det(A + B) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0.$$

**Теорема 7.2** (критерій оборотності матриці)**.** *Квадратна матриця $A$ є оборотною тоді і тільки тоді, коли* $\det A \neq 0$*, причому в цьому випадку*

$$\det A^{-1} = \frac{1}{\det A}.$$

*Доведення.* Якщо $A$ не є оборотною, то $\det A = 0$ (див. доведення теореми 7.1). Якщо $A$ є оборотною, то з рівності $AA^{-1} = E$ за теоремою про визначник добутку матриць маємо:

$$\det A \cdot \det A^{-1} = \det E = 1.$$

Отже, $\det A \neq 0$ і

$$\det A^{-1} = \frac{1}{\det A}. \qquad \square$$

## 7.3 Розклад визначника за рядком

У попередньому розділі ми розглянули метод обчислення визначника шляхом зведення матриці до ступінчастої форми. Розглянемо інший підхід, який дозволяє виразити визначник матриці порядку $n$ через визначники матриць порядку $n-1$. Це дає рекурсивний метод обчислення, зручний як для теоретичних міркувань, так і для обчислення визначників матриць невеликого порядку вручну.

**Означення 7.2.** Нехай $A = (a_{ij})$ — квадратна матриця порядку $n$. *Мінором* $M_{ij}$ елемента $a_{ij}$ називається визначник матриці порядку $n-1$, отриманої з $A$ викресленням $i$-го рядка та $j$-го стовпця:

$$M_{ij} = \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}.$$

*Алгебраїчним доповненням* елемента $a_{ij}$ називається число

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Знаки $(-1)^{i+j}$ алгебраїчних доповнень легко запам'ятати — вони утворюють шахову дошку:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Приклад 7.7.** Обчислимо алгебраїчні доповнення $A_{12}$, $A_{23}$, $A_{31}$ для матриці $A$:

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ -3 & 5 & 7 \end{pmatrix}, \quad \begin{aligned} A_{12} &= (-1)^{1+2} \det \begin{pmatrix} 2 & 0 \\ -3 & 7 \end{pmatrix} = -(14-0) = -14, \\ A_{23} &= (-1)^{2+3} \det \begin{pmatrix} 3 & 2 \\ -3 & 5 \end{pmatrix} = -(15+6) = -21, \\ A_{31} &= (-1)^{3+1} \det \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} = (0+1) = 1. \end{aligned}$$

**Теорема 7.3** (розклад визначника за рядком і стовпцем)**.** *Нехай $A = (a_{ij})$ — квадратна матриця порядку $n$. Тоді для будь-якого $i = 1, \ldots, n$ і $j = 1, \ldots, n$ виконуються формули*

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \quad \textit{(розклад за } i\textit{-м рядком)},$$
$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \quad \textit{(розклад за } j\textit{-м стовпцем)}.$$

*Доведення.* Запишемо $i$-й рядок матриці $A$ як суму рядків, кожен з яких містить лише один ненульовий елемент:

$$(a_{i1}, a_{i2}, \ldots, a_{in}) = (a_{i1}, 0, \ldots, 0) + (0, a_{i2}, \ldots, 0) + \cdots + (0, 0, \ldots, a_{in}).$$

За лінійністю визначника відносно рядків:

$$\det A = \sum_{j=1}^{n} \det A^{(j)},$$

де $A^{(j)}$ — матриця, яка збігається з $A$ в усіх рядках, крім $i$-го, а $i$-й рядок якої має вигляд $(0, \ldots, a_{ij}, \ldots, 0)$. Переставляючи $i$-й рядок матриці $A^{(j)}$

на перше місце за допомогою $i-1$ перестановок сусідніх рядків, а $j$-й стовпець — на перше місце за допомогою $j-1$ перестановок сусідніх стовпців, отримуємо матрицю блочного вигляду:

$$\det A^{(j)} = (-1)^{(i-1)+(j-1)} \begin{vmatrix} a_{ij} & 0 & \dots & 0 \\ * & & & \\ \vdots & & M_{ij} & \\ * & & & \end{vmatrix} = (-1)^{i+j} a_{ij} M_{ij}.$$

Підсумовуючи за всіма $j$, отримуємо потрібну формулу.

Розклад за стовпцем отримується застосуванням формули розкладу за рядком до матриці $A^T$ з урахуванням $\det A = \det A^T$. □

**Зауваження 7.3.** Формулу розкладу за рядком можна використати як індуктивне означення визначника. А саме, визначник матриці першого порядку покладають рівним $\det(a_{11}) = a_{11}$, а визначник матриці порядку $n$ визначають через визначники порядку $n-1$ за формулою розкладу за першим рядком:

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}.$$

Таке означення є еквівалентним до означення через підстановки і часто використовується як основне у курсах лінійної алгебри.

**Приклад 7.8.** Обчислимо визначник матриці

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 2 \\ 1 & -1 & 4 \end{pmatrix}.$$

Розкладемо за першим рядком:

$$\begin{aligned} \det A &= 2 \cdot A_{11} + 1 \cdot A_{12} + (-1) \cdot A_{13} = \\ &= 2 \begin{vmatrix} 3 & 2 \\ -1 & 4 \end{vmatrix} + 1 \cdot (-1) \begin{vmatrix} 0 & 2 \\ 1 & 4 \end{vmatrix} + (-1) \begin{vmatrix} 0 & 3 \\ 1 & -1 \end{vmatrix} = \\ &= 2(12+2) - 1(0-2) - 1(0-3) = 2 \cdot 14 + 2 + 3 = 33. \end{aligned}$$

**Приклад 7.9.** Обчислимо визначник матриці

$$A = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 1 & -1 & 3 & 2 \\ 0 & 2 & 2 & 1 \\ 1 & 3 & 0 & 4 \end{pmatrix}.$$

Розкладемо за третім стовпцем:

$$\det A = 3 \cdot A_{23} + 2 \cdot A_{33} = 3(-1)\begin{vmatrix} 2 & 1 & 3 \\ 0 & 2 & 1 \\ 1 & 3 & 4 \end{vmatrix} + 2\begin{vmatrix} 2 & 1 & 3 \\ 1 & -1 & 2 \\ 1 & 3 & 4 \end{vmatrix} =$$

$$= -3\big(2(8-3) - 1(0-1) + 3(0-2)\big) +$$

$$+ 2\big(2(-4-6) - 1(4-2) + 3(3+1)\big) = -35.$$

**Теорема 7.4** (про розклад за «чужим» рядком і стовпцем)**.** *Нехай* $A = (a_{ij})$ *— квадратна матриця порядку* $n$*. Для будь-яких двох різних індексів* $i \neq k$ *виконуються рівності:*

$$\sum_{j=1}^{n} a_{ij}A_{kj} = a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = 0,$$

$$\sum_{i=1}^{n} a_{ij}A_{il} = a_{1j}A_{1l} + a_{2j}A_{2l} + \cdots + a_{nj}A_{nl} = 0.$$

*Доведення.* Достатньо довести першу рівність. Розглянемо допоміжну матрицю $B$, отриману з матриці $A$ заміною $k$-го рядка на $i$-й рядок. Оскільки матриця $B$ містить два однакових рядки, її визначник дорівнює нулю. При цьому $B_{kj} = A_{kj}$ для всіх $j$, оскільки решта рядків не змінилась. Розкладаючи визначник матриці $B$ за $k$-м рядком, отримуємо:

$$0 = \det B = \sum_{j=1}^{n} b_{kj}B_{kj} = \sum_{j=1}^{n} a_{ij}A_{kj}.$$

Друга рівність доводиться аналогічно застосуванням першої до матриці $A^T$. □

**Приклад 7.10.** Проілюструємо розклад за «чужим» рядком для матриці

$$A = \begin{pmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & 2 & -5 \\ 2 & 0 & 3 & -4 \\ 5 & -1 & 0 & 2 \end{pmatrix}.$$

Розкладемо за алгебраїчними доповненнями першого рядка, але з елементами третього рядка:

$$a_{31}A_{11} + a_{32}A_{12} + a_{33}A_{13} + a_{34}A_{14} =$$
$$= 2\begin{vmatrix} 1 & 2 & -5 \\ 0 & 3 & -4 \\ -1 & 0 & 2 \end{vmatrix} + 3(-1)^{1+3}\begin{vmatrix} 0 & 1 & -5 \\ 2 & 0 & -4 \\ 5 & -1 & 2 \end{vmatrix} + (-4)(-1)^{1+4}\begin{vmatrix} 0 & 1 & 2 \\ 2 & 0 & 3 \\ 5 & -1 & 0 \end{vmatrix} =$$
$$= 2 \cdot (-1) + 3 \cdot (-14) + (-4) \cdot (-11) = -2 - 42 + 44 = 0.$$

Алгебраїчні доповнення та теореми про розклад визначника за своїм і чужим рядком дозволяють обчислити обернену матрицю за допомогою визначників.

**Означення 7.3.** *Приєднаною матрицею* до квадратної матриці $A = (a_{ij})$ порядку $n$ називається матриця $\operatorname{adj} A$, складена з алгебраїчних доповнень $A_{ij}$ елементів $a_{ij}$, взятих у транспонованому порядку:

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

**Приклад 7.11.** Знайдемо приєднану матрицю до матриці

$$A = \begin{pmatrix} 3 & -2 & 1 \\ 5 & 1 & 2 \\ -3 & 7 & 4 \end{pmatrix}.$$

Обчислимо всі дев'ять алгебраїчних доповнень:

$$A_{11} = \begin{vmatrix} 1 & 2 \\ 7 & 4 \end{vmatrix} = -10, \quad A_{12} = -\begin{vmatrix} 5 & 2 \\ -3 & 4 \end{vmatrix} = -26, \quad A_{13} = \begin{vmatrix} 5 & 1 \\ -3 & 7 \end{vmatrix} = 38,$$
$$A_{21} = -\begin{vmatrix} -2 & 1 \\ 7 & 4 \end{vmatrix} = 15, \quad A_{22} = \begin{vmatrix} 3 & 1 \\ -3 & 4 \end{vmatrix} = 15, \quad A_{23} = -\begin{vmatrix} 3 & -2 \\ -3 & 7 \end{vmatrix} = -15,$$
$$A_{31} = \begin{vmatrix} -2 & 1 \\ 1 & 2 \end{vmatrix} = -5, \quad A_{32} = -\begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix} = -1, \quad A_{33} = \begin{vmatrix} 3 & -2 \\ 5 & 1 \end{vmatrix} = 13.$$

Записуємо алгебраїчні доповнення у транспонованому порядку:

$$\operatorname{adj} A = \begin{pmatrix} -10 & -26 & 38 \\ 15 & 15 & -15 \\ -5 & -1 & 13 \end{pmatrix}^T = \begin{pmatrix} -10 & 15 & -5 \\ -26 & 15 & -1 \\ 38 & -15 & 13 \end{pmatrix}.$$

**Теорема 7.5.** *Для будь-якої квадратної матриці $A$ виконується:*

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot E.$$

*Доведення.* Розглянемо елемент $c_{ik}$ матриці $C = A \cdot \operatorname{adj} A$. За означенням добутку матриць:

$$c_{ik} = \sum_{j=1}^{n} a_{ij} (\operatorname{adj} A)_{jk} = \sum_{j=1}^{n} a_{ij} A_{kj}.$$

Згідно з теоремою про розклад за «своїм» та «чужим» рядком, ця сума дорівнює $\det A$, якщо $i = k$, і дорівнює 0, якщо $i \neq k$. Це означає, що матриця $C$ є діагональною з елементами $\det A$ на головній діагоналі, тобто $C = \det A \cdot E$. □

**Наслідок 7.3.** *Якщо* $\det A \neq 0$*, то матриця $A$ є оборотною і*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A.$$

**Приклад 7.12.** Знайдемо обернену до матриці з попереднього прикладу

$$A = \begin{pmatrix} 3 & -2 & 1 \\ 5 & 1 & 2 \\ -3 & 7 & 4 \end{pmatrix}.$$

Приєднана матриця вже була знайдена. Обчислимо визначник, розклавши за першим рядком:

$$\det A = 3 \cdot (-10) + (-2) \cdot (-26) + 1 \cdot 38 = -30 + 52 + 38 = 60.$$

Оскільки $\det A = 60 \neq 0$, матриця $A$ є оборотною і

$$A^{-1} = \frac{1}{60} \begin{pmatrix} -10 & 15 & -5 \\ -26 & 15 & -1 \\ 38 & -15 & 13 \end{pmatrix}.$$

## 7.4 Застосування визначників

У цьому розділі розглянемо три класичні застосування визначників: правило Крамера для розв'язання систем лінійних рівнянь, обчислення площ і об'ємів геометричних фігур та інтерполяцію многочленами.

Розглянемо систему $n$ лінійних рівнянь з $n$ невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{cases}$$

або у матричній формі $A\boldsymbol{x} = \boldsymbol{b}$, де $A = (a_{ij})$ — квадратна матриця коефіцієнтів, $\boldsymbol{x} = (x_1, \dots, x_n)$ — стовпець невідомих, $\boldsymbol{b} = (b_1, \dots, b_n)$ — стовпець вільних членів.

**Теорема 7.6** (правило Крамера)**.** *Нехай $A$ — оборотна матриця порядку $n$ і $\boldsymbol{b} \in \mathbb{R}^n$ — довільний вектор. Тоді система $A\boldsymbol{x} = \boldsymbol{b}$ має єдиний розв'язок, який визначається формулами*

$$x_j = \frac{\det A_j}{\det A}, \quad j = 1, \dots, n,$$

*де $A_j$ — матриця, отримана з $A$ заміною $j$-го стовпця на стовпець вільних членів $\boldsymbol{b}$.*

*Доведення.* Оскільки матриця $A$ є оборотною, то система має єдиний розв'язок $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. За формулою для оберненої матриці через приєднану (див. наслідок 7.3):

$$\boldsymbol{x} = \frac{1}{\det A}(\operatorname{adj} A)\boldsymbol{b}.$$

Обчислимо $j$-у компоненту вектора $\boldsymbol{x}$:

$$x_j = \frac{1}{\det A}\sum_{i=1}^{n} A_{ij}b_i.$$

Сума $\sum_{i=1}^{n} A_{ij}b_i$ є розкладом визначника матриці $A_j$ за $j$-м стовпцем: елементи $j$-го стовпця матриці $A_j$ є $b_1, \dots, b_n$, а їхні алгебраїчні доповнення збігаються з $A_{1j}, \dots, A_{nj}$, оскільки решта стовпців не змінилась. Отже,

$$x_j = \frac{\det A_j}{\det A}.$$

□

**Приклад 7.13.** Розв'яжемо систему методом Крамера:

$$\begin{cases} 2x_1 + 3x_2 - x_3 = 11, \\ x_1 - 2x_2 + 5x_3 = -9, \\ 4x_1 + x_2 + 2x_3 = 9. \end{cases}$$

Обчислимо визначник матриці системи:

$$\det A = \begin{vmatrix} 2 & 3 & -1 \\ 1 & -2 & 5 \\ 4 & 1 & 2 \end{vmatrix} = 2(-4-5) - 3(2-20) + (-1)(1+8) =$$
$$= -18 + 54 - 9 = 27.$$

Оскільки $\det A = 27 \neq 0$, система має єдиний розв'язок. Обчислимо визначники $\det A_1$, $\det A_2$, $\det A_3$:

$$\det A_1 = \begin{vmatrix} 11 & 3 & -1 \\ -9 & -2 & 5 \\ 9 & 1 & 2 \end{vmatrix} = 11(-4-5) - 3(-18-45) + (-1)(-9+18) =$$
$$= -99 + 189 - 9 = 81,$$
$$\det A_2 = \begin{vmatrix} 2 & 11 & -1 \\ 1 & -9 & 5 \\ 4 & 9 & 2 \end{vmatrix} = 2(-18-45) - 11(2-20) + (-1)(9+36) =$$
$$= -126 + 198 - 45 = 27,$$
$$\det A_3 = \begin{vmatrix} 2 & 3 & 11 \\ 1 & -2 & -9 \\ 4 & 1 & 9 \end{vmatrix} = 2(-18+9) - 3(9+36) + 11(1+8) =$$
$$= -18 - 135 + 99 = -54.$$

За правилом Крамера:

$$x_1 = \frac{81}{27} = 3, \quad x_2 = \frac{27}{27} = 1, \quad x_3 = \frac{-54}{27} = -2.$$

Визначники мають природну геометричну інтерпретацію через $n$-вимірний об'єм. Почнемо з випадку $n = 2$. Розглянемо паралелограм, побудований на векторах $\boldsymbol{a} = (a_1, a_2)$ і $\boldsymbol{b} = (b_1, b_2)$ у площині $\mathbb{R}^2$ (див. рис. 7.1). Його площа дорівнює

$$S = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cdot \sin\theta,$$

де $\theta$ — кут між векторами. Скористаємося тотожністю

$$(a_1b_2 - a_2b_1)^2 = |\boldsymbol{a}|^2|\boldsymbol{b}|^2 - (\boldsymbol{a}\cdot\boldsymbol{b})^2 = |\boldsymbol{a}|^2|\boldsymbol{b}|^2\sin^2\theta,$$

звідси $S = |a_1b_2 - a_2b_1|$. Отже, площа паралелограма дорівнює модулю визначника матриці, складеної з координат векторів:

$$S = \left|\det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right|.$$

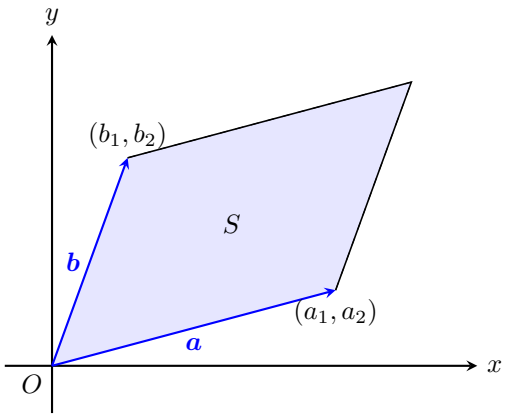



Рис. 7.1: Паралелограм, побудований на векторах $\boldsymbol{a}$ і $\boldsymbol{b}$.

Аналогічний результат має місце у просторі $\mathbb{R}^3$. Нехай $\boldsymbol{a} = (a_1, a_2, a_3)$, $\boldsymbol{b} = (b_1, b_2, b_3)$, $\boldsymbol{c} = (c_1, c_2, c_3)$ — три вектори. Об'єм паралелепіпеда, побудованого на цих векторах, виражається через мішаний добуток:

$$V = |(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}| = \left|\det\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\right|.$$

Ці спостереження узагальнюються на довільну розмірність.

**Означення 7.4.** *$n$-вимірним паралелепіпедом*, побудованим на векторах $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, називається множина

$$P = P(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = \{\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \cdots + \lambda_n\boldsymbol{a}_n \,|\, 0 \le \lambda_i \le 1, i = 1, \ldots, n\}.$$

Для паралелепіпеда в $\mathbb{R}^n$ визначається поняття $n$-вимірного об'єму, яке узагальнює площу при $n = 2$ та звичайний об'єм при $n = 3$.

**Теорема 7.7.** *$n$-вимірний об'єм паралелепіпеда, побудованого на векторах $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, дорівнює модулю визначника матриці $A$, стовпцями якої є вектори $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$:*

$$\mathrm{Vol}(P(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)) = |\det A|.$$

*Доведення.* Застосуємо до стовпців матриці $A$ елементарні перетворення, зводячи її до ступінчастої форми. Кожне елементарне перетворення стовпців відповідає геометричній операції над паралелепіпедом:

1. Перестановка двох стовпців не змінює паралелепіпеду: $V$ не змінюється, $|\det A|$ також не змінюється.

2. Множення стовпця на число $\lambda \neq 0$ розтягує паралелепіпед у $|\lambda|$ разів: $V$ множиться на $|\lambda|$, $|\det A|$ також множиться на $|\lambda|$.

3. Додавання до одного стовпця іншого, помноженого на число $\lambda$, не змінює об'єму паралелепіпеда, оскільки висота відносно гіперплощини решти векторів залишається незмінною. При цьому $|\det A|$ також не змінюється.

Оскільки $V$ та $|\det A|$ змінюються однаково при кожному елементарному перетворенні, достатньо перевірити рівність для ступінчастої форми матриці $A$.

Якщо ступінчаста форма містить нульовий стовпець, то один з векторів паралелепіпеда є нульовим, тому паралелепіпед вироджується: $V = 0$. При цьому $|\det A| = 0$.

Якщо ступінчаста форма не містить нульових стовпців, то вона є діагональною матрицею $D = \mathrm{diag}(d_1, \ldots, d_n)$ з ненульовими діагональними елементами. Паралелепіпед стає прямокутним зі сторонами $|d_1|, \ldots, |d_n|$, тому

$$V = |d_1| \cdots |d_n| = |\det D| = |\det A|. \qquad \square$$

**Зауваження 7.4.** Знак визначника $\det A$ також має геометричний зміст: він визначає орієнтацію системи векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Якщо $\det A > 0$, орієнтація збігається з орієнтацією стандартного базису $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$; якщо $\det A < 0$ — є протилежною.

**Приклад 7.14.** Знайдемо об'єм паралелепіпеда, побудованого на векторах $\boldsymbol{a}_1 = (2, -1, 4)$, $\boldsymbol{a}_2 = (1, 5, 3)$, $\boldsymbol{a}_3 = (3, 2, 1)$. За теоремою,

$$V = \left| \det \begin{pmatrix} 2 & 1 & 3 \\ -1 & 5 & 2 \\ 4 & 3 & 1 \end{pmatrix} \right| = |-62| = 62.$$

Ще одним важливим застосуванням визначників є задача поліноміальної інтерполяції.

**Означення 7.5** (Задача поліноміальної інтерполяції)**.** Нехай задано $n+1$ пару дійсних чисел $(x_0, y_0)$, $(x_1, y_1)$, $\ldots$, $(x_n, y_n)$, причому всі значення $x_i$ (вузли інтерполяції) є попарно різними: $x_i \neq x_j$ при $i \neq j$. Задача поліноміальної інтерполяції полягає у знаходженні коефіцієнтів $a_0, a_1, \ldots, a_n \in \mathbb{R}$ такого полінома

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

степеня не вище $n$, який у заданих вузлах набуває заданих значень:

$$p(x_i) = y_i, \quad i = 0, 1, \ldots, n.$$

Підставляючи координати точок у рівняння полінома, отримуємо систему $n+1$ лінійних рівнянь відносно невідомих коефіцієнтів $a_i$:

$$\begin{cases} a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n = y_0, \\ a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n = y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n = y_n. \end{cases}$$

Матриця цієї системи називається *матрицею Вандермонда*:

$$V(x_0, x_1, \ldots, x_n) = \begin{pmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{pmatrix}.$$

**Теорема 7.8** (визначник Вандермонда)**.**

$$\det V(x_0, x_1, \ldots, x_n) = \prod_{0 \leq i < j \leq n} (x_j - x_i).$$

*Доведення.* Доведення проведемо індукцією за $n$. При $n = 1$ маємо:

$$\det V(x_0, x_1) = \begin{vmatrix} 1 & x_0 \\ 1 & x_1 \end{vmatrix} = x_1 - x_0 = \prod_{0 \le i < j \le 1} (x_j - x_i).$$

Припустимо, що формула виконується для матриць Вандермонда порядку $n$ (тобто для $n$ вузлів). Розглянемо визначник $V(x_0, x_1, \ldots, x_n)$ порядку $n + 1$. Виконаємо такі елементарні перетворення стовпців: від кожного стовпця, починаючи з останнього, віднімемо попередній, помножений на $x_0$:

$$C_j \to C_j - x_0 C_{j-1}, \quad j = n, n-1, \ldots, 1.$$

Після цих перетворень перший рядок набуде вигляду $(1, 0, \ldots, 0)$. Отримаємо:

$$\det V(x_0, \ldots, x_n) = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_1 - x_0 & x_1^2 - x_0 x_1 & \cdots & x_1^n - x_0 x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_0 & x_n^2 - x_0 x_n & \cdots & x_n^n - x_0 x_n^{n-1} \end{vmatrix} =$$

$$= \begin{vmatrix} (x_1 - x_0) & x_1(x_1 - x_0) & \ldots & x_1^{n-1}(x_1 - x_0) \\ (x_2 - x_0) & x_2(x_2 - x_0) & \ldots & x_2^{n-1}(x_2 - x_0) \\ \vdots & \vdots & \ddots & \vdots \\ (x_n - x_0) & x_n(x_n - x_0) & \ldots & x_n^{n-1}(x_n - x_0) \end{vmatrix}.$$

Винесемо з кожного $i$-го рядка $(i = 1, \ldots, n)$ спільний множник $(x_i - x_0)$:

$$\det V(x_0, x_1, \ldots, x_n) = (x_1 - x_0)(x_2 - x_0) \ldots (x_n - x_0) \cdot \det V(x_1, \ldots, x_n).$$

Використовуючи припущення індукції для вузлів $x_1, \ldots, x_n$, маємо:

$$\det V(x_0, x_1, \ldots, x_n) = \prod_{k=1}^{n} (x_k - x_0) \cdot \prod_{1 \le i < j \le n} (x_j - x_i) = \prod_{0 \le i < j \le n} (x_j - x_i). \ \square$$

**Наслідок 7.4** (Існування та єдиність інтерполяційного полінома)**.** *Для будь-якого набору з $n + 1$ пари чисел $(x_0, y_0)$, $(x_1, y_1)$, $\ldots$, $(x_n, y_n)$ з попарно різними вузлами ($x_i \ne x_j$ при $i \ne j$) задача інтерполяції має єдиний розв'язок.*

*Доведення.* Як було показано вище, задача інтерполяції зводиться до системи лінійних рівнянь $V\boldsymbol{a} = \boldsymbol{y}$, де $V$ — матриця Вандермонда. Оскільки всі вузли $x_i$ за умовою є попарно різними, визначник матриці Вандермонда

$$\det V = \prod_{0 \le i < j \le n} (x_j - x_i)$$

не дорівнює нулю. Отже, згідно з правилом Крамера (або теоремою про обернену матрицю), система має єдиний розв'язок. Це доводить як існування, так і єдиність шуканого полінома. □

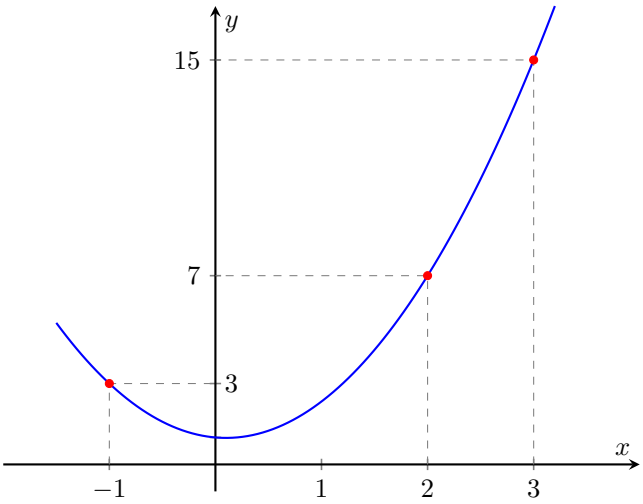



Рис. 7.2: Графік інтерполяційного полінома $p(x) = \frac{5}{3}x^2 - \frac{1}{3}x + 1$ та вузли інтерполяції.

**Приклад 7.15.** Знайдемо поліном степеня не вище 2, який проходить через точки $(-1, 3)$, $(2, 7)$, $(3, 15)$. Шукаємо поліном у вигляді $p(x) = a_0 + a_1 x + a_2 x^2$. Підставимо координати точок у рівняння полінома:

$$\begin{cases} a_0 - a_1 + a_2 = 3, \\ a_0 + 2a_1 + 4a_2 = 7, \\ a_0 + 3a_1 + 9a_2 = 15. \end{cases}$$

Розв'язуємо систему методом Крамера. Головний визначник системи (визначник Вандермонда для вузлів $x_0 = -1, x_1 = 2, x_2 = 3$):

$$\Delta = V(-1,2,3) = \begin{vmatrix} 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{vmatrix} = (2-(-1))(3-(-1))(3-2) = 3 \cdot 4 \cdot 1 = 12.$$

Обчислимо допоміжні визначники $\Delta_i$, замінюючи відповідні стовпці стовпцем вільних членів:

$$\Delta_0 = \begin{vmatrix} 3 & -1 & 1 \\ 7 & 2 & 4 \\ 15 & 3 & 9 \end{vmatrix} = 12, \quad \Delta_1 = \begin{vmatrix} 1 & 3 & 1 \\ 1 & 7 & 4 \\ 1 & 15 & 9 \end{vmatrix} = -4, \quad \Delta_2 = \begin{vmatrix} 1 & -1 & 3 \\ 1 & 2 & 7 \\ 1 & 3 & 15 \end{vmatrix} = 20.$$

Знаходимо коефіцієнти полінома:

$$a_0 = \frac{12}{12} = 1, \quad a_1 = \frac{-4}{12} = -\frac{1}{3}, \quad a_2 = \frac{20}{12} = \frac{5}{3}.$$

Отже, шуканий поліном: $p(x) = \frac{5}{3}x^2 - \frac{1}{3}x + 1$. Графік полінома $p(x)$ та вузли інтерполяції зображено на рис. 7.2.

## 7.5 Вправи

1. Обчисліть визначники:

а) $\begin{vmatrix} 9 & -2 \\ 7 & 4 \end{vmatrix}$; б) $\begin{vmatrix} 5 & 1 & 4 \\ 2 & 6 & -1 \\ 3 & 2 & 5 \end{vmatrix}$; в) $\begin{vmatrix} 1 & 4 & -3 \\ 6 & -2 & 1 \\ 2 & 5 & 4 \end{vmatrix}$;

г) $\begin{vmatrix} 4 & 1 & -2 & 3 \\ 2 & 3 & 1 & -1 \\ 1 & -2 & 4 & 2 \\ 3 & 4 & -1 & 5 \end{vmatrix}$; д) $\begin{vmatrix} 1 & -4 & 3 & 2 \\ 3 & 2 & -1 & 4 \\ 2 & 1 & 5 & -3 \\ -2 & 3 & 2 & 1 \end{vmatrix}$.

2. Розкладаючи за рядком або стовпчиком, обчисліть визначники:

а) $\begin{vmatrix} 2 & 0 & 3 & -1 \\ a & b & c & d \\ 1 & 5 & 1 & 4 \\ 0 & 3 & -2 & 2 \end{vmatrix}$; б) $\begin{vmatrix} 3 & -1 & -2 & 4 \\ 2 & 3 & 5 & 1 \\ 0 & 2 & 3 & -1 \\ 4 & 3 & -1 & 2 \end{vmatrix}$.

3. Обчисліть визначники порядку $n$:

а) $\begin{vmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -2 & -3 & \dots & 0 \end{vmatrix}$; б) $\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ -1 & 1 & 1 & \dots & 1 \\ -1 & -1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \dots & 1 \end{vmatrix}$;

в) $\begin{vmatrix} 2 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ 0 & 1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 \end{vmatrix}$; г) $\begin{vmatrix} 5 & 1 & 0 & \dots & 0 \\ 3 & 5 & 1 & \dots & 0 \\ 0 & 3 & 5 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 5 \end{vmatrix}$.

4. Знайдіть обернену матрицю за допомогою приєднаної матриці та перевірте результат:

а) $\begin{pmatrix} 4 & 7 \\ 2 & 5 \end{pmatrix}$; б) $\begin{pmatrix} 1 & -2 & 4 \\ 3 & 1 & -1 \\ 2 & 3 & 2 \end{pmatrix}$.

5. Розв'яжіть системи лінійних рівнянь методом Крамера:

а) $\begin{cases} 5x - 3y = 1, \\ 2x + 4y = 14; \end{cases}$ б) $\begin{cases} 2x_1 + x_2 - x_3 = 3, \\ x_1 + 2x_2 + x_3 = 2, \\ 3x_1 - x_2 + 2x_3 = 5. \end{cases}$

6. Використовуючи геометричний зміст визначника, розв'яжіть задачі:

   а) Знайдіть площу паралелограма, побудованого на векторах $(3, -2)$ та $(4, 5)$.

   б) Знайдіть площу трикутника з вершинами $A(1, 2)$, $B(4, -1)$, $C(3, 5)$.

   в) Знайдіть об'єм паралелепіпеда, побудованого на векторах $(2, 1, -3)$, $(1, -2, 4)$, $(3, 1, 2)$.

   г) Знайдіть об'єм тетраедра з вершинами $A(1, 2, 3)$, $B(3, 1, 2)$, $C(2, 3, 1)$, $D(4, 2, 3)$.

д) При яких значеннях параметра $k$ точки $A(1,3)$, $B(k,-1)$, $C(4,k)$ лежать на одній прямій?

7. Нехай $A$ — квадратна матриця порядку $n$, а $\lambda \in \mathbb{R}$. Доведіть, що $\det(\lambda A) = \lambda^n \det(A)$.

8. Нехай $A$ — кососиметрична матриця непарного порядку $n$. Доведіть, що $\det(A) = 0$.

9. Нехай $A$ — квадратна матриця порядку $n$. Доведіть, що

$$\det(\operatorname{adj} A) = \det(A)^{n-1}.$$

10. Доведіть властивості приєднаної матриці:

а) $\operatorname{adj}(AB) = \operatorname{adj}(B) \cdot \operatorname{adj}(A)$;

б) $\operatorname{adj}(A^T) = (\operatorname{adj} A)^T$;

в) $\operatorname{adj}(\operatorname{adj} A) = \det(A)^{n-2} A$ при $n \geq 2$.

# Розділ 8

# Многочлени, основні поняття

Нехай $F$ — деяке поле (наприклад, $\mathbb{R}$ або $\mathbb{C}$), $a_0$, $a_1$, $a_2$, $\ldots$, $a_{n-1}$, $a_n$ — елементи поля $F$, та $n \in \mathbb{N}$.

**Означення 8.1.** Многочленом або поліномом від змінної $x$ над полем $F$ називається наступний вираз

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0.$$

При такому заданні многочлена індекс $i$ коефіцієнта відповідає степеню змінної при якій він стоїть, тобто $a_i x^i$.

Часто також використовується інший, зворотній порядок нумерації коефіцієнтів, тоді многочлен записується у наступному вигляді

$$h(x) = c_0 x^n + c_1 x^{n-1} + \ldots + c_{n-2} x^2 + c_{n-1} x + c_n.$$

При такому записі індекс $j$ коефіцієнта відповідає степеню $n - j$ змінної при якій він стоїть, тобто $c_j x^{n-j}$.

**Означення 8.2.** Степенем многочлена є найбільший степінь $x$ для якого відповідний коефіцієнт $a_i$ не нульовий. Степінь многочлена позначають $\deg(f(x))$ або просто $\deg f$.

Наприклад,

$$\deg(-3x^5+0x^4+2x^3+7x^2+0x+4) = 5, \quad \deg(0x^4+0x^3+3x^2+8x+1) = 2.$$

Два многочлени називаються рівними, якщо у них рівні коефіцієнти при відповідних степенях $x$.

Окремо можна виділити випадок, коли многочлен явно не містить $x$. Якщо всі коефіцієнти многочлена $f(x)$ рівні 0, то такий многочлен називають *нульовим* і позначають його символом 0. Степінь нульового многочлена покладемо рівним $-\infty$. Якщо ж $f(x) = a_0 \neq 0$, то покладаємо $\deg(f(x)) = 0$. Таким чином многочленами нульового степеня є ненульові елементи поля $F$ (числа у випадку $\mathbb{R}$ або $\mathbb{C}$)

Нехай у многочлена $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$ коефіцієнт $a_n$ не дорівнює нулю. Нагадаємо, що коефіцієнт $a_n \neq 0$ називається *старшим коефіцієнтом*, а коефіцієнт $a_0$ називається *вільним членом*.

Множину всіх многочлнів від змінної $x$ над полем $F$ позначатимемо символом $F[x]$.

З многочленами можна виконувати операції додавання і множення задані формальним чином.

Розглянемо два многочлени

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0, \text{ де } a_n \neq 0,$$

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_2 x^2 + b_1 x + b_0, \text{ де } b_m \neq 0.$$

Припустимо, що $n > m$, тоді покладемо $b_i = 0$, якщо $i > m$. Таким чином

$$g(x) = 0x^n + \ldots + 0x^{m+1} + b_m x^m + b_{m-1}x^{m-1} + \ldots + b_2 x^2 + b_1 x + b_0.$$

Тоді *суму многочленів* формально обчислюють наступним чином

$$(f+g)(x) = f(x) + g(x) = (a_n + b_n)x^n + \ldots + (a_1 + b_1)x + (a_0 + b_0).$$

Нехай $c \in F$ деякий елемент поля, тоді

$$(cf)(x) = cf(x) = ca_n x^n + ca_{n-1}x^{n-1} + \ldots + ca_1 x + ca_0.$$

*Добуток многочленів* формально обчислюють наступним чином

$$(fg)(x) = f(x) \cdot g(x) = (a_n b_m)x^{n+m} + \ldots + a_0 b_0,$$

інакше можемо записати

$$(fg)(x) = c_{m+n}x^{n+m} + \ldots + c_0,$$

де

$$c_k = \sum_{i=0}^{k} a_i b_{k-1} = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0.$$

Далі запишемо як співвідносяться степені многочленів зі степенем їх добутку чи суми.

**Теорема 8.1.** *Для степеня добутку двох многочленів $f(x)$ та $g(x)$ справедлива рівність*

$$\deg(f(x)g(x)) = \deg(f(x)) + \deg(g(x)).$$

*Доведення.* Нехай

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 \quad \text{та} \quad g(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_0,$$

де $a_n \neq 0$ і $b_m \neq 0$. Тоді з правила множення многочленів

$$f(x)g(x) = a_n b_m x^{n+m} + \text{ члени менших степенів.}$$

Оскільки коефіцієнт $a_n b_m$ не дорівнює нулю, він є старшим коефіцієнтом та $\deg(f(x)g(x)) = n + m = \deg(f(x)) + \deg(g(x))$. □

**Наслідок 8.1.** *Для степеня добутку скінченної кількості многочленів*

$$f_1(x), \ldots, f_n(x)$$

*справедлива рівність*

$$\deg(f_1(x) \cdot \ldots \cdot f_n(x)) = \deg(f_1(x)) + \ldots + \deg(f_n(x)).$$

**Наслідок 8.2.** *Добуток не нульових многочленів не може дорівнювати нулю. Тобто з рівності $f(x)g(x) = 0$ випливає, що $f(x) = 0$ або $g(x) = 0$.*

**Теорема 8.2.** *Для степеня суми довільних многочленів в загальному випадку справедливе наступне співвідношення*

$$\deg(f(x) + g(x)) \leq \max\{\deg(f(x)), \deg(g(x))\}.$$

Така нерівність виникає оскільки у випадку додавання многочленів різних степенів, при зведенні подібних доданків, старший член одного з доданків залишається, а у випадку додавання многочленів одного степеня старші члени обох доданків можуть скоротитися і тоді степінь суми зменшиться.

**Приклад 8.1.** Знайти $\deg(f(x)+g(x))$, перевірити, що отримана відповідь задовольняє попередню теорему.

1. $f(x) = 2x^3 + 3x + 5$, $g(x) = -2x^3 + 7x^2 - x + 1$;
2. $f(x) = 3x^6 + 6x^5 + x + 2$, $g(x) = -4x^5 + 7x^2 - 5x - 2$;

*Розв'язання.* 1. Знайдемо суму многочленів $f(x)+g(x) = (2x^3+3x+5)+(-2x^3+7x^2-x+1) = 7x^2+2x+6$, тоді очевидно, що $\deg(f(x)+g(x)) = 2$. З умови $\deg(f(x)) = 3$ та $\deg(g(x)) = 3$. Тоді бачимо, що твердження попередньої теореми виконується, і справедливою є не строга нерівність

$$\deg(f(x) + g(x)) = 2 < 3 = \max\{3, 3\} = \max\{\deg(f(x)), \deg(g(x))\}.$$

2. Обчислюємо суму многочленів $f(x) + g(x) = (3x^6 + 6x^5 + x + 2) + (-4x^5+7x^2-5x-2) = 3x^6+2x^5+7x^2-4x$, тому $\deg(f(x)+g(x)) = 6$. З умови $\deg(f(x)) = 6$ та $\deg(g(x)) = 5$. Твердження попередньої теореми виконується, і у даному випадку справедливою є рівність

$$\deg(f(x) + g(x)) = 6 = \max\{6, 5\} = \max\{\deg(f(x)), \deg(g(x))\}. \quad \square$$

Далі сформулюємо основні властивості для операцій з многочленами.

**Твердження 8.1.** *Додавання, і множення многочленів є асоціативними*

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)),$$

$$(f(x) \cdot g(x)) \cdot h(x) = f(x) \cdot (g(x) \cdot h(x)),$$

*та комутативними діями*

$$f(x) + g(x) = g(x) + f(x),$$

$$f(x) \cdot g(x) = g(x) \cdot f(x).$$

**Твердження 8.2.** *Множення многочленів є дистрибутивним відносно додавання, тобто*

$$f(x) \cdot (g(x) + h(x)) = f(x) \cdot g(x) + f(x) \cdot h(x).$$

**Твердження 8.3.** *Для додавання нейтральним многочленом є нульовий многочлен, а для множення — тотожно рівний* 1.

**Твердження 8.4.** *Для кожного многочлена існує протилежний.*

**Твердження 8.5.** *Обернені для множення елементи існують лише для многочленів нульового степеня.*

**Твердження 8.6.** *Множина $F[x]$ відносно введених дій додавання і множення не є полем, але є комутативним кільцем з одиницею.*

**Твердження 8.7.** *У рівностях многочленів можна скорочувати на ненульовий спільний множник.*

*Доведення.* Можливість скорочувати означає, що для довільного ненульового многочлена $f(x)$, якщо виконується рівності $f(x)g(x) = f(x)h(x)$, то повинна виконуватися рівність $g(x) = h(x)$.

Якщо у рівності $f(x)g(x) = f(x)h(x)$ многочлен з правої частині рівності перенести у ліву, та винестиспільний множник $f(x)$, отримуємо рівносильну рівність $f(x)(g(x) - h(x)) = 0$. Оскільки $f(x)$ не нульовий многочлен, а в $F[x]$ добуток не нульових многочленів є не нульовим, то звідси випливає рівність $g(x) - h(x) = 0$. Тоді $g(x) = h(x)$, що й треба було довести. □

Кажемо, що многочлен $f(x)$ *має нуль*, якщо існує таке $x := a$, що $f(a) = 0$. Тобто $a$ *є нулем* многочлена $f(x)$.

## 8.1 Подільність многочленів

Доведемо теорему про ділення з остачею многочленів.

**Теорема 8.3.** *Нехай $f(x)$, $g(x)$ — довільні многочлени над полем $F[x]$, $g(x) \neq 0$. Тоді існують і єдині такі многочлени $q(x)$ і $r(x)$ над $F[x]$, що*

$$f(x) = g(x) \cdot q(x) + r(x), \text{ причому } \deg r(x) < \deg g(x).$$

*Доведення.* Зафіксуємо многочлен $g(x)$ який має степінь $\deg g(x) = m$, та $m \geq 0$. Доведемо твердження теореми для довільного многочлена $f(x)$.

Спочатку доведемо *існування* многочленів $q(x)$ і $r(x)$. Зауважимо, що у випадку $\deg f(x) < \deg g(x)$ (зокрема $f(x) = 0$) маємо рівність

$$f(x) = g(x) \cdot 0 + f(x),$$

тобто можемо покласти $q(x) = 0$, а $r(x) = f(x)$.

Проведемо індукцію за степенем $n$ многочлена $f(x)$.

Базою індукції є випадок $n = 0$. Потрібно розглянути лише випадок, коли $m = n = 0$. Тоді $f(x) = a_0$, а $g(x) = b_0$ для деяких $a_0, b_0 \in F$, $b_0 \neq 0$, і маємо рівність

$$f(x) = g(x)(a_0 b_0^{-1}) + 0.$$

Таким чином, можемо покласти $q(x) = a_0 b_0^{-1}$ і $r(x) = 0$.

Припустимо тепер, що твердження доведено для всіх многочленів $f(x)$, степені яких менші за $n$. Розглянемо многочлен $f(x)$ степеня $n$. Можна вважати, що $m \leq n$. Нехай

$$f(x) = a_n x^n + \ldots + a_0, \quad g(x) = b_m x^m + \ldots + b_0.$$

Визначимо многочлен

$$h(x) = f(x) - (a_n b_m^{-1}) x^{n-m} g(x).$$

Степінь многочлена $h(x)$ буде меншим за $n$ і тому до нього можемо застосувати припущення індукції. Таким чином, існують многочлени $q_1(x)$ і $r_1(x)$ такі, що

$$h(x) = g(x)q_1(x) + r_1(x) \quad \text{та} \quad \deg r_1(x) < \deg g(x).$$

Підставляючи в останню рівність значення $h(x)$, після очевидних перетворень отримаємо

$$f(x) = g(x)(q_1(x) + (a_n b_m^{-1}) x^{n-m}) + r_1(x).$$

Поклавши тепер $q(x) = q_1(x) + (a_n b_m^{-1}) x^{n-m}$ і $r(x) = r_1(x)$, отримаємо потрібне твердження для многочлена $f$.

Доведемо *єдиність*. Нехай

$$f = g(x)q_1(x) + r_1(x) = g(x)q_2(x) + r_2(x)$$

для деяких многочленів $q_1(x)$, $q_2(x)$, $r_1(x)$, $r_2(x)$, причому $\deg r_1(x) < \deg g(x)$ і $\deg r_2(x) < \deg g(x)$. Звідси отримуємо рівність

$$g(x)(q_1(x) - q_2(x)) = r_2(x) - r_1(x),$$

причому $\deg(r_2(x) - r_1(x)) < \deg g(x)$. Тому, враховуючи, що $g \neq 0$, остання рівність можлива лише тоді, коли $q_1(x) - q_2(x) = 0$, тобто $q_1(x) = q_2(x)$. Звідси $r_2(x) - r_1(x) = 0$ і $r_1(x) = r_2(x)$.

Теорему доведено. □

**Означення 8.3.** Многочлен $f(x) \in F[x]$ є дільником многочлена $g(x) \in F[x]$ або, що те ж саме, многочлен $g(x)$ ділиться на многочлен $f(x)$ (позначається $f(x) \mid g(x)$), якщо існує такий многочлен $h(x) \in F[x]$, що виконується рівність

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

Справедливими є властивості подільності для нульових многочленів та многочленів нульового степеня.

**Твердження 8.8.** *1. Нульовий многочлен є дільником лише самого себе.*

*2. Нульовий многочлен ділиться на будь-який многочлен.*

*3. Многочлени нульового степеня, і лише вони, є дільниками довільних многочленів.*

**Твердження 8.9.** *Для довільних многочленів $f(x), g(x), h(x) \in F[x]$ мають місце такі властивості*

1. *$f(x) \mid f(x)$ (рефлексивність);*
2. *якщо $f(x) \mid g(x)$ і $g(x) \mid h(x)$, то $f(x) \mid h(x)$ (транзитивність);*
3. *якщо $f(x) \mid g(x)$, то для довільного многочлена $h(x)$ справедливо, що $f(x) \mid (g(x)h(x))$;*
4. *якщо $f(x) \mid g(x)$ і $f(x) \mid h(x)$, то $f(x) \mid (g(x) + h(x))$.*

*Доведення.* 1. Рефлексивність випливає з рівності $f(x) = 1 \cdot f(x)$.

2. З умови випливає, що існують такі многочлени $u(x)$ та $v(x)$, що $g(x) = u(x)f(x)$ та $h(x) = v(x)g(x)$. Тоді підставивши вираз для $g(x)$ у вираз для $h(x)$, отримуємо $h(x) = v(x)u(x)f(x)$. Тобто бачимо, що $f(x) \mid h(x)$.

3. З умови випливає, що існує такий многочлен $u(x)$ такий, що $g(x) = u(x)f(x)$. Домножимо на многочлен $h(x)$ ліву і праву частину останньої рівності, тоді

$$g(x)h(x) = u(x)f(x)h(x) = (u(x)h(x))f(x) = v(x)f(x),$$

де $v(x) = u(x)h(x)$. Тобто існує такий многочлен $v(x)$, що $g(x)h(x) = v(x)f(x)$, тоді $f(x) \mid (g(x)h(x))$.

4. Існують такі многочлени $u(x)$ та $v(x)$, що ми можемо записати рівності $g(x) = u(x)f(x)$ та $h(x) = v(x)f(x)$. Тоді

$$g(x) + h(x) = u(x)f(x) + v(x)f(x) = (u(x) + v(x))f(x) = w(x)h(x),$$

де $w(x) = u(x) + v(x)$. Очевидно, що тоді $f(x) \mid (g(x) + h(x))$. □

**Означення 8.4.** Многочлени $f(x)$ і $g(x)$ називаються асоційованими, якщо кожен з них ділиться на інший.

**Твердження 8.10.** *Многочлени $f(x)$ і $g(x)$ є асоційованими тоді й лише тоді, коли існує ненульовий елемент $\lambda \in F$ такий, що $f(x) = \lambda g(x)$.*

*Доведення.* Достатність вказаної умови прямо випливає з означення подільності.

Перевіримо необхідність. Досить зробити це для ненульових многочленів. Нехай ненульові многочлени $f(x)$ і $g(x)$ є асоційованими, тобто $f(x) \mid g(x)$ і $g(x) \mid f(x)$. Тоді для деяких многочленів $q_1(x)$ і $q_2(x)$ мають місце рівності

$$f(x) = g(x)q_1(x), \quad g(x) = f(x)q_2(x).$$

Звідси маємо рівність

$$f(x) = f(x)q_1(x)q_2(x),$$

тобто $q_1(x)q_2(x) = 1$. Це означає, що обидва многочлени $q_1(x)$ і $q_2(x)$ є многочленами нульового степеня, звідки й отримуємо потрібне. □

**Означення 8.5.** Многочлен називається унітарним, якщо його старший коефіцієнт дорівнює 1.

**Твердження 8.11.** *1. Унітарні многочлени асоційовані тоді й лише тоді, коли вони рівні;*

*2. кожен ненульовий многочлен асоційований з рівно одним унітарним многочленом*

*3. множина всіх многочленів розбивається на класи попарно асоційованих між собою многочленів;*

*4. кожен клас, що складається з ненульових многочленів, містить єдиний унітарний многочлен.*

## 8.2 Найбільший спільний дільник многочленів

**Означення 8.6.** Найбільшим спільним дільником многочленів $f(x), g(x) \in F[x]$ називається такий многочлен $d(x) \in F[x]$, що:

1. $d(x) \mid f(x)$ і $d(x) \mid g(x)$;
2. для довільного многочлена $h(x) \in F[x]$ такого, що $h(x) \mid f(x)$ і $h(x) \mid g(x)$, також $h(x) \mid d(x)$.

Найбільший спільний дільник двох многочленів — це такий їх спільний дільник, який ділиться на кожен спільний дільник цих многочленів. Степінь найбільшого спільного дільника буде найбільшим серед степенів спільних дільників заданих многочленів. Коли один з многочленів нульовий, то найбільший спільний дільник існує і дорівнює іншому многочлену.

**Теорема 8.4.** *Для довільних многочленів $f(x), g(x) \in F[x]$ існує найбільший спільний дільник, причому його можна подати у вигляді*

$$m(x)f(x) + n(x)g(x)$$

*для деяких многочленів $m(x), n(x) \in F[x]$.*

*Доведення.* Якщо хоча б один з многочленів $f(x)$ і $g(x)$ нульовий, то твердження теореми виконується. Тому далі розглядаємо випадок ненульових многочленів $f(x)$ і $g(x)$.

Визначимо множину многочленів

$$I = \{m(x)f(x) + n(x)g(x) \mid m(x), n(x) \in F[x]\}.$$

Вона містить зокрема многочлени $f(x)$ і $g(x)$, тобто містить ненульові многочлени. Виберемо в ній деякий ненульовий многочлен $d(x)$ найменшого степеня. Тоді $d(x) = m_0(x)f(x) + n_0(x)g(x)$ для деяких многочленів $m_0(x)$, $n_0(x) \in F[x]$. Покажемо, що многочлен $d(x)$ буде найбільшим спільним дільником многочленів $f(x)$ і $g(x)$. Зауважимо, що $d(x)$ одразу матиме потрібний вигляд, як сказано в умові теореми. Поділимо $f(x)$ на $d(x)$ з остачею:

$$f(x) = q(x)d(x) + r(x)$$

для деяких многочленів $q(x), r(x) \in F[x]$ і $\deg r(x) < \deg d(x)$. Звідси

$$r(x) = (1 - q(x)m_0(x))f(x) + (-q(x)n_0(x))g(x), \text{ тобто } r(x) \in I.$$

Але степінь многочлена $d(x)$ є найменшим серед степенів ненульових многочленів з множини $I$. Тому многочлен $r(x)$ є нульовим та $d(x) \mid f(x)$. Аналогічно показується, що $d(x) \mid g(x)$.

Нехай тепер многочлен $h(x) \in F[x]$ такий, що $h(x) \mid f(x)$ і $h(x) \mid g(x)$. З властивостей подільності випливає, що тоді $h(x) \mid (m_0(x)f(x))$ і $h(x) \mid (n_0(x)g(x))$, тоді $h(x) \mid (m_0(x)f(x) + n_0(x)g(x))$, тобто $h(x) \mid d(x)$.

Теорему доведено. □

Зрозуміло, що всі найбільші спільні дільники двох заданих многочленів є асоційованими. Також многочлен, асоційований з найбільшим спільним дільником заданих многочленів, і сам буде найбільшим спільним дільником цих многочленів. Якщо розглядати найбільші спільні дільники двох многочленів, принаймні один з яких є ненульовим, то серед таких найбільших спільних дільників існує єдиний унітарний многочлен.

Тому далі завжди можемо вважати, що найбільший спільний дільник ненульових многочленів є унітарним многочленом.

## 8.3 Розширений алгоритм Евкліда для многочленів

Знайдемо найбільший спільний дільник двох ненульових многочленів $f(x)$ і $g(x)$. Можна вважати, що $\deg f(x) \geq \deg g(x)$.

**Лема 8.1.** *Нехай* $f(x) = q(x)g(x) + r(x)$. *Тоді*

$$\mathsf{НСД}(f(x), g(x)) = \mathsf{НСД}(g(x), r(x)).$$

*Доведення.* Припустимо, що многочлен $d(x)$, є спільним дільником для многочленів $f(x)$ і $g(x)$. Тоді за властивостями подільності многочленів $d(x)$ буде ділить і многочлен $r(x) = f(x) - q(x)g(x)$. Тобто $d(x)$ буде й спільним дільником для $g(x)$ і $r(x)$.

Тепер припустимо, що $d(x)$, є спільним дільником для многочленів $g(x)$ і $r(x)$. Тоді за властивостями подільності многочленів $d(x)$ буде ділить і многочлен $f(x) = q(x)g(x) + r(x)$. Тобто $d(x)$ буде й спільним дільником для $f(x)$ і $g(x)$.

Звідси випливає, що пари $(f(x), g(x))$ та $(g(x), r(x))$ мають однакові множини спільних дільників. Тоді найбільший спільний дільник буде однаковим для обох пар. □

Коротко алгоритм Евкліда для многочленів можна записати у наступному вигляді.

1. Поділимо многочлен $f$ на многочлен $g$ з остачею:

$$f = gq_1 + r_1, \text{ причому } \deg r_1 < \deg g.$$

Тоді з означення найбільшого спільного дільника і властивостей подільності отримуємо, що найбільший спільний дільник многочленів $f$ і $g$ буде також найбільшим спільним дільником многочленів $g$ і $r_1$, і навпаки.

2. Отже, пару многочленів $f$, $g$ можна замінити парою $g$, $r_1$, сума степенів яких є меншою за суму степенів початкових многочленів.

3. Якщо $r_1 = 0$, то шуканим найбільшим спільним дільником є многочлен $g$. Інакше переходимо на початок, тобто ділимо $g$ на $r_1$ з остачею і повторюємо вищенаведені міркування.

Оскільки на кожному кроці сума степенів многочленів зменшується, то рано чи пізно остача від ділення буде рівною 0. Найбільшим спільним дільником тоді буде остання ненульова остача. Записавши результати всіх виконаних ділень з остачею і послідовно виражаючи остачі через попередні, звідси можемо отримати подання найбільшого спільного дільника многочленів $f$ і $g$ у вигляді $mf + ng$.

Далі запишемо детальніше кроки даного алгоритму виписуючи відповідні формули.

Якщо один з многочленів $f(x)$ та $g(x)$ є нульовим, но найбільший спільний дільник цих многочленів дорівнює другому, не нульовому, многочлену. Тому розглядатимемо випадок коли обидва ці многочлени є ненульовими. Також вважатимемо, що $\deg f(x) \geq \deg g(x)$. Поділимо $f(x)$ на $g(x)$ з остачею

$$f(x) = g(x)q_1(x) + r_1(x),$$

тут $q_1(x)$ — частка, $r_1(x)$ — остача. Далі поділимо попередній дільник $g(x)$ поділимо на поточну остачу $r_1(x)$:

$$g(x) = r_1(x)q_2(x) + r_2(x),$$

тут $q_2(x)$ та $r_2(x)$ нові частка та остача відповідно. Наступним кроком знову ж поділимо попередній дільник $r_1(x)$ поділимо на поточну остачу $r_2(x)$:

$$r_1(x) = r_2(x)q_3(x) + r_3(x).$$

Продовжуючи так далі, з теореми про ділення з остачею отримуємо наступний ланцюг нерівностей

$$f(x) > g(x) > r_1(x) > r_2(x) > r_3(x) > \dots$$

Оскільки степені многочленів є невід'ємними числами, то цей ланцюг закінчиться маючи деяку скінченну кількість елементів. Тобто у деякий момент відбудеться ділення націло

$$\begin{aligned} r_{k-3}(x) &= r_{k-2}(x)q_{k-1}(x) + r_{k-1}(x), \\ r_{k-2}(x) &= r_{k-1}(x)q_k(x) + r_k(x), \\ r_{k-1}(x) &= r_k(x)q_{k+1}(x). \end{aligned}$$

**Теорема 8.5.** *Остання ненульова остача $r_k(x)$ з ланцюжка*

$$f(x),\, g(x),\, r_1(x), \dots,\, r_{k-3}(x),\, r_{k-2}(x),\, r_{k-1}(x),\, r_k(x)$$

*отриманого у результаті вище описаного алгоритму є найбільшим спільним дільником многочленів $f(x)$ та $g(x)$.*

*Доведення.* З останньої рівності для ділення многочленів, $r_{k-1}(x)$ ділиться на $r_k(x)$. Тоді $\mathsf{НСД}(r_{k-1}(x), r_k(x)) = r_k(x)$. З рівностей для ділення многочленів отриманих вище та попередньої леми маємо

$$\begin{aligned} r_k(x) = \mathsf{НСД}(r_{k-1}(x), r_k(x)) = \mathsf{НСД}(r_{k-2}(x), r_{k-1}(x)) = \dots = \\ = \mathsf{НСД}(r_1(x), r_2(x)) = \mathsf{НСД}(g(x), r_1(x)) = \mathsf{НСД}(f(x), g(x)). \quad \square \end{aligned}$$

Такий процес отримання найбільшого спільного дільника многочленів називають алгоритмом Евкліда, як бачимо він є аналогічним до пошуку найбільшого спільного дільника для двох цілих чисел.

Виписані вище рівності для ділення многочленів з алгоритма Евкліда використовуються для знаходження найбільшого спільного дільника у вигляді $\mathsf{НСД}(f(x), g(x)) = m(x)f(x) + n(x)g(x)$.

**Приклад 8.2.** Знайдемо найбільший спільний дільник многочленів

$$f(x) = x^4 + x^3 - 2x^2 + 3x - 3$$

і

$$g(x) = x^3 + x^2 - 2.$$

*Розв'язання.* Поділимо многочлен $f(x)$ на многочлен $g(x)$ з остачею

$$\begin{array}{r} x^4+x^3-2x^2+3x-3 = \big(x^3+x^2-2\big)\,x-2x^2+5x-3 \\ \underline{-\,x^4-x^3\qquad\quad +2x} \\ -\,2x^2+5x-3 \end{array}$$

Отримуємо частку $q_1(x)=x$ та остачу $r_1(x)=-2x^2+5x-3$.

Наступним кроком поділимо $g(x)$ на $r_1(x)$

$$\begin{array}{r} x^3\ +x^2\qquad\quad -2 = \big(-2x^2+5x-3\big)\big(-\frac{1}{2}x-\frac{7}{4}\big)+\frac{29}{4}x-\frac{29}{4} \\ \underline{-\,x^3+\frac{5}{2}x^2\ -\frac{3}{2}x} \\ \frac{7}{2}x^2\ -\frac{3}{2}x\ -2 \\ \underline{-\,\frac{7}{2}x^2+\frac{35}{4}x-\frac{21}{4}} \\ \frac{29}{4}x-\frac{29}{4} \end{array}$$

Отримуємо частку $q_2(x)=-\frac{1}{2}x-\frac{7}{4}$ та остачу $r_2(x)=\frac{29}{4}x-\frac{29}{4}$.

Ділимо $r_1(x)$ на $r_2(x)$

$$\begin{array}{r} -2x^2+5x-3 = \big(\frac{29}{4}x-\frac{29}{4}\big)\big(-\frac{8}{29}x+\frac{12}{29}\big) \\ \underline{2x^2-2x} \\ 3x-3 \\ \underline{-\,3x+3} \\ 0 \end{array}$$

Оскільки тут отримали нульову остачу, то попередня не нульова остача $r_2(x)$ і є найбільшим спільним дільником. Оскільки за домовленістю найбілишим спільним дільником вважаємо відповідний унітарний многочлен, то $\mathsf{НСД}(f(x),g(x))=x-1$. □

## 8.4 Корені многочленів

*Значенням многочлена*

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0\in F[x]$$

*в точці* $b\in F$ називається значення виразу

$$f(b)=a_nb^n+a_{n-1}b^{n-1}+\ldots+a_1b+a_0\in F.$$

Кожен многочлен $f\in F[x]$ визначає відображення з поля $F$ в $F$. Це відображення часто ототожнюють з самим цим многочленом. Для

нульового многочлена і многочленів нульового степеня це відображення є сталим і сам многочлен можна ототожнити з його значенням.

**Означення 8.7.** Коренем многочлена $f \in F[x]$ називається такий елемент $b \in F$, що $f(b) = 0$.

**Приклад 8.3.** 1. Коренем нульового многочлена є довільний елемент поля.

2. Многочлени нульового степеня не мають коренів.

3. Дійсне число $\sqrt{2}$ є коренем многочлена з дійсними коефіцієнтами $x^2 - 2$.

4. Многочлен $x^2 + 1$ не має дійсних коренів.

## 8.5 Зв'язок між відношенням подільності й поняттям кореня

**Теорема 8.6** (Теорема Безу)**.** *Значення многочлена $f(x)$ в точці $a$ дорівнює остачі від ділення $f(x)$ на $x - a$.*

*Доведення.* Поділимо многочлен $f(x)$ на $x - a$ з остачею:

$$f(x) = (x - a)q(x) + r(x), \text{ причому } \deg r < \deg(x - a),$$

тобто многочлен $r(x)$ або нульовий, або має нульовий степінь. Тому можна вважати, що $r(x) = r \in F$. Обчислюючи значення лівої та правої частин в точці $a$, отримуємо рівність $f(a) = r$, чим завершуємо доведення теореми. □

**Наслідок 8.3.** *Елемент $a \in F$ є коренем многочлена $f(x) \in F[x]$ тоді й тільки тоді, коли $f(x)$ ділиться на $x - a$.*

**Теорема 8.7.** *Кількість коренів многочлена степеня $n$ ($n \geq 0$) не перевищує $n$.*

*Доведення.* Індукція за $n$.

База індукції: при $n = 0$ маємо многочлен нульового степеня, у якого нема жодного кореня і тому твердження теореми справджується.

Нехай теорему доведено для всіх многочленів, степінь яких менший за $n$. Розглянемо многочлен $f(x)$ степеня $n$. Якщо він не має коренів, то твердження правильне. Інакше виберемо деякий його корінь $a$. Тоді

$f(x) = (x-a)f_1(x)$ для деякого многочлена $f_1(x)$ степеня $n-1$. Для довільного кореня $b$ многочлена $f(x)$, відмінного від $a$, маємо рівність $0 = f(b) = (b-a)f_1(b)$. Звідси $f_1(b) = 0$, тобто цей корінь буде коренем многочлена $f_1(x)$. Але за припущенням індукції цей многочлен має не більше $n-1$ кореня, звідки й отримуємо необхідне. □

**Означення 8.8.** Кратністю кореня $a$ многочлена $f(x)$ називається таке натуральне число $k$, що $(x-a)^k \mid f(x)$, але $(x-a)^{k+1}$ вже не є дільником $f(x)$.

Елемент $a$ — корінь $f(x)$ кратності $k$, якщо $f(x) = (x-a)^k f_1(x)$ для деякого многочлена $f_1(x)$, причому $f_1(a) \neq 0$. Означення є коректним, це випливає з наслідку. Корені многочлена, кратність яких рівна 1, будемо називати простими коренями, а корені вищої кратності — кратними. Будемо вважати, що елементи поля, які не є коренями многочлена, є його коренями кратності 0.

Наслідок з теореми Безу можна узагальнити з урахуванням кратності коренів

**Теорема 8.8.** *Нехай $a_1, \ldots, a_m$ — корені многочлена $f(x) \in F[x]$ кратностей $k_1, \ldots, k_m$ відповідно, то $f(x)$ має вигляд*

$$f(x) = (x-a_1)^{k_1} \ldots (x-a_m)^{k_m} g(x),$$

*причому жоден з елементів $a_1, \ldots, a_m$ не є коренем многочлена $g(x)$.*

**Теорема 8.9** (Теорема Вієта)**.** *Нехай многочлен*

$$f(x) = a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0$$

*можна розкласти на лінійні множники наступним чином*

$$f(x) = (x-b_1)(x-b_2)\ldots a_n(x-b_n).$$

*Тоді $b_1, b_2, \ldots, b_n$ є коренями даного многочлена і для них справедливі наступні формули, які пов'язують їх з коефіцієнтами многочлена*

$$b_1 + b_2 + \ldots + b_n = -\frac{a_{n-1}}{a_n},$$

$$b_1 b_2 + b_1 b_3 + \ldots + b_{n-1} b_n = \frac{a_{n-2}}{a_n},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\sum_{1\le i_1<i_2<\ldots<i_k\le n} b_{i_1}b_{i_2}\ldots b_{i_k} = (-1)^k\,\frac{a_{n-k}}{a_n},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$b_1b_2\ldots b_n = (-1)^n\,\frac{a_0}{a_n}.$$

*Доведення.* Для отримання шуканих формул достатньо перемножити даний в умові розклад многочлена $f(x)$. Тоді прирівняємо коефіцієнти при однакових степенях у розкритому добутку і у початковому многочлені $f(x)$. Запишемо які коефіцієнти отримано при відповідних степенях:

$$x^n: \quad a_n = a_n,$$

$$x^{n-1}: \quad a_{n-1} = -a_n\cdot(b_1+b_2+\ldots+b_n),$$

$$x^{n-2}: \quad a_{n-2} = a_n\cdot(b_1b_2+b_1b_3+\ldots+b_{n-1}b_n),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$x^{n-k}: \quad a_{n-k} = (-1)^k a_n\cdot \sum_{1\le i_1<i_2<\ldots<i_k\le n} b_{i_1}b_{i_2}\ldots b_{i_k},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$x^0: \quad a_0 = (-1)^n a_n\cdot b_1b_2\ldots b_n.$$

Звідси прямо отримуємо формули з твердження теореми. □

Отримані у теоремі формули, що пов'язують корені рівняння з коефіцієнтами рівняння, називають формулами Вієта.

## 8.6 Схема Горнера

Схема Горнера використовується для ділення з остачею многочлена

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \in F[x]$$

на многочлен $x - c$.

Для ділення довільного многочлена на двочлен $x-c$ можна скористатися наступним способом, який є простіший за використання ділення у стовпчик. Очевидно що у такому випадку результатом ділення буде многочлен степеня менше одиниці, тобто це буде деякий елемент поля $F$.

Розглянемо деякий многочлен та добуток, який отримано після ділення його на $x - c$

$$a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = (x - c)(b_{n-1}x^{n-1} + \ldots + b_1 x + b_0) + r,$$

де $r$ — остача від ділення, тобто деякий елемент поля, а $b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$ — частка від ділення. Тобто маємо у лівій і правій частині даної рівності два многочлени, вони будуть однаковими коли збігаються коефіцієнти при однакових степенях. Тому прирівняємо коефіцієнти при однакових степенях $x$ у лівій та правій частині та запишемо відповідні рівності:

$$\begin{gathered} x^n : a_n = b_{n-1}, \\ x^{n-1} : a_{n-1} = b_{n-2} - cb_{n-1}, \\ x^{n-2} : a_{n-2} = b_{n-3} - cb_{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^1 : a_1 = b_b - cb_1, \\ x^0 : a_0 = r - cb_0. \end{gathered}$$

З даних формул можемо отримати рекурентні формули для отримання коефіцієнтів $b_i$ многочлена частки та остачі $r$ через коефіцієнти початкового многочлена $a_{i+1}$ та $b_{i-1}$. Запишемо отримані рекурентні формули

$$\begin{gathered} b_{n-1} = a_n, \\ b_{n-2} = a_{n-1} + cb_{n-1}, \\ b_{n-3} = a_{n-2} + cb_{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ b_0 = a_1 + cb_1, \\ r = a_0 + cb_0. \end{gathered}$$

Початкові дані та результати зручно записувати як таблицю, таке подання називають схемою Горнена:

| | $a_n$ | $a_{n-1}$ | $a_{n-2}$ | $\ldots$ | $a_1$ | $a_0$ |
|---|---|---|---|---|---|---|
| $c$ | $b_{n-1}$ | $b_{n-2}$ | $b_{n-3}$ | $\ldots$ | $b_0$ | $r$ |

З отриманих рекурсивних рівностей випливає, що перший елемент другого рядка $b_{n-1}$ просто переписується з першого рядка, а кожен наступний елемент другого рядка обчислюється як сума числа, що стоїть над ним, та помноженого на $c$ числа зліва від нього.

**Приклад 8.4.** Поділимо з остачею многочлен $f(x) = x^3 + x^2 - 1$ на многочлен $x + 2$.

| | 1 | 1 | 0 | −1 |
|---|---|---|---|---|
| −2 | 1 | −1 | 2 | −5 |

Отже, частка рівна $x^2 - x + 2$, а остача рівна $-5$, $f(-2) = -5$, тому $-2$ — не корінь $f(x)$.

## 8.7 Похідна

**Означення 8.9.** Похідною многочлена

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-2}x^2 + a_{n-1}x + a_n \in F[x]$$

називається многочлен

$$f'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + \ldots + 2a_{n-2}x + a_{n-1} \in F[x].$$

Таке означення є формальним, на відміну від змістовного означення, яке дається за допомогою поняття границі. Тому воно підходить для дослідження многочленів з коефіцієнтами з довільного поля.

Зауважимо, що похідна многочлена нульового степеня, а також нульового многочлена, є нульовим многочленом. Також очевидно, що похідна многочлена додатного степеня має степінь, менший на одиницю: $\deg f'(x) = \deg f(x) - 1$, якщо $\deg f(x) > 0$.

**Теорема 8.10.** *Властивості похідної можна сформулювати у наступному вигляді:*

1. $(\lambda f(x))' = \lambda f'(x)$,
2. $(f(x) + g(x))' = f'(x) + g'(x)$,
3. $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$,
4. $(f_1f_2 \ldots f_k)'(x) = (f_1'f_2 \ldots f_k)(x) + (f_1f_2' \ldots f_k)(x) + \cdots + (f_1f_2 \ldots f_k')(x)$,
5. $((x-a)^k)' = k(x-a)^{k-1}$,
6. $(f^k(x))' = kf^{k-1}(x)f'(x)$.

*Доведення.* Доведення пунктів 1 та 2 випливає безпосередньо з означення.

Доведемо пункт 3, для цього записуємо добуток многочленів через суми відповідних елементів, далі обсимлюємо похідку суми за попередніми властивостями. Нехай

$$f(x)=\sum_{i=0}^{n}a_ix^i \quad \text{та} \quad g(x)=\sum_{j=0}^{m}b_jx^j.$$

Обчислимо похідну добутку цих многочленів

$$(f(x)g(x))'=\Big(\sum_{i=0}^{n}\sum_{j=0}^{m}a_ib_jx^{i+j}\Big)'=\sum_{i=0}^{n}\sum_{j=0}^{m}(i+j)a_ib_jx^{i+j-1}=$$

$$=\Big(\sum_{i=0}^{n}ia_ix^{i-1}\Big)\Big(\sum_{j=0}^{m}b_jx^j\Big)+\Big(\sum_{i=0}^{n}a_ix^i\Big)\Big(\sum_{j=0}^{m}jb_jx^{j-}\Big)=$$

$$=f'(x)g(x)+f(x)g'(x).$$

Пункт 4 випливає з пункта 3, якщо застосувати індукцію по $k$. Пункт 5 та 6 випливають з властивості 4. □

**Твердження 8.12.** *Нехай* $f(x)=a_0+a_1x+\ldots+a_nx^n$. *Тоді*

$$f'(x)=a_k\frac{k!}{0!}+a_{k+1}\frac{(k+1)!}{1!}x+\ldots+a_n\frac{n!}{(n-k)!}x^{n-k}=$$

$$=k!\left(a_k\binom{k}{0}+a_{k+1}\binom{k+1}{1}x+\ldots+a_n\binom{n}{n-k}x^{n-k}\right).$$

*Доведення.* Застосувавши властивість для похідної від суми многочленів, очевидно, що достатньо обчислити похідні кожного одночлена, який входить в многочлен $f(x)$. Оскільки кожен одночлен має вигляд $a_mx^m$, то за першою властивістю коефіцієнт $a_m$ виноситься за похідну. Тому залишається знайти вигляд похідної довільного порядку від $x^m$. Справджуються наступні рівності для $k$-ої похідної

$$(x^m)^{(k)}=m(x^{m-1})^{(k-1)}=m(m-1)(x^{m-2})^{(k-2)}=\ldots=$$

$$=m(m-1)\ldots(m-k+1)x^{m-k}=$$

$$=\frac{m!}{(m-k)!}x^{m-k}=k!\binom{m}{m-k}x^{m-k}.$$ □

**Теорема 8.11.** *Якщо корінь многочлена $f(x) \in F[x]$ має кратність $k$ і у полі $F$ виконана нерівність $k \neq 0$, то цей корінь є коренем кратності $k-1$ похідної цього многочлена. Зокрема, простий корінь многочлена не буде коренем його похідної.*

*Доведення.* Нехай елемент $a \in F$ є коренем кратності $k$ многочлена $f(x)$. Тоді $f(x) = (x-a)^k f_1(x)$ для деякого многочлена $f_1(x) \in F[x]$, причому елемент $a$ не є його коренем. Обчислимо похідну $f(x)$:

$$\begin{aligned} f'(x) = ((x-a)^k)' f_1(x) + (x-a)^k f_1'(x) = \\ = k(x-a)^{k-1} f_1(x) + (x-a)^k f_1'(x) = \\ = (x-a)^{k-1}(k f_1(x) + (x-a) f_1'(x)). \end{aligned}$$

Для многочлена $g(x) = k f_1(x) + (x-a) f_1'(x)$ маємо $g(a) = k f_1(a) \neq 0$. Це й означає, що $a$ є коренем кратності $k-1$ похідної многочлена $f(x)$. $\square$

**Теорема 8.12.** *Нехай $f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-2} x^2 + a_{n-1} x + a_n$ — многочлен з цілими коефіцієнтами, $p/q$ — його раціональний корінь. Тоді $p \mid a_n$, $q \mid a_0$.*

*Доведення.* Якщо корінь є нульовим, то твердження правильне.

Нехай далі $p \in \mathbb{Z}$, $p \neq 0$, $q \in \mathbb{N}$, причому $\mathsf{НСД}(p, q) = 1$. Спочатку розглянемо випадок, коли $a_0 = 1$. Підставимо в $f(x)$ корінь $p/q$:

$$\frac{p^n}{q^n} + a_1 \frac{p^{n-1}}{q^{n-1}} + \cdots + a_{n-2} \frac{p^2}{q^2} + a_{n-1} \frac{p}{q} + a_n = 0,$$

звідки

$$\frac{p^n}{q} + a_1 p^{n-1} + \cdots + a_{n-2} p^2 q^{n-3} + a_{n-1} p q^{n-2} + a_n q^{n-1} = 0.$$

Така рівність можлива лише тоді, коли $q = 1$, тобто корінь є цілим і рівним $p$. Тоді

$$p^n + a_1 p^{n-1} + \cdots + a_{n-2} p^2 + a_{n-1} p + a_n = 0,$$

звідки $p \mid a_n$.

В загальному випадку можна вважати, що $a_0 > 1$. Розглянемо многочлен

$$a_0^{n-1} f(x) = a_0^n x^n + a_1 a_0^{n-1} x^{n-1} + \cdots + a_{n-1} a_0^{n-2} a_0 x + a_n a_0^{n-1}.$$

Зробивши заміну змінної $y = a_0 x$, отримаємо многочлен

$$y^n + a_1 y^{n-1} + \cdots + a_{n-2} a_0^{n-3} y^2 + a_{n-1} a_0^{n-2} y + a_n a_0^{n-1}.$$

За доведеним, кожен раціональний корінь цього многочлена є цілим числом $p_1$, на яке ділиться $a_n a_0^{n-1}$. Тоді число $p_1/a_0$ буде коренем многочлена $f(x)$, і для нього виконується твердження теореми. □

## 8.8 Використання многочленів

Спеціальний тип многочленів виникає при обчислення визначників спеціального виду.

Нехай $a_1, \ldots, a_n$ — довільні елементи поля $F$, $n \geq 2$. Визначник

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

називається *визначником Вандермонда* і позначається $\Delta_n(a_1, \ldots, a_n)$.

**Приклад 8.5.** Визначник Вандермонда другого порядку обчислюється за означенням:

$$\Delta_2(a_1, a_2) = \begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = a_2 - a_1.$$

**Приклад 8.6.** Для обчислення визначника Вандермонда третього порядку спочатку віднімемо від третього стовпчика другий, домножений на $a_1$, потім від другого стовпчика перший, також домножений на $a_1$, а тоді розкладемо за першим рядком і винесемо спільні множники з обох рядків:

$$\Delta_3(a_1, a_2, a_3) = \begin{vmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_2 a_1 \\ 1 & a_3 - a_1 & a_3^2 - a_3 a_1 \end{vmatrix} =$$

$$= \begin{vmatrix} a_2 - a_1 & a_2^2 - a_2 a_1 \\ a_3 - a_1 & a_3^2 - a_3 a_1 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1) \begin{vmatrix} 1 & a_2 \\ 1 & a_3 \end{vmatrix} =$$

$$= (a_2 - a_1)(a_3 - a_1)\Delta_2(a_2, a_3) = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2).$$

**Теорема 8.13.**

$$\Delta_n(a_1,\dots,a_n)=\prod_{1\le i<j\le n}(a_j-a_i).$$

*Доведення.* Індукція за $n$.

База індукції: $n=2$. Рівність $\Delta_2(a_1,a_2)=a_2-a_1$ виконується.

Індуктивний крок. Для обчислення $\Delta_n(a_1,\dots,a_n)$ скористаємось тим же прийомом, що й при обчислення визначника Вандермонда третього порядку. А саме, від кожного стовпчика, починаючи з останнього, віднімемо попередній, домножений на $a_1$, тоді розкладемо за першим рядком, і винесемо з кожного рядка спільний множник його елементів.

$$\begin{vmatrix}1 & a_1 & a_1^2 & \cdots & a_1^{n-1}\\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1}\end{vmatrix}=\begin{vmatrix}1 & 0 & 0 & \cdots & 0\\ 1 & a_2-a_1 & a_2^2-a_2a_1 & \cdots & a_2^{n-1}-a_2^{n-2}a_1\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 1 & a_n-a_1 & a_n^2-a_na_1 & \cdots & a_n^{n-1}-a_n^{n-2}a_1\end{vmatrix}=$$

$$=(a_2-a_1)\dots(a_n-a_1)\begin{vmatrix}1 & a_2 & \cdots & a_2^{n-2}\\ 1 & a_3 & \cdots & a_3^{n-2}\\ \cdots & \cdots & \cdots & \cdots\\ 1 & a_n & \cdots & a_n^{n-2}\end{vmatrix}=$$

$$=(a_2-a_1)\dots(a_n-a_1)\Delta_{n-1}(a_2,\dots,a_n),$$

звідки, скориставшись припущенням індукції, отримуємо потрібну рівність. □

## 8.9 Інтерполяційна задача

Зафіксуємо поле $F$ і натуральне число $n$. Розглянемо $n$ пар елементів поля $F$

$$(a_1,b_1),(a_2,b_2),\dots,(a_n,b_n) \tag{8.1}$$

таких, що $a_i\ne a_j$ при $i\ne j$.

*Інтерполяційною задачею* з $n$ вузлами інтерполяції (8.1) називається задача знаходження такого многочлена $f$ над полем $P$, що $f(a_i)=b_i$, $1\le i\le n$. Коротко умову інтерполяційної задачі записують у вигляді таблиці

| $x$ | $a_1$ | $a_2$ | $\dots$ | $a_n$ |
|---|---|---|---|---|
| $f(x)$ | $b_1$ | $b_2$ | $\dots$ | $b_n$ |

**Теорема 8.14.** *Існує єдиний многочлен степеня, меншого за $n$, який є розв'язком інтерполяційної задачі з $n$ вузлами інтерполяції.*

*Доведення.* Потрібно довести існування і єдиність такого впорядкованого набору $(\alpha_1, \dots, \alpha_n) \in P^n$, що многочлен

$$f(x) = \alpha_1 + \alpha_2 x + \dots + \alpha_{n-1} x^{n-2} + \alpha_n x^{n-1}$$

в точці $a_i$ набуває значення $b_i$, $1 \le i \le n$.

Підставивши, отримаємо

$$\alpha_1 + \alpha_2 a_i + \dots + \alpha_{n-1} a_i^{n-2} + \alpha_n a_i^{n-1} = b_i, 1 \le i \le n.$$

Розглянемо ці рівності, як систему $n$ лінійних рівнянь з $n$ невідомими $\alpha_1, \dots, \alpha_n$. Основна матриця цієї системи має вигляд

$$\begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{pmatrix},$$

а її визначник — це визначник Вандермонда $\Delta_n(a_1, \dots, a_n)$. Оскільки числа $a_1, \dots, a_n$ попарно різні, то цей визначник не рівний нулю. Тому, за теоремою Крамера, система лінійних рівнянь має єдиний розв'язок. Це і означає існування і єдиність многочлена, вказаного в умові теореми. □

Застосувавши формули Крамера, можна обчислити коефіцієнти інтерполяційного многочлена. Проте можна вказати простіші способи знаходження інтерполяційного многочлена.

Далі розглянемо запис інтерполяційного многочлена у формі Лагранжа. Нехай задано інтерполяційну задачу

| $x$ | $a_1$ | $a_2$ | $\dots$ | $a_n$ |
|---|---|---|---|---|
| $f(x)$ | $b_1$ | $b_2$ | $\dots$ | $b_n$ |

.

Розглянемо таку інтерполяційну задачу

| $x$ | $a_1$ | $a_2$ | $\dots$ | $a_n$ |
|---|---|---|---|---|
| $L_1(x)$ | 1 | 0 | $\dots$ | 0 |

.

Оскільки $a_2, \dots, a_n$ — корені многочлена $L_1(x)$, то

$$L_1(x) = (x - a_2) \dots (x - a_n) h(x).$$

Оскільки існує інтерполяційний многочлен степеня $\leq n-1$, то $\deg h = 0$. Отже, $h(x) = \alpha \in P$ і з рівності $L_1(a_1) = 1$ отримуємо

$$\alpha = \frac{1}{(a_1 - a_2)\dots(a_1 - a_n)}.$$

Остаточно

$$L_1(x) = \frac{(x - a_2)\dots(x - a_n)}{(a_1 - a_2)\dots(a_1 - a_n)}.$$

Аналогічно для інтерполяційної задачі

| $x$ | $a_1$ | $a_2$ | $\dots$ | $a_n$ |
|---|---|---|---|---|
| $L_2(x)$ | 0 | 1 | $\dots$ | 0 |

отримуємо інтерполяційний многочлен

$$L_2(x) = \frac{(x - a_1)(x - a_3)\dots(x - a_n)}{(a_2 - a_1)(a_2 - a_3)\dots(a_2 - a_n)},$$

і т. д., для інтерполяційної задачі

| $x$ | $a_1$ | $a_2$ | $\dots$ | $a_n$ |
|---|---|---|---|---|
| $L_n(x)$ | 0 | 0 | $\dots$ | 1 |

отримуємо інтерполяційний многочлен

$$L_n(x) = \frac{(x - a_1)(x - a_2)\dots(x - a_{n-1})}{(a_n - a_1)(a_n - a_2)\dots(a_n - a_{n-1})}.$$

Тепер маємо інтерполяційний многочлен для початкової задачі:

$$f(x) = b_1 L_1(x) + b_2 L_2(x) + \dots + b_n L_n(x).$$

Тобто

$$f(x) = \sum_{i=1}^{n} b_i \frac{(x - a_1)\dots(x - a_{i-1})(x - a_{i+1})\dots(x - a_n)}{(a_i - a_1)\dots(a_i - a_{i-1})(a_i - a_{i+1})\dots(a_i - a_n)}.$$

Цей многочлен називається інтерполяційним многочленом у формі Лагранжа. Він має степінь, не більший за $n-1$, оскільки є сумою многочленів степеня $n-1$. При додаванні нових вузлів інтерполяції цей многочлен потрібно повністю перераховувати.

Також існує наступна форма запису для інтерполяційного многочлена. Розглянемо запис інтерполяційного многочлена у формі Ньютона. Нехай задано інтерполяційну задачу

| $x$ | $a_1$ | $a_2$ | $\ldots$ | $a_n$ |
|---|---|---|---|---|
| $f_n(x)$ | $b_1$ | $b_2$ | $\ldots$ | $b_n$ |

.

Розглянемо такі інтерполяційні задачі:

| $x$ | $a_1$ |
|---|---|
| $f_1(x)$ | $b_1$ |

,

| $x$ | $a_1$ | $a_2$ |
|---|---|---|
| $f_2(x)$ | $b_1$ | $b_2$ |

,

і т. д.,

| $x$ | $a_1$ | $a_2$ | $\ldots$ | $a_{n-1}$ |
|---|---|---|---|---|
| $f_{n-1}(x)$ | $b_1$ | $b_2$ | $\ldots$ | $b_{n-1}$ |

.

Тоді $f_1(x) = b_1$. Позначимо $c_1 = b_1$ і $f_n(x) = f(x)$. Для довільного $i$, $1 < i \leq n$ значення різниці многочленів $f_i - f_{i-1}$ в точках $a_1, \ldots, a_{i-1}$ рівні 0. Степінь цієї різниці не більший за $i - 1$. Тому

$$(f_i - f_{i-1})(x) = c_i(x - a_1) \ldots (x - a_{i-1})$$

для деякого $c_i \in P$. Послідовно виражаючи, отримуємо інтерполяційний многочлен

$$f(x) = c_1 + c_2(x - a_1) + c_3(x - a_1)(x - a_2) + \cdots + c_n(x - a_1)(x - a_2) \ldots (x - a_{n-1}).$$

Цей многочлен називається інтерполяційним многочленом у формі Ньютона. Коефіцієнти $c_1, c_2, \ldots, c_n$ у ньому не визначені. Їх знаходять, послідовно підставляючи вузли інтерполяції.

## 8.10 Вправи

1. Виконайте дії

$$(x^2 - x + 3)(x^3 - 1) - (x - 2)(3x^3 - 2x - 1).$$

2. Знайдіть степені і вкажіть старші коефіцієнти многочленів

   а) $0x^5 - 2x^3 + 5x^4 + 7x - 2$,

   б) $(2x^3 - 3x - 1)^7$.

3. Знайдіть степінь добутку многочленів

$$(x^5 + 2x^4 + 3x^3 + 4x^2 + 5x)(5x^3 - 4x^2 + 3x - 2).$$

4. Нехай $f(x) = 3x^2 - 4x + 2$ та $g(x) = 2x - 3$, обчисліть

   а) $(f + g)(x)$,

   б) $(f + 2g)(x)$,

   в) $(fg)(x)$.

5. Знайдіть вільний член та суму коефіцієнтів многочлена

$$(x^4 + x^3 - x - 1)^{2026}.$$

6. Доведіть тотожність

$$x^n - 1 = (x - 1)(x^{n-1} + x^{n-2} + \ldots + 1).$$

7. Нехай $c$ та $d$ є коренями многочлена $f(x) = x^2 + a_1x + a_2$ з невідомими коефіцієнтами $a_1$ та $a_2$. Знайдіть корені многочлена

$$h(x) = a_2x^2 + a_1(a_2 + 1)x + (a_2 - 1)^2 + a_1^2.$$

8. Нехай $f(x)$, $g(x)$, $h(x)$, $v(x)$ є ненульовими многочленами, для яких виконується рівність $f(x)g(x) = g(x)v(x)$. Доведіть, що $f(x) = v(x)$.

9. Знайдіть найбільший спільний дільник многочленів $f(x)$ та $g(x)$

   а) $f(x) = 2x^6 - 3x^4 + 4x^3 - 5x + 2$ та $g(x) = 3x^3 + 2x^2 + x - 1$;

   б) $f(x) = 2x^4 + 15x^3 + 31x^2 - 3x - 45$ та $g(x) = 8x^3 + 38x^2 + 27x - 45$.

10. За допомогою схеми Горнера поділіть з остачею многочлен $f(x)$ на $g(x)$, якщо

    а) $f(x) = 7x^5 + 7x^4 - 3x^2 - x - 1$ та $g(x) = x + 1$;

    б) $f(x) = x^6 + 2x^5 + 2x^4 + 4x^3 + 5x^2 + 11x + 3$ та $g(x) = x + 2$;

    в) $f(x) = x^4 - 3x^3 - 2x^2 + 10x - 6$ та $g(x) = x - 3$.

11. Визначити кратність кореня $a$ многочлена $f(x)$

а) $f(x) = x^4 - 3x^3 - 6x^2 + 28x - 24$ та $a = 1$;

б) $f(x) = x^4 - 8x^3 + 22x^2 - 24x + 9$ та $a = 3$.

12. Перевірити чи є незвідними многочлени над полем раціональних чисел:

а) $x^5 - 15x^3 + 25x^2 + 5x - 35$;

б) $x^4 - x^3 + 2x + 1$;

в) $x^{(p-1)p^k} + x^{(p-2)p^k} + \ldots + x^{p^k} + 1$, де $p$ — просте, а $k$ — натуральне число.

13. За яких умов многочлен $ax^n + bx^{n-1} + 1$ ділиться на многочлен $x^2 - 2x + 1$?

14. Нехай многочлен $a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ має корені $x_1$, $x_2$, $\ldots$, $x_n$. Які корені матимуть наступні многочлени

а) $a_n x^n - a_{n-1}x^{n-1} + a_{n-2}x^{n-2} - \ldots + (-1)^n a_0$;

б) $a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n$.

15. Сума двох коренів многочлена $2x^3 - x^2 - 7x + a$ дорівнює один. Знайдіть невідомий коефіцієнт $a$.

16. Знайдіть похідні многочлена $x^4 - 3x^3 - 6x^2 + 28x - 24$.

17. Довести, що 1 буде коренем кратності три многочлена

$$x^{2n} - nx^{n+1} + nx^{n-1} - 1$$

для довільного натурального $n$.

# Розділ 9

# Факторизація многочленів

**Означення 9.1.** Многочлени $f$ і $g$ називаються взаємно простими, якщо їх найбільший спільний дільник дорівнює 1.

**Теорема 9.1** (Критерій взаємної простоти многочленів)**.** *Многочлени $f$ і $g$ є взаємно простими тоді й тільки тоді, коли існують такі многочлени $m$ і $n$, що $mf + ng = 1$.*

*Доведення.* Необхідність випливає з означення взаємної простоти і теореми про існування найбільшого спільного дільника многочленів.

Доведемо достатність. Нехай існують такі многочлени $m$ і $n$, що $mf + ng = 1$ і многочлен $d$ є найбільшим спільним дільником $f$ і $g$. Тоді $d \mid f$, $d \mid g$, звідки $d \mid (mf)$, $d \mid (ng)$ і $d \mid (mf+ng)$. Отже, многочлени $d$ і 1 асоційовані, тобто многочлен 1 також є набільшим спільним дільником многочленів $f$ і $g$, що й завершує доведення. □

**Твердження 9.1.** *Якщо многочлен $f$ є взаємно простим з кожним із многочленів $g$ і $h$, то многочлени $f$ і $gh$ є взаємно простими.*

*Доведення.* За критерієм взаємної простоти для деяких многочленів $m_1$, $m_2$, $n_1$, $n_2$ мають місце рівності

$$1 = m_1 f + n_1 g \quad \text{і} \quad 1 = m_2 f + n_2 h.$$

Звідси

$$1 = m_1 f + n_1 g \cdot 1 = m_1 f + n_1 g(m_2 f + n_2 h) = (m_1 + n_1 g m_2) f + (n_1 n_2)(gh),$$

що за тим же критерієм і означає взаємну простоту многочленів $f$ і $gh$. □

**Твердження 9.2.** *Якщо многочлени $f$ і $g$ взаємно прості, а многочлен $h$ є таким, що $f \mid (gh)$, то $f \mid h$.*

*Доведення.* За критерієм взаємної простоти для деяких многочленів $m$, $n$ маємо рівність

$$1 = mf + ng.$$

Домножимо її на $h$:

$$h = (mh)f + n(gh).$$

Кожен з доданків у правій частині цієї рівності ділиться на $f$, звідки $f \mid h$. □

## 9.1 Незвідні многочлени

**Означення 9.2.** Многочлен $f(x) \in F[x]$ називається *незвідним* над $F$ якщо він не розкладається у добуток многочленів менших (додатніх) степенів над цим же $F$.

Зауважимо, що над різними множинами один й той самий многочлен може мати різні властивості щодо звідності чи незвідності.

Іншими словами, з рівності $f = gh$ випливає, що один з многочленів $g$, $h$ має нульовий степінь, а інший асоційований з $f$. Всі інші многочлени додатного степеня з $F[x]$ називаються звідними над $F$, тобто $f$ звідний, якщо $f = gh$, $\deg g > 0$, $\deg h > 0$.

Многочлен, асоційований з незвідним многочленом над полем $F$, сам буде незвідним над $F$.

Незвідність залежить від поля, над яким розглядаєтья многочлен: $x^2 + 1 = (x + i)(x - i)$ незвідний над $\mathbb{R}$, але звідний над $\mathbb{C}$.

**Приклад 9.1.** Кожен многочлен першого степеня є незвідним над полем, якому належать його коефіцієнти.

Справді, якби такий многочлен був звідним, то це означало б, що його степінь, число 1, можна записати, як суму двох натуральних чисел, а це неможливо.

**Приклад 9.2.** Кожен многочлен другого степеня над полем дійсних чисел, дискримінант якого від'ємний, є незвідним над полем $\mathbb{R}$.

Справді, нехай $f(x) = x^2 + bx + c \in \mathbb{R}$ і $D = b^2 - 4c < 0$. Припустимо, що такий многочлен звідний над $\mathbb{R}$, тобто $f(x) = (x + \alpha)(x + \beta)$ для деяких дійсних чисел $\alpha, \beta$. Тоді $b = \alpha + \beta$, $c = \alpha\beta$ і

$$D = (\alpha + \beta)^2 - 4\alpha\beta = (\alpha - \beta)^2 \geq 0.$$

Суперечність. Отже, многочлен $f$ є незвідним над полем $\mathbb{R}$.

**Лема 9.1.** *Нехай $f$ і $g$ — незвідні многочлени над полем $P$. Тоді вони або взаємно прості, або асоційовані.*

*Доведення.* Нехай $d$ — найбільший спільний дільник многочленів $f$ і $g$. Тоді $f = df_1$ і $g = dg_1$ для деяких многочленів $f_1, g_1$. З незвідності многочлена $f$ випливає, що або $\deg d = 0$, або $\deg d = \deg f > 0$. У першому випадку це означає, що многочлени $f$ і $g$ є взаємно простими. А у другому, враховуючи незвідність многочлена $g$, що $\deg f_1 = \deg g_1 = 0$, тобто кожен з многочленів $f$ і $g$ асоційований з $d$, а відтак $f$ і $g$ асоційовані. □

**Теорема 9.2** (Основна теорема арифметики для кільця многочленів)**.** *Кожен многочлен додатного степеня над полем $F$ розкладається в добуток незвідних многочленів над $F$, причому такий розклад однозначний з точністю до порядку множників та асоційованості.*

*Доведення.* Нехай $f(x) \in F[x]$, $\deg f > 0$. Доведемо, що вказаний розклад існує.

Проведемо індукцію за степенем многочлена $f$.

Базою індукції є значення степеня 1 і в цьому випадку сам многочлен є незвідним.

В загальному випадку, якщо многочлен $f$ є незвідним, то все доведено. Інакше існує розклад $f = f_1 f_2$, в якому степені обох множників додатні, а відтак обидва менші за степінь многочлена $f$. Тепер розглядаємо ці многочлени $f_1$ і $f_2$ меншого степеня і застосовуємо до них припущення індукції. Домноживши розклад одного з них на розклад іншого, отримаємо розклад для $f$.

Нехай $f = p_1 p_2 \dots p_k = q_1 q_2 \dots q_l$ для деяких незвідних многочленів $p_1,\ p_2,\ \dots,\ p_k,\ q_1,\ q_2,\ \dots,\ q_l,\ k, l \geq 1$. Тоді добуток $q_1 q_2 \dots q_k$ ділиться на многочлен $p_1$. За лемою многочлен $p_1$ або взаємно простий, або асоційований з кожним із многочленів $q_1, q_2, \dots, q_l$. Якщо припустити, що він взаємно простий із кожним з них, то він буде взаємно простим і з їхнім добутком, тобто не зможе бути дільником цього добутку і отримаємо суперечність. Отже, многочлен $p_1$ асоційований з, не обмежуючи загальності, многочленом $q_1$. Скорочуючи на $p_1$, отримаємо рівність добутків незвідних многочленів $p_2 \dots p_k = q_2' \dots q_l$.

Продовжуючи такі міркування далі, отримаємо рівність виду

$$1 = q_{k+1}' \dots q_l \text{ або } 1 = p_{l+1}' \dots p_k,$$

де всі множники є незвідними многочленами. Але це неможливо. Тому множників у початкових розкладах порівну, і кожному множнику одного розкладу відповідає асоційований з ним множник іншого так, що різним множникам першого розкладу відповідають різні множники іншого. □

**Теорема 9.3.** *Нехай $f(x)$ є многочленом з коефіцієнтами з деякого поля (наприклад $\mathbb{R}$ або $\mathbb{C}$) та степінь $f(x)$ дорівнює два або три. Тоді $f(x)$ є незвідним тоді і тільки тоді, коли не має нулів.*

*Доведення.* Нехай многочлен $f(x)$ має нуль $a$. Тоді його можна розкласти на множники наступним чином $f(x) = (x - a)h(x)$. Для випадку $\deg(f(x)) = 2$ маємо $\deg(h(x)) = 1$, та при $\deg(f(x)) = 3$ маємо $\deg(h(x)) = 2$. Таким чином бачимо, що $f(x)$ є звідним.

Тепер припустимо, що $f(x)$ є звідним. Тоді існують многочлени $h(x)$ та $g(x)$, такі, що $f(x) = h(x)g(x)$. Також з означення незвідного многочлена очевидно, що $\deg(h(x)) < \deg(f(x))$ і $\deg(g(x)) < \deg(f(x))$, та не втрачаючи загальності можемо вважати, що $\deg(g(x)) \leq \deg(h(x))$.

Скористаємося властивістю степеня добутку многочленів враховуючи, що $f(x) = h(x)g(x)$ маємо рівність $\deg(f(x)) = \deg(h(x)) + \deg(g(x))$. Оскільки $\deg(g(x)) \leq \deg(h(x))$, при $\deg(f(x)) = 2 = 1 + 1$ маємо $\deg(g(x)) = 1$, та $\deg(f(x)) = 3 = 2 + 1$ маємо $\deg(g(x)) = 1$. Тоді многочлен $g(x)$ є лінійним, тобто $g(x) = \alpha x + \beta$, де $\alpha \neq 0$. Тоді $a = -\beta/\alpha$ є нулем $g(x)$ а тоді й нулем $f(x)$. □

## 9.2 Многочлени з цілими коефіцієнтами

Розглянемо початкові відомості про многочлени з цілими коефіцієнтами.

Нехай $\mathbb{Z}[x]$ — кільце многочленів від однієї змінної з цілими коефіцієнтами. Многочлен $f(x) \in \mathbb{Z}[x]$ ділиться на $g(x) \in \mathbb{Z}[x]$, якщо існує $h \in \mathbb{Z}[x]$ такий, що $f(x) = g(x)h(x)$.

Зауважимо, що аналог теореми про ділення з остачею не виконується: $x$ не можна поділити на 2 ($x = 2 \cdot 1/2x$ в $\mathbb{Q}[x]$).

Многочлен $f \in \mathbb{Z}[x]$ — незвідний над $\mathbb{Z}$, якщо його не можна подати у вигляді добутку двох многочленів додатного степеня з цілими коефіцієнтами.

**Означення 9.3.** Вмістом многочлена

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 \in \mathbb{Z}[x]$$

називається найбільший спільний дільник його коефіцієнтів

$$\operatorname{cont}(f) = \mathsf{НСД}(a_0, a_1, \ldots, a_{n-2}, a_{n-1}, a_n).$$

Зауважимо, що вміст визначений однозначно з точністю до знака.

**Приклад 9.3.** $\operatorname{cont}(6x^3 - 24x^2 + 15) = \mathsf{НСД}(6, 24, 15) = 3$

**Твердження 9.3.** *$f(x) = \operatorname{cont}(f) \cdot f_1(x)$, причому $\operatorname{cont}(f_1) = 1$*

**Приклад 9.4.** $6x^3 - 24x^2 + 15 = 3 \cdot (2x^3 - 8x^2 + 5)$, де $\operatorname{cont}(2x^3 - 8x^2 + 5) = \mathsf{НСД}(2, 8, 5) = 1$

**Лема 9.2** (Лема Гауса)**.** *Для довільних $f, g \in \mathbb{Z}[x]$ має місце рівність*

$$\operatorname{cont}(f \cdot g) = \operatorname{cont}(f) \cdot \operatorname{cont}(g).$$

*Доведення.* З рівності

$$f \cdot g = \operatorname{cont}(f) f_1 \cdot \operatorname{cont}(g) g_1, \text{ де } \operatorname{cont}(f_1) = \operatorname{cont}(g_1) = 1,$$

випливає, що досить перевірити твердження леми для многочленів з вмістом 1.

Нехай маємо многочлени з цілими коефіцієнтами і вмістом 1:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0,$$

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_2 x^2 + b_1 x + b_0.$$

Позначимо $h(x) = f(x) \cdot g(x)$ і нехай

$$h(x) = c_{n+m} x^{n+m} + c_{n+m-1} x^{n+m-1} + \cdots + c_2 x^2 + c_1 x + c_0.$$

Позначимо $d = \operatorname{cont} h$ і припустимо, що $d > 1$. Тоді $d$ має прості дільники. Нехай $p$ — довільний простий дільник числа $d$. Тоді усі коефіцієнти многочлена $h$ діляться на $p$. В той же час, оскільки вміст кожного з початкових многочленів рівний 1, як у многочлена $f$, так і в многочлена $g$ знайдуться коефіцієнти, що не діляться на $p$.

Нехай $r$ — найменший такий номер, що коефіцієнт $a_r$ не ділиться на $p$, а $s$ — найменший такий номер, що коефіцієнт $b_s$ не ділиться на $p$. Зокрема, тоді на $p$ діляться всі коефіцієнти

$$a_{r-1}, a_{r-2}, \ldots, a_1, a_0 \text{ і } b_{s-1}, b_{s-2}, \ldots, b_1, b_0.$$

Але тоді в многочлена $h$ також знайдеться коефіцієнт, котрий не ділиться на $p$. Справді, коефіцієнт

$$c_{r+s} = a_r b_s + (a_{r+1}b_{s-1} + a_{r+2}b_{s-2} + \cdots) + (a_{r-1}b_{s+1} + a_{r-2}b_{s+2} + \cdots)$$

не ділиться на $p$.

Суперечність. Отже, $d = 1$, що й потрібно було довести. □

Завжди зручно використовувати теореми критерії, які дають необхідну і достатню умову для перевірки деякого факту. Так важливий і красивий результат сформульований Гаусом — це критерій розкладності на множники многочлена з цілими коефіцієнтами.

**Теорема 9.4.** *Нехай $f(x) \in \mathbb{Z}[x]$ — многочлен з цілими коефіцієнтами. Многочлен $f(x)$ можна розкласти у добуток двох многочленів степенів $n$ та $m$ з цілими коефіцієнтами тоді і тільки тоді, коли многочлен $f(x)$ можна розкласти у добуток двох многочленів тих самих степенів $n$ та $m$ з раціональними коефіцієнтами.*

*Доведення.* З незвідності многочлена з цілими коефіцієнтами над $\mathbb{Q}$ випливає незвідність над $\mathbb{Z}$, оскільки кожне ціле число є раціональним.

Доведемо в інший бік. А саме, покажемо, що зі звідності над $\mathbb{Q}$ випливає звідність над $\mathbb{Z}$. Розглянемо многочлен $f(x) \in \mathbb{Z}[x]$ додатного степеня. Нехай $f = gh$ для деяких $g(x), h(x) \in \mathbb{Q}[x]$. Без обмеження загальності можемо вважати, що $\operatorname{cont}(f) = 1$. Зафіксуємо таке натуральне число $m$, що $mg(x) \in \mathbb{Z}[x]$. Позначимо $n = \operatorname{cont}(mg)$, $n > 0$. Тоді для додатного раціонального числа $r = m/n$ виконуються умови: $rg(x) \in \mathbb{Z}[x]$ і $\operatorname{cont}(rg) = 1$. Аналогічно отримуємо додатне раціональне число $s$ таке, що: $sh(x) \in \mathbb{Z}[x]$ і $\operatorname{cont}(sh) = 1$. Покажемо, що тоді $rs = 1$. Справді, за лемою Гауса, $\operatorname{cont}(rg)\operatorname{cont}(sh) = \operatorname{cont}(rgsh)$, тобто $1 = \operatorname{cont}((rs)f)$. Звідси отримуємо розклад $f(x) = (rg(x))(sh(x))$, який є розкладом над $\mathbb{Z}$, що й потрібно було довести. □

**Наслідок 9.1.** *Нехай $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{Z}[x]$, де $a_0 \neq 0$. Якщо $f(x)$ має нуль в $\mathbb{Q}$, тоді $f(x)$ має корінь $b \in \mathbb{Z}$ та $b$ ділить $a_0$.*

*Доведення.* Якщо $b$ є нулем $f(x)$, то $(x - b)$ є лінійним дільником з $\mathbb{Q}[x]$. За теоремою 9.4 многочлен $f(x)$ має лінійний дільник з $\mathbb{Z}$, тобто

$$f(x) = (\alpha x + \beta)g(x).$$

Розглянемо старший коефіцієнт, отриманий з цього запису, і врахуємо, що старший коефіцієнт $f(x)$ дорівнює 1. Бачимо, що $\alpha$ має бути дільником 1, тобто $\alpha = \pm 1$. Оскільки можна замінити $g(x)$ на $-g(x)$ можемо вважати, що $\alpha = 1$. Покладемо $\beta = -a$, тоді

$$f(x) = (x - b)g(x).$$

Тоді $a \in \mathbb{Z}$ є нулем $f(x)$. Оскільки вільним членом $f(x)$ є $a_0$, то з запису $f(x) = (x - b)g(x)$ бачимо, що $b$ має бути дільником $a_0$. □

Дана ознака узагальнюється на випадок кілець з одиницею, які є факторіальними.

**Теорема 9.5.** *Розглянемо $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ є многочленом з цілими коефіцієнтами. Нехай існує просте число $p$ таке, що $p$ ділить $a_i$, при $i \leq n-1$, $p$ не ділить $a_n$ та $p^2$ не ділить $a_0$. Тоді $f(x)$ є незвідним у $\mathbb{Q}[x]$.*

*Доведення.* Доведемо від супротивного.

Нехай існує розклад $f(x) = g(x)h(x)$, де $g(x)$ та $h(x)$ є многочленами з цілими коефіцієнтами. Можемо записати

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0,$$
$$h(x) = c_h x^h + c_{h-1}x^{h-1} + \ldots + c_1 x + c_0.$$

Враховуючи останній запис та рівність $f(x) = g(x)h(x)$, маємо, що вільний член $f(x)$ дорівнює з одного боку $a_0$ а з іншого боку $b_0 c_0$. Оскільки $a_0$ ділиться на просте число $p$ та $a_0 = b_0 c_0$ хоча б одне з чисел $b_0$ або $c_0$ має ділитися на $p$. Не втрачаючи загальності розглянемо випадок, коли $b_0$ ділиться на $p$. Також з твердження теореми ми маємо, що $a_0$ не ділиться на $p^2$. Тоді оскільки $a_0 = b_0 c_0$ та $p \mid b_0$ отримуємо, що $c_0$ не ділиться на $p$.

Якщо б усі коефіцієнти многочлена $g(x)$ ділились на $p$, то усі коефіцієнти многочлена $f(x)$ також ділились би на $p$. Це суперечить твердженню теореми, оскільки старший коефіцієнт многочлена $f(x)$, який дорівнює $a_n$, не ділиться на $p$. Тому для виконання умов теореми необхідно, щоб існував хоча б один коефіцієнт, який не ділиться на $p$. Тоді розглянемо такий коефіцієнт $b_i$, який не ділиться на $p$ та має мінімальний номер $i$. Тоді розглянемо коефіцієнт $a_i$ многочлена $f(x)$, зауважимо, що $p \mid a_i$. Для даного коефіцієнта маємо запис

$$a_i = b_0 c_i + b_1 c_{i-1} + \ldots + b_{i-1} c_1 + b_i c_0$$

доданок $b_i c_0$ не ділиться на $p$, проте всі інші доданки діляться на $p$ (з вибору $b_i$, всі коефіцієнти $b_0$, $b_1$, ..., $b_{i-1}$, мають менші номери, тому діляться на $p$), тоді $a_i$ не ділиться на $p$. Отримали суперечність.

Тобто наше припущення, про те, що $f(x) = g(x)h(x)$ суперечить умовам теореми, тобто воно не вірне, та $f(x)$ є незвідним. □

## 9.3 Многочлени над $\mathbb{C}$ і $\mathbb{R}$

**Теорема 9.6** (Основна теорема алгебри)**.** *Кожен многочлен додатного степеня з комплексними коефіцієнтами має комплексний корінь.*

Також основна теорема алгебри може формулюватися наступним чином

**Теорема 9.7** (Основна теорема алгебри, альтернативне формулювання)**.** *Кожен многочлен додатного степеня $n$ з комплексними коефіцієнтами має рівно $n$ коренів враховуючи їх кратність.*

Існує багато способів доведення основної теореми алгебри. Далі розглянемо декілька теореми, які самі є цікавими і також можуть бути використані для доведення основної теореми алгебри.

Доведемо цю теорему трохи пізніше.

**Теорема 9.8** (Теорема Руше)**.** *Нехай $f$ та $g$ — многочлени та $\gamma$ — замкнена крива без самоперетинів на комплексній площині. Якщо*

$$|f(z) - g(z)| < |f(z)| + |g(z)|$$

*для всіх точок $z \in \gamma$, то в середині $\gamma$ лежить однакова кількість коренів многочленів $f$ та $g$ (враховуючи кратність).*

*Доведення.* Розглянемо на комплексній площині векторні поля $v(z) = f(z)$ та $w(z) = g(z)$. З умови $|f(z) - g(z)| < |f(z)| + |g(z)|$ випливає, що в жодній точні кривої $\gamma$ вектори $v$ і $w$ не є протилежно направленими. Розглянемо векторне поле $v_t = tv + (1-t)w$, так $v_0 = w$ та $v_1 = v$. Зауважимо, що у довільній точці $z \in \gamma$ вектор $v_t(z)$ не нульовий. Тоді індекс $ind(t)$ кривої $\gamma$, який дорівнює кількості оборотів вектора $v_t(z)$ при повному обході точки $z$ вздовж кривої $\gamma$, визначений відносно поля $v_t$. Ціле число $ind(t)$ неперервно залежить від $t$, тому $ind(t)$ є константою. Також індекси кривої $\gamma$ відносно векторних полів $v$ та $w$ співпадають.

Також, індекс кривої $\gamma$ відносно векторного поля $v$ дорівнює сумі особливих точок у яких $v(z) = 0$. Індекс особливої точки $z_0$ визначається як індекс кривої $|z - z_0| = \varepsilon$ при достатньо малому $\varepsilon$. Для векторного поля $v(z) = f(z)$ індекс особливої точки $z_0$ дорівнює кратності кореня $z_0$ многочлена $f$. Оскільки індекси кривої $\gamma$ співпадають відносно векторних полів $v(z) = f(z)$ та $w(z) = g(z)$, тоді в середині кривої $\gamma$ лежить однакова кількість коренів многочленів $f$ та $g$. □

Як приклад кривої без самоперетинів можемо розглянути коло, для такої кривої маємо наступну теорему, яка дає оцінку модуля довільного кореня многочлена $f$. Ця теорема також показує справедливість основної теореми алгебри і навіть є ширшою бо дає оцінку модулів отриманих коренів.

**Теорема 9.9.** *Нехай $f(z) = z^n + a_1 z^{n-1} + \ldots + a_n$, де $a_i \in \mathbb{C}$. Тоді у колі $|z| = 1 + \max|a_i|$ лежить рівно $n$ коренів многочлена $f$ (враховуючи кратність).*

*Доведення.* Нехай $a = \max|a_i|$. Многочлен $g(z) = z^n$ має у середині даного кола корінь 0 кратності $n$. Тому достатньо перевірити, що при $|z| = 1 + a$ маємо $|f(z) - g(z)| < |f(z)| + |g(z)|$.

Доведемо сильніше твердження: $|f(z) - g(z)| < |g(z)|$, тобто

$$|a_1 z^{n-1} + \ldots + a_n| < |z|^n.$$

Враховуючи, що $|z| = 1 + a$, маємо

$$|a_1 z^{n-1} + \ldots + a_n| \le a(|z|^{n-1} + \ldots + 1) = a\frac{|z|^n - 1}{|z| - 1} = |z|^n - 1 < |z|^n. \ \square$$

**Теорема 9.10** (Теорема Коші)**.** *Нехай $f(x) = x^n - b_1 x^{n-1} - \ldots - b_n$, де всі числа $b_i$ є невід'ємними і хоча б одне з них не нульове. Якщо найбільший спільний дільник індексів додатних коефіцієнтів $b_i$ дорівнює 1, то $f$ має єдиний додатний корінь $p$, а абсолютні значення інших коренів менші за $p$.*

*Доведення.* Нехай додатними будуть тільки коефіцієнти $b_{k_1}, b_{k_2}, \ldots, b_{k_m}$, та $k_1 < k_2 < \cdots < k_m$. Оскільки найбільший спільний дільник $k_1$, $k_2$, $\ldots$, $k_m$ Дорівнює одиниці, то існують цілі числа $t_1$, $t_2$, $\ldots$, $t_m$, для яких $t_1 k_1 + t_2 k_2 + \ldots + t_m k_m = 1$.

Розглянемо функцію

$$F(x) = \frac{b_{k_1}}{x^{k_1}} + \cdots + \frac{b_{k_m}}{x^{k_m}} - 1.$$

Рівняння $F(x) = 0$ має єдиний додатний розв'язок $p$. Нехай $x$ є довільним іншим не нульовим коренем многочлена $f$. Покладемо $q = |x|$, тоді

$$1 = \frac{b_{k_1}}{x^{k_1}} + \cdots + \frac{b_{k_m}}{x^{k_m}} \le \left|\frac{b_{k_1}}{x^{k_1}}\right| + \cdots + \left|\frac{b_{k_m}}{x^{k_m}}\right| = \frac{b_{k_1}}{q^{k_1}} + \cdots + \frac{b_{k_m}}{q^{k_m}},$$

тобто $F(q) \ge 0$.

Зрозуміло, що рівність $F(q) = 0$ можлива тільки якщо

$$\frac{b_{k_i}}{x^{k_i}} = \left|\frac{b_{k_i}}{x^{k_i}}\right| > 0 \text{ для всіх } i.$$

Проте у такому випадку

$$\frac{b_{k_1}^{t_1} \cdot \ldots \cdot b_{k_m}^{t_m}}{x} = \left(\frac{b_{k_1}}{x^{k_1}}\right)^{t_1} \cdot \ldots \cdot \left(\frac{b_{k_m}}{x^{k_m}}\right)^{t_m} > 0,$$

тобто $x > 0$. Це суперечить тому, що $x \ne p$ та $p$ є єдиним додатним коренем рівняння $F(x) = 0$. Тоді $F(q) > 0$, та $F(x)$ — монотонна для додатного $x$, звідси випливає $q < p$. □

З даної теореми можемо отримати наближення для абсолютних значень коренів многочлена з додатними коефіцієнтами.

**Теорема 9.11** (Енестрома-Какеї). *Нехай всі коефіцієнти многочлена $g(x) = a_0x^{n-1} + \ldots + a_{n-1}$ є додатними. Тоді для довільного кореня $\chi$ даного многочлена, справедливі нерівності*

$$\min_{1\le i\le n-1}\left\{\frac{a_i}{a_{i-1}}\right\} = \delta \le |\xi| \le \gamma = \max_{1\le i\le n-1}\left\{\frac{a_i}{a_{i-1}}\right\}.$$

*Доведення.* Розглянемо многочлен

$$(x-\gamma)g(x) = a_0x^n - (\gamma a_0 - a_1)x^{n-1} - \ldots - (\gamma a_{n-2} - a_{n-1})x - \gamma a_{n-1}.$$

За означенням $\gamma \ge \frac{a_i}{a_{i-1}}$, тобто $\gamma a_{i-1} - a_i \ge 0$. Тоді за теоремою Коші $\gamma$ є єдиним додатнім коренем многочлена $(x-\gamma)g(x)$, а модулі усіх інших коренів цього многочлена є більшими за $\gamma$.

Якщо $\xi$ є коренем многочлена $g$, то $\eta = \xi^{-1}$ є коренем многочлена $a_{n-1}y^{n-1} + \ldots + a_0$. Тому

$$|\xi|^{-1} = |\eta| = \max_{1\le i\le n-1}\left\{\frac{a_{i-1}}{a_i}\right\} = \left(\min_{1\le i\le n-1}\left\{\frac{a_i}{a_{i-1}}\right\}\right)^{-1}.$$

Тобто

$$|\xi| \ge \min_{1\le i\le n-1}\left\{\frac{a_i}{a_{i-1}}\right\}. \qquad \square$$

Узагальненням даної теореми є теорема Перрона фробеніуса яка дає обмеження на власні значення матриць (які є конерями характеристичного многочлена матриці).

Наслідки з основної теореми алгебри.

**Наслідок 9.2.** *1. Кожен многочлен, незвідний над полем комплексних чисел, має степінь 1.*

*2. Кожен многочлен додатного степеня з комплексними коефіцієнтами розкладається в добуток многочленів першого степеня з комплексними коефіцієнтами.*

*3. Сума кратностей комплексних коренів многочлена з комплексними коефіцієнтами дорівнює його степеневі.*

*Доведення.* 1. Випливає з основної теореми алгебри та наслідка з теореми Безу.

2. Випливає з першого і теореми про розклад многочлена на незвідні множники.

3. Випливає з другого. $\square$

Звідси можемо отримати і опис незвідних над полем дійсних чисел многочленів. Для цього многочлени з дійсними коефіцієнтами будемо розглядати як многочлени над полем комплексних чисел.

## 9.4 Многочлени Лаґерра

Нехай $z_1, z_2, \ldots, z_n \in \mathbb{C}$ це точки одиничної маси. Точка

$$\zeta = \frac{1}{n}(z_1 + \ldots + z_n)$$

називається *центром мас точок* $z_1, z_2, \ldots, z_n$.

Такий запис можна узагальнити наступним чином. Застосуємо дробово-лінійне перетворення $w$, яке спрямовує $z_0$ до $\infty$, тобто

$$w(z) = \frac{a}{z - z_0} + b.$$

Тепер знайдемо центр мас образів $z_1, z_2, \ldots, z_n$ а потім застосуємо обернене перетворення $w^{-1}$. Виконавши перетворення стає зрозуміло, що результат не залежить від $a$ та $b$, так ми отримуємо точку

$$\zeta_{z_0} = z_0 + n \cdot \frac{1}{\frac{1}{z_1 - z_0} + \ldots + \frac{1}{z_n - z_0}}$$

яка називається *центром мас точок* $z_1, z_2, \ldots, z_n$ *з урахуванням* $z_0$.

Очевидно, що центр мас точок $z_1, z_2, \ldots, z_n$ лежить в середині їх опуклої оболонки.

Це твердження легко можна узагальнити на випадок центру мас з урахуванням $z_0$. Треба тільки замінити лінії, що з'єднують $z_i$ та $z_j$ на кола, що проходять через $z_i$, $z_j$ та $z_0$. Точка $z_0$ відповідає $\infty$ лежить зовні опуклої оболонки.

**Теорема 9.12.** *Нехай* $f(z) = (z - z_1) \cdot \ldots \cdot (z - z_n)$. *Тоді центр мас коренів* $f$ *з урахуванням довільної точки* $z$ *задається формулою*

$$\zeta_z = z - n \frac{f(z)}{f'(z)}.$$

*Доведення.* Очевидно

$$\frac{f'(z)}{f(z)} = \frac{1}{z - z_1} + \ldots + \frac{1}{z - z_n}.$$

Тоді твердження теореми відразу випливає з формули для $\zeta_{z_0}$. □

**Теорема 9.13** (Лаґерра)**.** *Нехай* $f(z)$ *є многочленом степеня* $n$, *а* $x$ — *його простий корінь. Тоді центр мас усіх інших коренів многочлена* $f(z)$ *з урахуванням* $x$ *є точка*

$$X = x - 2(n - 1) \frac{f'(x)}{f''(x)}.$$

*Доведення.* Нехай $f(z) = (z-x)F(z)$. Тоді $f'(z) = F(z) + (z-x)F'(z)$ та $f''(z) = 2F'(z) + (z-x)F''(z)$. Тому $f'(x) = F(x)$ та $f''(x) = 2F'(x)$. Застосовуючи попередню теореми до многочлена $F$ степеня $n-1$ і покладаючи $z = x$ отримаємо твердження теореми. □

**Теорема 9.14** (Лаґерра)**.** *Нехай $f(x)$ є многочленом степеня $n$ та*

$$X(z) = z - 2(n-1)\frac{f'(z)}{f''(z)}.$$

*Нехай коло (чи лінія) $C$ проходить через простий корінь $z_1$ многочлена $f$, а інші корені $f$ належать одній з областей, на які $C$ розбиває площину. Тоді $X(z_1)$ належить тій самій області.*

*Доведення.* У звичному випадку центру має коло $C$ відповідає лінії, такій, що всі корені $f(z)$, крім $z_1$, лежать по один бік від неї. Центр має цих коренів лежить по той самий бік від цієї лінії. □

**Наслідок 9.3.** *Нехай $z_1$ є одним з простих коренів $f$ з максимальним абсолютним значенням. Тоді $|X(z_1)| \le |z_1|$, тобто*

$$\left| z_1 - 2(n-1)\frac{f'(z_1)}{f''(z_1)} \right| \le |z_1|.$$

*Доведення.* Усі корені $f$ лежать у диску $\{z \in \mathbb{C} : |z| < |z_1|\}$. Тому $X(z_1)$ також належить цьому диску. □

**Теорема 9.15.** *Нехай $f$ є многочленом з дійсними коефіцієнтами і задано*

$$\zeta_z = z - n\frac{f(z)}{f'(z)}.$$

*Усі корені $f$ є дійсними тоді і тільки тоді, коли $\operatorname{Im} z \cdot \operatorname{Im} \zeta_z < 0$ для довільного $z \in \mathbb{C} \setminus \mathbb{R}$.*

*Доведення.* Нехай всі корені $f$ є дійсними та $\operatorname{Im} z = a > 0$. Лінія, що складається з точок, з уявною частиною $\varepsilon$, де $0 < \varepsilon < a$, відділяє точку $z$ від усіх коренів $f$ оскільки вони належать дійсній осі. Тоді $\operatorname{Im} \zeta_z < \varepsilon$. Якщо спрямувати $\varepsilon$ до нуля, то $\operatorname{Im} \zeta_z < 0$.

Легко перевірити, що $\operatorname{Im} \zeta_z = 0$ не досягається. Нехай $\zeta_z \in \mathbb{R}$. Розглянемо коло, що проходить через $z$ і дотикається до дійсної осі у $\zeta_z$. Здійснюючи незначне пересування кола ми можемо отримати коло для

якого точки $z$ та $\zeta_z$ лежатимуть по один бік від нього, а усі інші корені $f$ лежатимуть по інший бік. Для випадку $\operatorname{Im}\zeta_z = a < 0$, міркування подібні.

Доведемо тепер в інший бік. Нехай $\operatorname{Im} z \cdot \operatorname{Im}\zeta_z < 0$ для довільного $z \in \mathbb{C} \setminus \mathbb{R}$. Покладемо $z_1$ коренем $f$ таким, що $\operatorname{Im}(z_1) \neq 0$. Тоді $\lim_{z\to z_1} \zeta_z = z_1$, а тому $\operatorname{Im} z \cdot \operatorname{Im}\zeta_z > 0$. □

## 9.5 Аполярні многочлени

Нехай $f(x)$ є многочленом степеня $n$, $\zeta$ — фіксоване число або $\infty$. Похідною многочлена $f(x)$ відносно точки $\zeta$ називають функцію

$$A_\zeta f(z) = \begin{cases} (\zeta - z)f'(z) + nf(z), & \text{при } \zeta \neq \infty; \\ f'(z), & \text{при } \zeta = \infty. \end{cases}$$

Для многочлена

$$f(z) = \sum_{k=0}^{n} \binom{n}{k} a_k z^k$$

маємо відповідну рівність для похідної

$$\frac{1}{n} f'(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} a_{k+1} z^k.$$

Тоді похідна многочлена $f(x)$ відносно точки $\zeta$ визначається з наступної рівності

$$\frac{1}{n} A_\zeta f(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} (a_k + a_{k+1}\zeta) z^k.$$

Нехай $z_1, \ldots, z_n$ є коренями многочлена

$$f(z) = \sum_{k=0}^{n} \binom{n}{k} a_k z^k,$$

а $\zeta_1, \ldots, \zeta_n$ є коренями многочлена

$$g(z) = \sum_{k=0}^{n} \binom{n}{k} b_k z^k.$$

Відповідно з формули для похідної многочленів такого виду відносно точки маємо, що

$$\frac{1}{n!}A_{\zeta_1}A_{\zeta_2}\dots A_{\zeta_n}f(z) = a_0 + a_1\sigma_1 + a_2\sigma_2 + \dots + a_n\sigma_n,$$

де

$$\sigma_1 = \zeta_1 + \zeta_2 + \dots + \zeta_n = -\binom{n}{1}\frac{b_{n-1}}{b_n},$$

$$\sigma_2 = \zeta_1\zeta_2 + \dots + \zeta_{n-1}\zeta_n = \binom{n}{2}\frac{b_{n-2}}{b_n},$$

$$\dots\dots\dots$$

$$\sigma_n = \zeta_1 \cdot \zeta_2 \cdot \dots \cdot \zeta_n = (-1)^n\,\frac{b_0}{b_n}.$$

Тоді рівність $A_{\zeta_1}A_{\zeta_2}\dots A_{\zeta_n}f(z) = 0$ еквівалентна рівності

$$a_0b_1 - \binom{n}{1}a_1b_{n-1} + \binom{n}{2}a_2b_{n-2} + \dots + (-1)^n a_nb_0 = 0.$$

Многочлени $f$ та $g$ коефіцієнти яких задовольняють останній формулі називають аполярними.

Круговою областю називають внутрішню або зовнішню частини круга або напівплощини.

**Лема 9.3.** *Нехай усі корені $z_1, \dots, z_n$ многочлена $f(z)$ лежать всередині кругової області $K$, а точка $\zeta$ лежить зовні $K$. Тоді всі корені многочлена $A_\zeta f(z)$ лежать всередині $K$.*

**Теорема 9.16** (Грейс)**.** *Нехай $f$ та $g$ є аполярними многочленами. Якщо усі корені многочлена $f$ лежать у круговій області $K$, то хоча б один корінь $g$ також лежить у області $K$.*

*Доведення.* Нехай $w_i$ є коренем многочлена $A_\zeta f(z)$. Тоді $\zeta$ є центром мас коренів многочлена $f(z)$ відносно точки $w_i$. Якщо $\zeta \neq \infty$, то рівність $A_\zeta f(w_i) = 0$ можна записати у наступному вигляді

$$(\zeta - w_i)f'(w_i) + nf(w_i) = 0,$$

тоді

$$\zeta = w_i - \frac{nf(w_i)}{f'(w_i)}.$$

При $\zeta = \infty$, маємо $f'(w_i) = A_\zeta f(w_i) = 0$, тому

$$\sum_{j=1}^{n} \frac{1}{z_i - w_i} = \frac{f'(w_i)}{f(w_i)} = 0.$$

Тоді центр мас точок $z_1, \ldots, z_n$ відносно точки $w_i$ лежить у точці

$$w_i + \left( \sum \frac{1}{z_j - w_i} \right)^{-1} = \infty.$$

З таких міркувань бачимо, що точка $w_i$ не може лежати зовні $K$. Дійсно, якщо б точка $w_i$ лежала зовні $K$, то центр мас точок $z_1, \ldots, z_n$ відносно точки $w_i$ лежав би у середині $K$, але це суперечить тому, що точка $\zeta$ лежить зовні $K$. □

Нехай

$$f(z) = \sum_{i=1}^{n} \binom{n}{i} a_i z^i \qquad \text{та} \qquad g(z) = \sum_{i=1}^{n} \binom{n}{i} b_i z^i.$$

Тоді многочлен

$$h(z) = \sum_{i=1}^{n} \binom{n}{i} a_i b_i z^i$$

називається композицією многочленів $f$ та $g$.

**Теорема 9.17** (Сегьо). *Для многочленів $f$ та $g$ степеня $n$. Нехай всі корені многочлена $f$ лежать у круговій області $K$. Тоді довільний корінь композиції $h$ многочленів $f$ та $g$ має вигляд $-\zeta_i k$, де $\zeta_i$ — деякий корінь многочлена $g$ та $k \in K$.*

*Доведення.* Нехай $\gamma$ є коренем многочлена $h$, тоді

$$\sum_{i=1}^{n} \binom{n}{i} a_i b_i \gamma^i = 0.$$

Тоді многочлени $f(z)$ та $G(z) = z^n g(-\gamma z^{-1})$ є аполярними. Тому за теоремою Грейса один з коренів $G(z)$ лежить у $K$. Нехай $g(-\gamma k^{-1}) = 0$, де $k \in K$. Тоді $-\gamma k^{-1} = \zeta_i$, де $\zeta_i$ є коренем многочлена $g$. □

Аналог теореми Грейса для многочленів з можливо різними степенями дає наступна теорема.

**Теорема 9.18.** *Нехай*

$$f(z)=\sum_{i=1}^{n}\binom{n}{i}a_i z^i \qquad та \qquad g(z)=\sum_{i=1}^{m}\binom{m}{i}b_i z^i.$$

*многочлени степенів $n$ та $m$ відповідно, також $m\le n$. Припустимо, що їх коефіцієнти пов'язані наступним співвідношенням*

$$\binom{m}{0}a_0 b_m-\binom{m}{1}a_1 b_{m-1}+\ldots+(-1)^m\binom{m}{m}a_m b_0=0.$$

*Тоді справедливі наступні твердження:*

1. *якщо корені многочлена $g(z)$ лежать у крузі $|z|\le r$, то хоча б один корінь многочлена $f(z)$ лежить у цьому крузі;*
2. *якщо всі корені многочлена $f(z)$ лежать зовні круга $|z|\le r$, то хоча б один корінь многочлена $g(z)$ лежить зовні цього круга.*

## 9.6 Розклад многочленів з дійсними коефіцієнтами

**Лема 9.4.** *Нехай комплексне число $c$ є коренем многочлена $f(x)$ з дійсними коефіцієнтами. Тоді число $\bar{c}$ також коренем цього многочлена.*

*Доведення.* Нехай $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n\in\mathbb{R}[x]$, $c\in\mathbb{C}$ і $f(c)=0$. Використовуючи властивості операції переходу до спряженого комплексного числа, маємо рівності

$$\begin{aligned}f(\bar{c})&=a_0\bar{c}^n+a_1\bar{c}^{n-1}+\ldots+a_{n-1}\bar{c}+a_n=\\&=a_0\overline{c^n}+a_1\overline{c^{n-1}}+\ldots+a_{n-1}\bar{c}+a_n=\bar{a}_0\overline{c^n}+\bar{a}_1\overline{c^{n-1}}+\ldots+\bar{a}_{n-1}\bar{c}+\bar{a}_n=\\&=\overline{a_0c^n+a_1c^{n-1}+\ldots+a_{n-1}c+a_n}=\overline{f(c)}=\bar{0}=0,\end{aligned}$$

тобто число $\bar{c}$ також є коренем многочлена $f(x)$. □

**Теорема 9.19.** *Кожен многочлен додатного степеня з дійсними коефіцієнтами розкладається в добуток многочленів з дійсними коефіцієнтами першого степеня та другого степеня з від'ємним дискримінантом.*

*Доведення.* Проведемо індукцію за степенем многочлена.

Базою індукції є випадок степеня 1, для якого твердження теореми виконується.

Нехай твердження теореми доведено для усіх многочленів, степінь яких менший за $n > 1$. Розглянемо многочлен $f(x)$ степеня $n$ з дійсними коефіцієнтами. Як многочлен з комплексними коефіцієнтами за основною теоремою алгебри він має комплексний корінь $c$.

Якщо $c \in \mathbb{R}$, то за наслідком з теореми Безу $f(x) = (x - c)f_1(x)$ для деякого многочлена $f_1(x)$ з дійсними коефіцієнтами степеня $n-1$. В цьому випадку залишається до многочлена $f_1(x)$ застосувати припущення індукції.

Якщо ж число $c$ не є дійсним, то за лемою спряжене з ним число $\bar{c} \neq c$ також є коренем многочлена $f(x)$. Тоді $f(x)$, як многочлен над полем комплексних чисел, ділиться як на $x - c$, так і на $x - \bar{c}$, а отже і на їх добуток

$$g(x) = (x - c)(x - \bar{c}) = x^2 - (c + \bar{c})x + c\bar{c} = x^2 - 2\operatorname{Re} c\, x + |c|^2,$$

який є многочленом з дійсними коефіцієнтами. Тобто $f(x) = g(x)f_1(x)$ для деякого $f_1(x) \in \mathbb{C}[x]$. Поділивши $f(x)$ на $g(x)$ з остачею, як многочлени над полем дійсних чисел, і скориставшись єдиністю частки і остачі, отримуємо, що всі коефіцієнти многочлена $f_1(x)$ дійсні. Якщо його степінь додатний, то застосуємо до нього припущення індукції. Інакше многочлени $f(x)$ і $g(x)$ асоційовані.

Зауважимо також, що для дискримінанта многочлена $g(x)$ маємо:

$$D = (-2\operatorname{Re} c)^2 - 4|c|^2 = 4(\operatorname{Re} c)^2 - 4((\operatorname{Re} c)^2 + (\operatorname{Im} c)^2) = -4(\operatorname{Im} c)^2 < 0,$$

бо число $c$ не є дійсним. Тому многочлен $g(x)$ незвідний.

Теорему доведено. □

**Наслідок 9.4.** *Кожен многочлен, незвідний над полем дійсних чисел, або має степінь 1, або має степінь 2 і від'ємний дискримінант.*

## 9.7 Теорема Гаусса-Люка

**Теорема 9.20.** *Корені похідної многочлена $f$ належать опуклій оболонці коренів самого многочлена $f$.*

*Доведення.* Нехай $f(z) = (z - z_1) \cdot \ldots \cdot (z - z_n)$. Легко бачити, що

$$\frac{f'(z)}{f(z)} = \frac{1}{z - z_1} + \ldots + \frac{1}{z - z_n}.$$

Нехай $f'(w) = 0$, $f(w) \neq 0$ та $w$ не належить опуклій оболонці точек $z_1, \ldots, z_n$. Тоді через точку $w$ можна провести пряму, яка не перетинає опуклу оболонку точек $z_1, \ldots, z_n$. Тому вектори $w - z_1, \ldots, w - z_n$ лежать в одній півплощині, що задана данною прямою. Звідси маємо, що вектори

$$\frac{1}{w - z_1}, \quad \ldots, \quad \frac{1}{w - z_n}$$

також лежать в одній півплощині, оскільки $\frac{1}{z} = \frac{\overline{z}}{|z|^2}$. Тоді

$$\frac{f'(w)}{f(w)} = \frac{1}{w - z_1} + \ldots + \frac{1}{w - z_n} \neq 0.$$

Отримали суперечність. Тому $w$ належить опуклій оболонці коренів многочлена $f$. □

**Теорема 9.21.** *Нехай $f(z) = (z - x_1) \cdot \ldots \cdot (z - x_n)$, де $x_1 < \ldots < x_n$. Тоді, якщо деякий корінь $x_i$ замінити на $x_i' \in (x_i, x_{i+1})$, то усі корені похідної многочлена $f$ збільшаться.*

*Доведення.* Нехай $z_1 < \ldots < z_{n-1}$ — корені похідної многочлена з коренями $x_1, \ldots, x_n$, а $z_1' < \ldots < z_{n-1}'$ — корені похідної многочлена з коренями $x_1, \ldots, x_i', \ldots, x_n$. Для коренів $z_k$ та $z_k'$ співвідношення

$$\frac{f'(z)}{f(z)} = \frac{1}{z - z_1} + \ldots + \frac{1}{z - z_n}$$

з доведення попередньої теореми набуває вигляду

$$\sum_{i=1}^{n} \frac{1}{z_k - x_i} = 0, \qquad \sum_{i=1}^{n} \frac{1}{z_k' - x_i'} = 0.$$

Припустимо, що твердження теореми неправильне, тобто існує деяке $k$, для якого $z_k' < z_k$. Тоді $z_k' - x_i < z_k - x_i$, зауважимо, що числа $z_k' - x_i$ та $z_k - x_i$ одного знаку. Дійсно $z_j < x_i$, $z_j' < x_i'$ при $j \leq i - 1$ та $z_j > x_i$, $z_j' > x_i'$ при $j \geq i$. Звідси

$$\frac{1}{z_k - x_i} < \frac{1}{z_k' - x_i'}$$

для всіх $i = 1, \ldots, n$. Але тоді рівності

$$\sum_{i=1}^{n} \frac{1}{z_k - x_i} = 0 \qquad \text{та} \qquad \sum_{i=1}^{n} \frac{1}{z_k' - x_i'} = 0$$

не можуть справджуватися одночасно. □

## 9.8 Корені похідної та фокуси еліпса

Для коренів похідної кубічного многочлена справедлива наступна геометрична інтерпретація.

**Теорема 9.22** (ван дер Берг)**.** *Нехай корені кубічного многочлена розташовані у вершинах трикутника $ABC$ на комплексній площині. Тоді корені похідної цього многочлена розташовані у фокусах еліпса, що дотикається сторін трикутника $ABC$ у їх серединах.*

*Доведення.* Зауважимо, що при $g(z) = f(z - z_0)$ виконується $g'(z) = f'(z - z_0)$. Тому як початок координат можна обрати довільну точку.

Довільне афінне перетворення площини можна подати у вигляді композиції руху, гомотетії та перетворення, яке має у деякій прямокутній системі координат вигляд $(x, y) \mapsto (x, y\cos\alpha)$. Тоді можемо вважати, що три кутник $ABC$ отримано перетворенням

$$\psi : z \mapsto \frac{z+\overline{z}}{2} + \frac{z-\overline{z}}{2}\cos\alpha = z\cos^2\frac{\alpha}{2} + \overline{z}\sin^2\frac{\alpha}{2}$$

з правильного трикутника з вершинами $w$, $\varepsilon w$ та $\varepsilon^2 w$, де $|w| = 1$ і $\varepsilon = \exp(2\pi i/3)$. Тоді напівосі $a$ і $b$ еліпса дорівнюють $\frac{1}{2}$ і $\frac{\cos\alpha}{2}$. Відстань між його фокусами $F_1$ і $F_2$ дорівнює $\sqrt{a^2 - b^2} = \frac{1}{2}\sin\alpha$. Точки $F_1$ і $F_2$ переходять у точки $(\pm 1, 0)$ при розтягненні з коефіцієнтом

$$\left(\frac{1}{2}\sin\alpha\right)^{-1} = \left(\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\right)^{-1}.$$

Застосовуючи перетворення $\psi$ і дане розтягнення отримуємо перетворення

$$z \mapsto z\operatorname{ctg}\frac{\alpha}{2} + \overline{z}\operatorname{tg}\frac{\alpha}{2}.$$

Покладемо $a = w\operatorname{ctg}\frac{\alpha}{2}$. Тоді многочлен з коренями $A$, $B$ та $C$ має вигляд

$$f(x) = \left(x - a - \frac{1}{a}\right)\left(x - a\varepsilon - \frac{1}{a\varepsilon}\right)\left(x - a\varepsilon^2 - \frac{1}{a\varepsilon^2}\right).$$

Тоді $f'(x) = 3x^2 + 3\varepsilon + 3\overline{\varepsilon} = 3x^2 - 3$, тому корені многочлена $f'$ дорівнюють $\pm 1$. □

## 9.9 Локалізація коренів похідної

Нехай $f$ є многочленом з дійсними коефіцієнтами. Для кожної пари спряжених коренів $z$ та $\overline{z}$ даного многочлена розглянемо відрізок з кінцями $z$ та $\overline{z}$ і побудуємо на даному відрізку як на діаметрі коло, такі кола називають колами Єнсена многочлена $f$.

**Теорема 9.23** (Єнсен)**.** *Довільний не дійсний корінь похідної многочлена $f$ лежить всередині чи на границі одного з кіл Єнсена.*

*Доведення.* Нехай $z_1, \dots, z_n$ є коренями многочлена $f$. Тоді

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^{n} \frac{1}{z - z_j}.$$

Покажемо, що якщо точка $z$ лежить зовні кола Єнсена з діаметром $z_p z_q$, то

$$\operatorname{sign} \operatorname{Im}\Big(\frac{1}{z - z_p} + \frac{1}{z - z_q}\Big) = -\operatorname{sign} \operatorname{Im} z.$$

Дійсно

$$\frac{1}{z - a - bi} + \frac{1}{z - a + bi} = \frac{2(z - a)((\overline{z} - a)^2 + b^2)}{|(z - a)^2 + b^2|^2}$$

та

$$\operatorname{Im}\Big((\overline{z} - a)|z - a|^2 + (z - a)b^2\Big) = \Big(b^2 - |z - a|^2\Big) \operatorname{Im} z.$$

Покажемо, що якщо $z \notin \mathbb{R}$ і $z_i = a \in \mathbb{R}$, то

$$\operatorname{sign} \operatorname{Im}\Big(\frac{1}{z - z_j}\Big) = -\operatorname{sign} \operatorname{Im} z.$$

Дійсно, виконуються рівності

$$\frac{1}{z - a} - \frac{1}{\overline{z} - a} = \frac{\overline{z} - z}{|z - a|^2} = \frac{-2 \operatorname{Im} z}{|z - a|^2}.$$

Тоді, якщо точка $z \notin \mathbb{R}$ розташована зовні усіх кіл Єнсена, то

$$\operatorname{sign} \operatorname{Im} \frac{f'(z)}{f(z)} = -\operatorname{sign} \operatorname{Im} z \neq 0.$$

Тому $f'(z) \neq 0$, тобто $z$ не корінь похідної. □

Для кількості коренів похідної, дійсна частина яких належить даному відрізку справедлива наступна теорема, сформульована Уолшем, яка дає оцінку, що уточнює теорему Єнсена.

**Теорема 9.24.** *Нехай $I = [\alpha, \beta]$, $K$ — об'єднання $I$ та кіл (кругів) Єнсена, які перетинаються з $I$. Тоді, якщо $K$ містить $k$ коренів многочлена $f(z)$, то кількість коренів многочлена $f'(z)$, які лежать у $K$, міститься у проміжку між $k-1$ та $k+1$.*

*Доведення.* Нехай $C$ межа найменшого прямокутника, сторони якого паралельні осям координат і який містить $K$. Розглянемо обмеження на $C$ відображення $z \mapsto e^{i\varphi}$, де $\varphi = \arg(f'(z)/f(z))$. З рівностей отриманих у попередній теоремі випливає, що образ частини $C$, що лежить у верхній півплощині, лежить на напівколі $|z| = 1$, $\operatorname{Im} z \le 0$. Образ частини $C$, яка лежить в нижній частині напівплощини лежить на напівколі $|z| = 1$, $\operatorname{Im} z \ge 0$. Тому кількість обертів образа кривої $C$ навколо початку координат дорівнює 0 або $\pm 1$. Тоді індекси кривої $C$ відносно векторних полів $f(z)$ та $f'(z)$ співпадають або відрізняються на $\pm 1$. Тобто кількості нулів функцій $f$ та $f'$, які лежать всередині кривої $C$ співпадають або відрізняються на $\pm 1$. □

**Теорема 9.25** (Уолш)**.** *Нехай корені многочленів $f_1$ і $f_2$ лежать у колах $K_1$ і $K_2$, радіуси яких дорівнюють $r_1$ і $r_2$, а центри знаходяться у $c_1$ і $c_2$. Тоді усі корені похідної многочлена $f = f_1 f_2$ лежать або у $K_1$, або у $K_2$, або у колі радіуса $\frac{n_2 r_1 + n_1 r_2}{n_1 + n_2}$ з центром у точці $\frac{n_2 c_1 + n_1 c_2}{n_1 + n_2}$, де $n_1 = \deg(f_1)$ та $n_2 = \deg(f_2)$.*

*Доведення.* Нехай $z$ є коренем похідної многочлена $f$, який лежить зовні $K_1$ і $K_2$. Тоді $f_1'(z) f_2(z) + f_1(z) f_2'(z) = 0$, де $f_1(z)$, $f_2(z)$, $f_1'$, $f_2'(z)$ не дорівнюють нулю.

Розглянемо точки $\zeta_1$ і $\zeta_2$, які є центрами мас коренів многочленів $f_1$ і $f_2$ відносно точки $z$. За теоремою 9.12

$$\zeta_1 = z - n_1 \frac{f_1(z)}{f_1'(z)}, \quad \zeta_2 = z - n_2 \frac{f_2(z)}{f_2'(z)}.$$

Тому

$$n_2\zeta_1 + n_1\zeta_2 = (n_1 + n_2)z - n_1 n_2 \Big( \frac{f_1(z)}{f_1'(z)} + \frac{f_2(z)}{f_2'(z)} \Big) = (n_1 + n_2)z,$$

тобто $z = \frac{n_2\zeta_1 + n_1\zeta_2}{n_1+n_2}$. Оскільки всі корені многочлена $f_i$ лежать у $K_i$, то $\zeta_i \in K_i$. Помітимо, що якщо точки $\zeta_1$ і $\zeta_2$ з масами $n_1$ і $n_2$ лежать у колах $K_1$ і $K_2$, то центр їх мас $z$ лежить в колі $K$. □

**Теорема 9.26** (Грейс-Хивуд)**.** *Якщо $z_1$ і $z_2$ різні корені многочлена $f$ степеня $n$, то у колі $|z-c| \le r$, де*

$$c = (z_1+z_2)/2 \qquad та \qquad r = (|z_1-z_2|/2)\operatorname{ctg}(\pi/n),$$

*знаходиться хоча б один корінь многочлена $f'$.*

*Доведення.* Нехай

$$f'(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} a_k z^k.$$

Тоді

$$0 = f(z_2) - f(z_1) = \int_{z_1}^{z_2} f'(z)\,dz = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} a_k b_{n-1-k},$$

де коефіцієнти $b_0, \dots, b_{n-1}$ залежать тільки від $z_1$ та $z_2$, а від коефіцієнтів $a_0, \dots, a_{n-1}$ вони не залежать. Таким чином, за даними $z_1$ та $z_2$ можна побудувати многочлен

$$g(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} b_k z^k$$

є аполярним до многочлена $f'(z)$.

Щоб отримати явний вираз для многочлена $g$, покладемо

$$a_k = (-1)^k x^{n-1-k},$$

тобто розглянемо многочлен $h(z) = (x-z)^{n-1}$. У такому випадку

$$g(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} b_{n-1-k} = \\ = \int_{z_1}^{z_2} (x-1)^{n-1}\,dz = \frac{(x-z_1)^n - (x-z_2)^n}{n}.$$

Корені многочлена $g$ мають вигляд $\zeta_k = \frac{z_1+z_2}{2} + i\,\frac{z_1-z_2}{2}\operatorname{ctg}\frac{k\pi}{n}$, де $k = 1, 2, \ldots, n-1$. Всі вони лежать на межі круга $|z-c| \le r$. Тоді за теоремою 9.16 у крузі $|z-c| \le r$ знаходиться хоча б один корінь многочлена $f'$. □

## 9.10 Результант

Розглянемо многочлени

$$f(x) = \sum_{i=0}^{n} a_i x^{n-i} \qquad \text{та} \qquad g(x) = \sum_{i=0}^{m} b_i x^{m-i},$$

де $a_0 \neq 0$ та $b_0 \neq 0$. Над алгебраїчно замкнутим полем многочлени $f$ та $g$ мають спільний дільник тоді і тільки тоді, коли вони мають спільний корінь. Якщо поле не є алгебраїчно замкненим, то спільний дільник може бути многочленом, що не має коренів.

Наявність у многочленів $f$ і $g$ спільного дільника еквівалентно тому, що існують такі многочлени $p$ і $q$, що $fq = gp$, де $\deg p \le n-1$ та $\deg q \le m-1$. Нехай $q = u_0 x^{m-1} + \ldots + u_{m-1}$ та $p = v_0 x^{n-1} + \ldots + v_{n-1}$. Рівність $fq = gp$ можна записати у вигляді системи рівнянь

$$a_0 u_0 = b_0 v_0,$$

$$a_1 u_0 + a_0 u_1 = b_1 v_0 + b_0 v_1,$$

$$a_2 u_0 + a_1 u_1 + a_0 u_2 = b_2 v_0 + b_1 v_1 + b_0 v_2,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Многочлени $f$ та $g$ мають спільний корінь тоді і тільки тоді, коли ця система має ненульовий розв'язок $(u_0, u_1, \ldots, v_0, v_1, \ldots)$. Матриця коефіцієнтів даної системи має вигляд

$$\begin{pmatrix} a_0 & 0 & 0 & \ldots & 0 & \ldots & -b_0 & 0 & 0 \\ a_1 & a_0 & 0 & \ldots & 0 & \ldots & -b_1 & -b_0 & 0 \\ a_2 & a_1 & a_0 & \ldots & 0 & \ldots & -b_2 & -b_1 & -b_0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & a_n & \ldots & 0 & 0 & -b_m \end{pmatrix}.$$

Якщо у такій матриці відкинути знаки при $b_i$ та транспонувати її, то отримаємо матрицю Сильвестра многочленів $f$ і $g$, позначимо її $S(f, g)$.

Визначник $\det S(f,g)$ називають результантом многочленів $f$ і $g$, позначимо його $R(f,g)$.

Наприклад, при $m=3$ та $n=2$ визначник матриці, яка записана вище набуватиме вигляд

$$\begin{vmatrix} a_0 & 0 & 0 & -b_0 & 0 \\ a_1 & a_0 & 0 & -b_1 & -b_0 \\ a_2 & a_1 & a_0 & -b_2 & -b_1 \\ 0 & a_2 & a_1 & -b_3 & -b_2 \\ 0 & 0 & a_2 & 0 & -b_3 \end{vmatrix} = \pm \begin{vmatrix} a_0 & a_1 & a_2 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix} =$$

$$= \pm \det S(f,g) = \pm R(f,g).$$

Зрозуміло, що $R(f,g)$ є однорідним многочленом степеня $m$ за змінними $a_i$ та степеня $n$ за змінними $b_j$. Многочлени $f$ і $g$ мають спільний дільник тоді і тільки тоді, коли визначник системи дорівнює нулю, тобто $R(f,g)=0$.

Результант має багато застосувань. Наприклад, коли задано поліноміальні співвідношення $P(x,z)=0$ і $Q(y,z)=0$, то за допомогою результанта можна отримати поліноміальне співвідношення $R(x,y)=0$. Дійсно, розглянемо многочлени $P(x,z)$ і $Q(y,z)$ як многочлени від $z$, вважаючи $x$ та $y$ постійними. Тоді результант $R(x,y)$ цих многочленів дає необхідне співвідношення $R(x,y)=0$.

Результант також дозволяє зводити розв'язання системи алгебраїчних рівнянь до знаходження коренів многочленів. Нехай $P(x_0,y_0)=0$ та $Q(x_0,y_0)=0$. Розглянемо $P(x,z)$ і $Q(x,y)$ як многочлени від $y$. При $x=x_0$ вони мають спільний корінь $y_0$. Тоді їх результант $R(x)$ дорівнює нулю при $x=x_0$.

**Теорема 9.27.** *Нехай $x_i$ — корені многочлена $f$, а $y_j$ — корені многочлена $g$. Тоді*

$$R(f,g) = a_0^m b_0^n \prod (x_i - y_j) = a_0^m \prod g(x_i) = b_0^n \prod f(y_j).$$

*Доведення.* Оскільки $f(x) = a_0(x-x_0)\cdot\ldots\cdot(x-x_n)$, то

$$a_k = \pm a_0 \sigma_k(x_1,\ldots,x_n),$$

де $\sigma_k$ — елементарна симетрична функція. Аналогічно

$$b_k = \pm b_0 \sigma_k(y_1,\ldots,y_m).$$

Результант є однорідним многочленом степеня $m$ за змінними $a_i$ та степеня $n$ за змінними $b_j$, тому

$$R(f,g) = a_0^m b_0^n P(x_1,\dots,x_n,y_1,\dots,y_m),$$

де $P$ — симетричний многочлен від $x_1,\dots,x_n$ та $y_1,\dots,y_m$, який перетворюється у нуль при $x_i = y_j$. Формула

$$x_i^k = (x_i - y_j)x_i^{k-1} + y_j x_i^{k-1}$$

показує, що

$$P(x_1,\dots,y_m) = (x_i - y_j)Q(x_1,\dots,y_m) + r(x_1,\dots,\widehat{x_i},\dots,y_m).$$

Підставимо у цю рівність $x_i = y_j$, отримуємо, що $r$ є нульовим многочленом.

Аналогічні міркування доводять, що многочлен $P$ ділиться на

$$S = a_0^m b_0^n \prod (x_i - y_j).$$

Оскільки

$$g(x) = b_0 \prod_{j=1}^{m} (x - y_j),$$

то

$$\prod_{i=1}^{n} g(x_i) = b_0^n \prod_{i,j} (x_i - y_j),$$

а тоді

$$S = a_0^m \prod_{i=1}^{n} g(x_i) = a_0^m \prod_{i=1}^{n} (b_0 x_i^m + b_1 x_i^{m-1} + \dots + b_m)$$

є однорідним многочленом степеня $n$ за змінними $b_0,\dots,b_m$. Для змінним $a_0,\dots,a_n$ міркування аналогічні.

Симетричний многочлен

$$a_0^m \prod_{i=1}^{n} (b_0 x_i^m + b_1 x_i^{m-1} + \dots + b_m)$$

є многочленом від $a_0,\dots,a_n,b_0,\dots,b_m$. Тоді $R(f,g) = R(a_0,\dots,b_m) = \lambda S$, де $\lambda$ — деяке число. З іншого боку коефіцієнт при $\prod x_i^m$ многочленів $a_0^m b_0^n P(x_1,\dots,y_m)$ і $S$ дорівнює $a_0^m b_0^n$, тому $\lambda = 1$. $\square$

**Наслідок 9.5.** $R(g,f) = (-1)^{\deg f \cdot \deg g} R(f,g)$.

**Наслідок 9.6.** *Якщо* $f = gq + r$, *то*

$$R(f,g) = b_0^{\deg f - \deg r} R(r,g),$$

*де* $b_0$ — *старший коефіцієнт многочлена* $g$.

*Доведення.* Нехай $y_j$ — корені многочлена $g$. Тоді $f(y_j) = r(y_j)$. Залишається скористатися тим, що

$$R(f,g) = b_0^{\deg f} \prod f(y_j) \qquad \text{та} \qquad R(r,g) = b_0^{\deg r} \prod f(y_j). \qquad \square$$

**Наслідок 9.7.** $R(f,gh) = R(f,g)R(f,h)$.

*Доведення.* Нехай $x_i$ — корені многочлена $f$, а коефіцієнт $a_0$ — його старший коефіцієнт. Тоді

$$\begin{aligned} R(f,gh) &= a_0^{\deg g + \deg h} \prod g(x_i)h(x_i), \\ R(f,g) &= a_0^{\deg g} \prod g(x_i), \\ R(f,h) &= a_0^{\deg h} \prod h(x_i). \end{aligned} \qquad \square$$

**Теорема 9.28.** *Нехай*

$$f(x) = \sum_{i=0}^{n} a_i x^{n-i} \qquad i \qquad g(x) = \sum_{i=0}^{m} b_i x^{m-i}.$$

*Тоді існують многочлени* $\varphi$ *і* $\psi$ *з цілими коефіцієнтами від змінних* $x, a_0, \ldots, a_n, b_0, \ldots, b_m$, *для яких виконується рівність*

$$\varphi(x,a,b)f(x) + \psi(x,a,b)g(x) = R(f,g).$$

*Доведення.* Нехай $c_0, \ldots, c_{n+m-1}$ — стовпці матриці Сильвестра $S(f,g)$ і $y_k = x^{m+n-k-1}$. Тоді $z_0c_0 + \ldots + y_{n+m-1}c_{n+m-1} = c$, де $c$ — стовпець $(x^{m-1}f(x), \ldots, f(x), x^{n-1}g(x), \ldots, g(x))$. Розглянемо цю рівність як систему лінійних рівнянь відносно $y_0, \ldots, y_{n+m-1}$ та скористаємося правилом Крамера щоб знайти $y_{n+m-1}$. В результаті отримуємо

$$y_{n+m-1} \det(c_0, \ldots, c_{n+m-1}) = \det(c_0, \ldots, c_{n+m-2}, c).$$

Зазначимо, що $y_{n+m-1} = 1$, $\det(c_0, \ldots, c_{n+m-1}) = R(f,g)$, а визначник $\det(c_0, \ldots, c_{n+m-2}, c)$ можна подати у потрібному вигляді. $\square$

## 9.11 Дискримінант

Нехай $x_1, \ldots, x_n$ — корені многочлена $f(x) = a_0x^n + \ldots + a_n$, де $a_0 \neq 0$. Величину

$$D(f) = a_0^{2n-2} \prod_{i<j} (x_i - x_j)^2$$

називають дискримінантом многочлена $f$.

**Теорема 9.29.** $R(f, f') = \pm a_0 D(f)$.

*Доведення.* За теоремою 9.27

$$R(f, f') = a_0^{n-1} \prod_i f'(x_i),$$

де $x_i$ корені многочлена $f$. Можна показати, що

$$f'(x_i) = a_0 \prod_{i \neq j} (x_j - x_i).$$

Тоді

$$R(f, f') = a_0^{2n-1} \prod_{i \neq j} (x_j - x_i) = \pm a_0^{2n-1} \prod_{i<j} (x_i - x_j)^2. \qquad \square$$

Очевидно, що вираз для результанта можемо переписати наступним чином пов'язавши його з дискримінантом

$$R(f, f') = -R(f, f') = (-1)^{n(n-1)/2} a_0 D(f).$$

**Наслідок 9.8.** *Дискримінант є многочленом з цілими коефіцієнтами від коефіцієнтів многочлена $f$.*

**Теорема 9.30.** *Нехай $f$, $g$ та $h$ — многочлени зі старшим коефіцієнтом* 1. *Тоді*

$$D(fg) = D(f)D(g)R^2(f, g),$$
$$D(fgh) = D(f)D(g)D(h)R^2(f, g)R^2(g, h)R^2(h, f).$$

*Доведення.* Нехай $x_1, \ldots, x_n$ — корені многочлена $f$, а $y_1, \ldots, y_m$ — корені многочлена $g$. Тоді

$$D(fg) = \prod (x_i - x_j)^2 \prod (y_i - y_j)^2 \prod (x_i - y_j)^2 = D(f)D(g)R^2(f, g).$$

Друга формула доводиться аналогічно. $\square$

**Теорема 9.31.** *Нехай $f$ — многочлен з дійсними коефіцієнтами степеня $n$, який не має дійсних коренів. Тоді* $\operatorname{sign} D(f) = (-1)^{n/2}$.

*Доведення.* Використавши розклад

$$f(x) = a_0(x - x_1) \cdot \ldots \cdot (x - x_n),$$

випливає, що

$$D((x - a)f(x)) = D(f(x))(f(a))^2.$$

Нехай $a$ і $\overline{a}$ — пара спряжених коренів многочлена $f$, тобто $f(x) = (x - a)(x - \overline{a})g(x)$. Тоді

$$D(f(x)) = D(g(x))(a - \overline{a})^2(f(a)f(\overline{a}))^2.$$

Очевидно, що $\operatorname{sign}(a - \overline{a})^2 = -1$ і $(f(a)f(\overline{a}))^2 = |f(a)|^4 > 0$. Тому $\operatorname{sign} D(f) = -\operatorname{sign} D(g)$. Далі твердження теореми випливає з міркувань індукції по $n$. □

**Теорема 9.32.** *Нехай $f(x) = x^n + a_1x^{n-1} + \ldots + a_n$ — многочлен з цілими коефіцієнтами. Тоді його дискримінант $D(f)$ має вигляд $4k$ або $4k + 1$, де $k$ — ціле число.*

*Доведення.* Нехай $x_1, \ldots, x_n$ — корені многочлена $f$. Тоді $D(f) = \delta^2(f)$, де

$$\delta(f) = \prod_{i<j}(x_i - x_j).$$

Розглянемо додатково

$$\delta_1(f) = \prod_{i<j}(x_i + x_j).$$

Оскільки $\delta_1(f)$ є симетричною функцією від коренів многочлена $f$, то $\delta_1(f)$ є цілим числом. Також

$$\delta_1^2(f) - \delta^2(f) = \prod_{i<j}((x_i - x_j)^2 + 4x_ix_j) - \prod_{i<j}(x_i - x_j)^2 = 4U(x_1, \ldots, x_n),$$

де $U$ — симетричний многочлен від $x_1, \ldots, x_n$ з цілими коефіцієнтами. Тому $D(f) = \delta_1^2(f) + 4k_1$, де $k_1$ — ціле число. Також зрозуміло, що $\delta_1^2(f) = 4k_2$ або $\delta_1^2(f) = 4k_2 + 1$. □

**Приклад 9.5.** Довести, що $D(x^n + a) = (-1)^{n(n-1)/2} n^n a^{n-1}$.

*Доведення.* Справедливою є формула

$$D(f) = (-1)^{n(n-1)/2} R(f, f') = (-1)^{n(n-1)/2} \prod_{i=1}^{n} f'(x_i),$$

де $x_1, \ldots, x_n$ — корені многочлена $f$. За умовою маємо $f(x) = x^n + a$, тоді $f' = nx^{n-1}$ та $\prod x_i = (-1)^n a$. Тоді $\prod x_i^{n-1} = (-1)^{n(n-1)} a^{n-1} = a^{n-1}$. □

**Приклад 9.6.** Довести, що $D(x^{n-1} + x^{n-2} + \ldots + 1) = (-1)^{(n-1)(n-2)/2} n^{n-2}$.

*Доведення.* Для даного в умові многочлена $f(x) = x^{n-1} + x^{n-2} + \ldots + 1$ виконується співвідношення $(x-1)f(x) = x^n - 1$. Тому

$$D(f)(f(1))^2 = D((x-1)f(x)) = D(x^n - 1) = (-1)^{(n-1)(n-2)/2} n^n.$$

Оскільки $f(1) = n$, то підставивши це значення у попередню формулу отримуємо те, що й треба було довести. □

## 9.12 Вправи

1. Довести, що у многочлена $f(x) = 1 - z + cz^n$, де $c \in \mathbb{C}$, є корінь в колі $|z - 1| \leq 1$.

2. Довести, що многочлен

   $$1 - z + c_1 z^{n_1} + \ldots + c_k z^{n_k},$$

   де $1 < n_1 < n_2 < \ldots < n_k$, має принаймні один корінь у колі

   $$|z| \leq \left(\left(1 - \frac{1}{n_1}\right) \cdot \ldots \cdot \left(1 - \frac{1}{n_k}\right)\right)^{-1}.$$

3. Припустимо, що корені кубічного многочлена розташовані у вершинах трикутника $ABC$ на комплексній площині. Довести, що корені похідної такого многочлена розташовані у фокусах еліпса, який дотикається сторін трикутника $ABC$ у його серединах.

4. Нехай $f_n(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}$. Довести, що

   $$D(n! f_n) = (-1)^{n(n-1)/2} (n!)^n.$$

5. Нехай $d = \mathsf{НСД}(r, s)$, $r_1 = r/d$ та $s_1 = s/d$. Доведіть, що
$$R(x^r - a,\, x^s - b) = (-1)^s (a^{s-1} - b^{r_1})^d.$$

6. Нехай $n > k > 0$, $d = \mathsf{НСД}(n, k)$, $n_1 = n/d$ та $k_1 = k/d$.
$$D(x^n + ax^k + b) = (-1)^{n(n-1)/2} b^{k-1} \times \\ \times (n^{n_1} b^{n_1 - k_1} + (-1)^{n_1+1} (n-k)^{n_1-k_1} k^{k_1} a^{n_1})^d.$$

7. Для яких дійсних значень параметра $a$ всі корені рівняння $z^3 - z^2 + a = 0$ задовольняють нерівність $|z| \le 1$?

8. Довести, що многочлен $f(x)$ ділиться на $f'(x)$ тоді і тільки тоді, коли $f(x) = a_0(x - x_0)^n$.

9. Нехай $n_1 < n_2 < \ldots < n_k$ — натуральні числа. Доведіть, що многочлен $1 + z^{n_1} + z^{n_2} + \ldots + z^{n_k}$ не має коренів у середині кола $|z| < \frac{1}{2}(\sqrt{5} - 1)$.

10. Розглянемо многочлен $p(t)$, який має $n$ різних дійсних нулів більших одиниці. Доведіть, що
$$q(t) = (t^2 + 1)p(t)p'(t) + t((p(t))^2 + (p'(t))^2)$$
має хоча б $2n - 1$ різних дійсних нулів.

11. Нехай $p(x)$ многочлен з дійсними коефіцієнтами, який має тільки дійсні корені. Доведіть, що
$$(n-1)(p'(x))^2 - np(x)p''(x) \ge 0.$$

12. Доведіть, що многочлен $a_1 x^{n_1} + a_2 x^{n_2} + \ldots + a_m x^{n_m}$ не має ненульових коренів кратності більшої ніж $m - 1$.

13. Знайти кількість дійсних коренів многочленів:

    а) $1 + x + \frac{x^2}{2} + \ldots + \frac{x^n}{n}$;

    б) $nx^n - x^{n-1} - \ldots - 1$.

14. Нехай $0 = m_0 < m_1 < \ldots < m_n$ та $m_i \equiv i \pmod 2$. Довести, що многочлен
$$a_0 + a_1 x^{n_1} + a_2 x^{n_2} + \ldots + a_m x^{n_m}$$
має не більше $m$ дійсних коренів.

15. Нехай числа $a_1, \ldots, a_n$ — попарно різні, а числа $b_1, \ldots, b_n$ — додатні числа. Доведіть, що тоді усі корені рівняння

$$\sum \frac{b_k}{x - a_k} = x - c, \quad c \in \mathbb{R}$$

є дійсними.

16. Скільки існує коренів многочлена $x^n + x^m - 1$ за модулем менших одиниці?

# Розділ 10

# Ознаки незвідності многочленів

Однією з найбільш відомих ознак незвідності многочленів є ознака Айзенштайна, яка вже розглядалася у попередніх розділах. Далі розглянемо інші ознаки незвідності многочленів.

## 10.1 Незвідність у полях лишків

Позначимо $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ поле лишків за модулем $p$, де $p$ є простим числом. Многочлени з цілими коефіцієнтами можна розглядати, як многочлени з коефіцієнтами з поля $\mathbb{Z}_p$. У такому випадку многочлен, який над $\mathbb{Z}$ є незвідним може виявитися звідним над полем $\mathbb{Z}_p$ для всіх $p$.

**Теорема 10.1.** *Многочлен $P(x) = x^4 + ax^2 + b^2$, де $a$, $b \in \mathbb{Z}$ є звідним над полем $\mathbb{Z}_p$ для всіх простих $p$.*

*Доведення.* При $p = 2$ існує чотири многочлени виду $P(x)$ з коефіцієнтами 0 та 1. Тобто це наступні многочлени $x^4$, $x^4 + x^2 = x^2(x^2+1)$, $x^4 + 1 = (x+1)^4$, $x^4 + x^2 + 1 = (x^2 + x + 1)^2$, які очевидно є звідними.

Нехай $p$ — непарне просте число. Тоді існує ціле число $s$ таке, що $a \equiv 2s \pmod p$. Тоді $P(x) = x^4 + ax^2 + b^2 \equiv (x^2+s)^2 - (s^2 - b^2) \equiv (x^2+b)^2 - (2b-2s)x^2 \equiv (x^2-b)^2 - (-2b-2s)x^2 \pmod p$. Тому достатньо довести, що одне з чисел $s^2 - b^2$, $2b-2s$, $-2b-2s$ є квадратичним лишком за модулем $p$.

З властивостей квадратичних лишків відомо, що добуток двох чисел, які не є квадратами за модулем $p$, є квадратом за модулем $p$. Якщо $2b-2s$

та $-2b-2s$ не є квадратами за модулем $p$, то їх добуток $4(s^2-b^2)$ буде квадратом за модулем $p$. Тоді $s^2-b^2$ теж є квадратом за модулем $p$. $\square$

## 10.2 Ознака Дюма

Нехай $p$ — фіксоване просте число,

$$f(x)=\sum_{i=1}^{n}A_ix^i$$

— многочлен з цілими коефіцієнтами та $A_0A_n\neq 0$. Запишемо ненульові коефіцієнти многочлена $f$ у вигляді $A_i=a_ip^{\alpha_i}$, де $a_i$ — ціле число, яке не ділиться на $p$. Кожному ненульовому коефіцієнту $a_ip^{\alpha_i}$ поставимо у відповідність точку на площині з координатами $(i,\,\alpha_i)$. За цими точками можна побудувати діаграму Ньютона многочлена $f$, що відповідає простому числу $p$.

Побудова відбувається наступним чином. Нехай $P_0=(0,\,\alpha_0)$ та $P_1=(i_1,\,\alpha_{i_1})$, де $i_1$ — найбільше ціле число, для якого нижче прямої $P_0P_1$ немає даних точок. Далі, нехай, $P_2=(i_2,\,\alpha_{i_2})$, де $i_2$ — найбільше ціле число, для якого нижче прямої $P_1P_2$ немає даних точок і т.д. Останній відрізок має вигляд $P_{r-1}P_r$, де $P_r=(n,\,\alpha_n)$. Якщо відрізки ламаної $P_0\dots P_r$ проходять через точки з координатами, які є цілими числами, то всі ці точки також вважаємо вершинами ламаної. Тоді до вершин $P_0,\,\dots,\,P_r$ додамо ще $s\geq 0$ вершин. Отриману в результаті ламану $Q_0\dots Q_{r+s}$, де $Q_0=P_0$ та $Q_{r+s}=P_r$, називають діаграмою Ньютона. Відрізки $P_lP_{l+1}$ та $Q_iQ_{i+1}$ називають відповідно сторонами та ланками діаграми Ньоютона, а вектори $\overrightarrow{Q_iQ_{i+1}}$ називають векторами ланок діаграми Ньютона.

Розглянемо систему векторів ланок діаграми Ньотона, розглядаючи кожен вектор з урахуванням його кратності, тобто стільки, скільки він входить у кількість векторів ланок.

**Теорема 10.2** (Дюм)**.** *Нехай $f=gh$, де $f$, $g$ та $h$ — многочлени з цілими коефіцієнтами. Тоді система векторів ланок для многочлена $f$ є об'єднанням систем векторів ланок для $g$ та $h$. Просте число $p$ для всіх многочленів обрано одне й те саме.*

*Доведення.* Нехай

$$f(x)=\sum_{i=0}^{n}a_ip^{\alpha_i}x^i,\quad g(x)=\sum_{j=0}^{m}b_jp^{\alpha_j}x^j\quad\text{та}\quad h(x)=\sum_{k=0}^{n}a_ip^{\alpha_i}x^i,$$

де числа $a_i$, $b_j$, $c_k$ не діляться на $p$. Розглянемо деяку сторону $P_lP_{l+1}$ діаграми Ньютона многочлена $f$, сторона $P_lP_{l+1}$ може складатися з декількох ланок діаграми Ньютона. Нехай точки $P_l$ та $P_{l+1}$ мають відповідно координати $i_-, \alpha_{i_-}$ та $i_+, \alpha_{i_+}$. Нахил сторони $P_lP_{l+1}$ дорівнює

$$M = \frac{\alpha_{i_+} - \alpha_{i_-}}{i_+ - i_-}.$$

Нехай $\alpha_{i_+} - \alpha_{i_-} = At$ та $i_+ - i_- = It$, де $t \geq 0$ дорівнює найбільшому спільному дільнику чисел $\alpha_{i_+} - \alpha_{i_-}$ та $i_+ - i_-$. Тоді $M = A/I$ та $\mathsf{НСД}(A, I) = 1$.

Дана сторона $P_lP_{l+1}$ діаграми Ньютона лежить на прямій $I\alpha - Ai = F$, де $F = I\alpha_{i_+} - Ai_+ = A\alpha_{i_-} - Ai_-$. За умовою усі точки $(i, \alpha_i)$, $i = 0, 1, \ldots, n$, лежать не нижче цієї прямої, тобто $I\alpha_i - Ai \geq F$. Остання нерівність є строгою при $i < i_-$ та при $i > i_+$ Назвемо число $I\alpha_i - Ai$ вагою монома $ap^\alpha x^i$, де $\mathsf{НСД}(a, p) = 1$. Числа $i_-$ та $i_+$ однозначно визначають найменший та найбільший показники степеня $x$ мономів многочлена $f$ з мінімальною вагою.

Для многочлена $g$ розглянемо

$$G = \min_{j=0,\ldots,m} \{I\beta_j - Aj\}$$

та покладемо $i_-$ та $i_+$ як найменший та найбільший індекси для яких $G = I\beta_{j_-} - Aj_- = I\beta_{j_+} - Aj_+$.

Аналогічно для многочлена $h$ розглянемо

$$H = \min_{k=0,\ldots,n-m} \{I\gamma_k - Ak\}$$

та покладемо $k_-$ та $k_+$ як найменший та найбільший індекси для яких $H = I\gamma_{k_-} - Ak_- = I\gamma_{k_+} - Ak_+$.

Тоді

$$a_{j_-+k_-} p^{\alpha_{j_-+k_-}} = \sum_{j+k=j_-+k_-} (b_j p^{\beta_j} x^j)(c_k p^{\gamma_k} x^k).$$

Вага добутку двох членів дорівнює сумі їх ваг, тому вага доданку з $j = j_-$ та $k = k_-$ дорівнює $G + H$. Ваги усіх інших доданків строго більші за $G + H$, оскільки для них $j < j_-$ або $k < k_-$. Якщо, наприклад, $j < j_-$, то вага члена $b_j p^{\beta_j} x^j$ строго більша за $G$, а вага члена $c_k p^{\gamma_k} x^k$ не менша за $H$.

Вага члена $(b_j p^{\beta_j} x^j)(c_k p^{\gamma_k} x^k)$ при $j + k = \text{const}$ монотонно зростає зі зростанням $\beta_j + \gamma_k$, оскільки $I > 0$. У нашому випадку $j + k = j_- + k_-$,

тому сума $\beta_j + \gamma_k$ є мінімальною при $j = j_-$ та $k = k_-$. Тоді вага члена $a_{j_-+k_-}p^{\alpha_{j_-+k_-}}$ дорівнює $G + H$. Також зрозуміло, що при $i < j_- + k_-$ вага члена $a_i p^{\alpha_i} x^i$ строго більша за $G + H$, а при $i \geq j_- + k_-$ вага члена $a_i p^{\alpha_i} x^i$ не менша за $G + H$. Тоді $G + H = F$ та $j_- + k_- = i_-$. Доведення того, шо $j_+ + k_+ = i_+$ є аналогічним. Тобто показано, що $i_+ - i_- = (j_+ - j_-) + (k_+ - k_-)$. Зокрема одне з чисел $j_+ - j_-$ та $k_+ - k_-$ є відмінним від нуля.

Якщо обидва числа $j_+ - j_-$ та $k_+ - k_-$ відмінні від нуля, то відрізок з кінцями $(j_-, \beta_{j_-})$ та $(j_+, \beta_{j_+})$ є стороною діаграми Ньютона многочлена $g$, а відрізок з кінцями $(k_-, \gamma_{k_-})$ та $(k_+, \gamma_{k_+})$ є стороною діаграми Ньютона многочлена $h$. Нахил сторін в обох випадках дорівнює $M = A/I$, оскільки

$$\frac{\beta_{j_+} - \beta_{j_-}}{j_+ - j_-} = \frac{A}{I} = \frac{\gamma_{k_+} - \gamma_{k_-}}{k_+ - k_-}$$

Зі співвідношення $i_+ - i_- = (j_+ - j_-) + (k_+ - k_-)$ бачимо, що сума довжин сторін з нахилом $M$ діаграми Ньютона многочленів $g$ та $h$ дорівнюють довжині сторони діаграми Ньютона многочлена $f$ з таким самим нахилом.

Якщо одне з чисел $j_+ - j_-$ та $k_+ - k_-$ дорівнює нулю, то у діаграми Ньютона одного з многочленів $g$ та $h$ є сторона з нахилом $M$, її довжина дорівнює довжині сторони діаграми Ньютона многочлена $f$, а у діаграми Ньютона другого многочлена сторін з нахилом $M$ немає.

Тобто, вектор сторони з нахилом $M$ діаграми Ньютона многочлена $f$ дорівнює сумі векторів сторін з тим самим нахилом $M$ діаграм Ньютона многочленів $g$ та $h$. Зі співвідношення $i_+ - i_- = (j_+ - j_-) + (k_+ - k_-)$, бачимо, що якщо у однієї з діаграм Ньютона многочленів $g$ та $h$ є сторона з деяким нахилом $M$, то у діаграми Ньютона многочлена $f$ теж буде сторона з таким нахилом. □

**Наслідок 10.1** (Ознака Дюма)**.** *Якщо для деякого простого числа $p$ діаграма Ньютона многочлена $f$ складається рівно з однієї ланки, то многочлен $f$ є незвідним.*

Тобто діаграма Ньютона з ознаки Дюма складається з одного відрізка, у середині якого немає точок з цілими координатами.

Зауважимо, що ознака Айзенштайна випливає з ознаки Дюма.

## 10.3 Многочлени з домінуючими коефіцієнтами

У деяких випадках можна стверджувати, що при достатньо великих коефіцієнтах многочлен буде звідним. Одні з кращих умов такого типу є критерій Перрона.

**Теорема 10.3.** *Нехай $f(x) = x^n + a_1 x^{n-1} + \ldots + a_n$ є многочленом з цілими коефіцієнтами таким, що $a_n \neq 0$.*

1. *Якщо $|a_1| > 1 + |a_2| + \ldots + |a_n|$, то $f$ незвідний.*
2. *Якщо $|a_1| \geq 1 + |a_2| + \ldots + |a_n|$ та $f(\pm 1) \neq 0$, то $f$ незвідний.*

*Доведення.* 1. Доведемо, що всі крім одного корені $f$ лежать у одиничному диску $|z| \leq 1$. Очевидно, що многчлен $g(x) = x^n + a_1 x_{n-1}$ задовольняє дану властивість, тобто всі корені $g$ крім одного лежать у одиничному диску $|z| \leq 1$. Тоді за теоремою 9.8 (Руше) достатньо довести, що для $|z| = 1$ виконується

$$|f(z) - g(z)| < |f(z)| + |g(z)|.$$

Але для $|z| = 1$ ми з одного боку маємо, що

$$|f(z) - g(z)| = |a_2 z^{n-2} + \ldots + a_n| \leq |a_2| + \ldots + |a_n| \leq |a_1| - 1,$$

а з іншого боку

$$|f(z)| + |g(z)| \geq |g(z)| = |z^n + a_1 z^{n-1}| = |z + a_1| \geq |a_1| - 1.$$

Доведемо в інший бік. Нехай многочлен $f$ розкладається у добуток многочленів $f_1$ та $f_2$ додатного степеня з цілими коефіцієнтами. Добуток коренів кожного з многочленів $f_1$ та $f_2$ є не нульовим цілим числом. Тому кожен з цих многочленів має корінь абсолютне значення якого не менше одиниці. Але многочлен $f$ має лише один такий корінь. Прийшли до суперечності.

2. За умови досягнення рівності $|a_1| = 1 + |a_2| + \ldots + |a_n|$, маємо не строгу нерівність $|f(z) - g(z)| \leq |a_1| - 1$. Та за умови $f(\pm 1) \neq 0$ маємо строгу нерівність $|f(z)| + |g(z)| > |a_1| - 1$. Для $|z| = 1$ рівність $|f(z)| + |g(z)| = |a_1| - 1$ може виконуватися тільки якщо $|f(z)| = 0$ та одночасно $|z + a_1| = |a_1| - 1$. Остання рівність виконується тільки при $z \in \mathbb{R}$. Оскільки $|z| = 1$, звідси випливає, що $z = \pm 1$. □

**Теорема 10.4.** *Нехай $a_1 \geq a_2 \geq \cdots \geq a_n$ є натуральними та $n \geq 2$. Тоді многочлен $p(x) = x^n - a_1x^{n-1} - a_2x^{n-2} - \ldots - a_n$ незвідний над $\mathbb{Z}$.*

*Доведення.* Розглянемо многочлен $f(x) = (x-1)p(x)$. Очевидно маємо $f(x) = x^{n+1} - b_1x^n - b_2x^{n-1} - \ldots - b_{n+1}$, де $b_1 = a_1 + 1$, $b_2 = a_1 - a_2$, $\ldots$, $b_n = a_{n-1} - a_n$, $b_{n+1} = a_n$. Числа $b_1, b_2, \ldots, b_n, b_{n+1}$ є натуральними та $b_1 = 1 + b_2 + \ldots + b_n + b_{n+1}$. Тому $f(x)$ задовольняє одну з умов критерія Перрона. Але не задовольняє іншу умову, оскільки $f(1) = 0$. Тому необхідні додаткові міркування.

Нехай $h(z) = b_1z^n - b_2z^{n-1} - \ldots - b_{n+1}$. Покажемо, що для всіх достатньо малих $\varepsilon > 0$ справедлива нерівність

$$|h(z)| > |z^{n+1}| = |f(z) + h(z)|$$

скрізь на колі $|z| = 1 + \varepsilon$. Тому, якщо $|z| = 1 + \varepsilon$, тоді

$$|h(z)| - |z^{n+1}| \geq b_1(1+\varepsilon)^n - b_2(1+\varepsilon)^{n-1} - \ldots - b_{n+1} - (1+\varepsilon)^{n+1} =$$

$$= \varepsilon(nb_1 - (n-1)b_2 - \ldots - 2b_{n-1} - b_n - (n+1)) + \ldots =$$

$$= \varepsilon(b_2 + 2b_3 + \ldots + (n-1)b_n + nb_{n+1} - 1) + \ldots$$

Коефіцієнт $\varepsilon$ є додатнім, тоді для достатньо малого $\varepsilon > 0$ маємо, що $|h(z)| - |z^{n+1}| > 0$. У такому випадку

$$|f(z) + h(z)| = |z^{n+1}| < |h(z)| \leq |f(z)| + |h(z)|.$$

Тоді за теоремою Руше многочлен $f(z)$ має стільки ж коренів у середині диску $|z| \leq 1 + \varepsilon$ як і многочлен $h(z)$. Проте усі курені $h(z)$ лежать строго у середині одиничного диску $|z| \leq 1$. Тоді, при $|z| \geq 1$ справедливі наступні нерівності

$$|h(z)| \geq b_1|z|^n - b_2|z|^{n-1} - \ldots - b_{n+1} \geq |z|^n(b_1 - b_2 - \ldots - b_{n+1}) = |z|^n > 0.$$

Поклавши $\varepsilon \to 0$ бачимо, що в середині та на межі одиничного диску є рівно $n$ коренів многочлена $f(x) = (x-1)p(x)$. Тоді рівно $n-1$ корінь многочлена $p(x)$ лежить в середині одиничного диску і хоча б один корінь лежить зовні. Тобто $p$ є незвідним. □

Ознака аналогічна ознаці Перрона, але з умовою не на коефіцієнт при $x^{n-1}$, а на вільний член, справедлива лише в тому випадку, коли вільний член є простим числом.

**Теорема 10.5.** *Нехай $f(x) = x^n + a_1x^{n-1} + \ldots + a_{n-1}x \pm p$ є многочленом з цілими коефіцієнтами та $p$ є простим числом.*

1. *Якщо $p > 1 + |a_1| + \ldots + |a_{n-1}|$, то многочлен $f$ є незвідним.*
2. *Якщо $p \geq 1+|a_1|+\ldots+|a_{n-1}|$ та серед коренів многочлена $f$ немає коренів з одиниці, то многочлен $f$ є незвідним.*

*Доведення.* Нехай $f(x) = g(x)h(x)$, де $g$ та $h$ є многочленами додатнього степеня з цілими коефіцієнтами. Добуток вільних членів многочленів $g$ та $h$ дорівнює $\pm p$. Оскільки число $p$ просте, то один з цих вільних членів дорівнює $\pm 1$. Тому добуток абсолютних величин коренів одного з многочленів $g$ та $h$ дорівнює 1. У цього многочлена має бути такий корінь $\alpha$, що $|\alpha| \leq 1$. Водночас $\alpha$ має бути також і коренем многочлена $f$. З рівності $f(\alpha) = 0$ випливає, що $p = |\alpha^n + a_1\alpha^{n-1} + \ldots + a_{n-1}\alpha| \leq 1 + |a_1| + \ldots + |a_{n-1}|$.

У першому випадку прийдемо до суперечності. У другому випадку, тобто коли серед коренів многочлена $f$ немає коренів з одиниці, виконується нерівність $|\alpha| < 1$, тому $p < 1 + |a_1| + \ldots + |a_{n-1}|$. Тобто і тут отримано суперечність. □

## 10.4 Ознака Пойя

**Лема 10.1.** *Нехай $g$ — многочлен степеня $k$ з цілими коефіцієнтами, $d_0 < d_1 < \ldots < d_k$ — цілі числа. Тоді $|g(d_i)| \geq k!2^{-k}$ для деякого $i$.*

*Доведення.* Розглянемо многочлен

$$G(x) = (x - d_0) \cdot \ldots \cdot (x - d_k) \sum_{i=0}^{k} \frac{g(d_i)}{x - d_i} \prod_{i \neq j} \frac{1}{d_i - d_j}.$$

Можна показати, що $G(d_i) = g(d_i)$ при $i = 0, \ldots, k$ та $\deg G \leq k$. Тоді маємо рівність $G(x) = g(x)$.

Старший коефіцієнт многочлена $G$ дорівнює

$$\sum_{i=0}^{k} g(d_i) \prod_{i \neq j} \frac{1}{d_i - d_j}.$$

За умовою він є ненульовим цілим числом, тому його абсолютна величина не менша за одиницю. Тоді одне з чисел $|g(d_i)|$ не менше за

$$\Big|\sum_{i=0}^{k}\prod_{i\neq j}\frac{1}{|d_i-d_j|}\Big|^{-1} \geq \Big|\sum_{i=0}^{k}\prod_{i\neq j}\frac{1}{|i-j|}\Big|^{-1} =$$

$$= \Big(\sum_{i=0}^{k}\frac{1}{i!(k-i)!}\Big)^{-1} = k!\Big(\sum_{i=0}^{k}\binom{k}{i}\Big)^{-1} = k!2^{-k}. \quad \square$$

**Теорема 10.6** (Ознака Пойя)**.** *Нехай $f$ є многочленом степеня $n$ з цілими коефіцієнтами та $m = [\frac{n+1}{2}]$. Якщо для попарно різних цілих чисел $a_1, \ldots, a_n$ виконуються нерівності $|f(a_i)| < 2^{-m}m!$ і числа $a_1, \ldots, a_n$ не є коренями многочлена $f$. Тоді многочлен $f$ є незвідним.*

*Доведення.* Нехай $f = gh$, де $g$ та $h$ є многочленами з цілими коефіцієнтами. Можемо вважати, що $\deg h \leq \deg g = k$. Тоді $m \leq k < n$. Зрозуміло, що $g(a_i) \neq 0$ та $g(a_i)$ ділить $f(a_i)$. Тоді $|g(a_i)| \leq |f(a_i)| < 2^{-m}m!$.

З іншого боку, за попередньою лемою для одного з членів $a_i$ виконується нерівність $|g(a_i)| \geq 2^{-k}k!$, якщо застосовуємо лему до $d_i = a_{i-1}$. Оскільки $k \geq m$, то $2^{-k}k! \geq 2^{-m}m!$. Тоді при $m = k + r$ маємо $\frac{m!}{k!} = (k+1)\cdot(k+2)\cdot\ldots\cdot(k+r) \leq 2^r = \frac{2^m}{2^k}$. $\square$

Можна сформулювати властивості для незвідності деяких тричленів специфічного виду $x^n \pm x^m \pm 1$.

**Теорема 10.7.** *Нехай $n \geq 2m$, $d = \mathsf{НСД}(n, m)$, $n_1 = n/d$ та $m_1 = m/d$. Тоді тричлен $g(x) = x^n + \varepsilon x^m + \varepsilon'$, де $\varepsilon = \pm 1$ та $\varepsilon' = \pm 1$, є незвідним за виключенням трьох випадків у яких $n_1 + m_1 \equiv 0 \pmod 3$:*

1. *$n_1$ та $m_1$ є непарними та $\varepsilon = 1$.*
2. *$n_1$ є парними та $\varepsilon' = 1$.*
3. *$m_1$ є парними та $\varepsilon' = \varepsilon$.*

*У цих випадках многочлен $g(x)$ є добутком $x^{2d} + \varepsilon^m \varepsilon'^n x^d + 1$ та незвідного многочлена.*

## 10.5 Вправи

1. Розглянемо многочлен $f \in \mathbb{Z}[x]$ має корені $a_1, \ldots, a_n$ та

$$M = \max_i |a_i|.$$

Нехай $f(x_0)$ є простим числом для деякого цілого числа $x_0$, такого, що $|x_0| > M + 1$. Доведіть, що тоді $f$ є незвідним.

2. Нехай $p$ є простим числом. Доведіть, що многочлен $x^p - x - a$, де $a$ — натуральне число, яке не ділиться на $p$, є незвідним.

3. Нехай для многочлена $f \in \mathbb{Z}[x]$ існує таке ціле число $n$, що:

   1) усі корені многочлена $f$ лежать в площині $\operatorname{Re} z < n - \frac{1}{2}$;
   2) $f(n-1) \neq 0$;
   3) $f(n)$ є простим числом.

   Доведіть, що многочлен $f$ незвідний.

4. Нехай $f(x) = x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + pa_n$ — многочлен з цілими коефіцієнтами та $p$ — просте число. Довести, що якщо

$$p > \sum_{i=0}^{n-1} |a_n|^{n-1-i} |a_i|,$$

   то многочлен $f$ є незвідним.

5. Нехай $f(x) = x^{p-1} + x^{p-2} + \ldots + x + 1$ та $p$ — просте число. Довести, що $f$ є незвідним над полем раціональних чисел.

6. Нехай $a_1, \ldots, a_n$ — попарно різні цілі числа Довести, що наступні многочлени є незвідними:

   1) $(x - a_1) \cdot (x - a_2) \cdot \ldots \cdot (x - a_n) - 1$;
   2) $(x - a_1)^2 \cdot (x - a_2)^2 \cdot \ldots \cdot (x - a_n)^2 + 1$.

7. Довести, що довільний многочлен з цілими коефіцієнтами розкладається у суму двох незвідних многочленів.

8. Нехай $F(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ та $f(x) = F(x, \ldots, x)$. Доведіть, що якщо многочлен $f$ незвідний, то многочлен $F$ теж є незвідним.

9. Нехай $p > 3$ — просте число, $n < 2p$. Доведіть, що многочлен $x^{2p} + px^n - 1$ є незвідним.

10. Нехай $p > 3$ — просте число, $n < 2p$ та $a_1 + \ldots + a_p = 2p$. Доведіть, що незвідним буде многочлен

$$x_1^{a_1} \cdot \ldots \cdot x_p^{a_p} + x_1^n + \ldots + x_p^n - 1.$$

# Бібліоґрафія


[1] Андрійчук В. І., Забавський Б. В. *Алгебра і теорія чисел.* Львів, 2005.

[2] Безущак О., Ганюшкін О. *Математична логіка.* Київ: Видавничо-поліграфічний центр «Київський університет», 2023.

[3] Безущак О. О., Ганюшкін О. Г., Кочубінська Є. А. *Навчальний посібник з лінійної алгебри для студентів механіко-математичного факультету.* Київ: ВПЦ «Київський університет», 2019.

[4] Боднарчук Ю. В., Олійник Б. В. *Основи дискретної математики.* Київ: Вид. дім «Києво-Могилянська академія», 2009.

[5] Бондаренко Є. В., Десятерик О. О. *Практикум з конкретної математики.* Київ, 2025.

[6] Курінний Г. Ч., *Многочлени.* Харків, 2015.

[7] Лукашова Т. Д., Друшляк М. Г. *Алгебра і теорія чисел. Частина 1.* Суми, 2022.

[8] Олійник А. С., Сущанський В. І. *Математична логіка.* Київ: Видавничо-поліграфічний центр «Київський університет», 2013.

[9] Шапочка І. *Курс лекцій з алгебри.* Ужгород: УжНУ «Говерла», 2013.

[10] Barbeau E. J. *Polynomials.* Springer-Verlag, 1989.

[11] Graham R. L., Knuth D. E., Patashnik O. *Concrete Mathematics. A Foundation for Computer Science.* Addison-Wesley, 1989.

[12] Lang S. *Undergraduate algebra.* Springer, 2004.

[13] Lay D. C. *Linear Algebra and Its Applications (4th ed.).* Addison Wesley, 2012.

[14] Poole D. *Linear Algebra: A Modern Introduction (4th ed.).* Cengage – Brooks/Cole, 2014.