Київський національний університет імені Тараса Шевченка
Міністерство освіти і науки України

Маринич Олександр Віталійович

Алгебраїчні структури, криптографія та захист інформації

Конспект лекцій

Київ – 2017

ЗМІСТ

Лекція 1

Означення групи, приклади груп

1.1 Бінарні операції.

Означення 1. Бінарною операцією на множині A називається правило за яким у відповідність кожній упорядкованій парі елементів множини A ставиться у відповідність деякий елемент множини A.

Приклад 2. На множині $\mathbb{N}$ операція $\min(a, b)$, яка ставить у відповідність парі натуральних чисел (a, b) менше з них, є бінарною.

Приклад 3. На множині $\mathbb{N}$ операція $*'$, що визначена рівністю $a *' b = \min(a, b) + 2$, є бінарною.

Приклад 4. На множині S дійснозначних функцій, що визначені на всій прямій $\mathbb{R}$, операція $\circ$, що визначена рівністю $(f \circ g)(x) = g(f(x))$, $x \in \mathbb{R}$, є бінарною.

Означення 5. Бінарна операція $*$ на множині A називається **комутативною**, якщо для довільних $a, b \in A$ маємо $a * b = b * a$. Бінарна операція $*$ на множині A називається **асоціативною**, якщо для довільних $a, b, c \in A$ маємо $(a * b) * c = a * (b * c)$.

ВПРАВА. Чи є бінарні операції в прикладах 2, 3,4 комутативними? асоціативними?

Бінарні операції на скінченних множинах зручно задавати у вигляді таблиць, як показано в такому прикладі.

$*$	a	b	c
a	b	c	b
b	a	c	b
c	c	b	a

Табл. 1.1: Бінарна операція $*$ на множині $\{a, b, c\}$.

Приклад 6. Таблиця 6 визначає бінарну операцію $*$ на множині $\{a, b, c\}$ за таким правилом:

$$(i\text{-ий елемент зліва}) * (j\text{-ий елемент зверху}) \\ = (\text{елемент на перетині } i\text{-го рядка та } j\text{-го стовпчика тіла таблиці}).$$

В попередньому прикладі $a*b = c$ та $b*a = a$, тому бінарна операція $*$ некомутативна. Зрозуміло, що в загальному випадку бінарна операція є комутативною тоді і тільки тоді, коли відповідна таблиця симетрична відносно діагоналі, що проходить з верхнього лівого в нижній правий кут.

1.2 Означення групи. Елементарні властивості.

Поняття бінарної операції узагальнює такі добре відомі з елементарної арифметики дії, як додавання та множення цілих чисел. Навчившись додавати та множити цілі числа, наступним кроком є спроба розв'язування рівнянь на кшталт $a + x = b$ та $ax = b$, де a, b задані числа, а x невідоме. Спробуємо проаналізувати, яким умовам має задовольняти множина A та бінарна операція $*$ на ній, щоб рівняння $a*x = b$, $a, b \in A$ (або $x*a = b$) мало розв'язки. Подивимось на таку послідовність дій при розв'язанні рівняння $5 + x = 2$ в множині цілих чисел:

$5 + x = 2$	задано
$-5 + (5 + x) = -5 + 2$	додаємо -5 до обох частин
$(-5 + 5) + x = -5 + 2$	використовуємо асоціативність
$0 + x = -5 + 2$	обчислюємо $-5 + 5$
$x = -5 + 2$	означення 0
$x = -3$	обчислюємо $-5 + 2$

Підкреслимо, що ці обчислення не показують, що -3 є розв'язком, а стверджують лише, що інших розв'язків існувати не може. Перевірка показує, що -3 дійсно є розв'язком. Аналогічним чином можна проаналізувати розв'язання рівняння $2x = 3$ в множині раціональних чисел:

$2x = 3$	задано
$\frac{1}{2}(2x) = \frac{1}{2} \cdot 3$	множимо на $\frac{1}{2}$ обидві частини
$\left(\frac{1}{2} \cdot 2\right) x = \frac{1}{2} \cdot 3$	використовуємо асоціативність
$1 \cdot x = \frac{1}{2} \cdot 3$	обчислюємо $\frac{1}{2} \cdot 2$
$x = \frac{1}{2} \cdot 3$	означення 1
$x = \frac{3}{2}$	обчислюємо $\frac{1}{2} \cdot 3$

З наведених прикладів випливає, що рівняння $a * x = b$ має розв'язок, якщо виконуються такі умови:

- в множині A повинен існувати елемент e такий, що $e * x = x$ для всіх $x \in A$;
- для кожного $a \in A$ повинен існувати елемент a' такий, що $a' * a = e$;
- операція $*$ є асоціативною.

Аналогічні умови можна виписати для рівняння $x * a = b$. Перелічені умови є визначальними для поняття **групи**.

Означення 7. Групою $(G, *)$ називається множина G та задана на ній бінарна операція $*$ такі, що виконуються умови:

($\mathcal{G}_1$) Бінарна операція $*$ є асоціативною.

($\mathcal{G}_2$) Існує елемент $e \in G$ такий, що $e * x = x * e = x$ для всіх $x \in G$. Такий елемент називається **одиницею**.

($\mathcal{G}_3$) Для кожного $a \in G$ існує $a' \in G$ такий, що $a' * a = a * a' = e$. Елемент a' називається **оберненим** елементом до a.

Твердження 8. *Нехай $(G, *)$ є групою, тоді мають місце* **закони скорочення**: *якщо $a * b = a * c$, то $b = c$; якщо $b * a = c * a$, то $b = c$;*

Доведення. Припустимо, що $a * b = a * c$. Згідно з властивістю $\mathcal{G}_3$ в G знайдеться елемент a', обернений до a. Маємо

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

За асоціативністю

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

За визначенням a'

$$e * b = e * c.$$

За визначенням e в $\mathcal{G}_2$

$$b = c.$$

Друга частина твердження доводиться аналогічно. □

Твердження 9. *В групі $(G, *)$ рівняння $a * x = b$ та $y * a = b$, де a, b є фіксованими елементами G, мають єдині розв'язки.*

Доведення. Маємо

$$a' * (a * b) \overset{\mathcal{G}_1}{=} (a' * a) * b \overset{\mathcal{G}_3}{=} e * b \overset{\mathcal{G}_2}{=} b.$$

Отже, $x = a' * b$ є розв'язком $a * x = b$. Аналогічно, $y = b * a'$ є розв'язком $y * a = b$. Щоб показати єдиність, припустимо, що $y * a = b$ та $y_1 * a = b$. Тоді $y * a = y_1 * a$ та $y = y_1$ за твердженням 8. Єдиність x перевіряється аналогічно. □

Підкреслимо, що розв'язки $x = a' * b$ та $y = b * a'$ можуть бути різними, якщо $*$ не є комутативною.

Означення 10. Група $(G, *)$ називається абелевою (або комутативною), якщо бінарна операція $*$ є комутативною.

Приклад 11. Мають місце такі твердження:

- $(\mathbb{N}, +)$ не є групою. В множині $\mathbb{N}$ не існує одиниці відносно операції $+$.
- $(\mathbb{N}_0, +)$ не є групою. В множині $\mathbb{N}$ не існує оберненого елемента до 2.
- $(\mathbb{Z}, +)$ є абелевою групою.

- $(\mathbb{N}, \times)$ не є групою. В множині $\mathbb{N}$ є одиниця відносно операції $\times$, але немає оберненого елемента до 3.

- $(\mathbb{Q}_+, \times)$ є абелевою групою.

Твердження 12. *В групі $(G, *)$ одиниця єдина. Для кожного $a \in G$ обернений елемент a' єдиний.*

Доведення. Припустимо, що для кожного $x \in G$

$$e * x = x * e = x \quad \text{та} \quad e_1 * x = x * e_1 = x.$$

Підставивши $x = e_1$ в першу рівність та $x = e$ в другу, отримаємо

$$e * e_1 = e_1 * e = e_1 \quad \text{та} \quad e_1 * e = e * e_1 = e.$$

Отже, $e_1 = e$.

Далі, припустимо, що

$$a' * a = a * a' = e \quad \text{та} \quad a'' * a = a * a'' = e.$$

Тоді

$$a * a' = a * a''.$$

Згідно з твердженням 8 $a' = a''$. □

1.3 Скінченні групи. Таблиця групи.

В попередніх прикладах всі групи складались з нескінченної кількості елементів. Спробуємо побудувати приклади груп, що мають скінченну кількість елементів. Оскільки кожна група повинна мати одиницю, то найменша множина, на якій можна задати структуру групи, це $\{e\}$. Єдина можлива бінарна операція на такій множині: $e * e = e$. Очевидно, $(\{e\}, *)$ є групою.

Побудуємо групу з двох елементів, задавши таблицю відповідної бінарної операції. Такі таблиці називаються **таблицями Келі**[1]. Оскільки група повинна містити одиницю, позначимо її елементи $\{e, a\}$. За визначенням e маємо

[1] Артур Келі, 1821 – 1895, англійський математик.

$*$	e	a
e	e	a
a	a	

Елемент a повинен мати обернений a' такий, що

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

В нашому випадку $a' = a$ або $a' = e$. Другий варіант, очевидно, не підходить, тому $a' = a$ та $a * a = e$. Отже, можемо заповнити таблицю в єдиний спосіб:

$*$	e	a
e	e	a
a	a	e

Таким чином, всі властивості групи виконуються, окрім, можливо, асоціативності. Безпосередньою перевіркою можна пересвідчитись, що бінарна операція, що задається наведеною таблицею, є асоціативною. Ми побудували групу на двох елементах та показали, що така група з точністю до перейменування елементів, єдина.

Проаналізуємо, які властивості повинна мати таблиця скінченної групи. Умова існування одиниці e накладає обмеження на рядок та стовпчик, що відповідає e. З умови існування оберненого елемента випливає, що кожен рядок та стовпчик таблиці має містити елемент e. Згідно з твердженням 9 рівняння $a * x = b$ та $y * a = b$ мають єдині розв'язки. Це означає, що кожний елемент $b \in G$ з'являється рівно один раз в кожному рядку і стовпчику.

З цих спостережень можемо побудувати таблицю групи з трьох елементів $\{e, a, b\}$.

$*$	e	a	b
e	e	a	b
a	a		
b	b		

$\Longrightarrow$

$*$	e	a	b
e	e	a	b
a	a	b	e
b	b		

$\Longrightarrow$

$*$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Можна перевірити, що побудована бінарна операція є асоціативною.

Приклад 13 (Групи з чотирьох елементів. 4-група Клейна.). Спробуємо побудувати групу (групи?) з чотирьох елементів $\{e, a, b, c\}$.

$*$	e	a	b	c
e	e	a	b	c
a	a	?		
b	b			
c	c			

На місці знака питання не можна поставити елемент a, тому там має бути одиниця e або один з елементів b, c.

Випадок $? = e$. Маємо дві можливості заповнити таблицю.

$*$	e	a	b	c
e	e	a	b	c
a	a	e		
b	b			
c	c			

$\Longrightarrow$

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c		
c	c	b		

$\Longrightarrow$

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

або

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

Випадок $? = b$ або $? = c$. Не зменшуючи загальності (оскільки елементи b, c можна поміняти місцями), розглянемо випадок $? = b$. Є лише один спосіб заповнити таблицю, щоб отримати групу.

$*$	e	a	b	c
e	e	a	b	c
a	a	b		
b	b			
c	c			

$\Longrightarrow$

$*$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c		
c	c	e		

$\Longrightarrow$

$*$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

Отримана таблиця задає таку саму групу, що й друга таблиця першого випадку. Щоб побачити це, достатньо перейменувати елементи останньої таблиці так: $\{e \mapsto e, a \mapsto b, b \mapsto a, c \mapsto c\}$. Кажуть, що групи, задані цими двома таблицями, **ізоморфні**. Детально ми розглянемо це поняття у наступних лекціях.

Підсумовуючи, ми показали, що існують дві групи порядку 4. Таблиці Келі цих груп:

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

та

$*$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

.

Група, що задається першою таблицею називається 4-групою Клейна[2] і позначається V_4 (від нім. Vierergruppe) або $\mathbb{Z}_2 \times \mathbb{Z}_2$. Група, що задається другою таблицею називається циклічною групою порядку 4 і позначається $\mathbb{Z}_4$. Обидві групи є комутативними.

Введемо ряд домовленостей щодо позначень. Не зважаючи на те, що група за означенням є парою об'єктів $(G, *)$, ми будемо іноді називати групою саму множину G, якщо бінарна операція $*$ зрозуміла з контексту. Замість запису $a * b$ ми будемо писати ab і називати це **добутком** елементів a та b. Якщо операція $*$ є комутативною, то замість $a * b = ab$ також вживається запис $a + b$ – **сума** елементів a та b. Обернений елемент a' позначається a^{-1} або $-a$ (в комутативному випадку). n-кратний добуток $a * a * \cdots * a$ позначається a^n або na (в комутативному випадку). Нарешті, покладемо $(a^n)^{-1} = (a^{-1})^n = a^{-n}$ або $-na$ (в комутативному випадку).

1.4 Підстановки. Групи підстановок.

Означення 14. Підстановкою ϕ на множині A називається довільна бієкція множини A на себе, тобто довільне відображення $\phi : A \mapsto A$, що задовольняє дві умови:

- якщо $a, b \in A$, $a \neq b$, то $\phi(a) \neq \phi(b)$ **(ін'єктивність)**;
- для кожного $a \in A$ існує $b \in A$, що $a = \phi(b)$ **(сюр'єктивність)**.

На множині підстановок існує природня бінарна операція – **суперпозиція (композиція)** $\circ$. Позначимо множину всіх підстановок на множині A через $\mathfrak{S}_A$. Якщо $\sigma, \tau \in \mathfrak{S}_A$, то $\sigma\tau := \tau \circ \sigma$ є підстановкою, що визначається рівністю

$$(\sigma\tau)(x) = \tau(\sigma(x)), \quad x \in A.$$

[2]Фелікс Клейн, 1849 – 1925, німецький математик.

Перевіримо, що $\sigma\tau \in \mathfrak{S}_A$, тобто дійсно є підстановкою. Нехай $x_1 \neq x_2$, тоді $\sigma(x_1) \neq \sigma(x_2)$, а тому $\tau(\sigma(x_1)) \neq \tau(\sigma(x_2))$. Отже, $\sigma\tau$ є ін'єктивним відображенням. Нехай $a \in A$, тоді існує $b \in A$, що $a = \tau(b)$. З іншого боку, існує $c \in A$, що $b = \sigma(c)$. Тому $a = \tau(\sigma(c))$, отже $\sigma\tau$ є сюр'єктивним.

Приклад 15. Нехай $A = \{1, 2, 3, 4, 5\}$. Підстановки скінченних множин зазвичай записуються у дворядковому вигляді. Наприклад,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix}$$

означає, що $\sigma(1) = 4, \sigma(2) = 2, \sigma(3) = 5, \sigma(4) = 3, \sigma(5) = 1$. Нехай

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 2 & 1 \end{pmatrix},$$

тоді

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 4 & 3 \end{pmatrix}.$$

Теорема 16. *Нехай $A \neq \varnothing$, тоді $(\mathfrak{S}_A, \circ)$, де $\circ$ позначає суперпозицію, є групою. Ця група називається* **симетричною групою** *множини A.*

Доведення. Нам потрібно перевірити властивості $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$ в означенні групи. Перевіримо асоціативність: для довільного $a \in A$ та $\sigma, \tau, \mu \in \mathfrak{S}_A$ маємо

$$((\sigma\tau)\mu)(a) = \mu((\sigma\tau)(a)) = \mu(\tau(\sigma(a))) = (\tau\mu)(\sigma(a)) = (\sigma(\tau\mu))(a).$$

Одиницею в $\mathfrak{S}_A$ є тотожнє відображення, тобто відображення $e : A \mapsto A$ для якого $e(a) = a$ для всіх $a \in A$. Оскільки кожне відображення $\sigma \in \mathfrak{S}_A$ є бієктивним, то існує обернене відображення σ^{-1} для якого має місце $\sigma^{-1}(\sigma(a)) = \sigma(\sigma^{-1}(a)) = a$ для всіх $a \in A$. Отже, властивості $\mathcal{G}_2$ та $\mathcal{G}_3$ виконуються. Доведення завершено. □

Зауваження 17. Нижче ми побачимо, що визначення композиції підстановок рівністю $(\sigma\tau)(x) = \sigma(\tau(x))$, для всіх $x \in A$, породжує группу, що ідентична $(\mathfrak{S}_A, \circ)$. Деталі буде наведено в прикладі 70.

Означення 18. Якщо A є скінченною множиною з n елементів, то $\mathfrak{S}_A$ позначається $\mathfrak{S}_n$ і називається симетричною групою на n літерах. Ця група має $n!$ елементів.

	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_1	ρ_1	ρ_2	ρ_0	μ_2	μ_3	μ_1
ρ_2	ρ_2	ρ_0	ρ_1	μ_3	μ_1	μ_2
μ_1	μ_1	μ_3	μ_2	ρ_0	ρ_2	ρ_1
μ_2	μ_2	μ_1	μ_3	ρ_1	ρ_0	ρ_2
μ_3	μ_3	μ_2	μ_1	ρ_2	ρ_1	ρ_0

Табл. 1.2: Таблиця Келі групи $\mathfrak{S}_3$.

Приклад 19 (Симетрична група на трьох літерах). Нехай $A = 1, 2, 3$. Шість підстановок, що утворюють групу $\mathfrak{S}_3$ такі:

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$
$$\mu_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \mu_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

а відповідна таблиця Келі наведена в таблиці 1.2. Підкреслимо, що ця група неабелева і є нашим першим прикладом скінченної неабелевої групи. Нижче ми побачимо, що єдина група на п'яти елементах є абелевою і, таким чином, $\mathfrak{S}_3$ є найменшою неабелевою групою.

Група $\mathfrak{S}_3$ тісно пов'язана з симетріями правильного трикутника. Занумеруємо його вершиним цифрами $\{1, 2, 3\}$. Існує шість перетворень площини, що залишають трикутник на місці: три повороти ρ_0, ρ_1, ρ_2 (на 0, 120 та 240 градусів, відповідно) та три симетрії μ_1, μ_2, μ_3 відносно кожної з висот. Це пояснює вибір позначень для елементів $\mathfrak{S}_3$.

Означення 20. Діедральна група D_n це група симетрій правильного n-кутника.

Приклад 21. Побудуємо діедральну групу D_4 симетрій квадрата. Кожен елемент цієї групи є деякої підстановкою на множині вершин $\{1, 2, 3, 4\}$. Група D_4, яка називається також октичною групою, складається з чотирьох поворотів (на 0, 90, 180 та 270 градусів), двох симетрій відносно серединних перпендикулярів до сторін та двох симетрій відносно діагоналей. Відповідні підстановки мають вигляд

	ρ_0	ρ_1	ρ_2	ρ_2	μ_1	μ_2	δ_1	δ_2
ρ_0	ρ_0	ρ_1	ρ_2	ρ_3	μ_1	μ_2	δ_1	δ_2
ρ_1	ρ_1	ρ_2	ρ_3	ρ_0	δ_2	δ_1	μ_1	μ_2
ρ_2	ρ_2	ρ_3	ρ_0	ρ_1	μ_2	μ_1	δ_2	δ_1
ρ_3	ρ_3	ρ_0	ρ_1	ρ_2	δ_1	δ_2	μ_2	μ_1
μ_1	μ_1	δ_1	μ_2	δ_2	ρ_0	ρ_2	ρ_1	ρ_3
μ_2	μ_2	δ_2	μ_1	δ_1	ρ_2	ρ_0	ρ_3	ρ_1
δ_1	δ_1	μ_2	δ_2	μ_1	ρ_3	ρ_1	ρ_0	ρ_2
δ_2	δ_2	μ_1	δ_1	μ_2	ρ_1	ρ_3	ρ_2	ρ_0

Табл. 1.3: Таблиця Келі групи D_4.

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \quad \rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \quad \rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \quad \rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$$
$$\mu_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad \delta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \quad \delta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}.$$

Відповідна таблиця Келі наведена в таблиці 1.3.

1.5 Цикли підстановок.

Означення 22. Підстановка σ на множині A називається **циклом** довжини n, якщо знайдуться такі $\{a_1, a_2, \ldots, a_n\} \subset A$, що

$$\sigma(a_1) = a_2, \sigma(a_2) = a_3, \ldots, \sigma(a_{n-1}) = a_n, \sigma(a_n) = a_1$$

та $\sigma(x) = x$ для $x \in A \setminus \{a_1, a_2, \ldots, a_n\}$. Такий цикл позначатимемо $(a_1, a_2, \ldots, a_n)$.

Приклад 23. Нехай $(1, 4, 5, 6), (2, 1, 5) \in \mathfrak{S}_6$. Тоді

$$(1,4,5,6)(2,1,5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & 5 & 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 3 & 4 & 2 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 3 & 2 & 6 & 5 \end{pmatrix}.$$

Набір циклів називається **диз'юнктним**, якщо жоден елемент A не входить до двох різних циклів набору. Множення диз'юнктних циклів, очевидно, комутативне.

Теорема 24. *Довільна підстановка σ на скінченній множині A може бути представлена у вигляді добутку диз'юнктних циклів.*

Доведення. Нехай $A = \{1, 2, \ldots, n\}$. Розглянемо послідовність

$$\sigma(1), \sigma^2(1), \sigma^3(1), \ldots.$$

Оскільки A скінченна, то ця послідовність не може містити всі різні елементи. Нехай $\sigma^r(1)$ перший член цієї послідовності, що з'являвся раніше. Тоді $\sigma^r(1) = 1$, оскільки у протилежному випадку ми мали б $\sigma^r(1) = \sigma^s(1)$ для деякого $1 \leq s < r$, а тому $\sigma^{r-s}(1) = 1$, що суперечить вибору r. Нехай $\tau_1 = (1, \sigma(1), \ldots, \sigma^{r-1}(1))$. Зрозуміло, що на всі елементи $1, \sigma(1), \ldots, \sigma^{r-1}(1) \in A$ підстановка τ_1 діє так само, як і σ.

Нехай i є першим з елементів A, що не входить в цикл τ_1, тоді наведений вище аргумент можна повторити для послідовності

$$i, \sigma(i), \sigma^2(i), \ldots$$

і побудувати цикл τ_2. Цикли τ_1, τ_2 диз'юнктні, оскільки якщо в них знайшовся б спільний елемент j, то вони співпали б в силу того, що кожний цикл можна породити, застосовуючи підстановку σ до j.

Аналогічно, будуємо τ_3, τ_4 і т.д. Цей процес скінченний, оскільки множина A скінченна. Якщо τ_m останній з побудованих циклів, то

$$\sigma = \tau_1 \tau_2 \cdots \tau_m.$$

Доведення завершено. □

Означення 25. Цикл довжини 2 називається **транспозицією**.

Оскільки

$$(a_1, a_2, \ldots, a_n) = (a_1, a_2)(a_1, a_3) \cdots (a_1, a_n),$$

то маємо такий наслідок з теореми 24.

Наслідок 26. Кожна підстановка на скінченній множині, що містить принаймні два елементи, може бути представлена у вигляді добутку транспозицій.

Підкреслимо, що в той час як представлення підстановки у вигляді добутку диз'юнктних циклів є єдиним з точністю до порядку співмножників, представлення у вигляді добутку транспозицій не є єдиним. Проте, має місце теорема, яка стверджує, що кількість транспозицій у різних записах завжди однакової парності.

Теорема 27. *Не існує підстановки на скінченній множині, яка може бути представлена у вигляді як парної, так і непарної кількості транспозицій.*

Доведення. Нехай $A = \{1, 2, \ldots, n\}$ та $n \geq 2$. Розглянемо тотожню підстановку e. Оскільки $e = (1, 2)(2, 1)$, то нам потрібно показати, що з представлення

$$e = \tau_1 \tau_2 \cdots \tau_k, \tag{1.1}$$

де всі τ_i є транспозиціями, випливає $k \equiv 0 (\mathrm{mod}\, 2)$. Виберемо деякий елемент m, що входить в одну з транспозицій представлення 1.1. Нехай τ_j є першою (якщо рахувати зліва направо) транспозицією, що містить m. Тоді $j < k$, оскільки в протилежному випадку ми мали б $e(m) \neq m$. Пара транспозицій $\tau_j \tau_{j+1}$ має один з чотирьох виглядів:

$$\begin{aligned}
(m, x)(m, x) &= e \\
(m, x)(m, y) &= (x, y)(m, x) \\
(m, x)(y, z) &= (y, z)(m, x) \\
(m, x)(x, y) &= (x, y)(m, y).
\end{aligned}$$

Підставивши правильний вигляд в формулу 1.1 ми або зменшимо кількість транспозицій на дві, або зсунемо першу появу m вправо на один крок. Повторимо цю процедуру, поки m не зникне з представлення 1.1, а потім повторимо для всіх інших елементів A, що з'являються в 1.1. Оскільки на кожному кроці k кількість транспозицій або не змінюється, або зменшується на два, то k має бути парним.

Нехай тепер σ є довільною підстановкою на множині A та нехай

$$\sigma = \tau_1 \tau_2 \cdots \tau_r = \tau'_1 \cdots \tau'_s$$

є два представлення σ у вигляді добутку транспозицій. Оскільки кожна транспозиція є оберненою до самої себе, то

$$e = \sigma \sigma^{-1} = \tau_1 \tau_2 \cdots \tau_r (\tau'_1 \cdots \tau'_s)^{-1} = \tau_1 \tau_2 \cdots \tau_r \tau'_s \cdots \tau'_1.$$

Число $r + s$ парне, а тому r та s повинні бути однакової парності. Доведення завершено. □

Означення 28. Підстановка на скінченній множині називається **парною (відповідно, непарною)**, якщо вона може бути представлена у вигляді добутку парної (відповідно, непарної) кількості транспозицій.

Приклад 29 (Альтернуюча група A_n). Нехай

$$A_n := \{\sigma \in \mathfrak{S}_n : \sigma \text{ – парна підстановка}\}.$$

Оскільки добуток парних підстановок парний, $e \in A_n$ та підстановка, обернена до парної, сама парна (чому?), то множина парних підстановок разом з операцією добутку підстановок є групою. Вона складається з $n!/2$ елементів і називається **альтернуючою** групою. Ця група є **підгрупою** $\mathfrak{S}_n$.

Лекція 2

Підгрупа, нормальна підгрупа, класи суміжності та фактор-групи

2.1 Означення підгрупи. Діаграми Гассе.

Означення 30. Нехай $(G, *)$ є групою, а $S \subset G$. Якщо для довільних $a, b \in S$ має місце $a * b \in S$, то S **замкнена відносно групової операції** $*$.

Означення 31. Нехай $(G, *)$ є групою, $H \subset G$ і H є замкненою відносно групової операції $*$. Якщо $(H, *)$ є групою, то кажуть, що H є підгрупою G і пишуть $H \leq G$, або $H < G$, якщо H є власною підмножиною G.

Приклад 32. В прикладі 13 ми побудували дві групи порядку 4. З таблиці Келі випливає, що єдиною замкненою підмножиною $\mathbb{Z}_4$ є $\{0, 2\}$. Ця множина є підгрупою $\mathbb{Z}_4$. 4-група Клейна має три власні замкнені підмножини: $\{e, a\}$, $\{e, b\}$ та $\{e, c\}$. Кожна з цих множин є підгрупою V_4, проте вони задають однакову групову структуру на двоелементній множині.

Нехай $H < G$ та $a \in H$. Згідно з твердженням 9 рівняння $ax = a$ повинно мати розв'язок в H. З іншого боку, це рівняння має розв'язок e в G. Це показує, що одиниця G є одиницею H. Аналогічно, розглянувши рівняння $ax = e$, бачимо, що обернений до a елемент a^{-1} в G є оберненим до a і в підгрупі H.

Теорема 33. *Нехай $H \subset G$. H є підгрупою G тоді й тільки тоді, коли:*

1) H є замкненою відносно групової операції G;

2) одиниця e групи G лежить в H;

3) для кожного $a \in H$ має місце $a^{-1} \in H$.

Доведення. Якщо $H < G$, то виконання умов 1,2,3 випливає з означення підгрупи та зауважень перед цією теоремою.

Припустимо тепер, що $H \subset G$ та виконуються умови 1,2,3. З умови 2 випливає виконання $\mathcal{G}_2$, а з умови 3 виконання умови $\mathcal{G}_3$. Оскільки для для довільних $a, b, c \in H$ маємо $(ab)c = a(bc)$ в G, а. отже, й в H, то $\mathcal{G}_1$ виконується. Доведення завершено. □

Відношення «бути підгрупою» на множині всіх підгруп групи є частковим порядком, тому підгрупи зручно зображати у вигляді **діаграми Гассе**. У такій діаграмі стрілка, що йде від деякої підгрупи $H_1 < G$ до деякої іншої підгрупи $H_2 < G$, означає, що $H_1 < H_2$. Вся група G розміщується вгорі діаграми, а тривіальна підгрупа $\{e\}$ – внизу. Чим більша є підгрупа, тим вище вона розміщена в діаграмі. На рисунках 2.1 та 2.2 зображені діаграми Гассе підгруп деяких груп, що нам зустрічались раніше.

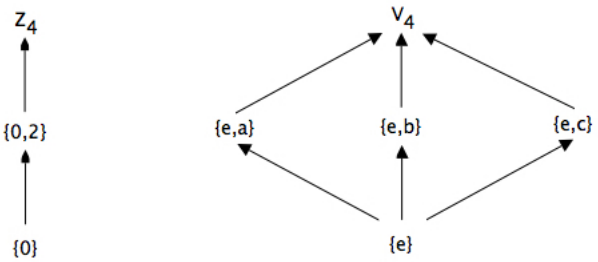

Рис. 2.1: Діаграми Гассе підгруп $\mathbb{Z}_4$ (зліва) та V_4 (справа).

2.2 Циклічні підгрупи.

Розглянемо довільний елемент $a \in G$, де G – група, та спробуємо побудувати найменшу підгрупу, що містить a. Зрозуміло, що a^n має лежати в такій підгрупі для всіх $n \in \mathbb{N}$. Також з теореми 33 випливає, що e та a^{-1} мають лежати в такій

підгрупі. Нарешті, a^{-n} повинно лежати в такій підгрупі для кожного $n \in \mathbb{N}$. Таким чином, підгрупа, що містить a, містить також множину

$$\langle a \rangle \ := \ \{a^n : n \in \mathbb{Z}\}.$$

Для елементів множини $\langle a \rangle$ виконується звичайний закон множення експонент: $a^m a^n = a^{m+n}$ при $m, n \in \mathbb{Z}$. Для $m, n \in \mathbb{N}_0$ або $-m, -n \in \mathbb{N}_0$ це очевидно. Якщо $m < 0, n \geq 0$, то маємо при $n \geq -m$

$$a^m a^n = \underbrace{a^{-1} \cdots a^{-1}}_{-m \text{ раз}} \underbrace{a \cdots a}_{n \text{ раз}} = a^{n-(-m)} = a^{n+m}.$$

Випадок $n < -m$ перевіряється аналогічно.

Теорема 34. *Нехай G є групою та $a \in G$. Тоді $\langle a \rangle$ є мінімальною підгрупою G, що містить a. Тобто, будь-яка інша підгрупа, що містить a, містить також $\langle a \rangle$.*

Доведення. Достатньо перевірити, що $\langle a \rangle$ підгрупа. Перевіримо умови 1,2,3 теореми 33. З рівності $a^r a^s = a^{r+s}$, $r, s \in \mathbb{Z}$, випливає, що $\langle a \rangle$ замкнена відносно групової операції G. Умови 2,3 випливають зі співвідношень $a^0 = e \in \langle a \rangle$ та $(a^r)^{-1} = a^{-r} \in \langle a \rangle$, $r \in \mathbb{Z}$. Мінімальність випливає з обговорення, що передувало теоремі. Доведення завершено. □

Означення 35. Підгрупа $\langle a \rangle$ називається **циклічною підгрупою** G, породженою $a \in G$.

Означення 36. Елемент a називається генератором групи G, якщо $G = \langle a \rangle$. Група G називається **циклічною**, якщо знайдеться $a \in G$, що $\langle a \rangle = G$.

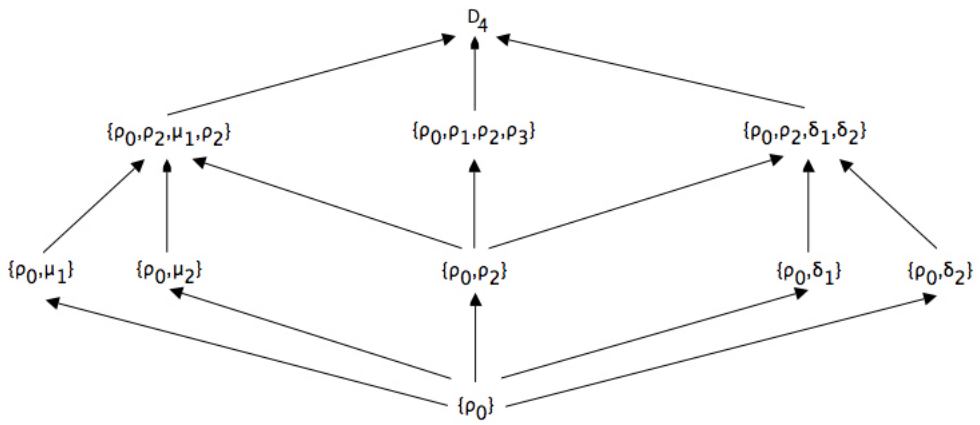

Рис. 2.2: Діаграма Гассе підгруп діедральної групи D_4.

Приклад 37. Мають місце такі твердження:

- Група $(\mathbb{Z}, +)$ є циклічною з генератором 1 (або -1).
- $\mathbb{Z}_4$ є циклічною, оскільки $\langle 1\rangle = \langle 3\rangle = \mathbb{Z}_4$.
- V_4 не є циклічною, оскільки $\langle a\rangle = \langle b\rangle = \langle c\rangle$ мають по два елементи, а $\langle e\rangle$ є тривіальною підгрупою.
- $n\mathbb{Z} := \langle n\rangle$ є циклічною підгрупою групи $(\mathbb{Z}, +)$. Якщо m ділить n, то $n\mathbb{Z} < m\mathbb{Z}$.

Наведемо деякі факти про циклічні групи та підгрупи. Наступне твердження очевидне.

Твердження 38. *Кожна циклічна група є абелевою.*

Теорема 39. *Кожна підгрупа циклічної групи є циклічною.*

Доведення. Нехай $G = \langle a\rangle$ і нехай $H < G$. Якщо $H = \{e\}$, то H є циклічною. Якщо $H \neq \{e\}$, то знайдеться $a^m \in H$. Нехай m є мінімальним натуральним числом з такою властивістю. Покажемо, що $H = \langle a^m\rangle$. Нехай $a^n \in H$ та $n = mp + q$ для деякого $p \in \mathbb{Z}$ та $0 \leq q < m$. Тоді

$$a^q = (a^m)^{-p}a^n \in H.$$

Оскільки $0 \leq q < m$, то $q = 0$ згідно з вибором m. Отже, $a^n = (a^m)^p \in H$. Доведення завершено. □

Наслідок 40. Єдиними підгрупами $(\mathbb{Z}, +)$ є $(n\mathbb{Z}, +)$ для $n \in \mathbb{N}$.

Означення 41. Елемент $a \in G$ має скінченний порядок, якщо циклічна підгрупа $\langle a\rangle$ скінченна. Кількість елементів в цій підгрупі називається **порядком** елемента a в G.

Теорема 42. *Нехай $G = \langle a\rangle$ є циклічною групою, що складається з n елементів, $b = a^s \in G$. Тоді $H := \langle b\rangle$ є циклічною підгрупою, що складається з n/НСД(n, s).*

Доведення. Те, що H є циклічною, випливає з теореми 39. Залишається показати, що $|H| = n/\text{НСД}(n, s)$. Нехай m є мінімальним натуральним числом таким, що $b^m = e$. Тоді $H = \{e, b, b^2, \ldots, b^{m-1}\}$, тобто $|H| = m$ та $a^{sm} = e$. З іншого боку, $a^n = e$ та $a^k \neq e$, якщо k не є кратним n. Отже, sm має бути кратним n. Мінімальне m з такою властивістю є $n/\text{НСД}(n, s)$. Доведення завершено. □

Наслідок 43. Якщо $G = \langle a \rangle$ є скінченною циклічною групою з n елементів, то $G = \langle a^r \rangle$ тоді і тільки тоді, коли r і n взаємнопрості.

Приклад 44. Знайдемо всі підгрупи $\mathbb{Z}_{18}$ та побудуємо відповідну діаграму Гассе. Всі підгрупи є циклічними. Згідно з попереднім наслідком, елементи $1, 5, 7, 11, 13, 17$ є генераторами. Елемент 2 породжує циклічну підгрупу

$$\langle 2 \rangle = \{0, 2, 4, 6, 8, 10, 12, 14, 16\}$$

порядку 9. Її генераторами є елементи вигляду $2h$, де h взаємнопросте з 9, тобто $h = 1, 2, 4, 5, 7, 8$ та $2h = 2, 4, 8, 10, 14, 16$. Елемент 6 породжує групу $\{0, 6, 12\}$, цю ж групу породжує елемент 12.

Таким чином, ми знайшли підгрупи породжені елементами 0, 1, 2, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 16. Залишається знайти підгрупи, породжені елементами $3, 9, 15$. Маємо

$$\langle 3 \rangle = \{0, 3, 6, 9, 12, 15\},$$

цю саму групу породжує елемент 15. Елемент 9 породжує підгрупу порядку 2: $\{0, 9\}$. Діаграма Гассе представлена на рисунку 2.3.

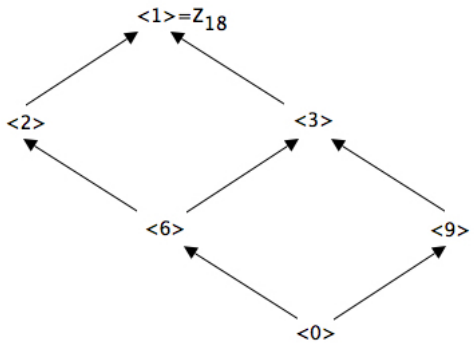

Рис. 2.3: Діаграма Гассе підгруп $\mathbb{Z}_{18}$.

2.3 Класи суміжності.

Уважний читач помітив, що таблиця 1.2 природнім чином розбивається на чотири блоки, кожен з яких складається лише з членів ρ_i або μ_i. Група $\mathfrak{S}_3$, таким чином, розбивається на два блоки B_μ та B_ρ однакового розміру, а множина $\{B_\mu, B_\rho\}$ є групою з таблицею Келі

	B_ρ	B_μ
B_ρ	B_ρ	B_μ
B_μ	B_μ	B_ρ

.

Назвемо будь-який елемент блоку його представником. Рівність $B_\rho B_\mu = B_\mu$ означає, що добуток **будь-яких** (в сенсі групової операції $\mathfrak{S}_3$) представників з B_ρ та B_μ, відповідно, є представником B_μ.

Нехай тепер G є довільною групою, яка розбита на блоки $\{B_i\}$. Спробуємо визначити, яким умовам має задовольняти це розбиття, щоб на множині $\{B_i\}$ групова операція G породжувала структуру групи. Іншими словами, ми шукатимемо умови, за яких добуток блоків не залежить від вибору їх представників.

Твердження 45. *Якщо група G може бути розбита на блоки $\{B_i\}$ так, що їх добуток коректно визначений і задає структуру групи на розбитті $\{B_i\}$, то блок B_e, що містить одиницю e групи G, повинен бути підгрупою G.*

Доведення. Оскільки $e \in B_e$ та добуток B_eB_e не залежить від вибору представників, то

$$B_eB_e = B_e,$$

що показує замкненість B_e відносно групової операції G. За визначенням $e \in B_e$. Нехай $a \in B_e$ та $a^{-1} \in B_k$ для деякого k, тоді B_eB_k містить e, а тому дорівнює B_e. З рівності $B_eB_k = B_eB_e$ випливає $B_k = B_e$, а тому $a^{-1} \in B_e$. Отже, B_e є підгрупою згідно з 33. Доведення завершено. □

Означення 46. Нехай $H < G$, $a \in G$. Лівим класом суміжності aH елемента a по підгрупі H називається множина

$$aH := \{ah : h \in H\}.$$

Аналогічно визначається правий клас суміжності Ha.

Твердження 47. *Якщо група G може бути розбита на блоки $\{B_i\}$ так, що їх добуток коректно визначений і задає структуру групи на розбитті $\{B_i\}$, то ці блоки мають бути лівими і одночасно правими класами суміжності по деякій підгрупі G. Зокрема, кожен лівий клас суміжності повинен бути правим класом суміжності.*

Доведення. З рівності $B_aB_e = B_a$ випливає, що $aB_e \subset B_a$. Нехай $a^{-1} \in B_k$ для деякого k, тоді $B_kB_a = B_e$, а тому $a^{-1}x =: b \in B_e$ для кожного $x \in B_a$. Звідси $x = ab$, а, отже, $B_a \subset aB_e$. Таким чином

$$B_a = aB_e.$$

Аналогічними міркуваннями перевіряється, що $B_a = B_ea$. Доведення завершено. □

Приклад 48. Знайдемо ліві класи суміжності в групі $(\mathbb{Z}, +)$ по підгрупі $3\mathbb{Z}$. Очевидно, що сама підгрупа $3\mathbb{Z}$ є лівим класом суміжності, оскільки $3\mathbb{Z} = 0 + 3\mathbb{Z}$. Іншими класами суміжності є $1 + 3\mathbb{Z}$ та $2 + 3\mathbb{Z}$. Перша з цих множин є множиною цілих чисел, які дають залишок 1 при діленні на 3, друга з цих множин є множиною цілих чисел, які дають залишок 2 при діленні на 3. Інших лівих класів суміжності немає. Зауважимо, що всі три класи суміжності не перетинаються.

Помітимо, що $b \in aH$ означає, що $a^{-1}b \in H$. Розглянемо бінарне відношення $\sim_H$ на множині елементів групи G, що визначене так: $a \sim_H b$ тоді і тільки тоді, коли $a^{-1}b \in H$. Має місце такий результат.

Теорема 49. *Нехай $H \leq G$. Відношення $\sim_H$ є відношенням еквівалентності на G. Класи еквівалентності по цьому відншенню є в точності лівими класами суміжності по підгрупі H. Всі ці класи суміжності рівнопотужні.*

Доведення. Нам потрібно показати, що $\sim_H$ є рефлексивним, симетричним та транзитивним відношенням.

РЕФЛЕКСИВНІСТЬ: Оскільки $a^{-1}a = e \in H$, то $a \sim_H a$.

СИМЕТРИЧНІСТЬ: Нехай $a \sim_H b$, тоді $a^{-1}b \in H$. Звідки $(a^{-1}b)^{-1} \in H$, оскільки H підгрупа. В силу того, що $(a^{-1}b)^{-1} = b^{-1}a$, маємо $b \sim_H a$.

ТРАНЗИТИВНІСТЬ: Нехай $a \sim_H b$ та $b \sim_H c$, тоді $a^{-1}b \in H$ та $b^{-1}c \in H$. Звідки, взявши добуток, $a^{-1}c \in H$, оскільки H підгрупа.

Таким чином, H є відношенням еквівалентності. Розглянемо клас еквівалентності $\overline{a}$, що містить a. Маємо

$$\overline{a} := \{x \in G : x \sim_H a\} = \{x \in G : a^{-1}x \in H\}$$
$$= \{x \in G : x = ah \text{ для деякого } h \in H\} = aH.$$

Отже, класи еквівалентності по відношенню $\sim_H$ є лівими класами суміжності по підгрупі H.

Залишається показати, що ліві класи суміжності рівнопотужні. Розглянемо відображення $\lambda_a : H \to aH$, що визначене так $\lambda_a(h) = ah$. Це відображення тривіальним чином є сюр'єктивним. Якщо $\lambda_a(h_1) = \lambda_a(h_2)$, то $ah_1 = ah_2$. Звідки $h_1 = h_2$ згідно з твердженнм 8. Отже, λ_a ін'єктивне і бієктивне. Це означає, що aH рівнопотужне H. Доведення завершено. □

Зауваження 50. Аналогічно можна визначити відношення еквівалентності $_H\sim$: $a \ {}_H\sim \ b$ тоді і тільки тоді, коли $ab^{-1} \in H$. Класами суміжності по цьому відношенню будуть праві класи суміжності по підгрупі H.

Приклад 51 (Множення лівих класів суміжності не завжди коректно визначене). Розглянемо підгрупу $H := \{\rho_0, \mu_1\}$ в $\mathfrak{S}_3$, див. приклад 19. Безпосереднім підрахунком пересвідчуємось, що ліві класи суміжності по підгрупі H такі:

$$H = \{\rho_0, \mu_1\}, \quad \rho_1 H = \{\rho_1, \mu_2\}, \quad \rho_2 H = \{\rho_2, \mu_3\}.$$

Запишемо таблицю Келі групи $\mathfrak{S}_3$, але її елементи згрупуємо по лівим класам суміжності: $\rho_0, \mu_1 | \rho_1, \mu_2 | \rho_2, \mu_3$.

	ρ_0	μ_1	ρ_1	μ_2	ρ_2	μ_3
ρ_0	ρ_0	μ_1	ρ_1	μ_2	ρ_2	μ_3
μ_1	μ_1	ρ_0	μ_3	ρ_2	μ_2	ρ_1
ρ_1	ρ_1	μ_2	ρ_2	μ_3	ρ_0	μ_1
μ_2	μ_2	ρ_1	μ_1	ρ_0	μ_3	ρ_2
ρ_2	ρ_2	μ_3	ρ_0	μ_1	ρ_1	μ_2
μ_3	μ_3	ρ_2	μ_2	ρ_1	μ_1	ρ_0

Як видно з затінення клітинок в наведеній таблиці, множення лівих класів суміжності неможливо коректно визначити. Множення елементів класу суміжності $\{\rho_0, \mu_1\}$ на елементи класу суміжності $\{\rho_1, \mu_2\}$ дає $\{\rho_1, \rho_2, \mu_1, \mu_2\}$, множину, яка не є класом суміжності.

Приклад 52. Розглянемо підгрупу $H := \{0, 2\} < \mathbb{Z}_4$. Класи суміжності по підгрупі H такі:

$$H = \{0, 2\}, \quad 1 + H = \{1, 3\}.$$

Таблиця Келі групи $\mathbb{Z}_4$, в якій елементи розбиті на класи суміжності $0, 2|1, 3$, така

	0	2	1	3
0	0	2	1	3
2	2	0	3	1
1	1	3	2	0
3	3	1	0	2

Бачимо, що множення (додавання) лівих класів суміжності коректно визначене і задає групу на $\{H, 1 + H\}$ з таблицею Келі

	H	1+H
H	H	1+H
1+H	1+H	H

.

Перш ніж описати ті підгрупи, за якими ліві класи суміжності утворюють групу, ми сформулюємо декілька важливих наслідків з вже доведених фактів.

2.4 Теорема Лагранжа та її наслідки.

Теорема 53. *Нехай G є групою з n елементів, $H < G$. Тоді $|H|$ ділить $n = |G|$.*

Доведення. Нехай $|H| = m$. Згідно з теоремою 49 набір лівих класів суміжності утворює диз'юнктне розбиття G. Оскільки всі ліві класи суміжності мають по m елементів, то $n = rm$, де r є числом лівих класів суміжності. Отже, m ділить n. □

Наслідок 54. Порядок елемента скінченної групи, див. означення 41, ділить порядок групи.

Означення 55. Нехай $H < G$. Число лівих класів суміжності по підгрупі H називається **індексом** підгрупи H в G і позначається $(G : H)$.

Наслідок 56. Припустимо, що $K < H < G$, а індекси $(G : H)$ та $(H : K)$ скінченні. Тоді $(G : K) = (G : H)(H : K)$.

2.5 Нормальні підгрупи. Фактор-групи.

Лема 57. *Нехай $H < G$ і нехай операція множення лівих (правих) класів суміжності по підгрупі H є коректно визначеною. Тоді множина лівих (правих) класів суміжності є групою.*

Доведення. Нагадаємо, що добуток $(aH)(bH)$ лівих класів суміжності визначається як клас суміжності, що містить добуток представників класів aH та bH. Коректність визначення добутку означає, що результат не залежить від вибору представників. Доведемо асоціативність:

$$[(aH)(bH)](cH) = (abH)(cH) = ((ab)c)H$$
$$= a(bc)H = (aH)(bcH) = (aH)[(bH)(cH)].$$

Одиницею групи є клас суміжності $eH = H$, оберненим елементом до aH є $a^{-1}H$. Доведення завершено. □

Центральним результатом теорії класів суміжності є наступна теорема.

Теорема 58. *Якщо $H < G$, то операція множення класів суміжності по підгрупі H є коректно визначеною тоді і тільки тоді, коли кожен правий клас суміжності є також лівим класом суміжності.*

Доведення. Припустимо, що операція множення класів суміжності по підгрупі H є коректно визначеною, тоді за попередньою лемою ліві класи суміжності утворюють групу. Тому, згідно з твердженням 47, ліві класи суміжності мають бути й правими класами суміжності.

Доведемо обернене твердження. Нехай кожен лівий клас суміжності aH є правим класом суміжності. Оскільки $g \in gH$ для кожного g, а правий клас суміжності, що містить g це Hg, то ми, фактично, припускаємо $gH = Hg$ для всіх

$g \in G$. Покажемо, що множення лівих класів суміжності коректно визначене. Нехай $a_1, a_2 \in aH$, $b_1, b_2 \in bH$. Перевіримо, що a_1b_1 та a_2b_2 лежать в одному класі суміжності. Маємо

$$a_1 = a_2h_1, \quad b_1 = b_2h_2,$$

для деяких $h_1, h_2 \in H$. Тоді $a_1b_1 = a_2h_1b_2h_2$. Оскільки $b_2H = Hb_2$, то $h_1b_2 = b_2h_3$ для деякого h_3. Отже, $a_1b_1 = a_2b_2h_3h_2$, а тому $a_1b_1 \in a_2b_2H$. Таким чином, a_1b_1 та a_2b_2 лежать в одному класі суміжності і операція їх множення коректна. Доведення завершено. □

Означення 59. Підгрупа H групи G називається **нормальною**, якщо $g^{-1}Hg = H$ для всіх $g \in G$. Те, що H є нормальною підгрупою в G, позначається $H \trianglelefteq G$.

Отже, ми довели, що саме нормальні підгрупи і тільки вони є тими підгрупами, класи суміжності за якими утворюють групу.

Маємо очевидий факт.

Твердження 60. *Кожна підгрупа абелевої групи є нормальною.*

Приклад 61. Як ми бачили у прикладі 51, підгрупа $\{\rho_0, \mu_1\}$ не є нормальною в $\mathfrak{S}_3$. Дійсно

$$(\rho_1)^{-1}\mu_1\rho_1 = \rho_2\mu_1\rho_1 = \mu_2 \notin \{\rho_0, \mu_1\}.$$

Означення 62. Нехай $N \trianglelefteq G$. Група класів суміжності по підгрупі N називається **фактор-групою** G по нормальній підгрупі N і позначається G/N. Класи суміжності називаються також **класами лишків** G **по модулю** N.

Приклад 63. Для довільного $n \in \mathbb{N}$ маємо $n\mathbb{Z} \trianglelefteq \mathbb{Z}$. Фактор група $\mathbb{Z}/n\mathbb{Z}$ складається з n класів лишків:

$$0 + n\mathbb{Z}, \quad 1 + n\mathbb{Z}, \quad \ldots, \quad (n-1) + n\mathbb{Z}$$

для яких

$$(i + n\mathbb{Z}) + (j + n\mathbb{Z}) = ((i + j) \bmod n) + n\mathbb{Z}, \quad i, j = 0, \ldots, n-1.$$

Приклад 64. Нехай $n \geq 3$, тоді[1] $A_n \trianglelefteq \mathfrak{S}_n$. Дійсно, оскільки парність σ^{-1} та σ однакова для довільної $\sigma \in \mathfrak{S}_n$, то $\sigma^{-1}\tau\sigma \in A_n$ для кожного $\tau \in A_n$ та $\sigma \in$

[1]Нагадаємо, що A_n є групою парних підстановок на $\{1, 2, \ldots, n\}$.

$\mathfrak{S}_n$. Фактор-група $\mathfrak{S}_n/A_n$ складається з двох класів лишків: парних та непарних підстановок і має структуру $\mathbb{Z}_2$.

Лекція 3

Поняття ізоморфізму, гомоморфізму. Теореми про ізоморфізм. Теорема Келі

3.1 Ізоморфізми

У попередніх підрозділах ми часто використовували словосполучення "група G_1 має таку саму структуру, що й група G_2". Надамо цій фразі строгий математичний зміст.

Означення 65. Ізоморфізмом груп $(G_1, *)$ та $(G_2, \times)$ називається бієкція $\phi : G_1 \mapsto G_2$ така, що

$$\phi(g * h) = \phi(g) \times \phi(h), \quad g, h \in G_1.$$

Ізоморфізм груп $(G_1, *)$ та $(G_2, \times)$ позначається $G_1 \cong G_2$.

Теорема 66. *Нехай $\phi : G_1 \mapsto G_2$ є ізоморфізмом, e – одиниця в G_1. Тоді $\phi(e)$ – одиниця в G_2 та $\phi(a^{-1}) = (\phi(a))^{-1}$.*

Доведення. Нехай $x_2 \in G_2$. Оскільки ϕ – бієкція, то знайдеться $x_1 \in G_1$, що $\phi(x_1) = x_2$. Маємо

$$x_2 = \phi(x_1) = \phi(x_1 e) = \phi(x_1)\phi(e) = x_2\phi(e)$$

та

$$x_2 = \phi(x_1) = \phi(ex_1) = \phi(e)\phi(x_1) = \phi(e)x_2.$$

Отже, $\phi(e)$ є одиницею в G_2. Нехай $a \in G_1$. Маємо

$$\phi(e) = \phi(a^{-1}a) = \phi(a^{-1})\phi(a) \quad \text{та} \quad \phi(e) = \phi(aa^{-1}) = \phi(a)\phi(a^{-1}).$$

Отже, $\phi(a^{-1})$ є оберненим до $\phi(a)$ в G_2. Доведення завершено. $\square$

Приклад 67. Групи $(\mathbb{R}, +)$ та $((0, \infty), \times)$ є ізоморфними. Ізоморфізм задається відображенням $x \mapsto e^x$.

Приклад 68. Кожна нескінченна циклічна група $G = \langle a \rangle$ ізоморфна групі $(\mathbb{Z}, +)$. Ізоморфізм задається відображенням $\phi(a^n) = n$.

Приклад 69. Групи $(\mathbb{Z}, +)$ та $(\mathbb{Q}, +)$ неізоморфні. Перша з груп циклічна, а друга – ні. Групи $(\mathbb{Q}_+^*, \times)$ та $((0, \infty), \times)$ неізоморфні тому, що множини $\mathbb{Q}_+^*$ та $(0, \infty)$ не є рівнопотужними.

Приклад 70 (Порядок підстановок при їх композиції). Нехай $A \neq \varnothing$ є довільною непорожньою множиною, а $(\mathfrak{S}_A, \circ)$ є симетричною групою, що була визначена в Розділі 1.4. Визначимо бінарну операцію $\circ_1$ на $\mathfrak{S}_A$ рівністю

$$(\sigma \circ_1 \tau)(x) = \sigma(\tau(x)), \quad x \in A.$$

Тоді $(\mathfrak{S}_A, \circ_1)$ є групою, що ізоморфна $(\mathfrak{S}_A, \circ)$. Ізоморфізм задається бієктивним відображенням $\phi(\sigma) = \sigma^{-1}$. Дійсно,

$$\begin{aligned}(\phi(\sigma \circ \tau))(x) = (\sigma \circ \tau)^{-1}(x) = (\tau^{-1} \circ \sigma^{-1})(x) = \sigma^{-1}(\tau^{-1}(x)) \\ = (\phi(\sigma) \circ_1 \phi(\tau))(x), \quad \sigma, \tau \in \mathfrak{S}_A, \quad x \in A.\end{aligned}$$

3.2 Теорема Келі

Проаналізувавши будь-яку з таблиць Келі, що нам зустрічались раніше, можна побачити, що кожному з елементів групи відповідає рядок (або стовпчик), що є підстановкою на множині елементів групи. З огляду на це, не є дивним, що принаймні кожна скінченна група G має бути ізоморфною деякій підгрупі підстановок на G, тобто деякій підгрупі $\mathfrak{S}_G$. Насправді, це твердження вірне також і для нескінченних груп і називається **теоремою Келі**.

Теорема 71. *Кожна група G ізоморфна деякій групі підстановок.*

Доведення. Розглянемо симетричну групу $\mathfrak{S}_G$ на множині елементів групи G і знайдемо підгрупу $G' < \mathfrak{S}_G$, ізоморфну G. Для фіксованого $a \in G$ розглянемо від-

ображення $\rho_a : G \mapsto G$, що визначене так

$$\rho_a(x) = xa.$$

Якщо $\rho_a(x) = \rho_a(y)$, то $xa = ya$, а тому $x = y$ і ρ_a – ін'єктивне. З іншого боку, для кожного $y \in G$,

$$\rho_a(ya^{-1}) = ya^{-1}a = y,$$

тому ρ_a – сюр'єктивне. Отже, ρ_a є підстановкою на G. Покладемо

$$G' := \{\rho_a : a \in G\}$$

і покажемо, що G' є групою.

Замкненість G' відносно композиції. Покажемо, що $\rho_a\rho_b = \rho_{ab}$. Маємо

$$(\rho_a\rho_b)(x) = \rho_b(\rho_a(x)) = (xa)b = x(ab) = \rho_{ab}(x)$$

для кожного $x \in G$.

Існування одиниці. Оскільки $\rho_e(x) = xe = x$, то $\rho_e -$ тотожня підстановка, яка є одиницею G'.

Існування оберненого елемента. З рівностей

$$\rho_a\rho_{a^{-1}} = \rho_{aa^{-1}} = \rho_e = \rho_{a^{-1}a} = \rho_{a^{-1}}\rho_a$$

випливає, що $\rho_{a^{-1}}$ є оберненим елементом до ρ_a. З того, що $a^{-1} \in G$ випливає, що $\rho_{a^{-1}} \in G'$.

Залишається перевірити, що відображення $\phi : G \mapsto G'$, що визначене рівністю

$$\phi(a) = \rho_a,$$

є ізоморфізмом. Перевіримо бієктивність. Якщо $\phi(a) = \phi(b)$, то $\rho_a = \rho_b$, зокрема $\rho_a(e) = \rho_b(e)$, а тому $a = b$. Сюр'єктивність випливає з означення. Нарешті,

$$\phi(ab) = \rho_{ab} = \rho_a\rho_b = \phi(a)\phi(b),$$

тому ϕ є ізоморфізмом. Доведення завершено. □

Зауваження 72. Розглянувши відображення $\lambda_a : G \mapsto G$, що визначене рівністю

$$\lambda_a(x) = ax,$$

можна показати, що $G \cong G'' := \{\lambda_a : a \in G\}$. Ізоморфізм $\psi : G \mapsto G''$ визначається рівністю

$$\psi(a) = \lambda_{a^{-1}}, \quad a \in G.$$

Означення 73. Група G', побудована в попередньому доведенні, називається **правим регулярним представленням** G. Аналогічно, група G'' з попереднього зауваження називається **лівим регулярним представленням** G

3.3 Гомоморфізми

Ізоморфізм груп був визначений як бієктивне відображення ϕ таке, що

$$\phi(ab) = \phi(a)\phi(b).$$

Якщо в цьому означенні відкинути вимогу бієктивності, то ми прийдемо до поняття гомоморфізму.

Означення 74. Гомоморфізмом груп $(G_1, *)$ та $(G_2, \times)$ називається відображення $\phi : G_1 \mapsto G_2$ таке, що

$$\phi(g * h) = \phi(g) \times \phi(h), \quad g, h \in G_1.$$

Приклад 75. Відображення $\gamma : \mathbb{Z} \mapsto \mathbb{Z}_n$, що визначене рівністю

$$\gamma(x) := r,$$

де r – залишок від ділення x на n, є гомоморфізмом груп $(\mathbb{Z}, +)$ та $(\mathbb{Z}_n, +)$. Якщо ототожнити $\mathbb{Z}_n$ з фактор-групою $\mathbb{Z}/n\mathbb{Z}$, то відображення γ ставить у відповідність числу x його клас лишків по модулю $n\mathbb{Z}$.

Твердження 76. *Якщо $N \trianglelefteq G$, то канонічне відображення $\gamma : G \mapsto G/N$, що визначене рівністю*

$$\gamma(a) = aN, \quad a \in G,$$

є гомоморфізмом.

Доведення. Це випливає з визначення добутку класів суміжності:

$$\gamma(ab) = (ab)N = (aN)(bN) = \gamma(a)\gamma(b).$$

□

Означення 77. Ядром гомоморфізму $\phi : G_1 \mapsto G_2$ називається множина

$$\ker \phi := \{g \in G_1 : \phi(g) = e_2\},$$

де e_2 – одиниця групи G_2.

Приклад 78. У прикладі 75 маємо $\ker \gamma = n\mathbb{Z}$, зокрема $\ker \gamma$ – нормальна підгрупа $\mathbb{Z}$ та

$$\mathbb{Z}/\ker \gamma = \mathbb{Z}_n.$$

Теорема 79. *Нехай $\phi : G_1 \mapsto G_2$ є гомоморфізмом. Тоді:*

(i) Якщо e_1 – одиниця в G_1, то $\phi(e_1)$ – одиниця в G_2; для довільного $a \in G_1$ маємо $\phi(a^{-1}) = (\phi(a))^{-1}$.

(ii) Якщо $H_1 < G_1$ ($H_1 \trianglelefteq G_1$), то $\phi(H_1) < G_2$ ($\phi(H_1) \trianglelefteq G_2$).

(iii) Якщо $H_2 < G_2$ ($H_2 \trianglelefteq G_2$), то $\phi^{-1}(H_2) < G_1$ ($\phi^{-1}(H_2) \trianglelefteq G_1$).

Доведення. Доведення (i). З рівності $\phi(a) = \phi(ae_1) = \phi(a)\phi(e_1)$, $a \in G_1$, випливає, що $\phi(e_1)$ – одиниця в G_2. З рівності $\phi(e_1) = \phi(aa^{-1}) = \phi(a)\phi(a^{-1})$ випливає, що $\phi(a^{-1}) = (\phi(a))^{-1}$.
Доведення (ii). Нехай $H_1 < G_1$ та $\phi(a), \phi(b) \in \phi(H_1)$. Тоді $\phi(a)\phi(b) = \phi(ab) \in \phi(H_1)$. Отже, $\phi(H_1)$ замкнена відносно групової операції G_2. З частини (i) випливає, що $\phi(e_1) \in \phi(H_1)$ та $(\phi(a))^{-1} = \phi(a^{-1}) \in \phi(H_1)$. Тому $\phi(H_1)$ є підгрупою G_2. Якщо H_1 є нормальною підгрупою G_1, то для довільних $g \in G_1, h \in H_1$

$$(\phi(g))^{-1}\phi(h)\phi(g) = \phi(g^{-1})\phi(h)\phi(g) = \phi(g^{-1}hg) = \phi(h_1) \in \phi(H_1).$$

В останній рівності ми використали те, що $g^{-1}hg \in H_1$, оскільки H_1 – нормальна.
Доведення (iii). Нехай $H_2 < G_2$ та $a, b \in \phi^{-1}(H_2)$. Тоді $\phi(a), \phi(b) \in H_2$, тому $\phi(ab) \in H_2$ і $ab \in \phi^{-1}(H_2)$. Одиниця $e_2 = \phi(e_1)$ групи G_2 лежить в H_2, тому $e_1 \in \phi^{-1}(H_2)$. Якщо $a \in \phi^{-1}(H_2)$, то $\phi(a) \in H_2$. Звідки

$$\phi(a^{-1}) = (\phi(a))^{-1} \in H_2,$$

а тому $a^{-1} \in \phi^{-1}(H_2)$. Отже, $\phi^{-1}(H_2)$ є підгрупою G_1. Якщо H_2 є нормальною підгрупою G_2, то для довільних $g_1 \in G_1, h \in \phi^{-1}(H_2)$

$$\phi(g_1^{-1}hg_1) = \phi(g_1^{-1})\phi(h)\phi(g_1) \in H_2,$$

звідки $g_1^{-1}hg_1 \in \phi^{-1}(H_2)$, що доводить нормальність $\phi^{-1}(H_2)$. Доведення завершено. □

3.4 Теореми про ізоморфізм

Нехай $\phi : G_1 \mapsto G_2$ є гомоморфізмом. З теореми 79 випливає, що $\ker \phi = \phi^{-1}(\{e_2\})$, де e_2 – одиниця G_2, є нормальною підгрупою, оскільки $\{e_2\}$ є нормальною підгрупою в G_2.

Теорема 80. *Нехай K є ядром гомоморфізма $\phi : G_1 \mapsto G_2$. Тоді $\phi(G_1)$ є групою, яка (канонічно) ізоморфна фактор-групі G_1/K.*

Доведення. Те, що $\phi(G_1)$ є групою, випливає з частини (ii) теореми 79. Визначимо відображення $\psi : G_1/K \mapsto \phi(G_1)$ формулою

$$\psi(aK) = \phi(a).$$

Перш за все ми маємо перевірити, що це визначення коректне, оскільки воно може залежати від вибору представника класу суміжності aK. Нехай $b \in aK$, тоді $b = ak_1$ для деякого $k_1 \in K$, а тому

$$\phi(b) = \phi(ak_1) = \phi(a)\phi(k_1) = \phi(a).$$

Отже, $\phi(a) = \phi(b)$ і $\psi(aK)$ не залежить від вибору представника класу суміжності aK. Перевіримо, що ψ є ізоморфізмом. Сюр'єктивність очевидна. Якщо $\psi(a_1K) = \psi(a_2K)$, то $\phi(a_1) = \phi(a_2)$, а тому

$$\phi(a_1a_2^{-1}) = \phi(a_1)(\phi(a_2))^{-1} = \phi(a_1)(\phi(a_1))^{-1} = e_2.$$

Це демонструє, що $a_1a_2^{-1} \in K$, а, отже, $a_1 \sim_K a_2$ і $a_1K = a_2K$. Отже, ψ є ін'єктивним відображенням. Рівність

$$\psi(aK)\psi(bK) = \phi(a)\phi(b) = \phi(ab) = \psi(abK) = \psi((aK)(bK))$$

демонструє, що ψ – ізоморфізм. Цей ізоморфізм є канонічним в такому розумінні: якщо γ – канонічний гомоморфізм $\gamma : G_1 \mapsto G_1/K$, то

$$\phi = \gamma\psi,$$

Говорять, що діаграма 3.1 комутує.

Доведення завершено. □

Теорему 80 також називають **першою теоремою про ізоморфізм**, оскільки її можна переформулювати у такому вигляді.

Теорема 81. *Нехай K є ядром гомоморфізма $\phi : G_1 \mapsto G_2$, а $\gamma : G_1 \mapsto G_1/K$ є канонічним гомоморфізмом, тоді існує єдиний ізоморфізм $\psi : G_1/K \mapsto \phi(G)$ такий, що $\phi(x) = \psi(\gamma(x))$ для всіх $x \in G_1$.*

Означення 82. Нехай $N, H < G$ тоді $HN := \{hn : n \in N, h \in H\}$, а $H \vee N$ є найменшою підгрупою G, що містить HN (еквівалентно, що містить N та H).

Лема 83. *Якщо $N \trianglelefteq G$ та $H < G$, то $NH = HN = N \vee H$. Якщо також $H \trianglelefteq G$, то $NH \trianglelefteq G$.*

Доведення. Покажемо, що HN є підгрупою G. Нехай $h_1, h_2 \in H$ та $n_1, n_2 \in N$. Оскільки $N \trianglelefteq G$, то існує $n_3 \in N$ такий, що $n_1 h_2 = h_2 n_3$. Тому $(h_1 n_1)(h_2 n_2) = h_1(h_2 n_3)n_2 \in HN$, а отже, HN замкнена відносно множення елементів. Очевидно, що $e = ee \in HN$. Для $h \in H, n \in N$ маємо $(hn)^{-1} n^{-1} h^{-1} = h^{-1} n_4$ для деякого $n_4 \in N$. Отже, $(hn)^{-1} \in HN$ і тому $HN < G$. Аналогічно перевіряється, що $NH < G$. Якщо $H \trianglelefteq G$, $h \in H$, $n \in N$ та $g \in G$, то $g^{-1} hng = (g^{-1} hg)(g^{-1} ng) \in HN$. Це демонструє, що $HN \trianglelefteq G$. Доведення завершено. □

Наступна теорема носить назву **другої теореми про ізоморфізм**.

Теорема 84. *Нехай $H < G$, $N \trianglelefteq G$. Тоді $(HN)/N \cong H/(H \cap N)$.*

Доведення. Оскільки $N \trianglelefteq G$, то $H \cap N \trianglelefteq H$. Дійсно, якщо $b \in H \cap N$, $a \in H$, то $a^{-1} ba \in N$ в силу $b \in N$. Очевидно, що $a^{-1} ba \in H$. Таким чином, $a^{-1} ba \in H \cap N$, звідки випливає $H \cap N \trianglelefteq H$.

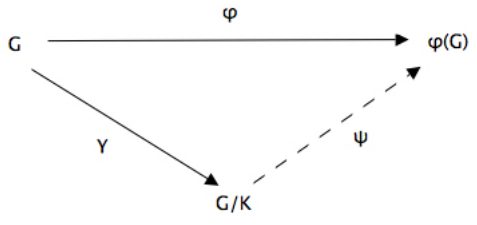

Рис. 3.1: Ілюстрація основної теореми про гомоморфізм.

Покажемо, що відображення $\phi : HN \mapsto H/(H \cap N)$, задане рівністю

$$\phi(hn) = h(H \cap N), \quad h \in H, n \in N,$$

є коректно визначеним. Нехай $h_1 \in H$, $n_1 \in N$ та $h_1 n_1 = hn$, тоді $h^{-1}h_1 = nn_1^{-1}$, а тому $h^{-1}h_1 \in H \cap N$. Звідси випливає, що $h(H \cap N) = h_1(H \cap N)$ і $\phi(h_1 n_1) = \phi(hn)$. Перевіримо, що ϕ є гомоморфізмом. Нехай $h_1, h_2 \in H$, $n_1, n_2 \in N$ та $n_1 h_2 = h_2 n_3$ для деякого $n_3 \in N$. Маємо

$$\begin{aligned}\phi((h_1 n_1)(h_2 n_2)) = \phi(h_1 h_2 n_3 n_2) = h_1 h_2 (H \cap N) \\ = h_1(H \cap N) h_2(H \cap N) = \phi(h_1 n_1)\phi(h_2 n_2).\end{aligned}$$

Очевидно, що $\phi(HN) = H/(H \cap N)$. Покажемо, що $\ker \phi = N$. Ядро ϕ складається з усіх таких $hn \in HN$, що $h \in H \cap N$. Отже, $\ker \phi = (H \cap N)N = N$. Таким чином, за першою теоремою про ізоморфізм $(HN)/N \cong H/(H \cap N)$. Доведення завершено. □

Теорема 85 (Третя теорема про ізоморфізм)**.** *Нехай $H, K \trianglelefteq G$, $K < H$. Тоді $G/H \cong (G/K)/(H/K)$.*

Доведення. Розглянемо відображення $\phi : G \mapsto (G/K)/(H/K)$, задане рівністю

$$\phi(a) = (aK)(H/K), \quad a \in G.$$

Очевидно, що $\phi(G) = (G/K)/(H/K)$. Для $a, b \in G$ маємо

$$\begin{aligned}\phi(ab) = [(ab)K](H/K) = [(aK)(bK)](H/K) \\ = [(aK)(H/K)][(bK)(H/K)] = \phi(a)\phi(b),\end{aligned}$$

отже, ϕ є гомоморфізмом. Його ядро – це всі такі $x \in G$, що $\phi(x) = H/K$, тобто $(xK)(H/K) = H/K$, а, отже, $x \in H$. Твердження теореми випливає з першої теореми про гомоморфізм. Доведення завершено. □

Твердження теореми 85 найкраще ілюструється комутативною діаграмою на Рис. 3.2.

В наступному підрозділі ми доведемо два фундаментальних результати теорії груп – теореми Шраєра та Жордана–Гьольдера. Для їх доведення нам знадобиться

ще один результат про ізоморфізм, що носить назву – **лема про метелика** або **лема Зассенхауса**[1]. Нехай $H, K < G$, $H^* \trianglelefteq H$, $K^* \trianglelefteq K$. З леми 83 випливає, що $H^*(H \cap K)$ є групою. Аналогічно, $H^*(H \cap K^*)$, $K^*(H \cap K)$ та $K^*(H^* \cap K)$ є групами. Також маємо $H^* \cap K \trianglelefteq H \cap K$ та $L := (H^* \cap K)(H \cap K^*) < H \cap K$. Ці співвідношення можна зобразити у вигляді діаграми Гассе, див. Рис. 3.3.

Має місце така лема.

Лема 86 (Лема про метелика)**.** *Нехай* $H, K < G$, $H^* \trianglelefteq H$, $K^* \trianglelefteq K$. *Тоді*

1) $H^*(H \cap K^*) \trianglelefteq H^*(H \cap K)$;

2) $K^*(H^* \cap K) \trianglelefteq K^*(H \cap K)$;

[1]Ганс Зассенхаус, 1912 – 1991, німецький математик.

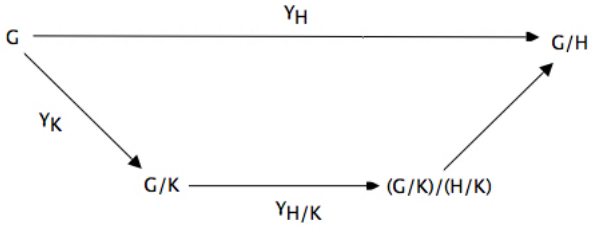

Рис. 3.2: Комутативна діаграма до Теореми 85.

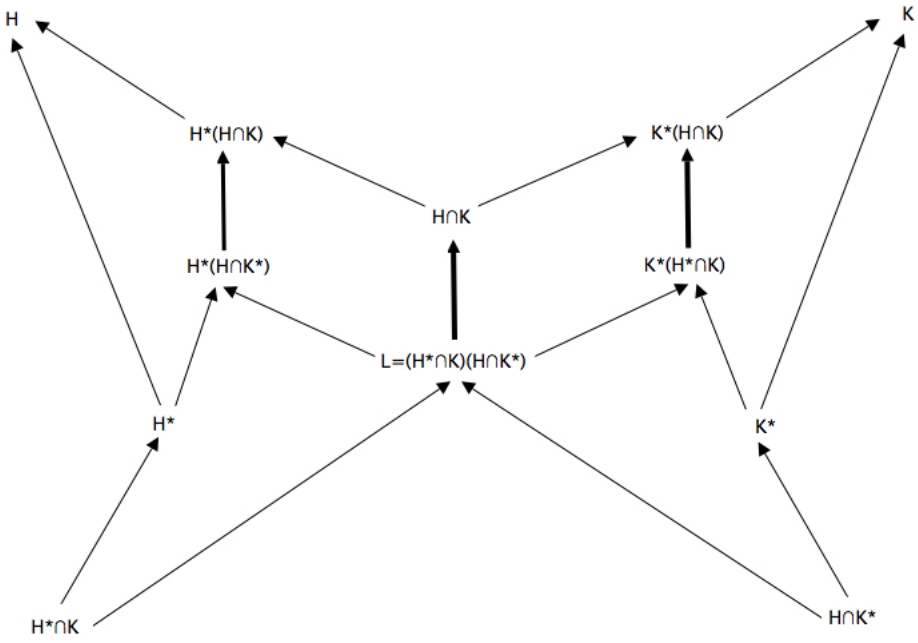

Рис. 3.3: Ілюстрація до леми 86.

3) $L \trianglelefteq H \cap K$;

4) $H^*(H \cap K)/H^*(H \cap K^*) \cong K^*(H \cap K)/K^*(H^* \cap K) \cong (H \cap K)/L$.

Іншими словами, кожна з трьох жирних стрілок на Рис. 3.3 задає відношення "є нормальною підгрупою", а всі три відповідні фактор-групи ізоморфні.

Доведення. Оскільки $H \cap K^* \trianglelefteq H \cap K$ та $H^* \cap K \trianglelefteq H \cap K$, то за лемою 83 маємо твердження частини 3) леми про метелика. Доведемо частину 1) та перше співвідношення частини 4). Частина 2) та друге співвідношення частини 4) будуть випливати з міркувань симетрії. Для доведення частини 1) достатньо побудувати гомоморфізм $\phi : H^*(H \cap K) \mapsto (H \cap K)/L$ такий, що $\phi(H^*(H \cap K)) = (H \cap K)/L$ та $\ker \phi = H^*(H \cap K^*)$. Перше співвідношення частини 4) тоді випливатиме з першої теореми про ізоморфізм.

Визначимо гомоморфізм $\phi : H^*(H \cap K) \mapsto (H \cap K)/L$ рівністю

$$\phi(hx) = xL, \quad h \in H^*, x \in H \cap K.$$

Перевіримо коректність визначення. Нехай $h_1, h_2 \in H^*$, $x_1, x_2 \in H \cap K$. Якщо $h_1 x_1 = h_2 x_2$, то $h_2^{-1} h_1 = x_2 x_1^{-1} \in H^* \cap (H \cap K) = H^* \cap K \subset L$, а тому $x_1 L = x_2 L$. Таким чином, відображення ϕ коректно визначене. Його сюр'єктивність очевидна. Оскільки $H^* \trianglelefteq H$, то знайдеться $h_3 \in H^*$ такий, що $x_1 h_2 = h_3 x_1$. Маємо

$$\phi((h_1 x_1)(h_2 x_2)) = \phi((h_1)(h_3)(x_1 x_2)) = x_1 x_2 L = (x_1 L)(x_2 L) = \phi(h_1 x_1)\phi(h_2 x_2).$$

Залишається перевірити, що $\ker \phi = H^*(H \cap K^*)$. Якщо $h \in H^*$, $x \in H \cap K$, то $\phi(hx) = xL = L$ тоді і тільки тоді, коли $x \in L$. Отже, $hx \in H^* L = H^*(H^* \cap K)(H \cap K^*) = H^*(H \cap K^*)$. Доведення завершено. □

3.5 Ряди груп. Теореми Шраєра та Жордана – Гьольдера.

Означення 87. Субнормальним рядом групи G називається скінченна послідовність $\{e\} = H_0, H_1, \ldots, H_n = G$ підгруп G така, що $H_i \trianglelefteq H_{i+1}$. Якщо всі H_i є нормальними підгрупами G, то ряд називається **нормальним**.

Приклад 88. Два ряди підгруп групи $(\mathbb{Z}, +)$:

$$\{0\} < 8\mathbb{Z} < 4\mathbb{Z} < \mathbb{Z}, \qquad \{0\} < 9\mathbb{Z} < \mathbb{Z}$$

є нормальними.

Приклад 89. Ряд підгруп групи D_4:

$$\{\rho_0\} < \{\rho_0, \mu_1\} < \{\rho_0, \rho_2, \mu_1, \mu_2\} < D_4$$

є субнормальним, але не є нормальним, оскільки $\{\rho_0, \mu_1\}$ не є нормальною підгрупою D_4.

Означення 90. Субнормальний (нормальний) ряд $\{K_j\}$ називається **ущільненням** субнормального (нормального) ряду $\{H_i\}$, якщо $\{H_i\} \subset \{K_j\}$.

Приклад 91. Ряд

$$\{0\} < 72\mathbb{Z} < 24\mathbb{Z} < 8\mathbb{Z} < 4\mathbb{Z} < Z$$

є ущільненням ряду

$$\{0\} < 72\mathbb{Z} < 8\mathbb{Z} < Z.$$

Означення 92. Два субнормальні ряди $\{H_i\}$ та $\{K_j\}$ називаються ізоморфними, якщо існує взаємнооднозначна відповідність між множинами фактор-груп $\{H_{i+1}/H_i\}$ та $\{K_{j+1}/K_j\}$ така, що фактор-групи, які відповідають одна одній, ізоморфні.

Приклад 93. Два ряди

$$\{0\} < \langle 5\rangle < \mathbb{Z}_{15}$$

та

$$\{0\} < \langle 3\rangle < \mathbb{Z}_{15}$$

ізоморфні, оскільки $\mathbb{Z}_{15}/\langle 5\rangle \cong \langle 3\rangle/\{0\}$ та $\mathbb{Z}_{15}/\langle 4\rangle \cong \langle 5\rangle/\{0\}$.

Ключовою є така **теорема Шраєра**[2].

Теорема 94. *Два субнормальні (нормальні) ряди групи G завжди мають ізоморфні ущільнення.*

Перед довденням наведемо приклад.

[2]Отто Шраєр, 1901 – 1929, австрійський математик.

Приклад 95. Спробуємо знайти ізоморфні ущільнення рядів

$$\{0\} < 8\mathbb{Z} < 4\mathbb{Z} < Z$$

та

$$\{0\} < 9\mathbb{Z} < \mathbb{Z}$$

з прикладу 88. Розглянемо ущільнення

$$\{0\} < 72\mathbb{Z} < 8\mathbb{Z} < 4\mathbb{Z} < \mathbb{Z}$$

ряду $\{0\} < 8\mathbb{Z} < 4\mathbb{Z} < Z$ та ущільнення

$$\{0\} < 72\mathbb{Z} < 18\mathbb{Z} < 9\mathbb{Z} < \mathbb{Z}$$

ряду $\{0\} < 9\mathbb{Z} < \mathbb{Z}$. Чотири фактор-групи в кожному з цих ущільнень ізоморфні $\mathbb{Z}_4, \mathbb{Z}_2, \mathbb{Z}_9$ та $72\mathbb{Z} \cong \mathbb{Z}$.

Доведення теореми Шраєра. Нехай G є групою з двома субнормальними рядами

$$\{e\} = H_0 < H_1 < H_2 < \cdots < H_n = G \tag{3.1}$$

та

$$\{e\} = K_0 < K_1 < K_2 < \cdots < K_m = G. \tag{3.2}$$

Сформуємо послідовність груп

$$\begin{aligned} H_i = H_i(H_{i+1} \cap K_0) &\leq H_i(H_{i+1} \cap K_1) \\ &\leq \cdots \leq H_i(H_{i+1} \cap K_m) = H_{i+1}, \quad 0 \leq i \leq n-1 \end{aligned}$$

і позначимо $H_{i,j} := H_i(H_{i+1} \cap K_j)$. Маємо ланцюжок (необов'язково різних!) груп:

$$\begin{aligned} \{e\} = H_{0,0} &\leq H_{0,1} \leq H_{0,2} \leq \cdots \leq H_{0,m-1} \leq H_{0,m} = H_{1,0} = H_1 \\ &\leq H_{1,1} \leq H_{1,2} \leq \cdots \leq H_{1,m-1} \leq H_{1,m} = H_{2,0} = H_2 \\ &\leq H_{2,1} \leq H_{2,2} \leq \cdots \leq H_{2,m-1} \leq H_{2,m} = H_{3,0} = H_3 \\ &\leq \cdots \leq \\ &\leq H_{n-1,1} \leq H_{n-1,2} \leq \cdots \leq H_{n-1,m-1} \leq H_{n-1,m} = G. \end{aligned} \tag{3.3}$$

Згідно з лемою про метелика цей ланцюжок є субнормальним рядом, тобто рядом в якому кожна група є нормальною підгрупою наступної. Побудований субнормальний ряд є ущільненням ряду (3.1).

Проробимо аналогічну процедуру з рядом (3.2). Покладемо $K_{j,i} := K_j(K_{j+1} \cap H_i)$ для $0 \le j \le m-1$ та $0 \le i \le n$. Це дасть субнормальний ряд

$$\begin{aligned}\{e\} = K_{0,0} &\le K_{0,1} \le K_{0,2} \le \cdots \le K_{0,n-1} \le K_{0,n} = K_{1,0} = K_1 \\ &\le K_{1,1} \le K_{1,2} \le \cdots \le K_{1,n-1} \le K_{1,n} = K_{2,0} = K_2 \\ &\le K_{2,1} \le K_{2,2} \le \cdots \le K_{2,n-1} \le K_{2,n} = K_{3,0} = K_3 \\ &\le \cdots \le \\ &\le K_{m-1,1} \le K_{m-1,2} \le \cdots \le K_{m-1,n-1} \le K_{m-1,n} = G. \end{aligned} \tag{3.4}$$

В цьому субнормальному ряді міститься $mn+1$ необов'язково різних груп та $K_{j,0} = K_j$. Побудований субнормальний ряд є ущільненням ряду (3.2).

Згідно з частиною 4 леми про метелика

$$H_i(H_{i+1} \cap K_{j+1})/H_i(H_{i+1} \cap K_j) \cong K_j(K_{j+1} \cap H_{j+1})/K_j(K_{j+1} \cap H_i)$$

або в наших позначеннях

$$H_{i,j+1}/H_{i,j} \cong K_{j,i+1}/K_{j,i}, \quad 0 \le i \le n-1, \quad 0 \le j \le m-1.$$

Таким чином, ряди (3.3) та (3.4) є ізоморфними: фактор-група, породжена символом $\le$ в i-му рядку та j-му стовпчику ряда (3.3), ізоморфна фактор-групі, породженій символом $\le$ в j-му рядку та i-му стовпчику ряда 3.4. Видаляючи з рядів (3.3) та (3.4) групи, що співпадають, отримуємо ізоморфні субнормальні ряди, що є ущільненнями субнормальних рядів (3.1) та (3.2), відповідно. Це завершує доведення теореми Шраєра для субнормальних рядів.

Для нормальних рядів твердження теореми випливає з того, що всі проміжні групи $H_{i,j}$ та $K_{j,i}$ є нормальними в G згідно з лемою 83. □

Щойно доведена теорема Шраєра є основним інгредієнтом у доведенні однієї з найважливіших теорем теорії груп – теореми Жордана[3] – Гьольдера[4]. Перш ніж сформулювати цю теорему, наведемо означення та декілька прикладів.

[3]Марі Енмóн Камі́ль Жордан, 1838 – 1922, французький математик.

[4]Отто Гьольдер, 1859 – 1937, німецький математик.

Означення 96. Субнормальний ряд $\{H_i\}$ групи G називається **композиційним рядом**, якщо кожна фактор-група H_{i+1}/H_i є **простою**, тобто не містить жодної невласної нормальної підгрупи. Ряд $\{H_i\}$ групи G, що є одночасно нормальним та композиційним називається **головним рядом** групи.

Приклад 97. Покажемо, що $(\mathbb{Z}, +)$ не містить композиційних (а тому і головних) рядів. Нехай

$$\{0\} = H_0 < H_1 < H_2 < \cdots < H_{n-1} < H_n = \mathbb{Z}$$

є субнормальним рядом. Ми знаємо, див. наслідок 40, що $H_1 = r\mathbb{Z}$ для деякого $r \in \mathbb{N}$, а тому $H_1/H_0 \cong r\mathbb{Z} \cong Z$. Отже, H_1/H_0 не є простою групою.

Приклад 98. Ряди в прикладі 93 є головними, оскільки групи $\mathbb{Z}_{15}/\langle 5\rangle \cong \langle 5\rangle/\{0\} \cong \mathbb{Z}_3$ та $\mathbb{Z}_{15}/\langle 3\rangle \cong \langle 3\rangle/\{0\} \cong \mathbb{Z}_5$ є простими тому, що мають простий порядок.

Сформулюємо тепер теорему Жордана – Гьольдера.

Теорема 99. *Будь-які два композиційні (головні) ряди групи G ізоморфні.*

Доведення. Нехай $\{H_i\}$ є композиційними (головним) рядом G. Спочатку покажемо, що $\{H_i\}$ не має нетривіальних ущільнень. Тобто для кожного i не існує такої підгрупи H_i', що

$$H_i \lhd H_i' \lhd H_{i+1}$$

Міркуємо від супротивного. Нехай така підгрупа H_i' знайдеться. Розгляднемо канонічний гомоморфізм $\gamma_{H_i} : H_{i+1} \mapsto H_{i+1}/H_i$. Згідно з частиною (ii) теореми 79 $\gamma_{H_i}(H_i')$ є нормальною підгрупою в H_{i+1}/H_i. Оскільки $H_i' \neq H_{i+1}$, то $\gamma_{H_i}(H_i') \neq H_{i+1}/H_i$. З іншого боку, $H_i' \neq H_i$, тому $\gamma_{H_i}(H_i') \neq H_i$. Отже, $\gamma_{H_i}(H_i')$ є невласною нормальною підгрупою H_{i+1}/H_i, що суперечить припущенню про простоту H_{i+1}/H_i.

Тепер твердження теореми Жордана – Гьольдера випливає безпосередньо з теореми Шраєра. Нехай $\{K_j\}$ є іншим композиційним (головним) рядом G. Він також не має нетривіальних ущільнень. Проте, згідно з теоремою Шраєра, $\{H_i\}$ та $\{K_j\}$ повинні мати ізоморфні ущільнення, а тому самі мають бути ізоморфними. Доведення завершено. □

Для скінченних груп теорема Жордана – Гьольдера є аналогом основної теореми арифметики, що стверджує єдиність розкладу натруального числа на прості

множники. Аналогічно до цієї теореми, теорема Жордана – Гьольдера стверджує, що кожна скінченна група може бути єдиним, з точністю до ізоморфізму, чином представлена у вигляді простих «множників» – простих фактор-груп.

Означення 100. Група G називається **розв'язною**, якщо вона має композиційний ряд $\{H_i\}$ в якому всі фактор-групи H_{i+1}/H_i абелеві.

Приклад 101. Група $\mathfrak{S}_3$ є розв'язною, оскільки

$$\{e\} \triangleleft A_3 \triangleleft \mathfrak{S}_3$$

є композиційним рядом та $A_3/\{e\} \cong \mathbb{Z}_3$, $\mathfrak{S}_3/A_3 \cong \mathbb{Z}_2$ є абелевими групами.

Лекція 4

Кільця та області цілісності

4.1 Означення та приклади кілець.

До цього моменту ми вивчали лише алгебраїчні структури з однієї бінарною операцією – групи. Приклади числових груп, які нам зустрічались: $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ та $(\mathbb{C}, +)$, наводять на думку про доцільність вивчення множин з декількома бінарними операціями. Найбільш загальною такою структурою є **кільце**.

Означення 102. Кільцем $(R, +, *)$ називається множина з двома бінарними операціями: додавання $+$ та множення $*$ такими, що

$(\mathcal{R}_1)$ $(R, +)$ є абелевою групою;

$(\mathcal{R}_2)$ бінарна операція $*$ є асоціативною;

$(\mathcal{R}_3)$ для довільних $a, b, c \in R$ виконуються дві властивості:

$$a * (b + c) = (a * b) + (a * c) \text{ (ліва дистрибутивність)},$$
$$(a + b) * c = (a * c) + (b * c) \text{ (права дистрибутивність)}.$$

У подальшому ми будемо писати ab замість $a * b$, так само як в теорії груп. Також ми будемо вважати, що операція множення виконується перед операцією додавання, що дозволить не писати зайвих дужок. Наприклад, ліва дистрибутивність запишеться так $a * (b + c) = a * b + a * c$. Як і раніше, ми іноді називатимемо кільцем саму множину R, якщо бінарні операції зрозумілі з контексту. Одиницю адитивної групи $(R, +)$ будемо позначати 0. Також ми будемо писати

$$n \cdot a = a + a + \cdots + a,$$

якщо $n \in \mathbb{N}$, та

$$n \cdot a = (-a) + (-a) + \cdots + (-a),$$

якщо $-n \in \mathbb{N}$. Нарешті, покладемо $0 \cdot a = 0$.

Приклад 103. Такі об'єкти є кільцями:

- $(\mathbb{Z}, +, *)$, $(\mathbb{Q}, +, *)$, $(\mathbb{R}, +, *)$ та $(\mathbb{C}, +, *)$;
- $\mathbb{Z}_n$ зі стандартним додаванням та множенням чисел по модулю n;
- $(n\mathbb{Z}, +, *)$

Приклад 104. Нехай R є кільцем. **Многочленом** над R називається вираз $a_0 + a_1x + \cdots + a_nx^n$, де $n \in \mathbb{N}_0$, $a_0, a_1, \ldots, a_n \in R$, а x позначає змінну, що набуває значень в R, або деякому кільці R_1, що містить R. Множину всіх таких многочленів позначимо $R[x]$ та введемо на цій множині операції поточкового додавання та множення многочленів. Множина $R[x]$ з такими операціями буде кільцем.

Теорема 105. *Нехай R є кільцем з адитивною одиницею 0. Для довільних $a, b \in R$ маємо:*

(i) $0a = a0 = 0$*;*

(ii) $a(-b) = (-a)b = -(ab)$*;*

(iii) $(-a)(-b) = ab$*.*

Доведення. Частина (i) випливає з рівностей

$$0a = (0 + 0)a = 0a + 0a \quad \text{та} \quad a0 = a(0 + 0) = a0 + a0$$

та правил скорочення в адитивній групі $(R, +)$. Для доведення частини (ii) запишемо

$$a(-b) + ab = a(-b + b) = a0 = 0 \quad \text{та} \quad (-a)b + ab = (-a + a)b = 0b = 0.$$

Це показує, що $a(-b)$ та $(-a)b$ є оберненими до ab елементами групи $(R, +)$. Частина (iii) випливає з рівностей

$$(-a)(-b) = -(a(-b)) = -(-(ab)) = ab.$$

Доведення завершено. □

Означення 106. Кільце $(R, +, *)$ з комутативним множенням $*$ називається **комутативним кільцем**. Кільце, в якому є мультиплікативна одиниця 1 така, що $1x = x1 = x$ для всіх $x \in R$, називається **кільцем з одиницею**.

Множина ненульових елементів R^* кільця з одиницею R, якщо вона замкнена відносно множення, утворює групу, якщо кожен елемент має обернений. Оберненим по множенню елементом до елемента $a \in R$, де R – кільце з одиницею 1, називається такий елемент $a^{-1} \in R$, що $aa^{-1} = a^{-1}a = 1$.

Означення 107. Елемент $a \in R$ називається **дільником одиниці**, якщо він має обернений елемент по множенню. Якщо множина дільників одиниці співпадає з R^*, множиною ненульових елементів R, то R називається **тілом**. Комутативне тіло називається **полем**.

Приклад 108 (Тіло кватерніонів). Класичний приклад структури, яка є тілом і не є полем, утворюють **кватерніони**. Нехай $\mathcal{Q} := \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$. Пара $(\mathcal{Q}, +)$, де $+$ є покоординатним додаванням, утворює абелеву групу. Введемо спеціальні позначення для одиничних векторів $\mathcal{Q}$:

$$1 := (1,0,0,0), \quad i := (0,1,0,0), \quad j := (0,0,1,0), \quad k := (0,0,0,1),$$

а також покладемо

$$a_1 := (a_1,0,0,0), \quad a_2 i := (0,a_2,0,0), \quad a_3 j := (0,0,a_3,0), \quad a_4 k := (0,0,0,a_4).$$

У введених позначеннях маємо

$$(a_1,a_2,a_3,a_4) = a_1 + a_2 i + a_3 j + a_4 k, \quad a_1,a_2,a_3,a_4 \in \mathbb{R}.$$

Визначимо операцію множення виразів вигляду $a_1 + a_2 i + a_3 j + a_4 k$, поклавши

$$1a = a1 = a, \quad a \in \mathcal{Q}, \quad i^2 = j^2 = k^2 = -1,$$

а також

$$ij = k, \quad jk = i, \quad ki = j, \quad ji = -k, \quad kj = -i, \quad ik = -j.$$

Будемо мати таке правило множення:

$$\begin{aligned}(a_1 + a_2 i + a_3 j + a_4 k)&(b_1 + b_2 i + b_3 j + b_4 k)\\ &= (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4) + (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)i\\ &= (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)j + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)k.\end{aligned}$$

Перевірка того, що множина $\mathcal{Q}$ з визначеними вище операціями, утворює тіло, є тривіальною, за винятком умови існування оберненого елемента до $a = a_1 + a_2 i + a_3 j + a_4 k$, де хоча б одна координата $a_i \neq 0$. Існування такого елемента випливає з рівності

$$(a_1 + a_2 i + a_3 j + a_4 k)(a_1 - a_2 i - a_3 j - a_4 k) = a_1^2 + a_2^2 + a_3^2 + a_4^2.$$

Поклавши

$$\bar{a} := a_1 - a_2 i - a_3 j - a_4 k, \quad |a|^2 = a_1^2 + a_2^2 + a_3^2 + a_4^2,$$

бачимо, що оберненим елементом до a буде $a^{-1} = \bar{a}/|a|^2$.

4.2 Дільники нуля. Області цілісності.

Стандартною властивістю звичайних числових систем є те, що добуток двох елементів дорівнює нулю тоді і тільки тоді, коли принаймні одне з цих чисел дорівнює нулю. Ця властивість використовується, наприклад, при розв'язуванні алгебраїчних рівнянь. Рівняння

$$x^2 - 5x + 6 = (x - 2)(x - 3) = 0$$

в множині дійсних чисел має рівно два корені $x = 2$ та $x = 3$ саме тому, що добуток $(a - 2)(a - 3)$ дорівнює нулю тоді і тільки тоді, коли $a - 2 = 0$ або $a - 3 = 0$.

Приклад 109. Безпосередньою перевіркою пересвідчуємось, що в кільці $\mathbb{Z}_{12}$ рівняння $x^2 - 5x + 6 = 0$ має чотири корені: $2, 3, 6, 11$. Оскільки в $\mathbb{Z}_{12}$ виконується $(11 - 2)(11 - 3) = (9)(8) = 0 = (6 - 2)(6 - 3) = (4)(3)$, то кажуть, що $3, 4, 8, 9$ є дільниками нуля в $\mathbb{Z}_{12}$. Можна перевірити, що $2, 6$ також є дільниками нуля в $\mathbb{Z}_{12}$.

Означення 110. Нехай $a, b \in R$ є такими, що $a \neq 0$, $b \neq 0$ та $ab = 0$. Елементи a, b називаються **дільниками нуля**. Зокрема, a називається лівим дільником нуля, а b називається правим дільником нуля.

Теорема 111. *В кільці $\mathbb{Z}_n$ дільниками нуля є в точності ті елементи, що не є взаємнопростими з n.*

Доведення. Нехай $m \in \mathbb{Z}_n$, $m \neq 0$ та НСД$(n, m) = d > 1$. Тоді

$$m\left(\frac{n}{d}\right) = n\left(\frac{m}{d}\right) = 0,$$

а тому m є дільником нуля. З іншого боку, якщо НСД$(n, m) = 1$ та $ms = 0$ для деякого $s \in \mathbb{Z}_n$, то n ділить ms, а тому ділить s. Отже, $s = 0$ і m не є дільником нуля. Доведення завершено. □

Наслідок 112. Якщо p є простим числом, то $\mathbb{Z}_p$ не містить дільників нуля.

Іншим підтвердженням важливості поняття дільників нуля є наступне твердження. Кажуть, що в кільці R виконується **правило скорочення**, якщо для довільних $a, b, c \in R$, $a \neq 0$, з рівності $ab = ac$ випливає $b = c$ та з рівності $ba = ca$ випливає $b = c$.

Твердження 113. *В кільці R виконується правило скорочення тоді і тільки тоді, коли R не має лівих та правих дільників нуля.*

Доведення. Нехай в R виконуються правила скорочення та $ab = 0$ для деяких $a, b \in R$. Якщо $a \neq 0$, то з рівності $ab = a0$ випливає $b = 0$. Якщо $b \neq 0$, то, аналогічно, $a = 0$. Отже, R не має лівих та правих дільників нуля. З іншого боку, якщо R не має дільників нуля та $ab = ac$, $a \neq 0$, то $a(b - c) = 0$, а отже $b = c$. Аналогічно перевіряється, що з $ba = ca$ та $a \neq 0$ випливає $b = c$. Доведення завершено. □

Означення 114. Областю цілісності називається комутативне кільце з одиницею, яке не має дільників нуля.

Теорема 115. *Кожне поле F є областю цілісності.*

Доведення. Нехай $a, b \in F$ та $a \neq 0$. Якщо $ab = 0$, то

$$b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0.$$

Отже, якщо $a \neq 0$, то $b = 0$, а отже F не має дільників нуля. Оскільки F є комутативним кільцем з одиницею, то F – область цілісності. □

Теорема 116. *Кожна скінченна область цілісності є полем.*

Доведення. Нехай $D = \{0, 1, a_1, \ldots, a_n\}$ є скінченною областю цілісності. Достатньо показати, що для кожного $a \in D$, $a \neq 0$, знайдеться $b \in D$, що $ab = 1$. Множина

$$\{a1, aa_1, \ldots aa_n\}$$

не містить 0, а також однакових елементів, оскільки це суперечить правилу скорочення в області цілісності. Отже, $\{a1, aa_1, \ldots aa_n\} = \{1, a_1, \ldots, a_n\}$. Зокрема, або $a1 = 1$, а тому $a = 1$, або $aa_i = 1$ для деякого $i = 1, \ldots, n$. Тому, a має обернений елемент в D. Доведення завершено. □

Наслідок 117. Якщо $p \in \mathbb{N}$ просте, то $\mathbb{Z}_p$ є полем.

4.3 Характеристика кільця. Теореми Ферма та Ейлера.

Означення 118. Нехай R є довільним кільцем. Мінімальне $n \in \mathbb{N}$ таке, що для довільного $a \in R$

$$n \cdot a = n + \cdots + n = 0,$$

називається **характеристикою кільця**. Якщо такого n не існує, то кажуть, що кільце має характеристику 0.

Приклад 119. Кільце $\mathbb{Z}_n$ має характеристику n. Кільця $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ мають характеристику 0.

Твердження 120. *Якщо R є кільцем з одиницею, то його характеристика n є мінімальним невід'ємним числом таким, що $n \cdot 1 = 0$.*

Доведення. Доведення випливає з рівностей

$$n \cdot a = a + a + \cdots + a = a(1 + 1 + \cdots + 1) = a(n \cdot 1) = a0 = 0$$

для довільного $a \in R$. □

Твердження 121. *Припустимо, що комутативне кільце R має характеристику p, яка є простим числом. Тоді мають місце такі рівності*

$$(ab)^{p^n} = a^{p^n} b^{p^n}, \quad (a + b)^{p^n} = a^{p^n} + b^{p^n}$$

для довільних $a, b \in R$ та $n \in \mathbb{N}$.

Доведення. Перша рівність очевидна в силу комутативності. Доведемо другу формулу індукцією по n. Для $n = 1$ маємо

$$(a+b)^p = a^p + p \cdot a^{p-1}b + \binom{p}{2} \cdot a^{p-2}b^2 + \cdots + b^p.$$

Отже, достатньо показати, що $\binom{p}{k}$ ділиться на p для довільного простого p та $1 \leq k < p$. Це випливає з представлення

$$\binom{p}{k} = \frac{p!}{k!(p-k)!},$$

в якому чисельник містить p в розкладі на прості множники, а знаменник – ні. Оскільки $\binom{p}{k}$ є цілим числом, то воно має ділитись на p. Отже, твердження вірне для $n = 1$. Крок індукції випливає з рівностей

$$(a+b)^{p^n} = ((a+b)^p)^{p^{n-1}} = (a^p + b^p)^{p^{n-1}} = (a^p)^{p^{n-1}} + (b^p)^{p^{n-1}} = a^{p^n} + b^{p^n},$$

де припущення індукції було використано в передостанній рівності. □

Наведемо деякі несподівані застосування вже розвиненої теорії кілець до теорії чисел.

Теорема 122 (Мала теорема Ферма[1])**.** *Нехай $a \in \mathbb{Z}$, а p є простим числом, що не ділить a, тоді $a^{p-1} \equiv 1 (\mathrm{mod}\, p)$.*

Доведення. Не зменшуючи загальності, можна вважати, що $a \in \mathbb{Z}_p$, $a \neq 1$. Множина ненульових елементів довільного поля утворює групу відносно операції множення. Зокрема, множина елементів $\{1, 2, \ldots, p-1\} \subset \mathbb{Z}_p$ є групою порядку $p-1$ відносно множення за модулем p. Оскільки за наслідком 54 з теореми Лагранжа порядок елемента $a \in \mathbb{Z}_p$ ділить $p-1$, то $a^{p-1} = 1$ в $\mathbb{Z}_p$. □

Наслідок 123. Для довільного $a \in \mathbb{Z}$ та простого p, $a^p \equiv a (\mathrm{mod}\, p)$.

Приклад 124. Обчислимо залишок від ділення 8^{103} на 13. Згідно з малою теоремою Ферма маємо

$$8^{103} \equiv (8^{12})^6 (8^7) \equiv 8^7 \equiv (-5)^7 \equiv 25^3(-5) \equiv (-1)^3(-5) \equiv 5 (\mathrm{mod}\, 13).$$

[1] П'єр Ферма , 1601 – 1665, французький математик.

Означення 125. Фі-функція Ейлера $\phi : \mathbb{N} \mapsto \mathbb{N}$ визначається так: для $n \in \mathbb{N}$ число $\phi(n)$ є кількістю чисел взаємнопростих з n, що не перевищують n.

Має місце таке узагальнення малої теореми Ферма.

Теорема 126 (Теорема Ейлера[2])**.** *Якщо* $a \in \mathbb{Z}$ *є взаємнопростим з* $n \in \mathbb{N}$*, то* $a^{\phi(n)} \equiv 1 (\mathrm{mod}\, n)$.

Ця теорема узагальнює малу теорему Ферма, оскільки $\phi(p) = p - 1$ для довільного простого числа p.

Для доведення теореми Ейлера нам знадобиться допоміжне твердження.

Твердження 127. *Множина* G_n *ненульових елементів* $\mathbb{Z}_n$*, які не є дільниками нуля, утворює групу відносно множення за модулем* n.

Доведення. Покажемо, що G_n замкнена відносно множення. Міркуємо від супротивного нехай $a, b \in G_n$, $ab \notin G_n$. Тоді знайдеться $c \neq 0$, що $(ab)c = 0$. Тоді $a(bc) = 0$, а тому $bc = 0$. Звідки $c = 0$, що дає суперечність. Підкреслимо, що це доведення працює для довільного кільця. Оскільки множення за модулем n асоціативне та $1 \in G_n$, залишається перевірити, що для довільного $a \in G_n$ знайдеться $b \in G_n$ таке, що $ab = 1$. Нехай $G_n = \{1, a_1, \ldots, a_r\}$. Елементи

$$a, aa_1, \ldots, aa_r$$

є різними. Дійсно, якщо $aa_i \neq aa_j$, то $a(a_i - a_j) = 0$. Тому $a_i - a_j = 0$. Отже, або $a1 = 1$, або $aa_j = 1$ для деякого j. Доведення завершено. □

Тепер доведемо теорему Ейлера.

Доведення теореми 126. Не зменшуючи загальності можемо вважати, що $1 \leq a < n$ та a взаємнопросте з n. Оскільки порядок групи G_n з попереднього твердження дорівнює $\phi(n)$, див. теорему 111, то $a^{\phi(n)} = 1$ в $\mathbb{Z}_n$. Доведення завершено. □

4.4 Поле часток.

Як ми побачили в попередньому розділі, будь-яке поле є областю цілісності, а будь-яка *скінченна* область цілісності є полем. Для нескінченних областей цілісності це твердження вже не виконується. Наприклад, нескінченне кільце $\mathbb{Z}$ є областю

[2] Леонард Ейлер, 1707 – 1783, швейцарський математик.

цілісності, але не є полем. Проте, існує загальна конструкція, яка дозволяє за даною областю цілісноті будувати найменшне поле, що її містить. Ця конструкція аналогічна побудові поля раціональних чисел $\mathbb{Q}$ з кільця $\mathbb{Z}$.

Нехай D – область цілісності, яку ми хочемо розширити до поля F, яке ми назвемо полем часток. Ми побудуємо поле F в чотири кроки:

- визначимо об'єкти, з яких складається F;
- визначимо операції додавання та множення елементів F;
- покажемо, що F з введеними опреаціями дійсно є полем;
- покажемо, що D ізоморфне деякому підкільцю F.

Крок 1. Розглянемо множину $S := D \times (D \setminus \{0\}) = \{(a, b) : a, b \in D, b \neq 0\}$ та введемо на множині S відношення еквівалентності $\sim$, поклавши $(a, b) \sim (c, d)$ тоді і тільки тоді, коли $ad = bc$. Зауважимо аналогію з раціональними числами: якщо $D = \mathbb{Z}$, то S є множиною всіх раціональних дробів, при цьому два дроби $(m_1, n_1) = \frac{m_1}{n_1}$ та $(m_2, n_2) = \frac{m_2}{n_2}$ еквіваленті, якщо вони представляють одне й те саме раціональне число, тобто $m_1 n_2 = m_2 n_1$. Покажемо, що $\sim$ є відношенням еквівалентності.

Рефлексивність. $(a, b) \sim (a, b)$, оскільки $ab = ba$ в силу комутативності множення в області цілісності D.

Симетричність. Нехай $(a, b) \sim (c, d)$, тоді $ad = bc$. Звідки в силу комутативності множення $cb = da$, а це означає $(c, d) \sim (a, b)$.

Транзитивність. Нехай $(a, b) \sim (c, d)$ та $(c, d) \sim (r, s)$, тоді $ad = bc$ та $cs = dr$. Маємо

$$asd = sad = sbc = bcs = bdr = brd.$$

Оскільки $d \neq 0$, а D – область цілісності, то з $asd = brd$ випливає $as = br$, тобто $(a, b) \sim (r, s)$.

Таким чином, множина S розбивається на класи еквівалентності за відношенням еквівалентності $\sim$. Визначимо множину F як сукупність відповідних класів еквівалентності. Будемо позначати клас еквівалентності (a, b) через $[(a, b)]$.

Крок 2. Визначимо додавання та множення класів еквівалентності так:

$$[(a,b)] + [(c,d)] = [(ad+bc, bd)]$$

та

$$[(a,b)][(c,d)] = [(ac, bd)].$$

Зауважимо, що так введені операції є аналогами додавання та множення раціональних дробів.

Помітимо, що $b \neq 0$ та $d \neq 0$. Оскільки D – область цілісності, то $bd \neq 0$. Це означає, що праві частини є елементами F. Залишається показати, що введені операції не залежать від вибору представників класів еквівалентності. Нехай $(a_1, b_1) \in [(a,b)]$ та $(c_1, d_1) \in [(c,d)]$. Покажемо, що

$$(a_1d_1 + b_1c_1, b_1d_1) \in [(ad+bc, bd)] \tag{4.1}$$

та

$$(a_1c_1, b_1d_1) \in [(ac, bd)]. \tag{4.2}$$

Маємо $(a_1, b_1) \in [(a,b)]$ та $(c_1, d_1) \in [(c,d)]$, тому $a_1b = b_1a$ та $c_1d = d_1c$. Домножимо першу рівність на d_1d, а другу на b_1b та додамо їх, отримаємо

$$a_1bd_1d + c_1db_1b = b_1ad_1d + d_1cb_1b,$$

звідки

$$(a_1d_1 + b_1c_1)bd = b_1d_1(ad+bc).$$

Отже,

$$(a_1d_1 + b_1c_1, b_1d_1) \sim (ad+bc, bd),$$

а тому (4.1) виконується. Для доведення (4.2) перемножимо рівності $a_1b = b_1a$ та $c_1d = d_1c$, отримаємо

$$a_1bc_1d = b_1ad_1c$$

та

$$a_1c_1bd = b_1d_1ac,$$

що означає

$$(a_1c_1, b_1d_1) \sim (ac, bd).$$

Отже, (4.2) виконується.

Крок 3. Наведені нижче твердження залишимо читачу для перевірки:

- додавання в F комутативне;
- додавання в F асоціативне;
- елемент $[(0,1)]$ є адитивною одиницею F;
- елемент $[(-a,b)]$ є оберненим по додаванню до елемента $[(a,b)] \in F$;
- множення в F комутативне;
- множення в F асоціативне;
- в F виконується закон дистрибутивності;
- елемент $[(1,1)]$ є мультиплікативною одиницею F;
- якщо $[(a,b)] \in F$ не є адитивною одиницею F, то $a \neq 0$ і $[(b,a)] \in F$ є оберненим по множенню до $[(a,b)]$.

Крок 4. Залишається показати, що D ізоморфне деякому підкільцю F. Розглянемо відображення $\mathcal{I} : D \mapsto F$, що визначене рівністю $\mathcal{I}(a) = [(a,1)]$ і покажемо, що воно є ізоморфізмом між кільцем D та підкільцем поля часток F, що складається з класів еквівалентності вигляду $[(a,1)]$.

Маємо

$$\mathcal{I}(a+b) = [(a+b,1)] = [(a,1)] + [(b,1)] = \mathcal{I}(a) + \mathcal{I}(b)$$

та

$$\mathcal{I}(ab) = [(ab,1)] = [(a,1)][(b,1)] = \mathcal{I}(a)\mathcal{I}(b).$$

Залишається показати, що $\mathcal{I}$ є бієкцією. Сюр'єктивність очевидна. Нехай $\mathcal{I}(a) = \mathcal{I}(b)$, тоді $[(a,1)] = [(b,1)]$, а тому $(a,1) \sim (b,1)$, звідки $a1 = 1b$ та $a = b$. Це доводить ін'єктивність.

Сформулюємо доведене твердження у вигляді теореми.

Теорема 128. *Довільна область цілісності D може бути розширена до поля F так, що елементи F можна представити у вигляді частки двох елементів D. Таке поле F називається* **полем часток**.

4.5 Ідеали та фактор-кільця.

У цьому підрозділі ми побудуємо аналоги в теорії кілець понять нормальних підгруп та фактор-груп в теорії груп, що обговорювались в лекції 2.

Нехай R є кільцем. Ми шукаємо диз'юнктні розбиття R на такі класи, що на їх множині можна визначити операції додавання та множення, які перетворять сукупність цих класів на кільце. Нагадаємо, що в теорії груп аналогічними класами були класи суміжності за нормальними підгрупами.

Твердження 129 (Аналог твердження 45)**.** *Якщо кільце R можна розбити на класи так, що їх додавання та множення коректно визначені та сукупність цих класів утворює кільце, то клас, який містить адитивну одиницю кільця R, повинен бути адитивною підгрупою N групи $(R,+)$. Понад це, N має задовольняти таку властивість: для довільних $r \in R$ та $n \in N$, добутки rn та nr мають лежати в N, тобто $rN, Nr \subset N$.*

Доведення. З результатів підрозділу 2.3 випливає, що N має бути нормальною підгрупою $(R,+)$. Оскільки $(R,+)$ є абелевою, то всі підгрупи нормальні, отже, N має бути нормальною підгрупою. Залишається показати, що $rN \subset N$ та $Nr \subset N$ для кожного $r \in R$. Зафіксуємо довільний елемент $r \in R$. Елемент r належить деякому класу A і за нашим припущенням добуток класів AN та NA на залежить від вибору представників. Зокрема, AN та NA містять $r0 = 0r = 0$, а тому співпадають з N, отже $AN = NA = N$. Таким чином, для довільних $r \in R$ та $n \in N$, $rn, nr \in N$. Доведення завершено. □

Безпосередньо з твердження 47 випливає такий результат.

Твердження 130 (Аналог твердження 47)**.** *Якщо кільце R можна розбити на класи так, що їх додавання та множення коректно визначені та сукупність цих класів утворює кільце, то класи повинні бути лівими (і одночано правими) класами суміжності по адитивній підгрупі $(N,+)$ адитивної групи $(R,+)$, де N є класом, що містить 0.*

Лема 131 (Аналог леми 57)**.** *Якщо $(N,+)$ є адитивною підгрупою групи $(R,+)$ кільця R, а відповідні операції додавання та множення класів суміжності $r+N$, $r \in R$, є коректно визначеними, то набір таких класів суміжності є кільцем.*

Доведення. Ми вже знаємо, що набір класів суміжності є абелевою групою (див. лему 57). Залишається перевірити, що множення класів суміжності є асоціативним та виконуються закони дистрибутивності. Ліва дистрибутивність випливає з рівностей

$$(r_1+N)\left((r_2+N)+(r_3+N)\right)=(r_1+N)(r_2+r_3+N)$$
$$=r_1(r_2+r_3)+N=r_1r_2+r_1r_3+N=(r_1r_2+N)+(r_1r_3+N).$$

Асоціативність та права дистрибутивність перевіряється аналогічно. □

Теорема 132 (Аналог леми 58)**.** *Якщо $(N,+)$ є адитивною підгрупою групи $(R,+)$ кільця R, то відповідні операції додавання та множення класів суміжності $r+N$, $r\in R$, є коректно визначеними тоді і тільки тоді, коли $rN\subset N$ та $Nr\subset N$ для всіх $r\in R$.*

Доведення. Пряма імплікація випливає з леми 131 та твердження 129. Доведемо протилежну імплікацію. Нехай $(N,+)$ є адитивною підгрупою групи $(R,+)$ кільця R такою, що $rN\subset N$ та $Nr\subset N$ для всіх $r\in R$. Додавання класів суміжності по підгрупі $(N,+)$ є коректно визначеним згідно з теоремою 58, що застосовується до підгрупи $(N,+)$ групи $(R,+)$. Для доведення того, що множення коректно визначене потрібно перевірити, що добуток $(r_1+N)(r_2+N)$ не залежить від вибору представників класів суміжності. Візьмемо два представники $r_1+n_1\in r_1+N$ та $r_2+n_2\in r_2+N$, $n_1,n_2\in N$ та покажемо, що $(r_1+n_1)(r_2+n_2)\in r_1r_2+N$. Маємо

$$(r_1+n_1)(r_2+n_2)=r_1r_2+n_1n_2+n_1r_2+r_1n_2.$$

З умов $rN\subset N$ та $Nr\subset N$ випливає, що $n_1n_2+n_1r_2+r_1n_2\in N$. Отже,

$$(r_1+n_1)(r_2+n_2)\in r_1r_2+N.$$

□

Ми довели, що в теорії кілець аналогами нормальних підгруп є такі адитивні підгрупи $(N,+)$ групи $(R,+)$, які задовольняють додаткову умову, що $rN\subset N$ та $Nr\subset N$ для всіх $r\in R$. Зауважимо, що з цих умов випливає, що N є підкільцем R, але не кожне підкільце їм задовольняє. Наприклад, $\mathbb{Q}$ є підкільцем $\mathbb{R}$, але $\pi\mathbb{Q}\not\subset\mathbb{Q}$.

Означення 133 (Аналог означення нормальної підгрупи)**.** Адитивна підгрупа $(N, +)$ кільця R, яка задовольняє властивість $rN \subset N$ та $Nr \subset N$ для всіх $r \in R$, називається *двостороннім ідеалом* кільця R. Односторонні ідеали визначаються аналогічно шляхом відкидання однієї з умов $rN \subset N$ або $Nr \subset N$.

Означення 134 (Аналог означення фактор-групи)**.** Якщо N є ідеалом кільця R, то кільце, що утворене множиною класів суміжності $r + N$, називається **фактор-кільцем** та позначається R/N. Класи суміжності низиваються також класами лишків за модулем N.

Приклад 135. В кільці $\mathbb{Z}$ єдиними адитивними підгрупами є $n\mathbb{Z}$. Очевидно, що $m(n\mathbb{Z}) \subset n\mathbb{Z}$ для довільного $m \in \mathbb{Z}$, тому $n\mathbb{Z}$ є ідеалом кільця $\mathbb{Z}$ для кожного n. Отже, класи суміжності $a + n\mathbb{Z}$ утворюють фактор-кільце $\mathbb{Z}/n\mathbb{Z}$.

Кожне кільце R має тривіальні ідеали R та $\{0\}$. Всі інші ідеали будемо називати **власними**. Наступна теорема та наслідок до неї вказують, що поняття ідеалу для полів не є цікавим.

Теорема 136. *Якщо R є кільцем з одиницею, а N є ідеалом в R, що містить хоча б один дільник одиниці, то $N = R$.*

Доведення. Нехай $u \in N \subset R$ є дільником одиниці, тоді $1 = u^{-1}u \in N$ в силу того, що $u^{-1}N \subset N$. З того, що $1 \in N$ випливає $N = R$. Доведення завершено. □

Наслідок 137. Оскільки в полі кожен ненульовий елемент є дільником одиниці, то кожен ідеал або не містить ненульових елементів, або співпадає з усім полем.

Лекція 5

Поля та їх розширення

5.1 Максимальні та прості ідеали.

Означення 138. Власний ідеал M кільця R називається **максимальним ідеалом**, якщо в R не існує власного ідеалу, що містить M.

Теорема 139. *Нехай R є комутативним кільцем з одиницею. Тоді M є максимальним ідеалом тоді і тільки тоді, коли R/M є полем.*

Доведення. Нехай $M \neq R$ є максимальним ідеалом, тоді R/M є, очевидно, комутативним кільцем з одиницею. Нехай $a + M \in R/M$ та $a \notin M$, тобто $a + M$ не є адитивною одиницею R/M. Ми маємо перевірити, що $a + M$ має обернений по множенню елемент в R/M. Покладемо

$$N := \{ra + m : r \in R, m \in M\}.$$

Тоді $(N, +)$ є абелевою групою. Оскільки для довільного $r_1 \in R$ маємо

$$r_1(ra + m) = (r_1 r)a + r_1 m \in N, \quad (ra + m)r_1 = (rr_1)a + mr_1 \in N,$$

то N є ідеалом в R. З того, що $M \subset N$, $a \in N$ та $a \notin M$, випливає $N = R$. Зокрема, $1 \in N$. Отже, існує $b \in R$ та $m \in M$, що $1 = ba + m$, $ba \in 1 + M$, а отже

$$(b + M)(a + M) = (a + M)(b + M) = 1 + M.$$

Доведемо обернене твердження. Нехай R/M є полем, а $\gamma_M : R \mapsto R/M$ є канонічним гомоморфізмом:

$$\gamma_M(a) = a + M, \quad a \in R.$$

Якби існував власний ідеал N такий, що $M \subset N \subset R$, то $\gamma_M(N)$ був би власним ідеалом в полі R/M. Це суперечить наслідку 137. □

Наслідок 140. Комутативне кільце з одиницею є полем тоді і тільки тоді, коли воно не містить власних ідеалів.

Доведення. Якщо комутативне кільце з одиницею є полем, то воно не містить власних ідеалів згідно з наслідком 137. В зворотній бік, якщо R не містить власних ідеалів, то $\{0\}$ є максимальним ідеалом, а отже $R \cong R/\{0\}$ є полем. □

Перейдемо тепер до питання для яких ідеалів відповідні фактор-кільця є областями цілісності.

Означення 141. Ідеал $N \neq R$ комутативного кільця R називається **простим**, якщо з $ab \in N$ випливає, що $a \in N$ або $b \in N$ для всіх $a, b \in R$.

Теорема 142. *Нехай R є комутативним кільцем з одиницею та $N \neq R$ є ідеалом. Тоді R/N є областю цілісності тоді і тільки тоді, коли N є простим ідеалом.*

Доведення. Фактор-кільце R/N є областю цілісноті тоді і тільки тоді, коли з рівностій $(a+N)(b+N) = N$ випливає, що $a+N = N$ або $b+N = N$. Це еквівалентно тому, що з $ab \in N$ випливає $a \in N$ або $n \in N$. Доведення завершено. □

Наслідок 143. В комутативному кільці з одиницею кожний максимальний ідеал є простим.

Доведення. Якщо M є максимальним ідеалом в R, то R/M є полем і, зокрема, областю цілісності. З попередньої теореми випливає, що M – простий ідеал. □

За допомогою понять максимального та простого ідеалу доведемно наступну важливу теорему, що демонструє фундаментальність полів $\mathbb{Z}_p$, де p є простим числом, та $\mathbb{Q}$.

Теорема 144. *Кожне поле F має або просту характеристику p та підполе, що ізоморфне $\mathbb{Z}_p$, або характеристику 0 та підполе, що ізоморфне $\mathbb{Q}$.*

Для доведення теореми 144 нам знадобиться допоміжне твердження.

Твердження 145. *Якщо R є комутативним кільцем з одиницею та характеристикою $n > 1$, то R містить підкільце, що ізоморфне $\mathbb{Z}_n$. Якщо R є комутативним кільцем з одиницею та характеристикою 0, то R містить підкільце, що ізоморфне $\mathbb{Z}$.*

Доведення. Відображення $\phi : \mathbb{Z} \to R$, що визначене так

$$\phi(m) = m \cdot 1,$$

є гомоморфізмом кілець. Його ядро K має бути ідеалом в $\mathbb{Z}$, а отже K має вигляд $s\mathbb{Z}$ для деякого $s \in \mathbb{Z}$. Якщо R має характеристику n, то з твердження 120 випливає, що $K = n\mathbb{Z}$. Згідно з основною теоремою про гомоморфізм

$$\mathbb{Z}_n \cong \mathbb{Z}/n\mathbb{Z} \cong \phi(\mathbb{Z}).$$

Таким чином, $\phi(\mathbb{Z})$ є підкільцем R, що ізоморфне $\mathbb{Z}_n$. Якщо R має характеристику 0, то $m \cdot 1 \neq 0$ для всіх $m \in \mathbb{Z}$. Отже, $K = \{0\}$ і $\phi(\mathbb{Z}) \cong \mathbb{Z}$. Доведення завершено. □

Доведення теореми 144. Якщо характеристика F не дорівнює нулю, то згідно з твердженням 145 F містить підкільце, ізоморфне $\mathbb{Z}_n$. Якщо n не є простим, то це підкільце (а тому й F!) має дільники нуля. Отже, n має бути простим. Якщо F має характеристику 0, то F містить підкільце, ізоморфне $\mathbb{Z}$. Мінімальне поле, що містить $\mathbb{Z}$, є $\mathbb{Q}$. Тому, F містить підполе, що ізоморфне $\mathbb{Q}$. □

Означення 146. Поля $\mathbb{Z}_p$, для простих p, та $\mathbb{Q}$ називаються **простими полями**.

5.2 Кільце многочленів.

У цьому підрозділі ми наведемо ряд найважливіших результатів про многочлени від однієї змінної з коефіцієнтами з довільного поля F. Нагадаємо, що $F[x]$ позначає множину многочленів від змінної x з коефіцієнтами з поля F. Будемо позначати $\deg f$ степінь многочлена $f(x) \in F[x]$.

Твердження 147. *Нехай*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

та

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_0$$

є елементами $F[x]$ *такими, що* $a_n \neq 0$, $b_m \neq 0$ *та* $m > 0$. *Тоді існують єдині многочлени* $q(x)$ *та* $r(x)$, *що* $f(x) = q(x)g(x) + r(x)$ *та* $\deg r < \deg g = m$. *Якщо* $r(x) = 0$, *то кажуть, що* $f(x)$ *ділиться на* $g(x)$.

Наслідок 148. Елемент $a \in F$ є нулем многочлена $f(x) \in F[x]$ тоді і тільки тоді, коли $f(x)$ ділиться на $x - a$.

Наслідок 149. Нехай $f(x) \in F[x]$ та f має степінь n. Тоді f може мати не більше ніж n нулів в F.

Означення 150. Многочлен $f(x) \in F[x]$ степеня вище одиниці називається незвідним над F, якщо з представлення $f(x) = g(x)h(x)$ випливає, що степінь $\deg g = \deg f$ або $\deg h = \deg f$.

Твердження 151. *Нехай* F *є полем. Тоді довільний многочлен* $f(x) \in F[x]$ *може бути розкладений у добуток незвідних многочленів. Цей розклад є єдиним з точністю до порядку множників та мультиплікативної константи.*

Маючи наведені вище базові факти, перейдемо тепер до вивчення структури ідеалів кільця $F[x]$.

Означення 152. Нехай R є комутативним кільцем з одиницею, $a \in R$. Множина $\langle a \rangle := \{ra : r \in R\}$ називається **головним ідеалом, породженим** a**.** Ідеал N називається **головним ідеалом**, якщо $N = \langle a \rangle$ для деякого $a \in R$.

Теорема 153. *Якщо* F *є полем, то кожен ідеал в* $F[x]$ *є головним.*

Доведення. Нехай N є ідеалом в $F[x]$. Якщо $N = \{0\}$, то $N = \langle 0 \rangle$. Припустимо, що $N \neq \{0\}$ і нехай $g(x) \in N$ є многочленом найменшого степеня в N. Якщо $\deg g = 0$, то $g(x) \in F$ є дільником одиниці в $F[x]$, а отже $N = F[x] = \langle 1 \rangle$. Нехай $\deg g > 0$. Розглянемо довільний елемент $f(x) \in N$ та запишемо $f(x) = q(x)g(x) + r(x)$. Оскільки N є ідеалом, то $r(x) = f(x) - q(x)g(x) \in N$. За визначенням $\deg r < \deg g$, а тому в силу вибору g, маємо $r(x) = 0$. Отже, $N = \langle g(x) \rangle$. □

Теорема 154. *Ідеал* $\langle p(x) \rangle \neq \{0\}$ *в* $F[x]$ *є максимальним тоді і тільки тоді, коли многочлен* $p(x)$ *є незвідним над* F.

Доведення. Нехай $\langle p(x)\rangle \neq \{0\}$ є максимальним ідеалом в $F[x]$. Тоді $\langle p(x)\rangle = F[x]$, а тому $p(x) \notin F$. Нехай $p(x) = f(x)g(x)$ є розкладом $p(x)$ в $F[x]$. Оскільки ідеал $\langle p(x)\rangle$ є максимальним, то він є простим. Отже, $f(x) \in \langle p(x)\rangle$ або $g(x) \in \langle p(x)\rangle$. Це означає, що або $f(x)$ ділиться на $p(x)$, або $g(x)$ ділиться на $p(x)$. Тому або $\deg f = \deg p$, або $\deg g = \deg p$.

Доведемо протилежне твердження. Нехай $p(x)$ є незвідним многочленом над F і нехай N є ідеалом в $F[x]$ таким, що $\langle p(x)\rangle \subset N \subset F[x]$. Згідно з теоремою 153 N є головним, тобто $N = \langle g(x)\rangle$ для деякого $g(x) \in F[x]$. Оскільки $p(x) \in N$, то $p(x) = g(x)q(x)$ для деякого $q(x) \in F[x]$. В силу припущення, що $p(x)$ незвідний, або $\deg g = 0$, або $\deg q = 0$. Якщо $\deg g = 0$, то $N = F[x]$. Якщо $\deg q = 0$, тобто $q(x) = c \in F$, то $g(x) = c^{-1}p(x)$, а тому $g(x) \in \langle p(x)\rangle$ та $N = \langle p(x)\rangle$. Доведення завершено. □

5.3 Розширення полів.

Означення 155. Поле E називається розширенням поля F, якщо F є підполем E. Баштою розширень називається довільний скінченний ланцюжок $F_1 \leq F_2 \leq \cdots \leq F_n$, в якому поле F_{m+1} є розширенням поля F_m, $1 \leq m < n$.

Ключовою теоремою теорії розширень полів є така теорема.

Теорема 156 (Теорема Кронекера[1])**.** *Нехай F є полем та $f(x) \in F[x]$ є многочленом ненульового степеня. Тоді існує розширення E поля F та $\alpha \in E$, що $f(\alpha) = 0$.*

Доведення. Згідно з твердженням 151 достатньо довести теорему для незвідного многочлена $p(x)$, що входить у розклад $f(x)$ на незвідні множники. Згідно з теоремою 154 ідеал $\langle p(x)\rangle$ є максимальним ідеалом в $F[x]$, тому $F[x]/\langle p(x)\rangle$ є полем. Покажемо, що F можна ототожнити з деяким підполем поля $F[x]/\langle p(x)\rangle$ шляхом відображення $\psi : F \mapsto F[x]/\langle p(x)\rangle$, яке визначене рівністю

$$\psi(a) = a + \langle p(x)\rangle, \quad a \in F.$$

Покажемо, що це відображення є ізоморфізмом поля F та підполя $\{a + \langle p(x)\rangle : a \in F\}$ поля $F[x]/\langle p(x)\rangle$. Сюр'єктивність очевидна. Нехай $\psi(a) = \psi(b)$, тоді

[1] Леопольд Кронекер, 1823 – 1891, німецький математик.

$a-b \in \langle p(x)\rangle$, тобто $a-b$ є кратним $p(x)$. Це можливо лише у випадку $a-b=0$. Отже, ψ є ін'єктивним. Рівності $\psi(ab)=\psi(a)\psi(b)$ та $\psi(a+b)=\psi(a)+\psi(b)$ очевидні з означення. Отже, $F[x]/\langle p(x)\rangle$ можна вважати розширенням поля F, позначимо це розширення E. Залишається показати, що існує $\alpha \in E$ таке, що $p(\alpha)=0$. Покладемо $\alpha = x+\langle p(x)\rangle$ та розглянемо відображення $\psi_\alpha : F[x] \mapsto E$, що визначене так

$$\psi_\alpha\left(a_0+a_1x+\cdots+a_nx^n\right)=a_0+a_1(x+\langle p(x)\rangle)+\cdots+a_n(x+\langle p(x)\rangle)^n, \qquad (5.1)$$

де $a_0, a_1 \ldots, a_n \in F$. Маємо

$$\psi_\alpha\left(a_0+a_1x+\cdots+a_nx^n\right)=a_0+a_1x+\cdots+a_nx^n+\langle p(x)\rangle=\langle p(x)\rangle.$$

Отже, $\psi_\alpha(p(x))=0$ в E, а тому $p(\alpha)=0$ в E. Доведення завершено. □

Означення 157. Відображення $\psi_\alpha : F[x] \mapsto E$, що визначене формулою (5.1), називається *гомоморфізмом обчислення*.

Приклад 158. Нехай $F=\mathbb{R}$ та x^2+1 є незвідним многочленом над F. Таким чином, поле $\mathbb{R}/\langle x^2+1\rangle$ є розширенням поля $\mathbb{R}$. Нехай $\alpha = x+\langle x^2+1\rangle$. Тоді

$$\alpha^2+1=(x+\langle x^2+1\rangle)^2+1+\langle x^2+1\rangle=x^2+1+\langle x^2+1\rangle=0.$$

Отже, α є нулем x^2+1 в $\mathbb{R}/\langle x^2+1\rangle$.

Приклад 159. Нехай $F=\mathbb{Q}$ та $f(x)=x^4-5x^2+6=(x^2-2)(x^2-3)$. Обидва многочлени x^2-2 та x^2-3 незвідні над $\mathbb{Q}$. Так само, як і в попередньому прикладі, ми можемо побудувати розширення E поля $\mathbb{Q}$, що містить нуль x^2-3, або розширення K поля $\mathbb{Q}$, що містить нуль x^2-2.

Означення 160. Елемент α розширення E поля F називається **алгебраїчним над** F, якщо $f(\alpha)=0$ для деякого $f(x) \in F[x]$. Якщо $\alpha \in E$ не є алгебраїчним, то α називається **трансцендентним над** F.

Приклад 161. Мають місце такі твердження:

- елементи $\sqrt{2}, \mathrm{i} \in \mathbb{C}$ є алгебраїчними над $\mathbb{Q}$, оскільки вони є нулями многочленів x^2-2 та x^2+1, відповідно;

- елементи $e, \pi \in \mathbb{R}$ є трансцендентними над $\mathbb{Q}$ (доведення не є простим!).

Нагадаємо, що многочлен $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in F[x]$ називається **нормованим**, якщо $a_n = 1$.

Теорема 162. *Нехай E є розширенням поля F та $\alpha \in E$ є алгебраїчним над F. Тоді існує єдиний нормований незвідний многочлен $p(x) \in F[x]$ такий, що $p(\alpha) = 0$. Якщо многочлен $f(x) \in F[x]$ такий, що $f(\alpha) = 0$, то $f(x)$ ділиться на $p(x)$.*

Доведення. Гомоморфізм обчислення $\psi_\alpha : F[x] \mapsto E$ має ядро K, яке є ідеалом в $F[x]$. Оскільки всі ідеали в $F[x]$ є головними, то $K = \langle p(x) \rangle$ для деякого $p(x) \in F[x]$. За визначенням, $K = \langle p(x) \rangle$ складається з тих і тільки тих многочленів з $F[x]$, що мають α нулем в E. Отже, всі елементи K діляться на $p(x)$. Залишається показати, що $p(x)$ незвідний. Якщо $p(x) = r(x)s(x)$, то $r(\alpha)s(\alpha) = 0$, тому або $r(\alpha) = 0$, або $s(\alpha) = 0$, оскільки E є полем. Отже, один з многочленів $r(x) \in K$ або $s(x) \in K$ ділиться на $p(x)$. Доведення завершено. □

Означення 163. Нехай E є розширенням поля F та $\alpha \in E$ є алгебраїчним над F. Єдиний нормований многочлен $p(x)$ в теоремі 162 називається **незвідним многочленом елемента α над F** і позначається $\operatorname{irr}(\alpha, F)$. Степінь многочлена $\operatorname{irr}(\alpha, F)$ називається **степенем алгебраїчного числа α над F**.

Приклад 164. Мають місце такі твердження:

- $\operatorname{irr}(\sqrt{2}, \mathbb{Q}) = x^2 - 2$;
- $\operatorname{irr}(\sqrt{1+\sqrt{3}}, \mathbb{Q}) = x^4 - 2x^2 - 2$;
- $\operatorname{irr}(\mathrm{i}, \mathbb{R}) = x^2 + 1$.

5.4 Прості розширення полів.

Нехай E є розширенням поля F, $\alpha \in E$ та $\psi_\alpha : F[x] \mapsto E$ є гомоморфізмом обчислення. Розглянемо два випадки.

ВИПАДОК 1. α Є АЛГЕБРАЇЧНИМ НАД F. Тоді ядром ψ_α є $\langle \operatorname{irr}(\alpha, F) \rangle$ і це ядро є максимальним ідеалом в $F[x]$. За основною теоремою про гомоморфізм, маємо

ізоморфізм полей $F[x]/\langle \mathrm{irr}\,(\alpha, F)\rangle \cong \psi_\alpha(F[x])$. Підполе $\psi_\alpha(F[x])$ є найменшим, що містить F та α. Позначимо це підполе $F(\alpha)$.

ВИПАДОК 2. α Є ТРАНСЦЕНДЕНТНИМ НАД F. Тоді $\psi_\alpha(f(x)) \neq 0$ для всіх неконстант $f(x) \in F[x]$, а отже $\ker \psi_\alpha = \{0\}$. За основною теоремою про гомоморфізм $F[x] \cong \psi_\alpha(F[x])$. Таким чином, $\psi_\alpha(F[x])$ не є полем, але є областю цілісності, яке ми позначимо $F[\alpha]$. Найменшим полем, що містить $F[\alpha]$ є поле часток цієї області цілісності. Це поле часток є найменшим полем, що містить F та α. Як і в попередньому випадку позначимо це поле $F(\alpha)$.

Приклад 165. Оскільки π є трансцендентним над $\mathbb{Q}$, то поле $\mathbb{Q}(\pi)$ ізоморфне полю $\mathbb{Q}(x)$ раціональних функції над $\mathbb{Q}$ від змінної x. Структурно це означає, що елемент, який трансцендентний над полем F, веде себе як незалежна змінна x над F.

Означення 166. Розширення E поля F називається **простим**, якщо $E = F(\alpha)$ для деякого $\alpha \in E$.

Теорема 167. *Нехай $E = F(\alpha)$ є простим розширенням поля F та α є алгебраїчним над F степеня $n \geq 1$. Тоді кожен елемент $\beta \in F(\alpha)$ можна єдиним чином подати у вигляді*

$$\beta = b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1}, \tag{5.2}$$

де $b_i \in F$.

Доведення. Як ми знаємо, кожний елемент поля $F(\alpha) = \psi_\alpha(F[x])$ має вигляд $\psi_\alpha(f(x)) = f(\alpha)$, тобто є многочленом від формальної змінної α з коефіцієнтами з поля F. Нехай

$$\mathrm{irr}\,(\alpha, F) = p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0.$$

Тоді $p(\alpha) = 0$, а отже

$$\alpha^n = -a_{n-1}\alpha^{n-1} - \cdots - a_0.$$

З цієї рівності випливає, що довільний многочлен від змінної α з коефіцієнтами з F можна подати у вигляді многочлена від змінної α степеня, що не перевищує $n-1$. Зокрема, кожний елемент $\beta \in F(\alpha)$ можна подати у вигляді (5.2). Для доведення

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

та

$*$	0	1	α	$1+\alpha$
0	0	0	0	0
1	0	1	α	$1+\alpha$
α	0	α	$1+\alpha$	1
$1+\alpha$	0	$1+\alpha$	1	α

Табл. 5.1: Таблиці додавання та множення поля з чотирьох елементів.

єдиності припустимо, що

$$b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1} = b'_0 + b'_1\alpha + \cdots + b'_{n-1}\alpha^{n-1}$$

для деяких $b_i, b'_i \in F$. Тоді многочлен

$$g(x) = (b'_0 - b_0) + (b'_1 - b_1)x + \cdots + (b'_{n-1} - b_{n-1})x^{n-1}$$

має α коренем. Оскільки його степінь $n-1$ менша за степінь $\operatorname{irr}(\alpha, F)$, то $g(x) \equiv 0$. Отже, $b_i = b'_i$. Доведення завершено. □

Приклад 168 (Побудова поля з чотирьох елементів). Многочлен $x^2+x+1 \in \mathbb{Z}_2[x]$ є незвідним над $\mathbb{Z}_2$. Згідно з теоремою 156 існує розширення E поля $\mathbb{Z}_2$, що містить нуль α многочлена $x^2 + x + 1$. Згідно з теоремою 167 поле $\mathbb{Z}_2(\alpha)$ складається з елементів

$$0 + 0\alpha = 0, 1 + 0\alpha = 1, 0 + 1\alpha = \alpha, 1 + 1\alpha = 1 + \alpha.$$

Таблиці множення і додавання представлені в таблиці 5.1. Дії виконуються з врахуванням співвідношення $1 + \alpha + \alpha^2 = 0$. Наприклад,

$$(1 + \alpha)(1 + \alpha) = 1 + \alpha + \alpha + \alpha^2 = \alpha + (1 + \alpha + \alpha^2) = \alpha.$$

Приклад 169. Многочлен $x^2 + 1 \in \mathbb{R}[x]$ є незвідним над $\mathbb{R}$. Поле $\mathbb{R}[x]/\langle x^2 + 1\rangle$ є розширенням $\mathbb{R}$. Нехай

$$\alpha = x + \langle x^2 + 1\rangle.$$

Тоді $\mathbb{R}(\alpha) = \mathbb{R}[x]/\langle x^2 + 1\rangle$ складається з елементів вигляду $a + b\alpha$, $a, b \in \mathbb{R}$. Перейменовуючи α на i, отримаємо класичне визначення поля $\mathbb{C}$ та $\mathbb{R}(\alpha) \cong \mathbb{C}$.

Теорема 170. *Нехай E є розширенням поля F та $\alpha \in E$ є алгебраїчним над F. Якщо $\deg \operatorname{irr}(\alpha, F) = n$, то $F(\alpha)$ є n-вимірним векторним простором над полем F з базисом $\{1, \alpha, \ldots, \alpha^{n-1}\}$. Кожен елемент $\beta \in F(\alpha)$ є алгебраїчним над F та $\deg \operatorname{irr}(\beta, F) \leq \deg \operatorname{irr}(\alpha, F)$.*

Доведення. Факт, що система $\{1, \alpha, \ldots, \alpha^{n-1}\}$ є повною[2] випливає з теореми 167. Припустимо, що ця система є лінійно залежною, тоді $b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1} = 0$ для деяких $b_i \in F$, що не дорівнюють нулю одночасно. Тобто α є коренем многочлена степеня $n-1$. Це суперечить припущенню, що $\deg \operatorname{irr}(\alpha, F) = n$. Таким чином, перша частина теореми доведена. Доведемо другу частину. Нехай $\beta \in F(\alpha)$. Система $\{1, \beta, \ldots, \beta^n\}$ складається з $n+1$ елемента, а тому є лінійно залежною над F. Отже, знайдуться $b_i' \in F$ такі, що

$$b_0' + b_1'\beta + \cdots + b_n'\beta^n = 0,$$

та не всі $b_i' \in F$ одночасно дорівнюють нулю. Це означає, що β є алгебраїчним над F та $\deg \operatorname{irr}(\beta, F) \leq n$. □

5.5 Скінченні розширення полів.

Означення 171. Розширення E поля F називається **алгебраїчним розширенням**, якщо кожний елемент E є алгебраїчним над F. Якщо розширення E поля F як векторний простір над F має скінченну розмірність n, то E називається **скінченним розширенням степеня** n поля F. Степінь n скінченного розширення будемо позначати $[E : F]$.

З теореми 170 випливає таке твердження.

Твердження 172. *Кожне скінченне розширення E поля F є алгебраїчним.*

Наступна теорема є аналогом теореми Лагранжа в теорії груп.

Теорема 173. *Якщо E є скінченним розширенням F, K є скінченнми розширенням E, то K є скінченним розширенням F, та*

$$[K : F] = [K : E][E : F].$$

Доведення. Нехай $(\alpha_i)_{i=1,\ldots,n}$ є базисом E над F, $(\beta_j)_{j=1,\ldots,m}$ є базисом K над E. Для доведення теореми достатньо показати, що $(\alpha_i\beta_j)_{i=1,\ldots,n,j=1,\ldots,m}$ буде базисом K над F. Нехай $\gamma \in K$. Можемо записати

$$\gamma = \sum_{j=1}^{m} b_j\beta_j$$

[2]Нагадаємо, це означає, що будь-який елемент $F(\alpha)$ лінійно виражається через $\{1, \alpha, \ldots, \alpha^{n-1}\}$.

для деяких $b_j \in E$. З іншого боку, для кожного $j = 1, \ldots, m$

$$b_j = \sum_{i=1}^{n} a_{ij}\alpha_i$$

для деяких $a_{ij} \in F$. Отже,

$$\gamma = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij}\alpha_i\beta_j,$$

а тому система $(\alpha_i\beta_j)_{i=1,\ldots,n,j=1,\ldots,m}$ є повною. Залишається показати, що вона лінійно незалежна над F. Нехай $\sum_{i,j} c_{ij}\alpha_i\beta_j = 0$ для деяких $c_{ij} \in F$. З рівності

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n} c_{ij}\alpha_i\right)\beta_j = 0$$

та лінійної незалежності $(\beta_j)_{j=1,\ldots,m}$ над F випливає, що

$$\sum_{i=1}^{n} c_{ij}\alpha_i = 0,$$

для всіх $j = 1, \ldots, m$. З лінійної незалежності $(\alpha_i)_{i=1,\ldots,n}$ над F випливає, що $c_{ij} = 0$. Доведення завершено. □

За індукцією маємо такий наслідок до попередньої теореми.

Наслідок 174. Якщо F_i є полем для кожного $i = 1, \ldots, r$, та F_{i+1} є скінченним розширенням F_i для всіх $i = 1, \ldots, r-1$, то F_r є скінченним розширенням F_1 та

$$[F_r : F_1] = [F_r : F_{r-1}][F_{r-1} : F_{r-2}] \cdots [F_2, F_1].$$

Нагадаємо, що $F(\alpha)$ позначає мінімальне розширення поля F, яке містить елемент $\alpha \in E$.

Означення 175. Нехай E є розширенням поля F та $\alpha_1, \alpha_2, \ldots, \alpha_n \in E$. Поле $F(\alpha_1, \alpha_2, \ldots, \alpha_n)$ – це мінімальне розширення F, що містить $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Теорема 176. *Нехай E є алгебраїчним розширенням поля F. Рівність $E = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$ для деякого скінченного набору $\alpha_1, \alpha_2, \ldots, \alpha_n \in E$ виконується тоді і тільки тоді, коли E є скінченним розширенням F.*

Доведення. Нехай $E = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Оскільки E є алгебраїчним розширенням поля F, то кожне α_i є алгебраїчним над F, а отже й над кожним розширенням F. Зокрема, α_j є алгебраїчним над $F(\alpha_1, \ldots, \alpha_{j-1})$ для всіх $j = 2, \ldots, n$. Отже, з наслідку 174, який застосований до послідовності полів

$$F, F(\alpha_1), F(\alpha_1, \alpha_2), \ldots, F(\alpha_1, \alpha_2, \ldots, \alpha_n) = E,$$

випливає, що E є скінченним розширенням F.

Доведемо протилежне твердження. Нехай E є скінченним алгебраїчним розширенням F. Якщо $[E : F] = 1$, то $E = F(1) = F$. Нехай $E \neq F$ та $\alpha_1 \in E \setminus F$ Тоді $[F(\alpha_1) : F] > 1$. Якщо $F(\alpha_1) = E$, то маємо необхідне твердження. В протилежному випадку візьмемо $\alpha_2 \in E \setminus F(\alpha_1)$ та розглянемо $F(\alpha_1, \alpha_2)$ і т.д. Оскільки $[E : F]$ скінченне, то цей процес завершиться за скінченну кількість кроків і ми отримаємо

$$E = F(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

□

5.6 Алгебраїчні замикання та алгебраїчно замкнені поля.

Теорема 177. *Нехай E є розширенням F. Множина*

$$\overline{F}_E := \{\alpha \in E : \alpha \text{ є алгебраїчним над } F\}$$

є підполем E. Воно називається **алгебраїчним замиканням поля F в E.**

Доведення. Нехай $\alpha, \beta \in \overline{F}_E$. Згідно з теоремою 176 $F(\alpha, \beta)$ є скінченним розширенням F, а за твердженням 172 $F(\alpha, \beta)$ є алгебраїчним. Тому $F(\alpha, \beta) \subset \overline{F}_E$, зокрема $\alpha \pm \beta$, $\alpha\beta$ та α/β ($\beta \neq 0$) належать $\overline{F}_E$. Отже, $\overline{F}_E$ є підполем E. □

Означення 178. Поле F називається **алгебраїчно замкненим**, якщо кожний многочлен ненульового степеня в $F[x]$ має корінь в F.

З наслідку 148 випливає, що поле є алгебраїчно замкненим тоді і тільки тоді, коли кожен многочлен ненульового степеня в $F[x]$ може бути розкладений в добуток лінійних множників.

Твердження 179. *Алгебраїчно замкнене поле F не має нетривіальних алгебраїчних розширень.*

Доведення. Нехай E є алгебраїчним розширенням F та $\alpha \in E$. Тоді $\operatorname{irr}(\alpha, F) = x - \alpha$, оскільки F є алгебраїчно замкненим. Отже, $\alpha \in F$ та $E = F$. □

Використовуючи *лему Цорна* можна довести таку теорему. Це доведення можна знайти в розділі 38.3 книги [6].

Теорема 180. *Кожне поле F має алгебраїчне замикання, тобто алгебраїчне розширення $\overline{F}$, яке є алгебраїчно замкненим.*

5.7 Структура скінченних полів.

Метою цього розділу є доведення факту, що кожне скінченне поле містить p^n елементів, де p – просте число, а $n \in \mathbb{N}$, а також оберненого твердження, що для кожного простого p та натурального $n \in \mathbb{N}$ існує поле, що містить p^n елементів. Нагадаємо, що в прикладі 168 було побудовано поле з $2^2 = 4$ елементів.

Легко бачити, що кожне скінченне поле повинно містити p^n елементів.

Теорема 181. *Нехай E є скінченним розширенням степеня n скінченного поля F. Якщо F містить q елементів, то E складається з q^n елементів.*

Доведення. Нехай $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ є базисом E над F. Тоді кожний $\beta \in E$ може бути єдиним чином записаний у вигляді

$$\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n$$

для деяких $b_i \in F$. Оскільки кожний коефіцієнт може приймати q різних значень, то всього існує q^n різних лінійних комбінацій елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$. □

З теореми 144 випливає, що кожне скінченне поле повинно мати просту характеристику p.

Наслідок 182. Якщо E є скінченним полем простої характеристики p, то E містить рівно p^n елементів для деякого $n \in \mathbb{N}$.

Доведення. З теореми 144 випливає, що E є скінченним розширенням простого поля $\mathbb{Z}_p$. Наслідок випливає з теореми 181. □

Означення 183. Нехай F є полем з алгебраїчним замиканням $\overline{F}$, $f(x) \in F[x]$. Підполе E поля $\overline{F}$ називається **полем розкладу** $f(x)$ **над** F, якщо E є мінімальним підполем $\overline{F}$, що містить F та всі нулі $f(x)$ в $\overline{F}$

Теорема 184. *Скінченне поле E з p^n елементів є (з точністю до ізоморфізму) полем розкладу многочлена $x^{p^n} - x$ над простим полем $\mathbb{Z}_p$.*

Доведення. Нехай E є скінченним полем з p^n елементів, де p є характеристикою E. Множина E^* ненульових елементів E утворює групу порядку $p^n - 1$ відносно операції множення в полі E. Для кожного $\alpha \in E^*$ маємо $\alpha^{p^n-1} = 1$ тому, що порядок α має ділити порядок групи E^*. Отже, $\alpha^{p^n} = \alpha$ і ми показали, що всі елементи E є нулями $x^{p^n} - x$. Оскільки цей многочлен може мати щонайбільше p^n нулів, то E повинен співпадати з полем розкладу $x^{p^n} - x$ над $\mathbb{Z}_p$. □

Для досягнення нашої мети залишається показати, що $x^{p^n} - x$ має p^n різних нулів в його полі розкладу над скінченним полем F характеристики p.

Лема 185. *Нехай F є скінченним полем характеристики p. Тоді $x^{p^n} - x$ має p^n різних нулів у полі розкладу $K < \overline{F}$ многочлена $x^{p^n} - x$ над F.*

Доведення. Очевидно, що 0 є нулем $x^{p^n} - x$ кратності 1. Нехай $\alpha \neq 0$ є нулем $x^{p^n} - x$, а отже, нулем $f(x) = x^{p^n-1} - 1$. Поділимо, $f(x)$ на $x - \alpha$ в $K[x]$:

$$\frac{f(x)}{x - \alpha} = x^{p^n-2} + \alpha x^{p^n-3} + \alpha^2 x^{p^n-4} + \cdots + \alpha^{p^n-2} =: g(x).$$

Підставивши $x = \alpha$ та використавши, що $\alpha^{p^n-2} = \frac{1}{\alpha}$, отримаємо

$$g(\alpha) = (p^n - 1) \cdot \frac{1}{\alpha} = -\frac{1}{\alpha} \neq 0.$$

Тут було використано, що $p^n \cdot x = 0$ для кожного $x \in K$, оскільки ми працюємо в полі характеристики p. Отже, $g(\alpha) \neq 0$, а тому $f(x)$ має $p^n - 1$ різних коренів в K. □

Теорема 186. *Скінченне поле $GF(p^n)$, що складається з p^n елементів, існує для кожного простого p та натурального $n \in \mathbb{N}$.*

Доведення. Нехай $K < \overline{\mathbb{Z}_p}$ є полем розкладу $x^{p^n} - x$ над $\mathbb{Z}_p$ і нехай підмножина $F \subset K$ складається з усіх нулів $x^{p^n} - x$ в K. З рівностей, див. твердження 121,

$$(\alpha \pm \beta)^{p^n} = \alpha^{p^n} \pm \beta^{p^n} = \alpha + \beta$$

та

$$(\alpha\beta)^{p^n} = \alpha^{p^n}\beta^{p^n} = \alpha\beta$$

випливає, що F замкнена відносно додавання, віднімання та множення. Очевидно, що $0, 1 \in F$. Якщо $\alpha \neq 0$, то з рівності $\alpha^{p^n} = \alpha$ випливає $(1/\alpha)^{p^n} = 1/\alpha$, а тому F містить обернений елемент до α. Таким чином, F є підполем K, яке містить $\mathbb{Z}_p$. За визначенням K є мінімальним розширенням $\mathbb{Z}_p$, що містить всі корені $x^{p^n} - x$, тому $K = F$. За попередньою лемою K складається з p^n різних елементів. Доведення завершено. □

Теорема 187. *Мультиплікативна група $\mathbb{F}_q^*$, що складається з ненульових елементів скінченного поля $\mathbb{F}_q$, $q = p^n$, є циклічною.*

Доведення. Припустимо, що скінченна група $\mathbb{F}_q^*$, $q = p^n$ не є циклічною. Порядок цієї групи дорівнює $q - 1$. Тоді існує число $s < q - 1$ таке, що $z^s = 1$ для кожного $z \in \mathbb{F}_q^*$. Це означає, що всі елементи поля $\mathbb{F}_q$ є коренями многочлена $x^{s+1} - x = 0$. Отже, многочлен степеня $< q$ має q коренів, що неможливо. □

Лекція 6

Основи теорії чисел

6.1 Основні арифметичні функції.

Нагадаємо деякі базові факти арифметики.

Твердження 188. *Мають місце такі твердження.*

(i) Кожне натуральне число $n \geq 1$ є або простим, або є добутком простих чисел.

(ii) Існує нескінченно багато простих чисел.

Доведення. Твердження (i) доведемо індукцією. Для $n = 2$ твердження очевидне. Припустимо, що твердження (i) виконується для всіх натуральних чисел $< n$. Якщо n просте, то (i) виконується. Якщо n не є простим, то n ділиться на деяке натуральне число $1 < d < n$, а отже $n = cd$ для деяких $1 < c, d < n$. Кожне з чисел c, d задовольняє (i) за припущенням індукції, тому n теж задовольняє (i).

Доведемо твердження (ii). Припустимо, що існує лише скінченне число простих чисел $p_1, p_2, \ldots, p_N$. Тоді число $n = p_1 p_2 \cdots p_N + 1$ є або простим, або добутком простих за частиною (i). Воно не може бути простим за припущенням, а отже є добутком простих. Але воно не ділиться на жодне з чисел $p_1, p_2, \ldots, p_N$. Отримане протиріччя доводить частину (ii). □

Теорема 189 (Основна теорема арифметики)**.** *Кожне натуральне число $n > 1$ може бути представлене у вигляді добутку простих чисел. Представлення єдине з точністю до порядку множників.*

Доведення. Доведемо твердження індукцією по n. Твердження, очевидно, вірне для $n = 2$. Припустимо, що твердження вірне для всіх натуральних чисел, більших за 1 та менших за n. Доведемо, що воно виконується для n. Якщо n просте, то доводити нічого. Якщо n складене та має два розклади

$$n = p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t, \tag{6.1}$$

то нам потрібно довести, що $s = t$ та кожне з q співпадає з деяким p. Оскільки p_1 ділить добуток $q_1 q_2 \cdots q_t$, то p_1 ділить одне з чисел $q_1, q_2, \ldots, q_t$. Перейменуємо $q_1, q_2, \ldots, q_t$ так, що p_1 ділить q_1. Тоді $p_1 = q_1$, оскільки ці числа прості. Скоротимо рівність (6.1) на p_1:

$$n/p_1 = p_2 p_3 \cdots p_s = q_2 q_3 \cdots q_t.$$

Якщо $s = 1$ та $t = 1$, то твердження доведене. Якщо $s > 1$ або $t > 1$, то $1 < n/p_1 < n$. Припущення індукції стверджує, що розклад числа n/p_1 єдиний з точністю до порядку множників. Отже, $s = t$ і розклади в (6.1) співпадають з точністю до порядку. Доведення завершено. □

Зауваження 190. Представлення числа $n > 1$ у вигляді добутку простих може містити однакові прості співмножники. Якщо в розкладі n на прості множники, $p_1, p_2, \ldots, p_s$ є різними простими співмножниками, та p_i входить a_i раз в розклад, то

$$n = \prod_{i=1}^{s} p_i^{a_i}.$$

Число 1 також можна записати в цьому вигляді з $a_i = 0$. Якщо записати прості числа у зростаючому порядку

$$p_1 = 2, \quad p_2 = 3, \quad p_3 = 5, \ldots$$

то кожне натуральне число n можна єдиним чином записати у вигляді

$$n = \prod_{i=1}^{\infty} p_i^{a_i},$$

де $a_i \geq 0$. Зрозуміло, що в такому записі $a_i > 0$ лише для скінченного числа іднексів i.

Твердження 191. *Мають місце такі твердження.*

(i) Якщо $n = \prod_{i=1}^{\infty} p_i^{a_i}$*, то множина всіх додатних дільників* n *співпадає з множиною чисел вигляду* $\prod_{i=1}^{\infty} p_i^{c_i}$*, де* $0 \le c_i \le a_i$*,* $i = 1, \dots, s$*.*

(ii) Якщо $a = \prod_{i=1}^{\infty} p_i^{a_i}$ *та* $b = \prod_{i=1}^{\infty} p_i^{b_i}$*, то*

$$\textit{НСД}\,(a,b) = \prod_{i=1}^{\infty} p_i^{\min(a_i,b_i)}.$$

(iii) Якщо $a = \prod_{i=1}^{\infty} p_i^{a_i}$ *та* $b = \prod_{i=1}^{\infty} p_i^{b_i}$*, то*

$$\textit{НСК}\,(a,b) = \prod_{i=1}^{\infty} p_i^{\max(a_i,b_i)}.$$

Теорема 192. *Показник степеня простого числа* p *в розкладі* $n!$ *на прості множники дорівнює*

$$\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \cdots + \left[\frac{n}{p^m}\right] + \cdots,$$

де $[\cdot]$ *позначає цілу частину.*

Доведення. Розглянемо множини

$$L_1 := \{k \in \mathbb{N} : 1 \le k \le n!, p \text{ ділить } k\}, \quad L_i := \{k \in L_{i-1} : p^i \text{ ділить } k\}, \quad i \ge 2.$$

Очевидно, що $|L_i| = [n/p^i]$. Твердження випливає сумуванням по всім i. □

6.2 Функції Мьобіуса та Ейлера.

Означення 193. Функція Мьобіуса $\mu : \mathbb{N} \mapsto \{-1, 0, 1\}$ визначається так: $\mu(1) = 1$, та, якщо $n = \prod_{i=1}^{s} p_i^{a_i}$ з $a_i > 0$, то

$$\mu(n) = (-1)^s, \quad \text{якщо } a_1 = a_2 = \cdots = a_s = 1;$$
$$\mu(n) = 0, \quad \text{інакше.}$$

Зокрема, $\mu(n) = 0$ тоді і тільки тоді, коли n ділиться на деякий повний квадрат > 1.

Теорема 194. *При* $n \in \mathbb{N}$ *маємо*

$$\sum_{d \textit{ ділить } n} \mu(d) = \left[\frac{1}{n}\right].$$

Доведення. Формула вірна для $n = 1$. Нехай $n = \prod_{i=1}^{s} p_i^{a_i}$. В сумі $\sum_{d \text{ ділить } n} \mu(d)$ ненульовими доданками є доданок при $d = 1$ та доданки, що відповідають дільникам d, які складаються з різних простих чисел. Тому,

$$\sum_{d \text{ ділить } n} \mu(d) = \mu(1) + \mu(p_1) + \cdots + \mu(p_s) + \mu(p_1 p_2) + \cdots + \mu(p_{s-1} p_s) + \cdots$$

$$+ \mu(p_1 p_2 \cdots p_s) = 1 + \binom{s}{1}(-1) + \binom{s}{2}(-1)^2 + \cdots + \binom{s}{s}(-1)^s = (1-1)^s = 0.$$

□

Нагадаємо, що фі-функція Ейлера була визначена в означенні 125. Як і для функції Мьобіуса має місце проста формула для суми $\sum_{d \text{ ділить } n} \phi(d)$.

Теорема 195. *При $n \in \mathbb{N}$ маємо*

$$\sum_{d \text{ ділить } n} \phi(d) = n.$$

Доведення. Розглянемо множину $S = \{1, 2, \ldots, n\}$ та представимо її у вигляді диз'юнктного об'єднання: $S = \cup_d A(d)$, де

$$A(d) := \{1 \le k \le n : \text{НСД}\,(k, n) = d\}.$$

За побудовою маємо

$$n = \sum_{d \text{ ділить } n} |A(d)|.$$

Рівність НСД $(k, n) = d$, $0 < k \le n$, виконується тоді і тільки тоді, коли НСД $(k/d, n/d) = 1$, $0 < k/d \le n/d$. Отже, якщо покласти $q = n/d$, то між множиною $A(d)$ та множиною тих q, для яких НСД $(q, n/d) = 1$ та $0 < q \le n/d$, існує бієкція. Потужність останньої з множин за визначенням дорівнює $\phi(n/d)$. Таким чином,

$$n = \sum_{d \text{ ділить } n} \phi(n/d).$$

Ця рівність еквівалентна рівності $\sum_{d \text{ ділить } n} \phi(d) = n$, оскільки, якщо d пробігає всі дільники n, то n/d також пробігає всі дільники n. Доведення завершено. □

Функції Ейлера та Мьобіуса пов'язані через таку формулу.

Теорема 196. *При $n \in \mathbb{N}$ маємо*

$$\phi(n) = \sum_{d \text{ ділить } n} \mu(d)\frac{n}{d}.$$

Доведення. Запишемо

$$\phi(n) = \sum_{k=1}^{n} \left[\frac{1}{\text{НСД}\,(n,k)}\right].$$

З теореми 194 випливає

$$\phi(n) = \sum_{k=1}^{n} \left[\frac{1}{\text{НСД}\,(n,k)}\right] = \sum_{k=1}^{n} \sum_{d \text{ ділить НСД}\,(n,k)} \mu(d)$$

$$= \sum_{k=1}^{n} \sum_{d \text{ ділить } n \text{ та } k} \mu(d) = \sum_{d \text{ ділить } n} \sum_{q=1}^{n/d} \mu(d) = \sum_{d \text{ ділить } n} \mu(d)\frac{n}{d}.$$

Передостання рівність в наведеній формулі випливає з того, що для довільного фіксованого дільника d числа n ми сумуємо по всім k, що кратні d, тобто по $k = qd$, де $q = 1, 2, \ldots, n/d$. □

Функцію Ейлера $\phi(n)$ легко обчислювати, знаючи розклад n на прості множники.

Теорема 197. *При $n \in \mathbb{N}$ маємо*

$$\phi(n) = n \prod_{p \text{ ділить } n} \left(1 - \frac{1}{p}\right).$$

Доведення. Для $n = 1$ формула очевидна, оскільки порожній добуток дорівнює одиниці. Нехай $n > 1$ та $p_1, p_2, \ldots, p_r$ є різними простими дільниками n. Маємо

$$\prod_{p \text{ ділить } n} \left(1 - \frac{1}{p}\right) = \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right) = 1 - \sum \frac{1}{p_i} + \sum \frac{1}{p_i p_j} + \cdots + (-1)^r \frac{1}{p_1 p_2 \cdots p_r}.$$

Помітимо, що кожний доданок в правій частині цієї формули має вигляд $\pm 1/d$, де d є дільником n, що складається з добутку різних простих чисел. Чисельник ± 1 в точності дорівнює $\mu(d)$. Таким чином,

$$\prod_{p \text{ ділить } n} \left(1 - \frac{1}{p}\right) = \sum_{d \text{ ділить } n} \frac{\mu(d)}{d},$$

а тому потрібна рівність випливає з теореми 196. □

З доведеної щойно теореми випливає ряд важливих наслідків, які ми наведемо в наступному твердженні.

Твердження 198. *Функція Ейлера ϕ має такі властивості:*

(i) $\phi(p^k) = p^k - p^{k-1}$ *для простого* p *та натурального* k*;*

(ii) $\phi(mn) = \phi(n)\phi(m)(d/\phi(d))$*, де* $d =$ *НСД* (n, m)*;*

(iii) $\phi(mn) = \phi(n)\phi(m)$*, якщо НСД* $(n, m) = 1$.

Доведення. Частина (i) випливає з теореми 197 з $n = p^k$. Для доведення частини (ii) запишемо

$$\frac{\phi(n)}{n} = \prod_{p \text{ ділить } n} \left(1 - \frac{1}{p}\right).$$

Помітимо, що кожний простий дільник добутку mn є або простим дільником n, або простим дільником m, а ті прості дільники, що ділять обидва числа n і m, ділять і $d =$ НСД (n, m). Отже,

$$\frac{\phi(mn)}{mn} = \prod_{p \text{ ділить } nm} \left(1 - \frac{1}{p}\right) = \frac{\prod_{p \text{ ділить } n} \left(1 - \frac{1}{p}\right) \prod_{p \text{ ділить } m} \left(1 - \frac{1}{p}\right)}{\prod_{p \text{ ділить } d} \left(1 - \frac{1}{p}\right)} = \frac{\frac{\phi(n)}{n}\frac{\phi(m)}{m}}{\frac{\phi(d)}{d}}.$$

Пункт (iii) є окремим випадком частини (ii). Доведення завершено. □

6.3 Формула Мьобіуса для обернення.

Формула

$$\phi(n) = \sum_{d \text{ ділить } n} \mu(d)\frac{n}{d}, \tag{6.2}$$

яку ми довели в теоремі 196 є окремим випадком формули Мьобіуса для обернення арифметичних згорток (згорток Діріхле), яку ми отримаємо нижче.

Означення 199. Для довільних функцій $f, g : \mathbb{N} \to \mathbb{C}$ їх **згорткою Діріхле** $f * g$ називається функція $h : \mathbb{N} \mapsto \mathbb{C}$, яка визанчена рівністю

$$h(n) = (f * g)(n) = \sum_{d \text{ ділить } n} f(d)g\left(\frac{n}{d}\right) = \sum_{ab=n} f(a)g(b),$$

де друга сума береться по всім парам натуральних чисел a, b, добуток яких дорівнює n.

Таким чином, формулу (6.2) можна записати

$$\phi = \mu * N,$$

де $N(n) = n$ є тотожною функцією. Має місце очевидне твердження

Твердження 200. *Згортка Діріхле є комутативною та асоціативною. Функція $I(n) = [1/n]$ є одиницею відносно згортки Діріхле, тобто*

$$f * I = I * f = f.$$

Розглянемо функцію $u : \mathbb{N} \mapsto \mathbb{C}$, що визначена рівністю $u(n) = 1$ для всіх $n \in \mathbb{N}$. Теорема 194 стверджує, що $\mu * u = I$.

Теорема 201. *З рівності*

$$f(n) = \sum_{d \text{ ділить } n} g(d) \tag{6.3}$$

випливає

$$g(n) = \sum_{d \text{ ділить } n} f(d)\mu\left(\frac{n}{d}\right). \tag{6.4}$$

Навпаки, з рівності (6.4) *випливає рівність* (6.3).

Доведення. Теорема є простим наслідком асоціативності згортки Діріхле та теореми 194. Рівність (6.3) можна записати у вигляді $f = g * u$. Домноживши на μ, отримаємо $f * \mu = (g * u) * \mu = g * (u * \mu) = g * I = g$, що є формулою (6.4). Навпаки, множення рівності $f * \mu = g$ на u дає (6.3). □

У якості застосування формули Мьобіуса для обернення отримаємо формулу для кількості незвідних нормованих многочленів над скінченним полем. Для цього нам знадобляться дві допоміжні леми.

Лема 202. *Нехай F – довільне поле. Мають місце такі твердження.*

(i) Многочлен $x^l - 1 \in F[x]$ ділить многочлен $x^n - 1 \in F[x]$ тоді і тільки тоді, коли l ділить n.

(ii) Якщо $a \in \mathbb{N}$, то $a^l - 1$ ділить $a^m - 1$ тоді і тільки тоді, коли l ділить m.

Доведення. Доведемо частину (i). Нехай $n = ql + r$, $0 \le r < l$. Тоді

$$\frac{x^n - 1}{x^l - 1} = x^r \frac{x^{ql} - 1}{x^l - 1} + \frac{x^r - 1}{x^l - 1}.$$

Оскільки $(x^{ql}-1)/(x^l-1) = (x^l)^{q-1}+(x^l)^{q-2}+\cdots+1$ в правій частині є многочленом, то ліва частина є многочленом тоді і тільки тоді, коли $\frac{x^r-1}{x^l-1}$ є многочленом. Це можливо лише у випадку $r = 0$.

Частина (ii) випливає з аналогічних міркувань після заміни змінної x на константу a. □

Помітимо, що в кільці $\mathbb{Z}_p[x]$ існує лише скінченна кількість многочленів заданого степеня, і позначимо через $F_d(x)$ (скінченний) добуток незвідних нормованих многочленів степеня d в кільці $\mathbb{Z}_p[x]$.

Лема 203. *Має місце рівність*

$$x^{p^n} - x = \prod_{d \text{ ділить } n} F_d(x).$$

Доведення. Помітимо, якщо деякий многочлен $f(x)$ ненульового степеня ділить $x^{p^n} - x$, то $f^2(x)$ не ділить $x^{p^n} - x$. Дійсно, якщо $x^{p^n} - x = f^2(x)g(x)$, то, взявши формально похідну, отримаємо

$$-1 = 2f(x)f'(x)g(x) + f^2(x)g'(x).$$

Отже, $f(x)$ ділить 1, що неможливо. Залишається довести: якщо $f(x)$ є нормованим незвідним многочленом степеня d в кільці $\mathbb{Z}_p[x]$, то $f(x)$ ділить $x^{p^n} - x$ тоді і тільки тоді, коли d ділить n.

Доведемо пряму імплікацію. Розглянемо просте розширення $K = \mathbb{Z}_p(\alpha)$ поля $\mathbb{Z}_p$, де α є коренем f. Оскільки f є незвідним многочленом степеня d над $\mathbb{Z}_p$, то K складається з p^d елементів, кожний з яких задовольняє умову $x^{p^d} - x = 0$.

Припустимо, що $x^{p^n}-x = f(x)g(x)$, тоді $\alpha^{p^n} = \alpha$. Розглянемо довільний елемент $b_1\alpha^{d-1} + b_2\alpha^{d-2} + \cdots + b_d$ поля K. Маємо

$$(b_1\alpha^{d-1} + b_2\alpha^{d-2} + \cdots + b_d)^{p^n} = b_1(\alpha^{p^n})^{d-1} + b_2(\alpha^{p^n})^{d-2} + \cdots + b_d$$
$$= b_1\alpha^{d-1} + b_2\alpha^{d-2} + \cdots + b_d.$$

Отже, всі елементи K задовольняють умову $x^{p^n} - x = 0$. Таким чином, $x^{p^d} - x$ ділить $x^{p^n} - x$, а отже згідно з лемою 202, d ділить n.

Доведемо обернене твердження. Припустимо, що d ділить n. Оскільки $\alpha^{p^d} = \alpha$ та $f(x)$ є мінімальним многочленом для α, то $f(x)$ ділить $x^{p^d} - x$. З припущення, що d ділить n, випливає, що $x^{p^d} - x$ ділить $x^{p^n} - x$ згідно з лемою 202. Отже, $f(x)$ ділить $x^{p^n} - x$.

□

Наслідок 204. Якщо позначити через N_d число незвідних нормованих многочленів степеня d в $\mathbb{Z}_p[x]$, то, прирівнюючи степені многочленів в попередній лемі, маємо

$$p^n = \sum_{d \text{ ділить } n} dN_d.$$

Використавши формулу Мьобіуса для обернення, маємо

$$N_n = \frac{1}{n} \sum_{d \text{ ділить } n} \mu(n/d)p^d.$$

Зауважимо, що сума в цій формулі не може дорівнювати нулю, оскільки є сумою степенів простого числа p з коефіцієнтами ± 1. Це означає, що $N_n \geq 1$ для всіх $n \in \mathbb{N}$. Таким чином, ми навели ще одне доведення того, що поле з p^n елементів існує для довільного простого p та натурального $n \in \mathbb{N}$.

6.4 Порівняння та системи порівнянь першого степеня.

Розглянемо найпростіше порівняння в кільці $\mathbb{Z}_m$:

$$ax \equiv b (\mathrm{mod}\, m). \tag{6.5}$$

Знаходження розв'язку цього порівняння еквівалентне знаходженню цілих розв'язків рівняння

$$ax + my = b. \tag{6.6}$$

Якщо b ділиться на НСД(a, m), то такі розв'язки можна знайти за допомогою алгоритму Евкліда. Якщо ж b не ділиться на НСД(a, m), то рівняння (6.6), а, отже, й порівняння (6.5) розв'язків не має. Продемонструємо алгоритм знаходження розв'язків на прикладі.

Приклад 205. Розв'яжемо рівняння

$$111x \equiv 75 (\mathrm{mod}\, 321).$$

Для цього розглянемо рівняння

$$111x + 321y = 75 \tag{6.7}$$

і розв'яжемо його в цілих числах. Застосовуючи алгоритм Евкліда для знаходження НСД $(321, 111)$, запишемо

$$321 = 2 \cdot 111 + 99, \quad 111 = 1 \cdot 99 + 12, \quad 99 = 8 \cdot 12 + 3, \quad 12 = 4 \cdot 3.$$

Отже, $d =$ НСД $(321, 111) = 3$ та

$$\begin{aligned} d = 3 = 99 - 8 \cdot 12 &= (321 - 2 \cdot 111) - 8 \cdot (111 - 1 \cdot 99) \\ &= 321 - 2 \cdot 111 - 8 \cdot 111 + 8(321 - 2 \cdot 111) = 9 \cdot 321 - 26 \cdot 111. \end{aligned}$$

Оскільки 75 ділиться на $d = 3$, то порівняння (6.7) має розв'язки. Домноживши рівність $3 = 9 \cdot 321 - 26 \cdot 111$ на $\frac{75}{d} = 25$, отримаємо

$$75 = 225 \cdot 321 - 650 \cdot 111.$$

Таким чином, для довільного $k \in \mathbb{Z}$

$$(-650 + k \cdot \frac{321}{d}) \cdot 111 \equiv 75 (\mathrm{mod}\, 321).$$

Три різні розв'язки порівняння (6.7) отримуємо при $k = 0, 1, 2$:

$$\begin{aligned} x_1 &= (-650) (\mathrm{mod}\, 321) = 313, \\ x_2 &= (-650 + 107) (\mathrm{mod}\, 321) = 99, \\ x_3 &= (-650 + 2 \cdot 107) (\mathrm{mod}\, 321) = 206. \end{aligned}$$

Розглянемо тепер систему лінійних порівнянь

$$\begin{cases} x \equiv c_1 (\mathrm{mod}\, m_1), \\ x \equiv c_2 (\mathrm{mod}\, m_2), \\ \quad \dots \\ x \equiv c_k (\mathrm{mod}\, m_k), \end{cases} \tag{6.8}$$

де m_i припустимо попарно взаємнопростими.

Теорема 206 (Китайська теорема про лишки)**.** *Покладемо* $M := m_1 m_2 \cdots m_k$, $M_s := M/m_s$. *Нехай* N_s *визначаються з рівностей*

$$M_s N_s \equiv 1 (\mathrm{mod}\, m_s), \quad s = 1, \ldots, k, \tag{6.9}$$

які мають розв'язки в силу співвідношення НСД $(M_s, m_s) = 1$*, яке є наслідком припущення, що* m_i *є попарно взаємнопростими. Тоді система порівнянь* (6.8) *має єдиний розв'язок в* $\mathbb{Z}_{m_1 m_2 \cdots m_k}$*:*

$$x = (M_1 N_1 c_1 + M_2 N_2 c_2 + \cdots + M_k N_k c_k) (\mathrm{mod}\, m_1 m_2 \cdots m_k).$$

Доведення. Для кожного $s = 1, \ldots, k$ помножимо рівність (6.9) на відповідне рівняння системи та отримаємо

$$x \equiv M_s N_s c_s (\mathrm{mod}\, m_s), \quad s = 1, \ldots, k.$$

Звідси

$$x \equiv M_s N_s c_s + \sum_{j \neq s} M_j N_j c_j (\mathrm{mod}\, m_s), \quad s = 1, \ldots, k,$$

оскільки $\sum_{j \neq s} M_j N_j c_j$ ділиться на m_s. Отже, кожен розв'язок системи (6.8) задовольняє

$$x \equiv x_0 (\mathrm{mod}\, m_s), \quad s = 1, \ldots, k, \tag{6.10}$$

де

$$x_0 = M_1 N_1 c_1 + M_2 N_2 c_2 + \cdots + M_k N_k c_k.$$

З рівностей (6.10) та в силу припущення, що m_i є попарно взаємнопростими, отримуємо

$$x \equiv x_0 (\mathrm{mod}\, m_1 m_2 \cdots m_k).$$

Доведення завершено. □

Зауваження 207. Кожне з порівнянь (6.9) можна розв'язати за допомогою алгоритму Евкліда.

6.5 Порівняння другого степеня. Символи Лежандра та Якобі.

Розглянемо порівняння

$$x^n \equiv a (\mathrm{mod}\, m), \quad \text{НСД}\,(a, m) = 1. \tag{6.11}$$

Означення 208. Якщо порівняння (6.11) має розв'язок, то a називається **лишком степеня** n **за модулем** m**.** Якщо це порівняння розв'язків не має, то a називається **нелишком степеня** n **за модулем** m**.** При $n = 2$ лишки та нелишки називаються **квадратичними.**

Будемо розглядати квадратичні порівняння за простим непарним модулем

$$x^2 \equiv a(\mathrm{mod}\, p), \quad \text{НСД}\,(a, p) = 1, \quad p \geq 3. \tag{6.12}$$

Теорема 209. *Якщо a є квадратичним лишком за модулем p, то порівняння* (6.12) *має рівно два розв'язки в $\mathbb{Z}_p$.*

Доведення. Нехай $a \neq 0$ є квадратичним лишком, тоді порівняння (6.12) має хоча б один розв'язок $x \equiv x_1(\mathrm{mod}\, p)$. Зрозуміло, що $x \equiv (p - x_1)(\mathrm{mod}\, p)$ теж буде розв'язком. Ці розв'язки різні, оскільки з рівності $x_1 \equiv p - x_1(\mathrm{mod}\, p)$ випливає $2x_1 \equiv 0(\mathrm{mod}\, p)$, а отже p ділить x_1, що суперечить припущенню $x_1^2 \equiv a(\mathrm{mod}\, p)$, $a \neq 0$. З іншого боку, многочлен $x^2 - a$ не може мати більше двох коренів в $\mathbb{Z}_p$. □

Теорема 210. *В множині $\mathbb{Z}_p \setminus \{0\}$, $p \geq 3$, існує рівно $(p-1)/2$ квадратичних лишків, та рівно $(p-1)/2$ квадратичних нелишків.*

Доведення. Серед чисел $1^2, 2^2, \ldots, (p-1)^2$ є не більше $(p-1)/2$ різних за модулем p чисел, оскільки $k^2 \equiv (p-k)^2(\mathrm{mod}\, p)$. З іншого боку, всі лишки за модулем p серед чисел $1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2$ є різними. Дійсно, якби $k^2 \equiv l^2(\mathrm{mod}\, p)$ для деяких $0 < k < l \leq \frac{p-1}{2}$, то порівняння $x^2 \equiv l^2(\mathrm{mod}\, p)$ мало б чотири різні корені $l, p-l, k, p-k$, що неможливо. □

Теорема 211. *Якщо a є квадратичним лишком за модулем p, то*

$$a^{\frac{p-1}{2}} \equiv 1(\mathrm{mod}\, p).$$

Якщо a є квадратичним нелишком за модулем p, то

$$a^{\frac{p-1}{2}} \equiv -1(\mathrm{mod}\, p).$$

Доведення. Згідно з малою теоремою Ферма $a^{p-1} \equiv 1(\mathrm{mod}\, p)$, а тому

$$\left(a^{\frac{p-1}{2}} + 1\right)\left(a^{\frac{p-1}{2}} - 1\right) \equiv 0(\mathrm{mod}\, p). \tag{6.13}$$

Тільки один з співмножників у лівій частині ділиться на p, оскільки в протилежному випадку, їх різниця, яка рівна 2, також ділилася б на p. Нехай a є квадратичним лишком, тоді $a \equiv x_0^2 (\mathrm{mod}\, p)$ для деякого x_0. Піднесемо це порівняння до степеня $\frac{p-1}{2}$:

$$a^{\frac{p-1}{2}} \equiv x_0^{p-1} \equiv 1 (\mathrm{mod}\, p).$$

Отже, квадратичні лишки є коренями рівняння $x^{\frac{p-1}{2}} - 1 = 0$ в $\mathbb{Z}_p$. Таких лишків рівно $\frac{p-1}{2}$, отже, інших коренів рівняння $x^{\frac{p-1}{2}} - 1 = 0$ не має. Зокрема, якщо a є квадратичним нелишком, то

$$a^{\frac{p-1}{2}} \not\equiv 1 (\mathrm{mod}\, p).$$

З рівності (6.13) випливає, що тоді $a^{\frac{p-1}{2}} \equiv -1 (\mathrm{mod}\, p)$. Доведення завершено. □

Означення 212. Нехай $a \in \mathbb{Z}$, а $p \geq 3$ є простим числом, та p не ділить a. **Символ Лежандра** $\left(\frac{a}{p}\right)$ визначається так:

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{якщо } a \text{ є квадратичним лишком за модулем } p, \\ -1, & \text{якщо } a \text{ є квадратичним нелишком за модулем } p. \end{cases}$$

Безпосередньо з теореми 211 випливає

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} (\mathrm{mod}\, p). \tag{6.14}$$

Твердження 213. *Символ Лежандра має такі властивості:*

(i) якщо $a \equiv a_1 (\mathrm{mod}\, p)$, *то* $\left(\frac{a}{p}\right) = \left(\frac{a_1}{p}\right)$;

(ii) $\left(\frac{1}{p}\right) = 1$;

(iii) $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$;

(iv) $\left(\frac{a_1 a_2 \cdots a_m}{p}\right) = \left(\frac{a_1}{p}\right)\left(\frac{a_2}{p}\right) \cdots \left(\frac{a_m}{p}\right)$.

Доведення. Частина (i) очевидна, оскільки з $a \equiv a_1 (\mathrm{mod}\, p)$ випливає $a^{\frac{p-1}{2}} \equiv a_1^{\frac{p-1}{2}} (\mathrm{mod}\, p)$. Оскільки порівняння $x^2 \equiv 1 (\mathrm{mod}\, p)$ має два корені $x = \pm 1$, то 1 є квадратичним лишком, що доводить (ii). Частина (iii) випливає з формули (6.14).

Для доведення (iv) запишемо з використанням формули (6.14):

$$\left(\frac{a_1a_2\cdots a_m}{p}\right) \equiv (a_1a_2\cdots a_m)^{\frac{p-1}{2}} = (a_1)^{\frac{p-1}{2}}(a_2)^{\frac{p-1}{2}}\cdots(a_m)^{\frac{p-1}{2}}$$
$$= \left(\frac{a_1}{p}\right)\left(\frac{a_2}{p}\right)\cdots\left(\frac{a_m}{p}\right) (\mathrm{mod}\, p).$$

□

Означення 214. Нехай $m \geq 3$ є непарним числом та $m = p_1p_2\cdots p_r$ є розкладом m на прості множники p_i (деякі з них можуть бути однаковими). Нехай a взаємно-просте з m. **Символом Якобі** називається вираз

$$\left(\frac{a}{m}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right).$$

Символ Якобі має властивості, аналогічні до властивостей символа Лежандра, хоча має і певні відмінності. Наприклад, у загальному випадку формула

$$\left(\frac{a}{m}\right) \equiv a^{\frac{m-1}{2}} (\mathrm{mod}\, m)$$

є *невірною*.

Твердження 215. *Символ Якобі має такі властивості:*

(i) якщо $a \equiv a_1 (\mathrm{mod}\, m)$, *то* $\left(\frac{a}{m}\right) = \left(\frac{a_1}{m}\right)$;

(ii) $\left(\frac{1}{m}\right) = 1$;

(iii) $\left(\frac{-1}{m}\right) = (-1)^{\frac{m-1}{2}}$;

(iv) $\left(\frac{a_1a_2\cdots a_l}{m}\right) = \left(\frac{a_1}{m}\right)\left(\frac{a_2}{m}\right)\cdots\left(\frac{a_l}{m}\right)$.

Доведення. Властивості (i), (ii) та (iv) є безпосередніми наслідками аналогічних властивостей символа Лежандра. Доведемо властивість (iii). Маємо

$$\left(\frac{-1}{m}\right) = \left(\frac{-1}{p_1}\right)\left(\frac{-1}{p_2}\right)\cdots\left(\frac{-1}{p_r}\right) = (-1)^{\frac{p_1-1}{2}}(-1)^{\frac{p_r-1}{2}}\cdots(-1)^{\frac{p_r-1}{2}}.$$

Залишається перевірити, що

$$(-1)^{\sum_{i=1}^{r}\frac{p_i-1}{2}} = (-1)^{\frac{m-1}{2}}.$$

Це випливає з формули

$$\frac{m-1}{2} = \frac{p_1 p_2 \cdots p_r - 1}{2} = \frac{\left(1+2^{\frac{p_1-1}{2}}\right)\left(1+2^{\frac{p_2-1}{2}}\right)\cdots\left(1+2^{\frac{p_r-1}{2}}\right)-1}{2}$$
$$= \sum_{i=1}^{r} \frac{p_i-1}{2} + 2N \quad (6.15)$$

для деякого цілого N. □

Наступна теорема дає інтерпретацію символа Лежандра в термінах підстановок.

Теорема 216. *Нехай p є непарним простим числом, a – взаємнопросте з p, $b \in \mathbb{Z}_p$. Визначимо підстановку $\pi_{a,b,p}$ на множині $\mathbb{Z}_p$ формулою $\pi_{a,b,p}(i) = (ai+b)(\mathrm{mod}\, p)$, $i \in \mathbb{Z}_p$. Тоді* $\mathrm{sign}\, \pi_{a,b,p} = \left(\frac{a}{p}\right)$, *де*

$$\mathrm{sign}\, \sigma := \begin{cases} 1, & \textit{якщо } \sigma \textit{ є парною підстановкою}, \\ -1, & \textit{якщо } \sigma \textit{ є непарною підстановкою}. \end{cases}$$

Доведення. Розглянемо многочлен

$$A(x_1, x_2, \ldots, x_p) = \prod_{1 \le i < j \le p} (x_i - x_j).$$

Під дією парної підстановки на множині аргументів $x_1, x_2, \ldots, x_p$ многочлен A не змінюється, а під дією непарної підстановки – змінює знак. Отже, для довільної $\sigma \in \mathfrak{S}_m$

$$\mathrm{sign}\, \sigma = \frac{A(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(p)})}{A(x_1, x_2, \ldots, x_p)}.$$

Зокрема, поклавши $x_i = i$ та $\sigma = \pi_{a,b,p}$, отримаємо

$$\mathrm{sign}\, \pi_{a,b,p} = \prod_{1 \le i < j \le p} \frac{\pi_{a,b,p}(i) - \pi_{a,b,p}(j)}{i-j}$$
$$\equiv \prod_{1 \le i < j \le p} \frac{ai - aj}{i-j} \equiv \prod_{1 \le i < j \le p} a \equiv a^{\frac{p(p-1)}{2}} (\mathrm{mod}\, p).$$

Оскільки, $a^p \equiv a (\mathrm{mod}\, p)$, то

$$\mathrm{sign}\, \pi_{a,b,p} \equiv a^{\frac{p-1}{2}} (\mathrm{mod}\, p) = \left(\frac{a}{p}\right).$$

Доведення завершено. □

Доведемо тепер так званий **квадратичний закон взаємності для символа Лежандра**.

Теорема 217 (Квадратичний закон взаємності для символа Лежандра)**.** *Нехай p, q є непарними різними простими числами. Тоді*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}.$$

Доведення. Розглянемо множину $X := \{0, 1, \ldots, pq-1\}$. Кожне число $x \in X$ можна подати єдиним чином у вигляді $x = ap + b$, де $0 \leq b < p$ та $0 \leq a < q$. Аналогічно, можемо єдиним чином записати $x = cq + d$, де $0 \leq c < p$, $0 \leq d < q$. Таким чином, відображення $\sigma_{p,q} : X \mapsto X$, що визначене рівністю

$$\sigma_{p,q}(b + aq) = a + bp, \quad 0 \leq a < p, \quad 0 \leq b < q,$$

є підстановкою на X.

Розглянемо множину $\overline{X} := \{(a, b) : 0 \leq a < p, 0 \leq b < q\}$. Відображення $\phi : X \mapsto \overline{X}$, що визначене рівністю

$$\phi(x) = (x(\mathrm{mod}\, p), x(\mathrm{mod}\, q))$$

є бієкцією згідно з китайською теоремою про лишки. Відображення μ, ν на $\overline{X}$, які задані формулами:

$$\mu(a, b) = (a, (a + pb)(\mathrm{mod}\, q)), \quad \nu(a, b) = ((qa + b)(\mathrm{mod}\, p), b),$$

де $0 \leq a < p$, $0 \leq b < q$, є підстановками на $\overline{X}$, оскільки p та q взаємнопрості. Покажемо, що

$$\phi \circ \nu^{-1} \circ \mu \circ \phi^{-1} = \sigma_{p,q}. \tag{6.16}$$

Для довільних $0 \leq a < p, 0 \leq b < q$ маємо

$$\begin{aligned}(\phi \circ \nu^{-1} \circ \mu \circ \phi^{-1})(qa + b) &= \phi^{-1}(\mu(\nu^{-1}(\phi(qa + b)))) \\ &= \phi^{-1}(\mu(\nu^{-1}((qa + b)(\mathrm{mod}\, p), b))) \\ &= \phi^{-1}(\mu(a, b)) \\ &= \phi^{-1}(a, (a + pb)(\mathrm{mod}\, q)) = a + bp = \sigma_{p,q}(b + aq).\end{aligned}$$

Це означає, що

$$\operatorname{sign}\sigma_{p,q} = \operatorname{sign}(\nu^{-1}\circ\mu) = \operatorname{sign}\nu \times \operatorname{sign}\mu. \tag{6.17}$$

Обчислимо знаки підстановок, що входять в формулу (6.17). Запишемо підстановку $\sigma_{p,q}$ у стандартному записі

$$\sigma_{p,q} = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & \cdots \\ y_1 & y_2 & y_3 & y_4 & \cdots \end{pmatrix}$$

Парність підстановки співпадає з парністю числа інверсій, тобто парністю кількості пар індексів i та j таких, що $x_i < x_j$ та $y_i > y_j$. Оскільки кожне x_i має вигляд $x_i = b_i + a_i q$, а кожне y_i має вигляд $y_i = a_i + b_i p$, то число інверсій співпадає з числом пар індексів i та j таких, що

$$b_i + a_i q < b_j + a_j q, \quad a_i + b_i p > a_j + b_j p,$$

або, еквівалентно,

$$b_i > b_j, \quad a_i < a_j.$$

Число пар індексів i та j, які задовольняють першу нерівність є $q(q-1)/2$, а другу – $p(p-1)/2$. Отже,

$$\operatorname{sign}\sigma_{p,q} = (-1)^{\frac{p(p-1)q(q-1)}{4}} = (-1)^{\frac{(p-1)(q-1)}{4}},$$

де остання рівність випливає з непарності числа pq. Для кожного фіксованого $0 \le a < p$, підстановка μ зводиться до підстановки $\pi_{p,a,q}$ з теореми 216, а тому має парність $\left(\frac{p}{q}\right)$. Оскільки, p непарне, то

$$\operatorname{sign}\mu = \left(\frac{p}{q}\right).$$

Аналогічно,

$$\operatorname{sign}\nu = \left(\frac{q}{p}\right).$$

Підставляючи отримані вирази в формулу (6.17), отримуємо потрібне твердження.

□

Наслідок 218 (Квадратичний закон взаємності для символа Якобі). Нехай m, n є непарними взаємнопростими натуральними числами. Тоді

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{(n-1)(m-1)}{4}}.$$

Доведення. Нехай $m = p_1 p_2 \cdots p_r$, $n = q_1 q_2 \cdots q_l$, де p_i, q_j є попарно різними непарними простими числами. Тоді

$$\left(\frac{m}{n}\right) = \left(\frac{p_1 p_2 \cdots p_r}{q_1 q_2 \cdots q_l}\right) = \prod_{j=1}^{l} \left(\frac{p_1 p_2 \cdots p_r}{q_j}\right) = \prod_{j=1}^{l} \prod_{i=1}^{r} \left(\frac{p_i}{q_j}\right),$$

де передостання рівність випливає з означення символа Якобі, а остання – з частини (iv) твердження 213. Далі, згідно з квадратичним законом взаємності для символа Лежандра

$$\left(\frac{m}{n}\right) = \prod_{j=1}^{l} \prod_{i=1}^{r} \left(\frac{p_i}{q_j}\right) = \prod_{j=1}^{l} \prod_{i=1}^{r} \left(\frac{q_j}{p_i}\right) (-1)^{\frac{(p_i-1)(q_j-1)}{4}} = \left(\frac{n}{m}\right) (-1)^{\sum_{j=1}^{l} \sum_{i=1}^{r} \frac{(p_i-1)(q_j-1)}{4}}.$$

Залишається перевірити, що

$$(-1)^{\sum_{j=1}^{l} \sum_{i=1}^{r} \frac{(p_i-1)(q_j-1)}{4}} = (-1)^{\frac{(n-1)(m-1)}{4}}.$$

Маємо

$$\left((-1)^{\sum_{i=1}^{r} \frac{p_i-1}{2}}\right)^{\sum_{j=1}^{l} \frac{q_j-1}{2}} = \left((-1)^{\frac{m-1}{2}}\right)^{\frac{n-1}{2}} = (-1)^{\frac{(n-1)(m-1)}{4}},$$

де передостання рівність випливає з формули (6.15). Доведення завершено. □

За допомогою квадратичного закону взаємності можна довести такий факт.

Твердження 219. *Нехай $m \geq 3$ – довільне непарне число. Тоді*

$$\left(\frac{2}{m}\right) = (-1)^{\frac{m^2-1}{8}}.$$

Доведення. При $m = 3$ це є безпосереднім наслідком рівностей $\left(\frac{2}{3}\right) \equiv 2^{\frac{3-1}{2}} - 1 (\mathrm{mod}\, 3) = -1$. Припустимо, що формула вірна для $k = m$ і доведемо, що вона вірна для $k = m + 2$. Маємо

$$\begin{aligned}
\left(\frac{2}{m+2}\right) &= \left(\frac{-1}{m+2}\right)\left(\frac{-2}{m+2}\right) = (-1)^{\frac{m+1}{2}} \left(\frac{m}{m+2}\right) \\
&= (-1)^{\frac{m+1}{2}} (-1)^{\frac{m-1}{2}\frac{m+1}{2}} \left(\frac{m+2}{m}\right) = (-1)^{\frac{m+1}{2}} \left(\frac{2}{m}\right) \\
&= (-1)^{\frac{m+1}{2}} (-1)^{\frac{m^2-1}{8}} = (-1)^{\frac{(m+2)^2-1}{8}},
\end{aligned}$$

де ми використали по черзі: властивість (iv) твердження 215; властивості (i) та (iii) цього ж твердження; квадратичний закон взаємності; властивість $(-1)^{n^2} = (-1)^n$ для кожного $n \in \mathbb{N}$ та властивість (i) твердження 215; припущення індукції. □

Розглняемо тепер, як використовувати символи Якобі та Лежандра для відповіді на питання про розв'язність квадратичного порівняння

$$x^2 \equiv a(\mathrm{mod}\, m), \quad \text{НСД}\,(a, m) = 1. \tag{6.18}$$

за довільним непарним модулем m. Якщо m – просте число, то (6.18) має розв'язок тоді і тільки тоді, коли символ Лежандра $\left(\frac{a}{p}\right) = 1$. Якщо m – складене, то аналогічний критерій з використанням символа Якобі вже не виконується. Наприклад, порівняння

$$x^2 \equiv 2(\mathrm{mod}\, 15)$$

розв'язків не має, як можна переконатись перебором, але

$$\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (2^{\frac{3-1}{2}}(\mathrm{mod}\, 3))(2^{\frac{5-1}{2}}(\mathrm{mod}\, 5)) = (-1)(-1) = 1.$$

Доведемо допоміжну лему.

Лема 220. *Нехай p – непарне просте число, a – квадратичний лишок за модулем p. Тоді порівняння*

$$x^2 \equiv a(\mathrm{mod}\, p^k)$$

має розв'язок для кожного $k \in \mathbb{N}$.

Доведення. Доведення проведемо індукцією по k. Для $k = 1$ твердження тривіальне. Нехай існує x_0 такий, що $x_0^2 \equiv a(\mathrm{mod}\, p^k)$, тобто $\frac{x_0^2 - a}{p^k} := x_1 \in \mathbb{Z}$. Будемо шукати розв'язок $x^2 \equiv a(\mathrm{mod}\, p^{k+1})$ у вигляді $x = x_0 + p^k t$, $t \in \mathbb{Z}$. Маємо

$$(x_0 + p^k t)^2 \equiv a(\mathrm{mod}\, p^{k+1}) \Leftrightarrow x_0^2 - a + 2x_0 t p^k \equiv 0(\mathrm{mod}\, p^{k+1}) \Leftrightarrow x_1 + 2x_0 t \equiv 0(\mathrm{mod}\, p).$$

Останнє порівняння, має розв'язок відносно невідомого параметра t, оскільки $2x_0$ та p взаємнопрості. □

Маючи на озброєнні щойно доведену лему та враховуючи, що порівняння

$$x^2 \equiv a(\mathrm{mod}\, pq),$$

де p, q – різні прості числа має розв'язок тоді і тільки тоді, коли мають розв'язки порівняння

$$x^2 \equiv a(\mathrm{mod}\, p) \quad \text{та} \quad x^2 \equiv a(\mathrm{mod}\, q),$$

можемо сформулювати такий критерій.

Твердження 221. *Нехай m є довільним непарним числом, a – взаємнопросте з m. Порівняння* (6.18) *має розв'язок тоді і тільки тоді, коли*

$$\left(\frac{a}{p_1}\right) = 1, \left(\frac{a}{p_2}\right) = 1, \ldots, \left(\frac{a}{p_r}\right) = 1,$$

де p_i, $i = 1, \ldots, r$ – всі прості дільники m. Зокрема, умова

$$\left(\frac{a}{m}\right) = 1$$

необхідна, але недостатня (якщо $r > 1$) для існування розв'язку порівняння (6.18).

Приклад 222. Порівняння $x^2 \equiv 219 (\mathrm{mod}\, 383)$ має два розв'язки. Це випливає з того, що число 383 просте, а символ Лежандра

$$\begin{aligned}\left(\frac{219}{383}\right) &= -\left(\frac{383}{219}\right) = -\left(\frac{164}{219}\right) = -\left(\frac{2 \cdot 2 \cdot 41}{219}\right) = -\left(\frac{41}{219}\right) \\ &= -\left(\frac{219}{41}\right) = -\left(\frac{14}{41}\right) = -\left(\frac{2}{41}\right)\left(\frac{7}{41}\right) = -\left(\frac{7}{41}\right) \\ &= -\left(\frac{41}{7}\right) = -\left(\frac{6}{7}\right) = -\left(\frac{-1}{7}\right) = -(-1)^{\frac{7-1}{2}} = 1.\end{aligned}$$

Зауважимо, що задача розв'язку порівняння (6.18) еквівалентна задачі факторизації m на прості множники, див. теорему 1 в [10], а тому є обчислювально складною і, наразі, нерозв'язною за прийнятний час для великих m.

Лекція 7

Основи криптографії

7.1 Базові поняття

Означення 223. Функція $f : X \mapsto Y$ називається **односторонньою**, якщо існує алгоритм «швидкого»[1] обчислення значення $f(x) \in Y$ для відомого $x \in X$, але не існує «швидкого» алгоритму знаходження $x \in X$ такого, що $f(x) = y$ при заданому $y \in Y$.

Приклад 224. Нехай $X = \mathbb{Z}$ та $Y = \mathbb{Z}_m$. Функція $f(x) = 3^x (\mathrm{mod}\, m)$, де m – деяке довге (≥ 128 десяткових знаків) ціле число, вважається односторонньою, оскільки наразі невідомий алгоритм швидкого обчислення x за відомим значенням $3^x (\mathrm{mod}\, m)$.

Приклад 225. Нехай $X = \mathcal{P} \times \mathcal{P}$ та $Y := \{n \in \mathbb{N} : n = pq$ для деяких $p, q \in \mathcal{P}\}$, де $\mathcal{P}$ – множина простих чисел. Функція $f(x, y) = xy$ вважається односторонньою, оскільки множення є швидкою операцією, а факторизація – ні.

Означення 226. Одностороння функція з порогом – це одностороння функція $f : X \mapsto Y$ разом з додатковою інформацією (**порогом**), яка дає змогу для довільного $y \in Y$ «швидко» обчислити прообраз $f^{-1}(y)$.

Приклад 227. В прикладі 225 порогом може бути, наприклад, знання одного з простих співмножників в $n = f(p, q) = pq$. Таке знання зводить факторизацію n до ділення довгих цілих чисел.

[1] Якщо X, Y – числові множини, то під «швидким» алгоритмом ми розуміємо алгоритм з поліноміальним (від довжини деякого позиційного запису вхідного числа) часом роботи.

Криптографією називають застосування математичного апарату для захисту інформації від несанкціонованого доступу, забезпечення цілісності інформації, аутентифікації адресатів тощо. **Криптосистемою** називають систему, яка забезпечує захист інформації.

Означення 228. Алфавітом A називається довільна непорожня скінченна множина символів. **Словом** в алфавіті A називається довільна скінченна послідовність символів з A. Множина A^* є сукупністю всіх слів в алфавіті A. **Повідомленням** m називається скінченна послідовність слів в A^*. **Простір повідомлень** M є множиною всіх повідомлень в алфавіті A. **Вихідним повідомленням** називається довільний елемент $m \in M$. **Простір шифортекстів** C – це множина всіх повідомлень в деякому алфавіті, відмінному від A. Елементи C називаються **шифротекстами**. **Функція шифрування** – це довільна бієкція $e : M \mapsto C$, обернене відображення $d : C \mapsto E$, $d = e^{-1}$, називається **функцією дешифрування**. **Шифрування вихідного повідомлення** m є обчисленням шифротексту $c = e(m)$. **Дешифрування** – це обчислення вихідного тексту $m = d(c)$. **Ключ** – це функція шифрування або функція дешифрування. **Простір ключей** – множина всіх ключів. **Схема шифрування (або шифр)** є пара (e, d). **Побудувати шифр** означає побудувати четвірку (M, C, e, d).

Адресатом називається санкціонований відправник або отримувач інформації. Канал – це засіб передачі інформації від одного адресата до іншого. Безпечний канал – це канал, який захищений від зовнішнього втручання.

Схема шифрування називається **криптографічно стійкою**, якщо без знання ключа (e, d) неможливо за прийнятний час розшифрувати будь-який шифротекст схеми.

Розрізняють шифри з **симетричним** та **відкритим** ключем.

Означення 229. Шифр (e, d) називається **шифром з симетричним ключем**, якщо ключ (e, d) можна «швидко» знайти, знаючи будь-яку з його компонент e або d.

Приклад 230. Нехай $A = \{a, b, c, \ldots, x, y, z\}$ є літерами англійського алфавіта. Словами є всі послідовності літер довжини п'ять. M – множина відповідних повідом-

лень, $C = M$. Ключом є довільна підстановка на множині A, наприклад:

$$e = \begin{pmatrix} a & b & c & d & \cdots & v & w & x & y & z \\ d & e & f & g & \cdots & y & z & a & b & c \end{pmatrix}.$$

Повідомлення «thisi sanic ebird» шифрується в повідомлення «wklvl vdqlf helug» і дешифрується за допомогою оберненої підстановки:

$$e = \begin{pmatrix} a & b & c & d & \cdots & v & w & x & y & z \\ x & y & z & a & \cdots & s & t & u & v & w \end{pmatrix}.$$

Приклад 231 (Шифр Вернама). Нехай алфавіт A є бінарним, $A = \{0, 1\}$. Повідомленням є послідовність бітів $m_1 m_2 \cdots m_t$, ключем є фіксоване бінарне слово $k_1 k_2 \cdots k_t$. Шифротекстом є бінарна послідовність $c_1 c_2 \cdots c_t$, де

$$c_j = (m_j + k_j) (\mathrm{mod}\, 2), \quad j = 1, \ldots, t.$$

Дешифрування можна здійснити за формулами

$$m_j = (c_j + k_j) (\mathrm{mod}\, 2), \quad j = 1, \ldots, t.$$

Означення 232. Шифр (e, d) є схемою шифрування з відкритим ключем, якщо функція шифрування e публікується (відкрита), але є односторонньою функцією: відповідна функція дешифрування d не обчислюється за прийнятний час.

Таблиця: Переваги та недоліки шифрів з симетричним та відкритим ключем.

	Симетричний ключ	Відкритий ключ
Переваги	швидке шифрування та дешифрування	лише частина ключа має зберігатись в таємниці
	короткі ключі	таємною частиною ключа адресати не обмінюються
	просте комбінування різних симетричних ключів	ключ може не змінюватись довгий час
		шифри з відкритим ключем можна використосувати для передачі симетричних ключів
Недоліки	ускладнена (таємна) передача ключів третім особам по каналам, на які може мати вплив супротивник	значно більш повільне шифрування та дешифрування у порівнянні з симетричними ключами
	необхідна часта зміна ключів	розміри ключів значно довші у порівнянні з симетричними ключами
		на даний момент немає доведення криптографічної стійкості

7.2 Протокол Діффі-Геллмана та проблема дискретного логарифма.

Протокол Діффі-Геллмана (англ. Diffie – Hellman key exchange) – це метод обміну ключами, який дає змогу адресатам узгоджувати ключ. Він дозволяє двом або більше участникам, що не мають жодних попередніх даних один про одного, отримати спільний секретний ключ із використанням незахищеного каналу зв'язку.

Алгоритм Діффі-Геллмана базується на складності задачі обчислення дискретного логарифма.

Означення 233. Нехай G – скінченна циклічна група порядку n. Нехай g є генератором G та $\beta \in G$. **Дискретний логарифм** $\log_g \beta$ елемента β за основою g – це єдине число $x \in \{1, 2, \ldots, n-1\}$ таке, що $g^x = \beta$.

Стандартна проблема дискретного логарифма – це задача знаходження цілого $x \in \{0, 1, \ldots, p-1\}$ для якого $g^x \equiv \beta (\mathrm{mod}\, p)$, де p – просте число, g – генератор мультиплікативної групи $\mathbb{Z}_p^*$ скінченного поля $\mathbb{Z}_p$, $\beta \in \mathbb{Z}_p^*$. Очевидним, але неефективним методом розв'язку проблеми дискретного лограрифма є простий перебір елементів g^0, g^1, g^2, доки не буде знайдений елемент β.

На даний момент невідомий поліноміальний (від довжини деякого позиційного запису p) алгоритм розв'язку стандартної проблеми дискретного логарифма. На цьому спостереженні базується алгоритм Діффі-Геллмана, який у найпростішій формі з двома адресатами описується схемою на Рис. 7.2[2]

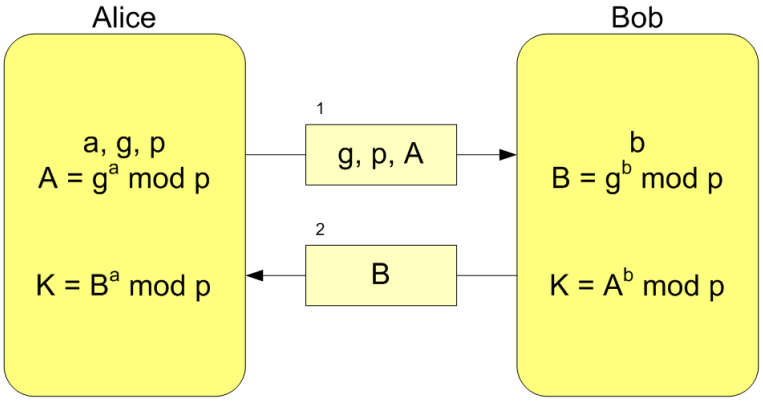

Рис. 7.1: Алгоритм Діффі – Геллмана.

[2]Зображення запозичено з `https://ru.wikipedia.org/wiki/IEEE_P1363`.

Нехай $G = \mathbb{Z}_p^*$, g – генератор G. Два адресати, Аліса і Боб, таємно обирають два випадкових цілих числа a та b з $\{1, 2, \ldots, p-1\}$ для яких обчислюють значення $A := g^a (\mathrm{mod}\, p)$ та $B := g^b (\mathrm{mod}\, p)$. Значеннями A та B вони обмінюються через незахищений канал передачі даних. Після цього Аліса та Боб обчислюють значення $a_{BA} = B^a (\mathrm{mod}\, p)$ та $a_{AB} = A^b (\mathrm{mod}\, p)$, відповідно, яке і буде їх спільним ключем $K = a_{AB} = a_{BA}$.

Приклад 234. Нехай Аліса і Боб домовились використивувати $p = 23$ та генератор $g = 5$ (ці дані можуть бути відкритими). Аліса обирає секретне число $a = 4$, та надсилає Бобу число $A = g^a (\mathrm{mod}\, p) = 5^4 (\mathrm{mod}\, 23) = 4$. Боб обирає секретне число $b = 3$, та надсилає Алісі число $B = g^b (\mathrm{mod}\, p) = 5^3 (\mathrm{mod}\, 23) = 10$. Аліса обчислює $a_{BA} = B^a (\mathrm{mod}\, p) = 10^4 (\mathrm{mod}\, 23) = 18$, а Боб обчислює $a_{AB} = A^b (\mathrm{mod}\, p) = 4^3 (\mathrm{mod}\, 23) = 18$. Спільне значення $K = a_{BA} = a_{AB} = 18$ є спільним таємним ключем.

Зауважимо, що лише a, b та $K = g^{ab} (\mathrm{mod}\, p)$ є секретними. Всі інші величини: p, g, $g^a (\mathrm{mod}\, p)$ та $g^b (\mathrm{mod}\, p)$ надсилаються відкрито.

Очевидно, що для надійного шифрування величини a, b та p повинні обиратись значно більшими. В цьому прикладі є лише 23 можливих значення $n (\mathrm{mod}\, 23)$, а тому значення a за відомими значеннями g, p та $g^a (\mathrm{mod}\, p)$ легко знайти перебором. Проте, якщо p є простим числом, що має > 256 десяткових знаків, то навіть найсучасніші комп'ютери не можуть знайти a з відомих g, p та $g^a (\mathrm{mod}\, p)$. Натомість, пряма задача: обчислення $g^a (\mathrm{mod}\, p)$ за відомими g, a та p є обчислювально простою, яка швидко розв'язується навіть для великих аргументів.

Розглянемо деякі алгоритми розв'язку проблеми дискретного логарифма.

ρ-АЛГОРИТМ ПОЛЛАРДА. Алгоритм Полларда для обчислення дискретного логарифма був запропонований у 1978 році і використовує так званий «парадокс днів народження» з теорії ймовірностей. Цей алгоритм вимагає в середньому $O(\sqrt{|G|})$ кроків для знаходження дискретного логарифма в скінченній групі G, тобто залишаєтсья експоненційним, але є значно швидшим за стандартний перебір, що потребує $O(|G|)$ кроків.

У найпростішому формулюванні «парадокс днів народження» стверджує, що

достатньо взяти групу лише з $23 \approx \sqrt{365}$ людей, щоб ймовірність події

$$\{\text{хоча б у двох людей день народження припадає на один день}\}$$

була більшою за $1/2$. Узагальнюючи, можна сказати, що потрібно зробити лише $O(\sqrt{n})$ виборів з поверненням з n-елементної множини, щоб хоча б один раз вибрати один і той самий елемент двічі з ймовірністю, яка перевищує $1/2$.

Ідея алгоритму Полларда полягає у побудові послідовності псевдовипадкових елементів циклічної групи G. Як тільки в цій послідовності з'явиться елемент, що вже зустрічався раніше, в послідовності виникає цикл. З цього циклу вдається обчислити значення дискретного логарифму. «Парадокс днів народження» стверджує, що з великою ймовірністю, знадобиться не більше ніж $O(\sqrt{|G|})$ кроків для знаходження циклу. Більш точно, для розв'язку рівняння $g^x = \beta$ алгоритм Полларда знаходить такі натуральні числа a, b, A, B, що $g^a\beta^b = g^A\beta^B$. Тоді γ, яке задовольняє $(B-b)\gamma \equiv (a-A) (\mathrm{mod}\, |G|)$, буде задовольняти $g^\gamma = \beta$. Для знаходження a, b, A, B алгоритм використовує метод Флойда (метод «черепахи та кролика») пошуку циклу, див. розділ 9.1 у Додатку, у послідовності $x_{i+1} = g^{a_{i+1}}\beta^{b_{i+1}} = f(x_i)$, де функція $f : G \mapsto G$ є «псевдовипадковою». Прикладом такої функції може бути:

$$f(x) = \begin{cases} \beta x, & x \in S_0, \\ x^2, & x \in S_1, \\ gx, & x \in S_2, \end{cases}$$

де $S_1 \cup S_2 \cup S_3 = G$ є деякми диз'юнктним розбиттям G на три підмножини приблизно одного розміру.

Алгоритм «малий крок – великий крок» (англ. «baby-step-giant-step»). Цей алгоритм є типовим представником алгоритмів з «просторово-часовою домовленістю» (англ. space – time tradeoff), в якому швидкодія досягається за рахунок використання більших об'ємів пам'яті. Алгоритм є досить простою модифікацією простого перебору.

Для розв'язання рівняння $g^x = \beta$ в циклічній групі G порядку n з генератором g, запишемо невідоме x у вигляді $x = im + j$, де $m := \lceil \sqrt{n} \rceil$, $0 \le i < m$, $0 \le j < m$.

Тоді

$$\beta(\alpha^{-m})^i = \alpha^j.$$

Алгоритм заздалегідь обчислює α^j для декількох значень j. Після цього, для фіксованого m, простою ітерацією обчислюється значення виразу в лівій частині і на кожному кроці порівнюється з кожним з заздалегідь обчислених значень α^j. Якщо для деякої пари індексів i, j рівність виконується, то шукане x знаходиться за формулою $x = im + j$.

Серед інших популярних алгоритмів згадаємо алгоритм Поліга-Геллмана, див. розділ 10.5.3 в [4], а також алгоритм обчислення порядку (англ. «the index calculus method»), див. розділ 3 в [12] та статтю [1]. Зокрема, якщо порядок групи G не є простим, але складається з простих множників, які є малими у порівнянні з $|G|$, то алгоритм Поліга-Геллмана, є найефективнішим з відомих.

7.3 Криптосистема RSA та проблема факторизації

У 1978 р. Р. Рівест, А. Шамір та Л. Адлеман запропонували схему шифрування з відкритим ключем, яка тепер відома під назвою RSA і використовується у величезній кількості сучасних протоколів захисту інформації таких, як PGP, $S/MIME$, TLS/SSL та інших. Криптографічна стійкість системи RSA базується на складності вирішення проблеми RSA, яка, в свою в чергу, базується на складності проблеми факторизації великих чисел.

Означення 235. Проблема RSA полягає у наступному: для заданих $n = pq$, де p, q – великі прості числа приблизно одного розміру, цілого e, яке взаємнопросте з $\phi = \phi(n) = (p-1)(q-1)$ та цілого $c \in \mathbb{N}$, знайти ціле число m таке, що $m^e \equiv c (\mathrm{mod}\, n)$.

Твердження 236. *В умовах проблеми RSA функція $f(x) = x^e (\mathrm{mod}\, n)$ є бієкцією, тому проблема RSA має єдиний розв'язок.*

Доведення. Достатньо показати, що відображення $f : \mathbb{Z}_n \mapsto \mathbb{Z}_n$ є сюр'єктивним. Іншими словами, рівняння $x^e \equiv y (\mathrm{mod}\, n)$ має розв'язок для всіх $y \in \mathbb{Z}_n$. Оскільки

НСД $(e, \phi) = 1$, то знайдуться $a, b \in \mathbb{Z}$ такі, що $ae + b\phi = 1$. Покажемо, що

$$x_0 \equiv y^a (\mathrm{mod}\, n) \tag{7.1}$$

буде шуканим розв'язком. Для цього перевіримо, що

$$y^{ea} \equiv y (\mathrm{mod}\, n).$$

Якщо НСД $(y, n) = 1$,то

$$y^{\phi b} \equiv 1 (\mathrm{mod}\, n), \tag{7.2}$$

оскільки мають місце рівності

$$(y^{b(q-1)})^{p-1} - 1 \equiv 0 (\mathrm{mod}\, p) \quad \text{та} \quad (y^{b(p-1)})^{q-1} - 1 \equiv 0 (\mathrm{mod}\, q),$$

які випливають з малої теореми Ферма. Домноживши обидві частини рівності (7.2) на y^{ea} та згадавши рівність $ae + b\phi = 1$, маємо

$$y \equiv y^{ea} \equiv x_0^a (\mathrm{mod}\, n).$$

Якщо НСД $(y, n) = p$, то достатньо перевірити, що

$$y^{ea} \equiv y (\mathrm{mod}\, q).$$

Ця рівність випливє з рівності $y^{\phi b} \equiv 1 (\mathrm{mod}\, q)$, яка також є наслідком малої теореми Ферма. Доведення завершено. □

Як випливає з доведення, знання значення $\phi = \phi(n)$ або, еквівалентно, знання факторизації $n = pq$, дозволяє швидко розв'язувати проблему RSA. На даний момент невідомий поліноміальний алгоритм розв'язку задачі факторизції і, як наслідок, проблеми RSA.

Розглянемо спочатку спрощену схему шифрування RSA, яка називається Kid-RSA.

СХЕМА KID-RSA. Нехай Аліса хоче надіслати Бобу повідомлення m, яке вважаємо натуральним числом. Боб генерує пару випадкових цілих чисел a, b та обчислює $M := ab - 1$. Потім генерує ще два числа A, B і обчислює $e = AM + a, d = BM + b, n = (ed - 1)/M$. Відкритим ключем Боба, який він передає Алісі, є пара (n, e), секретним ключем Боба є d. Аліса передає Бобу зашифроване повідомлення

$c = (me)(\mathrm{mod}\, n)$. Для його дешифрування Боб використовує свій секретний ключ d і обчислює $m = (dc)(\mathrm{mod}\, n)$.

Оскільки за побудовою $(de) \equiv 1(\mathrm{mod}\, n)$, то коректність дешифрування випливає з ланцюжка

$$(dc) \equiv (dme) \equiv (dem) \equiv m(\mathrm{mod}\, n).$$

Підкреслимо, що без знання секретного ключа d для дешифрування c потрібно знаходити $e^{-1}(\mathrm{mod}\, n)$.

Тепер розглянемо класичну схему RSA.

Схема RSA. Нехай Аліса хоче надіслати Бобу повідомлення $m \in \mathbb{N}$. Боб генерує пару різних простих чисел p, q приблизно одного розміру та обчислює $n = pq$, $\phi = \phi(n) = (p-1)(q-1)$. Потім Боб вибирає число e таке, що $1 < e < \phi$ та НСД$\,(e, \phi) = 1$ та обчислює за допомогою розширеного алгоритму Евкліда $d \in (1, \phi)$ таке, що $(de) \equiv 1(\mathrm{mod}\, \phi)$. Відкритим ключем Боба, який він передає Алісі є пара (n, e), секретним ключем Боба є d. Аліса передає Бобу зашифроване повідомлення $c = m^e(\mathrm{mod}\, n)$. Для його дешифрування Боб використовує свій секретний ключ d і обчислює $m = c^d(\mathrm{mod}\, n)$.

Доведення коректності. Маємо $de + b\phi = 1$ для деякого $b \in \mathbb{Z}$. Міркуючи так само, як при доведенні твердження 236, бачимо, що

$$m^{b\phi} \equiv 1(\mathrm{mod}\, n),$$

а тому

$$c^d \equiv m^{ed} \equiv m^{ed+b\phi} = m(\mathrm{mod}\, n).$$

□

Приклад 237 (Приклад на с. 176 в [4]). Нехай Аліса хоче передати Бобу повідомлення $t = NAB$. Боб генерує два простих числа $p = 499$, $q = 631$ та обчислює $n = pq = 314869$ та $\phi = (p-1)(q-1) = 313740$. Далі він обирає число $e = 305183$ таке, що НСД$\,(e, \phi) = 1$, та за допомогою розширеного алгоритму Евкліда знаходить $a = 181967 \in (1, \phi)$ таке, що $(ea) \equiv 1(\mathrm{mod}\, \phi)$. Відкритий ключ Боба $(314869, 305183)$ надсилається Алісі. Свій секретний ключ $a = 181967$ Боб зберігає в таємниці.

Аліса переводить текстове повідомлення $t = NAB$ у числову форму. Наприклад, записує його у вигляді числа в системі числення з основою 27 (число літер в англійському алфавіті + пробіл):

$$t = NAB \longrightarrow m = 14 \cdot 27^2 + 1 \cdot 27 + 2 = 10235.$$

Використовуючи відкритий ключ Боба, Аліса шифрує повідомлення m і отримує шифротекст c:

$$c = m^e (\mathrm{mod}\, n) = 10235^{305183} (\mathrm{mod}\, 314869) = 301085,$$

який надсилає Бобу.

Боб для дешифровки, використовуючи свій секретний ключ a, обчислює:

$$m = c^a (\mathrm{mod}\, n) = 301085^{181967} (\mathrm{mod}\, 314869) = 10235.$$

Отримане значення він переводить у текстову форму, записавши $10235 = 14 \cdot 27^2 + 1 \cdot 27 + 2$ та отримавши вихідний текст $t = NAB$.

Для знаходження секретного ключа Боба a супротивнику потрібно розв'язати порівняння $(ea) \equiv 1 (\mathrm{mod}\, \phi(n))$ або, що еквівалентно, знайти значення $\phi(n)$ за відомою частиною n відкритого ключа (n, e). Ця задача рівносильна задачі знаходження факторизації $n = pq$.

Розглянемо декілька алгоритмів факторизації.

ρ-АЛГОРИТМ ПОЛЛАРДА. Цей алгоритм схожий на алгоритм Полларда для знаходження дискретного логарифма, який був описаний вище. Алгоритм також базується на «парадоксі днів народження» і виглядає таким чином. Нехай N є складеним числом, для якого потрібно знайти нетривіальний дільник.

- Випадковим чином вибирається невелике початкове значення x_0, наприклад $x_0 = 2$.
- Будується послідовність $(x_n)_{n \geq 0}$, де $x_{n+1} = F(x_n) (\mathrm{mod}\, N)$ та $F : \mathbb{Z}_N \mapsto \mathbb{Z}_N$ є деяким простим для обчислення многочленом, який не є бієктивним, наприклад $F(x) = (x^2 + 1) (\mathrm{mod}\, N)$. Ідея такого вибору в тому, щоб сгенерована послідовність $(x_n)_{n \geq 0}$ була «псевдовипадковою» і до неї можна було б застосувати евритиску «парадоксу днів народження».

- На кожній ітерації обчислюється $d = \text{НСД}\,(N, |x_i - x_{2i}|)$.
- Якщо $d > 1$, то d є нетривіальним дільником N та $N = d \cdot (N/d)$.
- Якщо d або N/d не є простими, то за потреби алгоритм повторюється з N рівним N/d або d.

Коректність алгоритму випливає з таких міркувань. Нехай q – найменший нетривіальний дільник N. Послідовність $(z_n)_{n\geq 0}$, що визначена формулою

$$z_i = x_i (\mathrm{mod}\, q), \quad i \geq 0$$

може приймати лише скінченну кількість різних значень і також є «псевдовипадковою», оскільки

$$z_{i+1} = x_{i+1} (\mathrm{mod}\, q) = F(x_i) (\mathrm{mod}\, q) = F(x_i (\mathrm{mod}\, q)) = F(z_i), \quad i \geq 0.$$

Якби $(z_n)_{n\geq 0}$ була послідовністю незалежних випадкових чисел в $\mathbb{Z}_q$, то з ймовірністю більшою за $1/2$ вона почала б повторюватись після $O(\sqrt{q})$ кроків. Згідно з алгоритмом Флойда для пошуку циклів, ми шукаємо індекс i такий, що $z_i = z_{2i}$, тобто $(x_i - x_{2i}) \equiv 0 (\mathrm{mod}\, q)$. Це, в свою чергу, означає, що $\text{НСД}\,(x_i - x_{2i}, N) > 1$. Якщо $(x_i - x_{2i}) \not\equiv 0 (\mathrm{mod}\, N)$, то $\text{НСД}\,(x_i - x_{2i}, N)$ є нетривіальним дільником N. Якщо ж $x_i \equiv x_{2i} (\mathrm{mod}\, N)$ (ймовірність такої події дуже мала), то алгоритм не знаходить дільник і має бути виконаний спочатку з іншими x_0 та/або F. Наведені міркування також демонструють, що очікуваний час роботи алгоритму є $O(\sqrt{q})$, де q – найменший дільник N.

Метод Ферма. Метод базується на тривіальному факті.

Твердження 238. *Нехай $n \in \mathbb{N}$ є непарним. Тоді існує взаємнооднозначна відповідність між розкладами n на множники $n = ab$ та представленнями n у вигляді різниці квадратів $n = x^2 - y^2$. Ця відповідність задається формулами*

$$x = \frac{a+b}{2}, y = \frac{a-b}{2}, \quad a = x + y, \quad b = x - y.$$

Для факторизації цілого числа N алгоритм методу Ферма шукає таку пару цілих чисел (x, y), що $x^2 - y^2 = N$. Оскільки $x^2 - N = y^2$, тому $x \geq \lceil \sqrt{N} \rceil$. Далі простим

перебором шукається $k \in \mathbb{N}$ таке, що $(\lceil\sqrt{N}\rceil + k)^2 - N$ є повним квадратом. Якщо $(\lceil\sqrt{N}\rceil + k)^2 - N$ є повним квадратом y^2 для деякого $k \in \mathbb{N}$, то шуканий розклад є

$$N = (\lceil\sqrt{N}\rceil + k + y)(\lceil\sqrt{N}\rceil + k - y).$$

З опису методу Ферма зрозуміло, що він є найефективнішим у випадку, коли N має дільник, що близький до $\sqrt{N}$. Навпаки, якщо всі дільники N є або малими, або близькими до N, то метод Ферма є навіть гіршим за метод перебору дільників.

Ідея методу Ферма лежить в основі ряду сучасних алгоритмів факторизації, зокрема, алгоритму Діксона, див. [5], та алгоритму квадратичного решета. Спільною рисою цих методів є побудова пари (x, y) такої, що

$$x^2 \equiv y^2 (\mathrm{mod}\, N), \quad x \not\equiv \pm y (\mathrm{mod}\, N). \tag{7.3}$$

З цього співвідношення випливає, що спільний дільник НСД $(x - y, N)$ буде нетривіальним дільником N. Для побудови співвідношень (7.3) генерується достатня кількість порівнянь типу

$$U \equiv V (\mathrm{mod}\, N),$$

в яких повні факторизації U, V відомі. Після цього визначається така підмножина знайденого набору порівнянь, що їх множення дає співвідношення вигляду (7.3).

Алгоритм квадратичного решета. Алгоритм квадратичного решета складається з таких кроків.

- Спочатку вибираємо факторний базис $S := \{p_1, p_2, \ldots, p_t\}$, де $p_1 = -1$, p_j є $(j-1)$-им простим числом для якого N є квадратичним лишком[3] за модулем p_j, тобто $\left(\frac{N}{p_j}\right) = 1$. Число -1 включається в факторний базис, оскільки в алгоритмі потрібно факторизувати також від'ємні числа.
- Покладемо $m := \lfloor N \rfloor$.
- Будуємо $t + 1$ пару чисел (a_i, b_i). Покладемо $i = 1$. Обчислюємо при $z = 0, \pm 1, \pm 2, \ldots$ значення $q(z) = (m + z)^2 - N$ та перевіряємо, чи факторизується $q(z)$ в S (перебором дільників). Якщо $q(z)$ не факторизується в S, то

[3] Ми припускаємо, що НСД $(N, p_j) = 1$. В протилежному випадку факторизація тривіальна.

переходимо до наступного z. Якщо $q(z)$ факторизується в S, то покладаємо

$$a_i := z + m, \quad b_i := q(z) = \prod_{j=1}^{t} p_j^{e_{ij}}, \quad v_{ij} := e_{ij} (\mathrm{mod}\, 2).$$

та збільшуємо i на одиницю. Продовжуємо поки не буде побудовано $t+1$ пару (a_i, b_i).

- Набір векторів $\mathbf{v}_i = (v_{i1}, \ldots, v_{it}) \in \mathbb{Z}_2^t$, $1 \leq i \leq t+1$, лінійно залежний над $\mathbb{Z}_2$, оскільки є набором з $(t+1)$-го вектора в t-вимірному просторі. Тому будуємо підмножину $T \subset \{1, 2, \ldots, t+1\}$ таку, що $\sum_{i \in T} \mathbf{v_i} = \mathbf{0}$, знайшовши нетривіальний розв'язок системи лінійних рівнянь

$$\sum_{i=1}^{t+1} x_i \mathbf{v_i} = \mathbf{0}$$

над $\mathbb{Z}_2$.

- Покладемо $x = \prod_{i \in T} a_i (\mathrm{mod}\, N)$, $l_j := \frac{1}{2} \sum_{i \in T} e_{ij}$ при $1 \leq j \leq t$, та

$$y := \prod_{j=1}^{t} p_j^{l_j}.$$

За побудовою $b_i \equiv a_i^2 (\mathrm{mod}\, N)$ для всіх $i = 1, \ldots, t+1$, а тому

$$x^2 \equiv \prod_{i \in T} b_i \equiv y^2 (\mathrm{mod}\, N).$$

- Якщо $x \equiv \pm y (\mathrm{mod}\, N)$, то потірбно знайти іншу непорожню множину T таку, що $\sum_{i \in T} \mathbf{v_i} = \mathbf{0}$ і обчислити відповідні x та y. Якщо таку множину знайти не вдається, то потрібно замінити деякі з пар (a_i, b_i) на інші (для більших z).

- Якщо $x \not\equiv \pm y (\mathrm{mod}\, N)$, то $d :=$ НСД $(x - y, N)$ буде нетривіальним дільником N.

Зауваження 239. Еврістичними міркуваннями можна показати, що оптимальний розмір t факторного базису є

$$t \approx \exp\left\{ \frac{1}{2} \sqrt{\frac{1}{2} \log N \log\log N} \right\}.$$

При такому виборі t та з використанням оптимізованого для $\mathbb{Z}_2$ алгоритму Гаусса розв'язку системи лінійних рівнянь, можна отримати евристичну оцінку

$$O\left(\exp\left\{(1+o(1))\sqrt{\log N \log\log N}\right\}\right), \quad N \to \infty \tag{7.4}$$

для складносі алгоритму квадратичного решета. Деталі можна знайти на с. 45 в [9].

7.4 Криптосистема Ель-Гамаля

Криптографічна стійкість схеми Ель-Гамаля базується на складності задачі обчислення дискретного логарифма в мультиплікативній групі $\mathbb{F}_q^*$ скінченного поля $\mathbb{F}_q$, де $q = p^n$. Нагадаємо, що група $\mathbb{F}_q^*$ є циклічною, див. теорему 187. Схема була запропонована Тахером Ель-Гамалем в 1985 і є різновидом протоколу Діффі-Геллмана. Криптосистема Ель-Гамаля є криптосистемою з відкритим ключем.

Схема Ель-Гамаля. Нехай Аліса хоче надіслати Бобу повідомлення t. Боб генерує свій відкритий ключ:

- фіксується циклічна група $\mathbb{F}_{p^n}^*$ та її генератор g;
- вибирається випадкове число a, $1 \le a \le p^n - 2$;
- обчислюється елемент $y = g^a$;
- відкритим ключем є пара (g, y), а також структура групи $\mathbb{F}_{p^n}^*$, тобто числа p та n; секретним ключем є число a.

Аліса представляє повідомлення t у вигляді елемента $m \in \mathbb{F}_{p^n}^*$. Використовуючи відкритий ключ Боба (g, y), Аліса обчислює шифротекст $c = (\gamma, \delta)$, де

$$\gamma = g^k, \quad \delta = my^k,$$

а k є випадковим цілим числом з проміжку $1 \le k \le p^n - 2$ – «сеансовим ключем». Таке число генерується заново для кожного сеансу передачі.

Для дешифрування, використовуючи свій секретний ключ a, Боб обчислює γ^{-a}, а потім $m = \gamma^{-a}\delta$. Коректність дешифрування забезпечується рівностями:

$$\gamma^{-a}\delta = g^{-ka}\delta = y^{-k}\delta = m.$$

Приклад 240. Нехай $p = 13$, $n = 4$. Група $\mathbb{F}^*_{13^4}$ складається з $13^4 - 1 = 25860$ елементів. Елементами $\mathbb{F}^*_{13^4}$ є класи суміжності поліномів з $\mathbb{Z}_{13}[x]$ по незвідному поліному $p(x) = x^4 + x^2 + 2$. Можна показати, що $\mathbb{F}^*_{13^4} = \langle x + 5 + < p(x) > \rangle =: \langle g \rangle$. Боб генерує випадкове число $a = 2 \in [1, 28559]$ і обчислює

$$y = (x + 5 + < p(x) >)^2 = x^2 + 2x + 12 + < p(x) > .$$

Його відкритим ключем буде $(g, y) = (x + 5 + < p(x) >, x^2 + 2x + 12 + < p(x) >)$, а секретним – число 2.

Аліса представляє повідомідення $t = ZAM$ у вигляді елемента $m = \mathbb{F}^*_{13^4}$, наприклад, у такий спосіб. Спочатку t представляється у системі числення з основою 27 у вигляді $u = 26 \cdot 27^2 + 1 \cdot 27 + 13 = (18994)_{10}$, а потім це число переводить у систему числення з основою 13 у вигляді $m = (8851)_{13}$, або $m = 8x^3 + 8x^2 + 5x + 1 + < p(x) > \in \mathbb{F}^*_{13^4}$. Аліса генерує сеансовий ключ $k = 2134$ і за допомогою відкритого ключа Боба обчислює

$$\begin{aligned}
\gamma = g^k &= (x + 5)^{2134} + < p(x) > = 8x^3 + 9x^2 + 7x + 5 + < p(x) >, \\
y^k &= (x^2 + 2x + 12)^{2134} + < p(x) > = 10x^3 + 12x^2 + 3x + 1 + < p(x) >, \\
\delta = m \cdot y^k &= (8x^3 + 8x^2 + 5x + 1 + < p(x) >)(10x^3 + 12x^2 + 3x + 1 + < p(x) >) \\
&= 4x^3 + 6x^2 + 7x + 1 + < p(x) > .
\end{aligned}$$

Шифротекст $c = (8x^3 + 9x^2 + 7x + 5 + < p(x) >, 4x^3 + 6x^2 + 7x + 1 + < p(x) >)$ відправляється Бобу. За допомогою свого секретного ключа $a = 2$ Боб розшифровує повідомлення, обчисливши

$$\begin{aligned}
\gamma^a &= (8x^3 + 9x^2 + 7x + 5 + < p(x) >)^2 = 10x^3 + 12x^2 + 3x + 1 + < p(x) >, \\
\gamma^{-a} &= (10x^3 + 12x^2 + 3x + 1 + < p(x) >)^{-1} = 5x^3 + 7x^2 + 6x + 11 < p(x) >, \\
m = \gamma^{-a}\delta &= (5x^3 + 7x^2 + 6x + 11 < p(x) >)(4x^3 + 6x^2 + 7x + 1 + < p(x) >) \\
&= 8x^3 + 8x^2 + 5x + 1 + < p(x) > .
\end{aligned}$$

Як знайти генератор скінченної циклічної групи G порядку n? Ефективно це можна зробити лише знаючи факторизацію n. Як ми знаємо, в циклічній групі порядку n існує $\phi(n)$ генераторів. Тому вибраний навмання елемент групи g буде генератором з ймовірністю $\phi(n)/n$. Нехай відома факторизація $n = \prod_{i=1}^{k} p_k^{e_k}$,

тоді

$$\frac{\phi(n)}{n} = \prod_{i=1}^{k} \left(1 - \frac{1}{p_k}\right).$$

Якщо елемент $a \in G$ не є генератором, то його порядок є нетривіальним дільником n, а отже дільником n/p_i для деякого $i = 1, \ldots, k$. З цих спостережень випливає такий алгоритм пошуку генератора:

1) Вибрати випадковий елемент a групи G.

2) Для кожного $i = 1, \ldots, k$ обчислити $b = a^{n/p_i}$. Якщо $b = 1$, повернутись на крок 1.

3) Повернути a – генератор G.

7.5 Криптосистема Рабіна

Як і система RSA, криптосистема Рабіна базується на складності задачі факторизації великих чисел та проблемі квадратичних лишків. В той час як твердження про еквівалентність задачі взлому схеми RSA та задачі факторизації не доведена, аналогічне твердження для схеми Рабіна вірне.

Генерація ключей в схемі Рабіна виглядає так: вибираються два простих числа p, q приблизно одного розміру та обчислюється $n = pq$. Відкритим ключем є n, секретними ключем є пара (p, q). Повідомлення m представляється у вигляді цілого числа з проміжку $1, \ldots, n-1$ і обчислюється шифротекст $c \equiv m^2 (\mathrm{mod}\, n)$. Дешифрування зводиться до знаходження розв'яку рівняння $c \equiv x^2 (\mathrm{mod}\, n)$. Знання факторизації $n = pq$ (секретного ключа) робить цю задачу обчислювально простою, її розв'язок виглядає так:

1) Знайти два розв'язки $\pm r$ рівняння $c \equiv x^2 (\mathrm{mod}\, p)$ (див. нижче алгоритм Чіполли).

2) Знайти два розв'язки $\pm s$ рівняння $c \equiv x^2 (\mathrm{mod}\, q)$.

3) За допомогою розширенго алгоритму Евкліда знайти $u, v \in \mathbb{Z}$ такі, що $up + vq = 1$.

4) Повернути чотири[4] розв'язки $c \equiv x^2 (\mathrm{mod}\, n)$:

$$\pm(rvq + sup)(\mathrm{mod}\, n), \quad \pm(rvq - sup)(\mathrm{mod}\, n).$$

Те, що вказані числа будуть шуканими розв'язками випливає з ланцюжка рівностей

$$\begin{aligned}(\pm(rvq \pm sup))^2 (\mathrm{mod}\, n) &\equiv (r^2v^2q^2 + s^2u^2p^2 \pm 2rvsun)(\mathrm{mod}\, n) \\ &\equiv (r^2v^2q^2 + s^2u^2p^2)(\mathrm{mod}\, n) \equiv c(v^2q^2 + u^2p^2)(\mathrm{mod}\, n) \equiv c(\mathrm{mod}\, n),\end{aligned}$$

де передостання рівність випливає з порівнянь $r^2q \equiv cq(\mathrm{mod}\, n)$, $s^2p = cp(\mathrm{mod}\, n)$, а остання – з порівнянь

$$(v^2q^2 + u^2p^2) \equiv ((vq + up)^2 - 2vun)(\mathrm{mod}\, n) \equiv (vq + up)^2(\mathrm{mod}\, n) \equiv 1(\mathrm{mod}\, n).$$

Для розв'язку рівняння $x^2 \equiv a(\mathrm{mod}\, p)$, де p – непарне просте, існує ряд швидких алгоритмів. Ми розглянемо алгоритм Чіполли, іншим вживаним методом є алгоритм Тонеллі-Шенкса.

Алгоритм Чіполли. Алгоритм складається з трьох кроків і розв'язує рівняння $x^2 \equiv a(\mathrm{mod}\, p)$, де a – квадратичний лишок за модулем p.

1) Вибрати випадкове число b в полі $\mathbb{Z}_p$.

1') Якщо $b^2 - a = 0$ в полі $\mathbb{Z}_p$, повернути $\pm b$.

2) Якщо $\omega := b^2 - a$ є квадратичним лишком за модулем p, тобто символ Лежандра $\left(\frac{\omega}{p}\right) = 1$, повернутись на крок 1. Інакше перейти на крок 3.

3) Розв'язком порівняння $x^2 \equiv a(\mathrm{mod}\, p)$ будуть $\pm(b + \sqrt{\omega})^{\frac{p+1}{2}}$, де обчислення проводяться в простому розширенні $\mathbb{Z}_p(\sqrt{\omega})$ поля $\mathbb{Z}_p$.

Покажемо, що вказані числа будуть розв'язками. Оскільки ω є квадратичним нелишком, то згідно з теоремою 211

$$\omega^{\frac{p-1}{2}} \equiv -1(\mathrm{mod}\, p),$$

а тому в $\mathbb{Z}_p(\sqrt{\omega})$ маємо рівність

$$(\sqrt{\omega})^p = -\sqrt{\omega}.$$

[4]Далі буде продемонстровано, як знайти серед чотирьох розв'язків саме m, див. приклад 242.

Згадуючи, що поле $\mathbb{Z}_p(\sqrt{\omega})$ має характеристику p, отримуємо

$$\left(\pm(b+\sqrt{\omega})^{\frac{p+1}{2}}\right)^2 = (b+\sqrt{\omega})(b+\sqrt{\omega})^p = (b+\sqrt{\omega})(b^p+(\sqrt{\omega})^p)$$
$$(b+\sqrt{\omega})(b-\sqrt{\omega}) = b^2-\omega = b^2-(b^2-a) = a,$$

де ми використали те, що $x^p = x$ для всіх $x \in \mathbb{Z}_p$. Отже, $\pm(b+\sqrt{\omega})^{\frac{p+1}{2}}$ є двома нулями многочлена $x^2 - a = 0$ в $\mathbb{Z}_p(\sqrt{\omega})$. Цей многочлен другого степеня інших нулів в $\mathbb{Z}_p(\sqrt{\omega})$ мати не може. З іншого боку, він повинен мати два нулі в $\mathbb{Z}_p$, тому $\pm(b+\sqrt{\omega})^{\frac{p+1}{2}} \in \mathbb{Z}_p$.

Оскільки в $\mathbb{Z}_p^*$ є $(p-1)/2$ квадратичних нелишків та $(p-1)/2$ квадратичних лишків, то ймовірність вибрати b так, що $b^2 - n \neq 0$ є квадратичним нелишком дорівнює $1/2$.

Приклад 241. Розв'яжемо рівняння $x^2 \equiv 10 (\mathrm{mod}\, 13)$ алгоритмом Чіполло. Виберемо $b = 2$. Оскільки $b^2 - 10 \equiv 7 (\mathrm{mod}\, 13)$ та

$$\left(\frac{7}{13}\right) = \left(\frac{13}{7}\right) = \left(\frac{-1}{7}\right) = (-1)^{\frac{7-1}{2}} = -1,$$

то $b^2 - 10$ є квадратичним нелишком. Отже, розв'язками порівняння $x^2 \equiv 10 (\mathrm{mod}\, 13)$ будуть $x = \pm(2+\sqrt{7})^7 = \pm 6$.

Розглянемо тепер приклад застосування криптосистеми Рабіна.

Приклад 242. Нехай Аліса хоче відправити повідомлення $t = RAB$ Бобу. Боб генерує свій секретний ключ $(p, q) = (2131, 2437)$ та відкритий ключ $n = pq = 5193247$. Аліса переводить повідомлення $t = RAB$ у числовий вигляд, використавши, наприклад, систему числення з основною 27:

$$t = RAB \mapsto 18 \cdot 27^2 + 1 \cdot 26 + 2 = (13151)_{10} =: m.$$

Для того, щоб Боб міг відновити текст однозначно (нагадаємо, що квадратичне порівняння буде мати чотири розв'язки!), Аліса дописує до повідомлення m його останні два розряди і отримує повідомлення $m_1 = 1315151$. Повідомлення m_1 шифрується: $c = m_1^2 (\mathrm{mod}\, n) = 852957$. Шифротекст c відправляється Бобу.

Використовуючи свій секретний ключ $(p, q) = (2131, 2437)$, Боб дешифрує текст так:

- розв'язує рівняння $x^2 \equiv 852957 \equiv 557 (\mathrm{mod}\, 2131)$ і отримує $r = \pm 324$;
- розв'язує рівняння $x^2 \equiv 852957 \equiv 7 (\mathrm{mod}\, 2437)$ і отримує $s = \pm 829$;
- знаходить $(u, v) = (1107, -968)$ такі, що $1107 \cdot 2131 - 968 \cdot 2437 = 1$;
- знаходить чотири розв'язки $\pm(rvq \pm sup)$:

$$m^{(1)} = 1315151, \quad m^{(2)} = 3878096, \quad m^{(3)} = 2050346, \quad m^{(4)} = 3142901;$$

- оскільки лише $m^{(1)} = 1315151$ має в кінці розряди, що повторюються, то це і є оригінальним текстом, тому $m = m^{(1)} = 13151$.

7.6 Перевірка чисел на простоту

У попередніх підрозділах ми розглянули ряд криптосистем, чия складність базується на обчислювальній складності задачі факторизації великих чисел або задачі обчислення дискретного логарифма в скінченних циклічних групах. Одними з важливих підзадач в розглянутих нами алгоритмах є задачі перевірки великих чисел на простоту та генерації великих простих чисел. У цьому підрозділі ми розглянемо ряд детермінованих та ймовірнісних алгоритмів такого роду.

Нехай $a \in \mathbb{N}$ є фіксованим натуральним числом, яке потрібно перевірити на простоту. Розглянемо спочатку декілька тривіальних алгоритмів.

Простий перебір. Для кожного цілого $k = 2, \ldots, \lceil \sqrt{a} \rceil$ перевірити, чи ділиться a на k.

Критерій Вільсона. Алгоритм базується на теоремі Вільсона: число a є простим тоді і тільки тоді, коли $(a-1)! + 1 \equiv 0 (\mathrm{mod}\, a)$.

Доведення теореми Вільсона. Твердження очевидне для $a = 2$. Нехай $a = 2k + 1$ – непарне просте. Група $\mathbb{Z}_a^*$ ненульових елементів $\mathbb{Z}_a$ є циклічною. Нехай g є довільним генератором. Тоді множина $\{g^1, g^2, \ldots, g^{a-1}\}$ співпадає з $\{1, 2, \ldots, a-1\}$. Отже,

$$(a-1)! \equiv g^1 g^2 \cdots g^{a-1} = g^{\frac{a(a-1)}{2}} = g^{k(2k+1)} = (g^{2k})^k g^k = (g^{a-1})^k g^k \equiv g^k (\mathrm{mod}\, a).$$

Нарешті, $(g^k)^2 \equiv 1(\mathrm{mod}\, a)$, а тому $g^k \equiv \pm 1(\mathrm{mod}\, a)$. Оскільки g – генератор, то $g^k \not\equiv 1(\mathrm{mod}\, a)$, а тому

$$(a-1)! \equiv g^k \equiv -1(\mathrm{mod}\, a).$$

Обернене твердження майже очевидне. Нехай a складене, $a \neq 4$, тоді $(a-1)! \equiv 0(\mathrm{mod}\, a)$. При $a = 4$ маємо $(4-1)! \equiv 2(\mathrm{mod}\, 4)$. □

ТЕСТ ФЕРМА. Тест базується на малій теоремі Ферма: якщо a – просте, то $x^{a-1} \equiv 1(\mathrm{mod}\, a)$ для будь-якого x, яке не є кратним a. Проте, ця умова є лише необхідною, але не достатньою для простоти a: існують складені числа a такі, що $x^{a-1} \equiv 1(\mathrm{mod}\, a)$ для деяких x, взаємнопростих з a. Такі числа називаються **псевдопростими за основою** x. Тому при використанні тесту Ферма, зазвичай, беруть декілька різних x. Чим більшу кількість різних x перевірено, тим вища ймовірність, що a є простим. Нажаль, існують числа, які є псевдпростими за будь-якою основою x – **числа Кармайкла**.

Означення 243. Складене число n називається числом Кармайкла, якщо $x^{n-1} \equiv 1(\mathrm{mod}\, n)$ для кожного x, що є взаємнопростим з n. Еквівалентно $x^n \equiv x(\mathrm{mod}\, n)$ для всіх цілих n.

Існування чисел Кармайкла робить тест Ферма менш ефективним, оскільки такі числа завжди розпізнаються ним як прості. Доведемо результат, який дозволяє будувати конкретні приклади чисел Кармайкла.

Теорема 244 (Теорема Корсельта)**.** *Складене число n є числом Кармайкла тоді і тільки тоді, коли $n = p_1 p_2 \cdots p_r$, де всі p_i – різні прості числа та $p_i - 1$ ділить $n-1$ для всіх $i = 1, \ldots, r$.*

Доведення. Нехай $x^n \equiv x(\mathrm{mod}\, n)$ для всіх x. Покажемо, що x є вільним від квадратів, тобто в його розкладі всі прості співмножники є різними. Припустимо, що n ділиться на k^2 для деякого k, $1 < k < n$. Поклавши $x = k$, маємо $k^n \equiv k(\mathrm{mod}\, n)$, а тому $k^n \equiv k(\mathrm{mod}\, k^2)$, що неможливо. Тепер перевіримо, що $p-1$ ділить $n-1$, якщо p є простим дільником n. Розглянемо циклічну групу $\mathbb{Z}_p^*$ – мультиплікативну групу поля $\mathbb{Z}_p$ та деякий її генератор x. Оскільки $x^n \equiv x(\mathrm{mod}\, n)$, то $x^n \equiv x(\mathrm{mod}\, p)$. В силу взаємної простоти x та p, $x^{n-1} \equiv 1(\mathrm{mod}\, p)$. З іншого боку, $x^{p-1} \equiv 1(\mathrm{mod}\, p)$ в силу малої теореми Ферма. Отже, $n-1$ є кратним $p-1$.

Доведемо обернене твердження. Нехай n є вільним від квадратів та $p-1$ ділить $n-1$, якщо p є простим дільником n. Нехай x є довільним цілим числом. Спочатку покажемо, що $x^k \equiv x (\mathrm{mod}\, p)$ для кожного k такого, що $p-1$ ділить $k-1$. Нехай $k = 1 + q(p-1)$, $q \in \mathbb{Z}$. Маємо

$$x^k \equiv x^{q(p-1)+1} \equiv (x^p)^q x^{-q+1} \equiv x^q x^{-q+1} \equiv x (\mathrm{mod}\, p),$$

оскільки $x^p \equiv x (\mathrm{mod}\, p)$ згідно з малою теоремою Ферма. Зокрема, при $k = n$ маємо, що $x^n \equiv x (\mathrm{mod}\, p)$ для кожного простого дільника p числа n. Оскільки, n є вільним від квадратів, то $x^n \equiv x (\mathrm{mod}\, n)$. Доведення завершено. □

Приклад 245. З розкладів

$$561 = 3 \cdot 11 \cdot 17 \quad \text{та} \quad 1105 = 5 \cdot 13 \cdot 17$$

випливає, що 561 та 1105 є числами Кармайкла.

В роботі [2] доведено, що існує нескінченно багато чисел Кармайкла.

Розглянемо декілька більш складних тестів на простоту, які, на відміну від тесту Ферма, розрізняють числа Кармайкла.

Тест Соловея-Штрассена. Тест є ймовірнісним і дозволяє з будь-якою наперед заданою ймовірністю $p \in (0,1)$ дати відповідь на питання: «Чи є $a \in \mathbb{N}$ простим з ймовірністю, що перевищує p?». В основі тесту лежить таке твердження. Нагадаємо, що $\mathbb{Z}_a^*$ позначає мультиплікативну групу елементів, що мають обернені в кільці $\mathbb{Z}_a$. Порядок $\mathbb{Z}_a^*$ дорівнює $\psi(a)$.

Твердження 246. *Нехай a є складеним непарним числом. Позначимо через $E(a)$ підмножину тих чисел $k \in \mathbb{Z}_a^*$, які задовольняють двом умовам:*

- *НСД $(k, a) = 1$;*
- *$k^{\frac{a-1}{2}} \equiv \left(\frac{k}{a}\right) (\mathrm{mod}\, a)$, де $\left(\frac{k}{a}\right)$ є символом Якобі.*

Тоді $|E(a)| \leq \frac{|\mathbb{Z}_a^*|}{2} \leq \frac{a-1}{2}$.

Доведення. Покажемо спочатку, що знайдеться ціле k, $1 < k < a$, таке, що НСД $(k, a) = 1$, але $k^{\frac{a-1}{2}} \not\equiv \left(\frac{k}{a}\right) (\mathrm{mod}\, a)$, тобто $|E(a)| < a-1$. Нехай $a = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$

є розкладом a на прості непарні множники. Спочатку припустимо, що знайдеться $e_i > 1$. Не зменшуючи загальності, будемо вважати, що $e_1 > 1$. Покладемо

$$k = 1 + \frac{a}{p_1} = 1 + p_1^{e_1-1} p_2^{e_2} \cdots p_r^{e_r} =: 1 + q p_1^{e_1-1}, \quad q \in \mathbb{Z}, \tag{7.5}$$

тоді $k \equiv 1 (\mathrm{mod}\, p_i)$ для всіх $1 \leq i \leq r$, зокрема, НСД $(k, a) = 1$. Звідки

$$\left(\frac{k}{a}\right) = \left(\frac{k}{p_1}\right)^{e_1} \cdots \left(\frac{k}{p_r}\right)^{e_r} = 1.$$

Припустимо, що $k^{\frac{a-1}{2}} \equiv 1 (\mathrm{mod}\, a)$, а тому $k^{\frac{a-1}{2}} \equiv 1 (\mathrm{mod}\, p_1^{e_1})$. Нехай d є порядком k в $\mathbb{Z}^*_{p_1^{e_1}}$, тобто мінімальним натуральним числом з властивістю $k^d \equiv 1 (\mathrm{mod}\, p_1^{e_1})$. Тоді d ділить $\frac{a-1}{2}$, а тому d ділить $a - 1$. З формули (7.5) випливає, що $k \equiv (1 + q p_1^{e_1-1}) (\mathrm{mod}\, p_1^{e_1})$, а тому

$$1 \equiv k^d \equiv (1 + q p_1^{e_1-1})^d \equiv 1 + dq p_1^{e_1-1} (\mathrm{mod}\, p_1^{e_1}).$$

Це означає, що p_1 ділить dq. Оскільки q є взаємнопростим з p_1, то p_1 ділить d, а, отже, p_1 ділить $a - 1$. Це неможливо, оскільки p_1 є простим дільником a.

Розглянемо тепер випадок, коли a розкладається в добуток різних непарних простих множників першого степеня: $a = p_1 p_2 \cdots p_r$. Нехай s є квадратичним нелишком за модулем p_1, визначимо k з системи рівностей

$$\begin{cases} k \equiv s (\mathrm{mod}\, p_1), \\ k \equiv 1 (\mathrm{mod}\, p_2), \\ \quad \cdots \\ k \equiv 1 (\mathrm{mod}\, p_r). \end{cases}$$

Тоді

$$\left(\frac{k}{a}\right) = \left(\frac{k}{p_1}\right) \cdots \left(\frac{k}{p_r}\right) = \left(\frac{s}{p_1}\right) = -1.$$

З іншого боку, згідно з китайською теоремою про лишки $k^{\frac{a-1}{2}} \equiv 1 (\mathrm{mod}\, p_2 p_3 \cdots p_r)$, а тому порівняння $k^{\frac{a-1}{2}} \equiv -1 (\mathrm{mod}\, p_1 p_2 \cdots p_r)$ виконуватись не може. Оскільки за побудовою k не ділиться на жодне з p_i, то НСД $(k, a) = 1$.

Таким чином, існує принаймні одне число k, $1 < k < a$, таке, що дві умови твердження одночасно не виконуються. Залишається показати, що таких чисел не більше половини. Для цього помітимо, що $E(a)$ є підгрупою $\mathbb{Z}^*_a$. Оскільки

$|E(a)| < |\mathbb{Z}_a^*|$, то за теоремою Лагранжа про порядок підгрупи, $|E(a)| \leq \frac{\mathbb{Z}_a^*}{2} \leq \frac{a-1}{2}$. Доведення завершено. □

Сформулюємо тепер алгоритм Соловея-Штрассена. Для цього помітимо, якщо a є простим, то обидві умови твердження 246 виконуються в силу властивостей символа Лежандра для всіх $k \in \mathbb{Z}_a^*$. З іншого боку, якщо a є складеним, то знайдеться елемент $k \in \mathbb{Z}_a^*$ для якого принаймні одна з цих умов не виконується. Понад це, таких елементів не менше половини.

Алгоритм Соловея-Штрассена перевірки цілого непарного числа $a \in \mathbb{N}$ на простоту виглядає так:

1) Визначити цілу константу N, наприклад $N = 20$. Покласти $i := N$.

2) Вибрати випадкове ціле k, $1 < k < a$.

3) Якщо НСД $(a, k) > 1$, повернути «a є складеним». Інакше перейти на крок 4.

4) Якщо $k^{\frac{a-1}{2}} \not\equiv \left(\frac{k}{a}\right) (\mathrm{mod}\, a)$, повернути «$a$ є складеним». Інакше перейти на крок 5.

5) Покласти $i := i - 1$.

6) Якщо $i = 0$, то повернути «a є простим з ймовірністю $\geq 1 - 2^{-N}$». Інакше перейти на крок 2.

Алгоритм Міллера-Рабіна. Цей тест також є ймовірнісним, але, при рівній кількості ітерацій, дає значно кращу оцінку ймовірності розпізнавання простого числа, ніж тест Соловея-Штрассена.

Алгоритм Міллера-Рабіна базується на такому твердженні, доведення якого можна знайти в [11].

Твердження 247. *Нехай a є непарним складеним числом таким, що НСД $(6, a) = 1$. Визначимо цілі n, q рівністю $a - 1 = 2^n q$, де q є непарним. Позначимо через $\mathcal{S}(a)$ множину тих чисел $k \in \mathbb{Z}_a^*$ таких, що виконується одна з двох умов:*

- $k^q \equiv 1 (\mathrm{mod}\, a)$,
- $k^{2^j q} \equiv -1 (\mathrm{mod}\, a)$ *для деякого цілого* $0 \leq j < n$.

Тоді $|\mathcal{S}(a)| \leq \frac{a}{4}$.

Враховуючи наведене твердження, можемо сформулювати алгоритм Міллера-Рабіна перевірки цілого числа $a \in \mathbb{N}$, НСД $(6, a) = 1$, на простоту.

1) Знайти такі цілі n, q такі, що $a - 1 = 2^n q$, q є непарним.

2) Визначити цілу константу N, наприклад $N = 20$. Покласти $i := N$.

3) Покласти $i := i - 1$. Якщо $i = 0$, то повернути «a є простим з ймовірністю $\geq 1 - 4^{-(N-1)}$».

4) Вибрати випадкове ціле k, $1 < k < a$, та обчислити $b \equiv k^q (\mathrm{mod}\, a)$.

5) Якщо $b \equiv \pm 1 (\mathrm{mod}\, a)$ повернутись на крок 3. Інакше покласти $j := 0$ і перейти на крок 6.

6) Поки $i < n$ виконати:

   6.1) обчислити $b \equiv b^2 (\mathrm{mod}\, a)$;

   6.1) Якщо $b \equiv -1 (\mathrm{mod}\, a)$, повернутись на крок 3. Інакше покласти $j := j + 1$.

7) Повернути «a є складеним».

Лекція 8

Еліптичні криві

8.1 Проективна площина та однорідні координати.

Означення 248. Проективною площиною $\mathcal{P}_{\mathbb{R}^2}$ називається множина всіх прямих в $\mathbb{R}^3$, що проходять через початок координат. Еквівалентне визначення таке. Проективною площиною називається множина класів еквівалентності $\mathbb{R}^3 \setminus \{0\}$ по відношенню еквівалентності $\sim$, де $(x, y, z) \sim (x_1, y_1, z_1)$ тоді і тільки тоді, коли $\frac{x}{x_1} = \frac{y}{y_1} = \frac{z}{z_1}$.

Кожній прямій l, яка проходить через початок координат і не лежить в площині XOY, можна співставити точку $(x, y, 1)$, в якій l перетинає площину $Z = 1$. Якщо направляючий вектор l є (X, Y, Z), $Z \neq 0$, то $x = X/Z$ та $y = Y/Z$. Це означає, що проективна площина $\mathcal{P}_{\mathbb{R}^2}$ містить в собі звичайну площину $\mathbb{R}^2$, див. Рис. 8.1. Прямі, що лежать в площині $Z = 0$, не перетинають площину $Z = 1$. Тому вони відовідають не точкам на площині $Z = 1$, а *асимптотичним напрямкам* (пучкам паралельних прямих) в $\mathbb{R}^2$. Таким чином, проективну площину $\mathcal{P}_{\mathbb{R}^2}$ слід уявляти собі, як звичайну площину $\mathbb{R}^2$ до якої додано по одній точці на нескінченності в кожному з напрямків, який задається пучком паралельних прямих.

Означення 249. Будь-яка точка на проективній площині задається трійкою (X, Y, Z), що називається **однорідними координатами або проективними координатами точки**, де X, Y, Z не дорівнюють 0 одночасно.

З означення випливають такі властивості:

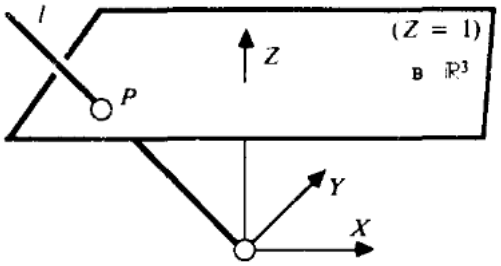

Рис. 8.1: Вкладення площини $\mathbb{R}^2$ в проективну площину $\mathcal{P}_{\mathbb{R}^2}$.

- точка, яка задається даним набором однорідних координат залишається незмінною, якщо координати помножити на ненульовий коефіцієнт;
- якщо $Z \neq 0$, то точці (X, Y, Z) відповідає точка $(X/Z, Y/Z)$ на Евклідовій площині.

Приклад 250. Як відомо, паралельний перенос в $\mathbb{R}^2$ не є лінійним перетворенням, а тому його не можна подати у вигляді множення вектора на деяку матрицю 2×2 зліва. Вкладення $\mathbb{R}^2$ в проективну площину $\mathcal{P}_{\mathbb{R}^2}$ дозволяє подати паралельний перенос у вигляді матричного множення. Нехай $\mathcal{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$, де $\mathcal{T}(x, y) = (x + x_0, y + y_0)$, є паралельним переносом на вектор (x_0, y_0) в $\mathbb{R}^2$. Розглянемо матрицю

$$T = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

та поставимо у відповідність точці $(x, y) \in \mathbb{R}^2$ точку $(x, y, 1) \in \mathcal{P}_{\mathbb{R}^2}$ проективної площини. Маємо

$$\begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + x_0 \\ y + y_0 \\ 1 \end{pmatrix}.$$

Ця рівність демонструє, що паралельний перенос можна реалізувати у вигляді матричного множення за допомогою однорідних координат. Аналогічно, за допомогою однорідних координат у вигляді матричного множення можна подати *перетворення перспективи*. Ці та інші властивості однорідних координат робить їх важливими у застосуваннях до комп'ютерної графіки.

Наведені означення без змін переносяться на випадок довільного поля K замість $\mathbb{R}$. Відповідну проективну площину будемо позначати $\mathcal{P}_{K^2}$.

Нехай K – довільне поле, і нехай $f \in K[x, y]$ є многочленом степеня m від змінних x, y з коефіцієнтами з поля K. Многочлену f відповідає однорідний многочлен F степеня m, що визначений рівністю $F(X, Y, Z) = Z^m f\left(\frac{X}{Z}, \frac{Y}{Z}\right)$ (зауважимо, що $\deg f = \deg F$).

Означення 251. Множина точок $C := \{(X, Y, Z) \in \mathcal{P}_{K^2} : F(X, Y, Z) = 0\}$, де F є однорідним многочленом,називається алгебраїчною кривою. Якщо $\deg F = m$, то кажуть, що F є кривою степеня m. Зокрема, при $m = 1$ алгебраїчна крива називається прямою, при $m = 2$ – квадратичною кривою (або конікою), при $m = 3$ – кубічною кривою (або кубікою).

Підкреслимо, що це визначення є коректним, оскільки $F(\lambda X, \lambda Y, \lambda Z) = \lambda^m F(X, Y, Z)$ для довільного $\lambda \neq 0$. Обмеження кривої C на K^2 співпадає з кривою $C^\circ \subset K^2$, що визначена рівністю

$$C^\circ = \{(x, y) \in K^2 : f(x, y) = 0\}.$$

Дійсно, якщо $Z \neq 0$, то $F(X, Y, Z) = 0$ тоді і тільки тоді, коли $f\left(\frac{X}{Z}, \frac{Y}{Z}\right) = 0$, тобто $\left(\frac{X}{Z}, \frac{Y}{Z}\right) \in C^\circ$.

Теорема Безу є фундаментальним результатом теорії алгебраїчних кривих.

Теорема 252 (Теорема Безу)**.** *Нехай C_1 та C_2 є двома алгебраїчними кривими над полем K степенів m та n, відповідно. Припустимо, шо криві не мають спільної компоненти, тобто не мають нескінченно багатьох спільних точок. Якщо поле K є алгебраїчно замкненим, то множина $C_1 \cap C_2 \subset \mathcal{P}_{K^2}$ їх точок перетину складається рівно з mn точок, якщо точки рахувати з кратностями.*

Ми доведемо теорему Безу лише в окремому випадку $m = 3$, $n = 3$. Проте, це доведення поширюється на загальний випадок без суттєвих змін. Інше доведення можна знайти, наприклад, в книзі [8].

Доведення випадку $m = 3$, $n = 3$. Нехай крива C_1 степеня $m = 3$ задана в однорідних координатах рівнянням

$$A(Z) := a_0 Z^3 + a_1 Z^2 + a_2 Z + a_3 = 0,$$

де a_i – є однорідними многочленом від X, Y степеня i. Аналогічно, нехай C_2 задана в однорідних координатах рівнянням

$$B(Z) := b_0 Z^3 + b_1 Z^2 + b_2 Z + b_3 = 0.$$

Точками перетину будуть розв'язки системи з цих двох рівнянь. Ці рівняння мають спільний корінь тоді і тільки, коли вони обидва діляться на деякий многочлен ненульового степеня. Тобто існують такі многочлени $P(Z)$ та $Q(Z)$ (їх коефіцієнти є многочленами від X, Y), що $A(Z)Q(Z) = B(Z)P(Z)$, при цьому $\deg Q \leq 2$ та $\deg P \leq 2$. Нехай

$$Q(Z) = u_0 Z^2 + u_1 Z + u_2, \quad P(Z) = v_0 Z^2 + v_1 Z + v_2.$$

Рівність $A(Z)Q(Z) = B(Z)P(Z)$ еквівалентна системі рівностей

$$\begin{aligned}
a_0 u_0 &= b_0 v_0, \\
a_1 u_0 + a_0 u_1 &= b_1 v_0 + b_0 v_1, \\
a_2 u_0 + a_1 u_1 + a_0 u_2 &= b_2 v_0 + b_1 v_1 + b_0 v_2, \\
a_3 u_0 + a_2 u_1 + a_1 u_2 &= b_3 v_0 + b_2 v_1 + b_1 v_2, \\
a_3 u_1 + a_2 u_2 &= b_3 v_1 + b_2 v_2, \\
a_3 u_2 &= b_3 v_2.
\end{aligned}$$

Цей набір є системою лійнійних однорідних рівнянь відносно 6 змінних u_0, u_1, u_2, v_0, v_1, v_2 з матрицею

$$M := \begin{pmatrix}
a_0 & 0 & 0 & -b_0 & 0 & 0 \\
a_1 & a_0 & 0 & -b_1 & -b_0 & 0 \\
a_2 & a_1 & a_0 & -b_2 & -b_1 & -b_0 \\
a_3 & a_2 & a_1 & -b_3 & -b_2 & -b_1 \\
0 & a_3 & a_2 & 0 & -b_3 & -b_2 \\
0 & 0 & a_3 & 0 & 0 & -b_3
\end{pmatrix}.$$

Система лінійних однорідних рівнянь має нетривільний розв'язок тоді і тільки

тоді, коли

$$\det M = \det M^T = -\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix} =: -\Delta(X,Y) = 0.$$

Нагадаємо, що для кожного i многочлени a_i та b_i є однорідними степеня i. Покажемо, що визначник в правій частині останньої рівності теж є однорідним многочленом степеня $mn = 3 \times 3 = 9$. Маємо для кожного $\lambda \neq 0$

$$\begin{aligned} &\Delta(\lambda X, \lambda Y) \\ &= \begin{vmatrix} a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 & 0 & 0 \\ 0 & a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 & 0 \\ 0 & 0 & a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 \\ b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 & 0 & 0 \\ 0 & b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 & 0 \\ 0 & 0 & b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 \end{vmatrix} = \lambda^{-6} \begin{vmatrix} a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 & 0 & 0 \\ 0 & \lambda a_0 & \lambda^2 a_1 & \lambda^3 a_2 & \lambda^4 a_3 & 0 \\ 0 & 0 & \lambda^2 a_0 & \lambda^3 a_1 & \lambda^4 a_2 & \lambda^5 a_3 \\ b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 & 0 & 0 \\ 0 & \lambda b_0 & \lambda^2 b_1 & \lambda^3 b_2 & \lambda^4 b_3 & 0 \\ 0 & 0 & \lambda^2 b_0 & \lambda^3 b_1 & \lambda^4 b_2 & \lambda^5 b_3 \end{vmatrix} \\ &= \lambda^{-6}\lambda\lambda^2\lambda^3\lambda^4\lambda^5\Delta(X,Y) = \lambda^9\Delta(X,Y). \end{aligned}$$

Оскільки поле K є алгебраїчно замкненим за припущенням, то многочлен $\Delta(u,1)$ розкладається в добуток mn лінійних множників. Отже,

$$\Delta(X,Y) = Y^{mn}\Delta(X/Y,1) = Y^{mn}\text{const}\prod_{k=1}^{mn}(X/Y - \alpha_k) = \text{const}\prod_{k=1}^{mn}(X - \alpha_k Y).$$

Це демонструє, що C_1 та C_2 мають рівно $mn = 9$ точок перетину в проективній площині $\mathcal{P}_{K^2}$. Доведення завершено. □

Ця теорема має ряд важливих наслідків, які нам знадобляться при вивченні еліптичних кривих.

Наслідок 253. Нехай точки $P_1, \ldots, P_8 \in \mathcal{P}_{K^2}$ є такими, що:

(i) жодні чотири не лежать на одній прямій (алебраїчній кривій першого порядку);

(ii) жодні сім не лежать на одній кривій другого порядку.

Тоді існує точка $P_9 \in \mathcal{P}_{K^2}$ така, що кожна кубічна крива, що проходить через $P_1, \ldots, P_8$, проходить також через P_9.

Ідея доведення. Кожна кубічна крива C може бути представлена як множина точок, що задовольняє рівняння

$$a_1X^3 + a_2X^2Y + a_3X^2Z + a_4XY^2 + a_5XYZ \\ + a_6XZ^2 + a_7Y^2Z + a_8YZ^2 + a_9Y^3 + a_{10}Z^3 = 0, \quad (8.1)$$

де $a_i \in K$. Такі рівняння можна додавати та множити на скаляри поля K, тому множина таких кубічних кривих утворює векторний простір S_3, розмірності 10 над K. Для кожної фіксованої точки P та кубічної кривої $C \in S_3$ умова $P \in C$ еквівалента тому, що деяка лінійна комбінація коефіцієнтів a_i, $i = 1, \ldots, 10$, кривої C дорівнює 0. Умови (i) та (ii) гарантують, що вісім умов $P_i \in C$ є лінійно незалежними, а тому згідно з теоремою Кронекера-Капеллі

$$\dim S_3(P_1, \ldots, P_8) = 10 - 8 = 2,$$

де $S_3(P_1, \ldots, P_8)$ є лінійним підпростором S_3, що складається з усіх кубічних кривих, які проходять через точки $P_1, \ldots, P_8$. Виберемо в $S_3(P_1, \ldots, P_8)$ базис F_1, F_2. Тоді з умови $G \in S_3(P_1, \ldots, P_8)$ випливає, що $G = \alpha F_1 + \beta F_2$. За теоремою Безу F_1 та F_2 перетинаються рівно в 9 точках, вісім з яких є $P_1, \ldots, P_8$. Позначимо дев'яту через P_9. Тоді $F_1(P_9) = F_2(P_9) = 9$, а тому $G(P_9) = 0$. Доведення завершено. □

8.2 Еліптичні криві над полем дійсних чисел.

Еліптична крива є геометричним об'єктом, точки якого мають природну групову структуру. Еліптичні криві можна визначити над довільним полем K. В криптографії використовуються еліптичні криві над скінченними полями. Спочатку розглянемо поняття еліптичної кривої над полем дійсних чисел $\mathbb{R}$.

Означення 254. Еліптичною кривою $\mathcal{E}$ над полем дійсних чисел називається множина точок, які задовольняють рівняння

$$y^2 = x^3 + ax + b, \quad a, b \in \mathbb{R}, \quad (8.2)$$

тобто $\mathcal{E} := \{(x, y) \in \mathbb{R}^2 : y^2 = x^3 + ax + b\}$.

Еліптична крива називається несингулярною, якщо $4a^3 + 27b^2 \neq 0$. Згадуючи формулу Кардано для коренів кубічного рівняння, бачимо, що умова несингулярності еліптичної кривої еквівалентна тому, що ця крива не має кратних коренів (не дотикається до осі абсцис).

До множини $\mathcal{E}$ додамо точку $\mathcal{O}$, яку будемо називати «нескінченно віддаленою точкою», і покладемо $\overline{\mathcal{E}} := \mathcal{E} \cup \{\mathcal{O}\}$.

Перейшовши до однорідних координат, можна ототожнити еліптичну криву $\mathcal{E}$ з кривою $\mathcal{E}_{\mathbb{R}^2}$ в проективній площині $\mathcal{P}_{\mathbb{R}^2}$, де

$$\mathcal{E}_{\mathbb{R}^2} := \{(X, Y, Z) \in \mathcal{P}_{\mathbb{R}^2} : ZY^2 = X^3 + aXZ^2 + bZ^3, Z \neq 0\}.$$

Відкинувши умову $Z \neq 0$, ми додамо до множини $\mathcal{E}_{\mathbb{R}^2}$ точку $(0, 1, 0) \in \mathcal{P}_{\mathbb{R}^2}$, яка відповідає «нескінченно віддаленій точкці» $\mathcal{O}$ в стандартній конструкції.

Таким чином, можна дати еквівалентне визначення еліптичної кривої, як кривої в проективному просторі.

Означення 255. Еліптичною кривою $\mathcal{E}_{\mathbb{R}^2}$ над полем дійсних чисел називається множина точок проективної площини $\mathcal{P}_{\mathbb{R}^2}$, які задовольняють рівняння

$$ZY^2 = X^3 + aXZ^2 + bZ^3, \quad a, b \in \mathbb{R}. \tag{8.3}$$

В подальшому ми будемо ототожнювати $\overline{\mathcal{E}}$ з множиною

$$\overline{\mathcal{E}}_{\mathbb{R}^2} := \{(X, Y, Z) \in \mathcal{P}_{\mathbb{R}^2} : ZY^2 = X^3 + aXZ^2 + bZ^3\}$$

та не писати індекс $\mathbb{R}^2$ в $\mathcal{E}_{\mathbb{R}^2}$ та $\overline{\mathcal{E}}_{\mathbb{R}^2}$.

Ключовий факт, на якому базується вся подальша теорія, полягає в тому, що на множині $\overline{\mathcal{E}}$ можна задати структуру абелевої групи. Розглянемо дві точки $P, Q \in \overline{\mathcal{E}}$ та проведемо через них пряму. Згідно з теоремою Безу, ця пряма перетне $\overline{\mathcal{E}}$ в деякій третій точці[1].

Задамо бінарну операцію $\oplus$ на $\overline{\mathcal{E}}$, як показано на Рис. 8.2. Також покладемо за визначенням $P \oplus \mathcal{O} = \mathcal{O} \oplus P = P$ для всіх $P \in \overline{\mathcal{E}}$. Будемо позначати через $-P$ точку, яка симетрична точці $P \in \overline{\mathcal{E}}$ відносно осі абсцис.

[1]Це виконується також для еліптичних кривих над довільним полем, навіть не алгебраїчно замкненим. Це випливає зі спостереження, що у кубічного рівняння над полем K, яке має в K два корені, третій корінь теж належить K. Точки $P, Q \in \overline{\mathcal{E}}$ за побудовою є коренями відповідного рівняння.

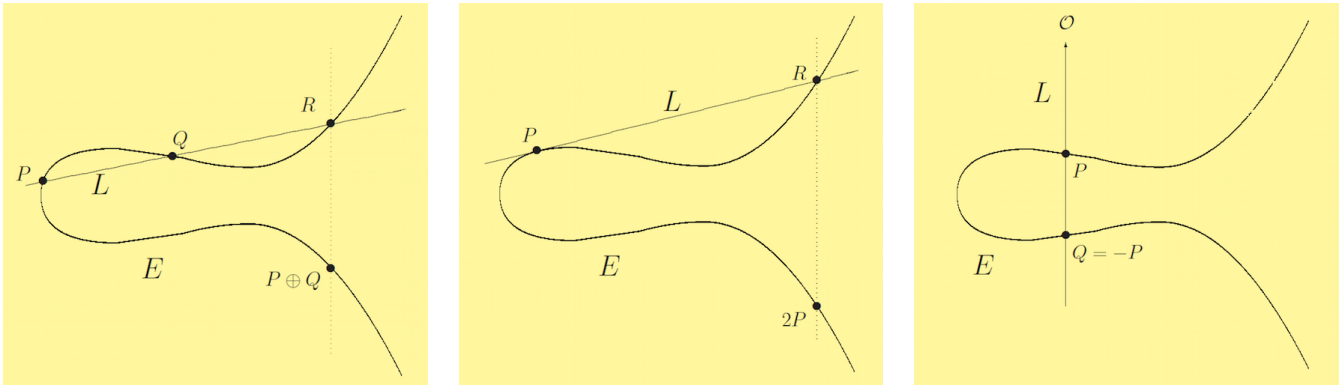

Рис. 8.2: Груповий закон $P \oplus Q$ на еліптичній кривій $y^2 = x^3 - 5x + 8$: $P \neq Q$ (зліва), $P = Q$ (в центрі), $P = -Q$ (справа)

Виведемо формули для координат (x_3, y_3) точки $P \oplus Q$ за відомими координатами (x_1, y_1) та (x_2, y_2) точок P, Q, відповідно. Розглянемо спочатку випадок $Q \neq P$, $Q \neq -P$, тобто $x_1 \neq x_2$.

Нехай $y = \alpha x + \beta$ є рівнянням прямої l, яка проходить через точки P та Q. Тоді $\alpha = \frac{y_2 - y_1}{x_2 - x_1}$, $\beta = y_1 - \alpha x_1$. Точка (x, y) цієї прямої лежить на еліптичній кривій $\mathcal{E}$ тоді і тільки тоді, коли $(\alpha x + \beta)^2 = x^3 + ax + b$ або $x^3 - (\alpha x + \beta)^2 + ax + b = 0$. Це кубічне рівняння має три корені: $(x_1, \alpha x_1 + \beta)$, $(x_2, \alpha x_2 + \beta)$ та (x_3, y_3). Згідно з теоремою Вієта

$$x_1 + x_2 + x_3 = \alpha^2 \quad \Rightarrow \quad x_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2, \tag{8.4}$$

а також

$$y_3 = \alpha x_3 + \beta = -y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_1 - x_3). \tag{8.5}$$

У випадку, коли $P = Q$, пряма l дотикається до еліптичної кривої $\mathcal{E}$ в точці (x_1, y_1). Взявши похідну, отримаємо

$$\alpha = \frac{3x_1^2 + a}{2y_1}.$$

Перехід до границі при $x_1 \to x_2$ в формулах (8.4) та (8.5) дає формули для точки $P \oplus Q = 2 \cdot P$:

$$x_3 = \left(\frac{3x_1^2 + a}{2y_1}\right)^2 - 2x_1 \tag{8.6}$$

та

$$y_3 = -y_1 + \left(\frac{3x_1^2 + a}{2y_1}\right)(x_1 - x_3). \tag{8.7}$$

Теорема 256. *Операція $\oplus$ на $\overline{\mathcal{E}}$ має такі властивості:*

(i) для всіх $P \in \overline{\mathcal{E}}$ маємо $P \oplus \mathcal{O} = \mathcal{O} \oplus P = P$;

(ii) для всіх $P \in \overline{\mathcal{E}}$ маємо $P \oplus (-P) = \mathcal{O}$;

(iii) для всіх $P, Q, R \in \overline{\mathcal{E}}$ маємо $P \oplus (Q \oplus R) = (P \oplus Q) \oplus R$;

(iv) для всіх $P, Q \in \overline{\mathcal{E}}$ маємо $P \oplus Q = Q \oplus P$.

Таким чином, $(\overline{\mathcal{E}}, \oplus)$ є абелевою групою.

Доведення. Властивості (i), (ii) є частиною означення. Властивість (iv) є очевидною за побудовою. Властивість (iii) можна довести з використанням явних формул (8.4), (8.5), (8.6) та (8.7), окремо розглянувши всі можливі випадки взаємного розташування точок $P, Q, R, P+Q, Q+R, P+(Q+R), (P+Q)+R \in \overline{\mathcal{E}}$. Ми наведемо доведення. що грунтується на теоремі Безу та наслідках з неї.

Достатньо перевірити, що $-(P \oplus (Q \oplus R)) = -((P \oplus Q) \oplus R)$. Розглянемо такі прямі:

$$\begin{aligned}
\mathcal{L}_1 &\quad \text{є пряма через точки} \quad P, Q, -(P \oplus Q);\\
\mathcal{L}_2 &\quad \text{є пряма через точки} \quad P \oplus Q, R, -((P \oplus Q) \oplus R);\\
\mathcal{L}_3 &\quad \text{є пряма через точки} \quad Q \oplus R, -(Q \oplus R), \mathcal{O};\\
\mathcal{N}_1 &\quad \text{є пряма через точки} \quad P \oplus Q, -(P \oplus Q), \mathcal{O};\\
\mathcal{N}_2 &\quad \text{є пряма через точки} \quad Q, R, -(Q \oplus R);\\
\mathcal{N}_3 &\quad \text{є пряма через точки} \quad P, Q \oplus R, -(P \oplus (Q \oplus R)).
\end{aligned}$$

Нехай відповідні рівняння є

$$\begin{gathered}
L_1 = L_1(X,Y,Z) = 0, \quad L_2 = L_2(X,Y,Z) = 0, \quad L_3 = L_3(X,Y,Z) = 0,\\
N_1 = N_1(X,Y,Z) = 0, \quad N_2 = N_2(X,Y,Z) = 0, \quad N_3 = N_3(X,Y,Z) = 0,
\end{gathered}$$

Розглянемо дві кубічні криві, які задані рівняннями:

$$L_1 \cdot L_2 \cdot L_3 = 0 \quad \text{та} \quad N_1 \cdot N_2 \cdot N_3 = 0.$$

За побудовою ці кубічні криві проходять через вісім точок множини

$$\mathcal{W}_8 := \{\mathcal{O}, P, Q, R, P \oplus Q, Q \oplus R, -(P \oplus Q), -(Q \oplus R)\}.$$

Припустимо, що ці точки задовольняють умовам (i) та (ii) наслідку 253. Тоді знайдеться точка S, яка лежить на кривих $L_1 \cdot L_2 \cdot L_3 = 0$, $N_1 \cdot N_2 \cdot N_3 = 0$, а також на кривій $\overline{\mathcal{E}}$, оскільки всі вони проходять через вказані вісім точок. Таким чином, десять точок $\mathcal{W}_8 \cup \{S, -(P \oplus (Q \oplus R))\}$ лежать на кривих $N_1 \cdot N_2 \cdot N_3 = 0$ та $\overline{\mathcal{E}}$. За теоремою Безу, таких точок може бути лише дев'ять, тому $S = -(P \oplus (Q \oplus R))$. Аналогічно, $S = -((P \oplus Q) \oplus R)$.

Залишається перевірити, що точки в $\mathcal{W}_8$ задовольняють умовам (i) та (ii) наслідку 253. Припустимо, що в $\mathcal{W}_8$ існують чотири точки, які лежать на одній прямій L, тоді L та $\overline{\mathcal{E}}$ перетинаються у чотирьох точках, що суперечить теоремі Безу. Аналогічно, теоремі Безу суперечить припущення, що в $\mathcal{W}_8$ існують сім точок, які лежать на одній кривій другого порядку. Доведення завершено. □

Приклад 257. Розглянемо точку $P = (1, 2)$ на еліптичній кривій (в звичайних координатах) $y^2 = x^3 - 5x + 8$ над полем $\mathbb{R}$. Використавши формули (8.6) та (8.7), отримаємо

$$2 \cdot P = P \oplus P = \left(-\frac{7}{4}, -\frac{27}{8}\right).$$

Нехай $Q = (-7/4, -27/8)$. Використавши формули (8.4) та (8.5), отримаємо

$$3 \cdot P = P \oplus Q = \left(-\frac{553}{121}, -\frac{11950}{1331}\right).$$

Аналогічно,

$$4 \cdot P = Q \oplus Q = \left(\frac{45313}{11664}, -\frac{8655103}{1259712}\right).$$

З формул (8.4), (8.5), (8.6) та (8.7) випливає такий алгоритм додавання точок еліптичної кривої.

АЛГОРИТМ ДОДАВАННЯ ТОЧОК $P_1 = (x_1, y_1)$ ТА $P_2 = (x_2, y_2)$ НА ЕЛІПТИЧНІЙ КРИВІЙ $y^2 = x^3 + ax + b$ НАД $\mathbb{R}$.

1) Якщо $P_1 \neq P_2$ та $x_1 = x_2$, повернути $P_1 \oplus P_2 = \mathcal{O}$.

2) Якщо $P_1 = P_2$ та $y_1 = 0$, повернути $P_1 \oplus P_2 = 2P_1 = \mathcal{O}$.

3) Якщо $P_1 \neq P_2$ та $x_1 \neq x_2$, покласти

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{та} \quad \nu = \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1}$$

та перейти на крок 5.

4) Якщо $P_1 = P_2$ та $y_1 \neq 0$, покласти

$$\lambda = \frac{3x_1^2 + a}{2y_1} \quad \text{та} \quad \nu = \frac{-x_1^3 + ax_1 + 2b}{2y_1}$$

та перейти на крок 5.

5) Повернути

$$P_1 \oplus P_2 = (\lambda - x_1 - x_2, -\lambda^3 + \lambda(x_1 + x_2) - \nu).$$

Зауваження 258. З алгоритму обчислення координат суми $P_1 \oplus P_2$ випливає таке важливе спостереження. Якщо параметри a, b еліптичної кривої, а також координати точок P_1, P_2 лежать в деякому полі K, то координати точки $P_1 \oplus P_2$ також лежать в полі K.

8.3 Еліптичні криві над полями характеристики $\neq 2, 3$.

Чому над полем $\mathbb{R}$ еліптична крива була визначена саме рівнянням (8.2), а не загальним поліномом третього степеня від змінних x, y? Першим кроком в поясненні цього факту є теорема Вейєрштрасса, яка стверджує, що довільну неособливу[2] криву третього порядку над довільним полем K можна привести до вигляду

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \quad a_i \in K. \tag{8.8}$$

Якщо характеристика поля K не дорівнює 2, це рівняння можна звести заміною змінної $y \mapsto y - (a_1x + a_3)/2$ до вигляду

$$y^2 = x^3 + Ax^2 + Bx + C. \tag{8.9}$$

Далі, якщо характеристика поля K не дорівнює 3, можна зробити заміну змінної $x \mapsto x - A/3$ і отримати знайому канонічну форму

$$y^2 = x^3 + ax + b.$$

Підсумовуючи, найбільш загальним виглядом еліптичної кривої над полем K є:

(i) якщо K має характеристику 2, то канонічна форма є (8.8);

[2] Алгебраїчна крива називається неособливою, якщо вона в кожній точці має дотичну. Дотичною називається пряма, точка перетину якої з заданою кривою має кратність більше за 1.

(ii) якщо K має характеристику 3, то канонічна форма є (8.9);

(iii) якщо K має характеристику $\neq 2, 3$, то канонічна форма є (8.2).

Формули (8.4), (8.5), (8.6) та (8.7), а також алгоритм обчислення суми $P_1 \oplus P_2$ з попереднього підрозділу, залишаються вірним в будь-якому полі характеристики $\neq 2, 3$. Схожі формули можна отримати і для полів характеристики 2 або 3, див. розділ 11 в [4].

Означення 259. Порядком еліптичної кривої E над полем K називається порядок групи $\overline{\mathcal{E}}(K) := (\overline{\mathcal{E}}, \oplus)$.

У криптографії використовуються еліптичні криві над скінченими полями $\mathbb{F}_q$, $q = p^n$. Очевидно, що порядок еліптичної кривої над скінченним полем є скінченним. На Рис. 8.3 представлена еліптична крива над скінченним полем $\mathbb{F}_7$.

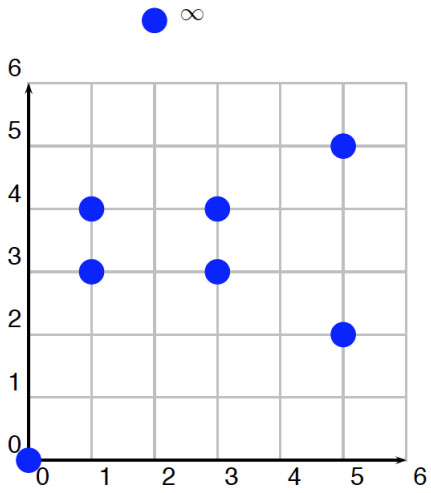

Рис. 8.3: Еліптична крива $y^2 = x^3 + x$ над полем $\mathbb{Z}_7$. Порядок цієї кривої дорівнює 8.

Задача обчислення порядку еліптичної кривої над скінченним полем є складною і загальна формула невідома. Для оцінки порядку використовується теорема Гассе, яку ми наводимо без доведення.

Теорема 260. *Нехай $\overline{\mathcal{E}}$ є еліптичною кривою над скінченним полем $\mathbb{F}_q$. Тоді*

$$||\overline{\mathcal{E}}(\mathbb{F}_q)| - q - 1| \leq 2\sqrt{q}.$$

В застосуваннях до криптографії корисиними є еліптичні криві, для яких відповідні криптосистеми є стійкими. Такими є еліптичні криві, порядок яких містить великі прості множники. Еліптичні криві, порядок яких розкладається на малі прості множники, криптографічно стійкими не є.

8.4 Криптосистема Ель-Гамаля над еліптичною кривою.

Параметри криптосистеми Ель-Гамаля над еліптичною кривою публікуються по аналогії з тим, як в класичній криптосистемі Ель-Гамаля інформація про просте поле $\mathbb{F}_q$ є загальнодоступною. До таких параметрів відносяться:

- опис скінченного поля $\mathbb{F}_q$;
- рівняння еліптичної кривої $\overline{\mathcal{E}}$ над $\mathbb{F}_q$ та її порядок;
- циклічна підгрупа G великого порядку групи $\overline{\mathcal{E}}(\mathbb{F}_q)$, а саме її порядок N та генератор P.

СХЕМА ЕЛЬ-ГАМАЛЯ НАД ЕЛІПТИЧНОЮ КРИВОЮ. Нехай Аліса хоче надіслати Бобу повідомлення t. Боб вибирає випадкове число $k \in \{1, 2, \ldots, N-1\}$ та обчислює

$$Y = k \cdot P = \underbrace{P \oplus \cdots \oplus P}_{k \text{ раз}}.$$

Четвірка $(\overline{\mathcal{E}}, N, P, Y)$ є відкритим ключем Боба. Число k є секретним ключем.

Аліса представляє повідомлення t у вигляді елемента $m \in \{1, 2, \ldots, N-1\}$, а потім елемент m представляє точкою M еліптичної кривої $\overline{\mathcal{E}}$. Використовуючи відкритий ключ Боба $(\overline{\mathcal{E}}, N, P, Y)$, Аліса обчислює шифротекст $c = (g, h)$, де

$$d = r \cdot Y, g = r \cdot P, h = M \oplus d,$$

а r є випадковим цілим числом з проміжку $\{1, 2, \ldots, N-1\}$ – «сеансовим ключем». Таке число генерується заново для кожного сеансу передачі.

Для дешифрування, використовуючи свій секретний ключ k, Боб обчислює $s = k \cdot g$, а потім $s_1 = -s$ та $M = s_1 \oplus h$ Коректність дешифрування забезпечується рівностями:

$$s_1 \oplus h = -s \oplus M \oplus d = -k \cdot g \oplus M \oplus r \cdot Y = -k \cdot r \cdot P \oplus M \oplus r \cdot k \cdot P = M.$$

8.5 Факторизація цілих чисел за допомогою еліптичних кривих.

Третім за швидкістю алгоритмом факторизації з відомих на сьогодні є алгоритм факторизації Ленстри[3] за допомогою еліптичних кривих.

[3]Гендрік Ленстра (нар. 1949) – нідерландський математик.

Нехай задане фіксоване складене натуральне число n, яке потрібно факторизувати. Розглянемо сукупність точок (x, y) кільця $\mathbb{Z}_n$, які задовольняють співвідношення

$$y^2 \equiv x^3 + ax + b (\mathrm{mod}\, n). \tag{8.10}$$

Нехай p, q – довільні два прості дільники n, тоді співвідношення (8.10) виконується також за модулями p та q. Розглянемо еліптичні криві $\overline{\mathcal{E}}_1$ та $\overline{\mathcal{E}}_2$ над простими полями $\mathbb{F}_p$ та $\mathbb{F}_q$, відповідно:

$$\overline{\mathcal{E}}_1 := \{(x, y) \in \mathbb{F}_p^2 : y^2 \equiv x^2 + ax + b (\mathrm{mod}\, p)\} \cup \{\mathcal{O}\},$$
$$\overline{\mathcal{E}}_2 := \{(x, y) \in \mathbb{F}_q^2 : y^2 \equiv x^2 + ax + b (\mathrm{mod}\, q)\} \cup \{\mathcal{O}\}.$$

Позначимо порядки $\overline{\mathcal{E}}_1$ та $\overline{\mathcal{E}}_2$ через N_p та N_q, відповідно. Нехай P – довільна точка, координати якої задовольняють співвідношення (8.10). Нехай k_p – мінімальне натуральне число таке, що

$$k_p \cdot P = \mathcal{O} \text{ на кривій } \overline{\mathcal{E}}_1.$$

Згідно з теоремою Лагранжа k_p ділить N_p. Аналогічно визначимо k_q, яке буде ділити N_q. Далі обчислюємо елементи послідовності

$$P, 2 \cdot P, 3 \cdot P, 4 \cdot P, \ldots \tag{8.11}$$

на кривих $\overline{\mathcal{E}}_1$ та $\overline{\mathcal{E}}_2$. Поки $k \cdot P$ задовольняє (8.10), $k \cdot P \neq \mathcal{O}$ на жодній з кривих. Якщо параметри a, b еліптичної кривої вибираються випадково, то з великою ймовірністю порядки N_p та N_q будуть мати багато різних простих дільників, а тому з великою ймовірністю знайдеться елемент $k \cdot P$ такий, що $k \cdot P = \mathcal{O}$ на кривій $\overline{\mathcal{E}}_1$, але не на кривій $\overline{\mathcal{E}}_2$, або навпаки. Це означає, що відповідна точка $(x_k, y_k) = k \cdot P$ більше не задовольняє (8.10), але лежить на $\overline{\mathcal{E}}_1$ або $\overline{\mathcal{E}}_2$. Припустимо, що вона лежить на $\overline{\mathcal{E}}_1$, тоді $w_k := y_k^2 - x_k^3 - ax_k - b$ ділиться на p, але не ділиться на n, а тому НСД (w_k, n) є нетривіальним дільником n.

Таким чином, алгоритм факторизації Ленстри полягає у послідовному обчисленні елементів послідовності (8.11) за модулем n. Якщо n складене, то на деякому кроці це зробити не вдасться тому, що $k \cdot P$ не буде задовольняти (8.10). З великою ймовірністю на цьому кроці буде знайдений нетривіальний дільник n.

Сформулюємо тепер послідовність кроків алгоритму Ленстри.

- Вибираємо випадкову еліптичну криву над $\mathbb{Z}_n$ вигляду $y^2 = x^2 + ax + b (\mathrm{mod}\, n)$ та вибираємо на цій кривій нетривіальну точку $P(x_0, y_0)$. Це можна зробити, наприклад, вибравши випадкові числа $x_0, y_0, a \in \mathbb{Z}_n$ та обчислюємо $b = y_0^2 - x_0^3 - ax_0 (\mathrm{mod}\, n)$.

- Вибирається деякий поріг e, наприклад $e = B!$ для деякого невеликого числа B.

- Обчислюється сума $e \cdot P$. При виконанні цієї операції на кожному кроці ми виконуємо ділення по модулю n, тобто обчислюємо НСД (v, n) для деякого $v \in \mathbb{Z}_n$. Якщо НСД $(v, n) \neq 1$, то нетривіальний дільник знайдено. У протилежному випадку ділення можна зробити і алгоритм продовжує роботу. Можливі такі варіанти:

  – якщо під час підрахунку $e \cdot P$ всі операції ділення вдалось виконати, то потрібно вибрати іншу еліптичну криву та інше початкову точку P та повторити алгоритм спочатку;

  – якщо $e_1 \cdot P = \mathcal{O}$ для деякого[4] $e_1 < e$, то потрібно вибрати іншу еліптичну криву та інше початкову точку P та повторити алгоритм спочатку;

  – якщо НСД $(v, n) \neq 1$, $v \in \mathbb{Z}_n$, то нетривіальний дільник n знайдено.

Еврістична оцінка швидкодії алгоритму Ленстри є

$$O\left(\exp\left\{(\sqrt{2} + o(1))\sqrt{\log p \log\log p}\right\}\right), \quad n \to \infty,$$

де p є найменшим простим дільником n. Якщо $p \approx \sqrt{n}$ (наприклад, як в схемі RSA), то ця оцінка співпадає з оцінкою (7.4) для алгоритму квадратичного решета.

[4]Це означає, що $e \cdot P = (e \mathrm{mod}\, e_1) \cdot P$, а тому послідовність $(k \cdot P)_{k \geq 0}$ буде періодичною з періодом, що менший за e.

Лекція 9

Додаток

9.1 Алгоритм Флойда пошуку циклів.

Нехай A є деякої фіксованою скінченною множиною, а f є довільним відображенням A на себе (не обов'язково бієктивним). Розглянемо послідовність елементів множини A

$$A \ni x_0, \quad x_1 = f(x_0), \quad x_2 = f(x_1), \quad x_3 = f(x_2), \ldots.$$

В силу припущення про скінченність A ця послідовність буде циклічною з деякого місця, тобто знайдуться $\lambda \in \mathbb{N}_0$ та $\tau \in \mathbb{N}$ такі, що

$$x_{\tau+n} = x_n, \quad n \geq \lambda. \tag{9.1}$$

Елементи $x_0, x_1, \ldots, x_{\lambda-1}$ називаються **підходом до цикла**, λ – **довжиною підхода**, найменше τ, для якого виконуються рівності (9.1) – **довжиною цикла**.

Алгоритм Флойда пошуку довжини цикла τ базується на простому спостереженні: якщо виконується рівність $x_n = x_{2n}$, то τ ділить n. Алгоритм знаходить індекс ν – мінімальний індекс n для якого $x_n = x_{2n}$. Після цього алгоритм починає перегляд послідовності спочатку і знаходить індекс λ такий, що $x_\lambda = x_{\lambda+\nu}$. Як тільки таке λ знайдене, τ знаходиться як мінімальний індекс i такий, що $x_{\lambda+i} = x_\lambda$.

Бібліографія

[1] L. ADLEMAN (1979). A subexponential algorithm for the discrete logarithm problem with applications to cryptography. In *20th Annual Symposium on Foundations of Computer Science*.

[2] W. R. ALFORD, A. GRANVILLE, C. POMERANCE (1994). There are Infinitely Many Carmichael Numbers. *Annals of Mathematics*, **139**, p. 703–722. Доступно за посиланням `https://www.math.dartmouth.edu/~carlp/PDF/paper95.pdf`

[3] T. APOSTOL (1976). *Introduction to Analytic Number Theory*. Springer Verlag.

[4] С. М. АВДОШИН, А. А. НАБЕБИН (2017). Дискретная математика: модулярная алгебра, криптография, кодирование. М.: ДМК Пресс.

[5] О. Н. ВАСИЛЕНКО (2003). Теоретико-числовые алгоритмы криптографии. М.:МЦНМО.

[6] J. FRALEIGH (1982). *A First Course in Abstract Algebra,* 3^{rd} *ed.* Addison-Wesley Publishing Company.

[7] K. IRELAND, M. ROSEN (1990). *A Classical Introduction to Modern Number Theory,* 2^{nd} *ed.* Springer Verlag.

[8] W. FULTON (2008). *Algebraic Curves,* 3^{rd} *ed.* Addison-Wesley Publishing Company.

[9] A. LENSTRA (2000). Integer factoring. *Designs, Codes and Cryptography*, **19**, p. 101 – 128. Доступно за посиланням `http://www.woodmann.com/yates/Cryptography/arjen_lenstra_factoring.pdf`

[10] M. O. RABIN (1979). Digitalized Signatures and Public Key Functions as Intractable as Factoring. Доступно за посиланнямhttp://publications.csail.mit.edu/lcs/pubs/pdf/MIT-LCS-TR-212.pdf

[11] M. O. RABIN (1980). Probabilistic algorithm for testing primality. *J. Numb. Theor.*, **12**, p. 128–138.

[12] C. STUDHOLME (2002). The discrete Log problem. Доступно за посиланням http://www.cs.toronto.edu/~cvs/dlog/research_paper.pdf.