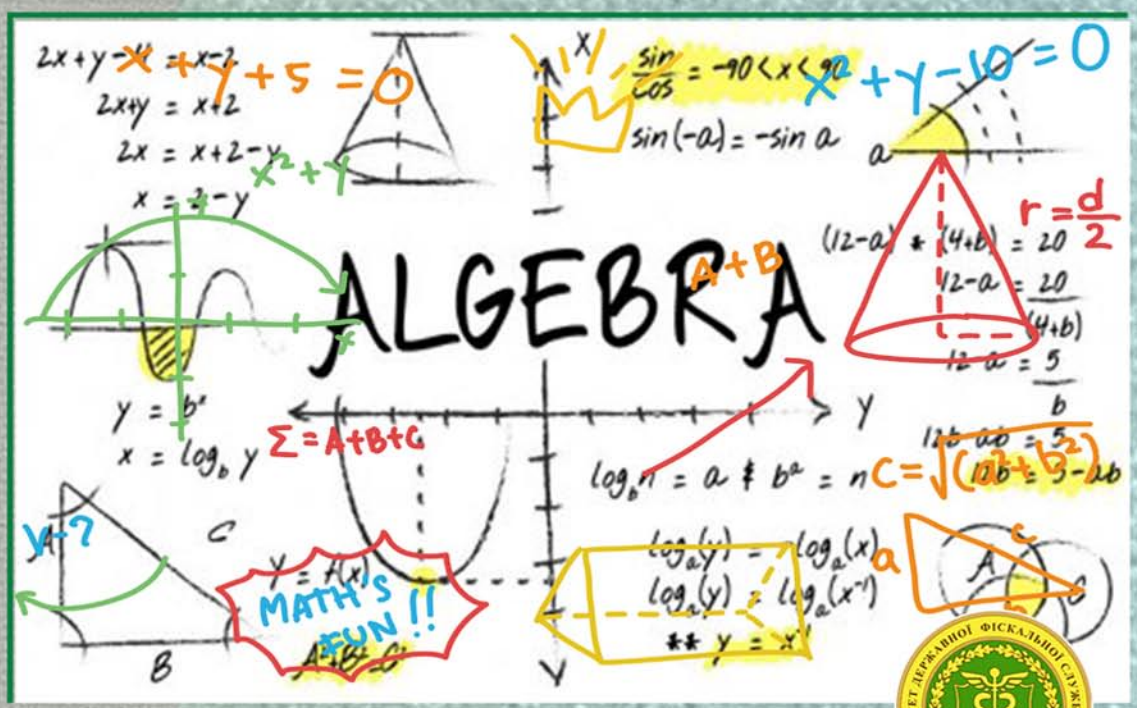
НА ДОПОМОГУ СТУДЕНТУ УДФСУ

100

М. М. Семко,
О. А. Ярова,
Л. В. Скасків

ОСНОВНІ ПОНЯТТЯ СУЧАСНОЇ АЛГЕБРИ

ОСНОВНІ ПОНЯТТЯ СУЧАСНОЇ АЛГЕБРИ



УНІВЕРСИТЕТ ДЕРЖАВНОЇ ФІСКАЛЬНОЇ СЛУЖБИ УКРАЇНИ
1921

СЕРІЯ «НА ДОПОМОГУ СТУДЕНТУ УДФСУ»

Серію «На допомогу студенту УДФСУ» засновано 2016 року.

У СЕРІЇ «НА ДОПОМОГУ СТУДЕНТУ УДФСУ» ВИЙШЛИ ДРУКОМ:

2016

«Методичні основи спеціальної фізичної та технічної підготовки студентів за розділом «Легка атлетика»
«Самостійна робота студента як одна з форм впливу на функціональну, фізичну та психологічну підготовленість»
«Організація роботи командира механізованого взводу»
«5.45-мм автомати Калашникова (АК-74, АКС-74, АК-74Н, АКС-74Н) та 5.45-мм ручні кулемети Калашникова (РПК-74, РПКС-74, РПК-74Н, РПКС-74Н)»
«Гранатомет підствольний ГП-25»
«Ручні гранати»
«Кулемети Калашникова – 7.62, ПК, УЖМ, ПКТ»
«Ручний протитанковий гранатомет РПГ-7»
«9-мм пістолет Макарова (ПМ)»

2017

«Вища та прикладна математика»
«Цивільний захист»
«Програмування мовою JAVA : практикум»
«Інформаційні системи і технології в юридичній практиці»
«Дослідження операцій : практикум»
«Чисельні методи»
«English for Students of Finance»
«Основи військової розвідки»

2018

«CASE-технології. Міждисциплінарне інформаційне моделювання»
«Економічна інформатика: практикум»
«Економічна теорія (політекономія, мікроекономіка, макроекономіка). Політекономія»
«Економічна теорія (політекономія, мікроекономіка, макроекономіка). Мікроекономіка»
«Економічна теорія (політекономія, мікроекономіка, макроекономіка). Макроекономіка»
«Охорона праці»
«Економіка і організація діяльності об'єднань підприємств»
«Основи християнської культури»
«Економіка підприємства»
«Фізика»
«Трудове право України»

2019

«Основи тактичної медицини»

«Аудит»

«Збірник задач. Вища та прикладна математика»

«Міжнародні розрахунки та валютні операції»

«Підготовка озброєння механізованого взводу до бойового застосування»

«Контролінг в управлінні підприємством»

«Актуальні питання судової експертизи (у питаннях і відповідях)»

«Інноваційний менеджмент»

«Організація навчальних занять з фізичного виховання за методом колового тренування»

«Теорія судових доказів (у таблицях і схемах)»

«Статистика»

«Фінансовий менеджмент проектів і програм»

«Ручний протитанковий гранатомет РПГ-7 (РПГ-7Д)»

«Організація роботи командира механізованого відділення»

«Психологія управління»

«Deutsch und Wirtschaft»

«Основи тактичної підготовки працівників правоохоронних органів»

«Основи кінології»

«Моделювання систем»

«UML. Уніфікована мова моделювання інформаційних систем»

«Історія держави і права України»

«Економічна теорія»

2020

«Менеджмент»

«Бізнес-аналітика та моделювання»

«Безпека життєдіяльності»

«Статистичний аналіз даних»

«Хімія»

«Мікробіологія»

«Медичне право»

«Бухгалтерський облік в управлінні підприємством»

«Загальна психологія»

«Хортинг у закладах вищої освіти»

«Теорія і методика викладання оздоровчого фітнесу»

«Організація охорони державної таємниці в Україні»

«Цивільне право України (особлива частина) »

«Цивільний процес України»

«Оздоровчий туризм»

«English Grammar for Philology Students»

«Фітнес-технології фізичного виховання у закладах вищої освіти»

«Конкурентна розвідка»

«Теорія та методика бойового хортингу»

«Логіка»

«Інституціональна економіка»

«Економетрика: прикладний аспект»

«Основні поняття сучасної алгебри»

УНІВЕРСИТЕТ ДЕРЖАВНОЇ ФІСКАЛЬНОЇ СЛУЖБИ УКРАЇНИ

СЕРІЯ «НА ДОПОМОГУ СТУДЕНТУ УДФСУ»
Заснована 2016 року

М. М. Семко,
О. А. Ярова,
Л. В. Скасків

ОСНОВНІ ПОНЯТТЯ СУЧАСНОЇ АЛГЕБРИ

Навчальний посібник

Ірпінь
2020

УДК 512.2
ББК 22.14
С 30

Рекомендовано до друку Вченою радою
Університету ДФС України
(протокол № 12 від 29 жовтня 2020 р.)

Рецензенти:
Бондаренко В. М., доктор фізико-математичних наук, професор (Інститут математики НАН України);
Требенко Д. Я., кандидат фізико-математичних наук, доцент (Національний педагогічний університет імені М.П. Драгоманова).

Семко М. М.

С 30 **Основні поняття сучасної алгебри** / М. М. Семко, О. А. Ярова, Л. В. Скасків. – Ірпінь : Університет ДФС України, 2020. – 128 с. – (Серія «На допомогу студенту УДФСУ» ; т. 81).
ISBN 978-966-337-608-0

У навчальному посібнику викладані основні базові поняття, результати та методи сучасного курсу алгебри, що необхідні для подальшого більш глибокого вивчення курсу алгебри.

Призначено для студентів математичного спрямування, викладачів та тих, кого цікавлять питання сучасної алгебри.

УДК 512.2
ББК 22.14

ISBN 978-966-337-608-0

ЗМІСТ

РОЗДІЛ 1
БІНАРНІ АЛГЕБРАЇЧНІ ОПЕРАЦІЇ, ЇХ ВЛАСТИВОСТІ

У математиці дуже мало понять, які були б більш первісні, ніж поняття бінарної алгебраїчної операції. Ми вже маємо досить багатий набір конкретних алгебраїчних операцій, досить багато знаємо про їх властивості. Ми маємо можливість осмислити цей конкретний матеріал з більш загальних позицій. Те загальне, що мають розглянуті вже операції, можна виділити у такому визначенні.

Нехай M – множина. *Бінарною алгебраїчною операцією (або законом композиції) на множині M будемо називати відображення*

$$f\colon M \times M \longrightarrow M$$

декартова квадрата множини M у цю ж множину. Таким чином, кожній впорядкованій парі (a, b) елементів множини M ставиться у відповідність однозначно визначений елемент f(a, b) тієї ж множини M. Елемент f(a, b) називається композицією елементів a і b щодо цієї операції.

Для запису композиції елементів a і b частіше всього виписують a та b у відповідному порядку, відділяючи їх спеціальним характеристичним знаком цієї операції, наприклад, a + b. З позначень знаків бінарних алгебраїчних операцій, які використовуються найбільш часто, вкажемо «+» та «·». Закон, що позначається знаком «+», називається *додаванням*, а відповідна композиція елементів a + b називається їх *сумою*; у цьому випадку говорять також про *адитивне* позначення бінарної алгебраїчної операції. Закон, що позначається знаком «·», називається *множенням*, а відповідна композиція елементів a·b називається їх *добутком*; у цьому випадку говорять також про *мультиплікативне* позначення бінарної алгебраїчної операції. В останньому випадку часто крапка опускається, тобто замість a·b пишуть просто ab. Слід пам'ятати, що у більшості випадків ці позначення досить умовні. Розглянемо деякі приклади бінарних алгебраїчних операцій.

1) додавання на множинах цілих, раціональних та дійсних чисел;

2) множення на множинах цілих, раціональних та дійсних чисел;

3) нехай M – множина. **P**(M) – множина всіх перетворень множини M, тоді композиція f ☺ g дає приклад бінарної алгебраїчної операції на множині **P**(M);

4) нехай M – множина, **B** (M) – його булеан. Відображення $(X, У) \longrightarrow X \cap У$, $(X, У) \longrightarrow X \cup У$, $(X, У) \longrightarrow X \setminus У$, $(X, У) \longrightarrow X \triangle У$, $X, У \subseteq M$, дають приклади бінарних алгебраїчних операцій на множині **B**(M);

5) додавання, множення та комутування на множині матриць $\mathbf{M}_n(R)$;

6) додавання та множення дійсних функцій (тобто перетворень множини R);

7) відображення $(n, k) \longrightarrow n^k$, $n, k \in N$, дає приклад бінарної алгебраїчної операції на множині натуральних чисел;

8) відображення $(n, k) \longrightarrow$ **НСД**(n, k), $(n, k) \longrightarrow$ **НСК**(n, k), $n, k \in Z$, дають приклади бінарних алгебраїчних операцій на множині цілих чисел (при цьому ми вважаємо, що **НСД**(n, k), **НСК**(n, k) ≥ 0);

9) додавання та векторне множення у просторі R^3.

Розглянемо деякі важливі властивості бінарних алгебраїчних операцій. При цьому будемо користуватись мультиплікативною формою запису операції і робити ілюстрації для адитивної форми запису.

Операція на множині M називається комутативною, якщо $ab = ba$ для кожної пари елементів $a, b \in M$.

Для адитивної форми запису це виглядає так: a + b = b + a.

Комутативними будуть операції додавання та множення на множинах цілих, раціональних та дійсних чисел; операції $\cap$, $\cup$ та $\triangle$ на булеані B(M); додавання матриць**;** додавання та множення дійсних функцій; операції **НСД**, **НСК** на Z; додавання векторів у просторі R^3.

Якщо маємо три елементи a, b, c $\in$ M, то з них можливо утворити наступні два добутки (якщо не змінювати порядок їх слідування): a(bc) та (ab)c.

Операція на множині M називається асоціативною, якщо

$$(ab)c = a(bc)$$

для кожної трійки елементів $a, b, c \in M$.

У адитивному запису: (a + b) + c = a + (b + c).

Для чотирьох елементів a, b, c, d добутків уже значно більше:

((ab)c)d, (ab)(cd), (a(bc))d, a(b(cd)), a((bc)d). Інакше кажучи, кожному добутку елементів a, b, c, d відповідає деяке розташування дужок між ними. Якщо операція асоціативна, то отримуємо, наприклад, ((ab)c)d = (ab)(cd). І така рівність має місце і в більш загальному випадку, як показує наступна теорема.

1.1. Теорема. *Якщо бінарна операція на множині M є асоціативною, то результат її послідовного застосування до кожного скінченого набору елементів множини M не залежить від розташування дужок між ними (порядок слідування елементів не змінюється!).*

Д о в е д е н н я. Розглянемо довільний скінчений набір елементів $a_1, a_2, \ldots, a_n$ множини M. Скористаємось індукцією за числом n. Якщо n = 1, 2, то доводити нічого, для n = 3 твердження випливає з закону асоціативності. Припустимо тепер, що n > 3 та твердження вже має місце для скінчених наборів, кількість елементів у яких менша ніж n. Покажемо, що кожний добуток елементів $a_1, a_2, \ldots, a_n$ (у цьому порядку слідування) збігається з деяким фіксованим добутком. Виділимо так званий *лівонормований добуток*:

$$(\ldots(((a_1\, a_2)\, a_3)\, a_4)\ldots)\, a_{n-1}))\, a_n.$$

Розглянемо довільний добуток елементів $a_1, a_2, \ldots, a_n$ (у цьому порядку слідування). При цьому будемо виписувати тільки

зовнішні дужки, оскільки для добутків, що складаються з меншої ніж n кількості елементів, твердження теореми є правильним і тому внутрішнє конкретне розташування дужок не має значення.

Отже, розглянемо добуток $(a_1 \ldots a_t)(a_{t+1} \ldots a_n)$. Мають місце такі два випадки: $t + 1 = n$ та $t + 1 < n$.

У першому випадку кожний добуток елементів $a_1, a_2, \ldots, a_{n-1}$ (у такому порядку слідування) збігається за індуктивною гіпотезою з їх лівонормованим добутком, а тому помноживши справа на елемент a_n отримуємо лівонормований добуток елементів $a_1, a_2, \ldots, a_n$ (у даному порядку слідування). Розглянемо тепер другий випадок. Застосовуючи закон асоціативності, будемо мати:

$$(a_1 \ldots a_t)(a_{t+1} \ldots a_n) = (a_1 \ldots a_t)((a_{t+1} \ldots a_{n-1})a_n) =$$
$$= ((a_1 \ldots a_t)(a_{t+1} \ldots a_{n-1}))a_n = (\ldots(((a_1 a_2)a_3)a_4)\ldots)a_{n-1}))a_n.$$

З теореми 1.1 випливає, що кожне розташування дужок при знаходженні добутку елементів $a_1, a_2, \ldots, a_n$ (у цьому порядку слідування) приводить до одного й того ж результату. Тому далі взагалі не будемо писати ніяких дужок і записувати добуток елементів $a_1, a_2, \ldots, a_n$ (у нашому порядку слідування) просто як $a_1 a_2 \ldots a_n$, зважаючи на ту обставину, що у кожній конкретній ситуації можливо розташувати дужки так, як це потрібно для даної конкретної ситуації.

Для скороченого запису добутку елементів $a_1, a_2, \ldots, a_n$ (слідування) будемо використовувати також таке позначення:

$$a_1 a_2 \ldots a_n = \prod_{1 \le i \le n} a_i.$$

У випадку, коли $a_1, a_2, \ldots, a_n = a$, добуток $a_1 a_2 \ldots a_n$ позначають через a^n і називають n-м ступенем елемента a.

1.2. Наслідок. *Якщо бінарна операція на множині M є асоціативною, то для кожного елемента $a \in M$ та довільних $n,m \in \mathbf{N}$ мають місце рівності:*

$$a^n a^m = a^{n+m}, (a^n)^m = a^{nm}.$$

У випадку, коли використовується адитивний запис для бінарної операції, замість $\prod_{1 \le i \le n} a_i$ використовується уже звичний для нас запис $\sum_{1 \le i \le n} a_i$, замість ступеню елемента говорять про його кратне:

$$na = \underbrace{a + \ldots + a}_{n}.$$

Твердження наслідку 1.1 для адитивного запису має вигляд:

$$na + ma = (n + m)a, \quad m(na) = (mn)a.$$

Елементи a, b називаються переставними або комутуючими, якщо $ab = ba$.

Для переставних елементів має місце рівність $(ab)^n = a^n b^n$ при довільному $n \in N$. Дійсно,

$$(ab)^2 = ab\ ab = a(ba)b = a(ab)b = (aa)(bb) = a^2 b^2,$$

а далі індукцією за числом n:

$$(ab)^n = (ab)^{n-1}(ab) = (a^{n-1} b^{n-1})ab = a^{n-1}(b^{n-1}a)b = a^{n-1}(b^{n-2}(ba))b =$$
$$= a^{n-1}b^{n-2}(ab)b = a^{n-1}(b^{n-2}a)\, b^2 = \ldots = a^{n-1}ab\, b^{n-1} = a^n\, b^n.$$

Використовуючи аналогічні аргументи, можна довести і більш загальне твердження.

1.3. Пропозиція. *Нехай на множині M задана бінарна асоціативна операція. Якщо $a_1, a_2, \ldots, a_n$ – такі елементи множини M, що $a_i a_j = a_j a_i$ $i, j, 1 \le i, j \le n$, то*

$$(a_1 a_2 \ldots a_n)^m = a_1^m a_2^m \ldots a2_n^m$$

для довільного $m \in \boldsymbol{N}$.

У адитивному запису ця рівність має вигляд:

$$m(a_1 + a_2 + \ldots + a_n) = ma_1 + ma_2 + \ldots + ma_n.$$

Нехай на множині M задана бінарна операція. *Елемент $z \in M$ називається центральним, якщо він переставний з кожним елементом множини M. Сукупність усіх центральних елементів називається центром M і буде позначатися через $\zeta(M)$.*

Непуста множина S, *на якій задана бінарна асоціативна операція, називається напівгрупою. Напівгрупа називається комутативною, якщо її операція ще й комутативна.*

Наведемо зараз приклади напівгруп.

Множини цілих, раціональних та дійсних чисел відносно операції додавання утворюють комутативну напівгрупу.

Множини цілих, раціональних та дійсних чисел відносно операції множення утворюють комутативну напівгрупу.

Нехай M – множина. Множина **P**(M) всіх перетворень множини M буде напівгрупою відносно композиції f ☺ g перетворень. Як ми вже бачили раніше, ця напівгрупа не є комутативною.

Нехай M – множина. Булеан **B**(M) цієї множини буде комутативною напівгрупою відносно операцій $(X, У) \longrightarrow X \cap У$, $(X, У) \longrightarrow X \cup У$, $(X, У) \longrightarrow X \setminus У$, $(X, Y) \longrightarrow X \triangle У$, $X, У \subseteq M$, за теоремою 1.1.

Множина квадратних матриць $\mathbf{M}_n(R)$ буде комутативною напівгрупою відносно операції додавання та напівгрупою відносно операції множення матриць.

Множина всіх дійсних функцій (тобто перетворень множини R) буде комутативною напівгрупою відносно звичайних операцій додавання та множення дійсних функцій.

Множина всіх цілих чисел буде комутативною напівгрупою відносно операцій $(n, k) \longrightarrow$ **НСД**(n, k), $(n, k) \longrightarrow$ **НСК**(n, k), $n, k \in Z$.

Векторний простір R^3 буде комутативною напівгрупою відносно операції додавання векторів.

Наведемо також ще один важливий приклад.

Нехай M – непуста множина, яку будемо називати *алфавітом*, а елементи цієї множини будемо називати *буквами*. Через $\mathbf{F_M}$ позначимо множину всіх скінчених наборів елементів множини M. Визначимо операцію на $\mathbf{F_M}$ за таким правилом:

$$(a_1, a_2, \ldots, a_n)(b_1, b_2, \ldots, b_m) = (a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_m).$$

Ця операція, яка називається ще конкатенацією, очевидно асоціативна. Отже, відносно цієї операції $\mathbf{F_M}$ стає напівгрупою.

$\mathbf{F_M}$ називають *вільною напівгрупою над алфавітом M*. Ми можемо ототожнити набір, що складається з одного елемента, з самим цим елементом, і отримати більш звичайний запис набору $(a_1, a_2, \ldots, a_n)$ у вигляді добутку $a_1\, a_2 \ldots a_n$. У цьому більш звичному вигляді і записують елементи вільної напівгрупи $\mathbf{F_M}$. Елемент $w = a_1\, a_2 \ldots a_n$ вільної напівгрупи $\mathbf{F_M}$ природно називати *словом*, а число n – *довжиною* цього слова. При цьому два слова $a_1\, a_2 \ldots a_n$ та $b_1\, b_2 \ldots b_m$ тоді і тільки тоді будуть рівними, коли $n = m$ та $a_t = b_t$ для кожного $1 \le t \le m$.

Нехай на множині M задана бінарна операція. *Елемент* $e \in M$ *називається нейтральним елементом відносно даної операції, якщо*

$$ae = ea = a$$

для кожного елементу a множини M.

Нейтральний елемент, якщо він існує, буде єдиним. Якщо операція на множині M записується мультиплікативно, то частіше використовується термін одиничний елемент; якщо операція записується адитивно, то говорять про нульовий елемент, у цьому випадку будемо позначати його через 0_M, так що визначення нульового елемента виглядає так: $a + 0_M = 0_M + a = a$ для кожного елемента $a \in M$.

Відносно операції додавання на множинах цілих, раціональних та дійсних чисел нульовим елементом буде число 0.

Відносно операції множення на множинах цілих, раціональних та дійсних чисел одиничним елементом буде число 1.

Нехай M – множина. **P**(M) – множина всіх перетворень множини M. Композиція f ☺ g має одиничний елемент, ним буде тотожня підстановка ε_M.

Нехай M – множина, **B**(M) – його булеан. Операції $\cap$, $\cup$, $\triangle$ також мають нейтральні елементи. Для операції $\cap$ ним буде множина M, а два двох інших операцій нейтральним елементом буде пуста множина.

Відносно операції додавання на множині матриць $\mathbf{M}_n(R)$ нульовим елементом буде нульова матриця O, відносно операції множення матриць одиничним елементом буде матриця E.

Відносно операції додавання дійсних функцій нульовим елементом буде функція, що дорівнює 0 для кожного значення аргументу; відносно операції множення дійсних функцій одиничним елементом буде функція, що дорівнює 1 для кожного значення аргументу.

Відносно операції (n, k) $\longrightarrow$ **НСД**(n, k) на множині цілих чисел нейтральним елементом буде число 1. Операція (n, k) $\longrightarrow$ **НСК**(n, k), n, k $\in$ Z, не має нейтрального елемента.

У просторі R^3 відносно операції додавання нульовим елементом буде нульовий вектор.

Напівгрупа, що має одиничний елемент, називається напівгрупою з одиницею.

Нехай на множині M задана бінарна операція. *Підмножина S множини M називається сталою відносно операції, якщо для кожної пари елементів $a, b \in S$ їх композиція ab також належить до S.*

Це означає, що звуження бінарної операції, заданої на множині M, на її підмножину S буде бінарною операцією на підмножині S. Наприклад, підмножина всіх парних чисел буде сталою підмножиною Z відносно операції додавання, водночас, як множина всіх непарних чисел не буде сталою відносно цієї операції.

1.4. Пропозиція. *Нехай на множині M задана бінарна операція. Якщо* S *– система сталих підмножин множини M, то перетин* $\cap$ S *підмножин цієї системи також буде підмножиною, сталою відносно даної операції.*

Це твердження очевидне.

Зазначимо, що об'єднання сталих підмножин не буде сталою підмножиною. Наприклад, підмножина 2Z всіх парних цілих чисел та підмножина 3Z всіх цілих чисел кратних 3 будуть сталими відносно операції додавання, а їх об′єднання не містить числа 5 = 2 + 3 і, отже, не буде сталою підмножиною.

Нехай C – довільна підмножина множини M, S – система сталих відносно операції підмножин, кожне з яких містить C. Тоді перетин $\cap$ S буде найменшою сталою підмножиною, яка містить у собі C. *Ця підмножина називається сталою підмножиною, породженою підмножиною C.*

Нехай на множині M задана бінарна операція, яка має одиничний елемент *e*. *Елемент a множини M має обернений, якщо існує такий елемент x (він і називається оберненим до a), що ax = xa = e.*

Якщо операція асоціативна, то обернений до a елемент буде єдиним. Дійсно, якщо y – елемент, що також має властивість ay = ya = e, то

$$y = ey = (xa)y = x(ay) = xe = x.$$

Тому далі єдиний обернений до a елемент будемо позначати через a^{-1}. З самого визначення видно, що $(a^{-1})^{-1} = a$.

Якщо операція записується адитивно, то замість елемента, оберненого до a, будемо говорити про елемент; *протилежний до a*, цей елемент будемо позначати через -a; отже визначення протилежного елементу має вигляд:

$$a + (-a) = (-a) + a = 0_M.$$

1.5. Пропозиція. *Нехай на множині M задана бінарна асоціативна операція, яка має одиничний елемент e. Якщо елементи $a_1, a_2, \ldots, a_n$ мають обернені у множині M, то їх добуток $a_1a_2 \ldots a_n$ також має обернений і:*

$$(a_1a_2 \ldots a_n)^{-1} = a_n^{-1}a_{n-1}^{-1} \ldots a_1^{-1}.$$

Дійсно, маємо:

$$(a_1a_2 \ldots a_n)(a_n^{-1}a^{-1}{}_{n-1} \ldots a_1^{-1}) = (a_1a_2 \ldots a_{n-1})(a_n a_n^{-1})(a^{-1}{}_{n-1}^{-1} \ldots a_1^{-1}) =$$
$$= (a_1a_2 \ldots a_{n-1})(a^{-1}{}_{n-1} \ldots a_1^{-1}) = \ldots = e.$$

1.6. Наслідок. *Нехай P – напівгрупа з одиницею. Тоді підмножина* U(P) *всіх тих елементів, що мають обернені, буде сталою.*

Наявність одиничного елемента та елемента, оберненого до a, дозволяє визначити всі цілі ступені цього елемента. Наприклад:

$$a^0 = e,\ a^{-n} = (a^{-1})^n,\ n \in N.$$

У адитивному запису має вигляд: $0a = 0_M$, $(-n)a = n(-a)$.

1.7. Пропозиція. *Якщо бінарна операція на множині M є асоціативною, то для кожного елемента $a \in M$ та довільних $n,m \in Z$ мають місце рівності:*

$$a^n a^m = a^{n+m},\ (a^n)^m = a^{nm}.$$

Д о в е д е н н я. Якщо n, m > 0, то рівності вже доведені у наслідку 9.2. Якщо хоча б одне з цих чисел дорівнює нулю, то рівності тривіальні. Нехай n, m < 0, тоді

$$a^n a^m = (a^{-1})^{-n}(a^{-1})^{-m} = (a^{-1})^{-n+(-m)} = (a^{-1})^{-(n+m)} = a^{n+m},$$
$$(a^n)^m = ((a^n)^{-1})^{-m} = (((a^{-1})^{-n})^{-1})^{-m} = ((a^{-1})^n)^{-m} = ((a^{-n}))^{-m} = a^{(-n)(-m)} = a^{nm}.$$

Нехай тепер n > 0, m < 0, n >-m, тоді

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

Якщо n > 0, m < 0, n < -m, тоді

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

Розглянемо тепер другу рівність. Нехай тепер n > 0, m < 0, тоді

$$(a^n)^m = ((a^n)^{-1})^{-m} = ((a^{-1})^n)^{-m} = ((a^{-1})^{n(-m)} = (a^{-1})^{-nm} = a^{-(-nm)} = a^{nm}.$$

Якщо n < 0, m > 0, то

$$(a^n)^m = (((a^{-1})^{-n})^m = (a^{-1})^{(-n)m} = (a^{-1})^{-nm} = a^{-(-nm)} = a^{nm}.$$

Нехай M, S – множини, на яких задані бінарні алгебраїчні операції, що позначаються символами □ та ✦ відповідно, та нехай $f : M \longrightarrow S$ – відображення. *Це відображення називають гомоморфізмом, якщо*

$$f(x \square y) = f(x) \ \✦\ f(y)$$

для довільних елементів $x, y \in M$.

Про відображення f говорять ще, що воно зберігає операції або узгоджено з операціями.

Ін'єктивний гомоморфізм називається мономорфізмом, сюр'єктивний гомоморфізм називається епіморфізмом, і бієктивний гомоморфізм називається ізоморфізмом.

Якщо $f : M \longrightarrow S$ – ізоморфізм, то з теореми 3.3 випливає існування оберненого відображення $f^{-1} : S \longrightarrow M$. Якщо u, v – довільні елементи множини S, то $u = f(x)$, $v = f(y)$ для деяких елементів $x, y \in M$. Тепер

$$f^{-1}(u \star v) = f^{-1}(f(x) \star f(y)) = f^{-1}(f(x \square y)) = x \square y = f^{-1}(u) \square f^{-1}(v).$$

Це означає, що відображення $f^{-1} : S \longrightarrow M$ також буде ізоморфізмом.

Множини M, S, на яких задані бінарні алгебраїчні операції, називаються ізоморфними, якщо існує ізоморфізм однієї з них на іншу; цей факт будемо позначати так $\boldsymbol{M \cong S}$.

Якщо M – множина, на якій задано бінарну алгебраїчну операцію, то її вивчення має дві сторони. Перша відповідає на питання, які об'єкти є елементами множини M, а друга вивчає властивості операції на цих об'єктах.

Відповідно до цього і вивчення M можна проводити з різних точок зору. Можна вивчати зв'язки між індивідуальними властивостями елементів та їх підмножин і їх властивостями по відношенню до операції. Такий підхід є присутнім при вивченні конкретних множин, таких, наприклад, як груп підстановок, груп перетворень площини та простору, груп симетрій та ін.

Однак можна вивчати і ті властивості, що цілком визначаються через властивості операцій. Цей підхід є характерним для загальної алгебри, він і приводить до важливого поняття ізоморфізму. Він виник досить давно, але входив у алгебру дуже повільно. Ще Р. Декарт говорив про суттєву спільність математичних дисциплін: "Хоча їх об'єкти різні, вони залишаються узгодженими між собою в тому, що розглядають у цих об'єктах тільки різні співвідношення або пропорції, які в них знаходяться".

Уточнюючи цю узгодженість, Г. Лейбніц вводить у розгляд загальне поняття ізоморфії (він називає це “подібність”) і можливість ототожнювати ізоморфні співвідношення або операції. Він наводить приклад ізоморфізму, який став класичним. Це відображення $x \longrightarrow \log x$ множини всіх додатних дійсних чисел з операцією множення на множину всіх дійсних чисел з операцією додавання.

Поняття ізоморфізму було знайомо Е. Галуа для груп підстановок. Але у загальному вигляді це поняття з’являється у теорії груп у середині XIX століття. Елементи множин M, S, що відповідають один одному при ізоморфізмі, мають однакові властивості відносно операцій. Загальна алгебра вивчає лише ті властивості, які зберігаються при ізоморфізмах.

РОЗДІЛ 2
ГРУПИ

У попередньому розділі ми вже привели всі поняття, необхідні для того, щоб дати одне *з фундаментальних понять алгебри – поняття групи.*

Напівгрупа з одиницею G, кожний елемент якої має обернений, називається групою. Інакше кажучи, група – це множина G, на якій задана бінарна алгебраїчна операція $(x, y) \longmapsto xy$, $x, y \in G$, *яка задовольняє наступним умовам (аксіомам групи):*

(G 1*) операція асоціативна, тобто* $x(yz) = (xy)z$ *для всіх елементів* $x, y, z \in G$;

(G 2*) G має одиничний елемент e , тобто елемент, що має властивість* $xe = ex = x$ *для кожного* $x \in G$;

(G 3*) кожен елемент* $x \in G$ *має обернений, тобто такий елемент* x^{-1}, *що* $xx^{-1} = x^{-1}x = e$.

Якщо операція в групі комутативна, то група називається абелевою (на честь норвезького математика Абеля). Як правило, у абелевій групі використовується адитивний запис операції, так що аксіоми абелевої групи мають такий вигляд:

(AG 1) операція комутативна, тобто $x + y = y + x$ *для всіх елементів* $x, y \in G$;

(AG 2) операція асоціативна, тобто $x + (y + z) = (x + y) + z$ *для всіх елементів* $x, y, z \in G$;

(AG 3) G має нульовий елемент 0_G *, тобто елемент, що має властивість* $x + 0_G = x$ *для кожного* $x \in G$;

(AG 4) кожен елемент $x \in G$ *має протилежний, тобто такий елемент* $-x$, *що* $x + (-x) = = 0_G$.

Самий термін “група” належить видатному французькому математику Евариcту Галуа. Серія досліджень, що приводить до поняття групи, це “теорія підстановок”, яка розвинулась з робіт

Лагранжа, Вандермонда, Гаусса, Руффіні, Коші, присвячених знаходженню коренів алгебраїчних рівнянь. У цих дослідженнях були отримані перші результати теорії груп.

Але реальним засновником теорії груп вважається Галуа, який звів вивчення алгебраїчних рівнянь до вивчення пов'язаних з ними груп підстановок, він перший ввів поняття нормальної підгрупи та зрозумів його важливість, він розглянув групи, що мають спеціальні властивості, йому належить ідея "лінійного представлення" груп, і цей факт показує, що він підійшов до поняття ізоморфізму.

Геніальні роботи Галуа довгий час залишались незрозумілими, розповсюдження його ідей проходило значно пізніше його трагічної загибелі у роботах Ж. Серре та "Трактаті про підстановки" К. Жордана. Досягнення теорії Галуа стимулювали інтенсивне вивчення скінченних груп підстановок, і на першому етапі розвиток теорії груп проходив спочатку у рамках теорії скінченних груп підстановок. Визначення "абстрактних" груп було дано А. Келі у середині XIX століття.

Однак навіть дослідження скінчених абстрактних груп досить довгий час сприймались як частина теорії підстановок. І лише у кінці XIX століття почався усвідомлений розвиток автономної теорії груп. Під впливом геометрії, топології, теорії диференціальних рівнянь виникла необхідність вивчення нескінченних груп перетворень. Відомий німецький геометр Ф. Клейн зв'язує геометрії з відповідними групами перетворень.

Ідея групи та пов'язані з нею ідеі інваріантності і симетрії усвідомлювалась досить довгий час. Усюди, де суттєву роль грає симетрія якогось об'єкта (алгебраїчного або диференціального рівняння, кристалічної решітки або геометричної фігури і т.і.), виникає група перетворень (рухів, замін перемінних, підстановок, індексів), що зберігають цей об'єкт. Групи виступають тут як міра симетрії і використовуються для класифікації цих об'єктів.

Це пояснює, чому теорія груп має надзвичайний успіх у геометрії, аналізі, теоретичній фізиці.

Визначення групи можна ослабити.

2.1. Теорема. *Напівгрупа G тоді і тільки тоді буде групою, коли вона має праву одиницю, тобто такий елемент* e_r, *що* $xe_r = x$ *для кожного елемента* $x \in G$, *та коли для кожного елементу* $x \in G$ *існує правий обернений, тобто елемент* x_r *із властивістю* $xx_r = e_r$.

Д о в е д е н н я. Якщо G – група, то її одиничний елемент буде і правою одиницею, а елемент, обернений до x, буде і правим оберненим. Навпаки, нехай напівтрупа G задовольняє умовам цієї теореми. Покажемо, що права одиниця буде одночасно і лівою одиницею. Маємо:

$$e_r x x_r = e_r e_r = e_r = x x_r.$$

Помножимо тепер обидві частини рівності $e_r x x_r = x x_r$ на правий обернений до x_r, і будемо мати $e_r x e_r = x e_r$, тобто $e_r x = x$. Отже, e_r – ліва одиниця. Нехай u – довільна права одиниця, v – довільна ліва одиниця, тоді v = vu = u. Це означає, що e_r – єдина права та ліва одиниця в G.

З рівності $x x_r = e_r$ отримуємо $x_r x x_r = x_r e = x_r$. Для елемента x_r існує правий обернений. Множимо обидві частини останньої рівності справа на цей елемент і отримаємо $x_r x e = x x_r = e$, тобто $x_r = x^{-1}$.

2.2. Теорема. *Напівгрупа G тоді і тільки тоді буде групою, коли для довільних елементів* $a, b \in G$ *рівняння* $ax = b$ *та* $xa = b$ *мають розв'язки.*

Д о в е д е н н я. Якщо G – група, то рівняння $ax = b$ має розв'язок $x = a^{-1}b$, а рівняння $xa = b$ має розв'язок $x = ba^{-1}$. Навпаки, нехай G – напівгрупа, в якій вказані рівняння мають розв'язки. Зокрема, рівняння $ax = a$ має розв'язок, який позначимо через e. Нехай b – довільний елемент G, позначимо через c розв'язок рівняння $xa = b$. Маємо тоді:

$$be = (ca)e = c(ae) = ca = b,$$

тобто e – права одиниця для G. Правим оберненим для довільного елемента $a \in G$ буде розв'язок рівняння $xa = e$, і ми можемо застосувати теорему 10.1.

Група G називається скінченою, якщо множина її елементів скінченна. Як і для множин, кількість елементів скінченої групи G будемо позначати через $|G|$ і називати порядком групи G. Група, що не є скінченою, називається нескінченною.

3.3. Теорема. *Скінченна напівгрупа G тоді і тільки тоді буде групою, коли для довільних елементів $a, b, c \in G$ з рівнянь $ab = ac$ та $ba = ca$ завжди випливає $b = c$.*

Д о в е д е н н я. Якщо G – група, то елемент a має обернений, тому з рівняння $ab = ac$ (відповідно $ba = ca$) випливає:

$$b = eb = (a^{-1}a)b = a^{-1}(ab) = a^{-1}(ac) = (a^{-1}a)c = ec = c$$
(відповідно $b = be = b(aa^{-1}) = (ba)a^{-1} = (ca)a^{-1} = c(aa^{-1}) = ce = c$).

Навпаки, нехай $G = \{g_1, \ldots, g_n\}$. З умов теореми випливає, що всі елементи $ag_1, \ldots, ag_n$ будуть попарно різними. Це означає, що $G = \{ag_1, \ldots, ag_n\}$, зокрема $b \in \{ag_1, \ldots, ag_n\}$. У свою чергу, це означає, що рівняння $ax = b$ має розв'язок у групі G. Аналогічно доводиться, що і рівняння $xa = b$ має розв'язок у групі G. Тепер можна застосувати теорему 2.2.

Раніше, ніж розглянути приклади груп, розглянемо важливе поняття підгрупи, а потім розглянемо разом і приклади груп і приклади їх підгруп.

Нехай G – група. Стала підмножина H групи G називається її підгрупою, якщо H буде групою відносно звуження на H операції, заданої в G. Той факт, що H – підгрупа G будемо позначати так: $H \leq G$.

2.4. Теорема (критерій підгрупи). *Непуста підмножина H групи G тоді і тільки тоді буде підгрупою, коли виконуються такі умови:*

(SG 1) *якщо $x, y \in H$, то $xy \in H$;*

(SG 2) *якщо $x \in H$, то і $x^{-1} \in H$.*

Д о в е д е н н я. Якщо H – підгрупа, то умова (SG 1) випливає з означення. Нехай e_H – одиничний елемент H, тоді для довільного елементу $x \in H$ будемо мати $xe_H = x$ (зокрема можна

покласти $x = e_H$). Елемент x має обернений у групі G, тому з останнього рівняння отримаємо:

$$x^{-1} xe_H = x^{-1}x = e.$$

Таким чином $e \in H$. Оскільки одиничні елементи G та H збігаються, то елемент, обернений до x в H, буде оберненим до x у G. Зокрема, має місце і умова (SG 2).

Навпаки, нехай підмножина H задовольняє умови (SG 1), (SG 2). З умови (SG 1) отримуємо, що H – стала підмножина G. Зокрема звуження на H операції, заданої в G, буде бінарною операцією на H. Ця операція асоціативна, оскільки асоціативною є операція в усій групі G. Якщо $x \in H$, то згідно (SG 2) і $x^{-1} \in H$. За умовою (SG 1) $e = xx^{-1} \in H$, зокрема, e буде одиничним елементом і для H. Нарешті остання аксіома групи випливає з умови (SG 2).

2.5. Наслідок. *Непуста підмножина H групи G тоді і тільки тоді буде підгрупою, коли виконується така умова:*

(SG 3) якщо $x, y \in H$, то $xy^{-1} \in H$.

Д о в е д е н н я. Покажемо, що умова (SG 3) рівносильна умовам (SG 1) та (SG 2). Очевидно (SG 3) випливає з обох умов (SG 1), (SG 2). Нехай має місце (SG 3). Якщо $x \in H$, то $e = xx^{-1} \in H$. Далі $x^{-1} = ex^{-1} \in H$, так що (SG 2) має місце.

Нарешті, нехай $y \in H$. Вже було доведено, що і $y^{-1} \in H$, а тому $xy = x(y^{-1})^{-1} \in H$, отже, і (SG 1) має місце.

Нехай G – абелева група з адитивною операцією. На G можна визначити *операцію віднімання* за таким правилом: $x – y = x + (-y)$. Наслідок 2.5 для цього випадку має такий вигляд:

Непуста підмножина H адитивної абелевої групи G тоді і тільки тоді буде підгрупою, коли виконується така умова:

(SG 3) *якщо $x, y \in H$, то $x – y \in H$.*

2.6. Наслідок. *Нехай G – група, H – її підгрупа. Підмножина $K \subseteq H$ тоді і тільки тоді буде підгрупою групи G, коли вона є підгрупою H.*

2.7. Наслідок. *Нехай G – група,* S *– система підгруп групи G. Перетин* $\cap$ S *підгруп цієї системи також буде підгрупою G.*

Д о в е д е н н я. Нехай S = $\cap$ S, x, y $\in$ S. Нехай U – довільна підгрупа з системи $\cap$ S, тоді $xy^{-1} \in U$, і отже xy^{-1} належить перетину всіх таких підгруп, тобто S. Залишилось застосувати наслідок 2.5.

Нехай G – група, H – її підгрупа, M – її деяка підмножина. Наприклад:

$$C_H(M) = \{x \in H \mid xg = gx \text{ для кожного } g \in M\}.$$

Підмножина $C_H(M)$ *називається централізатором підмножини M у підгрупі H.*

Якщо M = {a}, то замість $C_H(\{a\})$ будемо використовувати коротший запис $C_H(a)$. *Слід зазначити, що* $\zeta(G) = C_G(G)$.

2.8. Наслідок. *Нехай G – група, H – її підгрупа, M – підмножина. Централізатор* $C_H(M)$ *буде підгрупою групи G. Зокрема центр* $\zeta(G)$ *також буде підгрупою G.*

Д о в е д е н н я. Нехай x, y $\in C_H(M)$, g $\in$ M. Застосовуючи асоціативний закон, отримуємо (xy)g = x(yg) = x(gy) = (xg)y = (gx)y = g(xy), звідки xy $\in C_H(M)$. Далі, якщо обидві частини рівності xg = gx множити зліва на x^{-1}, то приходимо до рівності $x^{-1}xg = g = x^{-1}gx$. Знову множення обох частин цієї рівності на x^{-1} вже справа приводить до $gx^{-1} = x^{-1}g$, що доводить співвідношення $x^{-1} \in C_H(M)$. Залишилось застосувати теорему 10.4.

Зазначимо, що об′єднання підгруп не буде підгрупою. Воно не завжди буде сталою підмножиною, як це можна було побачити у попередньому параграфі.

Нехай M – множина. *Система L її підмножин називається локальною, якщо для кожної пари підмножин* $H, K \in L$ *існує така підмножина* $L \in L$, *що* $H, K \subseteq L$.

Спеціальним типом локальних систем є лінійно впорядковані системи підгруп. *Система L підмножин множини M називається лінійно впорядкованою, якщо для кожної пари підмножин* $H, K \in L$ *має місце одне з включень* $H \subseteq K$ *або* $K \subseteq H$.

2.9. Наслідок. *Нехай G – група, L – локальна система підгруп групи G. Тоді об'єднання* $\cup$ *L підгруп цієї системи також буде підгрупою G.*

Д о в е д е н н я. Нехай V = $\cup$ *L*, x, y $\in$ V. Існують такі підгрупи H, K з системи *L*, що x $\in$ H, y $\in$ K. Виберемо тепер підгрупу *L* $\in$ *L*, яка містить у собі обидві підгрупи H, K. Таким чином, x, y $\in$ L. Оскільки L – підгрупа, $xy^{-1} \in$ L за наслідком 2.5, і отже, $xy^{-1} \in$ V. Залишилось знову застосувати наслідок 2.5.

2.10. Наслідок. *Нехай G – група,* L *– лінійно впорядкована система підгруп групи G. Тоді об'єднання* $\cup$ *L підгруп цієї системи також буде підгрупою G.*

2.11. Наслідок. *Нехай G – група,* $H_1 \le H_2 \le \ldots \le H_n \le \ldots$ *– зростаюча система підгруп групи G. Тоді об'єднання* $\cup_{n \in N} H_n$ *підгруп цієї системи також буде підгрупою G.*

Нехай M – підмножина групи G, S – система всіх тих підгруп, які містять у собі M. Тоді перетин < M > = $\cap$ **S** буде за наслідком 2.7 підгрупою групи G. Цей перетин називають *підгрупою, породженою підмножиною M, а множина M називається системою породжуючих елементів для підгрупи* < *M* >. *Зокрема, якщо* < *M* > = *G, то будемо говорити, що M породжує всю групу G. Група G називається скінчено породженою, якщо вона породжується деякою своєю скінченою підмножиною.*

Якщо H – підгрупа групи G, що містить у собі підмножину M, то H містить у собі і підгрупу < M >, у цьому сенсі < M > – найменша з підгруп, які містять у собі підмножину M. Якщо M – підгрупа групи G, то < M > = M. Разом з цим ми маємо таке корисне твердження.

2.12. Пропозиція. *Нехай G – група, H – її підгрупа. Якщо* $M \subseteq H$, *то* $< M > \le H$.

Нехай x – елемент групи G, H – довільна підгрупа, до якої належить елемент x. З теореми 2.4 за допомогою звичайної індукції отримаємо співвідношення x^n, $e = x^0$, $x^{-n} \in$ H для будь якого n $\in$ N,

тобто $\{ x^n \mid n \in Z\} \subseteq H$. З іншого боку, з пропозиції 1.7 отримуємо рівність $x^n x^{-m} = x^{n-m}$, яка показує, що $\{ x^n \mid n \in Z \}$ – підгрупа групи G (див. наслідок 2.5). Це означає, що $\{ x^n \mid n \in Z \} = \langle x \rangle$ – підгрупа, породжена підмножиною {x}, або *підгрупа, породжена елементом x*. Цю підгрупу будемо називати також *циклічною*. *Елемент y з властивістю* $\langle y \rangle = \langle x \rangle$ *називається породжуючим елементом для підгрупи* $\langle x \rangle$. Згодом ми побачимо, що циклічна підгрупа може мати більше одного породжуючого елемента. *Група називається циклічною, якщо вона збігається з однією зі своїх циклічних підгруп.*

Розглянемо детальніше підгрупу $\langle x \rangle$. Можливі два випадки.

(I) $x^n \neq x^m$, якщо $n \neq m$.

У цьому випадку підгрупа $\langle x \rangle$ є нескінченною. *Будемо говорити, що елемент x має нескінченний порядок.*

(II) Існують такі цілі числа n, m, що $n \neq m$, але $x^n = x^m$. Одне з чисел n, m більше за інше. Нехай $n > m$. З рівності $x^n = x^m$ отримуємо $x^{n-m} = e$. Отже, серед додатних ступенів елементу x деякі дорівнюють одиничному елементу. Нехай t – найменше додатне число, для якого $x^t = e$. Довільне ціле число n поділимо на t (природно з остачею): $n = tq + r$, де $0 \leq r < t$. Маємо:

$$x^n = x^{tq+r} = x^{tq} x^r = (x^t)^q x^r = x^r.$$

За допомогою попередніх аргументів можна довести, що всі елементи $e = x^0, x = x^1, x^2, \ldots, x^{t-1}$ різні. Звідси випливає, що

$$\langle x \rangle = \{ x^n \mid n \in Z \} = \{ x^0, x^1, x^2, \ldots, x^{t-1} \}.$$

У цьому випадку будемо говорити, що *елемент x має скінченний порядок. Порядком елемента x називається найменше ненульове натуральне число t, яке має властивість* $x^t = e$. *Позначення:* $/x/ = t$. Із самого означення випливає рівність $|x| = |\langle x \rangle|$. Слід зазначити, що $|e| = 1$.

Група називається періодичною, якщо кожен її елемент має скінчений порядок.

Будемо говорити, що група не має скруту або вільна від скруту, якщо кожен її неодиничний елемент має нескінченний порядок.

Група називається мішаною, якщо вона має як неодиничні елементи скінченого порядку, так і неодиничні елементи нескінченного порядку.

Розглянемо тепер більш детально питання про породжуючі елементи циклічних груп.

2.13. Теорема. *Нехай $G = \langle g \rangle$ – циклічна група, $y \in G$.*

(1) Якщо G – нескінченна група, то рівність $\langle y \rangle = G$ має місце тоді і тільки тоді, коли або $y = g$, або $y = g^{-1}$.

*(2) Якщо група G скінченна та $|G| = t$, то рівність $\langle y \rangle = G$ має місце тоді і тільки тоді, коли $y = g^m$, де **НСД**$(m, t) = 1$.*

Д о в е д е н н я. Нехай спочатку $G = \langle g \rangle$ – нескінченна циклічна група, y – її інший породжуючий елемент. Оскільки $y \in \langle g \rangle$, то $y = g^m$ для деякого $m \in Z$. З іншого боку, $\langle y \rangle = G$, зокрема $g \in \langle y \rangle$, а тому і $g = y^d$, де $d \in Z$. Отже, маємо $g = y^d = (g^m)^d = g^{md}$. Множимо тепер обидві частини цієї рівності на елемент g^{-1} і отримаємо $g^{md - 1} = e$. Оскільки $\langle g \rangle$ – нескінченна циклічна група, це означає, що $md - 1 = 0$, тобто $m, d \in \{1, -1\}$. Для елемента y маємо тільки дві можливості: $y = g$ або $y = g^{-1}$. Розглянемо тепер випадок, коли група G – скінченна. Знову маємо $g^{md - 1} = e$, але в цьому випадку це означає, що $md - 1 = kt$ для деякого $k \in Z$. Звідси $md + kt = 1$, а отже **НСД**$(m, t) = 1$. Навпаки, нехай **НСД**$(m, t) = 1$. Існують такі цілі числа d, k, що $md + tk = 1$. Тоді

$$g = g^1 = g^{md + tk} = g^{md}\, g^{tk} = (g^m)^d\, (g^t)^k = y^d\, e^k = y^d\ .$$

Зокрема, $g \in \langle y \rangle$, і за пропозицією 2.12 $\langle g \rangle \le \langle y \rangle$. Та ж пропозиція 2.12 дає і протилежне включення $\langle y \rangle \le \langle g \rangle$, отже $\langle y \rangle = \langle g \rangle$.

2.14. Теорема. *Нехай $G = \langle g \rangle$ – циклічна група, H – її підгрупа. Тоді H також буде циклічною.*

Д о в е д е н н я. Оскільки H – підгрупа, то $e \in H$. Якщо $H = \{e\}$, то $H = \langle e \rangle$, зокрема H – циклічна. Припустимо тепер, що H містить неодиничні елементи. Всі елементи групи G вичерпуються цілими ступенями елемента g, тому існує таке ціле число m, що

$1 \neq g^m \in H$. Якщо $m < 0$, то, за умовою (SG 2), також $(g^m)^{-1} = g^{-m} \in H$. Це означає, що підгрупа H обов'язково містить додатні неодиничні ступені елементу g. Нехай $\Xi = \{ k > 0 \mid 1 \neq g^k \in H \}$. У множині Ξ вибираємо найменший елемент d. Зокрема $g^d \in H$, і за пропозицією 10.12 $\langle g^d \rangle \leq H$. Якщо x – довільний елемент підгрупи H, то $x = g^n$ для деякого $n \in Z$. Поділимо n на d (природно з остачею): $n = dq + r$, де $0 \leq r < d$. Звідси

$$x = g^n = g^{dq + r} = (g^d)^q (g^r) \text{ і } g^r = x (g^{dq})^{-1}.$$

За умовою (SG 3) $g^r \in H$, а тому $r \in \Xi$. З вибору числа d випливає тоді, що $r = 0$, отже, $x = g^n = (g^d)^q$. Це означає, що $H \leq \langle g^d \rangle$. Ми вже довели, що протилежне включення також має місце, отже $H = \langle g^d \rangle$.

Узагальнимо тепер розглянуту ситуацію.

2.15. Теорема. *Нехай G – група, M – її підмножина. Підгрупа $< M >$ складається з елементів, що мають вигляд $x_1 x_2 \ldots x_n$, де $x_j \in M$ або $x_j^{-1} \in M$, $1 \leq j \leq n$, і тільки з них.*

Д о в е д е н н я. Нехай Y – множина всіх елементів, що мають вигляд $x_1 x_2 \ldots x_n$, де $x_j \in M$ або $x_j^{-1} \in M$, $1 \leq j \leq n$. З теореми 2.4 випливає, що $Y \subseteq \langle M \rangle$. З іншого боку, з самого означення випливає включення $M \subseteq Y$. Покажемо, що Y задовольняє умову (SG 3), із наслідку 2.5 буде випливати тоді, що Y – підгрупа.

Отже, нехай $g_1, g_2 \in Y$, тоді $g_1 = x_1 x_2 \ldots x_n$, $g_2 = x_{n+1} x_{n+2} \ldots x_m$, де $x_j \in M$ або $x_j^{-1} \in M$, $1 \leq j \leq m$. Маємо:

$$g_1 (g_2)^{-1} = (x_1 x_2 \ldots x_n)(x_{n+1} x_{n+2} \ldots x_m)^{-1} = x_1 x_2 \ldots x_n x_m^{-1} \ldots x_{n+1}^{-1} \in Y.$$

Отже, Y – підгрупа, що містить у собі підмножину M. Це означає, що $\langle M \rangle \leq Y$. Оскільки і протилежне включення має місце, то $\langle M \rangle = Y$.

На закінчення розглянемо корисну конструкцію декартового (прямого) добутка скінченої множини груп. Нехай $G_1, \ldots, G_n$ – групи, $D = G_1 \times \ldots \times G_n$ – їх декартовий добуток як множин. На множині D визначимо операцію множення за правилом:

$$(g_1, \ldots, g_n)(x_1, \ldots, x_n) = (g_1 x_1, \ldots, g_n x_n), \text{ де } g_j, x_j \in G_j, 1 \leq j \leq n.$$

Оскільки g_j, x_j – елементи однієї і тієї ж групи G_j, то їх добуток визначений у цій групі G_j, $1 \le j \le n$, так що ми отримали дійсно операцію на D. Ця операція асоціативна:

$$((g_1, \ldots, g_n)(x_1, \ldots, x_n))\,(y_1, \ldots, y_n) = ((g_1x_1)\,y_1, \ldots, (g_nx_n)y_n) =$$
$$= (g_1(x_1y_1), \ldots, g_n(x_ny_n)) = (g_1, \ldots, g_n)((x_1, \ldots, x_n)\,(y_1, \ldots, y_n)).$$

Якщо e_j – одиничний елемент групи G_j, $1 \le j \le n$, то набір $(e_1, \ldots, e_n)$ буде, очевидно, одиничним елементом для D. Нарешті, також очевидно, що

$$(g_1, \ldots, g_n)^{-1} = (g_1^{-1}, \ldots, g_n^{-1}).$$

Отже, всі аксіоми групи для D виконані. Тому далі, коли йдеться про декартовий добуток груп, ми будемо розуміти під цим тільки що побудовану групу. Ця група називається також прямим добутком груп $G_1, \ldots, G_n$ (зазначимо, що поняття прямого та декартова добутка груп у загальному випадку відрізняються один від одного, але для випадку скінченої множини груп вони збігаються). Якщо всі групи $G_1, \ldots, G_n$ – абелеві, то і їх декартовий добуток буде абелевою групою:

$$(g_1, \ldots, g_n)(x_1, \ldots, x_n) = (g_1x_1, \ldots, g_nx_n) = (x_1g_1, \ldots, x_ng_n) =$$
$$= (x_1, \ldots, x_n)(g_1, \ldots, g_n).$$

Якщо у кожній з груп $G_1, \ldots, G_n$ операція записується адитивно, то адитивний запис використовується і для позначення групової операції в їх декартовому добутку:

$$(g_1, \ldots, g_n) + (x_1, \ldots, x_n) = (g_1 + x_1, \ldots, g_n + x_n).$$

РОЗДІЛ 3
ПРИКЛАДИ ГРУП ТА ПІДГРУП

Числові групи. Множина R усіх дійсних чисел буде групою відносно операції додавання. Ця група буде вільною від скруту. Множина Q всіх раціональних чисел та множина Z всіх цілих чисел будуть її підгрупами. Група Z дає нам приклад нескінченної циклічної групи, вона породжується числом 1. Для кожного $n \geq 0$ підмножина $nZ = \{nk \mid k \in Z\}$ буде підгрупою Z. Дійсно, $nk - nt = n(k - t)$, тобто nZ задовольняє умову (SG 3) і отже за наслідком 2.5 nZ буде підгрупою Z. З теореми 2.13 випливає, що довільна підгрупа Z збігається з однією з підмножин nZ, $n \geq 0$, і інших підгруп Z не має.

Нехай p – просте число, $Q_P = \{ m/(p^k) \mid m, k \in Z\}$. Дріб, що має вигляд $m/(p^k)$, називається p–ічним. Очевидна рівність

$$m/(p^k) - r/(p^s) = (mp^s - rp^k)/(p^{k+s})$$

показує, що Q_P задовольняє умову (SG 3), отже, за наслідком 10.5 Q_P буде підгрупою Q. Її називають (адитивною) групою p – ічних дробів.

Множина $R\backslash\{0\} = R^\times$ усіх ненульових дійсних чисел буде групою відносно операції множення. Множини $Q\backslash\{0\} = Q^\times$ усіх ненульових раціональних чисел та множина {1, -1} будуть її підгрупами, більш того, остання дає нам приклад скінченої циклічної групи. Зазначимо, що єдиними елементами групи $R^\times$, які мають скінченний порядок, є 1 та –1.

Булеан множини. Нехай A – довільна множина, B (A) – булеан множини A. З теореми 1.1 випливає, що B(A) буде групою відносно операції Δ, нагадаємо, що $C \Delta B = (C \setminus B) \cup (B \setminus C)$. Ця група буде абелевою, а з рівності $B \Delta Y = \varnothing$ отримуємо, що кожен елемент цієї групи має порядок 2.

У розділі 1 ми вже зазначали, що множина P(A) всіх перетворень множини A буде напівгрупою з одиницею. Зокрема, оскільки U(P(A)) = S(A), то множина S(A) буде групою відносно операції множення функцій. Її називають *групою підстановок*

множини A. Для групи S_n існує спеціальна назва: симетрична група ступеню n. Відомо, що $U(M_n(R)) = GL_n(R)$, зокрема, множина $GL_n(R)$ всіх невироджених квадратних матриць порядку n із дійсними коефіцієнтами буде групою відносно операції множення матриць. Її називають *загальною лінійною групою ступеня n над* R. Підмножина $GL_n(Q)$ буде підгрупою в $GL_n(R)$. Дійсно, з визначення операції множення матриць випливає, що добуток двох матриць з раціональними коефіцієнтами має тільки раціональні коефіцієнти. Якщо $A \in GL_n(Q)$, то з відомих результатів випливає, що $A^{-1} \in GL_n(Q)$. Таким чином, $GL_n(Q)$ задовольняє умови (SG 1) – (SG 2), отже, за теоремою 2.4 $GL_n(Q)$ буде підгрупою $GL_n(R)$.

Групи підстановок та перетворень. Нехай A – довільна множина. Ми вже зазначали, що множина **S**(A) всіх підстановок A буде групою. Розглянемо зараз важливі приклади підгруп цієї групи.

Стабілізатор підмножини. Нехай B – підмножина A,

$$St(B) = \{\pi \in S(A) \mid \pi(b) = b \text{ для кожного елемента } b \in B\}.$$

Якщо $\pi, \sigma \in St(B)$, то для кожного елемента $b \in B$ маємо:

$$\pi \☺ \sigma(b) = \pi(\sigma(b)) = \pi(b) = b,$$

тобто $\pi \☺ \sigma \in St(B)$, а також

$$b = \varepsilon_A(b) = (\pi^{-1} \☺ \pi)(b) = (\pi^{-1}(\pi(b)) = \pi^{-1}(b),$$

тобто і $\pi^{-1} \in St(B)$.

Отже, підмножина St(B) задовольняє обидві умови (SG 1), (SG 2) і за теоремою 2.4 буде підгрупою S(A). Зокрема, підгрупою буде St(a) для кожного елемента $a \in A$. Зазначимо також очевидну рівність $St(B) = \cap_{b \in B} St(b)$.

Нехай тепер $Inv(B) = \{ \pi \in S(A) \mid \pi(B) = B \}$. Якщо $\pi, \sigma \in Inv(B)$, $b \in B$, то $\sigma(b) = b_1 \in B$, так що маємо:

$$\pi \☺ \sigma(b) = \pi(\sigma(b)) = \pi(b_1) = b_2 \in B,$$

тобто $\pi \☺ \sigma \in \mathrm{Inv}(B)$. Оскільки $\pi(B) = B$, то існує такий елемент $b_3 \in B$, що $b = \pi(b_3)$, а тоді

$$b_3 = \varepsilon_A(b_3) = (\pi^{-1} \☺ \pi)(b_3) = (\pi^{-1}(\pi(b_3)) = \pi^{-1}(b),$$

тобто і $\pi^{-1} \in \mathrm{Inv}(B)$.

Відтак, підмножина Inv(B) задовольняє обидві умови (SG 1), (SG 2) і за теоремою 2.4 буде підгрупою S(A). Слід зазначити, що $\mathrm{St}(B) \le \mathrm{Inv}(B)$, водночас як $\mathrm{Inv}(a) = \mathrm{St}(a)$ для кожного елемента $a \in A$.

Фінітарні підстановки. Нехай $\pi \in \mathbf{S}(A)$. Підмножина

$$\mathrm{Supp}\ \pi = \{a \in A \mid \pi(a) \neq a\}$$

називається *носієм* підстановки π, а її доповнення $A \setminus \mathrm{Supp}\ \pi$ – *множиною нерухомих точок підстановки* π. Підстановка π називається *фінітарною*, якщо її носій $\mathrm{Supp}\ \pi$ є скінченним. Позначимо через FS(A) підмножину всіх фінітарних підстановок множини A. Якщо $\pi, \sigma \in \mathrm{FS}(A)$ і $b \in A \setminus (\mathrm{Supp}\ \pi \cup \mathrm{Supp}\ \sigma)$, то

$$\pi \☺ \sigma(b) = \pi(\sigma(b)) = \pi(b) = b,$$

тобто $\mathrm{Supp}\ (\pi \☺ \sigma) \subseteq \mathrm{Supp}\ \pi \cup \mathrm{Supp}\ \sigma$, зокрема $\mathrm{Supp}\ (\pi \☺ \sigma)$ скінченний, а отже, $\pi \☺ \sigma \in \mathrm{FS}(A)$. Далі, нехай $b \in A \setminus \mathrm{Supp}\ \pi$, тоді

$$b = \varepsilon_A(b) = (\pi^{-1} \☺ \pi)(b) = (\pi^{-1}(\pi(b)) = \pi^{-1}(b).$$

Це означає, що $\mathrm{Supp}\ \pi^{-1} \subseteq \mathrm{Supp}\ \pi$. Оскільки $\pi = (\pi^{-1})^{-1}$, то і $\mathrm{Supp}\ \pi \subseteq \mathrm{Supp}\ \pi^{-1}$, а тому $\mathrm{Supp}\ \pi^{-1} = \mathrm{Supp}\ \pi$. Зокрема підмножина $\mathrm{Supp}\ \pi^{-1}$ скінченна, отже, і $\pi^{-1} \in \mathrm{FS}(A)$. Відтак підмножина FS(A) задовольняє обидві умови (SG 1), (SG 2) і за теоремою 2.4 буде підгрупою S(A). FS(A) називають *групою фінітарних підстановок* множини A.

Групи симетрій. Нехай $E = R^n$, через d(x, y) будемо позначати відстань між точками x, y простору E. Відображення $f \in S(E)$ будемо називати *ізометрією* простору E, якщо f зберігає відстані між точками простору E, тобто $d(f(x), f(y)) = d(x, y)$ для кожної

пари точок x, y простору E. Сукупність всіх ізометрій простору E позначимо через Isom(E). Якщо $f, g \in Isom(E)$, $x, y \in E$, то

$$d(f \circ g(x), f \circ g(y)) = d(f(g(x)), f(g(y))) = d(g(x), g(y)) = d(x, y),$$

тобто $f \circ g \in Isom(E)$.

Далі

$$d(x, y) = d(f(f^{-1}(x)), f(f^{-1}(y))) = d(f^{-1}(x)), f^{-1}(y)),$$

тобто і $f^{-1} \in Isom(E)$.

Отже, підмножина Isom(E) задовольняє обидві умови (SG 1), (SG 2) і за теоремою 2.4 буде підгрупою S(E).

Нехай тепер M – підмножина простору E, визначемо:

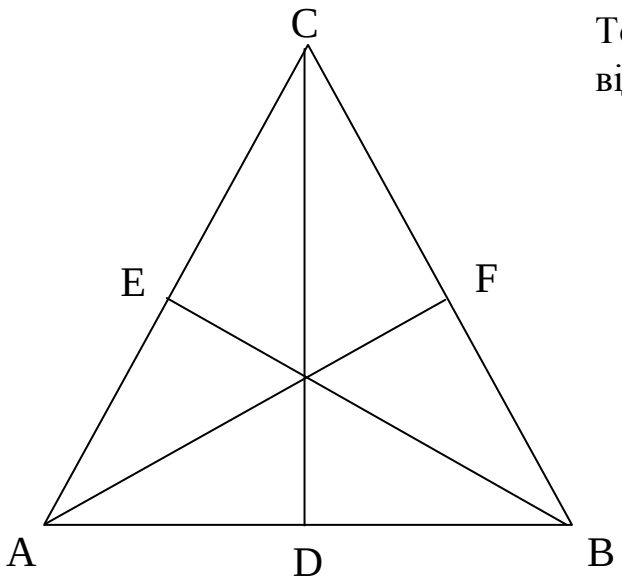
$$Sym(M) = Isom(E) \cap Inv(B),$$

інакше кажучи, Sym(M) складається з тих ізометрій простору E, які переводять підмножину M у себе. З наслідку 2.7 випливає, що Sym(M) – підгрупа Isom(E). Її називають *групою симетрій* підмножини M. Образно кажучи, група Sym(M) вимірює ступінь симетричності фігури M. Наприклад, той факт, що рівнобедрений трикутник є більш симетричним, ніж нерівнобедрений, можна виміряти тим, що група симетрій нерівнобедреного трикутника (як підмножини простору R^2 – дійсної площини) складається лише з однієї тотожної підстановки, водночас як група симетрій рівнобедреного трикутника складається з тотожної підстановки та відбиття відносно осі симетрії рівнобедреного трикутника. Якщо ж розгляднути групу симетрій рівностороннього трикутника, то вона складається уже з шести ізометрій: тотожної підстановки, поворотів на кути 120° та 240° навкруги центра трикутника та трьох відбиттів відносно трьох осей симетрій рівностороннього трикутника. У цьому сенсі рівносторонній трикутник буде більш симетричним, ніж рівнобедрений.

Групи симетрій скінченних об’єктів можна зв’язати з підгрупами групи підстановок скінченого ступеня. Розглянемо це детальніше на таких прикладах.

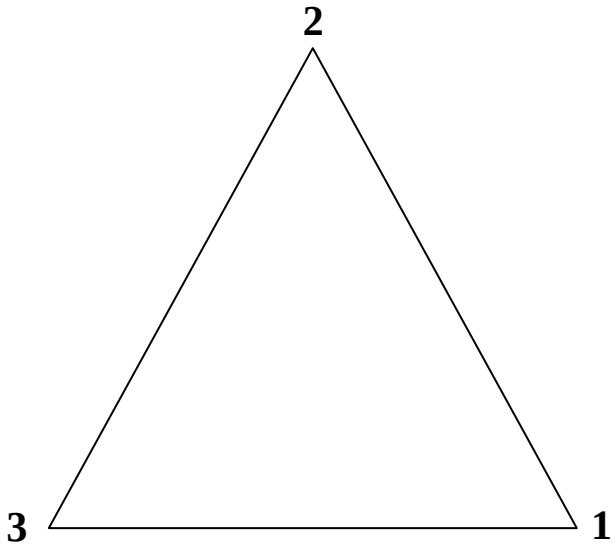
Нехай P_3 – правильний трикутник. Кожній вершині трикутника приписуємо деякий номер. Оскільки кожна симетрія переміщує вершини трикутника, то отримуємо деяку підстановку ступеня 3.

Вершині A приписуємо номер 1, вершині B – номер 2, а вершині C – номер 3.



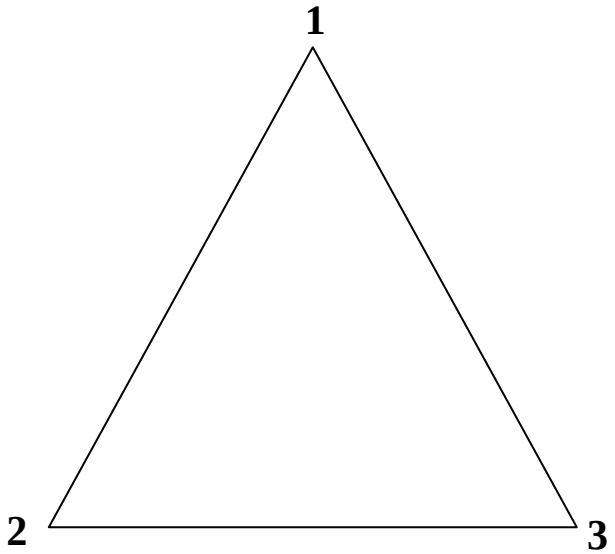
Тотожному перетворенню площини відповідає тотожня підстановка

$$\varepsilon = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$



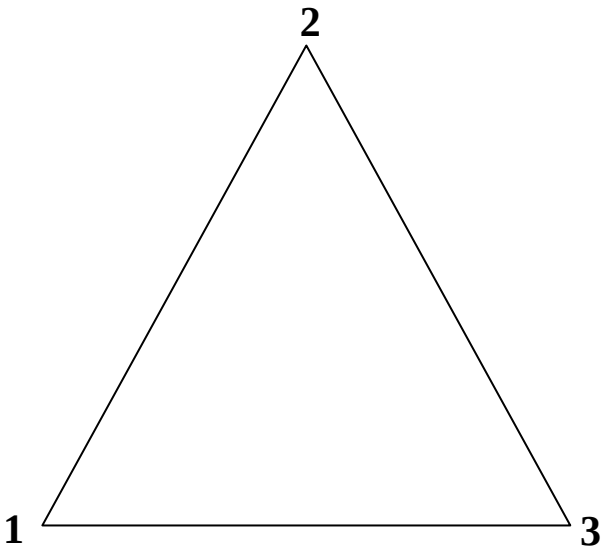
Повороту площини відносно центра трикутника на 120° відповідає підстановка

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$



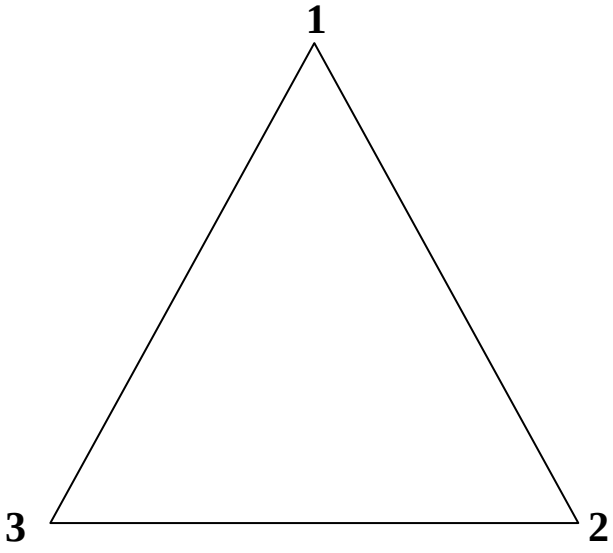
Повороту площини відносно центра трикутника на 240° відповідає підстановка

$$\pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$



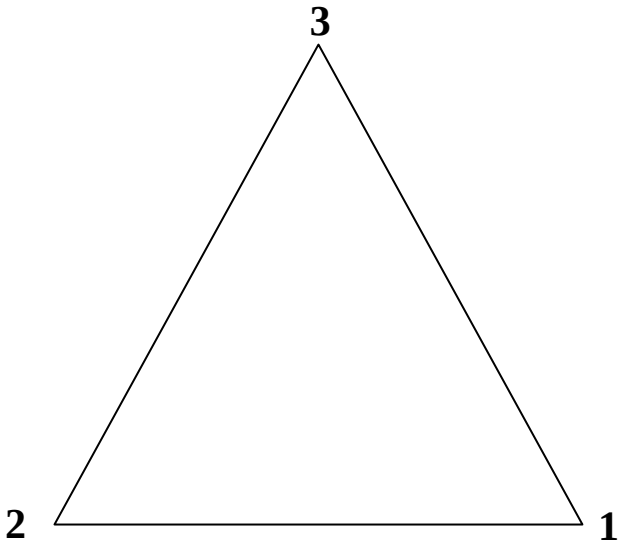
Відбиттю відносно осі AF трикутника відповідає підстановка

$$\pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$



Відбиттю відносно осі BE трикутника відповідає підстановка

$$\pi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$



Відбиттю відносно осі CD трикутника відповідає підстановка

$$\pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Отже, група симетрій правильного трикутника – це група $\mathbf{S_3}$.

Групу симетрій правильного чотирикутника не будемо розглядати так детально. Зазначимо, що вона складається з тотожної підстановки ε, поворотів площини відносно центра чотирикутника на кути 90°, 180°, 270°, яким відповідають підстановки:

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad \pi_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix},$$

з двох відбиттів відносно діагоналей чотирикутника, яким відповідають підстановки:

$$\pi_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} \quad \pi_5 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

з двох відбиттів відносно прямих, що проходять через середини протилежних сторін чотирикутника, яким відповідають підстановки:

$$\pi_6 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \quad \pi_7 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

Запишемо тепер таблицю множення для групи $\mathrm{Sym}(P_4)$ у вигляді матриці. У першому рядку і у першому стовпці запишемо всі елементи групи, а на перетині рядка, що починається з елемента x, та стовпця, що починається з елемента y, будемо записувати елемент xy.

Таку таблицю називають ще *таблицею Келі групи*.

	ε	π_1	π_2	π_3	π_4	π_5	π_6	π_7
ε	ε	π_1	π_2	π_3	π_4	π_5	π_6	π_7
π_1	π_1	π_2	π_3	ε	π_7	π_6	π_4	π_5
π_2	π_2	π_3	ε	π_1	π_5	π_4	π_7	π_6
π_3	π_3	ε	π_1	π_2	π_6	π_7	π_5	π_4
π_4	π_4	π_6	π_5	π_7	ε	π_2	π_1	π_3
π_5	π_5	π_7	π_4	π_6	π_2	ε	π_3	π_1
π_6	π_6	π_5	π_7	π_4	π_3	π_1	ε	π_2

π_7	π_7	π_4	π_6	π_5	π_1	π_3	π_2	ε

Зробимо невеликий аналіз. Елемент $\sigma = \pi_1$ має порядок 4, а $\pi_2 = \sigma^2$, $\pi_3 = \sigma^3$. Покладемо також $\tau = \pi_4$, тоді $\tau^2 = \varepsilon$ і з таблиці Келі можна побачити, що $Sym(P_4) = <\sigma, \tau>$. Елементи σ, τ не переставні: $\tau\sigma = \sigma^3\tau$. Групу, яку ми зараз отримали, називають *групою діедра*.

Лінійні групи. Лінійними групами (точніше скінчено – вимірними лінійними групами) називають групу $GL_n(R)$ та її підгрупи. Зараз і розглянемо деякі важливі її підгрупи. Ми вже зазначали, що множина $GL_n(Q)$ всіх невироджених квадратних матриць порядку n з раціональними коефіцієнтами буде підгрупою $GL_n(R)$. Встановлено

$$GL_n(Z) = \{A \in M_n(Z) \mid \det(A) \in \{1, -1\}\}.$$

З визначення операції множення матриць випливає, що добуток двох матриць з цілими коефіцієнтами має цілі коефіцієнти. Якщо $A \in GL_n(Z)$, то з відомих результатів випливає, що $A^{-1} \in GL_n(Z)$. Таким чином, $GL_n(Z)$ задовольняє умови (SG 1) – (SG 2), отже, за теоремою 2.4 $GL_n(Z)$ буде підгрупою $GL_n(R)$.

Визначимо тепер

$$SL_n(R) = \{A \in M_n(R) \mid \det(A) = 1\}.$$

Нехай $A, B \in SL_n(R)$. Тоді $\det(AB) = \det(A)\det(B) = 1$, що означає включення $AB \in SL_n(R)$. Далі отримаємо також, що

$$1 = \det(E) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

а тому і $\det(A^{-1}) = 1$, тобто $A^{-1} \in SL_n(R)$. Відтак, $SL_n(R)$ задовольняє умови (SG 1) – (SG 2), отже, за теоремою 2.4 $SL_n(R)$ буде підгрупою $GL_n(R)$. Її називають *спеціальною лінійною групою ступеня n над* R. Аналогічно визначаються підгрупи $SL_n(Q)$ та $SL_n(Z)$.

Позначимо тепер через $T_n(R)$ множину всіх невироджених трикутних (точніше верхньо – трикутних) матриць порядку n з дійсними коефіцієнтами. Нехай знову $A, B \in T_n(R)$, $A = \| a_{ij} \|_{1 \le i, j \le n}$, $B = \| b_{ij} \|_{1 \le i, j \le n}$, $C = AB = \| c_{ij} \|_{1 \le i, j \le n}$, та нехай $i > j$. Оскільки A та B – трикутні матриці, $a_{ij} = b_{ij} = 0$. Маємо тепер:

$$c_{i\,j} = a_{i\,1}b_{1\,j} + a_{i\,2}b_{2\,j} + \ldots + a_{i\,i-1}b_{i-1\,j} + a_{i\,i}b_{i\,j} + a_{i\,i+1}b_{i+1\,j} + \ldots + a_{i\,n}b_{n\,j}.$$

Оскільки

$$a_{i\,1} = a_{i\,2} = \ldots = a_{i\,i-1} = 0 \;\; i \;\; b_{i\,j} = b_{i+1\,j} = \ldots = b_{n\,j} = 0,$$

то також $c_{i\,j} = 0$. Це означає, що $AB \in T_n(R)$.

Слід відзначити також рівність

$$c_{i\,i} = a_{i\,1}b_{1\,i} + a_{i\,2}b_{2\,i} + \ldots + a_{i\,i-1}b_{i-1\,i} + a_{i\,i}b_{i\,i} + a_{i\,i+1}b_{i+1\,i} + \ldots + a_{i\,n}b_{n\,i} = a_{i\,i}\,b_{i\,i}.$$

Розглянемо тепер матрицю A^{-1}. Позначимо її коефіцієнти через $x_{i\,j}$, $1 \le i, j \le n$. Теорема 8.2 дає нам коефіцієнти $x_{i\,j} = A_{j\,i}/\det(A)$, $1 \le i, j \le n$. У свою чергу, $A_{j\,i} = (-1)^{j+i}\det(S_{j\,i})$, де $S_{j\,i} = \| y_{k\,m} \|_{1 \le k,\, m \le n-1}$ – підматриця матриці A, отримана викресленням з неї j-го рядка та i-го стовпця. Нехай $i > j$, розглянемо довільний k-й стовпець матриці $S_{j\,i}$. Якщо $k < j$, то k – й стовпець матриці $S_{j\,i}$ складається з елементів:

$$y_{1\,k} = a_{1\,k}, \ldots, y_{k-1\,k} = a_{k-1\,k}, y_{k\,k} = a_{k\,k}, y_{k\,k} = 0, y_{k+1\,k} = 0, \ldots, y_{n-1\,k} = 0.$$

Якщо $j = k$, то k-й стовпець матриці складається з елементів:

$$y_{1\,k} = a_{1\,k}, \ldots, y_{k-1\,k} = a_{k-1\,k}, y_{k\,k} = 0, y_{k+1\,k} = 0, \ldots, y_{n\,k} = 0.$$

Якщо $i > k > j$, то k-й стовпець матриці складається з елементів:

$$y_{1\,k} = a_{1\,k}, \ldots, y_{j-1\,k} = a_{j-1\,k}, y_{j\,k} = a_{j+1\,k}, \ldots, y_{k-1\,k} = a_{k\,k}, y_{k\,k} = 0, \ldots, y_{n-1\,k} = 0.$$

Якщо $k \ge i$, то k – й стовпець матриці складається з елементів:

$$y_{1\,k} = a_{1\,k+1}, \ldots, y_{j-1\,k} = a_{j-1\,k+1}, y_{j\,k} = a_{j+1\,k+1}, \ldots, y_{k-1\,k} = a_{k\,k+1}, y_{k\,k} = a_{k+1\,k+1}, y_{k+1\,k} = 0, \ldots, y_{n-1\,k} = 0.$$

Звідси видно, що матриця $S_{j\,i}$ трикутна. Якщо $j < n-1$, $y_{j\,j} = 0$, то $A_{j\,i} = (-1)^{j+i}\det(S_{j\,i}) = 0$. Якщо $j = n-1$, то $i = n$; у цьому випадку всі елементи останнього рядку матриці дорівнюють нулю. Це показує, що $A_{j\,i} = (-1)^{j+i}\det(S_{j\,i}) = 0$. Таким чином, якщо $i > j$, то $x_{i\,j} = A_{j\,i}/\det(A) = 0$. Це означає, що $A^{-1} \in T_n(R)$. Таким чином, $T_n(R)$

адовольняє умови (SG 1) – (SG 2), отже за теоремою 2.4 вона буде підгрупою (R). Встановимо T_n (Q) = GL_n (Q) $\cap$ T_n (R) та (Z) = GL_n (Z) $\cap$ T_n (R).

Позначимо через $UT_n(R)$ множину всіх унітрикутних матриць порядку n із дійсними коефіцієнтами. Якщо A, B $\in$ UT_n (R), $A = \| a_{ij} \|_{1 \le i, j \le n}$, $B = \| b_{ij} \|_{1 \le i, j \le n}$, $C = AB = \| c_{ij} \|_{1 \le i, j \le n}$, то, як ми тільки що довели, C – трикутна матриця. Ми також довели рівність $c_{ii} = a_{ii}\ b_{ii}$, з якої випливає, що $c_{ii} = 1$ для кожного i, $1 \le i \le n$. З цієї рівності також отримаємо, що матриця, обернена до унітрикутної, сама буде унітрикутною.

Відтак UT_n (R) задовольняє умови (SG 1) – (SG 2), отже, за теоремою 2.4 вона буде підгрупою GL_n (R). Визначимо GL_n (Q) $\cap$ UT_n (R) =UT_n (Q) та UT_n (Z) = GL_n (Z) $\cap$ UT_n (R).

Позначимо через $D_n(R)$ множину всіх невироджених діагональних матриць порядку n з дійсними коефіцієнтами. Якщо A, B $\in$ D_n (R), $A = \| a_{ij} \|_{1 \le i, j \le n}$, $B = \| b_{ij} \|_{1 \le i, j \le n}$, $C = AB = \| c_{ij} \|_{1 \le i, j \le n}$, то, застосовуючи попередні аргументи, неважко довести, що C – діагональна матриця. Більше того, $c_{ii} = a_{ii}\ b_{ii}$ для кожного i, $1 \le i \le n$. Зокрема, AB = BA. З цієї рівності також отримаємо, що матриця, обернена до діагональної, сама буде діагональною.

Відтак D_n (R) задовольняє умови (SG 1) – (SG 2), отже, за теоремою 2.4 вона буде абелевою підгрупою GL_n (R). Встановимо D_n (Q) = GL_n (Q) $\cap$ D_n (R) та D_n (Z) = GL_n (Z) $\cap$ D_n (R). Зокрема, якщо A $\in$ D_n (Z), то $a_{ii} \in \{1, -1\}$ для кожного i, $1 \le i \le n$. Звідси випливає, що D_n (Z) – скінченна абелева група і $| D_n (Z) | = 2^n$.

Матриця A $\in$ $M_n(R)$ називається *ортогональною*, якщо $AA^t = E$, тобто $A^{-1} = A^t$. З рівності $AA^t = E$ отримаємо:

$$1 = \det(E) = \det(AA^t) = \det(A)\det(A^t) = (\det(A))^2,$$

Зокрема, ортогональна матриця завжди невироджена.

Позначимо через $O_n(R)$ підмножину $M_n(R)$, що складається з усіх ортогональних матриць. Нехай A, B $\in$ O_n (R), тоді

$$(AB)^t = B^tA^t = B^{-1}A^{-1} = (AB)^{-1}, \text{ та } (A^{-1})^t = (A^t)^t = A = (A^{-1})^{-1}.$$

Інакше кажучи, $AB \in O_n(R)$, $A^{-1} \in O_n(R)$, а це, враховуючи теорему 2.4, означає, що $O_n(R)$ – підгрупа $GL_n(R)$.

Розділ 4
ВІДНОШЕННЯ ЕКВІВАЛЕНТНОСТІ НА МНОЖИНАХ. РОЗКЛАД ГРУПИ В ОБ'ЄДНАННЯ СУМІЖНИХ КЛАСІВ

Ми вже зазначали раніше, що поняття відповідності і її часткового випадку – поняття відношення є одним із самих загальних понять математики. Конкретні важливі типи відношень визначаються тими чи іншими властивостями. Розглянемо зараз одне з важливих відношень – відношення еквівалентності.

Нагадаємо, що *бінарне відношення на множині A – це підмножина Φ декартового добутку $A \times A$. Якщо $\alpha = (x, y) \in \Phi$, то будемо говорити, що елементи x, y (у вказаному порядку) відповідають один одному при Φ.* Замість запису $(x, y) \in \Phi$ будемо використовувати також і інший запис $x\Phi y$.

Бінарні відношення таким чином є множинами, а тому їх вивчення це є, по суті, вивчення підмножин декартового добутку $A \times A$. Зокрема, можна говорити про включення одного бінарного відношення у інше, про об'єднання чи перетин бінарних відношень та ін.

Випишемо зараз найбільш важливі властивості бінарних відношень.

Рефлективність. *Бінарне відношення Φ називається рефлексивним, якщо для кожного елементу $a \in A$ має місце $(a, a) \in \Phi$ (або має місце $a \Phi a$ для кожного $a \in A$).*

Транзитивність. *Бінарне відношення Φ називається транзитивним, якщо з включень $(a, b), (b, c) \in \Phi$ завжди випливає $(a, c) \in \Phi$, $a, b, c \in A$ (або з $a \Phi b$ та $b \Phi c$ випливає $a \Phi c$).*

Симетричність. *Бінарне відношення Φ називається симетричним, якщо з включення $(a, b) \in \Phi$ завжди випливає $(b, a) \in \Phi$, $a, b \in A$ (або з $a \Phi b$ випливає $b \Phi a$).*

Антисиметричність. *Бінарне відношення Ф називається антисиметричним, якщо з включень (a, b), (b, a) ∈ Ф завжди випливає a = b, a, b ∈ A (або з aФb та bФa випливає a = b).*

Відношення еквівалентності пов'язані та виникають з різних розбиттів множини A.

Система S *підмножин множини A називається покриттям A, якщо A =* ∪S. *Покриття* S- *називається розбиттями множини A, якщо кожні дві різні підмножини з системи* S *мають пустий перетин.*

Нехай S – розбиття множини A. Розглянемо бінарне відношення Ξ(S) на множині A, що задається за наступним правилом:

*(x, y) ∈ Ξ(*S*) тоді і тільки тоді, коли елементи x та y належать до однієї множини S з системи* S .

Укажемо зараз властивості відношення Ξ(S).

Оскільки A = ∪ S, то для кожного елементу x ∈ A існує підмножина S ∈ S, до якої належить елемент x. Зокрема, (x, x) ∈ Ξ(S). Це означає, що відношення Ξ(S) рефлексивне.

Відношення Ξ(S) симетричне, це очевидно.

Нарешті, нехай (x, y), (y, z) ∈ Ξ(S). Це означає, що існують такі підмножини S, R ∈ S, що x, y ∈ R та y, z ∈ R. Зокрема, y ∈ S ∩ R, і з означення розбиття отримуємо рівність S = R. Отже, елементи x, z належать до підмножини S, яка є елементом розбиття S. Інакше кажучи, відношення Ξ(S) є транзитивним.

Бінарне відношення Ξ на множині A називається відношенням еквівалентності або просто еквівалентністю, якщо воно рефлексивне, симетричне, транзитивне.

Нехай Ξ – відношення еквівалентності на множині A, x ∈ A. *Підмножину[x]*$_{Ξ}$ *= {y ∈ A / (x, y) ∈ Ξ} будемо називати класом еквівалентності елемента x.*

Слід зразу зазначити, що клас еквівалентності визначається кожним своїм елементом. Дійсно, нехай y ∈ [x]$_{Ξ}$. Якщо z ∈ [y]$_{Ξ}$, то (y, z) ∈ Ξ. Разом з (x, y) ∈ Ξ це дає, враховуючи транзитивність Ξ співвідношення (x, z) ∈ Ξ, тобто z ∈ [x]$_{Ξ}$. Інакше кажучи,

$[y]_{\Xi} \subseteq [x]_{\Xi}$. Повторюючи ті ж самі аргументи, можна довести, що і $[x]_{\Xi} \subseteq [y]_{\Xi}$, тобто $[y]_{\Xi} = [x]_{\Xi}$.

Оскільки $(x, x) \in \Xi$, то $x \in [x]_{\Xi}$. Звідси випливає, що сукупність усіх класів еквівалентності утворює покриття множини A. Розглянемо перетин двох класів еквівалентності $[x]_{\Xi}$ та $[y]_{\Xi}$ і припустимо, що цей перетин непустий. Нехай $z \in [x]_{\Xi} \cap [y]_{\Xi}$. Тоді, як ми вже зазначали вище, $[z]_{\Xi} = [x]_{\Xi}$ та $[z]_{\Xi} = [y]_{\Xi}$, а отже, $[y]_{\Xi} = [x]_{\Xi}$. Звідси випливає, що різні класи еквівалентності мають пустий перетин, а тому система всіх класів еквівалентності утворює розбиття $\mathbf{P}(\Xi)$ множини A.

4.1. Теорема. *Нехай A – множина. Відображення* $\rho: \Xi \mapsto \mathbf{P}(\Xi)$ *буде бієкцією множини всіх відношень еквівалентності, заданих на A, на множину всіх розбиттів множини A.*

Д о в е д е н н я. Нехай Ξ, Θ – два різних відношення еквівалентності на множині A. Тоді існує пара (x, y), $x, y \in A$, яка належить до одного з цих відношень і не належить до іншого. Нехай для визначеності $(x, y) \in \Xi \setminus \Theta$. Оскільки $(x, y) \in \Xi$, то $y \in [x]_{\Xi}$, і оскільки $(x, y) \notin \Theta$, то $y \notin [x]_{\Theta}$.

Припустимо, що $P(\Xi) = P(\Theta)$, тоді знайдеться такий елемент $z \in A$, що $[z]_{\Theta} = [x]_{\Xi}$. Зокрема, $x \in [z]_{\Theta}$. Ми вже зазначали, що клас еквівалентності визначається кожним своїм елементом, тобто $[z]_{\Theta} = [x]_{\Theta}$. Отже, $[x]_{\Theta} = [x]_{\Xi}$, і ми отримали протиріччя, оскільки $y \notin [x]_{\Theta}$. Це протиріччя показує, що $P(\Xi) \neq P(\Theta)$, тобто відображення ρ ін'єктивне.

Нехай S – довільне розбиття множини A. За цим розбиттям вище ми побудували відношення $\Xi(S)$, яке є відношенням еквівалентності. Якщо x – довільний елемент множини A, то $x \in S$ для деякої підмножини $S \in S$. З визначення відношення $\Xi(S)$ випливає, що кожен елемент підмножини S належить до класу еквівалентності елемента x.

Навпаки, нехай y – такий елемент A, що $(x, y) \in \Xi(S)$. Це означає, що існує така підмножина $Q \in S$, що $x, y \in Q$. У цьому випадку $x \in S \cap Q$, і з означення розбиття отримуємо рівність $Q = S$.

Зокрема, $y \in S$. Отже, клас еквівалентності елемента x відносно $\Xi(S)$ співпадає з елементом розбиття S. Звідси і отримуємо рівність $P(\Xi) = S$. Інакше кажучи, відображення ρ буде також і сюр'єктивним, а тому і бієктивним.

Наведемо тепер приклади відношень еквівалентності. Поза математикою відношення еквівалентності виникають всякий раз, коли потрібно проводити класифікацію об'єктів тієї чи іншої природи. Так, наприклад, множину всіх людей можна розбити на нації, держави та ін.; множину всіх тварин на види і т.д.

Якщо A – довільна множина, то можна розглянути два крайніх випадки: $\Xi = A \times A$ та $\Xi = \{(x, x) \mid x \in A\}$ (діагональ декартового добутку $A \times A$). Всі інші відношення еквівалентності розташовані між цими двома крайніми позиціями.

У геометрії прикладом відношення еквівалентності є відношення паралельності, що визначене на множині всіх прямих площини або простору або на множині всіх площин простору.

Інший приклад: на множині $P = \{(x, y) \mid x > 0, y > 0\}$ розглянемо множину всіх гіпербол $G_a = \{(x, y) \mid xy = a\}$. Очевидно, що система $\{ G_a \mid a > 0 \}$ дає розбиття множини P, так що, дійсно, приходимо до відношення еквівалентності.

Нехай M – множина всіх фундаментальних послідовностей $s = (x_n)_{n \in N}$, що складаються з раціональних чисел. Розглянемо відношення ς на множині M, яке визначається за таким правилом: $(s, r) \in \varsigma$ тоді і тільки тоді, коли

$$\lim_{n \to \infty} (x_n - y_n) = 0,$$

тут $r = (y_n)_{n \in N}$. Неважко упевнитись у тому, що ς – відношення еквівалентності.

Нехай $M = [0, 1]$, визначимо відношення P за таким правилом: $(x, y) \in P$ тоді і тільки тоді, коли $x – y$ – раціональне число. Легко бачити, що P також буде відношенням еквівалентності.

Наведемо ще один важливий арифметичний приклад.

Два цілих числа a, b називаються конгруентними за модулем m, якщо a – b ділиться на m; цей факт ми будемо позначати так: $a \equiv b \pmod m$.

Неважко упевнитись у тому, що відношення конгруентності буде відношенням еквівалентності. Пізніше ми більш детально розглянемо властивості відношення конгруентності.

Цей приклад має таке важливе узагальнення.

Нехай G – група, H – її підгрупа. *Визначимо відношення Σ_H за таким правилом. Для елементів $x, y \in G$ включення $(x, y) \in \Sigma_H$ має місце тоді і тільки тоді, коли $xy^{-1} \in H$.*

Відношення Σ_H рефлексивне: $xx^{-1} = e \in H$, тобто $(x, x) \in \Sigma_H$.

Відношення Σ_H симетричне: якщо $xy^{-1} \in H$, то оскільки H – підгрупа, вона містить також і елемент $(xy^{-1})^{-1} = (y^{-1})^{-1}x^{-1} = yx^{-1}$, тобто $(y, x) \in \Sigma_H$.

Відношення Σ_H транзитивне: якщо $xy^{-1}, yz^{-1} \in H$, то оскільки H – підгрупа, вона містить також і добуток цих елементів: $(xy^{-1})(yz^{-1}) = xz^{-1}$, тобто $(x, z) \in \Sigma_H$.

Отже, Σ_H – відношення еквівалентності.

З'ясуємо тепер, з яких елементів складається клас еквівалентності елементу x за відношенням Σ_H. Нехай $(x, y) \in \Sigma_H$, тоді $xy^{-1} = h \in H$. Множимо обидві частини рівності $xy^{-1} = h$ спочатку справа на y, а потім зліва на h^{-1}, отримаємо $y = h^{-1}x$.

Визначимо тепер

$$Hx = \{ux \mid u \in H\}.$$

Підмножина Hx називається правим суміжним класом (елемента x) групи G за підгрупою H.

Відтак, кожний елемент, еквівалентний x (відносно Σ_H), належить до Hx. Навпаки, якщо $z \in Hx$, то $z = ux$ для деякого елементу $u \in H$. Маємо тепер $xz^{-1} = x(ux)^{-1} = x\,x^{-1}\,u^{-1} = u^{-1} \in H$, а це означає, що $(x, z) \in \Sigma_H$. Отже, класи еквівалентності за відношенням Σ_H – це праві суміжні класи за підгрупою H і тільки вони. Звідси випливає, що правий суміжний клас Hx визначається кож-

ним своїм елементом (представником цього суміжного класу): якщо $y \in Hx$,то $Hy = Hx$; два правих суміжних класа або збігаються або мають пустий перетин, і група буде об'єднанням усіх правих суміжних класів. Інакше кажучи, система всіх правих суміжних класів за підгрупою H буде розбиттям групи. Якщо $H = \langle e \rangle$, то $xH = \{x\}$ для кожного елемента $x \in G$, так що отримуємо найбільше розбиття групи G, яке складається з усіх одноелементних підмножин; якщо ж $H = G$, то отримуємо найменше розбиття, що складається з однієї множини – самої групи G.

Аналогічно, *нехай тепер*

$$xH = \{xu \mid u \in H\}.$$

Підмножина xH називається лівим суміжним класом (елемента x) групи G за підгрупою H.

Ліві суміжні класи – це точно класи еквівалентності за відношенням Λ_H, що визначається за таким правилом:

Для елементів $x, y \in G$ включення $(x, y) \in \Lambda_H$ має місце тоді і тільки тоді, коли $y^{-1}x \in H$.

Тому кожний лівий суміжний клас xH визначається кожним своїм елементом: якщо $y \in xH$, то $yH = xH$; два лівих суміжних класа або збігаються або мають пустий перетин, і група буде об'єднанням усіх лівих суміжних класів. Інакше кажучи, система всіх лівих суміжних класів за підгрупою H також буде розбиттям групи.

Зазначимо, що з такими розбиттями в групах ми вже зустрічались. Раніше ми розглядали у групі S_n підмножини $P_i = \{\pi \in S_n \mid \pi(n) = i\}$ і показали, що $P_i \cap P_j = \varnothing$, якщо $i \neq j$ та

$$S_n = \cup_{1 \le i \le n} P_i = P_1 \cup P_2 \cup \ldots \cup P_n.$$

Оскільки $P_n = St(n)$, то P_n – підгрупа S_n. Також ми показали фактично, що $P_i = \tau_{i\,n} \circledcirc P_n$.

Між розбиттями на ліві та праві суміжні класи існує такий зв'язок.

4.2. Пропозиція. *Нехай G – група, H – її підгрупа. Відображення* ν*: Hx* $\mapsto$ $x^{-1}H$ *буде бієкцією множини всіх правих суміжних класів за підгрупою H на множину всіх лівих суміжних класів за підгрупою H.*

Д о в е д е н н я. Спочатку слід показати, що ν – дійсно відображення, тобто, що воно не залежить від вибору представника суміжного класу. Нехай y – інший представник суміжного класу Hx, тоді y = ux для деякого елементу $u \in H$. За пропозицією 1.5 $y^{-1} = x^{-1}u^{-1}$, а тому $y^{-1} \in x^{-1}H$, отже $y^{-1}H = x^{-1}H$. Далі, відображення ν – ін′єктивне: припустимо, що $x^{-1}H = \nu(Hx) = \nu(Hy) = y^{-1}H$. Тоді $y^{-1} = x^{-1}v$ для деякого елементу $v \in H$, а тому $y = v^{-1}x \in Hx$. Це означає, що Hx = Hy.

Відображення ν – сюр’єктивне: розглянемо довільний лівий суміжний клас zH. Маємо $\nu(Hz^{-1}) = (z^{-1})^{-1}H = zH$. Таким чином, ν – бієктивне відображення.

Нехай G – група, H – її підгрупа. У кожному лівому (відповідно правому) суміжному класі групи G за підгрупою H виберемо по одному представнику і об’єднаємо їх у підмножину ***l*** T (G, H) (відповідно r T (G, H)), яку будемо називати *повною множиною представників лівих (відповідно правих) суміжних класів групи G за підгрупою H або лівою (відповідно правою) трансферсаллю підгрупи H у групі G.*

Отже, $G = \cup_{x \in \mathbf{l}\,T(G,\,H)} xH$ (відповідно $G = \cup_{x \in \mathbf{r}T(G,\,H)} Hx$)

і рівність xH = yH (відповідно Hx = Hy) для $x, y \in \mathbf{l}$ T (G, H) відповідно $x, y \in \mathbf{r}$ T (G, H)) означає, що x = y.

Нехай G – група, H – її підгрупа. *Якщо множина всіх правих суміжних класів групи G за підгрупою H є скінченою, то говорять, що підгрупа H має в групі G скінчений індекс, а кількість всіх (різних) правих суміжних класів називається індексом підгрупи H в групі G.* З пропозиції 4.2 випливає, що це число дорівнює кількості всіх (різних) лівих суміжних класів групи G за підгрупою H, тобто *індекс підгрупи H в групі G може бути визначений також як кількість всіх (різних) лівих суміжних класів*

групи G за підгрупою H. Індекс підгрупи H в групі G будемо позначати через $|G : H|$. Інакше кажучи,

$$|G : H| = |\mathbf{r}\, T(G, H)| = |\mathbf{l}\, T(G, H)|.$$

У випадку, коли H = G маємо $|G : H| = 1$, і навпаки. Якщо група G – скінченна і $H = \langle e \rangle$, то $|G : H| = |G|$.

Будемо говорити, що підгрупа H має в групі G нескінченний індекс, якщо множина всіх (правих) суміжних класів групи G за підгрупою H нескінченна.

4.3. Теорема. *Нехай G – група, H, K – такі її підгрупи, що* $K \leq H$. *Нехай* $T = \mathbf{l}\, T(G, H)$, $U = \mathbf{l}\, T(H, K)$. *Тоді підмножина* $R = \{ tu \mid x \in T, u \in U \}$ *буде повною множиною представників (лівих) суміжних класів групи G за підгрупою K.*

Д о в е д е н н я. Маємо

$$G = \cup_{t \in T} tH,\ H = \cup_{u \in U} uK.$$

Якщо x – довільний елемент групи G, то $x \in tH$ для деякого $t \in T$. Інакше кажучи, $x = th$, де $h \in H$. Знайдеться такий елемент $u \in U$, що $h \in uK$, тобто $h = uv$ для деякого $v \in K$. Отже, $x = tuv$. Звідси випливає рівність

$$G = \cup_{t \in T,\, u \in U} tuK.$$

Нехай $tuK = zwK$, де $t, z \in T$, $u, w \in U$. Зокрема $tu = zwv$ для деякого елемента $v \in K$. Оскільки $u, w, v \in H$, то звідси випливає, що суміжні класи tH та zH мають непустий перетин, тобто співпадають. З визначення множини T випливає рівність $t = z$. Множимо обидві частини рівності $tu = zwv$ зліва на елемент t^{-1} і отримаємо $u = wv$. Це означає, що $u \in wK$, а оскільки u належить також і до суміжного класу uK, то суміжні класи uK та wK мають непустий перетин, тобто співпадають. З визначення множини U випливає рівність $u = w$.

Інакше кажучи, рівність $tuK = zwK$ можлива тільки у випадку, коли $t = z$ і $u = w$. Разом з рівністю $G = \cup_{y \in R} yK$ це означає, що підмножина R буде повною множиною представників (лівих) суміжних класів групи G за підгрупою K.

4.4. Наслідок. *Нехай G – група, H, K – такі її підгрупи, що K ≤ H. Підгрупа K тоді і тільки тоді має скінчений індекс в G, коли скінченними будуть індекси* $|G : H|$ *та* $|H : K|$. *У цьому випадку має місце рівність* $|G : K| = |G : H|\,|H : K|$.

4.5. Наслідок (теорема Лагранжа). *Нехай G – скінченна група, H – її підгрупа. Тоді* $|G| = |G : H|\,|H|$. *Інакше кажучи, порядок підгрупи скінченної групи завжди буде дільником порядку всієї групи.*

Дійсно, покладемо у цьому випадку K = < e >, тоді $|H : K| = |H|$.

4.6. Наслідок. *Порядок елемента скінченої групи буде дільником порядку всієї групи.*

Дійсно, порядок елемента співпадає з порядком циклічної підгрупи, яку породжує цей елемент.

4.7. Наслідок. *Нехай G – скінченна група. Якщо* $|G|$ *– просте число, то G – циклічна група.*

Дійсно, нехай $e \neq g \in G$. Тоді підгрупа <g> має хоча б один неодиничний елемент, а тому $|< g >| > 1$. Оскільки $|G|$ – просте число, з теореми Лагранжа випливає рівність $|< g >| = |G|$, з якої отримуємо, що $< g > = G$.

4.8. Наслідок. *Нехай G – група, H, K – її підгрупи. Якщо індекси* $|G : H|$ *та* $|G : K|$ *скінченні, то скінченним буде також індекс* $|G : H \cap K|$. *Більше того,* $|G : H \cap K| \leq |G : H|\,|G : K|$.

Д о в е д е н н я. Нехай $L = H \cap K$, $T = \mathbf{l}\,т(H, L)$, $x, y \in T$, $x \neq y$. Припустимо, що xK = yK, тоді $x^{-1}y \in K$. Оскільки $x, y \in H$, то $x^{-1}y \in H \cap K = L$, а тому xL = yL. З вибору елементів x, y випливає рівність x = y. Отримане протиріччя показує, що суміжні класи xK та yK будуть різними. З того факту, що індекс $|G : K|$ скінченний, випливає, що скінченним буде індекс $|H : L|$, більше того $|H : L| \leq |G : K|$. З наслідку 2.4 отримуємо

$$|G : H \cap K| = |G : H|\,|H : H \cap K| \le |G : H|\,|G : K|.$$

4.9. Наслідок (теорема Пуанкаре). *Перетин скінченної множини підгруп, кожна з яких має скінченний індекс, буде підгрупою скінченного індексу.*

Скористаємось тепер наведеними вище результатами, щоб отримати більш детальну інформацію про підгрупи циклічних груп. Раніше вже довели теорему 2.14, яка показує, що кожна підгрупа циклічної групи сама буде циклічною.

4.10. Теорема. *Нехай $G = \langle g \rangle$ – циклічна група, H – її підгрупа.*

(1) Якщо G – нескінченна, $H = \langle g^n \rangle$ і $n \neq 0$, то H також нескінченна, $|G : H| = n$; якщо $n = 0$, то індекс $|G : H|$ – нескінченний.

(2) Якщо G має скінченний порядок r, то $|H|$ – дільник числа r. Навпаки, для кожного дільника d числа r існує одна і тільки одна підгрупа порядку d, яка співпадає з $\langle g^{r/d} \rangle$.

Д о в е д е н н я. (1) Нехай x – довільний елемент групи G, тоді $x = g^k$ для деякого $k \in \mathbf{Z}$. Ділимо k на n: $k = nq + s$, де $0 \le s < n$. Маємо

$$g^k = g^{nq + s} = (g^n)^q g^s \in g^s H.$$

Звідси випливає рівність

$$G = \cup_{0 \le s < n} g^s H.$$

Припустимо тепер, що $g^s H = g^t H$, де $0 \le s \le t < n$. Маємо тоді $g^t = g^s u$ для деякого $u \in H$, а оскільки $u = g^{nm}$ для деякого $m \in \mathbf{Z}$, то

$$g^t = g^s g^{nm} = g^{s + nm}.$$

З того факту, що G – нескінченна циклічна група, отримуємо рівність $t = s + nm$, а з вибору чисел s, t випливає $m = 0$, тобто $s = t$. Таким чином, всі суміжні класи $g^s H$, де $0 \le s < n$, різні, а це означає, що $|G : H| = n$.

Нехай тепер $|G| = r < \infty$. З теореми Лагранжа (наслідок 4.5) отримуємо, що $|H|$ є дільником числа r. Навпаки, нехай d – дільник числа r: r = db. Якщо припустити, що $|g^b| = u < d$, то $g^{bu} = e$ і $bu < bd = r$, що суперечить рівності $|g| = r$. Отже, $|g^b| = d$. Нехай $< g^v >$ – інша підгрупа порядку d. Тоді $g^{vd} = e$, отже, vd ділиться на r, а тому v ділиться на b. У свою чергу, звідси випливає включення $<g^v> \le <g^b>$. Але ж $d = |< g^v >|$ і $d = |< g^b >|$, так що $< g^v > = < g^b >$.

4.11. Наслідок. *Група G тоді і тільки тоді має тільки дві підгрупи (тобто $< e >$ та G), коли $|G|$ – просте число.*

Д о в е д е н н я. Нехай g – неодиничний елемент групи G, тоді $<g> = G$, зокрема група G – циклічна. Якщо припустити, що $|g|$ – нескінченний, то $< e > \neq < g^2 > \neq <g>$. Це протиріччя показує, що $|g|$ – скінченний. Враховуючи теорему 4.10, отримуємо, що він мусить бути простим числом.

З теореми 4.10 випливає, що кожна неодинична підгрупа нескінченної циклічної групи має скінченний індекс. У роботі [1] Ю. Г. Федоров довів, що і навпаки, якщо в нескінченній групі G кожна неодинична підгрупа має скінченний індекс, то G – циклічна група.

РОЗДІЛ 5
НОРМАЛЬНІ ПІДГРУПИ. ФАКТОРГРУПИ

У попередньому розділі розглянули розбиття групи, що складаються з правих та лівих суміжних класів за деякою підгрупою. Природно виникає питання: чи будуть ці два розбиття (за однією і тією ж самою підгрупою) суттєво різними, чи вони можуть співпадати? Інакше кажучи, чи будуть еквівалентності Σ_H та Λ_H різними? Або точніше: для яких підгруп H еквівалентності Σ_H та Λ_H (а значить і розбиття на праві та ліві суміжні класи за підгрупою H) співпадають, а для яких вони різні? Зразу видно, що для абелевих груп має місце співпадання. Але на простому прикладі групи S_3 можна упевнитись у тому, що обидві можливості мають місце.

Виберемо підгрупи:

$$\left\{ \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \right\}$$

$H = \mathbf{A_3} = \varepsilon, \pi = \begin{matrix} 2 & 3 & 1 \end{matrix}, \pi^2 = \begin{matrix} 3 & 1 & 2 \end{matrix} \}$, та $K = < \tau_{12} >$.

Оскільки $|H| = 3$, то з теореми Лагранжа (наслідок 4.5) випливає, що $|\mathbf{S_3} : H| = 2$.

5.1. Пропозиція. *Нехай G – група, H – її підгрупа. Якщо* $|G : H| = 2$, *то для кожного* $x \notin H$ *має місце рівність* $xH = Hx = G \setminus H$.

Д о в е д е н н я. Дійсно, як праве так і ліве розбиття на суміжні класи за підгрупою H складається з двох підмножин, одна з яких – сама підгрупа H. Отже, другий елемент цього розбиття співпадає з $G \setminus H$.

Таким чином, множина всіх лівих суміжних класів групи $\mathbf{S_3}$ за підгрупою $\mathbf{A_3}$ співпадає з множиною всіх таких правих суміжних класів. Щодо підгрупи K, то для неї будемо мати три таких правих суміжних класів:

$$K, K \circledcirc \tau_{13} = \{\tau_{13}, \pi^2\}, K \circledcirc \tau_{23} = \{\tau_{23}, \pi\}$$

та три таких лівих суміжних класів:

$$K, \tau_{13} \circledcirc K = \{\tau_{13}, \pi\}, \tau_{23} \circledcirc K = \{\tau_{23}, \pi^2\}.$$

Звідси видно, що множина всіх лівих суміжних класів групи $\mathbf{S_3}$ за підгрупою K не співпадає з множиною всіх правих суміжних класів за підгрупою K.

Нехай G – група, H – така її підгрупа, що множина всіх лівих суміжних класів групи G за підгрупою H співпадає з множиною всіх таких правих суміжних класів. Якщо $x \in G$, то існує такий елемент y, що $xH = Hy$. Оскільки $x \in xH$, то $x \in Hy$, а значить $Hy = Hx$, тобто $xH = Hx$.

Підгрупа H групи G називається нормальною (в G), якщо $xH = Hx$ для кожного елементу $x \in G$.

Якщо A, B – дві підмножини групи G, то отримаємо:

$$AB = \{ab \mid a \in A, b \in\}.$$

У випадку, коли підмножина A (відповідно B) складається з одного елементу a (відповідно b) замість{a}B (відповідно A{b}) будемо використовувати більш короткий запис aB (відповідно Ab). Аналогічно можна визначити добуток будь-якого скінченного набору підмножин групи. Оскільки множення елементів групи є асоціативним, то і множення підмножин також буде асоціативним.

Дві підгрупи A, B називаються переставними, якщо AB = BA. Наступне твердження показує, яку роль відіграє ця властивість.

5.2. Пропозиція. *Нехай G – група, A, B – її підгрупи. Добуток AB тоді і тільки тоді буде підгрупою, коли підгрупи A і B – переставні. У цьому випадку $AB = < A \cup B > = < A, B >$.*

Д о в е д е н н я. Припустимо, що AB – підгрупа G. Оскільки $A \leq AB$ і $B \leq AB$, то $BA \leq AB$. Нехай x – довільний елемент AB. Оскільки AB – підгрупа, то, згідно з теоремою 2.4, разом з елементом x вона містить і елемент x^{-1}. Таким чином, $x^{-1} = uv$, для деяких елементів $u \in A$, $v \in B$. Маємо

$$x = (x^{-1}\}^{-1} = (uv)^{-1} = v^{-1} u^{-1}.$$

Знову, використовуючи теорему 2.4, отримаємо включення $u^{-1} \in A$, $v^{-1} \in B$, з яких випливає, що $v^{-1} u^{-1} \in BA$, тобто і $AB \leq BA$, що доводить рівність $AB = BA$.

Навпаки, нехай $AB = BA$. Якщо $x, y \in AB$, то $x = a_1b_1$, $y = a_2b_2$, для деяких елементів $a_1, a_2 \in A$, $b_1, b_2 \in B$. Звідси

$$xy^{-1} = (a_1b_1)(a_2b_2)^{-1} = (a_1b_1)(b_2^{-1}a_2^{-1}) = a_1 (b_1b_2^{-1}) a_2^{-1}.$$

Оскільки B – підгрупа, то за наслідком 2.5 вона містить елемент $b_1b_2^{-1}$, отже, $(b_1b_2^{-1}) a_2^{-1} \in BA = AB$, а тому $(b_1b_2^{-1}) a_2^{-1} = a_3b_3$, де $a_3 \in A$, $b_3 \in B$. Звідси

$$xy^{-1} = a_1 (a_3 b_3) = (a_1a_3)b_3 \in AB.$$

Згідно з наслідком 2.5 AB буде підгрупою.

Наступне твердження є досить корисним.

5.3. Пропозиція. *Нехай G – група, A, B, C – такі її підгрупи, що $A \leq C$. Тоді $(AB) \cap C = A (B \cap C)$.*

Д о в е д е н н я. Дійсно, $A (B \cap C) \subseteq AB$ та $A(B \cap C) \subseteq AC = C$, отже, і $A(B \cap C) \subseteq (AB) \cap C$. Навпаки, нехай $x \in (AB) \cap C$, тоді $x = ab$, де $a \in A$, $b \in B$. Маємо $b = a^{-1}x \in AC = C$, зокрема $b \in B \cap C$ і $x \in A(B \cap C)$, що доводить включення $(AB) \cap C \subseteq A(B \cap C)$, а з ним і потрібну нам рівність

$$(AB) \cap C = A (B \cap C).$$

Слід зазначити, що з включення $A \leq C$ отримуємо, що $A \cap C = A$, і доведена вище рівність може бути записана у такому вигляді

$$(AB) \cap C = (A \cap C) (B \cap C).$$

На жаль, у загальному випадку рівність $(AB) \cap C = (A \cap C) (B \cap C)$ виконується далеко не завжди.

Підгрупа групи G, яка переставна з кожною підгрупою G, називається *переставною або квазінормальною*. Найбільш важливим прикладом переставних підгруп будуть нормальні підгрупи, до яких зараз повернемось. Наведемо основні ознаки нормальних підгруп.

5.4. Пропозиція. *Нехай G – група, H – її підгрупа. Наступні твердження рівносильні:*

(1) H – нормальна підгрупа G;

(2) (xH)(yH) = xyH для будь-яких елементів x, y $\in G$*;*

(3) $x^{-1}Hx \subseteq H$ *для кожного елементу* $x \in G$*;*

(4) $x^{-1}Hx = H$ *для кожного елементу* $x \in G$.

Д о в е д е н н я. (1) ⇨ (2) Маємо

$$(xH)(yH) = x(Hy)H = x(yH)H = (xy)(HH).$$

Оскільки H – підгрупа, то для довільних елементів u, v ∈ H маємо, зважаючи на теорему 10.4, uv ∈ H, тобто $HH \subseteq H$. З іншого боку, оскільки e ∈ H, то $H = eH \subseteq HH$, а отже, HH = H. Таким чином, якщо підгрупа H нормальна, то (xH)(yH) = xyH.

(1) ⇨ (3) Для кожного x ∈ G маємо

$$x^{-1}Hx = (x^{-1}Hx)e \subseteq (x^{-1}Hx)H = (x^{-1}H)(xH) = x^{-1}x\,H = eH = H.$$

(3) ⇨ (4)Нехай x ∈ G. Тоді $x^{-1}Hx \subseteq H$. Але ж і для елемента x^{-1} подібне включення має місце: $(x^{-1})^{-1}Hx^{-1} = xHx^{-1} \subseteq H$, Тому

$$H = (x^{-1}x)H(x^{-1}x) = x^{-1}(xHx^{-1})x \subseteq x^{-1}Hx,$$

і отже, $x^{-1}Hx = H$.

(4) ⇨ (1) Маємо

$$xH = x(x^{-1}Hx) = (xx^{-1})Hx = eHx = Hx.$$

Наведемо деякі стандартні нормальні підгрупи. Зразу домовимось використовувати для запису того факту, що підгрупа H нормальна в групі G, наступне позначення $H \lhd G$.

Очевидно, $< e > \lhd G$ та $G \lhd G$. Якщо група G не має інших нормальних підгруп, то G називається *простою групою*. Антипод простої групи – група, всі підгрупи якої нормальні. Такі групи називають зараз *дедекіндовими* на честь Р. Дедекінда, який ще в 1897 р. отримав опис скінчених груп з цією властивістю [2]. Очевидно, кожна абелева група є дедекіндовою. З наслідку 4.11 випливає, що *проста абелева група – це група простого порядку*. Також слід зазначити, що кожна підгрупа центра $\zeta(G)$ групи G буде нормальною.

5.5. Наслідок. *Нехай G – група, S – система нормальних підгруп групи G . Перетин* $\cap S$ *підгруп цієї системи також буде нормальною підгрупою G.*

Д о в е д е н н я. Нехай $S = \cap S$. З наслідку 2.7 випливає, що S – підгрупа. Нехай $x \in S$, $g \in G$. Якщо U – довільна підгрупа із системи S, тоді $x \in U$, а оскільки U – нормальна, вона разом з елементом x містить і $g^{-1}xg$. Звідси випливає, що $g^{-1}xg$ належить до перетину всіх підгруп системи S, тобто до S. Залишилось застосувати пропозицію 5.4.

5.6. Наслідок. *Нехай G – група, L – локальна система нормальних підгруп групи G. Тоді об'єднання* $\cup L$ *підгруп цієї системи також буде нормальною підгрупою G.*

Д о в е д е н н я. Нехай $V = \cup L$. З наслідку 2.9 випливає, що V буде підгрупою. Нехай $x \in V$, $g \in G$. Виберемо тепер підгрупу $L \in L$, яка містить у собі елемент x. Оскільки L – нормальна підгрупа, то разом з елементом x вона містить і $g^{-1}xg$. Звідси випливає, що $g^{-1}xg \in V$, і за пропозицією 5.4 підгрупа V також нормальна.

5.7. Наслідок. *Нехай G – група, L – лінійно впорядкована система нормальних підгруп групи G. Тоді об'єднання* $\cup L$ *підгруп цієї системи також буде нормальною підгрупою G.*

5.8. Наслідок. *Нехай G – група,*

$$H_1 \leq H_2 \leq \ldots \leq H_n \leq \ldots \quad -$$

зростаюча система нормальних підгруп групи G. Тоді об'єднання $\cup_{n \in N} H_n$ *підгруп цієї системи також буде нормальною підгрупою G.*

Нехай M – підмножина групи G, S – система всіх тих нормальних підгруп, які містять у собі M. Тоді перетин $< M >^G = \cap S$ буде за наслідком 5.5 нормальною підгрупою групи G. Цей перетин називають *нормальною підгрупою, породженою підмножиною M,*

або нормальним замкненням множини M. Якщо H – нормальна підгрупа групи G, що містить у собі підмножину M, то H містить у собі і підгрупу $< M >^G$, у цьому сенсі $< M >^G$ – найменша з нормальних підгруп, які містять у собі підмножину M. Зрозуміло, що якщо M – нормальна підгрупа групи G, то $< M >^G = M$. Разом з цим маємо таке корисне твердження.

5.9. Пропозиція. *Нехай G – група, H – її нормальна підгрупа. Якщо* $M \subseteq H$, *то* $< M >^G \leq H$.

Якщо H – підгрупа групи G, то для кожного елементу $x \in G$ підмножина $x^{-1}Hx$ буде підгрупою G. Дійсно, нехай $a, b \in x^{-1}Hx$. Тоді

$$a = x^{-1}ux,\ b = x^{-1}vx$$

для деяких елементів $u, v \in H$. Маємо тепер

$$ab^{-1} = x^{-1}ux\,(x^{-1}vx)^{-1} = x^{-1}ux\,(x^{-1}v^{-1}x) = x^{-1}uv^{-1}x.$$

Оскільки H – підгрупа, $uv^{-1} \in H$. Таким чином $x^{-1}Hx$ задовольняє умову (SG 3) і за наслідком 2.5 буде підгрупою.

Будемо говорити далі, що підгрупи H і $x^{-1}Hx$ спряжені за допомогою елемента x.

З пропозиції 5.4 випливає, що *підгрупа H тоді і тільки тоді нормальна, коли вона співпадає з усякою спряженою до неї підгрупою*.

5.10. Пропозиція. *Нехай G – група, H – її підгрупа. Тоді*

$$\bigcap_{x \in G} x^{-1}Hx = \mathbf{Core}_G(H) \text{ –}$$

нормальна підгрупа групи G. Якщо H містить у собі нормальну підгрупу K, то $K \leq \mathbf{Core}_G(H)$. *Інакше кажучи,* $\mathbf{Core}_G(H)$ *– найбільша нормальна підгрупа, що міститься в підгрупі H.*

Д о в е д е н н я. Як тільки що довели, $x^{-1}Hx$ – підгрупа G для кожного $x \in G$. З наслідку 2.7 випливає, що $\mathrm{Core}_G(H)$ – підгрупа.

Нехай тепер $u \in \mathbf{Core}_G(H)$, $g \in G$. Маємо $u \in (xg^{-1})^{-1}H(xg^{-1}) = gx^{-1}H\,xg^{-1}$, тобто $u = gx^{-1}v\,xg^{-1}$ для деякого елемента $v \in H$. Звідси отримаємо $g^{-1}\,u\,g = x^{-1}\,v\,x \in x^{-1}H\,x$. Оскільки x – довільний елемент групи G, то $g^{-1}\,u\,g \in \mathrm{Core}_G(H)$. З пропозиції 5.4 випливає тепер, що $\mathbf{Core}_G(H)$ – нормальна підгрупа. Далі, оскільки підгрупа K – нормальна, то $K = x^{-1}K\,x \le x^{-1}H\,x$, тобто $K \le \bigcap_{x \in G} x^{-1}Hx = \mathbf{Core}_G(H)$.

За аналогією з підгрупами *будемо говорити, що елементи g, y спряжені у групі G, якщо існує такий елемент $u \in G$, що $g = u^{-1}yu$.* Точніше у цьому випадку будемо говорити, що *елементи g та y спряжені за допомогою елемента u.*

Слід зазначити, що відношення спряженості є відношенням еквівалентності. Дійсно, $g = e^{-1}ye$, так що спряженість є рефлексивною. Якщо $g = u^{-1}yu$, то

$$(u^{-1})^{-1}\,g\,u^{-1} = u\,g\,u^{-1} = u\,(u^{-1}\,y\,u)\,u^{-1} = (uu^{-1})\,y\,(uu^{-1}) = y,$$

тобто елемент y спряжений з елементом g за допомогою елемента u^{-1}, отже, спряженість є симетричною. Нарешті, нехай $g = u^{-1}yu$ та $y = v^{-1}zv$, тоді маємо

$$g = u^{-1}yu = g = u^{-1}(v^{-1}zv)\,u = u^{-1}\,v^{-1}\,z\,vu = (vu)^{-1}z\,(vu),$$

і це показує, що спряженість є транзитивною.

Якщо x – елемент групи G, то *клас еквівалентності елементу x за відношенням спряженості називається класом спряженості елементу x і позначається x^G*, інакше кажучи, $x^G = \{g^{-1}\,x\,g \mid g \in G\}$. Класи спряжених елементів складуть розбиття групи G.

5.11. Пропозиція. *Нехай G – група, $x \in G$. Елементи $u^{-1}xu$ та $v^{-1}xv$ співпадають тоді і тільки тоді, коли $u\,v^{-1} \in C_G(x)$.*

Д о в е д е н н я. Дійсно, якщо $u^{-1}xu = v^{-1}xv$, то отримаємо $vu^{-1}\,xuv^{-1} = x$, або $(uv^{-1})^{-1}\,x(uv^{-1}) = x$. А це означає, що $u\,v^{-1} \in C_G(x)$. Повторюючи все в зворотньому порядку, отримаємо достатність.

5.12. Наслідок. *Нехай G – група, x ∈ G. Існує бієкція множини x^G на множину $\{g\, C_G(x) \mid g \in G\}$.*

Д о в е д е н н я. Дійсно, розглянемо відображення ϕ, яке визначене за таким правилом: ϕ: $u^{-1}xu \longrightarrow u\, C_G(x)$, $u \in G$. Відображення ϕ – ін'єктивне: якщо $uC_G(x) = vC_G(x)$, то $uv^{-1} \in C_G(x)$, і пропозиція 5.11 доводить рівність $u^{-1}xu = v^{-1}xv$. Сюр'єктивність відображення ϕ очевидна.

5.13. Наслідок. *Нехай G – група, x ∈ G. Якщо множина x^G скінченна, то скінченною буде і індекс $|G : C_G(x)|$. Більше того, $|x^G| = |G : C_G(x)|$.*

Нехай G – група. Покладемо ***FC(G)*** *= { x ∈ G* $\mid x^G$ *– скінченна}.*

5.14. Наслідок. *Нехай G – група. Підмножина **FC**(G) є підгрупою.*

Д о в е д е н н я. Дійсно, з рівності $(g^{-1}xg)^{-1} = g^{-1}x^{-1}g$ випливає, що $(x^{-1})^G = (x^G)^{-1}$, зокрема якщо клас x^G є скінченним, то скінченним буде і клас $(x^{-1})^G$.

Далі рівність $g^{-1}(xy)g = g^{-1}xg\ g^{-1}yg$ доводить включення $(xy)^G \subseteq x^G\, y^G$, зокрема, якщо класи спряженості x^G та y^G є скінченними, то скінченним буде і $(xy)^G$. Застосування теореми 2.4 доводить тепер, що FC(G) буде підгрупою.

Розглянемо тепер множину **L** (G) усіх підгруп групи G. Як і для елементів, можна довести, що відношення "підгрупи H і K є спряженими" буде відношенням еквівалентності на множині **L** (G). Як і для елементів, будемо говорити *про класи спряжених підгруп, клас спряженості підгрупи H буде позначатися через* $\mathbf{cl}_G(H)$ (тут зазначається, що розглядається клас спряженості H у всій групі G, бо можливо розглядати і класи спряженості у підгрупах, які містять у собі H). З пропозиції 5.4 випливає, що підгрупа H тоді і тільки тоді нормальна в G, коли $\mathbf{cl}_G(H) = \{H\}$.

Нехай G – група, H – її підгрупа. *Підмножина*

$$N_G(H) = \{ x \in G \mid x^{-1}Hx = H \}$$

називається нормалізатором підгрупи H в групі G.

5.15. Пропозиція. *Нехай G – група, H – її підгрупа. Нормалізатор $N_G(H)$ буде підгрупою групи G.*

Д о в е д е н н я. Нехай $x \in N_G(H)$, тоді $x^{-1}Hx = H$. Звідси отримуємо, що і $xHx^{-1} = (x^{-1})^{-1}Hx^{-1} = H$, тобто $x^{-1} \in N_G(H)$. Нехай $x, y \in N_G(H)$, маємо

$$(xy)^{-1}\ H(xy) = y^{-1}x^{-1}H\ xy = y^{-1}(x^{-1}Hx)y\ =\ y^{-1}Hy = H,$$

тобто разом з елементами x, y підмножина $N_G(H)$ містить і їх добуток. За теоремою 2.4 $N_G(H)$ буде підгрупою. Тообто, $N_G(H)$ – це найбільша підгрупа, в якій H нормальна.

5.16. Пропозиція. *Нехай G – група, $x \in G$. Елементи $u^{-1}Hu$ та $v^{-1}Hv$ співпадають тоді і тільки тоді, коли $uv^{-1} \in N_G(H)$.*

Доведення аналогічне доведенню пропозиції 5.11.

5.17. Наслідок. *Нехай G – група, $H \leq G$. Існує бієкція множини $\boldsymbol{cl}_G(H)$ на множину $\{g\ N_G(H)\ \mid g \in G\}$.*

Д о в е д е н н я. Дійсно, розглянемо відображення ϕ, яке визначене за таким правилом:

$$\phi:\ u^{-1}xu \longrightarrow u\ N_G(H),\ u \in G.$$

Відображення ϕ – ін'єктивне: якщо $uN_G(H) = vN_G(H)$, то $uv^{-1} \in N_G(H)$, і пропозиція 5.16 доводить рівність $u^{-1}Hu = v^{-1}Hv$. Сюр'єктивність відображення ϕ очевидна.

13.18. Наслідок. *Нехай G – група, $H \leq G$. Якщо множина $\boldsymbol{cl}_G(H)$ скінченна, то скінченою буде і індекс $|G : N_G(H)|$. Більше того, $|\boldsymbol{cl}_G(H)| = |G : N_G(H)|$.*

З п. (2) пропозиції 5.4 випливає, що множина всіх суміжних класів групи G за нормальною підгрупою H буде сталою відносно операції множення підмножин. Більше того, відносно звуження цієї операції ця множина буде групою. Дійсно, асоціативність операції

множення підмножин вже зазначали вище, зокрема множення суміжних класів є асоціативним. Одиничним елементом буде сама підгрупа H:

$$(xH)H = xHH = xH,\ H(xH) = (Hx)H = (xH)H = xH.$$

Нарешті, оберненим елементом до суміжного класу xH буде елемент $x^{-1}H$:

$$(xH)(x^{-1}H) = (xx^{-1})H = eH = (x^{-1}x)H = (x^{-1}H)(xH).$$

Цю групу називають факторгрупою групи G за нормальною підгрупою H і позначають через G/H.

Слід зазначити, що деякі властивості групи наслідуються усіма її факторгрупами. Наприклад, якщо G – абелева, то і кожна її факторгрупа буде абелевою. Дійсно, маємо xH yH = xyH = yx H = yH xH для кожної нормальної підгрупи H і довільних елементів x, y групи G.

А ось інші властивості, наприклад властивість “бути нескінченною”, не переносяться на всі факторгрупи. У цьому можна упевнитись, розглянувши (адитивну) групу Z цілих чисел та її підгрупу n**Z**. З теореми 4.10 випливає рівність $|Z/nZ| = n$, яка показує, що нескінченна група Z має скінченні факторгрупи. З іншого боку, очевидно, що будь-яка факторгрупа скінченної групи буде скінченною. Таким чином, одне з перших природних питань, яке виникає при розгляді кожного конкретного класу груп **X,** це питання про те, чи буде **X** замкненим відносно формування факторгруп? З факторгрупами пов’язане також і інше питання. Якщо H = <e>, то для кожного елементу $x \in G$ маємо xH = x<e> = {x}, а також xH yH = {x}{y} = {xy}. Це показує, що факторгрупа G/<e> практично нічим не відрізняється від групи G (використовуючи термінологію наступного параграфа, можна сказати, що G та G/<e> ізоморфні). Зокрема вони мають однакові (алгебраїчні) властивості.

Факторгрупа G/H називається власною, якщо H – неодинична нормальна підгрупа.

Якщо x, y – елементи групи G, то *елемент* $[x, y] = x^{-1}y^{-1}xy$ *будемо називати комутатором елементів x та y*.

Якщо $xy = yx$, то звідси послідовно отримуємо $y^{-1}xy = x$ та потім $x^{-1}y^{-1}xy = 1$; навпаки, повторюючи теж саме у зворотньому порядку, з рівності $[x, y] = 1$ отримаємо $xy = yx$. Таким чином, *елементи x та y тоді і тільки тоді переставні, коли* $[x, y] = 1$.

Підгрупу групи G, породжену підмножиною $\{[x, y] \mid x, y \in G\}$, *називають комутантом або похідною підгрупою групи G та позначають через* $[G, G]$.

З очевидної рівності $g^{-1}[x, y]g = [g^{-1}xg, g^{-1}yg]$ випливає, що $D = [G, G]$ – нормальна підгрупа групи G. На роль комутанта вказує таке твердження.

5.19. Пропозиція. *Нехай G – група.*

(1) Факторгрупа G/[G, G] – абелева.

(2) Якщо H – така нормальна підгрупа групи G, що G/H – абелева, то $[G, G] \leq H$.

Д о в е д е н н я. Дійсно, розглянемо факторгрупу G/D, де $D = [G, G]$. Для кожної пари суміжних класів xD, yD маємо

$$[xD, yD] = (xD)^{-1}(yD)^{-1}(xD)(yD) = (x^{-1}D)(y^{-1}D)(xD)(yD) = x^{-1}y^{-1}xyD = [x, y]D = D,$$

що доводить абелевість G/D. Нехай тепер H – така нормальна підгрупа групи G, що факторгрупа G/H – абелева. Аналогічним чином можна довести рівність $H = [x, y]H$, з якої випливає включення $[x, y] \in H$ для кожної пари елементів $x, y \in G$. Звідси випливає включення $[G, G] \leq H$.

РОЗДІЛ 6
ГОМОМОРФІЗМИ ГРУП

Розглянемо спочатку факторизацію звичайних відображень. Нехай A – множина, на якій задано відношення еквівалентності Ξ, і введемо нову *множину A/Ξ, елементами якої будуть класи еквівалентності за відношенням еквівалентності Ξ. Цю множину будемо називати фактор-множиною множини A за відношенням еквівалентності Ξ.* Разом з тим виникає таке відображення

$$\sigma_{\Xi} : A \longrightarrow A/\Xi,$$

що визначається за правилом:

$$\sigma_{\Xi}(a) = [a]_{\Xi},\ a \in A.$$

Відображення σ_{Ξ} називається канонічною сюр'єкцією A на фактор-множину A/Ξ.

Нехай A, B – множини, f: A $\longrightarrow$ B – відображення. Пов'яжемо із цим відображенням бінарне відношення $\Delta(f)$, яке визначається за правилом: $(x, y) \in \Delta(f)$ тоді і тільки тоді, коли $f(x) = f(y)$, $x, y \in A$. Відношення $\Delta(f)$, очевидно, рефлексивне: $f(x) = f(x)$; симетричне: рівності $f(x) = f(y)$ та $f(y) = f(x)$ рівносильні; транзитивне: з рівностей $f(x) = f(y)$ та $f(y) = f(z)$ випливає $f(x) = f(z)$, $x, y, z \in A$. Таким чином, $\Delta(f)$ – відношення еквівалентності на множині A.

Розглянемо фактор-множину $A/\Delta(f)$ та пов'яжемо з нею відображення $\psi_f : A/\Delta(f) \longrightarrow \mathrm{Im}\ f$, яке визначається за правилом: $\psi_f([a]_{\Delta(f)}) = f(a)$ для кожного класу еквівалентності $[a]_{\Delta(f)} \in A/\Delta(f)$.

Перш за все потрібно переконатися у тому, що ψ_f визначено коректно, тобто не залежить від вибору представника класу еквівалентності. Дійсно, нехай c – елемент множини A, для якого $[c]_{\Delta(f)} = [a]_{\Delta(f)}$. Тоді за самим визначенням відношення $\Delta(f)$ маємо $f(a) = f(c)$, що і доводить коректність.

Відображення ψ_f є бієктивним: для кожного елементу $b \in \mathrm{Im}\ f$ існує елемент $a \in A$ з властивістю $b = f(a)$, так що маємо

$\psi_f([a]_{\Delta(f)}) = f(a) = b$, тобто ψ_f – сюр'єктивне; якщо $\psi_f([a]_{\Delta(f)}) = \psi_f([c]_{\Delta(f)})$, то за визначенням ψ_f маємо $f(a) = f(c)$, а це означає, що $(a, c) \in \Delta(f)$ і отже, $[c]_{\Delta(f)} = [a]_{\Delta(f)}$, таким чином ψ_f – ін'єктивне, а отже, і бієктивне.

Нарешті розглянемо наступний добуток відображень $j_D \circ \psi_f \circ \sigma_{\Delta(f)}$, де $D = \operatorname{Im} f$, а j_D – канонічна ін'єкція D в B. Для довільного елементу $a \in A$ маємо

$$j_D \circ \psi_f \circ \sigma_{\Delta(f)}(a) = j_D(\psi_f(\sigma_{\Delta(f)}(a))) = j_D(\psi_f([a]_{\Delta(f)})) = j_D(f(a)) = f(a).$$

Оскільки область визначення відображення $j_D \circ \psi_f \circ \sigma_{\Delta(f)}$ співпадає з A, а область значень – з B, то отримаємо рівність $f = j_D \circ \psi_f \circ \sigma_{\Delta(f)}$. Підводячи підсумок, сформулюємо все доведене вище у вигляді теореми.

6.1. Теорема. *Нехай A, B – множини, $f: A \longrightarrow B$ – відображення.*

(1) відношення $\Delta(f)$, визначене за правилом:

$$(x, y) \in \Delta(f) \text{ тоді і тільки тоді, коли } f(x) = f(y),\ x, y \in A,$$

є відношенням еквівалентності на множині A.

(2) відображення $\psi_f : A/\Delta(f) \longrightarrow \operatorname{Im} f$, визначене за правилом: $\psi_f([a]_{\Delta(f)}) = f(a)$ для кожного класу еквівалентності $[a]_{\Delta(f)} \in A/\Delta(f)$, є біективним.

(3) $f = j_D \circ \psi_f \circ \sigma_{\Delta(f)}$.

Зокрема, довільне відображення є добутком сюр'єкції, біекції та ін'єкції.

Розклад відображення, отриманий у теоремі 6.1, називають канонічним.

Повернемось тепер до груп та їх відображень. Природно будемо розглядати не будь-які відображення, а ті, що зберігають операцію, тобто гомоморфізми груп. Почнемо з елементарних властивостей групових гомоморфізмів.

6.2. Пропозиція. *Нехай G, U – групи, f: G ⟶ U – гомоморфізм.*

(1) f(e) = e_U – одиничний елемент групи U.

(2) Якщо f(x) = u, то $f(x^{-1}) = u^{-1}$.

(3) Якщо H – підгрупа G, то її образ f(H) буде підгрупою в U. Зокрема f(G) =Im f – підгрупа U.

(4) Якщо V – підгрупа U, то її повний прообраз $f^{-1}(V)$ – підгрупа в G.

(5) Якщо V – нормальна підгрупа U, то її повний прообраз $f^{-1}(V)$ – нормальна підгрупа в G. Зокрема,

Ker $f = \{ x \in G \mid f(x) = e_U \} = f^{-1}(< e_U >)$ *– нормальна підгрупа G.*

Д о в е д е н н я. (1) За визначенням одиничного елемента маємо ex = x для кожного елемента $x \in G$, зокрема ee = e. Звідси отримуємо:

$$f(e) = f(ee) = f(e)f(e).$$

Кожний елемент групи U, зокрема f(e), має обернений; множимо на цей обернений обидві частини останньої рівності і приходимо до рівності

$$e_U = e_U f(e) = f(e).$$

(2) За визначенням елементу, оберненого до x, маємо $xx^{-1} = e = x^{-1}x$. Звідси отримуємо

$$f(x)f(x^{-1}) = f(x\, x^{-1}) = f(e) = e_U = f(e) = f(x^{-1}x) = f(x^{-1})\, f(x).$$

Ці рівності показують, що $u^{-1} = f(x^{-1})$.

(3) Нехай $u, v \in f(H)$, тоді знайдуться елементи $a, c \in H$, для яких u = f(a), v = f(c). Маємо $uv^{-1} = f(a)\, f(c^{-1}) = f(a\, c^{-1}) \in f(H)$, оскільки H – підгрупа G. Згідно з наслідком 2.5 f(H) буде підгрупою U.

(4) Нехай $x, y \in f^{-1}(V)$, тоді f(x), f(y) ∈ V. Знову застосуємо критерій підгрупи: $f(xy^{-1}) = f(x)\, f(y^{-1}) = f(x)(f(y))^{-1} \in V$ і отже, $xy^{-1} \in f^{-1}(V)$. За наслідком 2.5 $f^{-1}(V)$ буде підгрупою G.

(5) Нехай g – довільний елемент групи G, x $\in$ $f^{-1}(V)$. Розглянемо елемент $g^{-1}xg$, точніше його образ відносно f:

$$f(g^{-1}xg) = f(g^{-1})f(x)f(g) = (f(g))^{-1} f(x)\, f(g) \in V,$$

оскільки V $\lhd$ U. Це означає, що $g^{-1}xg \in f^{-1}(V)$, а отже, $f^{-1}(V)$ – нормальна підгрупа G.

Слід зазначити, що образ нормальної підгрупи не завжди буде нормальною підгрупою; у цьому можна упевнитись, розглянувши канонічну ін'єкцію ненормальної підгрупи в групу.

Нормальну підгрупу ***Ker*** *f називають ядром гомоморфізма f.*

Розглянемо тепер класичні теореми про гомоморфізми.

6.3. Теорема (теорема про мономорфізми). *Нехай G, U – групи. Гомоморфізм f: G* $\longrightarrow$ *U тоді і тільки тоді буде мономорфізмом, коли* ***Ker*** *f = <e>. Якщо f: G* $\longrightarrow$ *U – мономорфізм, то G* $\cong$ Im f.

Д о в е д е н н я. Дійсно, якщо f – мономорфізм, то з x $\neq$ e випливає $f(x) \neq f(e) = e_U$. Це означає, що жодний неодиничний елемент x не може належати до **Ker** f, отже **Ker** f = <e>. Навпаки, нехай **Ker** f = <e>, x, y – такі елементи групи G, що f(x) = f(y). Тоді маємо

$$e_U = f(e) = f(xy^{-1}) = f(x)f(y^{-1}) = f(x)(f(y))^{-1}, \text{ і отже,}$$

$xy^{-1} \in$ **Ker** f. Це означає, що $xy^{-1} = e$, тобто x = y, а значить, f – ін'єктивне відображення.

6.4. Теорема (теорема про епіморфізми). *Нехай G, U – групи, f: G* $\longrightarrow$ *U – епіморфізм. Тоді група U ізоморфна до факторгрупи G/****Ker*** *f.*

Д о в е д е н н я. Покладемо для зручності H = **Ker** f. Розглянемо відношення еквівалентності $\Delta(f)$. Як і в попередній теоремі, можна показати, що рівність f(x) = f(y) рівносильна $f(xy^{-1}) = e$, інакше кажучи, $(x, y) \in \Delta(f)$ рівносильне $xy^{-1} \in H$, тобто $\Delta(f) = \Sigma_H$. З результатів розділу 4 отримуємо $[x]_{\Delta(f)} = xH$ для кожного x $\in$ G, а значить $G/\Delta(f) = G/H$.

Розглянемо тепер відображення ψ_f: G/H $\longrightarrow$ U, яке визначається за правилом ψ_f (xH) = f(x). З теореми 6.1 випливає, що ψ_f – бієкція, так що залишається довести тільки, що ψ_f – гомоморфізм. Маємо

$$\psi_f (xH\ yH) = \psi_f (xyH) = f(xy) = f(x)f(y) = \psi_f(xH)\ \psi_f (yH).$$

6.5. Теорема (теорема про гомоморфізми). *Нехай G, U – групи, f: G $\longrightarrow$ U – гомоморфізм. Тоді G/**Ker** f $\cong$ **Im** f $\leq$ U.*

Д о в е д е н н я. Дійсно, звуження f до відображення G $\longrightarrow$ **Im** f буде епіморфізмом, так що з **теореми 14.4** отримуємо ізоморфізм **Im** f $\cong$ G/**Ker** f. Нарешті з пропозиції 6.2 випливає, що **Im** f є підгрупою в U.

Розглянемо тепер деякі застосування отриманих результатів, і як перший крок отримаємо опис циклічних груп.

14.6. Теорема. *Нехай G = < g > – циклічна група.*

(1) Якщо група G – нескінченна, то вона ізоморфна адитивній групі Z цілих чисел.

(2) Якщо G – скінченна і $\left|G\right|$ = m, то G $\cong$ Z /m Z.

Д о в е д е н н я. Розглянемо відображення f: Z $\longrightarrow$ G, що визначається за правилом f(n) = g^n, n $\in$ Z. З властивостей цілих ступенів отримуємо

$$f(n + k) = g^{n+k} = g^n\ g^k = f(n)\ f(k),$$

так що f – гомоморфізм. Оскільки кожен елемент групи G є цілим ступенем елемента g, то f буде епіморфізмом.

Нехай група G є нескінченною, тоді з нерівності n $\neq$ k випливає, що f(n) = $g^n \neq g^k$ = f(k), тобто відображення f – ін'єктивне, а отже, буде ізоморфізмом.

Нехай тепер G – скінченна група. У цьому випадку маємо | g | = m, зокрема g^m = e, а значить m $\in$ Ker f. З пропозиції 2.2 отримуємо включення <m> = mZ $\leq$ Ker f. З теореми 2.14 отримуємо, що Ker f = <t> = tZ, а оскільки m $\in$ tZ, то m = ts. З теореми 6.4 отримуємо ізоморфізм G $\cong$ G/Ker f, зокрема $\left|\text{G/Ker f}\right|$ = m.

З іншого боку, теорема 4.10 дає рівності $|Z/mZ| = m$ та $|Z/tZ| = t$, які показують, що $t = m$. Таким чином, $G \cong Z/mZ$.

Розглянемо інші цікаві приклади.

6.7. Приклад. Нехай a – дійсне число, $a > 1$, Визначимо відображення $f: R \longrightarrow R^{\times}$ за таким правилом: $f(x) = a^x$ для кожного $x \in R$. Маємо тепер $f(x + y) = a^{x+y} = a^x a^y = f(x)f(y)$, так що f – гомоморфізм. Очевидно, що f – ін'єктивне та **Im** f складається з усіх додатних дійсних чисел, тобто f – ізоморфізм адитивної групи всіх дійсних чисел на мультиплікативну групу всіх строго додатних дійсних чисел.

6.8. Приклад. Визначимо відображення $f: \mathbf{GL_n}(R) \longrightarrow R^{\times}$ за таким правилом: $f(A) = \mathbf{det}(A)$ для кожної матриці $A \in \mathbf{GL_n}(R)$. За відомою теоремою випливають рівності

$$f(AB) = \mathbf{det}(AB) = \mathbf{det}(A)\ \mathbf{det}(B) = f(A)\ f(B),$$

які показують, що f – гомоморфізм. Далі

$$\mathbf{Ker}\ f = \{\ A \in \mathbf{GL_n}(R)\ |\ \mathbf{det}(A) = 1\ \} = \mathbf{SL_n}(R),$$

Зокрема $\mathbf{SL_n}(R)$ – нормальна підгрупа в $\mathbf{GL_n}(R)$. Очевидно, що відображення f – ін'єктивне, і з теореми 6.4 отримуємо

$$\mathbf{GL_n}(R)/\mathbf{SL_n}(R) \cong R^{\times}.$$

Аналогічно

$$\mathbf{GL_n}(Q)/\ \mathbf{SL_n}(Q) \cong Q^{\times} \quad \text{і} \quad \mathbf{GL_n}(Z)/\ \mathbf{SL_n}(Z) \cong \{1, -1\}.$$

6.9. Приклад. Визначимо відображення $f: \mathbf{T_n}(R) \longrightarrow \mathbf{D_n}(R)$ за таким правилом:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1\,n-1} & a_{1\,n} \\ 0 & a_{22} & a_{23} & \dots & a_{2\,n-1} & a_{2\,n} \\ 0 & 0 & a_{33} & \dots & a_{3\,n-1} & a_{3\,n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & a_{n\,n} \end{pmatrix} \xrightarrow{\ f\ } \begin{pmatrix} a_{11} & a_{12} & 0 & \dots & 0 & 0 \\ 0 & a_{22} & 0 & \dots & 0 & 0 \\ 0 & 0 & a_{33} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & a_{n\,n} \end{pmatrix}$$

З результатів розділу 3 випливає, що f – гомоморфізм (*f називають стираючим гомоморфізмом*). Очевидно, відображення f – сюр'єктивне і **Ker** f = $\mathbf{UT_n}(R)$, зокрема $\mathbf{UT_n}(R)$ – нормальна підгрупа в $\mathbf{T_n}(R)$. З теореми 6.4 отримуємо ізоморфізм $\mathbf{T_n}(R)/\mathbf{UT_n}(R) \cong \mathbf{D_n}(R)$. Доречно також зазначити, що $\mathbf{UT_n}(R)$ не буде нормальною підгрупою в усій групі $\mathbf{GL_n}(R)$.

6.10. Приклад. Нехай G – група. Гомоморфізм ϕ: G ⟶ G будемо називати *ендоморфізмом* групи G. Множину всіх ендоморфізмів групи G позначимо через **End**(G). Ендоморфізм ϕ групи G, який одночасно є підстановкою G, називається *автоморфізмом* групи G. Множину всіх автоморфізмів групи G позначимо через **Aut**(G). Зразу слід сказати, що **Aut**(G) – підгрупа групи підстановок **S**(G). Дійсно, нехай $\phi, \psi \in$ **Aut**(G), тоді

$$(\phi \circ \psi)(xy) = \phi(\psi(xy)) = \phi(\psi(x)\psi(y)) = \phi(\psi(x))\,\phi(\psi(y)) =$$
$$= (\phi \circ \psi)(x)\,(\phi \circ \psi)(y),$$

так що $\phi \circ \psi \in$ **Aut**(G). У розділі 1 вже показували, що відображення, обернене до ізоморфізму, само буде ізоморфізмом, зокрема, відображення, обернене до автоморфізму, буде авто морфізмом. Таким чином, **Aut**(G) задовольняє всі умови критерію підгрупи (теорема 2.4), отже, **Aut**(G) ≤ **S**(G).

У групі G візьмемо довільний елемент g і пов'яжемо з ним відображення $\mathbf{in}_g$: G ⟶ G, що визначається за правилом: $\mathbf{in}_g(x) = g^{-1}xg$

для кожного $x \in G$. Отримуємо:

$$\mathbf{in}_g(xy) = g^{-1}xyg = g^{-1}xgg^{-1}yg = \mathbf{in}_g(x)\ \mathbf{in}_g(y),$$

так що $\mathbf{in}_g(x)$ – ендоморфізм. Якщо x – довільний елемент групи G, то з рівностей $\mathbf{in}_g(gxg^{-1}) = g^{-1}(g\,x\,g^{-1})g = x$ випливає, що $\mathbf{in}_g(x)$ – сюр'єкція; якщо $\mathbf{in}_g(x) = g^{-1}xg = g^{-1}yg = \mathbf{in}_g(y)$, то помноживши справа на g^{-1}, а зліва на g, отримаємо x = y, так що $\mathbf{in}_g(x)$ – ін'єкція, тобто $\mathbf{in}_g(x)$ – автоморфізм. *Відображення $\boldsymbol{in}_g$ називають внутрішнім автоморфізмом, індукованим елементом g.*

Тепер розглянемо нове відображення Φ: G $\longrightarrow$ **Aut**(G), визначене за таким правилом: $\Phi(g) = \mathbf{in}_g$-1 для кожного елементу $g \in G$. З наступної допоміжної рівності

$$\mathbf{in}_{gh}(x) = (gh)^{-1}\, x\, (gh) = h^{-1}g^{-1}\, x\, gh = h^{-1}\, (g^{-1}\, x\, g)\, h =$$
$$\mathbf{in}_h(\mathbf{in}_g\ (x)) \ = \ (\mathbf{in}_h \ \textcircled{\scriptsize\smile}\ \mathbf{in}_g)(x)$$

випливає, що $\mathbf{in}_{gh} = \mathbf{in}_h \ \textcircled{\scriptsize\smile}\ \mathbf{in}_g$. Тепер маємо

$$\Phi(gh) = \mathbf{in}_{(gh)}\text{-}1 = \mathbf{in}_h\text{-}1_g\text{-}1 = \mathbf{in}_g\text{-}1 \ \textcircled{\scriptsize\smile}\ \mathbf{in}_h\text{-}1 = \Phi(g) \ \textcircled{\scriptsize\smile}\ \Phi(h),$$

і це показує, що Φ – гомоморфізм. Нехай $u \in$ **Ker** Φ, тоді $\Phi(u) = \mathbf{in}_u$-1= ε_G, отже

$$\mathbf{in}_u\text{-}1(x) = g\, x\, g^{-1} = \ \varepsilon_G(x) = x$$

для кожного $x \in G$. У свою чергу, рівність $g\, x\, g^{-1} = x$ еквівалентна $gx = xg$ для кожного $x \in G$, а це показує, що $g \in \zeta(G)$, тобто **Ker** $\Phi \le \zeta(G)$. Протилежне включення очевидне, тому отримуємо рівність **Ker** $\Phi = \zeta(G)$. **Im** Φ позначимо через **Inn**(G). За пропозицією 6.2 **Inn**(G) – підгрупа в **Aut**(G), *її називають групою внутрішніх автоморфізмів групи G*. З теореми 6.5 отримуємо ізоморфізм **Inn**(G) $\cong G/\zeta(G)$.

6.11. Приклад. Нехай G – група, H – її підгрупа, $\mathbf{R}_H = \{xH \mid x \in G\}$ – розбиття G на ліві суміжні класи за підгрупою H. Якщо $g \in G$, то визначимо відображення τ_g: $\mathbf{R}_H \longrightarrow \mathbf{R}_H$ за таким правилом: $\tau_g(xH) = (gx)H$, Очевидно воно визначено коректно.

Відображення τ_g: – ін'єктивне: $xH = g(g^{-1}xH) = \tau_g(g^{-1}xH)$.

Відображення τ_g: – ін'єктивне: якщо $\tau_g(x) = \tau_g(y)$, то $gxH = gyH$. Звідси $(gx)^{-1}(gy) \in H$, але $(gx)^{-1}(gy) = x^{-1}g^{-1}gy = x^{-1}y$, і це означає, що $xH = yH$. Таким чином, τ_g – підстановка множини $\mathbf{R}_H$. Розглянемо тепер відображення Ψ: G $\longrightarrow$ **S**($\mathbf{R}_H$), яке визначається за правилом: $\Psi(g) = \tau_g$ для кожного $g \in G$.

Маємо

$$\tau_{gv}\,(xH) = (gv)(xH) = g(vxH) = \tau_g\,(vxH) = \tau_g\,(\tau_v\,(xH)) = \tau \ \textcircled{\scriptsize\smile}\ \tau_v(xH),$$

тобто $\tau_{gv} = \tau_g \circledcirc \tau_v$. Тому$\Psi(gv) = \tau_{gv} = \tau_g \circledcirc \tau_v = \Psi(g) \circledcirc \Psi(v)$, так що Ψ – гомоморфізм. Нехай $K = \mathbf{Ker}\ \Psi$. Якщо $g \in \mathbf{Ker}\ \Psi$, $\Psi(g) = \tau_g = \varepsilon$, де ε – тотожна підстановка множини $\mathbf{R}_H$. Інакше кажучи, $\tau_g(xH) = gxH = xH$ для кожного $x \in G$. Зокрема, для $x = e$ отримуємо $gH = H$, тобто $g \in H$. Таким чином, $K \le H$.

Нехай тепер L – нормальна підгрупа G, яка міститься в підгрупі H. Якщо $g \in L$, то для довільного $x \in G$ маємо $x^{-1}gx = g_1 \in L$. Звідси $gx = xg_1$, а тому $\tau_g(xH) = gxH = xg_1H = xH$, тобто $L \le K$.

Відтак K – найбільша нормальна підгрупа, яку містить у собі H, тобто $K = \mathbf{Core}_G(H)$.

Якщо H = <e>, то і K = <e>, і за теоремою 6.3 Ψ – мономорфізм.

6.12. Теорема (теорема Келі). *Кожна група G ізоморфна підгрупі групи підстановок деякої множини.*

6.13. Наслідок. *Якщо група G – скінченна, $|G| = n$, то G ізоморфна підгрупі групи $\mathbf{S}_n$.*

6.14. Наслідок. *Нехай група G містить у собі підгрупу H скінченного індексу, $|G: H| = n$. Тоді H містить у собі підгрупу K, що нормальна в усій групі G, причому $|G/K| \le n!$.*

РОЗДІЛ 7
КІЛЬЦЯ, ПІДКІЛЬЦЯ, ЇХ ЕЛЕМЕНТАРНІ ВЛАСТИВОСТІ. ПРИКЛАДИ КІЛЕЦЬ

Уже розглядали множини, на яких задано одну бінарну алгебраїчну операцію. Проте в математиці часто доводиться зустрічатися з множинами, на яких задано кілька алгебраїчних операцій, причому ці операції ще певним чином пов'язані між собою. Як правило, зв'язок між операціями забезпечується законом дистрибутивності. Одним з перших прикладів є множини Z, Q, R множина дійсних функцій, а також множини матриць, $M_n(Z)$, $M_n(Q)$, $M_n(R)$. Історично, витоки теорії кілець та полів з'являються в теорії чисел та в теорії поліномів, основні поняття теорії полів з'являються в теорії Галуа. Сформувавшись, це поняття знайшло численні застосування у різних розділах алгебри й інших маремати чних дисциплінах.

Множина R, на якій задано дві бінарні алгебраїчні операції – додавання та множення, називається кільцем, якщо вона задовольняє такі умови:

(R 1) *відносно операції додавання R – абелева група.*

Це означає, що операція додавання *комутативна*, тобто

$$a + b = b + a \text{ для будь-яких } a, b \in R;$$

операція додавання *асоціативна*:

$$a + (b + c) = (a + b) + c \text{ для будь-яких } a, b, c \in R;$$

існує *нульовий елемент* O_R:

$$a + O_R = a \text{ для будь-якого } a \in R;$$

для кожного елемента $a \in R$ існує протилежний елемент $-a$:

$$a + (-a) = O_R.$$

(R 2) *операції додавання і множення зв'язані законами дистрибутивності:*

$$(a + b)c = ac + bc,\ a(b + c) = ab + ac \text{ для довільних } a, b, c \in R.$$

Розглядаючи R тільки з операцією додавання, будемо говорити про адитивну групу кільця R.

Наявність протилежного елемента дозволяє ввести в кільці операцію *віднімання* за правилом $\boldsymbol{a - b = a + (-b)}$.

Зазначимо деякі елементарні наслідки з аксіом кільця.

7.1. Пропозиція. *Нехай R – кільце. Мають місце такі рівності:*

$a \cdot O_R = O_R \cdot a = O_R$ *для всіх* $a \in R$.

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

$a(b - c) = ab - ac$; $(a - b)c = ac - bc$ *для довільних* $a, b, c \in R$.

Д о в е д е н н я. Дійсно, для кожного $b \in R$ маємо $b + O_R = b$. Скориставшись законами дистрибутивності, одержимо

$$ab = a(b + O_R) = ab + a \cdot O_R \quad \text{і} \quad ba = (b + O_R)a = ba + O_R \cdot a.$$

Оскільки для елемента ab кільця R існує протилежний елемент, то, додаючи його до обох частин цих рівностей, одержуємо $a \cdot O_R = O_R \cdot a = O_R$.

З означення протилежного елемента і законів дистрибутивності одержуємо

$$O_R = a \cdot O_R = a(b+(-b)) = ab + a(-b); \; O_R = O_R \cdot b = (a + (-a)) \cdot b = ab + (-a) \cdot b.$$

Ці рівності показують, що кожен з добутків $a(-b)$ і $(-a)b$ – це елемент, протилежний ab.

Доведені щойно рівності показують, що операції віднімання і множення теж зв'язані законами дистрибутивності:

$$a(b - c) = ab - ac, \; (a - b)c = ac - bc.$$

Кільце R називається асоціативним, якщо множення в R асоціативне:

(R 3) $a(bc) = (ab)c$ *для всіх* $a, b, c \in R$.

Кільце R називається комутативним, якщо множення в R комутативне:

(R 4) $ab = ba$ *для будь-яких* $a, b \in R$.

Будемо говорити також, що R – кільце з одиницею, якщо R має одиничний елемент e відносно операції множення:

(R 5) ae = ea = a для кожного $a \in R$.

Надалі будемо розглядати лише асоціативні кільця з одиницею. Треба сказати, що і теорія деяких типів неасоціативних кілець (у певному розумінні близьких до асоціативних) розвинена досить глибоко. Відзначимо найважливіші із цих типів кілець.

Кільце R називається лієвим кільцем (або кільцем Лі), якщо воно задовольняє такі умови:

(LR 1) $a^2 = O_R$ *для будь-якого* $a \in R$;

(LR 2) $(ab)c + (bc)a + (ca)b = O_R$ *для всіх* $a, b, c \in R$.

Умова (LR 2) називається *тотожністю Якобі.*

Кільце R називається йордановим кільцем, якщо воно комутативне і задовольняє таку умову:

(JR) $((a \cdot a)b)a = (a \cdot a)(b \cdot a)$ *для будь-яких елементів* $a, b \in R$.

Кільце R називається альтернативним, якщо воно задовольняє умови

(AR 1) $(aa)b = a(ab)$;

(AR 2) $(ba)a = b(aa)$

для довільних елементів $a, b \in R$.

Починаючи з цього моменту, говорячи "кільце", будемо розуміти під цим асоціативне кільце з одиницею.

Нехай R – кільце, в якому $O_R = e$. Тоді маємо $a = ae = aO_R = O_R$ для кожного $a \in R$. Таким чином, *кільце, в якому нульовий і одиничний елементи рівні, складається лише з одного нульового елемента.* Такі кільця будемо називати тривіальними. А надалі будемо розглядати тільки кільця, **у яких нульовий і одиничний елементи різні**.

З визначення кільця R випливає, що відносно операції множення воно буде напівгрупою з одиницею. Тому можемо розглянути групу U(R) всіх інвертованих елементів цієї напівгрупи. Далі *U(R) будемо називати групою інвертованих елементів кільця R.*

Відзначимо зразу ж, що $O_R \notin U(R)$.

Кільце R називається тілом, якщо кожен його ненульовий елемент інвертований, тобто $U(R) = R \setminus \{O_R\}$. *Комутативне тіло називається полем.*

Ненульовий елемент a кільця R називається лівим (відповідно правим) дільником нуля, якщо в R існує такий ненульовий елемент b, що

$ab = O_R$ *(відповідно* $ba = O_R$ *).*

Для комутативного кільця поняття правого і лівого дільника нуля збігаються.

Нехай $ab = O_R$ і припустимо, що $a \in U(R)$. Тоді

$$O_R = a^{-1} \cdot O_R = a^{-1}(ab) = (a^{-1}a)b = e \cdot b = b.$$

Звідси випливає, що інвертований елемент не може бути дільником нуля.

7.2. Пропозиція. *Нехай R – кільце,* $a \in R$. *Якщо елемент a не є лівим (відповідно правим) дільником нуля, то з рівності* $a[= ay$ *(відповідно* $xa = ya$*) завжди випливає* $x = y$.

Насправді, $ax = ay$ приводить до рівності $ax - ay = O_R$. Використовуючи закони дистрибутивності, одержуємо $a(x - y) = O_R$. Оскільки a не є лівим дільником нуля, то це означає, що $x - y = O_R$, тобто $x = y$.

7.3. Наслідок. *Нехай R – кільце, що не має дільників нуля, a – його ненульовий елемент. Тоді з рівності* $ax = ay$ *(відповідно* $xa = ya$*) завжди випливає* $x = y$.

Будемо говорити, що кільце R є областю цілісності, якщо воно комутативне і не має дільників нуля.

Нехай R – кільце, H – його підмножина. *Будемо говорити, що H – підкільце R, якщо H є сталою підмножиною R відносно обох операцій і H є кільцем відносно звужень цих операцій.*

7.4. Теорема. *Непуста підмножина H кільця R тоді і тільки тоді є підкільцем, коли вона задовольняє такі умови:*

(SR 1) якщо $x, y \in H$, *то* $x - y \in H$;

(SR 2) якщо $x, y \in H$, *то* $xy \in H$.

Д о в е д е н н я. Нехай H – підкільце в R. Зокрема H є сталою підмножиною відносно операції додавання і буде групою відносно звуження цієї операції. Інакше кажучи, H – підгрупа адитивної групи кільця R. Наслідок 2.5 показує, що умова (SR 1) буде необхідною. Умова (SR 2) показує, що H – стала підмножина відносно операції множення.

Навпаки, нехай підмножина задовольняє вказані умови. Наслідок 2.5 показує, що відносно операції додавання H буде підгрупою R. Зокрема H є сталою підмножиною відносно операції додавання і буде групою відносно звуження цієї операції. Умова (SR 2) показує, що звуження операції множення на H буде бінарною операцією на підмножині H. Закони дистрибутивності та асоціативності виконуються для всіх елементів з R, зокрема вони виконуються і для всіх елементів з H. Отже, H задовольняє всі умови (R 1) – (R 3), а тому буде підкільцем у R.

Слід зазначити, що, на відміну від підгруп, підкільце не обов'язково містить одиничний елемент усього кільця. *Підкільце H кільця R називається унітарним, якщо H містить одиничний елемент кільця R.*

Надалі той факт, що H – підкільце у кільці R буде записуватись так: $H \leq R$.

7.5. Наслідок. *Нехай R – кільце, S – система підкілець кільця R. Перетин* $\cap S$ *підкілець цієї системи також буде підкільцем R.*

Д о в е д е н н я. Нехай $S = \cap S$, $x, y \in S$, U – довільне підкільце із системи S, тоді $x - y$, $xy \in U$, і отже, $x - y$, xy належить перетину всіх таких підгруп, тобто S. Залишилось застосувати теорему 7.4.

Слід зазначити, що об'єднання підкілець не завжди буде підкільцем, оскільки і об'єднання підгруп не завжди буде підгрупою, як вже бачили це раніше.

7.6. Наслідок. *Нехай R – кільце, L – локальна система підкілець кільця R. Тоді об'єднання* $\cup$ *L підкілець цієї системи також буде підкільцем R.*

Д о в е д е н н я. Нехай V = $\cup$ L, x, y $\in$ V. Існують такі підкільця H, K з системи L, що x $\in$ H, y $\in$ K. Виберемо тепер підкільце L $\in$ L, яке містить у собі обидва підкільця H, K. Таким чином, x, y $\in$ L. Оскільки L – підкільце, x – y, xy $\in$ L за теоремою 7.4, і отже, x – y, xy $\in$ V. Залишилось знову застосувати теорему 7.4.

7.7. Наслідок. *Нехай R – кільце, L – лінійно впорядкована система підкілець кільця R. Тоді об'єднання* $\cup$ *L підкілець цієї системи також буде підкільцем R.*

7.8. Наслідок. *Нехай R – кільце,* $H_1 \leq H_2 \leq \ldots \leq H_n \leq \ldots$ *– зростаюча система підкілець кільця R. Тоді об'єднання* $\cup_{n \in N} H_n$ *також буде підкільцем R.*

Нехай M – підмножина кільця R, S – система всіх тих підкілець, які містять у собі M. Тоді перетин r(M) = $\cap$ S буде за наслідком 7.5 підкільцем кільця R. Цей перетин називають *підкільцем, породженим підмножиною M, а множина M називається системою породжуючих елементів для підкільця r(M). Зокрема, якщо r(M) = R, то будемо говорити, що M породжує все кільце R. Кільце R називається скінченно породженим, якщо воно породжується деякою своєю скінченною підмножиною.*

Якщо H – підкільце в R, що містить у собі підмножину M, то H містить у собі і підкільце r(M), у цьому сенсі r(M) – найменше з підкілець, які містять у собі підмножину M. Зрозуміло, що якщо M – підкільце кільця R, то r(M) = M. Разом з тим маємо таке корисне твердження.

7.9. Пропозиція. *Нехай R – кільце, H – його підкільце. Якщо* $M \subseteq H$, *то* $\mathbf{r}(M) \leq H$.

Нехай T – тіло, H – його підмножина. *Будемо говорити, що H – підтіло T, якщо H – унітарне підкільце, яке є тілом відносно операцій додавання і множення, що задані в T.*

7.10. Теорема. *Нехай T – тіло. Підмножина H тіла T тоді і тільки тоді є підтілом T, коли H задовольняє такі умови:*

(SF 1) якщо $x, y \in H$, *то* $x - y, xy \in H$;

(SF 2) $e \in H$;

(SF 3) якщо $x \in H, x \neq 0_T$, *то* $x^{-1} \in H$.

Твердження теореми майже очевидне. Наступні наслідки аналогічні відповідним твердженням для підкілець, тому їх доведення опускаємо.

7.11. Наслідок. *НехайT – тіло. Підмножина H тіла T тоді і тільки тоді буде його підтілом, коли H задовольняє такі умови:*

(SF 4) H містить ненульові елементи;

(SF 5) якщо $x, y \in H$, *то* $x - y, xy \in H$;

(SF 6) якщо $x, y \in H, y \neq 0_T$, *то* $xy^{-1} \in H$.

7.12. Наслідок. *Нехай T – тіло, S – система підтіл T. Перетин* $\cap$ *S підтіл, що належать до цієї системи, також буде підтілом T.*

Як і для підкілець, об'єднання підтіл не завжди буде підтілом.

7.13. Наслідок. *Нехай T – тіло, L – локальна система підтіл T. Тоді об'єднання* $\cup$ *L підтіл, що належать до цієї системи, також буде підтілом T.*

7.14. Наслідок. *Нехай T – тіло, L – лінійно впорядкована система підтіл T. Тоді об'єднання* $\cup$ *L підтіл, що належать до цієї системи, також буде підтілом T.*

7.15. Наслідок. *Нехай T – тіло,* $H_1 \leq H_2 \leq \ldots \leq H_n \leq \ldots$ *– зростаюча система підтіл T. Тоді об'єднання* $\cup_{n \in N} H_n$ *також буде підтілом T.*

Із наслідку 7.12 випливає, що перетин T_0 усіх тіл підтіла T є підтілом, яке вже не містить власних підтіл.

Тіло T називається простим, якщо воно не містить у собі власних (тобто відмінних від T) підтіл.

Нехай T – тіло, H – його підтіло. Якщо H комутативне, то будемо говорити, що H – підполе T.

Як і для груп розглянемо корисну конструкцію декартового (прямого) добутку скінченної множини кілець. Нехай $R_1, \ldots, R_n$ – кільця, $D = R_1 \times \ldots \times R_n$ – їх декартовий добуток як множин. На множині D визначимо операції додавання та множення за правилами

$$(y_1, \ldots, y_n) + (x_1, \ldots, x_n) = (y_1 + x_1, \ldots, y_n + x_n), \text{ де } y_j, x_j \in R_j, 1 \le j \le n,$$
$$(y_1, \ldots, y_n)(x_1, \ldots, x_n) = (y_1x_1, \ldots, y_nx_n), \text{ де } y_j, x_j \in R_j, 1 \le j \le n.$$

Оскільки y_j, x_j – елементи одного і того ж кільця R_j, то їх сума та добуток визначені у цьому кільці R_j, $1 \le j \le n$, так що отримали дійсно операції на D. Застосовуючи ті ж самі аргументи, що використовувались при розгляді декартових добутків груп, можна довести, що операції додавання та множення асоціативні, додавання комутативне, має нульовий елемент, яким буде набір $(0_1, \ldots, 0_n)$, де 0_j – нульовий елемент кільця R_j, $1 \le j \le n$, і також

$$-(y_1, \ldots, y_n) = (-y_1, \ldots, -y_n).$$

Якщо e_j – одиничний елемент кільця R_j, $1 \le j \le n$, то набір $(e_1, \ldots, e_n)$ буде одиничним елементом для D. Нарешті покажемо, що додавання та множення пов'язані законами дистрибутивності. Дійсно

$$((y_1, \ldots, y_n) + (x_1, \ldots, x_n))(z_1, \ldots, z_n) =$$
$$(y_1 + x_1, \ldots, y_n + x_n)(z_1, \ldots, z_n) =$$
$$((y_1 + x_1)z_1, \ldots, (y_n + x_n)z_n) = (y_1z_1 + x_1z_1, \ldots, y_nz_n + x_nz_n) =$$
$$(y_1z_1, \ldots, y_nz_n) + (x_1z_1, \ldots, x_nz_n) =$$
$$(y_1, \ldots, y_n)(z_1, \ldots, z_n) + (x_1, \ldots, x_n)(z_1, \ldots, z_n).$$

Інший закон дистрибутивності доводиться аналогічно. Таким чином, усі аксіоми кільця для D виконані. Тому далі, коли йтиметься про декартовий добуток кілець, будемо розуміти під цим тільки що побудоване кільце. Це кільце називається також

прямою сумою кілець $R_1, \ldots, R_n$ (як і для груп, поняття прямої суми та декартового добутку кілець у загальному випадку відрізняються один від одного, але для випадку скінченої множини кілець вони співпадають). Якщо всі кільця $R_1, \ldots, R_n$ – комутативні, то і їх декартовий добуток буде комутативним кільцем.

Розглянемо тепер будову групи U(D). Нехай $(y_1, \ldots, y_n) \in U(D)$. Тоді існує такий набір $(z_1, \ldots, z_n)$, що

$$(y_1, \ldots, y_n)(z_1, \ldots, z_n) = (y_1z_1, \ldots, y_nz_n) = (e_1, \ldots, e_n).$$

Інакше кажучи, $y_j z_j = e_j$, тобто $y_j \in U(R_j)$ для кожного j, $1 \le j \le n$. Нехай тепер $y_j \in U(R_j)$ для кожного j, $1 \le j \le n$. Повторюючи попередні аргументи у зворотньому порядку, покажемо, що $(y_1, \ldots, y_n) \in U(D)$. Таким чином, отримали рівність

$$U(R_1 \times \ldots \times R_n) = U(R_1) \times \ldots \times U(R_n).$$

Розглянемо тепер деякі важливі приклади кілець і підкілець.

Числові кільця. Множина R усіх дійсних чисел є комутативним кільцем. Його унітарні підкільця – множина Q усіх раціональних чисел і множина Z усіх цілих чисел. Більше того, Q і R є полями. Z дає нам також приклад скінченно породженого кільця, воно породжується числом 1. У розділі 3 вже відмічали, що для кожного $n \ge 0$ підмножина $nZ = \{ nk \mid k \in Z\}$ буде підгрупою Z, зокрема вона задовольняє умову (SR 1). Далі, (nk)(nt) = n(nkt), тобто nZ задовольняє умову (SR 2) і отже, за теоремою 7.4 nZ, буде підкільцем Z. З теореми 2.13 випливає, що довільна підгрупа Z співпадає з однією з підмножин nZ, $n \ge 0$, і інших підгруп Z не має. Це означає, що множина всіх підкілець Z співпадає з $\{n Z \mid n \ge 0\}$.

Нехай p – просте число, розглянемо адитивну групу Q_p p-ічних дробів. Маємо $(m/(p^k))(r/(p^s)) = (mr)/(p^{k+s})$, а це показує, що Q_p задовольняє умову (SR 2), отже, за теоремою 15.4 Q_p, буде підкільцем Q, більше того, унітарним підкільцем. Його називають кільцем p-ічних дробів.

Матричні кільця. Нехай R – цілісності. Через $M_n(R)$ позначимо сукупність усіх квадратних матриць порядку n, коефіцієнти яких належать кільцю R. Якщо $A = \| a_{ij} \|_{1 \le i, j \le n}$, $B = \| b_{ij} \|_{1 \le i, j \le n}$ – дві матриці з множини $M_n(R)$, то їх *сумою A + B будемо називати матрицю* $\| c_{ij} \|_{1 \le i, j \le n}$, *коефіцієнти якої визначаються за правилом*

$$c_{ij} = a_{ij} + b_{ij} \text{ для кожної пари індексів } i, j,\ 1 \le i, j \le n,$$

а добутком AB цих матриць будемо називати матрицю $\| d_{ij} \|_{1 \le i\ j \le n}$, *коефіцієнти якої визначаються за правилом*

$$d_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{in} b_{nj} = \sum_{1 \le k \le n} a_{ik} b_{kj}$$

для кожної пари індексів $i, j,\ 1 \le i, j \le n.$

Інакше кажучи, ми зберігаємо ті визначення суми та добутку, які вже були введені для числових матриць. Легко показати, що $M_n(R)$ буде кільцем відносно цих операцій. Слід зазначити, що, як і для числових матриць, це кільце не комутативне.

Як і для числових матриць, можемо ввести визначення визначника матриці $A = \| a_{ij} \|_{1 \le i, j \le n} \in M_n(R)$ за правилом

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sign} \pi\ a_{1\,\pi(1)}\, a_{2\,\pi(2)} \ \ldots\ a_{n\,\pi(n)}.$$

Неважко впевнитися в тому, що всі властивості визначників залишаються справедливими для області цілісності R. Нехай $A \in U(M_n(R))$, $a = \det(A)$. Оскільки

$$e = \det(E) = \det(AA^{-1}) = \det(A)\ \det(A^{-1}) = a\ \det(A^{-1}),$$

то $a \in U(R)$. Навпаки, нехай A – така матриця, що $\det(A) \in U(R)$. Розглянемо тепер матрицю $B = \| b_{ij} \|_{1 \le i\ j \le n} \in M_n(R)$, коефіцієнти якої визначаються за правилом $b_{i\,j} = A_{j\,i}/\det(A)$, $1 \le i,\ j \le n$. Неважко пересвідчитися в тому, що $AB = BA = E$, тобто матриця B буде оберненою для матриці A. Отже:

$$U(M_n(R)) = \{\ A \in M_n(R) \mid \det(A) \in U(R)\ \}.$$

Як і для числових матриць, цю групу називають *загальною лінійною групою ступеня n над кільцем R*. Зокрема, якщо R – поле, то $U(M_n(R))$ – це знову множина всіх невироджених матриць.

Знову позначимо через E_{km} матрицю, у якої на (k, m) – му місці стоїть одиничний елемент, а всі інші місця зайняті нулями. Як і для числових матриць, неважко довести таку рівність:

$$E_{km}\, E_{r\,s} = \begin{cases} E_{k\,s}, \text{ якщо } m = r, \\ O, \text{ якщо } m \neq r, \end{cases}$$

зокрема, $E_{1\,2}\, E_{2\,3} = E_{1\,3}$, $E_{2\,3}\, E_{1\,2} = O$.

Звідси випливає, що кільце $M_n(R)$ – некомутативне (якщо $n \geq 2$) і містить дільники нуля.

Позначимо тепер через $T_n°(R)$ підмножину $M_n(R)$, що складається з усіх трикутних (точніше верхньо-трикутних) матриць. Очевидно, різниця двох трикутних матриць є трикутною матрицею. Як і є в розділі 3, можна показати, що добуток двох трикутних матриць є матрицею трикутною. Із теореми 7.4 одержуємо, що $T_n°(R)$ – підкільце $M_n(R)$.

Трикутну матрицю будемо називати нультрикутною, якщо всі елементи її головної діагоналі нульові.

Множину всіх нультрикутних матриць порядку n з коефіцієнтами з кільця R позначимо через NT_n(R. Можна легко пересвідчитися, що підмножина $NT_n(R)$ задовольняє умови (SR 1) і (SR 2), так що з теореми 7.4 одержуємо, що $NT_n(R)$ – підкільце $M_n(R)$.

Позначимо через $D_n°(R)$ множину всіх діагональних матриць порядку n, коефіцієнти яких належать до кільця R. Далі покладемо

$$RE = \{\lambda E \mid \lambda \in R\},$$

тобто RE – підмножина всіх скалярних матриць. З теореми 7.4 знову одержуємо, що $D_n°(R)$ і RE – підкільця $M_n(R)$. Із самого означення випливає, що $T_n°(R)$, $D_n°(R)$ і RE – унітарні підкільця $M_n(R)$.

Нарешті, нехай

$$RE_{ii} = \{ \lambda E_{ii} \mid \lambda \in R \}.$$

Маємо $\lambda E_{ii} - \mu E_{ii} = (\lambda - \mu)E_{ii}$, $\lambda E_{ii}\, \mu E_{ii} = \lambda\mu E_{ii}$. Із теореми 7.4 випливає, що RE_{ii} – підкільце $M_n(R)$. Далі, $\lambda E_{ii}\, eE_{ii} = eE_{ii}\, \lambda E_{ii} = \lambda E_{ii}$.

Ця рівність показує, що $eE_{ii} = E_{ii}$ – одиничний елемент підкільця RE_{ii}. Таким чином, підкільце RE_{ii} хоча і не є унітарним , але має свій одиничний елемент.

Функціональні кільця. Нехай R – довільне кільце, M – довільна множина. У множині R^M усіх функцій, визначених на M, що набувають значень у кільці R, визначимо суму і добуток за правилом:

$$(f + g)(a) = f(a) + g(a),\ (f\, g)(a) = f(a)\, g(a)$$

для будь-якого елемента $a \in M$.

Множина R^M є абелевою групою відносно операції додавання: якщо $f, g \in R^M$, то

$$(f + g)(a) = f(a) + g(a) = g(a) + f(a) = (g + f)(a)$$

для кожного елемента $a \in M$. Звідси випливає, що $f + g = g + f$.

Аналогічно, для $f, g, h \in R^M$ маємо

$$(f + (g + h))(a) = f(a) + (g + h)(a) = f(a) + (g(a) + h(a)) =$$
$$= (f(a) + g(a)) + h(a) = (f + g)(a) + h(a) = ((f + g) + h)(a)$$

для довільного елемента $a \in M$, так що $f + (g + h) = (f + g) + h$.

Задамо відображення θ правилом $\theta(a) = O_R$ для кожного $a \in M$. Тоді

$$(f + \theta)(a) = f(a) + \theta(a) = f(a) + O_R = f(a),\ \text{тобто}\ f + \theta = f,$$

для будь-якого $f \in R^M$.

Це означає, що θ – нульовий елемент R^M. Нарешті, відображення $-f$ визначимо за правилом $(-f)(a) = -f(a)$. Очевидно,

$f + (-f) = \theta$. Отже, всі аксіоми абелевої групи виконані.

Нехай $f, g, h \in R^M$, $a \in M$, тоді

$$(f(g + h))(a) = f(a)((g + h))(a) = f(a)(g(a) + h(a)) = f(a)g(a) + f(a)h(a) = \\ = (fg)(a) + (fh)(a) = (fg + fh)(a),$$

тобто $f(g + h) = fg + fh$.

Аналогічно доводиться, що і $(f + g)h = fh + gh$.

Множення функцій асоціативне:

$$(f(gh))(a) = f(a)(gh)(a) = f(a)(g(a)h(a)) = (f(a)g(a))h(a) = (fg)(a)h(a) = \\ ((fg)h)(a).$$

Одиничним елементом буде функція I, яка визначається за правилом $I(a) = e$ для кожного елемента $a \in M$: $(fI)(a) = f(a)I(a) = f(a)e = f(a)$, тобто $fI = f$. І аналогічно $If = f$.

Таким чином, відносно введених вище операцій додавання та множення R^M буде кільцем. Якщо R комутативне, то і кільце R^M комутативне:

$$(fg)(a) = f(a)g(a) = g(a)f(a) = (gf)(a)$$

Розглянемо групу $U(R^M)$. Зрозуміло, що вона складається з усіх таких функцій f, що $f(a) \in U(R)$ для всіх $a \in M$.

Кільце R^M має дільники нуля. Наприклад, нехай L – непуста підмножина M. Задамо функції f, g таким чином:

$$f(a) = \begin{cases} e, \text{ якщо } a \in L, \\ O_R, \text{ якщо } a \notin L, \end{cases} \qquad g(a) = \begin{cases} O_R, \text{ якщо } a \in L \\ e, \text{ якщо } a \notin L. \end{cases}$$

Функції f, g – ненульові, проте $f(a)g(a) = O_R$ для будь-якого $a \in M$.

В аналізі найчастіше зустрічається приклад, коли $M = [a, b]$ – відрізок прямої, або $M = R$, а $R = R$ – поле дійсних чисел. Інакше кажучи, $R^M = R^{[a, b]}$ (відповідно $R^M = R^R$) – кільце всіх функцій, заданих на [a, b] (відповідно на R) із значеннями в полі дійсних чисел. Тут виникає досить багато підкілець: підкільце всіх неперервних функцій, підкільце всіх диференційованих функцій, підкільце всіх двічі диференційованих функцій тощо.

Кільце ендоморфізмів абелевої групи. Нехай A – абелева група, операція в якій записується адитивно. У розділі 6 уже розглядали множину End(A) усіх ендоморфізмів групи A. Задамо додавання ендоморфізмів таким правилом: якщо f, g ∈ End(A), то покладемо

$$(f + g)(a) = f(a) + g(a) \text{ для будь-якого елемента } a \in A.$$

Маємо

$$(f + g)(a + b) = f(a + b) + g(a + b) = f(a) + f(b) + g(a) + g(b) = f(a) + g(a) + f(b) + g(b) = (f + g)(a) + (f + g)(b).$$

Це означає, що f + g – ендомофізм A. Таким чином, відображення

$$(f, g) \longrightarrow f + g,\ f, g \in \mathrm{End}(A)$$

є бінарною операцією на множині End(A).

Множина End(A) є абелевою групою відносно цієї операції:

$$(f + g)(a) = f(a) + g(a) = g(a) + f(a) = (g + f)(a);$$

$$(f + (g + h))(a) = f(a) + (g + h)(a) = f(a) + (g(a) + h(a));$$

$$((f + g) + h)(a) = (f(a) + g(a)) + h(a) = (f(a) + g(a)) + h(a)$$

для будь-якого елемента a ∈ A; звідси одержуємо, що

$$f + g = g + f,\ f + (g + h) = (f + g) + h$$

для довільних f, g, h ∈ End(A).

Задамо відображення θ правилом $\theta(a) = 0_A$ для кожного a ∈ A.

Тоді (f + θ)(a) = f(a) + θ(a), тобто f + θ = f для будь-якого f ∈ End(A). Це означає, що θ – нульовий елемент End(A). Нарешті, покладемо (-f)(a) = - f(a). Очевидно, f + (-f) = θ. Отже, всі аксіоми абелевої групи виконані.

Нехай f, g ∈ End(A), a, b ∈ A. Маємо

$$(f \text{ ☺ } g)(a + b) = f(g(a + b)) = f(g(a) + g(b)) = f(g(a)) + f(g(b)) = (f \text{ ☺ } g)(a) + (f \text{ ☺ } g)(b).$$

Ця рівність показує, що $f \circledcirc g$ – ендоморфізм A. Таким чином, відображення

$$(f, g) \longrightarrow f \circledcirc g,\ f, g \in \mathrm{End}(A),$$

є бінарною операцією на множині End(A).

Нехай $f, g, h \in \mathrm{End}(A)$, $a \in A$.

$$(f \circledcirc (g + h))(a) = f(g + h)(a)) = f(g(a) + h(a)) = f(g(a)) + f(h(a)) = (f \circledcirc g)(a) + (f \circledcirc h)(a) = (f \circledcirc g + f \circledcirc h)(a).$$

Звідси випливає, що

$$f \circledcirc (g + h) = f \circledcirc g + f \circledcirc h.$$

Аналогічно доводиться другий закон дистрибутивності. Як вже знаємо, множення відображень асоціативне, а підстановка ε_A, що задається правилом $\varepsilon_A(a) = a$ для кожного $a \in A$, є одиничним елементом відносно операції множення. Таким чином, **End(**A) – кільце. Його називають *кільцем ендоморфізмів абелевої групи A.*

Булеві кільця. Нехай M – довільна множина, A = B (M) – її булеан. Якщо $a, b \in A$, то покладемо

$$a + b = (a \cup b) \setminus (a \cap b),\ a\,b = a \cap b.$$

Неважко перевірити, що **B** (M) – кільце. Більше того, у цьому кільці для кожного елемента a маємо $a^2 = a$, $2a = \varnothing = 0_A$. Кільця з такими властивостями називають *булевими***.**

Центр кільця. Якщо R – кільце, то через $\zeta(R)$ позначимо множину всіх елементів R, що переставні з кожним елементом R. Нехай $a, b \in \zeta(R)$, $x \in R$. Маємо

$$(a - b)x = ax - bx = xa - xb = x(a - b),$$

$$(ab)x = a(bx) = a(xb) = (ax)b = (xa)b = x(ab).$$

Ці рівності показують, що $a - b,\ ab \in \zeta(R)$. З теореми 7.4 випливає, що $\zeta(R)$ – підкільце R. Підкільце $\zeta(R)$ *називається центром кільця R.*

Нехай $Ze = \{n\,e \mid n \in Z\}$. Знову

$$ne - ke = (n - k)e,\ (ne)\,(ke) = (nk)e.$$

З теореми 7.4 знову випливає, що Ze – підкільце R, а оскільки

$$(ne)a = n(ea) = na = a(ne),$$

то $ne \in \zeta(R)$. Таким чином, $Ze \leq \zeta(R)$.

Нільпотентні елементи кільця. Нехай R – кільце. Елемент a називають нільпотентним, якщо $a^n = 0_R$ для деякого $n \in N$. Якщо R – комутативне кільце, то сукупність його нільпотентних елементів буде підкільцем. Справді, знову застосовуємо теорему 7.4. Нехай $a^n = 0_R$, $b^t = 0_R$ для деяких $n, t \in N$. Розглянемо $(a - b)^{n+t}$. Цей елемент є лінійною комбінацією елементів $a^k b^m$ з деякими цілими коефіцієнтами, причому $k + m = n + t$. Тому якщо $k \geq n$, то $a^k b^m = 0_R$. Якщо ж $k < n$, то $m > t$ і знову $a^k b^m = 0_R$. Таким чином, $(a - b)^{n+t} = 0_R$.

Далі,

$$(a\,b)^{n+t} = a^{n+t} b^{n+t}.(a - b)^{n+t} = 0_R.$$

РОЗДІЛ 8
ІДЕАЛИ КІЛЕЦЬ. ФАКТОР-КІЛЬЦЕ

Як уже зазначалося, одним з основних понять теорії груп є поняття нормальної підгрупи. Нормальні підгрупи дозволяють отримувати з даної групи нові – її факторгрупи, а також тісно пов'язані з іншим важливим поняттям – поняттям гомоморфізмів. Аналог нормальних підгруп виникає і в кільцях, цим аналогом є ідеали кілець. Поняття ідеалу виникло із задач, що зовсім не є аналогічними тим, з яких виникло поняття нормальної підгрупи. Це також і обумовило і різницю в термінах. Поняття ідеалу виникло при вивченні числових кілець, у яких розклад на прості елементи не є однозначним. Намагання якимось чином повернути однозначність хоча б у більш слабкій формі і привело до виникнення того поняття, яке зараз називається ідеалом. І тільки пізніше почала виявлятись аналогія між поняттями нормальної підгрупи групи та ідеалу кільця.

Підкільце H кільця R називається ідеалом (в R), якщо для кожного елементу $x \in R$ та кожного елементу $h \in H$ добутки xh та hx містяться в H.

Це поняття, як правило, поділяється на два.

Підкільце H кільця R називається лівим (відповідно правим) ідеалом (в R), якщо для кожного елементу $x \in R$ та кожного елементу $h \in H$ добуток xh відповідно hx) міститься в H.

Враховуючи критерій підкільця (теорема 7.4), отримуємо наступний критерій ідеалу.

8.1. Пропозиція. *Непуста підмножина H кільця R тоді і тільки тоді є ідеалом, коли вона задовольняє такі умови:*

(І 1) якщо $x, y \in H$, то $x - y \in H$;

(І 2) якщо $x \in R$, $h \in H$, то добутки xh та hx містяться в H.

Кожна підмножина, що є одночасно правим і лівим ідеалом, буде ідеалом кільця. Тому говорять ще про "двосторонній ідеал" кільця. Якщо ж R – комутативне кільце, то всі ці поняття збігаються.

У будь-якому кільці R підмножини $< 0_R >$ та R є ідеалами.

Кільце R називається простим, якщо його єдиними ідеалами є лише $< 0_R >$ та R.

Наступні наслідки аналогічні відповідним твердженням для підкілець, тому їх доведення опускаємо.

8.2. Наслідок. *Нехай R – кільце, S – система ідеалів кільця R. Перетин $\cap$ S ідеалів, що належать до цієї системи також буде ідеалом R.*

8.3. Наслідок. *Нехай R – кільце, L – локальна система ідеалів кільця R. Тоді об'єднання $\cup$ L ідеалів, що належать до цієї системи, також буде ідеалом R.*

8.4. Наслідок. *Нехай R – кільце, L – лінійно впорядкована система ідеалів кільця R. Тоді об'єднання $\cup$ L ідеалів, що належать до цієї системи, також буде ідеалом R.*

8.5. Наслідок. *Нехай R – кільце,*

$$H_1 \leq H_2 \leq \ldots \leq H_n \leq \ldots -$$

зростаюча система ідеалів кільця R. Тоді об'єднання $\cup_{n \in N} H_n$ *також буде ідеалом R.*

Нехай M – підмножина кільця R, S – система всіх тих ідеалів, які містять у собі M. Тоді перетин id(M) = $\cap$ S буде за наслідком 8.2 ідеалом кільця R. Цей перетин називають *ідеалом, породженим підмножиною M, а множина M називається системою породжуючих елементів для ідеала id(M).* Якщо M = {a}, то, скорочуючи, замість id({ a }) будемо писати id(a) і говорити про ідеал, породжений елементом a. Цей ідеал частіше називають *головним ідеалом* кільця R.

Розглянемо випадок, коли кільце R – комутативне. Якщо x $\in$ R, то ax $\in$ id(a). Покладемо aR = { ax $\mid$ x $\in$ R}. Наша мета зараз – довести рівність aR = id(a). Уже доведено включення aR $\leq$ id(a). З

іншого боку, покажемо, що aR – ідеал кільця R, тобто aR задовольняє умови (I 1) (I 2). Дійсно, якщо $u, v \in aR$, то $u = ay$, $v = az$ для деяких елементів $y, z \in R$. Маємо

$$u - v = ay - az = a(y - z) \in aR,$$
$$gu = ug = (ay)g = a(yg) \in aR$$

для довільного елементу g кільця R. Отже, aR – ідеал в R. Оскільки $a = ae$, то $a \in aR$, а тому і $id(a) \leq aR$. Це останнє включення доводить рівність $aR = id(a)$.

Раніше вже відзначалося, що будь-яке підкільце кільця Z має вигляд nZ, де $n \geq 0$. Таким чином, кожне підкільце Z є ідеалом, причому головним ідеалом.

У розділі 5 визначили добуток підгруп у довільній групі. За аналогією можна визначити суму підкілець кільця. Нехай R – кільце, H, K – його підкільця.

Покладемо $H + K = \{ x + y \mid x \in H, y \in K \}$. *Підмножину* $H + K$ *будемо називати сумою підкілець H, K.*

Із самого визначення випливає включення $H, K \subseteq H + K$. Як і в ситуації груп, сума підкілець не завжди буде підкільцем. Але якщо хоча б одне з них, наприклад H, є ідеалом, то $H + K$ – підкільце. Дійсно, нехай $a, b \in H + K$, тоді $a = x + y$, $b = x_1 + y_1$ для деяких елементів $x, x_1 \in H$, $y, y_1 \in K$. Маємо

$$a - b = (x + y) - (x_1 + y_1) = (x - x_1) + (y - y_1) \in H + K,$$

$$ab = (x + y)(x_1 + y_1) = xx_1 + xy_1 + yx_1 + yy_1 .$$

Оскільки H – ідеал, то $xx_1 + xy_1 + yx_1 \in H$, а тому

$$(xx_1 + xy_1 + yx_1) + yy_1 \in H + K.$$

Отже, $H + K$ задовольняє умови (SR 1) та (SR 2), і з теореми 7.4 випливає, що $H + K$ – підкільце R. Більше того, якщо H, K – ідеали R, то і їх сума $H + K$ буде ідеалом: якщо $x \in H$, $y \in K$, $z \in R$, то

$$z(x + y) = zx + zy \in H + K,$$

$$(x + y)z = xz + yz \in H + K,$$

так що H + K задовольняє умови (I 1), (I 2) і за пропозицією 8.1 буде ідеалом кільця R. Зауважимо також, що якщо ідеали H, K містяться у більшому ідеалі L, то і $H + K \le L$..

8.6. Пропозиція. *Нехай R – кільце, H – його ідеал. Якщо H містить інвертований елемент кільця R, то H = R.*

Дійсно, нехай $x \in H \cap U(R)$. Оскільки H – ідеал, то $e = x^{-1} x \in H$. Якщо тепер a – довільний елемент R, то $a = a\, e \in H$. Це і означає, що H = R.

8.7. Наслідок. *Будь-яке тіло є простим кільцем.*

Однак наступна теорема показує, що обернене твердження не є правильним.

8.8. Теорема. *Кільце матриць $\mathbf{M}_n(F)$ над полем F є простим.*

Д о в е д е н н я. Нехай L – ненульовий ідеал кільця $\mathbf{M}_n(F)$, $B = \| b_{i\,j} \|_{1 \le i\ j \le n}$ – ненульова матриця, що належить до L. Існує така пара індексів k, m, що $b_{k\,m} \ne \mathbf{O}_F$. Розглянемо послідовно такі добутки:

$$E_{s\,t}\, B = E_{s\,t} \Big(\sum_{1 \le i\ j \le n} b_{i\,j}\, E_{i\,j} \Big) = \sum_{1 \le i\ j \le n} b_{i\,j}\, E_{s\,t}\, E_{i\,j} =$$

$$= \sum_{1 \le j \le n} b_{t\,j}\, E_{s\,t}\, E_{i\,j} = \sum_{1 \le j \le n} b_{t\,j}\, E_{s\,j};$$

$$E_{s\,t}\, B\, E_{q\,r} = \Big(\sum_{1 \le j \le n} b_{t\,j}\, E_{s\,j} \Big) E_{q\,r} = \sum_{1 \le j \le n} b_{t\,j}\, E_{s\,j}\, E_{q\,r} = b_{t\,q}\, E_{s\,r}.$$

Оскільки L – ідеал $\mathbf{M}_n(F)$, то $E_{s\,t}\, B\, E_{q\,r} \in L$ для довільних індексів s, t, q, r. Нехай тепер $A = \| a_{i\,j} \|_{1 \le i\ j \le n}$ – довільна матриця. З доведеного вище отримуємо, що

$$a_{i\,j}\, E_{i\,j} = \big(a_{i\,j}\, a_{ik\,m}^{-1} E \big) \big(E_{i\,k}\, B\, E_{m\,j} \big) \in L$$

для довільної пари індексів i, j, а це показує, що $A = \sum_{1 \le i\ j \le n} a_{ij}\, E_{ij} \big) \in L$, отже, $L = \mathbf{M}_n(F)$.

Як видно з попереднього розділу, $\mathbf{M}_n(F)$ містить дільники нуля, а тому не є полем. Але у випадку комутативних кілець твердження, обернене до наслідку 8.7, є правильним:

8.9. Теорема. *Якщо R – комутативне просте кільце, то R є полем.*

Д о в е д е н н я. Дійсно, нехай a – довільний ненульовий елемент кільця R. Оскільки $a \in aR$, то aR – ненульовий ідеал кільця K, а тому $aR = R$. Зокрема, існує такий елемент x, що $ax = e$. Отже, довільний ненульовий елемент a має обернений, і отже, кільце R є полем.

Нехай R – кільце, H – його ідеал. Оскільки H – підгрупа адитивної групи кільця R, то множина всіх суміжних класів $\{x + H \mid x \in R\}$ є абелевою групою відносно операції додавання, яка визначається за правилом

$$(x + H) + (y + H) = x + y + H.$$

Визначимо множення суміжних класів за таким правилом

$$(x + H)(y + H) = xy + H.$$

Це означення коректне: насправді, нехай x_1, y_1 – такі елементи кільця R, що $x + H = x_1 + H$, $y + H = y_1 + H$, тоді $x_1 = x + u$, $y_1 = y + v$ для деяких елементів $u, v \in H$ і

$$x_1 y_1 = (x + u)(y + v) = xy + uy + xv + uv.$$

Оскільки H – ідеал, то $uy, xv, uv \in H$, тому $xy + H = x_1 y_1 + H$.

Далі

$$(x + H)(y + H + z + H) = (x + H)(y + z + H) = x(y + z) + H = xy + xz + H =$$

$$xy + H + xz + H = (x + H)(y + H) + (x + H)(z + H);$$

і аналогічно

$$(x + H + y + H)(z + H) = (x + H)(z + H) + (y + H)(z + H);$$

$$(x + H)((y + H)(z + H)) = (x + H)(yz + H) = x(yz) + H =$$
$$(xy)z + H = (xy + H)(z + H) = ((x + H)(y + H))(z + H);$$

$$(e + H)(x + H) = ex + H = x + H = xe + H = (x + H)(e + H),$$

так що відносно цих операцій додавання і множення множина суміжних класів є кільцем.

Це кільце називається фактор-кільцем кільця R за ідеалом H і позначається символом R/H.

Зазначимо, якщо кільце R – комутативне, то кожне його фактор-кільце також комутативне.

Розглянемо відображення σ_H: R $\longrightarrow$ R/H, що визначається рівністю $\sigma_H(x) = x + H$, $x \in R$. Тоді

$$\sigma_H(x + y) = x + y + H = x + H + y + H = \sigma_H(x) + \sigma_H(y);$$

$$\sigma_H(xy) = xy + H = (x + H)(y + H) = \sigma_H(x)\ \sigma_H(y).$$

Отже, як і у випадку груп, приходимо до відображень, що зберігають операції.

Нехай R, S – кільця. Відповідно до загального означення *відображення f: R $\longrightarrow$ S називається кільцевим гомоморфізмом, якщо воно задовільняє такі умови*

$$f(x + y) = f(x) + f(y),\ f(xy) = f(x)\ f(y)$$

для довільних елементів $x, y \in R$.

Оскільки у цій частині будемо розглядати тільки кільця, то далі замість слів “кільцевий гомоморфізм” будемо писати “гомоморфізм”.

Бієктивний гомоморфізм називається ізоморфізмом.

Якщо f: R $\longrightarrow$ S - ізоморфізм, то, як було показано у розділі 1, f^{-1}: S $\longrightarrow$ R також буде ізоморфізмом. *Кільця R і S називаються ізоморфними, якщо існує ізоморфізм одного з них на другий*, цей факт будемо позначати так R $\cong$ S.

Очевидно, тотожна підстановка ε_R: R ⟶ R є ізоморфізмом.

Якщо f: R ⟶ S, g: S ⟶ U – гомоморфізми, то добуток g ☺ f є також гомоморфізмом.

8.10. Пропозиція. *Нехай R, S – кільця, f: R ⟶ S – гомоморфізм.*

(1) $f(O_R) = O_S$.

(2) $f(–x) = – f(x)$ для будь-якого $x \in R$.

(3) Якщо H – підкільце R, то його образ f(H) буде підкільцем в S. Зокрема, f(R) =Im f – підкільце S.

(4) Якщо V – підкільце S, то його повний прообраз $f^{-1}(V)$ – підкільце в R.

(5) Якщо V – ідеал S, то його повний прообраз $f^{-1}(V)$ – ідеал в R. Зокрема,

$$Ker\ f = \{ x \in R \mid f(x) = O_S \} = f^{-1}(<O_S>) \text{ – ідеал } R.$$

(6) f(e) – одиничний елемент підкільця Im f.

(7) якщо R – комутативне, то і Im f – комутативне.

Д о в е д е н н я. Оскільки f – гомоморфізм адитивних груп R і S, то твердження (1), (2) випливають з пропозиції 6.2. Із цієї ж теореми випливає, що f(H) – підгрупа адитивної групи S, а $f^{-1}(V)$ – підгрупа адитивної групи кільця R. Нехай x, y ∈ H, a = f(x), b = f(y), тоді ab = f(x) f(y) = f(xy) ∈ f(H). Із теореми 7.4 випливає, що f(H) є підкільцем S.

Нехай x, y ∈ $f^{-1}(V)$, тоді f(x), f(y) ∈ $f^{-1}(V)$, і оскільки V – підкільце, то f(x) f(y) = f(xy) ∈ V, тобто x y ∈ $f^{-1}(V)$. З теореми 7.4 випливає тепер, що $f^{-1}(V)$ є підкільцем R. Отже, твердження (3), (4) також доведені. Нехай V – ідеал S, x ∈ $f^{-1}(V)$, y ∈ R, тоді f(xy) = f(x) f(y) ∈ V, тобто x y ∈ $f^{-1}(V)$. Аналогічно доводиться, що і y x ∈ $f^{-1}(V)$, а це і означає, що $f^{-1}(V)$ – ідеал кільця R, так що і (5) доведено.

Якщо a ∈ Im f, то a = f(x) для деякого x ∈ R, тому

a = f(x) = f(xe) = f(x) f(e) = a f(e); a = f(x) = f(ex) = f(e) f(x) = f(e) a,

тобто f(e) – одиничний елемент Im f.

Нарешті, нехай R – комутативне, a, b $\in$ Im f, a = f(x), b = f(y) для деяких x, y $\in$ R. Тепер

$$ab = f(x)\,f(y) = f(xy) = f(yx) = f(y)\,f(x) = ba.$$

Ідеал Ker f знову назвемо ядром гомоморфізму f.

Ін'єктивний гомоморфізм називається мономорфізмом; сюр'єктивний гомоморфізм називається епіморфізмом.

Гомоморфізм кільця R у себе називається ендоморфізмом. Ізоморфізм кільця R на себе називається автоморфізмом.

Як і для груп доводяться такі твердження.

8.11. Теорема (теорема про мономорфізми). *Нехай R, S – кільця. Гомоморфізм f: R $\longrightarrow$ S тоді і тільки тоді буде мономорфізмом, коли Ker f = $< O_R >$. Якщо f: R $\longrightarrow$ S – мономорфізм, то R $\cong$ Im f.*

8.12. Теорема (теорема про епіморфізми). *Нехай R, S – кільця, f: R $\longrightarrow$ S – епіморфізм. Тоді кільце S ізоморфне до фактор-кільця R/Ker f.*

8.13. Теорема (теорема про гомоморфізми). *Нехай R, S – кільця, f: R $\longrightarrow$ S – гомоморфізм. Тоді R/Ker f $\cong$ Im f $\leq$ S.*

8.14. Наслідок (теорема про гомоморфізми для тіл). *Нехай T, Z – тіла, f: T $\longrightarrow$ Z – ненульовий гомоморфізм (тобто Ker f $\neq$ T). Тоді f(e) – одиничний елемент Z, Im f – підтіло Z і f – мономорфізм, Зокрема, T ізоморфне деякому підтілу Z.*

Д о в е д е н н я. Дійсно, припустимо, що f(e) = O_Z. Для кожного елементу x $\in$ T маємо тоді

$$f(x) = f(xe) = f(x)\,f(e) = f(x)\,O_Z = O_Z,$$

що суперечить умовам про f. Таким чином, f(e) – ненульовий елемент Z, і отже, f(e) має обернений в Z.

З визначення одиничного елементу отримаємо $f(e) = f(ee) = f(e)\, f(e)$, і помноживши обидві частини цієї рівності на елемент, обернений до $f(e)$, отримаємо, що $f(e) = e_Z$ – одиничний елемент Z. З пропозиції 8.10 випливає, що Im f – підкільце Z, більш того, тільки що ми довели, що Im f – унітарне підкільце Z. Нехай u – ненульовий елемент Im f, тоді $u = f(y)$ для деякого елементу y кільця R. Очевидно y – ненульовий елемент R, так що для нього існує обернений елемент y^{-1}. Покажемо, що елемент $v = f(y^{-1})$ буде оберненим до u, з теореми 7.10 буде випливати тоді, що Im f – підтіло Z. Маємо

$$uv = f(y)\, f(y^{-1}) = f(y\, y^{-1}) = f(e) = = e_Z,$$
$$vu = f(y^{-1})\, f(y) = f(y^{-1}\, y) = f(e) = e_Z.$$

За пропозицією 8.10, Ker f – ідеал кільця R. Оскільки Ker $f \neq$ T, то з наслідку 8.7 отримаємо рівність Ker $f = \{O_R\}$. Залишається застосувати теорему 8.11.

Розглянемо деякі застосування отриманих результатів.

Характеристика кільця.

Нехай R – кільце, розглянемо відображення $f : Z \longrightarrow R$, що задається правилом $f(n) = ne$, $n \in Z$. Очевидно,

$$f(n + k) = (n + k)e = ne + ke = f(n) + f(k),$$

$$f(nk) = (nk)e = n(ke) = (ne)(ke) = f(n)\, f(k)$$

для всіх $n, k \in Z$.

З пропозиції 8.10 випливає, що $\text{Im } f = \{ne \mid n \in Z\} = Z\, e$ – підкільце R. Із теореми 8.13 отримуємо, що $Z\, e = \text{Im } f \cong Z / \text{Ker } f$. Оскільки Ker f – ідеал кільця Z за пропозицією 8.10, то з опису ідеалів кільця Z маємо Ker $f = n\, Z$, де або $n \geq 0$.

Якщо Ker $f = <0>$, то $Z\, e \cong Z$. *У цьому випадку будемо говорити, що кільце R має характеристику 0 і писати char R = 0.* Це означає також, що одиничний елемент e має в адитивній групі кільця R нескінченний порядок.

Якщо Ker f = n Z, де n > 0, то Z e ≅ Z /n Z. *У цьому випадку будемо говорити, що кільце R має характеристику n і писати char R = n.* Це означає, що одиничний елемент e має в адитивній групі кільця R порядок n. Якщо a – довільний елемент R, то na = (ne)a = O_R a = O_R, так що порядок кожного елемента (в адитивній групі кільця) є дільником числа n.

Якщо n не є простим числом, то n = kt, де 1 < k < n, 1 < t < n. Зокрема, k, t ∉ nZ, тому ke ≠ O_R, te ≠ O_R. Разом з тим

$$(ke)\ (te) = (kt)e = ne = O_R.$$

Отже, якщо char R не є простим числом, то кільце R містить дільники нуля.

8.15. Пропозиція. *Якщо кільце R не має дільників нуля, то або char R = 0, або char R = p, де p – просте число. Зокрема, якщо T – тіло, то або char T = 0, або char T = p, де p – просте число.*

Розглянемо будову простих підтіл.

16.16. Пропозиція. *Якщо T – тіло, тоді його центр ζ(T) є підполем.*

Д о в е д е н н я. Оскільки O_T, e ∈ ζ(T), то ζ(T) – ненульова підмножина T. Уже зазначалося, що ζ(T) – підкільце T, зокрема умови (SF 1) і (SF 2) виконані для ζ(T). Нехай $O_T \neq x \in \zeta(T)$, y – довільний елемент T. Тоді

xy = yx. Маємо

$$y\,x^{-1} = e\,y\,x^{-1} = (x^{-1}x)\,y\,x^{-1} = x^{-1}(xy)\,x^{-1} = x^{-1}(yx)\,x^{-1} = (x^{-1}y)(x\,x^{-1}) = (x^{-1}y)\,e = x^{-1}y.$$

Ця рівність показує, що $x^{-1} \in \zeta(T)$, тобто і умова (SF 3) також виконана.

8.17. Наслідок. *Просте підтіло T_0 тіла T є його підполем.*

Наведемо приклади простих полів.

Поле Q раціональних чисел.

Нехай P – деяке підполе Q. Із (SF 2) випливає, що $1 \in P$. Оскільки P – підкільце Q, то для будь-якого $r \in Z$ маємо $r\,1 = r \in P$. Якщо $0 \neq k \in P$ і $k \in Z$, то із (SF 3) отримуємо, що $1/k \in P$. Тепер для всіх $r, k \in Z$, $k \neq 0$, маємо $r/k = r\,(1/k) \in P$. Це співвідношення доводить рівність $P = Q$.

Поле F_p $= Z/pZ$, де p – просте число. Покажемо спочатку, що фактор-кільце Z/pZ є полем. Нехай $pZ \neq x + pZ \in Z/pZ$. Оскільки $x \notin pZ$, то числа x і p взаємно прості. Тому знайдуться такі цілі числа k, r, що $xk + pr = 1$. Тепер маємо

$$1 + pZ = xk + pr + pZ = xk + pZ = (x + pZ)(k + pZ).$$

Якщо P – підполе F_p, то P – підкільце F_p, а тому адитивна група P є підгрупою адитивної групи F_p. Із теореми Лагранжа (наслідок 4.5) одержуємо, що $|P|$ ділить $|F_p| = p$. Оскільки p – просте число і $|P| \geq 2$, то це означає, що $|P| = p$. Отже, $P = F_p$.

Наступна теорема свідчить, що інших прикладів простих тіл немає.

8.18. Теорема. *Нехай T – тіло, T_0 – його просте підтіло.*
(1) Якщо char T = p – просте число, то $T_0 \cong F_p$.
(2) Якщо char T = 0, $T_0 \cong Q$.

Д о в е д е н н я. Розглянемо відображення $f : Z \longrightarrow T$, що задається правилом $f(n) = ne$, $n \in Z$. Як було показано вище, це відображення є гомоморфізмом. Якщо char T = p – просте число, то Ker f = pZ. Відображення $\psi: Z/pZ \longrightarrow T$, що визначається за правилом $\psi(n + pZ) = f(n)$, $n \in Z$, також буде гомоморфізмом, і цей гомоморфізм буде ненульовим, оскільки $\psi(1 + pZ) = f(1) = e \neq O_T$. Очевидно, Im ψ = Im f = Z e і залишається застосувати пропозицію 8.14.

Нехай char T = 0. У цьому випадку Ker f = <0>, тобто f – мономорфізм. Оскільки $e \in T_0$, то Z e – підкільце T_0. Розширимо відображення f до відображення f_1: $Q \longrightarrow T$. Нехай m/n –

довільний елемент Q, тоді $n \neq 0$, і отже, $f(n) = ne \neq O_T$. Оскільки T – тіло, його ненульовий елемент ne має в T обернений. Отримаємо: $f_1(m/n) = (me)(ne)^{-1}$. Покажемо, що f_1 визначено коректно, нехай $k/t = m/n$. Маємо такий ланцюжок рівностей:

$$kn = tm \Rightarrow (ke)(ne) = (te)(me) \Rightarrow (me)(ne)^{-1} = (ke)(te)^{-1} \Rightarrow$$
$$\Rightarrow f_1(m/n) = (me)(ne)^{-1} = (ke)(te)^{-1} = f_1(k/t),$$

так що визначення f_1 є коректним. Відображення f_1 продовжує f: якщо $n \in Z$, то $n = nk/k$ для довільного $0 \neq k \in Z$, отже,

$$f_1(n) = f_1(nk/k) = (nke)(ke)^{-1} = (ne)(ke)(ke)^{-1} = ne = f(n).$$

Відображення f_1 є гомоморфізмом:

$$f_1((m/n) + (k/t)) = f_1((mt + kn)/(nt)) = ((mt + kn)e)((nt)e)^{-1} =$$
$$= ((mt)e + (kn)e)((ne)(te))^{-1} = ((me)(te) + (ke)(ne))(ne)^{-1}(te)^{-1} =$$
$$= (me)(te)(ne)^{-1}(te)^{-1} + (ke)(ne))(ne)^{-1}(te)^{-1} =$$
$$= (me)(ne)^{-1} + (ke)(te)^{-1} = f_1(m/n) + f_1(k/t);$$

$$f_1((m/n)\,(k/t)) = f_1((mk)/(nt)) = ((mk)e)((nt)e)^{-1} = ((me)(ke)((ne)(te))^{-1} =$$
$$= (me)(ke)(ne)^{-1}(te)^{-1} = (me)(ne)^{-1}\,(ke)(te)^{-1} = f_1(m/n)\,f_1(k/t).$$

За наслідком 16.14 $Q \cong \mathrm{Im}\, f_1$, а оскільки $\mathrm{Im}\, f_1$ – підполе T, то $\mathrm{Im}\, f_1 \geq T_0$. З іншого боку, $ne \in T_0$ для кожного $n \in Z$, і якщо $n \neq 0$, то $(ne)^{-1} \in T_0$, отже, $(me)(ne)^{-1} \in T_0$ для кожної пари $m, n \in Z$, $n \neq 0$. Таким чином, навпаки $\mathrm{Im}\, f_1 \leq T_0$, а тому $\mathrm{Im}\, f_1 = T_0$.

Наступний результат дає можливість отримувати розширення кілець.

8.19. Теорема. *Нехай R, K – кільця, $f: R \longrightarrow K$ – мономорфізм. Тоді існує кільце S, ізоморфне K, яке містить R як підкільце. Якщо e_R, e_K – одиничні елементи кілець R та K відповідно і $f(e_R) = e_K$, то e_R – одиничний елемент S. Нарешті, якщо R та K – тіла, то R – підтіло S.*

Д о в е д е н н я. Можна вважати, що $K \cap R = \varnothing$, якщо це не так, то замість K розглянемо його ізоморфний образ, що має пустий перетин з R. Наприклад, покладемо $K_1 = K \times \{1\}$ і визначимо операції за правилом

$$(x, 1) + (y, 1) = (x + y, 1); (x, 1)(y, 1) = (xy, 1).$$

Нехай $U = \operatorname{Im} f$. З теореми 8.11 випливає, що U – підкільце K і U ізоморфне R. Нехай $A = K \setminus U$, $S = R \cup U$. Задамо відображення $g: S \longrightarrow K$ правилом

$$g(x) = \begin{cases} f(x), \text{ якщо } x \in R \\ \\ x, \text{ якщо } x \in A. \end{cases}$$

Неважко перевірити, що g – бієкція S на K. Визначимо операції + і ✕ в S за правилом

$$x + y = g^{-1}(g(x) + g(y)),\ x \times y = g^{-1}(g(x)\, g(y))$$

для будь-яких $x, y \in S$. З цього означення випливає, що

$$g(x + y) = g(g^{-1}(g(x) + g(y))) = g(x) + g(y),$$
$$g(x \times y) = g(g^{-1}(g(x)\, g(y))) = g(x)\, g(y).$$

Будемо мати

$$x + y = g^{-1}(g(x) + g(y)) = g^{-1}(g(y) + g(x)) = y + x;$$

$$(x + y) + z = g^{-1}(g(x + y)) + g(z)) = g^{-1}(g(x) + g(y)) + g(z)) =$$
$$= g^{-1}(g(x) + (g(y) + g(z)) = g^{-1}(g(x) + g(y + z)) = x + (y + z);$$

$$(x + O_R) = g^{-1}(g(x) + g(O_R)) = g^{-1}(g(x) + O_K) = g^{-1}(g(x)) = x;$$

якщо $x \in R$, то

$$x + (-x) = g^{-1}(g(x) + g(-x))) = g^{-1}(g(x + (-x)) = g^{-1}(g(O_R)) = O_R.$$

Якщо $x \in A$, то $-x \notin U$, бо U - підкільце; тому

$$x + (-x) = g^{-1}(g(x) + g(-x))) = g^{-1}(x + (-x)) = g^{-1}(O_K)) = O_R.$$

Отже, відносно операції додавання S є абелевою групою.

Далі

$$(x + y) \times z = g^{-1}(g(x + y))\ g(z)) = g^{-1}(g(x) + g(y))\ g(z)) = g^{-1}(g(x)\ g(z)$$
$$g(y)\ g(z)) = g^{-1}(g(x \times z) + g(y \times z)) = g^{-1}(g(x \times z) + g(y \times z)) = (x \times z) +$$
$$(y \times z)).$$

Аналогічно доводиться рівність

$$x \times (y + z) = (x \times y) + (x \times z).$$

Продовжуємо перевірку:

$$(x \times y) \times z = g^{-1}(g(x \times y))\ g(z)) = g^{-1}(g(x)\ g(y))\ g(z)) =$$
$$= g^{-1}(g(x)\ (g(y)\ g(z)) = g^{-1}(g(x)\ g(y \times z)) = x \times (y \times z);$$

$$x \times g^{-1}(e_K) = g^{-1}(g(x)\ g^{-1}(g(e_K))) = g^{-1}(g(x)\ e_K) = g^{-1}(g(x)) = x;$$

$$g^{-1}(e_K) \times x = g^{-1}(g^{-1}(g(e_K))\ g(x)) = g^{-1}(e_K\, g(x)) = g^{-1}(g(x)) = x,$$

так що $g^{-1}(e_K)$ – одиничний елемент кільця S. Зокрема, якщо $f(e_R) = g(e_R) = e_K$, то e_R – одиничний елемент кільця S. Отже, S – кільце. Уже було показано, що

$$g(x + y) = g(x) + g(y),\quad g(x \times y) = g(x)\ g(y),$$

і тому g – ізоморфізм.

Якщо $x, y \in R$, то

$$x + y = g^{-1}(g(x) + g(y)) = g^{-1}(f(x) + f(y)) = g^{-1}(f(x + y)) = g^{-1}(g(x + y)) = x + y;$$

$$x \times y = g^{-1}(g(x)\ g(y)) = g^{-1}(f(x)\ f(y)) = g^{-1}(f(x\ y)) = g^{-1}(g(x\ y)) = x\ y,$$

тобто звуження нових операцій на R збігаються з початковими операціями на R. Із теореми 7.4 випливає також, що R – підкільце в S.

Нарешті, нехай R, K – тіла. Теорема 7.10 показує, що залишилось тільки перевірити умову (SF 3). Якщо x – ненульовий елемент R, то він має обернений в R. Оскільки одиничний елемент R є також одиничним елементом S, то цей обернений буде оберненим до x і в тілі S.

Ця теорема показує, що для побудови деякого розширення E кільця R достатньо побудувати мономорфізм кільця R у яке-небудь кільце K, що ізоморфне E. Зазначимо деякі застосування здобутих результатів.

У розділі 15 було розглянуто тільки кільця, в яких є одиничний елемент. Наступна теорема дає обгрунтування для цієї домовленості.

8.20. Теорема. *Для кожного кільця R існує кільце K з одиницею, яке містить R як підкільце.*

Д о в е д е н н я. Якщо кільце R має одиничний елемент, то все доведено. Тому припустимо, що одиничного елементу R немає. З огляду на теорему 8.18 досить побудувати мономорфізм $f: R \longrightarrow K$, де K – кільце з одиницею.

Нехай $K = R \times Z$, задамо операції в K правилами

$$(x, n) + (y, k) = (x + y, n + k),$$

$$(x, n)\,(y, k) = (xy + kx + ny, nk)$$

для всіх $x, y \in R$, $n, k \in Z$.

Оскільки додавання в K зводиться до додавання в R і Z, то воно успадковує властивості додавання в R і Z. Зокрема, K є абелевою групою. При цьому

$$O_K = (O_R, 0), \; -(x, n) = (-x, -n).$$

Далі

$$((x, n) + (y, k))(z, m) = (x + y, n + k)(z, m) = ((x + y)z + (n + k)z + m(x + y),$$

$$(n + k)m) = (xz + yz + nz + kz + mx + my, nm + km),$$

$$(x, n)(z, m) + (y, k)(z, m) = (xz + nz + mx, nm) + (yz + kz + my, km) =$$

$$= (xz + nz + mx + yz + kz + my, nm + km),$$

тобто

$$((x, n) + (y, k))(z, m) = (x, n)(z, m) + (y, k)(z, m).$$

Аналогічно доводиться рівність

$$(x, n)\,((y, k) + (z, m)) = (x, n)(y, k) + (x, n)(z, m).$$

Множення в К асоціативне:

$$((x, n)\,(y, k))(z, m) = (xy + ny + kx, nk)\,(z, m) =$$

$$= ((xy + ny + kx)z + m(xy + ny + kx) + (nk)z, (nk)m) =$$

$$= ((xy)z + n(yz) + k(xz) + m(xy) + (mn)y + (mk)x + (nk)z, (nk)m);$$

$$(x, n)\,((y, k)(z, m)) = (x, n)\,(yz + kz + my, km) =$$

$$= (x(yz + kz + my) + n(yz + kz + my) + (km)x, n(km)) =$$

$$= (x(yz) + k(xz) + m(xy) + n(yz) + (nk)z + (nm)y + (km)x, n(km)).$$

Оскільки множення в R і Z асоціативне, то

$$((x, n)\,(y, k))(z, m) = (x, n)\,((y, k)(z, m)).$$

Нарешті,

$$(x, n)\,(O_R, 1) = (x, n) = (O_R, 1)\,(x, n),$$

отже, $(O_R, 1)$ – одиничний елемент кільця K.

Задамо відображення $f: R \longrightarrow K$ за правилом $f(x) = (x, 0)$, $x \in R$. Очевидно,

$$f(x + y) = (x + y, 0) = (x, 0) + (y, 0) = f(x) + f(y);$$

$$f(xy) = (xy, 0) = (x, 0)\,(y, 0) = f(x)\,f(y)$$

для будь-яких $x, y \in R$, а також $\mathrm{Ker}\, f = < O_R >$. Отже, f – шуканий мономорфізм.

У теоремі 6.12 зазначалось, що будь-яка група вкладається в групу підстановок деякої множини. Наступна теорема показує, що для кілець роль такого “універсального об’єкта” відіграють кільця ендоморфізмів абелевих груп.

8.21. Теорема. *Для кожного кільця R існує мономорфізм у кільце ендоморфізмів деякої абелевої групи.*

Д о в е д е н н я. На підставі теореми 8.19 можна вважати, що кільце R має одиничний елемент. Нехай $a \in R$, розглянемо відображенн τ_a: $R \longrightarrow R$, що визначається рівністю $\tau_a(x) = ax$, $x \in R$. Маємо

$$\tau_a(x + y) = a(x + y) = ax + ay = \tau_a(x) + \tau_a(y),$$

і це означає, що τ_a – ендоморфізм адитивної групи кільця R.

Розглянемо відображення $f: R \longrightarrow \text{End } R$, яке задається правилом $f(a) = \tau_a$, для кожного $a \in R$. Якщо $x \in R$, то

$$\tau_{a+b}(x) = (a + b)x = ax + bx = \tau_a(x) + \tau_b(x) = (\tau_a + \tau_b)(x);$$

$$\tau_{a\,b}(x) = (ab)x = a(bx) = \tau_a(bx) = \tau_a(\tau_b(x)) = \tau_a \text{ ☺ } \tau_b(x);$$

Звідси випливають рівності

$$f(a + b) = \tau_{a+b} = \tau_a + \tau_b = f(a) + f(b);$$

$$f(ab) = \tau_{a\,b} = \tau_a \text{ ☺ } \tau_b = f(a) \text{ ☺ } f(b)$$

для будь-яких $a, b \in R$. Ці рівності показують, що f – гомоморфізм. Нарешті, якщо $a \neq b$, то

$$\tau_a(e) = ae = a \neq b = be = \tau_b(e),$$

так що $f(a) = \tau_a \neq \tau_b = f(b)$. Тому f – шуканий мономорфізм.

РОЗДІЛ 9
КІЛЬЦЕ ПОЛІНОМІВ ТА ФОРМАЛЬНИХ СТУПЕНЕВИХ РЯДІВ

Кільце поліномів є одним з більш старих та добре вивчених кілець. Це одне з найбільш важливих кілець, його роль у теорії комутативних кілець подібна до ролі груп підстановок для теорії скінчених груп. У термінах поліномів формулюються та вирішуються найрізноманітніші проблеми не тільки в алгебрі, а і в багатьох інших математичних дисциплінах. Розглянемо одну задачу, що приводить до цього кільця. Нехай K – комутативне кільце, R – його унітарне цілісне підкільце, M – підмножина K. Розглянемо сукупність M тих підкілець, які містять у собі як підкільце R, так і підмножину M. Нехай $R[M] = \cap$ M. За наслідком 7.5 R[M] – підкільце. З його означення випливає, що R[M] – найменше підкільце, яке містить у собі підкільце R і підмножину M. Будемо говорити, що *R[M] – підкільце, породжене над R підмножиною M.*

Найпростійший випадок, коли M складається з єдиного елемента y. У цьому випадку замість R[M] будемо писати R[y]. З'ясуємомо, з яких елементів складається R[y]. З теореми 7.4 випливає, що $y^n \in R[y]$ для кожного $n \in N$. З цієї ж теореми отримуємо, що $ay^n \in R[y]$ для кожного $a \in R$, і отже, усі можливі суми вигляду

$$a_0 + a_1 y + a_2 y^2 + \ldots + a_n y^n \text{ (поліноми від y)},$$

де $a_0, a_1, a_2, \ldots, a_n \in R$, містяться у R[y]. Більш того, покажемо, що всі ці суми вичерпують підкільце R[y]. Дійсно, розглянемо дві довільні такі суми $(a_0 + a_1 y + a_2 y^2 + \ldots + a_n y^n)$ і $(b_0 + b_1 y + b_2 y^2 + \ldots + b_k y^k)$ та нехай, для визначеності $n \geq k$. Маємо

$$(a_0 + a_1 y + a_2 y^2 + \ldots + a_n y^n) + (b_0 + b_1 y + b_2 y^2 + \ldots + b_k y^k) =$$
$$= (a_0 + b_0) + (a_1 + b_1) y + b_2 y^2 + \ldots (a_k + b_k) y^k + a_{k+1} y^{k+1} + \ldots + a_n y^n;$$

$$(a_0 + a_1 y + a_2 y^2 + \ldots + a_n y^n)(b_0 + b_1 y + b_2 y^2 + \ldots + b_k y^k) =$$
$$= (a_0 + b_0) + (a_0 b_1 + a_1 b_0) y + (a_0 b_2 + a_1 b_1 + a_2 b_0) y^2 + \ldots + a_n b_k y^{n+k}.$$

З теореми 7.4 отримуємо, що всі такі суми утворюють підкільце. Оскільки воно містить у собі як R, так і y, то воно збігається з R[y]. Потрібно зазначити, що y – деякий елемент кільця K, тому різні за виглядом поліноми від y можуть збігатися. Тому щоб прийти до звичного поняття поліному, потрібно звільнитися від усіх подібних співвідношень. На цій дорозі і з'являється поняття "невідомої". Але, у свою чергу, тут ми мусимо якось її визначити. Тому віддамо перевагу іншому підходу, де основний акцент переноситься на операції над поліномами, що виникли вище. Кожен поліном повністю визначається через задання своїх коефіцієнтів. З набору коефіцієнтів та операцій і будемо виходити. Більш того, кільце поліномів отримаємо як підкільце у деякому більшому кільці – кільці формальних ступеневих рядів.

Нехай R – область цілісності. Розглянемо множину R[[X]] (поки що це просто формальне позначення) усіх послідовностей :

$$(a_n)_{n \in N0} = (a_0, a_1, \ldots, a_n, \ldots),$$

що складаються з елементів кільця R.

Дві послідовності $(a_n)_{n \in N0}$ *та* $(b_n)_{n \in N0}$ *називаються рівними, якщо* $a_n = b_n$ *для кожного* $n \in N_0$.

Операції додавання і множення послідовностей задаються правилами, що були вже вказані вище, а саме:

$$(a_n)_{n \in N0} + (b_n)_{n \in N0} = (c_n)_{n \in N0}, \text{ де } c_n = a_n + b_n \text{ для кожного } n \in N_0;$$

$$(a_n)_{n \in N0} \; (b_n)_{n \in N0} = (d_n)_{n \in N0},$$

$$\text{де } d_n = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0 =$$

$$= \sum_{0 \le j \le n} a_j b_{n-j} = \sum_{j+k=n} a_j b_k \text{ для кожного } n \in N_0.$$

Додавання послідовностей зводиться до додавання відповідних елементів кільця R, через це неважко впевнитися у тому, що воно комутативне і асоціативне. Нульовим елементом, очевидно, буде послідовність

$$O = (O_R, O_R, \ldots, O_R, \ldots).$$

Для кожної послідовності $(a_n)_{n \in N0}$ існує протилежний елемент – це послідовність $(- a_n)_{n \in N0}$. Отже, для операції додавання виконані всі аксіоми абелевої групи.

Неважко упевнитись у тому, що множення послідовностей комутативне, тому треба доводити тільки один закон дистрибутивності:

$$((a_n)_{n \in N0} + (b_n)_{n \in N0}) (c_n)_{n \in N0} =$$
$$= (a_n)_{n \in N0} (b_n)_{n \in N0} + (a_n)_{n \in N0} (c_n)_{n \in N0}.$$

Позначимо

$$(u_n)_{n \in N0} = ((a_n)_{n \in N0} + (b_n)_{n \in N0}) (c_n)_{n \in N0},$$
$$(v_n)_{n \in N0} = (a_n)_{n \in N0} (b_n)_{n \in N0} + (a_n)_{n \in N0} (c_n)_{n \in N0}.$$

Маємо для будь-якого $n \in N_0$

$$u_n = \sum_{j + k = n} (a_j + b_j) c_k = \sum_{j + k = n} a_j c_k + \sum_{j + k = n} b_j c_k ;$$
$$v_n = \sum_{j + k = n} a_j c_k + \sum_{j + k = n} b_j c_k .$$

Тобто $u_n = v_n$ при кожному $n \in N_0$, отже, дистрибутивний закон доведено.

Множення асоціативне:

$$((a_n)_{n \in N0} (b_n)_{n \in N0}) (c_n)_{n \in N0} = (a_n)_{n \in N0} ((b_n)_{n \in N0} (c_n)_{n \in N0}).$$

Дійсно, нехай

$$((a_n)_{n \in N0} (b_n)_{n \in N0}) (c_n)_{n \in N0} = (w_n)_{n \in N0},$$

$$(a_n)_{n \in N0} ((b_n)_{n \in N0} (c_n)_{n \in N0}) = (z_n)_{n \in N0}.$$

Маємо тепер при довільному $n \in N_0$

$$w_n = \sum_{m + t = n} \Big(\sum_{j + k = m} a_j b_k \Big) c_t = \sum_{m + j + k = n} (a_j b_k)c_t ;$$

$$z_n = \sum_{j + m = n} a_j \Big(\sum_{k + t = m} a_j b_k \Big)c_t = \sum_{m + j + k = n} a_j (b_k c_t).$$

Тобто $w_n = z_n$ при кожному $n \in N_0$, отже, асоціативність множення доведена.

Відносно операції множення існує одиничний елемент, це, як неважко показати,

$$e = (e, O_R, \ldots, O_R, \ldots).$$

Отже, R[[X]] – комутативне кільце з одиницею. Більше того, R[[X]] – область цілісності. Насправді, нехай $(a_n)_{n \in N0} \neq O$, $(b_n)_{n \in N0} \neq O$. Тоді знайдуться такі індекси j, k, що $a_j \neq O_R$, але $a_n = O_R$ при $0 \leq n \leq j$, $b_k \neq O_R$, але $b_n = O_R$ при $0 \leq n \leq k$. Якщо $(a_n)_{n \in N0}$ $(b_n)_{n \in N0} = (d_n)_{n \in N0}$, то

$$d_{j+k} = a_0 b_{j+k} + \ldots + a_{j-1} b_{k+1} + a_j b_k + a_{j+1} b_{k-1} + \ldots + a_{j+k} b_0 = a_j b_k \neq O_R.$$

Послідовність $(a_n)_{n \in N0}$ назвемо поліномом, якщо знайдеться таке число t, що $a_n = O_R$ при $n \geq t$. Ступенем полінома $(a_n)_{n \in N0}$ називають таке число m, що $a_m \neq O_R$, але $a_n = O_R$ при $n > m$. При цьому a_m називають старшим коефіцієнтом полінома. Якщо $a_n = O_R$ для будь-якого $n \in N_0$, то ніякого степеня поліному не присвоюють, іноді нульовому елементу присвоюють степінь $-\infty$.

Підмножину всіх поліномів позначимо через R[X]. Нехай $(a_n)_{n \in N0}$, $(b_n)_{n \in N0}$ – два поліноми і m, t – такі числа, що $a_m \neq O_R$, але $a_n = O_R$ при $n > m$, $b_t \neq O_R$, але $b_n = O_R$ при $n > t$.

Тоді $a_n - b_n = O_R$ при $n > \max\{m, t\}$; $a_0 b_n + a_1 b_{n-1} + \ldots + a_j b_{n-j} + \ldots + a_n b_0 = O_R$ при $n > m = t$. З теореми 7.4 випливає, що R[X] – підкільце кільця R[[X]]. Очевидно також, що R[X] – унітарне підкільце.

Нехай $X = (O_R, e, O_R, \ldots, O_R, \ldots)$. З означення множення одержуємо рівність

$$X^n = (O_R, O_R, \ldots, O_R, e, O_R, \ldots, O_R, \ldots).$$
$$\downarrow$$
$$n$$

Для довільного елемента $a \in R$ позначимо $\bar{a} = (a, O_R, O_R, \ldots, O_R, \ldots)$. Тоді

$$\bar{a}\, X^n = (\, O_R, O_R, \ldots, O_R, \underset{\underset{n}{\downarrow}}{a}, O_R, \ldots, O_R, \ldots).$$

Нехай $(a_n)_{n \in N0}$ – поліном ступеня m. Тепер можна переписати його в звичайному вигляді:

$$(a_n)_{n \in N0} = (a_0, O_R, O_R, \ldots, O_R, O_R, O_R, \ldots) +$$
$$+ (O_R, a_1, O_R, \ldots, O_R, O_R, O_R, \ldots) + \ldots +$$
$$+ (\, O_R, O_R, O_R, \ldots, O_R, a_m, O_R, \ldots) = \bar{a}_0 + \bar{a}_1 X + \ldots + \bar{a}_m X^m.$$

Далі,

$$\bar{a} + \bar{b} = (a + b, O_R, \ldots, O_R, \ldots) = \overline{a + b};$$

$$\bar{a}\,\bar{b} = (a\,b, O_R, \ldots, O_R, \ldots) = \overline{a\,b}.$$

Ці рівності показують, що відображення $a \longrightarrow \bar{a}$, $a \in R$, є гомоморфізмом R в $R[[X]]$. Очевидно, цей гомоморфізм є бієктивним, тобто він є мономорфізмом. Зокрема, R і його образ ізоморфні. Оскільки в алгебрі ізоморфні об'єкти ототожнюються, можемо ототожнити R з його образом і замість $\bar{a}$ писати a, так що приходимо до звичайного запису полінома $f(X) = a_0 + a_1 X + \ldots + a_m X^m$. Надалі ступінь полінома $f(X)$ будемо позначати через $\deg f(X)$. З уже зазначеного вище випливають співвідношення

$$\deg (f(X) \pm g(X)) \le \max \{\deg f(X), \deg g(X)\},$$

$$\deg (f(X)\, g(X)) = \deg f(X) + \deg g(X)\}.$$

З другої рівності одразу випливає, що $U(R[X]) = U(R)$. Також з цієї рівності випливає, що $R[X]$ не має дільників нуля.

Повернемося до формальних ступеневих рядів. Зокрема, отримаємо і для них запис у вигляді нескінченної суми. Для випадку довільного кільця не маємо поняття границі послідовності, тому тут буде необхідно зробити деякі домовленості, щоб мати можливість додавати нескінченну множину послідовностей.

Набір $\{ A_j \mid j \in N_0\}$ *послідовностей* $A_j = (a_{j\,n})_{n \in N0}$ *називається сумовим, якщо для будь-якого* $n \in N_0$ *множина* $\{ a_{j\,n} \mid j \in N_0 \}$ *містить тільки скінчену підмножину ненульових елементів. У цьому випадку сумою* $\sum_{j \in N0} A_j$ *набору послідовностей* $\{A_j \mid j \in N_0\}$ *називають послідовність* $(b_n)_{n \in N0}$, *де* b_n *– це сума тих елементів множини* $\{ a_{j\,n} \mid j \in N_0 \}$, *які відмінні від нуля.*

Прокоментуємо це означення. Щоб з'ясувати, чи буде набір послідовностей $\{ A_j \mid j \in N_0\}$ сумовим, треба подумки записати ці послідовності одну під одною і перевірити, чи буде кожен стовпчик цієї нескінченної матриці містити лише скінчену множину ненульових елементів. Підсумовувати таку послідовність треба по стовпчиках, додаючи тільки ненульові елементи кожного стовпчика. Тепер

$$(a_n)_{n \in N0} = (a_0 , O_R, O_R, \ldots , O_R, O_R, O_R, \ldots) +$$
$$+ (O_R, a_1, O_R, \ldots , O_R, O_R, O_R, \ldots) + \ldots +$$
$$+ (O_R, O_R, O_R, \ldots , O_R, a_m , O_R, \ldots) + \ldots$$
$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots .$$

Це показує, що набір $\{ a_n X^n \mid n \in N_0 \}$ є сумовим і його сума дорівнює $(a_n)_{n \in N0}$. Отже, і для формального ступеневого ряду $(a_n)_{n \in N0}$ отримуємо звичний запис у вигляді $\sum_{n \in N0} a_n X^n$.

Нарешті, з'ясуємо, з яких рядів складається група U(R[[X]]). Нехай $f(X) = \sum_{n \in N0} a_n X^n$, $g(X) = \sum_{n \in N0} b_n X^n$ і припустимо, що $f(X)\, g(X) = e$. Звідси одразу ж випливає, що $a_0\, b_0 = e$, тобто $a_0 \in U(R)$.

Навпаки, припустимо, що $a_0 \in U(R)$. Якщо з рівності $f(X)\, g(X) = e$ можна визначити всі коефіцієнти ряду g(X), то це буде означати, що $f(X) \in U(R[[X]])$. Маємо $a_0\, b_0 = e$, тобто $b_0 = a_0^{-1}$. Далі, $a_0\, b_1 + a_1\, b_0 = O_R$. Звідси $b_1 = a_1 a_0^{-2}$. Припустимо, що коефіцієнти $b_0, b_0, \ldots , b_n$ уже визначені. Тоді коефіцієнт b_{n+1} можна визначити з рівності

$$a_0\, b_{n+1} + a_1 b_n + \ldots + a_n\, b_1 + a_{n+1} = O_R,$$

а саме:

$$b_{n+1} = (- a_1 b_n - \ldots - a_n\, b_1 - a_{n+1}) a_0^{-1}.$$

Отже, група U(R[[X]]) складається з тих і тільки тих рядів $\sum_{n \in N0} a_n X^n$, у яких $a_0 \in U(R)$. Зокрема, якщо R – поле, то U(R[[X]]) складається з тих і тільки тих рядів $\sum_{n \in N0} a_n X^n$, у яких $a_0 \neq O_R$.

Розглянемо питання про ідеали в кільцях поліномів на формальних ступеневих рядів. Будову ідеалів у кільці поліномів отримаємо з наступної важливої теореми.

9.1. Теорема. *Нехай F – поле, f(X), g(X) ∈ F[X], причому g(X) ≠ O_F. Тоді знайдуться такі поліноми q(X), r(X) ∈ F[X], що f(X) = q(X)g(X) + r(X), де або r(X) = O_F або deg r(X) < deg g(X). Таке представлення полінома f(X) є єдиним.*

Д о в е д е н н я. Нехай

$$f(X) = a_0 + a_1X + \ldots + a_nX^n,$$

$$g(X) = b_0 + b_1X + \ldots + b_kX^k,$$

де $b_k \neq O_F$. Застосуємо індукцію за числом n. Якщо deg g(X) > deg f(X), то покладемо r(X) = f(X), q(X) = O_F. Тому розглянемо випадок, коли deg g(X) ≥ deg f(X). Якщо deg f(X) = 0, то r(X) = O_F, $q(X) = a_0 b_0^{-1}$. Нехай n > 0 і припустимо, що теорема доведена для поліномів, ступінь яких менший ніж n. Поліном $a_n b_k^{-1} X^{n-k} g(X)$ має ступінь n і коефіцієнт його при X^n дорівнює a_n. Тому ступінь полінома $f(X) - a_n b_k^{-1} X^{n-k} g(X) = f_1(X)$ уже менша від n. З урахуванням індуктивного припущення $f_1(X) = q_1(X)g(X) + r(X)$, де або r(X) = O_F, або deg r(X) < deg g(X). Тепер f(X) = q(X)g(X) + r(X), де $q(X) = q_1(X) + a_n b_k^{-1} X^{n-k}$.

Нехай маємо також ще такий вираз: $f(X) = q_2(X)g(X) + r_2(X)$, де або $r_2(X) = O_F$, або deg $r_2(X)$ < deg g(X). Тоді

$$q(X)g(X) + r(X) = q_2(X)g(X) + r_2(X).$$

Звідси $g(X)(q(X) - q_2(X)) = r_2(X) - r(X)$. Поліном, розташований у правій частині цієї рівності, або нульовий, або степінь його менша deg g(X). Якщо $q(X) - q_2(X)) \neq O_F$, то

$$\deg (g(X)(q(X) - q_2(X)) = \deg g(X) + \deg (q(X) - q_2(X)) \geq \deg g(X).$$

Це означає, що $q(X) - q_2(X)) = O_F$. Звідси виходить, що $r_2(X) = r(X)$.

9.2. Наслідок. *Нехай F – поле. Тоді кожен ідеал кільця F[X] є головним.*

Д о в е д е н н я. Нехай H – ідеал кільця F[X]. Якщо $H = \{O_F\}$, то $H = O_F\ F[X]$. Припустимо, що H містить ненульові поліноми. Серед них виберемо поліном g(X) найменшого можливого ступеня. Нехай f(X) – довільний елемент H. Із теореми 9.1 отримуємо розклад $f(X) = q(X)g(X) + r(X)$, де або $r(X) = O_F$, або $\deg r(X) < \deg g(X)$. Якщо припустити, що $r(X) \neq O_F$, то $r(X) = f(X) - q(X)g(X) \in H$. Але це суперечить вибору поліному g(X). Одержана суперечність показує, що $r(X) = O_F$, тобто $f(X) = q(X)g(X)$. Звідси випливає, що $H = g(X)\ F[X]$.

Ще один приклад комутативного кільця, де кожен ідеал буде основним, отримаємо з наступної теореми.

9.3. Теорема. *Нехай F – поле. Тоді кожен ідеал кільця F[[X]] є головним.*

Д о в е д е н н я. Нехай H – ідеал кільця F[[X]]. Якщо $H = \{O_F\}$, то $H = O_F\ F[[X]]$. Припустимо, що H містить ненульові ступеневі ряди. Якщо $\sum_{n \in N0} a_n X^n \in H$ і $a_0 \neq O_F$, то як ми бачили вище,

$$\sum_{n \in N0} a_n X^n \in U(F[[X]])$$

і із пропозиції 8.6 випливає, що $H = F[[X]]$, тобто $H = e\ F[[X]]$.

Розглянемо випадок, коли коефіцієнти при X^0 кожного ряду з H дорівнюють O_F. Це означає, що кожен елемент f(X) із H має вигляд $X\ g(X)$ для деякого $g(X) \in F[[X]]$. Нехай m – натуральне число, яке має такі властивості: якщо $f(X) \in H$, то $f(X) = X^m\ g(X)$ для деякого ряду g(X) і існує такий ряд $d(X) = \sum_{n \in N0} b_n X^n \in H$, що $b_0 = b_1 = \ldots = b_{m-1} = O_F$, але $b_m \neq O_F$. Звідси отримуємо, що $d(X) = X^m\ d_1(X)$, причому коефіцієнт при X^0 ряду $d_1(X)$ дорівнює b_1. Це означає, що $d_1(X) \in U(F[[X]])$.

Нехай $d_2(X)$ - такий ряд, що $d_1(X)\ d_2(X) = e$. Оскільки H – ідеал кільця F[[X]], то $d(X)\ d_2(X) = X^m\ d_1(X)\ d_2(X) = X^m\ e = X^m \in H$. Тоді $X^m\ F[[X]] \leq H$. З іншого боку, кожен ряд f(X) з H має вигляд $f(X) = X^m\ g(X)$, що доводить рівність $H = X^m\ F[[X]]$.

Повернемось до розгляду кільця R[y], з якого ми починали цей розділ. Нехай K – комутативне кільце, R – його унітарне цілісне підкільце, y ∈ K. Якщо f(X) ∈ R[X], f(X) = $a_0 + a_1X + \ldots + a_nX^n$, $a_0, a_1, \ldots, a_n \in R$, то позначимо $f(y) = a_0 + a_1y + \ldots + a_ny^n$. Очевидно, f(y) ∈ R[y].

Будемо говорити, що y – корінь полінома f(X), якщо f(y) = O_R.

Корінь може як належати кільцю R, так і знаходитись поза ним. Наприклад, поліном $X^2 - 2$ не має коренів у полі раціональних чисел, його корені належать полю R дійсних чисел; поліном $X^2 + 1$ не має коренів навіть у полі R.

9.4. Пропозиція. *Нехай K – комутативне кільце, R – його унітарне цілісне підкільце, y ∈ K. Відображення η: R[X] ⟶ K, що визначається за правилом η(f(X)) = f(y), f(X) ∈ R[X], є гомоморфізмом.*

Це твердження доводиться безпосередньою перевіркою.

Розглянемо тепер Im η. З пропозиції 8.10 випливає, що Im η – підкільце K. Якщо f(X) = a – поліном нульового ступеня, то η(f(X)) = a. Звідси випливає включення R ≤ Im η. З іншого боку,

$$\text{Im } \eta = \{\eta(f(X)) \mid f(X) \in R[X]\} = \{f(y) \mid f(X) \in R[X]\},$$

а вже зазначалося, що f(y) ∈ R[y] для будь - якого f(X) ∈ R[X]. Це означає, що Im η = R[y] = {f(y) ∣ f(X) ∈ R[X]}. Розглянемо тепер Ker η.

Будемо говорити, що елемент y трансцендентний над кільцем R, якщо y не є коренем жодного ненульового полінома з коефіцієнтами з R.

Іншими словами, якщо елемент y трансцендентний над R, то Ker η= $\{O_F\}$. За теоремою 8.11 η – мономорфізм, тобто R[X] ≅ R[y].

Нехай Ker η≠ $\{O_F\}$. З пропозиції 8.10 випливає, що Ker η – ідеал кільця поліномів R[X]. Ker η складається з тих і тільки тих поліномів f(X), для якихf(y) = O_F. Інакше кажучи, Ker η складається з усіх поліномів, що мають y своїм коренем.

Будемо говорити, що елемент y алгебраїчний над R, якщо в кільці поліномів R[X] знайдеться ненульовий поліном, що має у своїм коренем.

Якщо R – поле, то з наслідку 9.2 випливає, що Ker η = h(X) R[X] для деякого полінома h(X) $\in$ R[X]. Якщо r(X) – інший поліном з властивістю Ker η = r(X) R[X], то маємо

$$h(X) = r(X)\ u(X) \quad i \quad r(X) = h(X)\ v(X)$$

для деяких u(X), v(X) $\in$ R[X]. Звідси одержуємо

$$h(X) = r(X)\ u(X) = h(X)\ v(X)\ u(X),$$

тобто v(X) u(X) = e. Оскільки, як ми бачили вище, U(R[X]) = U(R), то це означає, що v(X) = c $\in$ U(R), u(X) = c^{-1}. Отже, r(X) = c h(X), де c $\in$ U(R).

Серед усіх поліномів, що породжують ідеал Ker η, виберемо той, старший коефіцієнт якого дорівнює одиничному елементу. Цей поліном позначимо через $m_y(X)$ і будемо називати *мінімальним поліномом для елемента a над полем R.* Із теореми 8.13 отримуємо ізоморфізм

$$R[y\} = Im\ \eta \cong R[X]/Ker\ \eta,$$

а якщо ще R – поле, то Ker η = $m_y(X)$ R[X].

Нехай F – поле, y $\in$ F. Із теореми 9.1 одержуємо розклад f(X) = q(X)(X – y) + r(X), де або r(X) = O_F, або deg r(X) < deg (X - y) = 1. Отже, в будь-якому випадку r(X) $\in$ F, тобто r(X) = b $\in$ F. З пропозиції 17.4 випливає, що f(y) = q(y) (y – y) + b = b. Отже, f(X) = (X – y) q(X) + f(y).

Будемо говорити, що поліном f(X) ділиться на поліном g(X), якщо f(X) = g(X)q(X) для деякого полінома q(X).

Отже, твердження доведене.

9.5. Пропозиція. *Нехай F – поле, $f(X) \in F[X]$. Елемент $y \in F$ у тому і тільки у тому разі є коренем полінома $f(X)$, коли $f(X)$ ділиться на $(X - y)$.*

9.6. Наслідок. *Нехай F – поле, $f(X) \in F[X]$. Тоді $f(X) = (X - c_1)^{k}1 \ldots (X - c_m)^{k}m\, q(X)$, де поліном $q(X)$ не має коренів у полі F, $c_j \neq c_t$ при $j \neq t$, $k_1, \ldots, k_m \in N$.*

9.7. Наслідок. *Нехай F – поле, $f(X) \in F[X]$, $M = \{ a \in F \mid a$ – корінь $f(X)\}$. Тоді $\mid M \mid \leq \deg f(X)$.*

Поле F називається алгебраїчно замкнутим, якщо будь-який поліном $f(X) \in F[X]$ має в полі F хоча б один корінь.

9.8. Наслідок. *Поле F у тому і тільки тому разі є алгебраїчно замкнутим, коли для будь-якого полінома $f(X) \in F[X]$ має місце рівність*

$$f(X) = (X - c_1)^{k}1 \ldots (X - c_m)^{k}m,$$

де $c_1, \ldots, c_m \in F$, $c_j \neq c_t$ при $j \neq t$, $k_1, \ldots, k_m \in N$.

РОЗДІЛ 10
ПОЛЕ КОМПЛЕКСНИХ ЧИСЕЛ

Іншим важливим прикладом, що потребує окремого розгляду, є поле комплексних чисел. Комплексні числа виникли в зв'язку із задачами знаходження коренів поліномів з дійсними коефіцієнтами. Прикладом найпростішого полінома, що не має коренів у полі дійсних чисел, буде $X^2 + 1$. Це приводить до необхідності розглядати вираз $\sqrt{-1}$. Перші спроби введення комплексних чисел були здійснені італійськими математиками. Уже Кардано формально оперував (із будь-якими застереженнями) з виразами, що містять квадратні корені з від'ємних величин, а його послідовник Бомбеллі виклав числення з комплексними числами у формі, що дуже близька до сучасної. Проте визнання комплексних чисел відбувалося досить повільно. Історія математики бачила довгу боротьбу між прихильниками та супротивниками комплексних чисел. Основне питання тут було пов'язане з пошуками природного пояснення для величини $\sqrt{-1}$. Лише Гаусу зміг дати переконливе для багатьох обгрунтування. Подальший розвиток маремематики, а також використання у механіці і фізиці зробили поняття комплексного числа "законним" і звичним.

Нехай P, F – поля. *Якщо F – підполе поля P, то будемо говорити, що P – розширення поля F.*

Нехай F – підполе поля P, M – деяка підмножина P. Розглянемо сукупність *M* тих підполів, які містять у собі як підполе F, так і підмножину M. Нехай $F(M) = \cap$ M. За наслідком 7.12 F(M) – підполе. З його означення випливає, що F(M) – найменше підполе, яке містить у собі підполе F і підмножину M.

Підполе F(M) називається розширенням поля F, одержаним приєднанням до F підмножини M.

Найпростійший випадок, коли M складається з єдиного елемента y. У цьому випадку замість F(M) будемо писати F(y).

Розширення, одержане приєднанням до F одного елемента y, називається простим розширенням.

Очевидно, $F[M] \leq F(M)$, зокрема $F[a] \leq F(a)$.

У термінах розширень можна зараз сформулювати визначення поля комплексних чисел.

Полем комплексних чисел називається розширення поля дійсних чисел, одержане приєднанням до нього кореня полінома $X^2 + 1$.

Поле комплексних чисел будемо позначати C, а для кореня поліному $X^2 + 1$ будемо використовувати його звичайне позначення i. У таких позначеннях $C = R(\mathbf{i})$. Далі потрібно звернути увагу на таку обставину. Коли давали визначення розширення поля, що отримане приєднанням деякої множини елементів (зокрема, одного елемента), вважали, що і поле, яке розширюється, і множина, що приєднується, містяться в одному більш широкому полі. Визначаючи вище поле комплексних чисел, не знаємо (поки що), чи існує поле, підполем у якому буде поле дійсних чисел, і яке містить корінь полінома $X^2 + 1$. Тому перше питання, яке тут виникає, це питання про існування поля комплексних чисел. Але спочатку припустимо, що таке поле існує і знайдемо вигляд його елементів. Отже, нехай F – поле, підполем якого буде R, і нехай F містить елемент i – корінь полінома $X^2 + 1$. Якщо K – підполе F, яке містить у собі R та елемент i, то разом з кожним дійсним числом x підполе K містить і елемент xi. Отже, якщо $x, y \in R$, то $x + yi \in K$. Таким чином, підполе C, яке є перетином всіх підполів K, що містять у собі C та елемент i, містить всі елементи $x + yi$. Нехай $S = \{ x + yi \mid x, y \in R\}$, тоді маємо включення $S \subseteq C$. Покажемо, що підмножина S є підполем. Скористаємось для цього критерієм підполя. Нехай $\alpha = x + yi$, $\beta = u + vi$, тоді маємо

$$\alpha - \beta = (x - u) + (y - v)i \in S, \alpha\beta = (x + yi)(u + vi) = (xu - yv) + i\,(yu + xv) \in S.$$

З означення підполя випливає, що одиничний елемент підполя R, тобто число 1, буде одиничним елементом усього поля F, але ж $1 = 1 + 0i \in S$. Аналогічно $i = 0 + 1i \in S$. Нарешті, нехай $0 \neq \alpha = x + yi \in S$. Зокрема, хоча б одне з чисел x або y – ненульове, а тому $x^2 + y^2 \neq 0$.

Маємо

$$(x + yi)(x - yi) = x^2 + y^2.$$

Це підказує таку рівність

$$(x + yi)^{-1} = (x/(x^2 + y^2) + (-y)/(x^2 + y^2)i) \in S.$$

Таким чином, усі умови теореми 7.10 виконуються для S, отже, S – підполе поля F. За самим визначенням S містить у собі всі дійсні числа, а також елемент i. Це означає, що S містить у собі C = R (i). З іншого боку, ми довели вже включення $S \subseteq C$, яке доводить рівність $C = \{x + yi \mid x, y \in R\}$.

Припустимо, що $x + yi = u + vi$ для деяких $x, y, u, v \in R$. Тоді $x - u = (v - y)i$. Оскільки $i \notin R$, то ця рівність можлива лише в тому випадку, коли $v - y = 0$, $x - u = 0$, тобто $x = u$, $y = v$.

Таким чином, будь-яке комплексне число можна записати однозначно у вигляді $x + yi$, де x, y – дійсні числа. Доданок x називається його *дійсною частиною*, доданок yi – його *уявною частиною*.

Скористаємось тепер усім цим, щоб довести існування поля комплексних чисел. Почнемо зі звичайної трактовки комплексних чисел як точок звичайної площини. Це так звана геометрична модель поля комплексних чисел.

Отже, комплексною площиною будемо називати множину $R \times R$ усіх точок $\alpha = (x, y)$, $x, y \in R$, на якій операції додавання та множення задаються такими рівностями: якщо $\beta = (u, v)$, то

$$\alpha + \beta = (x + u, y + v), \alpha\beta = (xu - yv, yu + xv).$$

Як бачимо, ці формули підказані нам попереднім розглядом. Розглянемо тепер властивості цих операцій.

Додавання комутативне:

$$\alpha + \beta = (x + u, y + v) = (u + x, v + y) = \beta + \alpha.$$

Додавання ассоціативне:

нехай $\gamma = (z, w)$, тоді

$$(\alpha + \beta) + \gamma = ((x + u) + z, (y + v) + w) = (x + (u + z), y + (v + w)) = \alpha + (\beta + \gamma).$$

Відносно операції додавання існує нульовий елемент, ним буде пара (0, 0):

$$(x, y) + (0, 0) = (x + 0, y + 0) = (x, y).$$

Для кожної пари (x, y) існує протилежний елемент, ним буде (-x, -y) :

$$(x, y) + (-x, -y) = (x + (-x), y + (-y)) = (0, 0).$$

Множення комутативне:

$\alpha\beta = (xu - yv, yu + xv)$, $\beta\alpha = (ux - vy, uy + vx)$, тобто $\alpha\beta = \beta\alpha$.

Множення ассоціативне:

$$(\alpha\beta)\gamma = (xu - yv, yu + xv)\,(z, w) =$$

$$= ((xu - yv)z - (yu + xv)w, (yu + xv)z + (xu - yv)w) =$$

$$= (xuz - yvz - yuw - xvw, yuz + xvz + xuw - yvw);.$$

$$\alpha(\beta\gamma) = (x, y)\,(uz - vw, vz + uw) = (x(uz - vw) - y(vz + uw),$$

$$y(uz - vw) + x(vz + uw)) =$$

$$(xuz - xvw - yvz - yuw, yuz - yvw + xvz + xuw),$$

отже, $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.

Відносно операції множення існує одиничний елемент, ним буде пара (1, 0):

$$(x, y)\,(1, 0) = (x1 - y\,0, y\,1 + x\,0) = (x, y).$$

Для кожної пари $(x, y) \neq (0, 0)$ існує обернений елемент, ним буде $(x/(x^2 + y^2), -y/(x^2 + y^2))$:

$$(x, y)\,(x/(x^2 + y^2), -y/(x^2 + y^2))) =$$
$$((x^2 + y^2)/(x^2 + y^2), (yx - xy)/(x^2 + y^2))) = (1, 0).$$

Додавання та множення пов'язані законом дистрибутивності:

$$(\alpha + \beta)\gamma = ((x, y) + (u, v))(z, w) = (x + u, y + v)(z, w) =$$

$$= ((x + u)z - (y + v)w), (y + v)z + (x + u)w)) =$$

$$= (xz + uz - yw - vw, yz + vz + xw + uw);$$

$$\alpha\gamma + \beta\gamma = (x, y)(z, w) + (u, v)(z, w) =$$

$$= (xz - yw, yz + xw) + (uz - vw, vz + uw) =$$

$$= (xz - yw + uz - vw, yz + xw + vz + uw),$$

отже, $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$.

Отже, відносно введених вище операцій додавання та множення $R \times R$ стає полем.

Підмножина $R \times \{0\} = \{ (x, 0) \mid x \in R \}$ є підполем. Дійсно

$$(x, 0) - (y, 0) = (x - y, 0) \in R \times \{0\};$$

$$(x, 0)(y, 0) = (xy - 00, 0\,y + x\,0) = (xy, 0) \in R \times \{0\};$$

$$(1, 0) \in R \times \{0\};$$

якщо $(x, 0) \neq (0, 0)$, то $x \neq 0$ і $(x, y)^{-1} = (x^{-1}, 0) \in R \times \{0\}$.

Розглянемо тепер відображення $f: R \longrightarrow R \times \{0\}$, що визначається за правилом $f(x) = (x, 0)$. Маємо

$$f(x + y) = (x + y, 0) = (x, 0) + (y, 0) = f(x) + f(y);$$

$$f(xy) = (xy, 0) = (x, 0)(y, 0) = f(x)\,f(y).$$

Очевидно, що f є бієктивним, тому f – ізоморфізм. Інакше кажучи, підполе $R \times \{0\}$ ізоморфно до поля R дійсних чисел. Це дає нам можливість ототожнювати пару $(x, 0)$ з дійсним числом x.

Маємо очевидну рівність

$$(x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0).$$

Також $(0, 1)^2 = (-1, 0)$. Отже, якщо покласти $i = (0, 1)$, то $i^2 = (-1, 0)$, так що, ототожнюючи пару $(-1, 0)$ з дійсним числом -1, отримаємо

$$(x, y) = (x, 0) + (0, 1)(y, 0) = x + yi.$$

Отже, наше побудоване поле містить у собі як підполе поле R дійсних чисел (точніше його ізоморфну копію), а також містить корінь полінома $X^2 + 1$. Отже, воно містить у собі підполе $R(i) = R$. Більш того, з попередніх розглядів видно, що воно збігається з C.

Наведемо іншу модель поля комплексних чисел, так звану *матричну модель*. Розглянемо підмножину P кільця $M_2(R)$, яка складається з матриць, що мають такий вигляд

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

Рівності

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix} + \begin{pmatrix} u & v \\ -v & u \end{pmatrix} = \begin{pmatrix} x + u & y + v \\ -y - v & x + u \end{pmatrix},$$

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix} + \begin{pmatrix} u & v \\ -v & u \end{pmatrix} = \begin{pmatrix} xu - yv & xv + yu \\ -xv - yu & xu - yv \end{pmatrix},$$

показують, що підмножина P є сталою відносно операцій додавання та множення матриць. Додавання елементів множини P є кому-тативним та асоціативним, оскільки воно комутативне та асоціативне для всіх елементів матричного кільця $M_2(R)$, нульова матриця належить до P, зокрема вона буде нульовим елементом для P. Нарешті,

$$-\begin{pmatrix} x & y \\ -y & x \end{pmatrix} = \begin{pmatrix} -x & -y \\ y & -x \end{pmatrix},$$

і це показує, що відносно операції додавання P буде абелевою групою. Операції додавання та множення пов'язані законами дистрибутивності, оскільки вони мають місце для всіх матриць з $M_2(R)$, з тієї ж причини множення в P є асоціативним. Далі маємо

$$\begin{pmatrix} x & -y \\ -y & x \end{pmatrix}\begin{pmatrix} u & v \\ -v & u \end{pmatrix} = \begin{pmatrix} xu - yv & xv + yu \\ -xv - yu & xu - yv \end{pmatrix},$$

$$\begin{pmatrix} u & v \\ -v & u \end{pmatrix}\begin{pmatrix} x & y \\ -y & x \end{pmatrix} = \begin{pmatrix} ux - vy & vx + uy \\ -vx - uy & ux - vy \end{pmatrix},$$

а це доводить комутативність множення елементів P. Нехай тепер

$$Y = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

довільний ненульовий елемент P. Тоді хоча б одне з чисел x або y – ненульове. Маємо **det**(Y) = $x^2 + y^2 \neq 0$. Інакше кажучи, ненульові елементи з P будуть невиродженими матрицями, і тому кожна з них має обернену. Застосовуючи правило знаходження оберненої матриці, отримаємо

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix}^{1} \qquad \begin{pmatrix} u & v \\ -v & u \end{pmatrix}$$

де $u = x/(x^2 + y^2)$, $v = (-y)/(x^2 + y^2)$. Це показує, що $Y^{-1} \in P$. Отже, P є полем.

Розглянемо відображення f: C $\longrightarrow$ P, що визначено за правилом

$$f(x + yi) = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

Очевидно воно бієктивне, а також, якщо $\alpha = x + y\mathbf{i}$, $\beta = u + v\mathbf{i}$, то

$$f(\alpha + \beta) = f((x + u) + i(y + v)) = \begin{pmatrix} x + u & y + v \\ -y - v & x + u \end{pmatrix},$$

$$f(\alpha) + f(\beta) = f(x + iy) + f(u + iv) = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}\begin{pmatrix} u & v \\ -v & u \end{pmatrix},$$

а отже, $f(\alpha + \beta) = f(\alpha) + f(\beta)$.

Далі

$$f(\alpha\beta) = f((xu - yv) + i\,(yu + xv)) = \begin{pmatrix} xu - yv & xv + yu \\ -xv - yu & xu - yv \end{pmatrix} = \begin{pmatrix} x & y & u & v \\ -y & x & -v & u \end{pmatrix} =$$

$$= f(x + yi)\ f(u + vi) = f(\alpha)\ f(\beta).$$

Усі ці рівності показують, що відображення f зберігає операції додавання та множення, а значить буде ізоморфізмом. Отже поле C комплексних чисел ізоморфно до побудованого нами поля P. При цьому f(C) = RE – множина всіх скалярних матриць, яка очевидно буде підполем P,

$$f(i) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathbf{I}.$$

Отже, можемо ототожнити скалярну матрицю xE з дійсним числом x. Також

$$\mathbf{I}^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -E,$$

зокрема, поле P містить корінь матричного полінома $EX^2 + E$. Тепер знову можемо отримати звичайну форму комплексного числа в його матричній моделі:

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} + \begin{pmatrix} 0 & y \\ -y & 0 \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} + \begin{pmatrix} y & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = xE + (yE)\mathbf{I}.$$

Абсолютним значенням комплексного числа x + yi називають невід'ємне дійсне число

$$\sqrt{x^2 + y^2} = \overline{|x + yi|}$$

10.1. Пропозиція. *Відображення* $\Delta: x + yi \mapsto x - yi$ *є автоморфізмом поля* C. *При цьому* $\Delta^2 = \varepsilon_C$, $\Delta(a) = a$ *для кожного* $a \in C$.

Д о в е д е н н я. Дійсно,

$$\Delta(x + yi + u + vi) = \Delta((x + u) + (y + v)i) = (x + y) - (y + v)i =$$
$$= (x - yi) + (u - vi) = \Delta(x + yi) + \Delta(u + vi);$$

$$\Delta((x + yi)(u + vi)) = \Delta((xu - yv) + (xv + yu)i) = (xu - yv) - (xv + yu)i =$$
$$= (x - yi)(u - vi) = \Delta(x + yi)\ \Delta(u + vi);$$

Ін'єктивність і сюр'єктивність відображення Δ очевидні. Нарешті,

$$\Delta^2(x + yi) = \Delta(\Delta(x + yi)) = \Delta(x - yi) = x + yi.$$

Із означення Δ випливає, що $\Delta(a) = a$ для будь-якого $a \in R$.

Числа $x + yi$ і $x - yi$ називають *комплексно спряженими*. У подальшому будемо писати

$$\Delta(x + yi) = \overline{x + yi}.$$

Очевидно, $(x + yi)(x - yi) = x^2 + y^2 = |x + yi|^2$.

Нехай $r = |x + yi|$. Запишемо комплексне число $x + yi$ у вигляді

$$x + yi = r((x/r) + (y/r)i).$$

Тоді

$$(x^2/r^2) + (y^2/r^2) = (x^2 + y^2)/(x^2 + y^2) = 1.$$

Тому знайдеться такий кут ϕ, що $\cos\phi = x/r$, $\sin\phi = y/r$. Цей кут α називається *аргументом комплексного числа* $x + yi$ *і позначається через* $arg(x + yi)$. Зрозуміло, що кут ϕ визначається з точністю до доданка, кратного 360°. Отже, маємо $x + yi = r(\cos\phi + i\sin\phi)$. *Ця форма називається тригонометричною формою запису комплексного числа.*

Тригонометрична форма запису виявляється досить зручною при діях, пов'язаних з множенням. Дійсно, нехай

$$x + yi = r(\cos\phi + i\sin\phi),$$

$$u + vi = q(\cos\psi + i\sin\psi),$$

тоді

$$(x + yi)(u + vi) = r(\cos\phi + i\sin\phi)\, q(\cos\psi + i\sin\psi) =$$
$$= rq((\cos\phi\cos\psi - \sin\phi\sin\psi) + i(\sin\phi\cos\psi + \cos\phi\sin\psi)) =$$
$$= rq(\cos(\phi + \psi) + i\sin(\phi + \psi)).$$

Таким чином, виходять досить зручні формули

$$|(x + yi)(u + vi)| = |x + yi|u + vi|,$$

$$\arg((x + yi)(u + vi)) = \arg(x + yi) + \arg(u + vi).$$

Очевидно, ці формули можна розповсюдити на будь-яку скінченну кількість множників. Зокрема,

$$(r(\cos\phi + i\sin\phi))^k = r^k(\cos k\phi + i\sin k\phi))^k,\ k \in N,$$

або в іншому записі

$$|(x + yi)^k| = |x + yi|^k;\ \arg((x + yi)^k) = k\arg(x + yi).$$

Ці формули називають *формулами Муавра*. Використовуючи їх, одержуємо можливість здійснити обернену операцію – добування кореня.

10.2. Пропозиція. *Для будь-якого комплексного числа*

$$x + yi = r(\cos\phi + i\sin\phi)$$

існує рівно k різних комплексних чисел

$$u + vi = q(\cos\xi + i\sin\xi)$$

із властивістю $(u + vi)^k = x + yi$ *(коренів k-го степеня з числа* $x + vi$*). Їх дає формула* $q = \sqrt[k]{r}$ *(тут* $\sqrt[k]{r}$ *означає арифметичне значення кореня k-го степеня з додатного дійсного числа r),*

$$\xi = (\phi + 2\pi t)/k, \text{ де } 0 \le t < k.$$

Д о в е д е н н я. Дійсно, з формули Муавра отримуємо $q^k = r$, $k\xi = \phi + 2\pi s$, $s \in N$. Звідси $q = \sqrt[k]{r}$, $\xi = (\phi + 2\pi s)/k$. Число s можна подати у вигляді $s = km + t$, де $0 \le t \le k - 1$. Тому

$$(\phi + 2\pi s)/k = (\phi + 2\pi(km + t))/k = (\phi + 2\pi t)/k + 2\pi m,$$

$$\cos((\phi + 2\pi s)/k) = \cos((\phi + 2\pi t)/k),$$

$$\sin((\phi + 2\pi s)/k) = \sin((\phi + 2\pi t)/k),$$

де $0 \le t \le k - 1$. Якщо $\cos\chi = \cos\psi$ та водночас $\sin\chi = \sin\psi$, то $\chi - \psi = 2\pi j$ для деякого цілого j. Це означає, що всі числа

$$\sqrt[k]{r}\,(\cos((\phi + 2\pi t)/k) + i\sin((\phi + 2\pi t)/k)),$$

де $0 \le t \le k - 1$, будуть різними, тобто дійсно отримуємо рівно k різних коренів k-го степеня з числа $x + yi$.

10.3. Наслідок. *Корені k-го степеня з одиниці знаходяться за формулою*

$$\varepsilon_k = \cos(2\pi t/k) + i\sin(2\pi t/k), \text{ де } 0 \le t < k.$$

10.4. Наслідок. *Множина C_k усіх коренів k-го степеня є циклічною підгрупою U(C), причому* $|C_k| = k$.

Дійсно, якщо $\alpha^k = \beta^k = 1$, то очевидно $(\alpha\beta^{-1})^k = 1$, із наслідку 2.5 отримаємо, що C_k є підгрупою U(C). Далі, з формули Муавра випливає рівність $\varepsilon_k = \varepsilon_1^k$, а це означає, що $C_k = \langle \varepsilon_1 \rangle$.

Корінь e k-го степеня з одиниці називається примітивним або первісним, якщо він не є коренем з одиниці ступеня d для $d < k$.

Інакше кажучи, k є найменшим натуральним числом, для якого $e^k = 1$, тобто порядок елементу *e* у групі U(C) дорівнює k. Як зазначалося у розділі 10, $|\langle e \rangle| = |e|$, а тому $\langle e \rangle = C_k$. Отже, корінь *e* k-го степеня з одиниці тоді і тільки тоді буде примітивним, коли $\langle e \rangle = C_k$.

10.5. Наслідок. *Корінь e k-го степеня з одиниці тоді і тільки тоді буде примітивним, коли $e = \varepsilon_m$, де НСД(m, k) = 1.*

Використовуючи поле комплексних чисел, побудуємо два важливих алгебраїчних об'єкта.

Квазіциклічна група (p-група Прюфера)

Нехай p – просте число, позначимо

$$C_{p^\infty} = \{x \in C \mid x^{p^k} = 1 \text{ для деякого } k \in N_0\}.$$

Покажемо, що C_{p^∞} – підгрупа U(C). Нехай x, y $\in C_{p^\infty}$, k, t – такі числа, що $x^{p^m} = 1, y^{p^k} = 1$. Покладемо m = max {k, t}. Тоді маємо $(x/y^{p^t}) = 1$. З наслідку 10.5 випливає, що C_{p^∞} – підгрупа U(C).

Нехай x $\in C_{p^\infty}$, k – найменше число, для якого $x^{p^k} = 1$. Інакше кажучи, $|x| = p^k$. Але тоді x – примітивний корінь з одиниці ступеня p^k, тобто $C_p^k = \langle x \rangle$. Очевидно,

$$C_p \le C_p^2 \le \ldots \le C_p^k \ldots, \bigcup_{k \in N} C_p^k = C_{p^\infty}.$$

Нехай H – підгрупа C_{p^∞}. Можливі дві ситуації:

1) Порядки елементів підгрупи H обмежені в сукупності.

2) Порядки елементів H не обмежені.

Розглянемо перший випадок. Тоді знайдеться таке число k, що $y^{p^k} = 1$, для кожного y $\in$ H, і H містить такий елемент x, що $|x| = p^k$. Це означає, з одного боку, що $H \le C_p^k$, а з іншого, що $\langle x \rangle = C_p^k$. Інакше кажучи, $H = C_p^k$.

Розглянемо другий випадок. Тоді для будь-якого числа k $\in$ N знайдеться такий елемент z $\in$ H, що $|z| = p^m$, де m ≥ k. Із рівності $|z| = p^m$ випливає, що $\langle z \rangle = C_{p^m}$, а оскільки m ≥ k, то

$$C_p^k \le C_p^m = \langle z \rangle \le H.$$

Отже, для будь-якого k $\in$ N підгрупа H містить у собі C_p^k. Це означає, що $H = C_{p^\infty}$. *Таким чином, будь-яка власна підгрупа C_{p^∞} є скінченою і циклічною.*

Тому групу C_{p^∞} називають квазіциклічною. Інша її назва – *p*-група Прюфера. У 1938 р. О.Ю. Шмідт поставив питання про опис нескінченних груп, усі власні підгрупи яких скінченні [3]. Довго не вдавалося побудувати відмінних від C_{p^∞} груп такого типу і проблема О.Ю. Шмідта трансформувалася в проблему

існування відмінних від C_{p^∞} груп, усі власні підгрупи яких циклічні. Порівняно недавно серії таких груп були побудовані О.Ю. Ольшанським [4, §28].

Тіло кватерніонів. Розглянемо кільце матриць $M_2(C)$, а в ньому підмножину H, що складається з матриць, які мають такий вигляд:

$$\begin{pmatrix} x & y \\ -\Delta(y) & \Delta(x) \end{pmatrix}, \text{ де } x, y \in C,$$

Покажемо, що H – підкільце $M_2(C)$. Дійсно,

$$\begin{pmatrix} x & y \\ -\Delta(y) & \Delta(x) \end{pmatrix} - \begin{pmatrix} u & v \\ -\Delta(v) & \Delta(u) \end{pmatrix} = \begin{pmatrix} x-u & y-v \\ -(\Delta(y)-\Delta(v)) & \Delta(x)-\Delta(u) \end{pmatrix} =$$

$$= \begin{pmatrix} x-u & y-v \\ -(\Delta(y-v)) & \Delta(x-u) \end{pmatrix} \in H.$$

$$\begin{pmatrix} x & y \\ -\Delta(y) & \Delta(x) \end{pmatrix} \begin{pmatrix} u & v \\ -\Delta(v) & \Delta(u) \end{pmatrix} = \begin{pmatrix} xu - y\Delta(v) & xv + y\Delta(u) \\ -\Delta(y)u - \Delta(x)\Delta(v)) & -\Delta(y)v + \Delta(x)\Delta(u) \end{pmatrix}$$

$$= \begin{pmatrix} xu - y\Delta(v) & xv + y\Delta(u) \\ -(\Delta(y)\Delta(\Delta(u)) + \Delta(x)\Delta(v)) & -\Delta(y)\Delta(\Delta(v)) + \Delta(x)\Delta(u) \end{pmatrix}$$

$$= \begin{pmatrix} xu - y\Delta(v) & xv + y\Delta(u) \\ -(\Delta(y\Delta(u)) + \Delta(xv)) & -\Delta(y\Delta(v)) + \Delta(xu) \end{pmatrix}$$

$$= \begin{pmatrix} xu - y\Delta(v) & xv + y\Delta(u) \\ -\Delta(y\Delta(u) + xv)) & \Delta(xu - y\Delta(v))) \end{pmatrix} \in H.$$

З теореми 7.4 одержуємо, що H – підкільце $M_2(C)$.

Розглянемо довільний ненульовий елемент H:

$$\alpha = \begin{pmatrix} x & \\ -\Delta(y) & \end{pmatrix} \neq \begin{pmatrix} y & 0\ 0 \\ \Delta(x) & 0\ 0 \end{pmatrix}.$$

Маємо $\det(\alpha) = x\,\Delta(x) + y\,\Delta(y) = |x|^2 + |y|^2 \neq 0$, зокрема, матриця α не вироджена, і отже, має обернену. Застосуємо для її знаходження відому теорему, і отримаємо

$$\alpha^{-1} = \begin{pmatrix} \Delta(x)/\det(\alpha) & -y/\det(\alpha) \\ \Delta(y)/\det(\alpha) & x/\det(\alpha) \end{pmatrix} \begin{pmatrix} \Delta(x)/\det(\alpha) & -y/\det(\alpha) \\ -\Delta(-y/\det(\alpha)) & \Delta(\Delta(x)/\det(\alpha)) \end{pmatrix} \in H.$$

Отже, H – тіло. Отримаємо звичайну форму запису для елементів H. Нехай $x = a + bi$, $y = c + di$, тоді

$$\begin{pmatrix} x & y \\ -\Delta(y) & \Delta(x) \end{pmatrix} = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} =$$

$$= \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \quad \begin{pmatrix} bi & 0 \\ 0 & -bi \end{pmatrix} \quad \begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & di \\ -di & 0 \end{pmatrix} =$$

$$= a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + c \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Нехай

$$\boldsymbol{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \boldsymbol{I} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \ \boldsymbol{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ \boldsymbol{K} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Шляхом безпосереднього підрахунку для матриць *E, I, J, K* одержимо таку таблицю множення:

	E	***I***	***J***	***K***
E	***E***	***I***	***J***	***K***
I	***I***	***-E***	***K***	***-J***
J	***J***	***-K***	***-E***	***I***
K	***K***	***J***	***-I***	***-E***

Зокрема, IJ = K, JI = - K, тобто множення некомутативне. Отже, H – некомутативне тіло. H *називають тілом кватерніонів, а його елементи – кватерніонами, (точніше – дійсними кватерніонами).*

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Федоров Ю.Г. О бесконечных группах, все нетривиальные подгруппы которых имеют конечный индекс / Ю.Г. Федоров // Успехи мат. наук. – 1951. – 6, № 1. – С. 187–189.

2. Dedekind R. Über Gruppen, deren sammtliche Teiler Normalteiler sind / R. Dedekind // Math. Annalen. – 1897. – 48. – S. 548–561.

3. Черников С. Н. О проблеме Шмидта / С. Н. Черников // Укр. мат. журн. – 1971. – 23, № 5.– С. 598–603.

4. Ольшанский А. Ю. Геометрия определяющих соотношений в группах / А. Ю. Ольшанский. – М. : Наука, 1989.

5. Александров П. С. Введение в теорию групп / П. С. Александров .– М.: Наука, 1980.

6. Ван дер Варден Б. Л. Алгебра / Б. Л. Ван дер Варден. – М.: Наука, 1976.

7. Алгебра і теорія чисел. Практикум / С. Т. Завало, С. С. Левіщенко, В. В. Пилаєв, І. О. Рокицький. – К. : Вища школа, 1983. – Частина 1.

8. Алгебра і теорія чисел. Практикум / Завало С. Т., Левіщенко С. С, Пилаєв В. В, Рокицький І. О. – К. : Вища школа, 1986. – Частина 2.

9. Калужнин Л. А. Введение в общую алгебру. – М.: Наука, 1973.

10. Кострикин А. И. Введение в алгебру. – М. : Наука, 1977.

11. Кострикин А. И. Введение в алгебру. Основы алгебры. – М. : Наука, 1994.

12. Курош А. Г. Курс высшей алгебры / А. Г. Курош. – М. : Наука, 1971.

13. Курош А. Г. Лекции по общей алгебре / А. Г. Курош. – М. : Наука, 1973.

14. Курош А. Г. Общая алгебра / А. Г. Курош. – М. : Наука, 1974.

15. Ленг С. Алгебра / С. Ленг. – М. : Мир, 1968.

16. Мальцев А. И. Основы линейной алгебры / А. И. Мальцев. – М. : Наука, 1970.

17. Скорняков Л. А. Элементы алгебры / Л. А. Скорняков. – М. : Наука, 1980.

18. Скорняков Л. А. Элементы общей алгебры / Л. А. Скорняков. – М. : Наука, 1983.

19. Фаддеев Д. К. Лекции по алгебре / Д. К. Фаддеев. – М. : Наука, 1984.

20. Фаддеев Д. К. Сборник задач по высшей алгебре / Д. К. Фаддеев, И. С. Соминский. – М. : Наука, 1974.

21. Шафаревич И. Р. Основные понятия алгебры. Итоги науки и техники / И. Р. Шафаревич. – М. : ВИНИТИ, 1986. – Том 11. – (Серия «Современные проблемы математики. Фундаментальные направления»).

СЕРІЯ «НА ДОПОМОГУ СТУДЕНТУ УДФСУ»

Семко Микола Миколайович,
Ярова Оксана Анатоліївна,
Скасків Лілія Василівна

ОСНОВНІ ПОНЯТТЯ СУЧАСНОЇ АЛГЕБРИ

Навчальний посібник

Відповідальний за випуск *Д. Ф. Салахова*

Відповідальний редактор *І. Є. Врублевська*
Редактори *Н. І. Грицюк,*
Л. Б. Дьомена

Форматування та
комп'ютерна верстка *Д. П. Завальніцька*

Здано до друку 14.12.20. Формат 60×84/14.
Папір офсетний № 1. Гарнітура «Times New Roman».
Ум. друк. арк. 7.3.
Наклад 300 прим. Замовлення № 344

Підготовлено до друку Видавничо-поліграфічним центром
Університету ДФС України
08205, вул. Університетська, 31, м. Ірпінь, Київська область,
Україна

Свідоцтво про внесення суб'єкта видавничої справи
до державного реєстру видавців, виготовлювачів і
розповсюджувачів видавничої продукції
Серія ДК № 5104 від 20.05.2016