С. Т. Завало
В. М. Костарчук
Б. І. Хацет

АЛГЕБРА І ТЕОРІЯ ЧИСЕЛ

2

С. Т. ЗАВАЛО
В. М. КОСТАРЧУК
Б. І. ХАЦЕТ

АЛГЕБРА І ТЕОРІЯ ЧИСЕЛ

Частина друга

*Допущено Міністерством освіти УРСР
як підручник для студентів фізико-математичних
факультетів педагогічних інститутів*

Видавниче об'єднання «Вища школа»
Головне видавництво
Київ — 1976

УДК 512 + 511 (075.8)

Алгебра и теория чисел, ч. 2. Завало С. Т., Костарчук В. Н., Хацет Б. И. Издательское объединение «Вища школа», 1976, 384 с. (на украинском языке).

Книга является продолжением первой части учебника, изданной в 1974 г. Она составлена в соответствии с действующей программой курса «Алгебра и теория чисел» для математических отделений педагогических институтов.

Изложение теории чисел органически связано с рассмотрением сведений об алгебраических структурах (группы, кольца, поля), что дает возможность осветить теорию делимости чисел и многочленов с единой точки зрения.

Кроме традиционного материала по алгебре многочленов учебник содержит элементы теории разрешимости алгебраических уравнений в квадратных радикалах с применением их к классическим конструктивным задачам геометрии. Особое внимание уделено вопросам, которые нашли отражение в новой программе и факультативных курсах по математике средней школы.

Учебник предназначен для студентов физико-математических факультетов педагогических институтов. Он будет полезен также для учителей средних школ.

Табл. 30. Ил. 7. Список лит. 45.

Редакція літератури з математики і фізики
Зав. редакцією *А. С. Макуха*

З $\frac{20203-202}{М211(04)-76}$ 28—76

ПЕРЕДМОВА

Підручник складено відповідно до діючої програми. Основним змістом його є такі питання: елементи теорії груп і кілець, основи теорії чисел, алгебра многочленів.

Означення і деякі загальні властивості основних алгебраїчних структур (група, кільце, поле) були розглянуті в першій частині підручника. Тепер відомості про групи і кільця розширюються і поглиблюються, вивчаються важливі класи цих структур і їх специфічні особливості, порівняно докладно викладається загальна теорія подільності в кільці та її конкретизація для окремих типів кілець (кільце головних ідеалів та евклідове кільце). Ця абстрактна теорія подільності широко використовується далі при викладі теорії подільності многочленів. В свою чергу сама вона будується за зразком елементарної і значною мірою відомої читачеві теорії подільності в кільці цілих чисел за допомогою належного узагальнення відповідних понять. У зв'язку з цим подільність цілих чисел розглядається у розділі II перед викладом абстрактної теорії. Разом з тим розділ II є вступом до теоретико-числової частини курсу. Згідно з програмою підручник охоплює переважно ті питання теорії чисел, які стосуються подільності цілих чисел; саме ці питання найбільш пов'язані як з шкільним курсом математики, так і з загальними алгебраїчними поняттями і методами.

Теоретико-числовий матеріал курсу охоплює такі питання: системи числа та системи числення, прості числа і їх розподіл у натуральному ряді, скінченні ланцюгові дроби (розділ II); теорія конгруенцій та застосування її до арифметики, основні числові функції (розділ IV). У цій частині книги деякі відомості відповідно до рекомендації програми подано без доведення.

Решту розділів (V—VIII) присвячено алгебрі многочленів та її застосуванням, зокрема до розв'язування алгебраїчних рівнянь та їх систем. Значна частина цього матеріалу традиційна для курсу вищої алгебри в педагогічних інститутах і не потребує коментарів. До особливостей викладу алгебри многочленів у цьому підручнику слід віднести таке.

1. Відповідно до програми кільце многочленів та поле раціональних дробів розглядаються не лише над числовими, а й над довільними абстрактними полями. У зв'язку з цим основу викладу становить не функціональне, а алгебраїчне тлумачення многочленів. Разом з тим автори постійно приділяють належну увагу питанню про умови рівносильності функціонального та алгебраїчного підходів, оскільки це питання важливе для усвідомлення зв'язку алгебри з аналізом та іншими предметами навчального плану.

2. Істотно змінено традиційний виклад теорії подільності многочленів. Для многочленів однієї змінної теорію подільності подано майже без доведень, оскільки вона є просто конкретизацією загальної теорії подільності в кільці головних ідеалів. Для многочленів же від кількох змінних дано доведення основних властивостей подільності (зокрема, теореми про єдиність розкладу на незвідні множники), що раніше в курс алгебри для педагогічних інститутів не включалося. Зауважимо, що цей матеріал важливий не лише сам по собі, а й для усвідомлення нетривіальності факту однозначності розкладу на прості множники.

3. Досить докладно розглянуто алгебраїчні розширення числових полів та застосування їх до питання про розв'язність алгебраїчних рівнянь у квадратних радикалах (розділ VIII).

При викладі другої частини курсу ми виходили з того, що читач знайомий з матеріалом першої частини. Тому, зокрема, ми використовуємо символи математичної логіки та основні правила перетворення логічних формул. При посиланнях на першу частину підручника вживається запис виду (1, § 16).

Завершуючи роботу над цим підручником з нового для пединститутів навчального предмета «Алгебра і теорія чисел», автори усвідомлюють, що ця перша спроба далеко не досконала. Вони висловлюють глибоку подяку професору М. С. Бродському, доцентам Л. М. Вивальнюку, Г. Е. Кисилевському, А. В. Нестерчуку, кандидату фізико-математичних наук М. О. Примаку, А. В. Мержеєвському за цінні поради і важливі зауваження щодо змісту підручника.

Відгуки і побажання просимо надсилати на адресу:

252054, Київ, 54, Гоголівська, 7, Головне видавництво видавничого об'єднання «Вища школа», редакція літератури з математики і фізики.

Автори

СИСТЕМИ ЛІНІЙНИХ НЕРІВНОСТЕЙ

§ 1. СИСТЕМИ ЛІНІЙНИХ НЕРІВНОСТЕЙ ТА ЇХ ГЕОМЕТРИЧНИЙ СМИСЛ

Поряд з рівняннями істотну роль у всіх розділах сучасної математики відіграють нерівності. Розв'язування багатьох задач зводиться до розв'язування нерівностей або їх систем. В цьому розділі мова йтиме про деякі питання теорії лінійних нерівностей та її застосування при розв'язуванні планово-економічних задач.

1.1. Нерівності та системи нерівностей. Якщо число a більше від числа b, то пишуть $a > b$, а якщо a менше від b, то пишуть $a < b$. Якщо число a більше від числа b або дорівнює b, то пишуть $a \geqslant b$ і говорять, що a не менше b, а якщо a менше від b або дорівнює b, то пишуть $a \leqslant b$ і говорять, що a не більше від b. Якщо відомо, що числа a і b нерівні, але невідомо, яке з них більше, а яке менше, то в цьому випадку пишуть $a \neq b$.

Співвідношення $a > b$, $a < b$, $a \geqslant b$, $a \leqslant b$, $a \neq b$, де a і b — деякі дійсні числа, називають числовими нерівностями.

Нерівності $a > b$ і $a < b$ називають *строгими*, а нерівності $a \geqslant b$ і $a \leqslant b$ — *нестрогими*.

Розв'язуючи рівняння

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

відшукують системи значень невідомих (аргументів) $x_1, x_2, \ldots, x_n$, при яких функції $F(x_1, x_2, \ldots, x_n)$ і $\Phi(x_1, x_2, \ldots, x_n)$ набувають однакових значень. Однак нерідко буває необхідно знайти значення аргументів $x_1, x_2, \ldots, x_n$, при яких значення однієї з цих функцій, наприклад $F(x_1, x_2, \ldots, x_n)$, більші від відповідних значень другої. Тоді говорять, що треба розв'язати нерівність

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

з невідомими $x_1, x_2, \ldots, x_n$.

Отже, *нерівність $F(x_1, x_2, \ldots, x_n) > \Phi(x_1, x_2, \ldots, x_n)$ — це символічний запис задачі про знаходження систем значень аргументів (невідомих) $x_1, x_2, \ldots, x_n$, при яких значення функції $F(x_1, x_2, \ldots, x_n)$ більші від відповідних значень функції $\Phi(x_1, x_2, \ldots, x_n)$*. Читач самостійно легко сформулює означення нерівностей

$$F(x_1, x_2, \ldots, x_n) < \Phi(x_1, x_2, \ldots x_n),\ F(x_1, x_2, \ldots, x_n) \geqslant$$
$$\geqslant \Phi(x_1, x_2, \ldots, x_n),\quad F(x_1, x_2, \ldots, x_n) \leqslant \Phi(x_1, x_2, \ldots, x_n).$$

Нерівність з n невідомими $x_1, x_2, \ldots, x_n$ в загальному вигляді записують так:

$$F(x_1, x_2, \ldots, x_n) \vee \Phi(x_1, x_2, \ldots, x_n), \qquad (1)$$

де $\vee$ — символ, під яким можна розуміти будь-який із знаків $>$, $\geqslant$, $<$, $\leqslant$. При цьому символ $\wedge$ означає відповідно знак $<$, $\leqslant$, $>$, $\geqslant$.

Функцію $F(x_1, x_2, \ldots, x_n)$ називають л і в о ю, а $\Phi(x_1, x_2, \ldots, x_n)$ — п р а в о ю частиною нерівності. Далі вважатимемо, що ліві й праві частини нерівностей — дійсні функції дійсних змінних, і, отже, всі нерівності розглядатимемо у полі д і й с н и х ч и с е л.

Кожну систему значень невідомих $x_1, x_2, \ldots, x_n$, при яких ліва й права частини нерівності визначені, називають *допустимою системою значень невідомих*. Множина всіх допустимих систем значень невідомих називається *областю допустимих систем значень* невідомих (ОДЗ) нерівності або *областю визначення нерівності*. Нерівність, до якої входять невідомі, може бути справедлива для одних допустимих систем значень невідомих і несправедлива для інших. Так нерівність $x + y > 2$ справедлива при $x = 1$, $y = 2$ і несправедлива при $x = 1$, $y = 1$.

Кожна допустима система значень невідомих $x_1 = k_1$, $x_2 = k_2$, ... , $x_n = k_n$, при яких справджується нерівність (1), називається *розв'язком* цієї нерівності. Інакше кажучи, *допустиму систему значень невідомих* $x_1 = k_1$, $x_2 = k_2$, ... , $x_n = k_n$ *називають розв'язком нерівності* (1), *якщо*

$$F(k_1, k_2, \ldots, k_n) \vee \Phi(k_1, k_2, \ldots, k_n).$$

Як і у випадку рівнянь, розв'язок нерівності з невідомими $x_1, x_2, \ldots, x_n$ будемо записувати як упорядковану множину чисел $(k_1, k_2, \ldots, k_n)$, що є значеннями невідомих, або як сукупність рівностей $x_1 = k_1$, $x_2 = k_2$, ... , $x_n = k_n$. Якщо система чисел $(k_1, k_2, \ldots, k_n)$ є розв'язком нерівності (1), то кажуть, що вона *задовольняє* нерівність (1). А якщо $(k_1, k_2, \ldots, k_n)$ не є розв'язком нерівності (1), то кажуть, що ця система чисел *не задовольняє* нерівність (1).

Якщо нерівність (1) справджується при всіх допустимих системах значень невідомих $x_1, x_2, \ldots, x_n$, то кажуть, що вона справджується тотожно, і називають її *тотожною нерівністю*. Тотожною, наприклад, є нерівність $x_1^2 + x_2^2 + \cdots + x_n^2 \geqslant 0$.

Розв'язати нерівність — це значить знайти множину всіх її розв'язків. Нехай $\boldsymbol{L}_n$ — деякий n-вимірний векторний простір над полем дійсних чисел $\mathbf{R}$. Припустимо, що в цьому просторі зафіксовано деякий базис $B\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$. Тоді, як відомо, кожен вектор $\boldsymbol{a}$ простору $\boldsymbol{L}_n$ однозначно визначається своїми координатами $\alpha_1, \alpha_2, \ldots, \alpha_n$ у цьому базисі.

Будемо вважати, що в нерівності (1) $x_1, x_2, \ldots, x_n$ позначають координати вектора $\boldsymbol{x} \in \boldsymbol{L}_n$. Тоді ліва й права частини нерівності (1) будуть функціями, заданими на просторі $\boldsymbol{L}_n$, значення яких належать полю дійсних чисел $\mathbf{R}$. Трактуючи так ліву й праву частини нерівності (1), говорять, що нерівність (1) розглядається над дійсним векторним простором $\boldsymbol{L}_n$. Нерівність (1), очевидно, можна розглядати над будь-яким n-вимірним дійсним векторним простором. Всюди далі, ради визначеності, нерівності з n невідомими розглядатимемо над простором V_n n-вимірних векторів $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ з компонентами x_i із поля дійсних чисел $\mathbf{R}$. У відповідності до прийнятих вище

означень вектор $\boldsymbol{l} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ називатимемо *допустимим* для нерівності (1), якщо його компоненти $\lambda_1, \lambda_2, \ldots, \lambda_n$ утворюють допустиму систему значень невідомих $x_1, x_2, \ldots, x_n$ цієї нерівності. Допустимий вектор $\boldsymbol{l} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ називатимемо *розв'язком* нерівності (1), якщо його компоненти $\lambda_1, \lambda_2, \ldots, \lambda_n$ задовольняють цю нерівність.

Аналогічно до того, як це робилось для рівнянь, означаємо поняття системи нерівностей. Нехай дано нерівності $F_i(x_1, x_2, \ldots, x_n) \vee \vee \Phi_i(x_1, x_2, \ldots, x_n)$ $(i = 1, 2, \ldots, m)$. Ми говоритимемо, що ці нерівності утворюють *систему* нерівностей

$$\begin{cases} F_1(x_1, x_2, \ldots, x_n) \vee \Phi_1(x_1, x_2, \ldots, x_n), \\ F_2(x_1, x_2, \ldots, x_n) \vee \Phi_2(x_1, x_2, \ldots, x_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ F_m(x_1, x_2, \ldots, x_n) \vee \Phi_m(x_1, x_2, \ldots, x_n), \end{cases} \tag{2}$$

якщо треба знайти всі допустимі вектори, кожен з яких є розв'язком всіх нерівностей (2). Вектор $\boldsymbol{g} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$, що є розв'язком кожної з нерівностей системи (2), називають *розв'язком* цієї *системи нерівностей*.

Розв'язок $\boldsymbol{g} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ системи нерівностей (2) називають д о д а т н и м, якщо всі компоненти $\gamma_1, \gamma_2, \ldots, \gamma_n$ — додатні, його називають н е в і д ' є м н и м, якщо всі компоненти — невід'ємні.

Розв'язати систему нерівностей — це означає знайти множину всіх її розв'язків. Якщо Q_i — множина всіх розв'язків нерівності

$$F_i(x_1, x_2, \ldots, x_n) \vee \Phi_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, m),$$

а Q — множина всіх розв'язків системи нерівностей (2), то

$$Q = Q_1 \cap Q_2 \cap \cdots \cap Q_m. \tag{3}$$

Система (2) називається с у м і с н о ю, якщо вона має принаймні один розв'язок; в противному разі вона називається н е с у м і с н о ю.

1.2. Рівносильність нерівностей і систем нерівностей. Нехай дано нерівності

$$f(x_1, x_2, \ldots, x_n) \vee \varphi(x_1, x_2, \ldots, x_n) \tag{4}$$

і

$$F(x_1, x_2, \ldots, x_n) \vee \Phi(x_1, x_2, \ldots, x_n). \tag{5}$$

Якщо всі розв'язки нерівності (4) задовольняють нерівність (5), то кажуть, що нерівність (5) є *наслідком* нерівності (4).

Аналогічно нерівність

$$F(x_1, x_2, \ldots, x_n) \vee \Phi(x_1, x_2, \ldots, x_n) \tag{6}$$

називають *наслідком* системи нерівностей

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) \vee \varphi_1(x_1, x_2, \ldots, x_n), \\ f_2(x_1, x_2, \ldots, x_n) \vee \varphi_2(x_1, x_2, \ldots, x_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_m(x_1, x_2, \ldots, x_n) \vee \varphi_m(x_1, x_2, \ldots, x_n), \end{cases} \tag{7}$$

якщо кожний розв'язок системи (7) є розв'язком і нерівності (6).

Систему нерівностей

$$\begin{cases} F_1(x_1, x_2, \dots, x_n) \vee \Phi_1(x_1, x_2, \dots, x_n), \\ F_2(x_1, x_2, \dots, x_n) \vee \Phi_2(x_1, x_2, \dots, x_n), \\ \dots\dots\dots\dots\dots\dots\dots \\ F_s(x_1, x_2, \dots, x_n) \vee \Phi_s(x_1, x_2, \dots, x_n), \end{cases} \tag{8}$$

називають наслідком системи нерівностей (7), якщо кожний розв'язок системи (7) є розв'язком і системи (8).

Дві нерівності (системи нерівностей) називаються рівносильними, якщо множини розв'язків їх збігаються. Інакше кажучи, дві нерівності (системи нерівностей) називаються рівносильними, якщо кожна з них є наслідком іншої.

У процесі розв'язування задач часто доводиться посилатися на такі твердження про наслідки з нерівностей та про рівносильність нерівностей і їх систем.

Теорема 1. *Якщо на деякій множині M систем значень невідомих $x_1, x_2, \dots, x_n$ справджуються нерівності*

$$F_1(x_1, x_2, \dots, x_n) \wedge \Phi_1(x_1, x_2, \dots, x_n) \tag{9}$$

і

$$F_2(x_1, x_2, \dots, x_n) \wedge \Phi_2(x_1, x_2, \dots, x_n). \tag{10}$$

то на множині M справджується й нерівність

$$F_1(x_1, x_2, \dots, x_n) + F_2(x_1, x_2, \dots, x_n) \wedge \\ \wedge \Phi_1(x_1, x_2, \dots, x_n) + \Phi_2(x_1, x_2, \dots, x_n), \tag{11}$$

тобто нерівність (11) *є наслідком нерівностей* (9) *і* (10).

Доведення. Нехай $x_1 = a_1, x_2 = a_2, \dots, x_n = a_n$ — довільна система значень невідомих, що належить множині M. Тоді справедливі числові нерівності

$$F_1(a_1, a_2, \dots, a_n) \wedge \Phi_1(a_1, a_2, \dots, a_n),$$
$$F_2(a_1, a_2, \dots, a_n) \wedge \Phi_2(a_1, a_2, \dots, a_n),$$

а тому справедлива також нерівність

$$F_1(a_1, a_2, \dots, a_n) + F_2(a_1, a_2, \dots, a_n) \wedge \Phi_1(a_1, a_2, \dots, a_n) + \\ + \Phi_2(a_1, a_2, \dots, a_n).$$

Отже, для кожної системи значень невідомих $x_1 = a_1, x_2 = a_2, \dots, x_n = a_n$, що належить множині M, справджується нерівність (11). Цим теорему доведено.

Аналогічними міркуваннями доводять і таку теорему.

Теорема 2. *Якщо на деякій множині M систем значень невідомих $x_1, x_2, \dots, x_n$ справджуються нерівності*

$$0 < F_1(x_1, x_2, \dots, x_n) < \Phi_1(x_1, x_2, \dots, x_n) \tag{12}$$

і

$$0 < F_2(x_1, x_2, \dots, x_n) < \Phi_2(x_1, x_2, \dots, x_n), \tag{13}$$

то на множині M справджуються й нерівності

$$0 < F_1(x_1, x_2, \dots, x_n)\, F_2(x_1, x_2, \dots, x_n) < \\ < \Phi_1(x_1, x_2, \dots, x_n)\, \Phi_2(x_1, x_2, \dots, x_n), \tag{14}$$

тобто нерівності (14) *є наслідком нерівностей* (12) *і* (13).

Теорема 3. *Якщо до обох частин нерівності*

$$F(x_1, x_2, \dots, x_n) \vee \Phi(x_1, x_2, \dots, x_n) \tag{14'}$$

додати функцію $\omega(x_1, x_2, \dots, x_n)$, яка визначена при всіх допустимих системах значень невідомих нерівності (14'), *то дістанемо нерівність*

$$F(x_1, x_2, \dots, x_n) + \omega(x_1, x_2, \dots, x_n) \vee \\ \vee \Phi(x_1, x_2, \dots, x_n) + \omega(x_1, x_2, \dots, x_n), \tag{14''}$$

рівносильну заданій.

Доведення. Справді, якщо вектор $(k_1, k_2, \dots, k_n)$ є розв'язком нерівності (14'), то

$$F(k_1, k_2, \dots, k_n) \vee \Phi(k_1, k_2, \dots, k_n).$$

Звідси

$$F(k_1, k_2, \dots, k_n) + \omega(k_1, k_2, \dots, k_n) \vee \Phi(k_1, k_2, \dots, k_n) + \\ + \omega(k_1, k_2, \dots, k_n)$$

і, отже, вектор $(k_1, k_2, \dots, k_n)$ є також розв'язком і нерівності (14''). Навпаки, якщо $(k_1, k_2, \dots, k_n)$ є розв'язком нерівності (14''), то

$$F(k_1, k_2, \dots, k_n) + \omega(k_1, k_2, \dots, k_n) \vee \Phi(k_1, k_2, \dots, k_n) + \\ + \omega(k_1, k_2, \dots, k_n)$$

і, отже,

$$F(k_1, k_2, \dots, k_n) \vee \Phi(k_1, k_2, \dots, k_n),$$

а це означає, що $(k_1, k_2, \dots, k_n)$ є розв'язком нерівності (14').

Отже кожний розв'язок будь-якої з нерівностей (14') і (14'') є розв'язком і другої.

Наслідок. *Будь-який доданок з однієї частини нерівності можна перенести в другу з протилежним знаком.*

Теорема 4. *Якщо функція $\omega(x_1, x_2, \dots, x_n)$ додатна при всіх допустимих системах значень невідомих, то нерівності*

$$F(x_1, x_2, \dots, x_n) \vee \Phi(x_1, x_2, \dots, x_n) \tag{15}$$

і

$$\omega(x_1, x_2, \dots, x_n)\, F(x_1, x_2, \dots, x_n) \vee \\ \vee \omega(x_1, x_2, \dots, x_n)\, \Phi(x_1, x_2, \dots, x_n) \tag{16}$$

рівносильні.

Доведення. Справді, якщо система значень невідомих $x_1 = k_1, x_2 = k_2, \dots, x_n = k_n$ — розв'язок нерівності (15), тобто

$$F(k_1, k_2, \dots, k_n) \vee \Phi(k_1, k_2, \dots, k_n),$$

то, оскільки $\omega(k_1, k_2, \dots, k_n) > 0$,

$$\omega(k_1, k_2, \dots, k_n) F(k_1, k_2, \dots, k_n) \vee \vee \omega(k_1, k_2, \dots, k_n) \Phi(k_1, k_2, \dots, k_n)$$

і, отже, $(k_1, k_2, \dots, k_n)$ є також розв'язком і нерівності (16). Навпаки, якщо $(k_1, k_2, \dots, k_n)$ є розв'язком нерівності (16), тобто

$$\omega(k_1, k_2, \dots, k_n) F(k_1, k_2, \dots, k_n) \vee \vee \omega(k_1, k_2, \dots, k_n) \Phi(k_1, k_2, \dots, k_n),$$

то, оскільки $\omega(k_1, k_2, \dots, k_n) > 0$,

$$F(k_1, k_2, \dots, k_n) \vee \Phi(k_1, k_2, \dots, k_n)$$

і, отже, $(k_1, k_2, \dots, k_n)$ є розв'язком нерівності (15). Таким чином, кожний розв'язок однієї з нерівностей (15) і (16) є розв'язком і другої. Аналогічними міркуваннями доводять справедливість такої теореми:

Теорема 5. *Якщо функція $\omega(x_1, x_2, \dots, x_n)$ від'ємна при всіх допустимих системах значень невідомих, то нерівності*

$$F(x_1, x_2, \dots, x_n) \vee \Phi(x_1, x_2, \dots, x_n)$$

і

$$\omega(x_1, x_2, \dots, x_n) F(x_1, x_2, \dots, x_n) \wedge \omega(x_1, x_2, \dots, x_n) \Phi(x_1, x_2, \dots, x_n)$$

рівносильні.

Теорема 6. *Якщо в системі нерівностей*

$$\begin{cases} F_1(x_1, x_2, \dots, x_n) \vee \Phi_1(x_1, x_2, \dots, x_n), \\ \dots\dots\dots\dots\dots\dots \\ F_s(x_1, x_2, \dots, x_n) \vee \Phi_s(x_1, x_2, \dots, x_n), \\ \dots\dots\dots\dots\dots\dots \\ F_m(x_1, x_2, \dots, x_n) \vee \Phi_m(x_1, x_2, \dots, x_n) \end{cases} \quad (17)$$

будь-яку нерівність замінено рівносильною їй нерівністю, то дістанемо систему нерівностей, рівносильну заданій.

Доведення. Припустимо, що в системі (17) нерівність

$$F_s(x_1, x_2, \dots, x_n) \vee \Phi_s(x_1, x_2, \dots, x_n)$$

замінено рівносильною їй нерівністю

$$f(x_1, x_2, \dots, x_n) \vee \varphi(x_1, x_2, \dots, x_n).$$

Утворена в результаті цієї заміни система нерівностей

$$\begin{cases} F_1(x_1, x_2, \dots, x_n) \vee \Phi_1(x_1, x_2, \dots, x_n), \\ \dots\dots\dots\dots\dots\dots \\ F_{s-1}(x_1, x_2, \dots, x_n) \vee \Phi_{s-1}(x_1, x_2, \dots, x_n), \\ f(x_1, x_2, \dots, x_n) \vee \varphi(x_1, x_2, \dots, x_n), \\ F_{s+1}(x_1, x_2, \dots, x_n) \vee \Phi_{s+1}(x_1, x_2, \dots, x_n), \\ \dots\dots\dots\dots\dots\dots \\ F_m(x_1, x_2, \dots, x_n) \vee \Phi_m(x_1, x_2, \dots, x_n) \end{cases} \quad (18)$$

рівносильна системі (17).

Справді, системи (17) і (18) відрізняються одна від одної лише s-ю нерівністю. Оскільки s-ті нерівності цих систем рівносильні і, отже, мають одні й ті самі розв'язки, то кожний розв'язок $x_1 = k_1$, $x_2 = k_2$, ... , $x_n = k_n$ однієї з систем (17) і (18) буде розв'язком другої.

З доведеної теореми безпосередньо випливає такий наслідок: *якщо в заданій системі нерівностей кілька її нерівностей замінимо рівносильними їм нерівностями, то дістанемо систему нерівностей, рівносильну заданій.*

Примітка. Зауважимо, що поняття нерівності й системи нерівностей можна також означити й з допомогою понять математичної логіки.

Все, що в I,17.2 було сказано про рівняння і системи рівнянь з n невідомими, можна дослівно повторити про нерівності і системи нерівностей з n невідомими (замінюючи лише в кожному рівнянні символ «=» на символ «$\leqslant$»). На нерівності повністю переносяться означення рівносильності та поняття наслідку, наведені в I, 17.2 для випадку рівнянь і систем рівнянь.

1.3. Системи лінійних нерівностей та їх геометричний смисл. Нерівність

$$F(x_1, x_2, \dots, x_n) \vee \Phi(x_1, x_2, \dots, x_n) \quad (19)$$

називають *алгебраїчною*, якщо функції $F(x_1, x_2, \dots, x_n)$ і $\Phi(x_1, x_2, \dots, x_n)$ — многочлени. Причому, якщо $F(x_1, x_2, \dots, x_n)$ і $\Phi(x_1, x_2, \dots, x_n)$ — многочлени відповідно степеня m і s й $m > s$, то нерівність (19) називають нерівністю степеня m з n невідомими. Зокрема, якщо $F(x_1, x_2, \dots, x_n)$ і $\Phi(x_1, x_2, \dots, x_n)$ — лінійні функції, тобто многочлени 1-го степеня, то нерівність (19) називають *лінійною*. Інакше кажучи, лінійними нерівностями називають нерівності, до яких невідомі входять лише в першому степені. Такими, наприклад, є нерівності $2{,}7x + 3{,}2y - 7{,}11 \leqslant 0$ і $x + 2{,}32y - z - \sqrt{3} \geqslant 0$. Перша з них це лінійна нерівність з двома, а друга — з трьома невідомими. Лінійну нерівність з n невідомими в загальному вигляді записуватимемо так:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n - a \leqslant 0. \quad (20)$$

Будь-яку лінійну нерівність за наслідком з теореми 3 можна записати в такій формі.

В цьому записі $a_1, a_2, \dots, a_n$ — задані дійсні числа, деякі з них можуть дорівнювати нулю.

Числа $a_1, a_2, \dots, a_n$ називають *коефіцієнтами* нерівності (20), а число a — *вільним членом*. Лінійну нерівність, у якої вільний член дорівнює нулю, тобто нерівність вигляду

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \leqslant 0,$$

називають *лінійною однорідною нерівністю*.

Систему нерівностей, що складається з лінійних нерівностей, називають *системою лінійних нерівностей*.

Систему лінійних нерівностей з n невідомими в загальному вигляді записуватимемо так:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - a_1 \leqslant 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - a_2 \leqslant 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - a_m \leqslant 0. \end{cases}$$

Система нерівностей, що містить лише лінійні однорідні нерівності, називається *однорідною системою лінійних нерівностей*.

З'ясуємо геометричний смисл системи лінійних нерівностей. Це допоможе нам краще зрозуміти, в чому ж суть задачі розв'язання системи лінійних нерівностей.

Геометричний смисл системи нерівностей з двома невідомими. Розглянемо спочатку одну лінійну нерівність з двома невідомими x і y:

$$ax + by - c \leqslant 0$$

Будемо вважати x і y декартовими координатами точки площини. Як відомо читачеві з курсу геометрії, співвідношення $ax + by - c = 0$ є рівняння прямої, а нерівність $ax + by - c \leqslant 0$ визначає одну з півплощин, тобто справджується в точках однієї з півплощин, на які пряма, що задається рівнянням $ax + by - c = 0$, ділить координатну площину і не справджується у решті точок площини. (Пряма $ax + by - c = 0$ відноситься до кожної з півплощин).

Припустимо тепер, що задана деяка система лінійних нерівностей з невідомими x і y:

$$\begin{cases} a_1x + b_1y - c_1 \leqslant 0, \\ a_2x + b_2y - c_2 \leqslant 0, \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my - c_m \leqslant 0. \end{cases} \qquad (21)$$

Сукупність розв'язків кожної з нерівностей цієї системи геометрично зображається множиною точок деякої півплощини. Отже, сукупність всіх розв'язків системи (21) геометрично зображається множиною точок перетину (спільної частини) цих півплощин. Перетин скінченного числа півплощин є деяка область K. Границя області K, взагалі кажучи, складається з відрізків прямих, тому говорять, що K є *многокутна область*. Область K називають *областю розв'язків* системи нерівностей (21). Якщо область K обмежена, то її називають *многокутником розв'язків* системи нерівностей (21). Зауважимо, що область розв'язків може бути многокутником, нескінченною областю, відокремленою деякою незамкненою ламаною лінією, смугою, що міститься між двома паралельними прямими, прямою, променем, відрізком, точкою, або порожньою множиною.

Останній випадок має місце тоді, коли система (9) не має розв'язків, тобто несумісна.

Півплощина, як відомо читачеві з курсу геометрії, є о п у к л а множина[1]. Перетин опуклих множин, як легко перевірити, є опукла множина. Тому область розв'язків K є опукла множина, а многокутник розв'язків — о п у к л и й многокутник.

Оскільки всі нерівності системи (21) справджуються одночасно в точках многокутної області K і тільки в них, то говорять, що система нерівностей (9) *визначає многокутну область* K.

Приклад. Знайти область розв'язків системи нерівностей

$$\begin{cases} 2x - 3y + 13 \geqslant 0, \\ x + y - 6 \geqslant 0, \\ 4x - y - 19 \leqslant 0. \end{cases}$$

Запишемо цю систему нерівностей так:

$$\begin{cases} y \leqslant \frac{2}{3}x + \frac{13}{3}, \\ y \geqslant -x + 6, \\ y \geqslant 4x - 19. \end{cases}$$

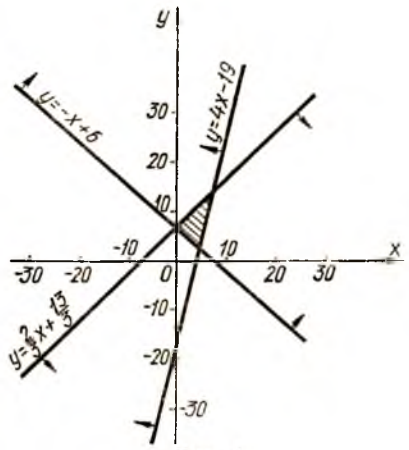


Рис. 1.

Перша нерівність системи визначає півплощину, розташовану нижче від прямої $y = \frac{2}{3}x + \frac{13}{3}$; друга — півплощину, розташовану вище від прямої $y = -x + 6$; третя нерівність визначає півплощину, розташовану вище від прямої $y = 4x - 19$ (рис. 1). На рисунку стрілки показують ті півплощини, які визначаються відповідними нерівностями. Областю розв'язків розглядуваної системи нерівностей є трикутник.

Геометричний смисл системи нерівностей з трьома невідомими. Розглянемо тепер лінійну нерівність з трьома невідомими

$$ax + by + cz - d \leqslant 0. \qquad (22)$$

Вважатимемо x, y і z декартовими координатами точки звичайного простору E_3. З'ясуємо, що становить собою множина всіх точок простору, координати яких задовольняють нерівність (22).

Припустимо, що в нерівності (22) $c \neq 0$. Розв'язавши нерівність (22) відносно z і позначивши $-\frac{a}{c}$ через k, $-\frac{b}{c}$ через l і $\frac{d}{c}$ через p, дістанемо нерівність

$$z \geqslant kx + ly + p, \qquad (23)$$

якщо $c < 0$, і нерівність

$$z \leqslant kx + ly + p, \qquad (24)$$

якщо $c > 0$.

[1] Опукла множина — це множина, яка разом з будь-якими двома своїми точками A і B містить увесь відрізок AB.

Співвідношення $z = kx + ly + p$, як відомо, є рівняння деякої площини П, яка ділить простір E_3 на два півпростори: півпростір E_3', що лежить «над» площиною П, і півпростір E_3'', що лежить «під» площиною П. Площину П ми відноситимемо до кожного з півпросторів E_3' і E_3''. Легко бачити, що координати будь-якої точки (x_1, y_1, z_1) півпростору E_3' задовольняють нерівність (23), а координати будь-якої точки (x_2, y_2, z_2) півпростору E_3'' — нерівність (24). Очевидно також, що й навпаки, кожна точка, координати якої задовольняють нерівність (23), належить півпростору E_3', а кожна точка, координати якої задовольняють нерівність (24), належить півпростору E_3''. Отже, множина всіх розв'язків нерівності (23) геометрично зображається сукупністю всіх точок півпростору E_3', а нерівності (24) — півпростору E_3''.

Якщо коефіцієнт $c = 0$, а один з коефіцієнтів a або b є відмінним від нуля, то міркування проводять аналогічно. З щойно викладеного випливає, що нерівність (22) справджується в одному з півпросторів, на які площина, що задається рівнянням $ax + bx + cx - d = 0$, ділить простір E_3. Таким чином, нерівність (22) визначає один з півпросторів E_3' і E_3'', на які площина $ax + bx + cx - d = 0$ ділить простір E_3.

Припустимо тепер, що дано систему лінійних нерівностей з трьома невідомими x, y і z

$$\begin{cases} a_1x + b_1y + c_1z - d_1 \leqslant 0, \\ a_2x + b_2y + c_2z - d_2 \leqslant 0, \\ \dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz - d_m \leqslant 0. \end{cases} \tag{25}$$

Кожна нерівність системи (25) визначає деякий півпростір. Отже, сукупність всіх розв'язків системи (25) геометрично зображається множиною точок перетину (спільної частини) півпросторів, що визначаються нерівностями системи (25). Перетин скінченного числа півпросторів є деяка опукла многогранна область M, оскільки півпростір є опукла множина.

Як і у випадку систем нерівностей з двома невідомими, область M ми називатимемо *областю розв'язків системи нерівностей* (25).

Перетин півпросторів, що визначаються нерівностями системи (25), може виявитися необмеженою областю, може також трапитися, що він буде порожньою множиною; це означатиме, що система (25) розв'язків не має, тобто що вона суперечлива.

Якщо область розв'язків M обмежена, то вона являє собою опуклий многогранник, який зокрема може бути многокутником, відрізком чи точкою.

Приклад. Побудувати многогранник розв'язків системи нерівностей

$$\begin{cases} 2x - 3y + 2z - 6 \leqslant 0, \\ x \geqslant 0, \\ y \leqslant 0. \\ z \geqslant 0. \end{cases}$$

Як відомо, рівняння $x = 0$, $y = 0$, $z = 0$ — це рівняння координатних площин. Тому нерівність $z \geqslant 0$ визначає півпростір, який лежить над площиною xOy (включаючи й цю площину), нерівність $x \geqslant 0$ визначає півпростір, що лежить справа від площини yOz (включаючи й саму площину yOz). Нерівність $y \leqslant 0$ визначає півпростір, що лежить за координатною площиною xOz (включаючи й саму площину xOz).

Перерізом цих півпросторів є, очевидно, о к т а н т, утворений координатними площинами, який міститься між додатними частинами Ox і Oz і від'ємною частиною Oy.

Розглянемо тепер нерівність

$$2x - 3y + 2z - 6 \leqslant 0.$$

Перенісши в праву частину вільний член і поділивши на нього всю нерівність, дістанемо

$$\frac{x}{3} + \frac{y}{-2} + \frac{z}{3} \leqslant 1.$$

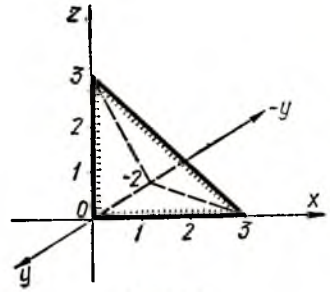


Рис. 2.

Ця нерівність визначає півпростір, який міститься під площиною, що відтинає на координатних осях Ox, Oy і Oz відповідно відрізки 3, -2 і 3. Внаслідок перерізу всіх заданих півпросторів маємо многогранник розв'язків, показаний на рис. 2.

Геометричний смисл системи лінійних нерівностей з n невідомими. Раніше, ніж перейти до з'ясування геометричного смислу системи лінійних нерівностей з n невідомими, нагадаємо деякі відомості з геометрії.

Нехай дана деяка множина $\mathfrak{A}$, елементи якої називатимемо точками і позначатимемо буквами A, B, C, ... , M, ... Нехай дано також деякий векторний простір L. Припустимо, що кожній упорядкованій парі точок з $\mathfrak{A}$ поставлено у відповідність вектор з простору L. Якщо парі точок A, B відповідає вектор $\boldsymbol{x}$, то писатимемо: $\boldsymbol{x} = AB$; таким чином AB є лише нове позначення вектора $\boldsymbol{x}$. Точку A називатимемо початком, а точку B — кінцем вектора AB.

Означення 1. *Множина $\mathfrak{A}$ точок A, B, C, ... , M, ... , співставлена з векторним простором L, називається афінним простором, якщо виконуються такі вимоги:*

1. *Для кожної точки $A \in \mathfrak{A}$ і кожного вектора $\boldsymbol{x} \in L$ знайдеться і притому тільки одна точка $B \in \mathfrak{A}$ така, що $AB = \boldsymbol{x}$.*

2. *Якщо $AB = \boldsymbol{x}$, $BC = \boldsymbol{y}$, то $AC = \boldsymbol{x} + \boldsymbol{y}$.*

Якщо векторний простір L — n-вимірний, то пов'язаний з ним афінний простір $\mathfrak{A}$ називають *n-вимірним* і позначають символом $\mathfrak{A}_n$. Афінний простір $\mathfrak{A}$ називають *дійсним*, якщо відповідний йому векторний простір L — дійсний; його називають *комплексним*, якщо простір L — комплексний.

Введемо в n-вимірному афінному просторі $\mathfrak{A}_n$ так звану афінну систему координат.

Для цього виберемо в просторі $\mathfrak{A}_n$ довільну точку O, яку називатимемо *початком координат*, а у відповідному векторному просторі L_n зафіксуємо який-небудь базис $B\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$. Сукупність $\{O, \boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$, складену з точки O і векторів базису $B\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$, називатимемо *афінною системою координат* простору $\mathfrak{A}_n$.

Нехай M — довільна точка простору $\mathfrak{A}_n$. Разом з початком координат O точка M визначає вектор $OM \in L_n$ який називають *радіусом-вектором* точки M.

Радіус-вектор OM має в базисі $B\{e_1, e_2, \ldots, e_n\}$ цілком визначені координати $\gamma_1, \gamma_2, \ldots, \gamma_n$. Числа $\gamma_1, \gamma_2, \ldots, \gamma_n$ називають координатами точки M в афінній системі координат $\{O, e_1, e_2, \ldots, e_n\}$.

Нехай в афінному просторі $\mathfrak{A}_n$ зафіксовано деяку точку A, а у відповідному векторному просторі L_n — деякий s-вимірний підпростір L_s.

Означення 2. *Множина всіх точок M афінного простору $\mathfrak{A}_n$, для яких $AM \in L_s$ називається s-вимірною площиною простору $\mathfrak{A}_n$, що проходить через точку A в напрямку підпростору L_s.*

Площина розмірності $n-1$ називається *гіперплощиною*. Нехай $\mathfrak{A}_n$ є n-вимірний дійсний афінний простір. Зафіксуємо в цьому просторі яку-небудь афінну систему координат $\{O, e_1, e_2, \ldots, e_n\}$.

Розглянемо тепер нерівність

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - a \leqslant 0. \tag{26}$$

Вважатимемо, що $x_1, x_2, \ldots, x_n$ — це координати точки простору $\mathfrak{A}_n$ у вибраній системі координат, і з'ясуємо, що являє собою множина всіх точок простору $\mathfrak{A}_n$, координати яких задовольняють нерівність (26), тобто що визначає нерівність (26). Як відомо з курсу геометрії, лінійне рівняння

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - a = 0, \tag{27}$$

не всі коефіцієнти якого — нулі, визначає в просторі $\mathfrak{A}_n$ деяку гіперплощину. Крім точок, які належать гіперплощині (27) в просторі $\mathfrak{A}_n$, очевидно, знайдуться точки, в яких виконується нерівність

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - a < 0 \tag{28}$$

і точки, в яких виконується нерівність

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - a > 0. \tag{29}$$

Таким чином, у просторі $\mathfrak{A}_n$ відносно гіперплощини (27) можна виділити дві частини: область, в точках якої виконується нерівність (28), і область, в точках якої виконується нерівність (29). Ці області називають відкритими *півпросторами*.

Якщо до кожного з цих півпросторів приєднаємо гіперплощину (27), то дістанемо так звані замкнуті півпростори. Один із цих півпросторів складається з точок, координати яких задовольняють нерівність

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - a \leqslant 0,$$

а другий — з точок, координати яких задовольняють нерівність

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - a \geqslant 0.$$

Отже, нерівність (26) визначає деякий замкнутий півпростір простору $\mathfrak{A}_n$.

Після всього сказаного легко з'ясувати, який геометричний смисл системи лінійних нерівностей з n невідомими. Нехай дана система лінійних нерівностей з n невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - a_1 \leqslant 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - a_2 \leqslant 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - a_m \leqslant 0. \end{cases} \tag{30}$$

Кожна з нерівностей системи (30) визначає в просторі $\mathfrak{A}_n$ деякий замкнутий півпростір, який є опуклою множиною. Система (30) визначає деяку область M, що є перетином півпросторів, які визначаються її нерівностями.

Область M, як перетин опуклих множин, є опукла множина. За аналогією з двовимірним і тривимірним випадками область M називають *опуклою многогранною областю*. Отже, множина точок простору $\mathfrak{A}_n$, координати яких задовольняють систему нерівностей (30), є опукла многогранна область M. Область M називають *областю розв'язків* системи нерівностей (30). Якщо область M є обмежена множина, то її називають *опуклим многогранником*.

§ 2. ОСНОВНІ ВЛАСТИВОСТІ СИСТЕМ ЛІНІЙНИХ НЕРІВНОСТЕЙ

2.1. Однорідні системи лінійних нерівностей. В цьому параграфі буде викладено кілька важливих теорем теорії лінійних нерівностей. Умовимося систему лінійних нерівностей з n невідомими в загальному вигляді записувати так:

$$\begin{cases} f_1(x) - a_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - a_1 \leqslant 0, \\ f_2(x) - a_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - a_2 \leqslant 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_m(x) - a_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - a_m \leqslant 0, \end{cases} \tag{1}$$

або скорочено

$$\{f_i(x) - a_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - a_i \leqslant 0 \quad (i = 1, 2, \ldots, m). \tag{1'}$$

Однорідну систему лінійних нерівностей з n невідомими в загальному вигляді записуватимемо так:

$$\begin{cases} f_1(x) = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leqslant 0, \\ f_2(x) = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leqslant 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_m(x) = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leqslant 0, \end{cases} \tag{2}$$

або скорочено

$$\{f_i(x) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant 0 \quad (i = 1, 2, \ldots, m). \tag{2'}$$

Нехай дана однорідна система лінійних нерівностей (2) і лінійна однорідна нерівність

$$f(x) = b_1x_1 + b_2x_2 + \cdots + b_nx_n \leqslant 0. \tag{3}$$

Теорема 1 (Мінковського). *Якщо нерівність* $f(x) = b_1x_1 + b_2x_2 + \cdots + b_nx_n \leqslant 0$ *є наслідок системи нерівностей* (2), *то існують такі невід'ємні числа* $p_1, p_2, \ldots, p_m$, *що справджується тотожне відносно* $x = (x_1, x_2, \ldots, x_n) \in V_n$ *співвідношення*

$$f(x) = \sum_{i=1}^{m} p_i f_i(x).$$

Доведення. Насамперед покажемо, що коли виконується умова теореми, то існують такі дійсні числа $q_1, q_2, \ldots, q_m$, що

$$f(x) = \sum_{i=1}^{m} q_i f_i(x).$$

Справді, припустимо, що таких чисел не існує, тобто що $f(x)$ не є лінійною комбінацією функцій $f_1(x), f_2(x), \ldots, f_m(x)$. Тоді, як легко бачити, ранг розширеної матриці системи лінійних рівнянь

$$\begin{cases} f_i(x) = 0 & (i = 1, 2, \ldots, m), \\ f(x) = 1 \end{cases} \tag{4}$$

дорівнює рангу матриці цієї системи, а саме $r + 1$, де r — ранг матриці системи нерівностей (2) і тому система (4) сумісна. Нехай $(\gamma_1, \gamma_2, \ldots, \gamma_m)$ — деякий розв'язок системи (4). Цей розв'язок, очевидно, задовольняє систему нерівностей (2) і не задовольняє нерівність $f(x) \leqslant 0$, а це суперечить умові теореми. Отже, припущення, що $f(x)$ не є лінійною комбінацією функцій $f_1(x), f_2(x), \ldots, f_m(x)$, приводить до суперечності і тому воно неправильне. Тепер перейдемо безпосередньо до доведення теореми. Доводитимемо її індукцією по числу нерівностей системи (2). Теорема, очевидно, справедлива, якщо число нерівностей $m = 1$. Припустимо, що теорема справедлива для кожної системи з числом нерівностей меншим ніж m і доведемо, що тоді вона справедлива і для кожної системи нерівностей з числом нерівностей m.

Справді, нехай система (2) — будь-яка система нерівностей, що задовольняє умову теореми і складається з m нерівностей. Тоді, за доведеним вище, існують такі дійсні числа $q_1, q_2, \ldots, q_m$, що

$$f(x) = \sum_{i=1}^{m} q_i f_i(x). \tag{5}$$

Якщо всі числа q_i невід'ємні, то для системи (2), яка складається з m нерівностей, теорема справедлива. Припустимо, що принаймні одне з чисел $q_1, q_2, \ldots, q_m$ від'ємне. Не втрачаючи загальності міркувань, вважатимемо, що числа $q_1, q_2, \ldots, q_s$ — від'ємні, а числа $q_{s+1}, q_{s+2}, \ldots, q_m$ — невід'ємні (тут $m > 1$); не виключено, що $s = m$.

Нерівність

$$f^*(x) = q_1f_1(x) + \sum_{s<i<m} q_i f_i(x) \leqslant 0 \tag{6}$$

є наслідком системи нерівностей (2). Справді, з співвідношень (5) і (6) випливає, що

$$f(x) = f^*(x) + \sum_{1<i<s} q_i f_i(x), \tag{7}$$

звідки $f^*(x) = f(x) - \sum_{1<i<s} q_i f_i(x)$, тобто

$$f^*(x) = f(x) + \sum_{1<i<s} (-q_i) f_i(x). \tag{8}$$

Оскільки нерівності $f(x) \leqslant 0$ і $(-q_i) f_i(x) \leqslant 0$ $(1 < i \leqslant s)$ є наслідками системи нерівностей (2), то за рівністю (8) і нерівність $f^*(x) \leqslant 0$ є наслідок системи нерівностей (2). Нерівність (6), очевидно, є також наслідок системи нерівностей

$$\begin{cases} -f_1(x) \leqslant 0, \\ f_i(x) \leqslant 0 \quad (i = 2, 3, \ldots, m). \end{cases} \tag{9}$$

Оскільки нерівність (6) є наслідок системи нерівностей (2) і системи нерівностей (9), то вона є наслідок і системи нерівностей

$$\{ f_i(x) \leqslant 0 \quad (i = 2, 3, \ldots, m), \tag{10}$$

бо кожен розв'язок системи (10) є розв'язком або нерівності $f_1(x) \leqslant 0$, або нерівності $-f_1(x) \leqslant 0$, а отже, і розв'язком або системи нерівностей (2), або системи нерівностей (9). Тому, за індуктивним припущенням, існують такі невід'ємні числа $q'_2, q'_3, \ldots, q'_m$, що

$$f^*(x) = \sum_{i=2}^{m} q'_i f_i(x).$$

Підставивши цей вираз у рівність (7), дістанемо:

$$f(x) = \sum_{1<i<s} (q_i + q'_i) f_i(x) + \sum_{s<i<m} q'_i f_i(x).$$

Таким чином, ми знайшли для $f(x)$ нове зображення вигляду (5), в якому число від'ємних коефіцієнтів, очевидно, менше ніж s. Виходячи з цього зображення, повторимо попередні міркування; дістанемо нове зображення вигляду (5) для $f(x)$, в якому число від'ємних коефіцієнтів ще зменшиться. Зрозуміло, що, повторивши такі міркування кілька разів, ми дістанемо для $f(x)$ зображення вигляду (5), у якого всі коефіцієнти невід'ємні. Отже, для будь-якої системи, що складається з m нерівностей, теорема справедлива. Тому, за принципом індукції, теорема справедлива для системи, що складається з будь-якого числа нерівностей. Теорему доведено.

2.2. **Критерій несумісності системи лінійних нерівностей.** Як відомо, система лінійних нерівностей (1) називається *несумісною*, якщо вона не має жодного розв'язку, тобто якщо область її розв'яз-

ків є порожня множина. На перший погляд здається, що несумісні системи лінійних нерівностей не можуть становити ні теоретичного, ні практичного інтересу. Насправді це не так: властивості таких систем не лише цікаві самі по собі, а й дають можливість з'ясувати ряд важливих фактів. Зокрема, як ми побачимо, основна теорема теорії лінійного програмування, так звана теорема двоїстості, виводиться з деяких властивостей несумісних систем. Тому несумісні системи лінійних нерівностей заслуговують на таку саму увагу, як і сумісні системи. З теореми Мінковського випливає теорема, яка є критерієм несумісності системи лінійних нерівностей. Доведемо цю теорему.

Теорема 2. *Система лінійних нерівностей* (1) *несумісна тоді і тільки тоді, коли існують такі невід'ємні числа* $p_1, p_2, \dots, p_m$, *що справджується тотожне відносно* $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in V_n$ *співвідношення* $\sum_{i=1}^{m} p_i f_i(\boldsymbol{x}) = 0$ *і нерівність* $\sum_{i=1}^{m} p_i a_i < 0$.

Доведення. Припустимо, що система (1) несумісна. Розглянемо однорідну систему лінійних нерівностей

$$\{f_i(\boldsymbol{x}) - a_i t = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n - a_i t \leqslant 0 \quad (i = 1, 2, \dots, m). \tag{11}$$

Нехай $x_1 = \gamma_1$, $x_2 = \gamma_2$, ..., $x_n = \gamma_n$, $t = \beta$ — деякий довільно вибраний розв'язок[1] цієї системи нерівностей.

Тоді справджуються нерівності

$$a_{i1}\gamma_1 + a_{i2}\gamma_2 + \dots + a_{in}\gamma_n - a_i\beta \leqslant 0 \quad (i = 1, 2, \dots, m). \tag{12}$$

У розв'язку $(\gamma_1, \gamma_2, \dots, \gamma_n, \beta)$ число β не може бути додатним, оскільки при $\beta > 0$ із нерівностей (12) випливали б нерівності

$$a_{i1}\frac{\gamma_1}{\beta} + a_{i2}\frac{\gamma_2}{\beta} + \dots + a_{in}\frac{\gamma_n}{\beta} - a_i \leqslant 0 \quad (i = 1, 2, \dots, m),$$

а це означало б, що система нерівностей (1) сумісна.

Отже, кожен розв'язок $(\gamma_1, \gamma_2, \dots, \gamma_n, \beta)$ системи нерівностей (11) є також розв'язком і нерівності $t \leqslant 0$. Нерівність $t \leqslant 0$, таким чином, є наслідком системи нерівностей (11). Тому, за теоремою Мінковського, існують такі невід'ємні числа $p_1, p_2, \dots, p_m$, що справджується тотожне відносно $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in V_n$ і $t \in \mathbf{R}$ співвідношення

$$t = \sum_{i=1}^{m} p_i [f_i(\boldsymbol{x}) - a_i t].$$

Звідси при $t = 0$ дістаємо тотожне відносно $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in V_n$ співвідношення

$$\sum_{i=1}^{m} p_i f_i(\boldsymbol{x}) = 0, \tag{13}$$

[1] Система (11) має принаймні нульовий розв'язок $x_1 = 0$, $x_2 = 0$, ..., $x_n = 0$, $t = 0$.

а при $t = 1$, беручи до уваги співвідношення (13), дістаємо рівність

$$1 = -\sum_{i=1}^{m} p_i a_i,$$

з якої випливає справедливість нерівності

$$\sum_{i=1}^{m} p_i a_i < 0.$$

Цим необхідність умов теореми доведено. Доведемо тепер їх достатність. Припустимо, що існують такі невід'ємні числа $p_1, p_2, \dots, p_m$, що справджується тотожне відносно $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in V_n$ співвідношення $\sum_{i=1}^{m} p_i f_i(\boldsymbol{x}) = 0$ і нерівність $\sum_{i=1}^{m} p_i a_i < 0$. Тоді нерівність

$$\sum_{i=1}^{m} p_i [f_i(\boldsymbol{x}) - a_i] \leqslant 0, \tag{14}$$

яку, очевидно, можна записати так:

$$\sum_{i=1}^{m} p_i f_i(\boldsymbol{x}) - \sum_{i=1}^{m} p_i a_i \leqslant 0,$$

не має розв'язків. А тому не має розв'язків і система нерівностей (1), бо кожен розв'язок системи (1) був би розв'язком і нерівності (14). Отже, система нерівностей (1) несумісна. Достатність умов теореми також доведено.

Існування невід'ємних чисел $p_1, p_2, \dots, p_m$, для яких справджуються тотожне відносно $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in V_n$ співвідношення

$$\sum_{i=1}^{m} p_i f_i(\boldsymbol{x}) = 0 \tag{15}$$

і нерівність

$$\sum_{i=1}^{m} p_i a_i < 0, \tag{16}$$

рівносильне існуванню невід'ємного розв'язку системи лінійних рівнянь

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \dots + a_{mj}y_m = 0 \quad (j = 1, 2, \dots, n), \\ a_1 y_1 + a_2 y_2 + \dots + a_m y_m = -1. \end{cases} \tag{17}$$

Справді, якщо існують невід'ємні числа $p_1, p_2, \dots, p_m$ такі, що справджується тотожне співвідношення (15) і нерівність (16), то для будь-якого $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in V_n$

$$\sum_{i=1}^{m} p_i (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n) =$$

$$= \sum_{k=1}^{n} (a_{1k}p_1 + a_{1k}p_2 + \dots + a_{mk}p_m) x_k = 0 \tag{18}$$

і

$$\sum_{i=1}^{m} p_i a_i = -c, \text{ де } c > 0. \tag{19}$$

Узявши в співвідношенні (18) $x_j = 1$, $(j = 1, 2, \ldots, n)$, а $x_i = 0$ $(i \neq j)$, дістанемо:

$$a_{1j}p_1 + a_{2j}p_2 + \cdots + a_{mj}p_m = 0 \qquad (j = 1, 2, \ldots, n). \tag{20}$$

З рівностей (19) і (20) випливає, що $\left(\frac{p_1}{c}, \frac{p_2}{c}, \ldots, \frac{p_n}{c}\right)$ є невід'ємний розв'язок системи рівнянь (17). Навпаки, якщо $(p_1, p_2, \ldots, p_n)$ — невід'ємний розв'язок системи рівнянь (17), то

$$a_{1j}p_1 + a_{2j}p_2 + \cdots + a_{mj}p_m = 0 \qquad (j = 1, 2, \ldots, n),$$

$$\sum_{i=1}^{m} p_i a_i = -1 < 0,$$

звідки для будь-якого $x_j \in \mathbf{R}$

$$(a_{1j}p_1 + a_{2j}p_2 + \cdots + a_{mj}p_m)\, x_j = 0 \quad (j = 1,2, \ldots, n),$$

і тому для будь-якого $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in V_n$

$$\sum_{j=1}^{n} (a_{1j}p_1 + a_{2j}p_2 + \cdots + a_{mj}p_m)\, x_j = \sum_{i=1}^{m} p_i (a_{i1}x_1 + a_{i2}x_2 +$$

$$+ \cdots + a_{in}x_n) = \sum_{i=1}^{m} p_i f_i(\boldsymbol{x}) = 0.$$

Отже, існують невід'ємні числа $p_1, p_2, \ldots, p_m$ такі, що справджується тотожне співвідношення (15) і нерівність (16). З викладеного випливає, що критерій несумісності системи лінійних нерівностей можна також сформулювати у вигляді такого твердження.

Теорема 2'. *Система лінійних нерівностей*

$$\{ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant a_i \quad (i = 1, 2, \ldots, m)$$

несумісна тоді і тільки тоді, коли система лінійних рівнянь

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m = 0 \quad (j = 1, 2, \ldots, n), \\ a_1y_1 + a_2y_2 + \cdots + a_my_m = -1 \end{cases}$$

має невід'ємний розв'язок.

Зауважимо, що теорему 2' можна довести й безпосередньо, без посилання на теорему 2, повторивши майже дослівно міркування, які ми проводили при доведенні теореми 2.

2.3. Невід'ємні розв'язки системи лінійних рівнянь. Другим важливим наслідком з теореми Мінковського є теорема про існування невід'ємних розв'язків системи лінійних рівнянь.

Нехай дано систему лінійних рівнянь

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m).$$

Теорема 3. *Система лінійних рівнянь*

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \qquad (i = 1, 2, \ldots, m) \tag{21}$$

має невід'ємний розв'язок тоді і тільки тоді, коли система нерівностей

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \leqslant 0 \quad (j = 1, 2, \ldots, n), \\ b_1y_1 + b_2y_2 + \cdots + b_my_m > 0 \end{cases} \tag{22}$$

несумісна.

Доведення. Припустимо, що система нерівностей (22) має розв'язок $y = (\delta_1, \delta_2, \ldots, \delta_m)$. Тоді справджуються нерівності

$$a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m \leqslant 0 \qquad (j = 1, 2, \ldots, n),$$

$$b_1\delta_1 + b_2\delta_2 + \cdots + b_m\delta_m > 0. \tag{23}$$

Покажемо, що система рівнянь (21) в цьому випадку невід'ємних розв'язків не має. Нехай система (21) має невід'ємний розв'язок $\boldsymbol{x} = (\eta_1, \eta_2, \ldots, \eta_n)$. Тоді справджуються рівності

$$(a_{i1}\eta_1 + a_{i2}\eta_2 + \cdots + a_{in}\eta_n)\,\delta_i = b_i\delta_i \qquad (i = 1, 2, \ldots, m),$$

а отже, і рівність

$$\sum_{i=1}^{m} (a_{i1}\eta_1 + a_{i2}\eta_2 + \cdots + a_{in}\eta_n)\,\delta_i = \sum_{i=1}^{m} b_i\delta_i,$$

яку можна записати ще так:

$$\sum_{j=1}^{n} (a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m)\,\eta_j = \sum_{i=1}^{m} \beta_i\delta_i.$$

Але, оскільки числа $\eta_1, \eta_2, \ldots, \eta_n$ — невід'ємні, то за нерівностями (23)

$$\sum_{j=1}^{n} (a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m)\,\eta_j \leqslant 0, \text{ а } \sum_{i=1}^{m} b_i\delta_i > 0$$

і тому

$$\sum_{j=1}^{n} (a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m)\,\eta \neq \sum_{i=1}^{m} \beta_i\delta_i.$$

Отже, припущення, що система (21) має невід'ємний розв'язок, приводить до суперечності і тому воно неправильне. Припустимо тепер, що система нерівностей (22) несумісна. Доведемо, що в цьому випадку система рівнянь (21) має невід'ємний розв'язок. Справді, оскільки система (22) несумісна, то кожен розв'язок $\boldsymbol{y} = (\gamma_1, \gamma_2, \ldots, \gamma_m)$ системи нерівностей [1]

$$\{a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \leqslant 0 \quad (j = 1, 2, \ldots, n) \tag{24}$$

не задовольняє нерівність $b_1y_1 + b_2y_2 + \cdots + b_my_m > 0$ і тому задовольняє нерівність

$$b_1y_1 + b_2y_2 + \cdots + b_my_m \leqslant 0. \tag{25}$$

[1] Система нерівностей (24) має принаймні нульовий розв'язок $y_1 = 0$, $y_2 = 0$, ..., $y_m = 0$.

Отже, нерівність (25) є наслідок системи нерівностей (24). Тому, за теоремою Мінковського, існують такі невід'ємні числа $p_1, p_2, \ldots, p_n$, що справджується тотожне відносно $y = (y_1, y_2, \ldots, y_m) \in V_n$ співвідношення

$$b_1y_1 + b_2y_2 + \cdots + b_my_m = \sum_{i=1}^{n} p_i (a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{im}y_m).$$

Звідси при $y_i = 1$ і $y_j = 0$ $(j \neq i)$ дістаємо рівності

$$a_{i1}p_1 + a_{i2}p_2 + \cdots + a_{in}p_n = b_i \quad (i = 1, 2, \ldots, m),$$

які означають, що система невід'ємних чисел $p_1, p_2, \ldots, p_n$ є розв'язок системи рівнянь (21). Теорему доведено.

2.4. Невід'ємні розв'язки системи лінійних нерівностей. Доведемо теорему про невід'ємні розв'язки системи нерівностей, на яку далі нам також доведеться посилатися.

Теорема 4. *Система лінійних нерівностей*

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = 1, 2, \ldots, m) \tag{26}$$

має невід'ємний розв'язок тоді і тільки тоді, коли система нерівностей

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant 0 \quad (j = 1, 2, \ldots, m), \\ b_1y_1 + b_2y_2 + \cdots + b_my_m < 0 \end{cases} \tag{27}$$

не має невід'ємного розв'язку.

Доведення. Нехай система (26) має невід'ємний розв'язок $x = (\gamma_1, \gamma_2, \ldots, \gamma_n)$. Тоді справджуються нерівності

$$a_{i1}\gamma_1 + a_{i2}\gamma_2 + \cdots + a_{in}\gamma_n \leqslant b_i \quad (i = 1, 2, \ldots, m). \tag{28}$$

Покажемо, що система (27) невід'ємного розв'язку не має.

Припустимо, що система (27) має невід'ємний розв'язок $y = (\eta_1, \eta_2, \ldots, \eta_m)$. Тоді справджуються нерівності

$$\begin{gathered} a_{1j}\eta_1 + a_{2j}\eta_2 + \cdots + a_{mj}\eta_m \geqslant 0 \quad (j = 1, 2, \ldots, n), \\ b_1\eta_1 + b_2\eta_2 + \cdots + b_m\eta_m < 0. \end{gathered} \tag{29}$$

Оскільки числа $\eta_1, \eta_2, \ldots, \eta_m$ — невід'ємні, то, з одного боку, з нерівностей (28) випливає справедливість нерівності

$$\sum_{i=1}^{m} (a_{i1}\gamma_1 + a_{i2}\gamma_2 + \cdots + a_{in}\gamma_n)\eta_i \leqslant \sum_{i=1}^{m} b_i\eta_i,$$

яку можна записати ще так:

$$\sum_{j=1}^{n} (a_{1j}\eta_1 + a_{2j}\eta_2 + \cdots + a_{mj}\eta_m)\gamma_j \leqslant \sum_{i=1}^{m} b_i\eta_i. \tag{30}$$

З другого боку, оскільки $\gamma_1, \gamma_2, \ldots, \gamma_n$ — невід'ємні числа, то з нерівностей (29) випливає, що

$$\sum_{j=1}^{n} (a_{1j}\eta_1 + a_{2j}\eta_2 + \cdots + a_{mj}\eta_m)\gamma_j \geqslant 0, \text{ а } \sum_{i=1}^{m} b_i\eta_i < 0$$

і тому нерівність (30) не може справджуватися. Отже, припущення, що система (27) має невід'ємний розв'язок, приводить до суперечності, і тому воно неправильне.

Припустимо тепер, що система (26) не має невід'ємного розв'язку. Доведемо, що в цьому випадку система нерівностей (27) має невід'ємний розв'язок. Справді, якщо система нерівностей (26) не має невід'ємного розв'язку, то система рівнянь

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + 0z_1 + \cdots + 0z_{i-1} + z_i + 0z_{i+1} + \cdots + 0z_m = b_i \quad (i = 1, 2, \ldots, m) \tag{31}$$

також не має невід'ємного розв'язку, бо якби система (31) мала невід'ємний розв'язок $x_1 = \beta_1, x_2 = \beta_2, \ldots, x_n = \beta_n, z_1 = \lambda_1, z_2 = \lambda_2, \ldots, z_m = \lambda_m$, то система невід'ємних чисел $\beta_1, \beta_2, \ldots, \beta_n$ задовольняла б систему нерівностей (26). Оскільки система рівнянь (31) не має невід'ємного розв'язку, то за теоремою 3 система нерівностей

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \leqslant 0 \quad (j = 1, 2, \ldots, n), \\ y_i \leqslant 0 \qquad (i = 1, 2, \ldots, m), \\ b_1y_1 + b_2y_2 + \cdots + b_my_m > 0 \end{cases}$$

має деякий розв'язок $y = (\gamma_1, \gamma_2, \ldots, \gamma_m)$; тут $\gamma_i \leqslant 0$ $(i = 1, 2, \ldots, m)$. Але тоді $y = (-\gamma_1, -\gamma_2, \ldots, -\gamma_m)$ є невід'ємний розв'язок системи нерівностей

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant 0 \qquad (j = 1, 2, \ldots, n), \\ y_i \geqslant 0 \qquad (i = 1, 2, \ldots, m), \\ b_1y_1 + b_2y_2 + \cdots + b_my_m < 0, \end{cases}$$

тобто $y = (-\gamma_1, -\gamma_2, \ldots, -\gamma_m)$ є невід'ємний розв'язок системи нерівностей (27). Теорему доведено.

Примітка. Теореми 2—4 не є такими ефективними критеріями, як критерії для систем лінійних рівнянь, про які йшлося в першій частині підручника в розділах VI, VII. Однак ці теореми встановлюють зв'язки між різними фактами (сумісність систем лінійних нерівностей, існування невід'ємних розв'язків систем лінійних рівнянь і систем лінійних нерівностей), що істотно використовуватиметься в дальшому викладі.

§ 3. ЗАДАЧІ ЛІНІЙНОГО ПРОГРАМУВАННЯ

3.1. Приклади задач лінійного програмування. Теорія лінійних нерівностей знаходить численні застосування; зокрема вона широко застосовується в галузі планово-економічних розрахунків. Здебільшого планово-економічні задачі — це екстремальні задачі на знаходження *найбільш вигідного варіанту*.

Розглянемо кілька прикладів таких задач.

Задача про використання сировини. Підприємство випускає продукцію n видів: $П_1, П_2, \ldots, П_n$, для виготовлення якої треба використовувати сировину m видів: $S_1, S_2, \ldots, S_m$. Припустимо, що запаси сировини кожного виду становлять відповідно $b_1, b_2, \ldots, b_m$ умовних

одиниць. Кількість одиниць сировини, яка необхідна для виготовлення одиниці продукції кожного з видів $П_j$, задається такою таблицею:

Види сировини	Запаси сировини	Кількість одиниць сировини, яка необхідна для виготовлення одиниці продукції			
		$П_1$	$П_2$	...	$П_n$
S_1	b_1	a_{11}	a_{12}	...	a_{1n}
S_2	b_2	a_{21}	a_{22}	...	a_{2n}
.	.	.	.	.	.
.	.	.	.	.	.
S_m	b_m	a_{m1}	a_{m2}	...	a_{mn}

Тут a_{ij} $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$ позначає кількість одиниць сировини S_i, яка необхідна для виготовлення одиниці продукції $П_j$. Відомо, що від реалізації одиниці продукції $П_j$ $(j = 1, 2, \ldots \ldots, n)$ підприємство одержує прибуток c_j крб. Треба скласти такий план випуску продукції, при якому прибуток підприємства від реалізації всієї продукції був би максимальним. Побудуємо математичну модель поставленої задачі. Припустимо, що підприємство випускає x_j одиниць продукції $П_j$. Тоді прибуток підприємства виражатиметься формулою

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n.$$

Загальна кількість сировини S_i $(i = 1, 2, \ldots, m)$, що використовується при виготовленні продукції всіх видів, дорівнює $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$. Вона не може перевищувати всього запасу цієї сировини b_i. Отже, повинні виконуватися нерівності

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = 1, 2, \ldots, m).$$

З самого смислу величин x_i очевидно, що $x_i \geqslant 0$. Таким чином, математично задачу можна сформулювати так.

Дано систему лінійних нерівностей

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leqslant b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leqslant b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leqslant b_m, \\ x_i \geqslant 0, \qquad (i = 1, 2, \ldots, n) \end{cases} \tag{1}$$

і лінійну функцію

$$f(x_1, x_2, \ldots, x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n. \tag{2}$$

Треба серед розв'язків системи нерівностей (1) вибрати такий, при якому лінійна функція $f(x_1, x_2 \ldots, x_n)$ набуває найбільшого значення (максимізується), або, інакше кажучи, треба знайти максимум лінійної функції (2) в області розв'язків системи нерівностей (1).

Транспортна задача. В m родовищах $P_1, P_2, \ldots, P_m$ щомісяця видобувають відповідно $a_1, a_2, \ldots, a_m$ тонн вугілля. Все це вугілля треба вивозити в пункти споживання $П_1, П_2, \ldots, П_s$, причому в кожен з цих пунктів щомісяця повинно завозитися відповідно $b_1, b_2, \ldots, b_s$ тонн. Вартість перевезення однієї тонни вугілля з родовища P_i в пункт $П_j$ $(i = 1,2,\ldots, m;\ j = 1,2, \ldots, s)$ дорівнює c_{ij} крб. Необхідно скласти такий план перевезень, при якому загальна вартість їх була б найменшою. Всі дані цієї задачі запишемо у вигляді такої таблиці:

Пункти споживання / Шахти	Вартість перевезень 1-ї тонни				Видобуток
	$П_1$	$П_2$	...	$П_m$	
P_1	c_{11}	c_{12}	.	c_{1m}	a_1
P_2	c_{21}	c_{22}	...	c_{2m}	a_2
.	.	.	.	.	
.	.	.	.	.	
P_s	c_{s1}	c_{s2}	...	c_{sm}	a_s
Потреба у вугіллі	b_1	b_2	...	b_m	

Зауважимо, що, оскільки все вугілля, яке видобувається в родовищах $P_1, P_2, \ldots, P_s$, повинно вивозитися в пункти $П_1, П_2, \ldots, П_m$, то

$$a_1 + a_2 + \cdots + a_s = b_1 + b_2 + \cdots + b_m.$$

Позначимо через x_{ij} $(i = 1, 2, \ldots, s;\ j = 1,2, \ldots, m)$ кількість вугілля в тоннах, що вивозиться з родовища P_i в пункт $П_j$, а через f — вартість перевезення всього вугілля. Складемо схему перевезень.

Пункти споживання / Шахти	$П_1$	$П_2$	...	$П_m$	Всього вивезено
P_1	x_{11}	x_{12}	...	x_{1m}	a_1
P_2	x_{21}	x_{22}	...	x_{2m}	a_2
.	.	.	.	.	.
.	.	.	.	.	.
P_s	x_{s1}	x_{s2}	...	x_{sm}	a_s
Всього привезено	b_1	b_2	...	b_m	

Загальна кількість вугілля, що завозиться з усіх родовищ в пункт $П_j$, дорівнює $x_{1j} + x_{2j} + \cdots + x_{sj}$. Оскільки в пункт $П_j$ має вивозитися b_j тонн вугілля, то

$$x_{1j} + x_{2j} + \cdots + x_{sj} = b_j \quad (j = 1, 2, \ldots, m).$$

Загальна кількість вугілля, що вивозиться в усі пункти з родовища P_i, дорівнює $x_{i1} + x_{i2} + \cdots + x_{im}$.

З умови задачі випливає, що

$$x_{i1} + x_{i2} + \cdots + x_{im} = a_i \quad (i = 1, 2, \ldots, s).$$

Вартість f перевезення всього вугілля, очевидно, виражається формулою

$$f = c_{11}x_{11} + c_{12}x_{12} + \cdots + c_{1m}x_{1m} + c_{21}x_{21} + c_{22}x_{22} + \cdots + c_{2m}x_{2m} + \cdots + c_{s1}x_{s1} + c_{s2}x_{s2} + \cdots + c_{sm}x_{sm}.$$

Таким чином, ми приходимо до такої математичної задачі.

Задано систему лінійних рівнянь

$$\begin{cases} x_{1j} + x_{2j} + \cdots + x_{sj} = b_j \quad (j = 1, 2, \ldots, m), \\ x_{i1} + x_{i2} + \cdots + x_{im} = a_i \quad (i = 1, 2, \ldots, s) \end{cases} \tag{3}$$

і лінійну функцію

$$f = \sum_{j=1}^{m} \sum_{i=1}^{s} c_{ij}x_{ij}.$$

Треба серед всіх невід'ємних розв'язків системи рівнянь (3) вибрати такий, при якому функція f набуває найменшого значення (мінімізується).

3.2. Різні форми задачі лінійного програмування. Розглянуті нами задачі є прикладами так званих задач лінійного програмування. У цих задачах з математичної точки зору треба відшукати невід'ємні значення кількох невідомих (змінних), які задовольняли б деяку систему лінійних рівнянь або нерівностей, і при яких деяка лінійна функція від цих змінних набувала б мінімального чи максимального значення. У задачі про сировину шукані значення невідомих повинні задовольняти деяку систему нерівностей, а в транспортній — систему рівнянь. Проте зустрічаються задачі практичного характеру, подібні до розглянутих, в яких шукані значення невідомих повинні задовольняти і рівняння і нерівності, тобто систему рівнянь і нерівностей. Не завжди також у подібних задачах мають бути невід'ємними шукані значення всіх невідомих, значення деяких з них, а іноді й усіх, можуть бути від'ємними. Всі ці окремі випадки охоплюються таким загальним формулюванням задачі лінійного програмування: *задана система рівнянь і нерівностей*

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, s), \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = s+1, s+2, \ldots, m), \\ x_j \geqslant 0 \qquad (j = 1, 2, \ldots, t; \quad t \leqslant n) \end{cases} \tag{4}$$

і лінійна функція

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n; \tag{5}$$

треба знайти розв'язок системи рівнянь і нерівностей (4), *при якому функція f максимізується, тобто досягає максимуму (або мінімізується, тобто досягає мінімуму).*

Сформульовану задачу називають *загальною задачею лінійного програмування.*

Систему рівнянь і нерівностей (4) називають *системою обмежень*, або просто *обмеженнями* даної задачі. Лінійну функцію (5) називають *цільовою функцією даної задачі.*

Якщо в системі обмежень загальної задачі $s = 0$, а $t = n$, то загальну задачу запишемо так: *дана система нерівностей*

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = 1, 2, \ldots, m) \tag{6}$$

і цільова функція

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n;$$

треба знайти невід'ємний розв'язок системи рівнянь (6), *тобто розв'язок, що задовольняє нерівності $x_1 \geqslant 0$, $x_2 \geqslant 0$, ..., $x_n \geqslant 0$, при якому цільова функція f максимізується (або мінімізується).* Цю задачу називають *стандартною задачею* лінійного програмування[1].

Якщо в загальній задачі $s = m$, а $t = n$, то її записують так: *дана система рівнянь*

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m) \tag{7}$$

і цільова функція

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n;$$

треба знайти невід'ємний розв'язок системи рівнянь (7), *який максимізує (мінімізує) цільову функцію f.* Задачу лінійного програмування в такому формулюванні називають *канонічною.*

Зауважимо, що система обмежень стандартної задачі, точно кажучи, складається з системи нерівностей (6) і нерівностей $x_1 \geqslant 0$, $x_2 \geqslant 0$, ..., $x_n \geqslant 0$, а канонічної — з системи рівнянь (7) і нерівностей $x_1 \geqslant 0$, $x_2 \geqslant 0$, ..., $x_n \geqslant 0$. Проте в стандартній і канонічних задачах умови $x_1 \geqslant 0$, $x_2 \geqslant 0$, ..., $x_n \geqslant 0$ не вважають обмеженнями, оскільки вони не є характерними для даної стандартної чи канонічної задачі, а є загальними для всіх стандартних і канонічних задач. Під системою обмежень стандартної задачі розуміють систему нерівностей (6), а канонічної — систему рівнянь (7).

Умовимося вектор $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$, всі компоненти якого невід'ємні: $a_1 \geqslant 0$, $a_2 \geqslant 0$, ..., $a_n \geqslant 0$, називати *невід'ємним*, і записуватимемо $\boldsymbol{a} \geqslant \boldsymbol{0}$, де $0 = (0, 0, \ldots, 0)$. Вектор $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ вважатимемо більшим від вектора $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$, якщо $a_1 \geqslant b_1$, $a_2 \geqslant b_2$, ..., $a_n \geqslant b_n$, і записуватимемо $\boldsymbol{a} \geqslant \boldsymbol{b}$.

Вектор $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, що задовольняє обмеження даної задачі лінійного програмування, зокрема й обмеження $\boldsymbol{x} \geqslant 0$ в стандартній і канонічній задачі, називають *допустимим розв'язком* або *допустимим вектором* цієї задачі. Задачу, що має допустимий розв'язок,

[1] Задачу лінійного програмування, в якій системою обмежень є система нерівностей виду

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geqslant b_i \ (i = 1, 2, \ldots, n),$$

також називають *стандартною.*

називають *допустимою*. Допустимий розв'язок, який мінімізує (максимізує) цільову функцію, називають *оптимальним розв'язком* або *оптимальним вектором*. Значення мінімуму (відповідно максимуму) цільової функції називають *значенням задачі лінійного програмування*. Зауважимо, що оптимальний розв'язок задачі лінійного програмування (якщо він взагалі існує) не обов'язково єдиний. Може трапитися так, що задача матиме нескінченну множину оптимальних розв'язків. Задачу лінійного програмування, в якій ідеться про відшукання допустимого розв'язку, що максимізує цільову функцію, називають *задачею максимізації*, а задачу, в якій треба знайти допустимий розв'язок, який мінімізує цільову функцію, називають *задачею мінімізації*. Оскільки $\max f = -\min(-f)$ і $\min f = -\max(-f)$, то одна з цих задач, очевидно, зводиться до другої заміною цільової функції f на $-f$.

З'ясуємо тепер питання про взаємозв'язок між різними формами задачі лінійного програмування. Насамперед покажемо, що канонічну задачу можна перевести в стандартну і навпаки. Справді, рівняння $a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n = b_s$, очевидно, рівносильне системі нерівностей

$$\begin{cases} a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n \leqslant b_s, \\ -a_{s1}x_1 - a_{s2}x_2 - \cdots - a_{sn}x_n \leqslant -b_s. \end{cases}$$

Якщо в канонічній задачі кожне рівняння

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$$

системи (7) замінимо рівносильною йому системою нерівностей

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i, \\ -a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n \leqslant -b_i, \end{cases}$$

то дістанемо стандартну задачу.

З другого боку, нерівність $a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n \leqslant b_k$, очевидно, рівносильна рівнянню $a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n + x_{n+k} = b_k$, де x_{n+k} — додаткове невідоме, що задовольняє умову $x_{n+k} \geqslant 0$. Якщо в системі обмежень (6) стандартної задачі кожну нерівність $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i$ замінимо рівносильним їй рівнянням

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + x_{n+i} = b_i,$$

ввівши додаткові невід'ємні невідомі $x_{n+1}, x_{n+2}, \ldots, x_{n+m}$, то стандартна задача набуде вигляду канонічної задачі. Проілюструємо це на прикладі розглянутої вище задачі про сировину. Ця задача записана у формі стандартної задачі. Її система обмежень така:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leqslant b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leqslant b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leqslant b_m. \end{cases}$$

Замінимо кожну нерівність цієї системи рівносильним рівнянням, ввівши додаткові невід'ємні невідомі $x_{n+1}, x_{n+2}, \ldots, x_{n+m}$. Тоді система обмежень задачі набуде вигляду

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + x_{n+1} = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + x_{n+2} = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + x_{n+m} = b_m. \end{cases} \tag{8}$$

Задачу про сировину тепер, очевидно, можна формулювати так: треба знайти невід'ємний розв'язок системи рівнянь (8), який максимізує функцію

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n + 0 \cdot x_{n+1} + 0 \cdot x_{n+2} + \cdots + 0 \cdot x_{n+m}.$$

Зрозуміло, що в шуканому розв'язку нас цікавлять лише значення $x_1, x_2, \ldots, x_n$.

Стандартна і канонічна задачі, як відомо, є окремими випадками загальної задачі. Покажемо, що, навпаки, загальну задачу можна записати у вигляді стандартної, а отже, і канонічної.

Для того щоб загальну задачу записати у вигляді стандартної, замінимо в системі обмежень (4) кожне рівняння $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ рівносильною системою нерівностей

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i, \\ -a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n \leqslant -b_i. \end{cases}$$

Тоді система обмежень (4) загальної задачі набере вигляду

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = 1, 2, \ldots, s), \\ -a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n \leqslant -b_i \quad (i = 1, 2, \ldots, s), \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = s+1, s+2, \ldots, m), \\ x_j \geqslant 0 \quad (j = 1, 2, \ldots, t;\ t \leqslant n). \end{cases} \tag{9}$$

Введемо тепер нові невід'ємні невідомі $y_{t+1}, y_{t+2}, \ldots, y_n$ і $z_{t+1}, z_{t+2}, \ldots, z_n$, які пов'язані з невідомими $x_{t+1}, x_{t+2}, \ldots, x_n$ співвідношеннями

$$x_{t+1} = y_{t+1} - z_{t+1},\ x_{t+2} = y_{t+2} - z_{t+2},\ \ldots,\ x_n = y_n - z_n.$$

Підставимо вирази невідомих $x_{t+1}, x_{t+2}, \ldots, x_n$ через нові невідомі в систему обмежень (9); тоді система обмежень набуде вигляду

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{it}x_t + a_{i\,t+1}(y_{t+1} - z_{t+1}) + \cdots + \\ + a_{in}(y_n - z_n) \leqslant b_i \quad (i = 1, 2, \ldots, s), \\ -a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{it}x_t - a_{i\,t+1}(y_{t+1} - z_{t+1}) - \cdots - \\ - a_{in}(y_n - z_n) \leqslant -b_i \quad (i = 1, 2, \ldots, s), \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{it}x + a_{it+1}(y_{t+1} - z_{t+1}) + \cdots + \\ + a_{in}(y_n - z_n) \leqslant b_i \quad (i = s+1, \ldots, m), \end{cases} \tag{10}$$

$$\begin{cases} x_j \geqslant 0 \quad (j = 1, 2, \ldots, t;\ t \leqslant n); \\ y_r \geqslant 0,\ z_r \geqslant 0 \quad (r = t+1, t+2, \ldots, n), \end{cases} \tag{11}$$

а цільова функція — вигляду

$$f = c_1x_1 + c_2x_2 + \cdots + c_tx_t + c_{t+1}(y_{t+1} - z_{t+1}) + \cdots + c_n(y_n - z_n). \quad (12)$$

Легко бачити, що невід'ємний вектор

$$w^\circ = (x_1^\circ, x_2^\circ, \ldots, x_t^\circ, y_{t+1}^\circ, y_{t+2}^\circ, \ldots, y_n^\circ, z_{t+1}^\circ, z_{t+2}^\circ, \ldots, z_n^\circ)$$

тоді і тільки тоді задовольняє систему нерівностей (10) і (11) й максимізує (мінімізує) функцію (12), коли вектор $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ задовольняє систему обмежень (4) і максимізує (мінімізує) цільову функцію (5).

Отже, загальна задача зводиться до такої стандартної задачі: дана система нерівностей (10) — (11) і цільова функція (12); треба знайти невід'ємний вектор

$$w = (x_1, x_2, \ldots, x_t, y_{t+1}, y_{t+2}, \ldots, y_n, z_{t+1}, z_{t+2}, \ldots, z_n),$$

який задовольняє систему нерівностей (10) і максимізує (мінімізує) цільову функцію (12).

3.3. **Геометричний смисл задачі лінійного програмування.** Кожну задачу лінійного програмування можна звести до стандартної задачі. Стандартну ж задачу легко інтерпретувати геометрично.

Нехай $\mathfrak{A}_n$ — n-вимірний евклідів афінний простір[1]. Як показано в п. 1.3, областю невід'ємних розв'язків системи обмежень (6) стандартної задачі є деяка опукла многогранна область (многогранник) M. Вона складається з усіх точок простору $\mathfrak{A}_n$, координати яких невід'ємні й задовольняють систему нерівностей (6).

Значення цільової функції

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

в точці $x^\circ(x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ тобто $f_0 = c_1x_1^\circ + c_2x_2^\circ + \cdots + c_nx_n^\circ$, можна розглядати як відхилення точки $x^\circ(x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ від гіперплощини

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0. \quad (13)$$

Відхилення точки $X(x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ від гіперплощини (13) в такому розумінні пропорціональне віддалі від точки X до цієї гіперплощини. Таким чином, з геометричної точки зору стандартна задача лінійного програмування полягає у знаходженні в многогранній області (многограннику) M точки, яка найбільше (найменше) відхилена від гіперплощини (13).

Якщо в стандартній задачі число невідомих $n = 2$, то областю невід'ємних розв'язків системи обмежень (6) є деяка опукла многокутна область (многокутник) Q, розташована в першому квадранті координатної площини, а рівняння $c_1x_1 + c_2x_2 = 0$ є рівняння деякої прямої, що проходить через початок координат. Рисунки 3—6 ілюструють різні випадки, які при цьому можуть мати місце. На цих рисунках стрілки на сторонах многокутників показують півплощини, що визначаються відповідними нерівностями системи обмежень, а стрілки на прямих $c_1x_1 + c_2x_2 = 0$ показують ті півплощини, в точках яких справджується нерівність $c_1x_1 + c_2x_2 \geqslant 0$. З геометричних міркувань

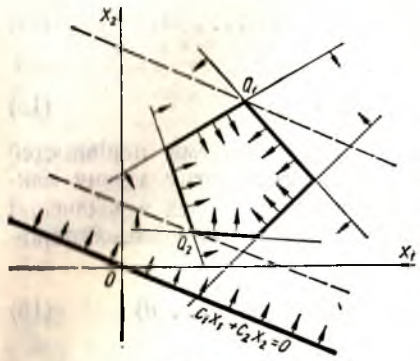


Рис. 3.

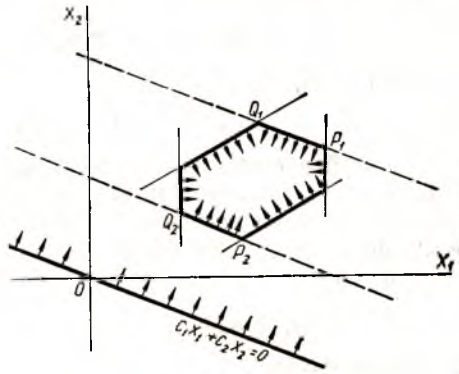


Рис. 4.

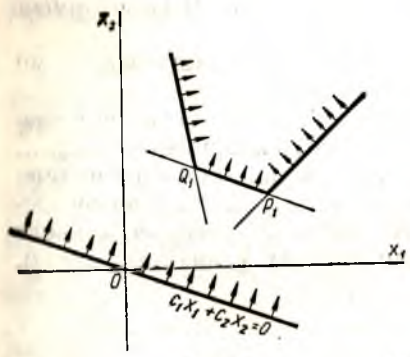


Рис. 5.

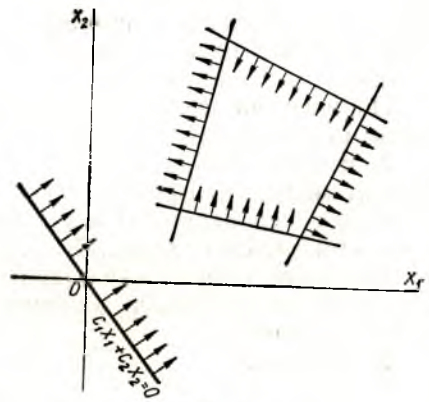


Рис. 6.

зрозуміло, що оптимальним розв'язком стандартної задачі буде, взагалі кажучи, якась вершина (її координати) многокутника Q. У випадку, що ілюструється рис. 3, оптимальним розв'язком задачі максимізації є вершина Q_1, а мінімізації — Q_2, причому ці розв'язки єдині. У випадку, що ілюструється рис. 4, і задача максимізації, і задача мінімізації мають нескінченну множину оптимальних розв'язків: кожна точка відрізка Q_1P_1 є оптимальним розв'язком задачі максимізації, а кожна точка відрізка Q_2P_2 — задачі мінімізації. Рисунок 5 ілюструє випадок, коли задача мінімізації має безліч оптимальних розв'язків (кожна точка відрізка Q_1P_1 є оптимальним розв'язком

[1] n-вимірним евклідовим афінним простором називають афінний простір $\mathfrak{A}_n$, якому відповідає евклідів векторний простір E_n.

задачі мінімізації), а задача максимізації оптимального розв'язку не має, бо цільова функція необмежена.

Рисунок 6 ілюструє випадок, коли ні задача максимізації, ні задача мінімізації розв'язків не мають, оскільки система обмежень несумісна.

3.4. Взаємно двоїсті задачі лінійного програмування. Нехай дано систему лінійних нерівностей

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \ (i = 1, 2, \ldots, m) \tag{14}$$

і лінійну форму

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n; \tag{15}$$

треба знайти розв'язок $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \geqslant 0$ системи нерівностей (14), який максимізує лінійну форму (15). Цю стандартну задачу максимізації будемо називати *вихідною* або задачею I. Поряд з задачею I розглянемо таку стандартну задачу мінімізації: дана система нерівностей

$$\{a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant c_j \ (j = 1, 2, \ldots, n) \tag{16}$$

і лінійна форма

$$\varphi = b_1y_1 + b_2y_2 + \cdots + b_my_m; \tag{17}$$

треба знайти розв'язок $\boldsymbol{y} = (y_1, y_2, \ldots, y_m) \geqslant 0$ системи нерівностей (16), який мінімізує лінійну форму (17). Цю задачу ми називатимемо *двоїстою* відносно вихідної або задачею I′.

Порівнюючи задачі I і I′, легко помітити, що двоїсту задачу ми дістаємо з вихідної задачі за такими правилами:

1. Вільні члени $b_1, b_2, \ldots, b_m$ обмежень (14) вихідної задачі є коефіцієнтами цільової функції (17) двоїстої задачі, а коефіцієнти $c_1, c_2, \ldots, c_n$ цільової функції (15) вихідної задачі є вільними членами обмежень (16) двоїстої задачі. Отже, число змінних у двоїстій задачі дорівнює числу обмежень вихідної задачі, тобто дорівнює m, а число обмежень двоїстої задачі дорівнює числу змінних у вихідній задачі, тобто дорівнює n.

2. Матриця

$$A' = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}.$$

коефіцієнтів обмежень (16) двоїстої задачі одержується з матриці

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

коефіцієнтів обмежень (14) вихідної задачі транспонуванням останньої.

3. Система обмежень вихідної задачі складається з нерівностей типу $\leqslant$, а система обмежень двоїстої задачі — з нерівностей протилежного типу $\geqslant$.

4. Максимізація цільової функції (15) вихідної задачі замінюється мінімізацією цільової функції (17) двоїстої задачі.

Будемо тепер вважати вихідною не задачу I, а задачу I′. Запишемо її так, як була записана задача I, тобто щоб система обмежень її складалася з нерівностей типу $\leqslant$ й щоб у задачі йшлося про максимізацію цільової функції. Для цього, очевидно, достатньо кожну нерівність системи (16) помножити на —1, а задачу мінімізації функції φ замінити рівносильною задачею максимізації функції $-\varphi$, оскільки $\min \varphi = -\max(-\varphi)$.

Отже, запишемо задачу I′ так: дана система нерівностей

$$\{-a_{1j}y_1 - a_{2j}y_2 - \cdots - a_{mj}y_m \leqslant -c_j \ (j = 1, 2, \ldots, n) \tag{18}$$

і лінійна форма

$$-\varphi = -b_1y_1 - b_2y_2 - \cdots - b_my_m; \tag{19}$$

треба знайти розв'язок системи нерівностей (18), який максимізує лінійну форму (19).

Для записаної таким способом задачі I′ сформулюємо, згідно з правилами 1—4, двоїсту їй задачу: дана система нерівностей

$$\{-a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n \geqslant -b_i \ (i = 1, 2, \ldots, m) \tag{20}$$

і лінійна форма

$$-c_1x_1 - c_2x_2 - \cdots - c_nx_n = -f; \tag{21}$$

треба знайти розв'язок $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \geqslant 0$ системи нерівностей (20), який мінімізує лінійну форму (21).

Якщо в цій задачі кожну нерівність системи обмежень (20) помножити на —1, а задачу мінімізації функції $-f$ замінити рівносильною їй задачею максимізації функції f, то матимемо задачу I. Отже, двоїстою задачею для задачі I′ є задача I. Як бачимо, відношення двоїстості є взаємним: задача I′ двоїста задачі I, а задача I двоїста задачі I′. Тому задачі I и I′ називають *взаємно двоїстими*. Доведемо одне твердження, що стосується взаємно двоїстих задач, на яке далі нам доведеться посилатися.

Лема. *Якщо* $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ *є допустимий розв'язок задачі* I *(тобто невід'ємний розв'язок системи нерівностей* (14)*), а* $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ *— допустимий розв'язок двоїстої задачі* I′ *(невід'ємний розв'язок системи нерівностей* (16)*), то*

$$c_1x_1^\circ + c_2x_2^\circ + \cdots + c_nx_n^\circ \leqslant b_1y_1^\circ + b_2y_2^\circ + \cdots + b_my_m^\circ.$$

Доведення. Оскільки $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ є розв'язки відповідно систем нерівностей (14) і (16), то справджуються нерівності

$$a_{i1}x_1^\circ + a_{i2}x_2^\circ + \cdots + a_{in}x_n^\circ \leqslant b_i \ (i = 1, 2, \ldots, m) \tag{21}$$

і

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ \geqslant c_j \ (j = 1, 2, \ldots, n). \qquad (22)$$

Помножимо кожну з нерівностей (21) на невід'ємне число y_i° і потім всі одержані нерівності додамо. Дістанемо нерівність

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_j^\circ y_i^\circ \leqslant \sum_{i=1}^{m} b_i y_i^\circ. \qquad (23)$$

Аналогічно, помноживши кожну з нерівностей (22) на невід'ємне число x_j° і додавши всі одержані нерівності, дістанемо нерівність

$$\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}y_i^\circ x_j^\circ \geqslant \sum_{j=1}^{n} c_j x_j^\circ. \qquad (24)$$

Ліві частини нерівностей (23) і (24), очевидно, однакові. Тому

$$\sum_{j=1}^{n} c_j x_j^\circ \leqslant \sum_{i=1}^{m} b_i y_i^\circ.$$

Лему доведено.

З доведеної леми безпосередньо випливає справедливість такого твердження.

Теорема 1 (критерій оптимальності). *Якщо існують такі допустимі розв'язки* $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ *і* $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ *задачі* I *і двоїстої їй задачі* I′, *що*

$$c_1x_1^\circ + c_2x_2^\circ + \cdots + c_nx_n^\circ = b_1y_1^\circ + b_2y_2^\circ + \cdots + b_my_m^\circ,$$

то ці допустимі розв'язки є оптимальними розв'язками відповідних задач.

Доведення. Припустимо, що $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ не є оптимальним розв'язком задачі I. Тоді існує допустимий розв'язок $\boldsymbol{x}' = (x_1', x_2', \ldots, x_n')$ задачі I′ такий, що

$$c_1x_1^\circ + c_2x_2^\circ + \cdots + c_nx_n^\circ < c_1x_1' + c_2x_2' + \cdots + c_nx_n'.$$

Але за лемою 1

$$c_1x_1' + c_2x_2' + \cdots + c_nx_n' \leqslant b_1y_1^\circ + b_2y_2^\circ + \cdots + b_my_m^\circ,$$

тому

$$c_1x_1^\circ + c_2x_2^\circ + \cdots + c_nx_n^\circ < b_1y_1^\circ + b_2y_2^\circ + \cdots + b_my_m^\circ,$$

а це суперечить умові теореми. Отже, припущення, що $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ не є оптимальним розв'язком задачі I, неправильне. Аналогічними міркуваннями доводиться, що $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ є оптимальний розв'язок задачі I′.

Для взаємно двоїстих стандартних задач лінійного програмування справедлива така теорема.

Теорема 2 (теорема двоїстості). *Якщо кожна з двох взаємно двоїстих стандартних задач допустима, то обидві вони мають оптимальні розв'язки і значення їх збігаються. Якщо хоч одна з цих задач не є допустимою, то жодна з них не має оптимального розв'язку.*

Доведення. Припустимо, що і стандартна задача максимізації I і двоїста їй задача I′ допустимі, тобто що системи нерівностей

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \ (i = 1, 2, \ldots, m) \qquad (14')$$

і

$$a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant c_j \ (j = 1, 2, \ldots, n) \qquad (16')$$

мають невід'ємні розв'язки. Доведемо, що задачі I і I′ мають оптимальні розв'язки і значення їх збігаються. Для цього достатньо показати, що існують невід'ємні розв'язки $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ систем (14′) і (16′), які задовольняють також нерівність

$$\sum_{j=1}^{n} c_jx_j - \sum_{i=1}^{m} b_iy_i \geqslant 0. \qquad (25)$$

Справді, якщо невід'ємні розв'язки $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ систем (14′) і (16′) задовольняють нерівність (25), то

$$\sum_{j=1}^{n} c_jx_j \geqslant \sum_{i=1}^{m} b_iy_i^\circ; \qquad (26)$$

з другого боку, за лемою,

$$\sum_{j=1}^{n} c_jx_j^\circ \leqslant \sum_{i=1}^{m} b_iy_i^\circ. \qquad (27)$$

З нерівностей (26) і (27) випливає, що $\sum_{j=1}^{n} c_jx_j^\circ = \sum_{i=1}^{m} b_iy_i^\circ$. Тому, за теоремою 1, допустимі розв'язки $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ задач I і I′ — оптимальні, а значення цих задач $\sum_{j=1}^{n} c_jx_j^\circ$ і $\sum_{i=1}^{m} b_iy_i^\circ$ збігаються. Отже, покажемо, що системи нерівностей (14′) і (16′) мають невід'ємні розв'язки, які задовольняють також нерівність (25), тобто що система нерівностей

$$\begin{cases} \sum_{j=1}^{n} a_{ij}x_j \leqslant b_i \ (i = 1, 2, \ldots, m), \\ \sum_{i=1}^{m} a_{ij}y_i \geqslant c_j \ (j = 1, 2, \ldots, n), \\ \sum_{j=1}^{n} c_jx_j - \sum_{i=1}^{m} b_iy_i \geqslant 0 \end{cases} \qquad (28)$$

має невід'ємний розв'язок.

Припустимо, що система (28) не має невід'ємних розв'язків; тоді не має невід'ємних розв'язків також і рівносильна їй система нерівностей

$$\begin{cases} \sum_{j=1}^{n} a_{ij}x_j \leqslant b_i \quad (i = 1, 2, \ldots, m), \\ \sum_{i=1}^{m} (-a_{ji})\, y_i \leqslant -c_j \quad (j = 1, 2, \ldots, n), \\ \sum_{j=1}^{n} (-c_j)\, x_j + \sum_{i=1}^{m} b_i y_i \leqslant 0. \end{cases}$$

Тому за теоремою 4, п. 2.4, знайдеться такий невід'ємний вектор $t = (\zeta_1, \zeta_2, \ldots, \zeta_m, \omega_1, \omega_2, \ldots, \omega_n, \theta)$, що справджуватимуться нерівності

$$a_{1j}\zeta_1 + a_{2j}\zeta_2 + \cdots + a_{mj}\zeta_m - c_j\theta \geqslant 0 \quad (j = 1, 2, \ldots, n),$$
$$(-a_{i1})\,\omega_1 + (-a_{i2})\,\omega_2 + \cdots + (-a_{in})\,\omega_n + b_i\theta \geqslant 0 \quad (i = 1, 2, \ldots, m),$$
$$b_1\zeta_1 + b_2\zeta_2 + \cdots + b_m\zeta_m - c_1\omega_1 - c_2\omega_2 - \cdots - c_n\omega_n < 0,$$

тобто нерівності

$$\sum_{i=1}^{m} a_{ij}\zeta_i \geqslant c_j\theta \quad (j = 1, 2, \ldots, n), \tag{29}$$

$$\sum_{j=1}^{n} a_{ij}\omega_j \leqslant b_i\theta \quad (i = 1, 2, \ldots, m), \tag{30}$$

$$\sum_{i=1}^{m} b_i\zeta_i - \sum_{j=1}^{n} c_j\omega_j < 0. \tag{31}$$

Невід'ємне число θ не дорівнює 0. Справді, нехай $\theta = 0$. За припущенням, системи нерівностей (14′) і (16′) мають невід'ємні розв'язки; нехай $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $y^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ деякі з таких розв'язків.

Помноживши кожну з нерівностей (29) на невід'ємне число x_j і додавши одержані нерівності, дістанемо нерівності

$$0 \leqslant \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}\zeta_i x_j^\circ = \sum_{i=1}^{m} \zeta_i \sum_{j=1}^{n} a_{ij}x_j^\circ \leqslant \sum_{i=1}^{m} b_i\zeta_i, \tag{32}$$

оскільки $\sum_{j=1}^{n} a_{ij}x_j^\circ < b_i$ й $\zeta_i \geqslant 0$.

Помноживши кожну з нерівностей (30) на невід'ємне число y_i° й додавши одержані нерівності, дістанемо нерівності

$$0 > \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\omega_j y_i^\circ = \sum_{j=1}^{n} \omega_j \sum_{i=1}^{m} a_{ij}y_i^\circ > \sum_{j=1}^{n} c_j\omega_j, \tag{33}$$

бо

$$\sum_{j=1}^{n} a_{ij}y_i^\circ > c_i \text{ й } \omega_j \geqslant 0.$$

З нерівностей (32) і (33) випливає справедливість нерівності

$$\sum_{i=1}^{m} b_i\zeta_i - \sum_{j=1}^{n} c_j\omega_j > 0,$$

а це суперечить нерівності (31). Тому припущення, що $\theta = 0$, неправильне і, отже, $\theta > 0$.

Оскільки $\theta > 0$, то, як випливає з нерівностей (30) і (29), вектори $\frac{1}{\theta}w = \left(\frac{\omega_1}{\theta}, \frac{\omega_2}{\theta}, \ldots, \frac{\omega_n}{\theta}\right)$ і $\frac{1}{\theta}z = \left(\frac{\zeta_1}{\theta}, \frac{\zeta_2}{\theta}, \ldots, \frac{\zeta_m}{\theta}\right)$ є допустимі розв'язки відповідно задачі I і I′. Тому за лемою $c_1\frac{\omega_1}{\theta} + c_2\frac{\omega_2}{\theta} + \cdots + c_n\frac{\omega_n}{\theta} \leqslant b_1\frac{\zeta_1}{\theta} + b_2\frac{\zeta_2}{\theta} + \cdots + b_m\frac{\zeta_m}{\theta}$, тобто

$$\sum_{i=1}^{m} b_i\zeta_i - \sum_{j=1}^{n} c_j\omega_j \geqslant 0,$$

що знову суперечить нерівності (31). Таким чином, припущення, що система нерівностей (28) не має невід'ємних розв'язків, приводить до суперечності, і тому воно неправильне. Отже, система нерівностей (28) має невід'ємний розв'язок, задачі I і I′ мають оптимальні розв'язки й їх значення збігаються. Припустимо тепер, що одна з задач, наприклад задача I, не є допустимою, тобто що система нерівностей (14′) не має невід'ємного розв'язку. В такому разі ця задача безперечно не має оптимального розв'язку.

Якщо задача I′ не є допустимою, то вона також не має оптимального розв'язку. Нехай задача I′ допустима. Оскільки система нерівностей (14′) не має невід'ємних розв'язків, то за теоремою 4, п. 42.4 існує такий невід'ємний вектор $z = (\zeta_1, \zeta_2, \ldots, \zeta_m)$, що справджуються нерівності

$$a_{1j}\zeta_1 + a_{2j}\zeta_2 + \cdots + a_{mj}\zeta_m \geqslant 0 \quad (j = 1, 2, \ldots, n), \tag{33′}$$
$$b_1\zeta_1 + b_2\zeta_2 + \cdots + b_m\zeta_m < 0. \tag{34}$$

З другого боку, оскільки задача I′ допустима, існує такий невід'ємний вектор $y^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$, що справджуються нерівності

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ \geqslant c_j \quad (j = 1, 2, \ldots, n). \tag{35}$$

З нерівностей (35) і (33′) випливає, що для будь-якого додатного числа λ вектор

$$y^\circ + \lambda z = (y_1^\circ + \lambda\zeta_1, y_2^\circ + \lambda\zeta_2, \ldots, y_m^\circ + \lambda\zeta_m)$$

є допустимим розв'язком задачі I′. Але, за нерівністю (34)

$$b_1\zeta_1 + b_2\zeta_2 + \cdots + b_m\zeta_m < 0.$$

Тому вираз

$$b_1(y_1^\circ + \lambda\zeta_1) + b_2(y_2^\circ + \lambda\zeta_2) + \cdots + b_m(y_m^\circ + \lambda\zeta_m) = (b_1y_1^\circ + b_2y_2^\circ + \cdots + b_my_m^\circ) + \lambda(b_1\zeta_1 + b_2\zeta_2 + \cdots + b_m\zeta_m)$$

може стати як завгодно малим і тому він не має мінімуму. Отже, задача I′ не має оптимального розв'язку. Теорему доведено.

На цьому ми закінчимо розгляд питання про взаємно двоїсті стандартні задачі лінійного програмування.

Для канонічної й загальної задач лінійного програмування двоїсті задачі також існують. З'ясуємо спочатку питання про двоїсту задачу для загальної задачі.

Загальна задача, як відомо, формулюється так: знайти n-вимірний вектор $x = (x_1, x_2, \ldots, x_n)$, який задовольняє систему обмежень

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, s), \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \quad (i = s+1, s+2, \ldots, m), \\ x_j \geqslant 0 \quad (j = 1, 2, \ldots, t;\ t \leqslant n) \end{cases} \tag{36}$$

і максимізує цільову функцію

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n.$$

Називатимемо цю задачу вихідною або задачею II.

Запишемо задачу II у вигляді стандартної задачі лінійного програмування. В системі обмежень (36) кожне рівняння $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ замінимо рівносильною йому системою нерівностей

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i, \\ -a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n \leqslant -b_i \end{cases}$$

і введемо нові невід'ємні невідомі $w_{t+1}, w_{t+2}, \ldots, w_n$ і $z_{t+1}, z_{t+2}, \ldots, z_n$, які пов'язані з невідомими $x_{t+1}, x_{t+2}, \ldots, x_n$ співвідношеннями

$$x_{t+1} = w_{t+1} - z_{t+1}, \quad x_{t+2} = w_{t+2} - z_{t+2}, \quad \ldots, \quad x_n = w_n - z_n.$$

Задача II запишеться так: дана система нерівностей

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{it}x_t + a_{i\,t+1}(w_{t+1} - z_{t+1}) + \cdots + \\ \quad + a_{in}(w_n - z_n) \leqslant b_i \quad (i = 1, 2, \ldots, s) \\ -a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{it}x_t - a_{i\,t+1}(w_{t+1} - z_{t+1}) - \cdots - \\ \quad - a_{in}(w_n - z_n) \leqslant -b_i \quad (i = 1, 2, \ldots, s) \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{it}x_t + a_{i\,t+1}(w_{t+1} - z_{t+1}) + \cdots + \\ \quad + a_{in}(w_n - z_n) \leqslant b_i \quad (i = s+1, \ldots, m) \end{cases} \tag{37}$$

і цільова функція

$$f = c_1x_1 + c_2x_2 + \cdots + c_tx_t + c_{t+1}(w_{t+1} - z_{t+1}) + \cdots + c_n(w_n - z_n), \tag{38}$$

треба знайти невід'ємний вектор

$$w = (x_1, x_2, \ldots, x_n, w_{t+1}, w_{t+2}, \ldots, w_n, z_{t+1}, z_{t+2}, \ldots, z_n),$$

який задовольняє систему нерівностей (37) і максимізує цільову функцію (38). Двоїстою до стандартної задачі (37) — (38), очевидно, є така задача: дана система нерівностей

$$\begin{cases} a_{1j}(u_1 - v_1) + a_{2j}(u_2 - v_2) + \cdots + a_{sj}(u_s - v_s) + a_{s+1\,j}y_{s+1} + \\ \quad + \cdots + a_{mj}y_m \geqslant c_j \quad (j = 1, 2, \ldots, t), \qquad (39) \\ a_{1j}(u_1 - v_1) + a_{2j}(u_2 - v_2) + \cdots + a_{sj}(u_s - v_s) + a_{s+1\,j}y_{s+1} + \\ \quad + \cdots + a_{mj}y_m \geqslant c_j \quad (j = t+1, \ldots, n), \qquad (40) \\ -a_{1j}(u_1 - v_1) - a_{2j}(u_2 - v_2) - \cdots - a_{sj}(u_s - v_s) - a_{s+1\,j}y_{s+1} - \\ \quad - \cdots - a_{mj}y_m \geqslant -c_j \quad (j = t+1, \ldots, n) \qquad (41) \end{cases}$$

і цільова функція

$$\varphi = b_1(u_1 - v_1) + b_2(u_2 - v_2) + \cdots + b_s(u_s - v_s) + b_{s+1}y_{s+1} + b_{s+2}y_{s+2} + \cdots + b_my_m, \tag{42}$$

треба знайти невід'ємний вектор

$$y = (u_1, v_1, u_2, v_2, \ldots, u_s, v_s, y_{s+1}, y_{s+2}, \ldots, y_m),$$

який задовольняє систему нерівностей (39) — (41) і мінімізує цільову функцію (42).

В цій задачі систему нерівностей (40) і (41) замінимо рівносильною їй системою рівностей, різницю $u_i - v_i$ при $i = 1, 2, \ldots, s$ позначимо через [1] y_i й сформулюємо її так: знайти m-вимірний вектор $y = (y_1, y_2, \ldots, y_m)$, який задовольняє умови

$$\begin{cases} a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant c_j \quad (j = 1, 2, \ldots, t), \\ a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m = c_j \quad (j = t+1, t+2, \ldots, n), \\ y_i \geqslant 0 \quad (i = s+1, s+2, \ldots, m) \end{cases} \tag{43}$$

і мінімізує цільову функцію

$$\varphi = b_1y_1 + b_2y_2 + \cdots + b_my_m. \tag{44}$$

Задачу (43) — (44) вважатимемо двоїстою для задачі II й позначатимемо її символом II′. Якщо для задачі II′ ми сформулюємо щойно описаним способом двоїсту задачу, то дістанемо задачу II. Отже задачі II і II′ — взаємно двоїсті.

Задача II′ одержується з задачі II за такими правилами:

1) невід'ємній змінній $x_j \geqslant 0$ $(j = 1, 2, \ldots, t)$ задачі II відповідає нерівність $a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant c_j$ $(j = 1, 2, \ldots, t)$ в задачі II′;

2) нерівності $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i$ $(i = s+1, s+2, \ldots, m)$ в задачі II відповідає невід'ємна змінна y_i $(i = s+1, s+2, \ldots, m)$ в задачі II′;

3) рівнянню $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ $(i = 1, 2, \ldots, s)$ в задачі II відповідає змінна y_i $(i = 1, 2, \ldots, s)$ довільного знака в задачі II′;

4) змінній x_j $(j = t+1, \ldots, n)$ довільного знака в задачі II відповідає рівняння $a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m = c_j$ $(j = t+1, \ldots, n)$ в задачі II′.

Канонічна задача є окремим випадком загальної задачі. Якщо $s = m$ і $t = n$, то загальна задача являє собою канонічну задачу: знайти невід'ємний вектор $x = (x_1, x_2, \ldots, x_n)$, який задовольняє систему рівнянь

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m) \tag{45}$$

---

[1] Очевидно, що y_i $(i = 1, 2, \ldots, s)$ може бути додатним, від'ємним і рівним 0.

і максимізує цільову функцію

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n. \tag{46}$$

Двоїстою для цієї задачі, очевидно, є задача II′ при $s = m$ і $t = n$, тобто така задача: знайти вектор $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$ (без обмежень за знаком), який задовольняє систему нерівностей

$$a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant c_j \quad (j = 1, 2, \ldots, n) \tag{47}$$

і мінімізує цільову функцію

$$\varphi = b_1y_1 + b_2y_2 + \cdots + b_my_m. \tag{48}$$

Двоїста задача для канонічної, як бачимо, не є канонічною. Але ці задачі, як випливає з викладеного вище, також є взаємно двоїстими.

Для канонічної і загальної задач також справедлива теорема двоїстості.

Теорема 3 (загальна теорема двоїстості). *Якщо кожна з двох взаємно двоїстих задач допустима, то обидві вони мають оптимальні розв'язки і значення цих задач збігаються. Якщо принаймні одна з цих задач не є допустимою, то жодна з них не має оптимального розв'язку.*

Для стандартних взаємно двоїстих задач справедливість теореми доведено. Для загальної задачі теорема доводиться зведенням загальної задачі до стандартної. З справедливості ж теореми для загальної задачі випливає справедливість її й для канонічної, оскільки канонічна задача є окремим випадком загальної.

Доведемо ще одну теорему, яка стосується канонічної задачі й двоїстої для неї й відіграє важливу роль в лінійному програмуванні.

Теорема 4. *Нехай* $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ *— невід'ємний розв'язок системи рівнянь*

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m), \tag{45′}$$

а $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ *— розв'язок системи нерівностей*

$$a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \geqslant c_j \quad (j = 1, 2, \ldots, n). \tag{47′}$$

Для того щоб одночасно $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ *максимізував лінійну форму*

$$f = c_1x_1 + c_2x_2 + \cdots + c_nx_n, \tag{46′}$$

а $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ *мінімізував лінійну форму*

$$\varphi = b_1y_1 + b_2y_2 + \cdots + b_my_m, \tag{48′}$$

необхідно й достатньо, щоб $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ *і* $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ *задовольняли умову: якщо* $a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ > c_j$, *то* $x_j^\circ = 0$.

Доведення. Припустимо, що виконується умова: якщо $a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ > c_j$, то $x_j = 0$. Тоді $x_j^\circ[c_j - (a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ)] = 0$ $(j = 1, 2, \ldots, n)$, оскільки принаймні один із множників лівої частини рівності дорівнює нулю. Отже, $x_jc_j = x_j^\circ(a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ)$ і тому

$$\sum_{j=1}^{n} c_jx_j^\circ = \sum_{j=1}^{n} x_j^\circ \sum_{i=1}^{m} a_{ij}y_i^\circ = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_j^\circ y_i^\circ = \sum_{i=1}^{m} y_i^\circ \left(\sum_{j=1}^{n} a_{ij}x_j^\circ\right) = \sum_{i=1}^{m} b_iy_i^\circ,$$

тобто

$$\sum_{j=1}^{n} c_jx_j^\circ = \sum_{i=1}^{m} b_iy_i^\circ. \tag{49}$$

З рівності (49) випливає, що $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ максимізує лінійну форму (46′), а $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ мінімізує форму (48′). Справді, вектори $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ є допустимими розв'язками канонічної задачі (45) — (46) і двоїстої їй задачі (47) — (48). Тому, за теоремою 3, ці задачі мають оптимальні розв'язки $\boldsymbol{x} = (\xi_1, \xi_2, \ldots, \xi_n)$ і $\boldsymbol{y} = (\eta_1, \eta_2, \ldots, \eta_m)$ й значення їх збігаються: $\sum_{j=1}^{n} c_j\xi_j = \sum_{i=1}^{m} b_i\eta_i$. Якби вектор $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ не максимізував лінійну форму (46′), то було б $\sum_{j=1}^{n} c_jx_j < \sum_{j=1}^{n} c_j\xi_j = \sum_{i=1}^{m} b_i\eta_i \leqslant \sum_{i=1}^{m} b_iy_i^\circ$, тобто $\sum_{j=1}^{n} c_jx_j^\circ < \sum_{i=1}^{m} b_iy_i^\circ$, а це суперечить рівності (49). Отже, $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ максимізує форму (46′). Аналогічними міркуваннями доводять, що $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ мінімізує форму (48′). *Достатність умови* доведено. Доведемо тепер *необхідність умови*. Нехай вектор $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ максимізує форму (46′), а вектор $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ мінімізує форму (48′), тобто $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ і $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ є оптимальні розв'язки відповідно канонічної задачі (45) — (46) і двоїстої їй задачі (47) — (48). Тоді, за теоремою 3,

$$\sum_{j=1}^{n} c_jx_j^\circ = \sum_{i=1}^{m} b_iy_i^\circ.$$

Отже,

$$\sum_{j=1}^{n} c_jx_j^\circ = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j^\circ\right) y_i^\circ = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}y_i^\circ\right) x_j^\circ,$$

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}y_i^\circ - c_j\right) x_j^\circ = 0.$$

Оскільки $\sum_{i=1}^{m} a_{ij}y_i^\circ - c_j \geqslant 0$ і $x_j^\circ \geqslant 0$, то кожен з доданків у лівій частині останньої рівності невід'ємний і тому дорівнює нулю. А це

буде тільки тоді, коли $x_j^\circ = 0$, якщо $\sum_{i=1}^{m} a_{ij} y_i^\circ > c_j$. Теорему доведено.

3.5. Базисні розв'язки. Нехай дано систему лінійних рівнянь

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m). \tag{50}$$

Вважаючи стовпці розширеної матриці системи (50) m-вимірними векторами й позначивши їх відповідно символами $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n, \boldsymbol{b}$ запишемо цю систему так:

$$x_1\boldsymbol{p}_1 + x_2\boldsymbol{p}_2 + \cdots + x_n\boldsymbol{p}_n = \boldsymbol{b}. \tag{50'}$$

Припустимо, що ранг системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ дорівнює r. Тоді кожен базис системи векторів S складається з r векторів. Нехай B — один із них. Не втрачаючи загальності міркувань, вважатимемо, що базис B складається з векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_r$.

Невідомі $x_1, x_2, \ldots, x_r$, на які в рівнянні (50') відповідно множаться вектори $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_r$, називатимемо *базисними невідомими* (відносно базису B); а невідомі $x_{r+1}, x_{r+2}, \ldots, x_n$ називатимемо вільними. Припустимо, що система рівнянь (50) сумісна, тоді, як відомо (див. I, 29.3), її можна розв'язати відносно базисних невідомих $x_1, x_2, \ldots, x_r$. Загальний розв'язок при цьому матиме вигляд:

$$x_i = l_i + c_{i1}x_{r+1} + c_{i2}x_{r+2} + \cdots + c_{i\,n-r}x_n \quad (i = 1, 2, \ldots, r). \tag{51}$$

Якщо $x_{r+1} = x_{r+2} = \cdots = x_n = 0$, то за формулами (51)

$$x_1 = l_1, \; x_2 = l_2, \; \ldots, \; x_r = l_r.$$

Вектор $\boldsymbol{x}^\circ = (l_1, l_2, \ldots, l_r, 0, 0, \ldots, 0)$ є частинний розв'язок системи рівнянь (50). Його називають *базисним розв'язком* системи (50), відповідним базису $B\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_r\}$.

Оскільки при заданих значеннях вільних невідомих значення базисних невідомих за формулами (51) визначаються однозначно, то базису $B\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ відповідає тільки один базисний розв'язок.

Якщо ми виберемо інший базис B' системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$, то він також визначить певну систему базисних і вільних невідомих і йому відповідатиме певний базисний розв'язок. Таким чином, кожному базису системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ відповідає певна система базисних, а отже, і вільних невідомих і, якщо система рівнянь (50) сумісна, єдиний базисний розв'язок.

Підкреслимо, що *базисним розв'язком, відповідним базису* $B\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_r\}$ *називається частинний розв'язок, у якому значення вільних невідомих* $x_{r+1}, x_{r+2}, \ldots, x_n$ *дорівнюють нулю.* Базисні розв'язки, що відповідають всім різним базисам системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$, називають просто базисними розв'язками системи рівнянь (50).

Поняття базисного розв'язку системи лінійних рівнянь можна означити й інакше.

Нехай $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_k, \ldots, \gamma_n)$ — деякий розв'язок системи рівнянь (50). Вважатимемо, що розв'язок $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_k, \ldots, \gamma_n)$

залежить від підсистеми $F = \{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}\}$ системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$, якщо його компоненти $\gamma_{i_1}, \gamma_{i_2}, \ldots, \gamma_{i_k}$ відмінні від нуля, а всі інші компоненти дорівнюють нулю. Кожен розв'язок системи рівнянь (50), очевидно, залежить від деякої підсистеми системи $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$.

Розв'язок $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ системи (50) називається *базисним*, якщо він залежить від деякої лінійно незалежної підсистеми $F = \{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}\}$ системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$.

Звичайно, базисний розв'язок $\boldsymbol{x}^\circ$, що залежить від лінійно незалежної підсистеми F, вважають відповідним підсистемі F.

Друге означення базисного розв'язку рівносильне першому. Справді, нехай $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ — базисний розв'язок, відповідний деякому базису $B\{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_r}\}$ за першим означенням, тоді він, очевидно, залежить або від самого базиса $B\{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_r}\}$, або від деякої його підсистеми, тобто залежить від деякої лінійно незалежної підсистеми системи $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ і, таким чином, є базисним розв'язком за другим означенням. Нехай тепер $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ — базисний розв'язок за другим означенням; в такому разі він залежить від деякої лінійно незалежної підсистеми $F = \{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}\}$ системи $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$. Якщо $k = r$, то підсистема $F = \{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}\}$ є базисом системи $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$, а $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ є базисним розв'язком, відповідним базису $B\{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}\}$. Якщо $k < r$, то доповнимо лінійно незалежну підсистему $F = \{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}\}$ до деякого базиса $B\{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}, \boldsymbol{p}_{i_{k+1}}, \ldots, \boldsymbol{p}_{i_r}\}$; вектор $\boldsymbol{x}^\circ = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ є, очевидно, базисним розв'язком, відповідним базису $B\{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \ldots, \boldsymbol{p}_{i_k}, \ldots, \boldsymbol{p}_{i_r}\}$, за першим означенням.

Якщо система (50) має розв'язок, тобто якщо вона сумісна, то, як зазначалось вище, вона має і базисний розв'язок. Число базисних розв'язків системи (50) дорівнює числу базисів системи векторів $S = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ і тому воно не більше, ніж число A_n^r розміщень із n елементів по r елементів.

Доведемо тепер таку теорему.

Теорема 5. *Якщо система лінійних рівнянь* (50) *має невід'ємний розв'язок, то вона має також і базисний невід'ємний розв'язок.*

Доведення. Доводитимемо теорему методом математичної індукції. При $n = 1$ теорема справедлива: система рівнянь (50) зводиться до одного рівняння $a_{11}x_1 = b_1$, і невід'ємний розв'язок цього рівняння $x_1 = \frac{b_1}{a_1}$ буде і базисним його розв'язком. Припустимо, що теорема справедлива для будь-якого $k < n$, і доведемо, що в такому разі вона справедлива й для n. Тоді, за принципом математичної індукції, вона справедлива для будь-якого натурального n.

Нехай $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ — невід'ємний розв'язок рівняння (50'). Якщо в цьому розв'язку принаймні одна з компонент, наприклад x_n°,

дорівнює нулю, то вектор $\bar{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_{n-1}^\circ)$ є невід'ємний розв'язок рівняння

$$x_1 p_1 + x_2 p_2 + \cdots + x_{n-1} p_{n-1} = b. \qquad (52)$$

Тоді, за індуктивним припущенням, рівняння (52) має невід'ємний, базисний розв'язок $\bar{x}' = (x_1', x_2', \ldots, x_{n-1}')$.

Нехай розв'язок $\bar{x}'$ залежить від лінійно незалежної підсистеми $F' = \{p_{i_1}, p_{i_2}, \ldots, p_{i_k}\}$, $(k \leqslant n-1)$ системи векторів $S = \{p_1, p_2, \ldots, p_n\}$. Не втрачаючи загальності міркувань, можемо вважати, що $F' = \{p_1, p_2, \ldots, p_k\}$ $(k \leqslant n-1)$. В такому разі $\bar{x}' = (x_1', x_2', \ldots, x_k', 0, 0, \ldots, 0)$. Тоді n-вимірний вектор $x' = (x_1', x_2', \ldots, x_k', 0, 0, \ldots, 0)$, що є невід'ємним розв'язком рівняння (50'), є невід'ємним базисним розв'язком цього рівняння, оскільки він залежить від лінійно незалежної системи векторів $F' = (p_1, p_2, \ldots, p_k)$.

Нехай у розв'язку x° всі компоненти додатні. Якщо система векторів $S = \{p_1, p_2, \ldots, p_n\}$ лінійно незалежна, то $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ є невід'ємний базисний розв'язок рівняння (50'). Якщо ж система векторів лінійно залежна, то існують такі числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, які не всі дорівнюють нулю, що

$$\lambda_1 p_1 + \lambda_2 p_2 + \cdots + \lambda_n p_n = 0. \qquad (53)$$

Оскільки серед чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ є відмінні від нуля, то ми можемо вважати, що деякі з них додатні. В противному разі розглядатимемо рівність $(-\lambda_1) p_1 + (-\lambda_2) p_2 + \cdots + (-\lambda_n) p_n = 0$.

Розглянемо тепер відношення $\frac{\lambda_1}{x_1^\circ}, \frac{\lambda_2}{x_2^\circ}, \ldots, \frac{\lambda_n}{x_n^\circ}$.

Позначимо буквою ξ найбільше з цих відношень:

$$\xi = \max\left\{\frac{\lambda_1}{x_1^\circ}, \frac{\lambda_2}{x_2^\circ}, \ldots, \frac{\lambda_n}{x_n^\circ}\right\}.$$

Оскільки серед чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ є додатні і $x_i^\circ > 0$ $(i = 1, 2, \ldots, n)$, то $\xi > 0$. Не втрачаючи загальності міркувань, вважатимемо, що $\xi = \frac{\lambda_1}{x_1^\circ}$ (цього завжди можна досягти, змінивши нумерацію векторів p_i і невідомих x_i).

Запишемо тепер рівність (53) так:

$$\frac{\lambda_1}{x_1^\circ} x_1^\circ p_1 + \frac{\lambda_2}{x_2^\circ} x_2^\circ p_2 + \cdots + \frac{\lambda_n}{x_n^\circ} x_n^\circ p_n = 0.$$

Це можна зробити, оскільки в розглядуваному випадку

$$x_i^\circ > 0 \quad (i = 1, 2, \ldots, n).$$

Помноживши останню рівність на відмінне від нуля число $\frac{1}{\xi}$, матимемо:

$$\frac{1}{\xi}\left(\frac{\lambda_1}{x_1^\circ} x_1^\circ p_1 + \frac{\lambda_2}{x_2^\circ} x_2^\circ p_2 + \cdots + \frac{\lambda_n}{x_n^\circ} x_n^\circ p_n\right) = 0. \qquad (54)$$

Оскільки $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ — розв'язок рівняння (50'), то

$$x_1^\circ p_1 + x_2^\circ p_2 + \cdots + x_n^\circ p_n = b. \qquad (55)$$

Рівність (55) запишемо так:

$$\frac{1}{\xi}(\xi x_1^\circ p_1 + \xi x_2^\circ p_2 + \cdots + \xi x_n^\circ p_n) = b. \qquad (56)$$

Віднімемо тепер почастинно рівність (54) від рівності (56). Дістанемо:

$$\frac{1}{\xi}\left(\xi - \frac{\lambda_1}{x_1^\circ}\right) x_1^\circ p_1 + \frac{1}{\xi}\left(\xi - \frac{\lambda_2}{x_2^\circ}\right) x_2^\circ p_2 + \cdots + \frac{1}{\xi}\left(\xi - \frac{\lambda_n}{x_n^\circ}\right) x_n^\circ p_n = b. \qquad (57)$$

Позначимо в рівності (57) число $\frac{1}{\xi}\left(\xi - \frac{\lambda_i}{x_i^\circ}\right) x_i^\circ$ $(i = 1, 2, \ldots, n)$ символом γ_i. Тоді матимемо:

$$\gamma_1 p_1 + \gamma_2 p_2 + \cdots + \gamma_n p_n = b. \qquad (58)$$

Рівність (58) означає, що вектор $g = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ є розв'язком рівняння (50').

За означенням числа ξ, число $\gamma_1 = 0$, а числа $\gamma_2, \gamma_3, \ldots, \gamma_n$ невід'ємні. Отже, рівняння (50') має невід'ємний розв'язок $g = (0, \gamma_2, \ldots, \gamma_n)$, у якого перша компонента дорівнює нулю, і тому, за доведеним вище, рівняння (50') має невід'ємний базисний розв'язок. Теорему доведено.

Повернемося тепер до канонічної задачі лінійного програмування. Нагадаємо, що система обмежень в канонічній задачі складається з лінійних рівнянь, тобто є системою лінійних рівнянь. Тому правомірним буде говорити про базисні розв'язки системи обмежень канонічної задачі, а отже, і про базисні розв'язки канонічної задачі.

Означення. *Оптимальний розв'язок канонічної задачі називається базисним, якщо він є базисним розв'язком системи обмежень цієї задачі.*

Справедлива така теорема.

Теорема 6. *Якщо канонічна задача лінійного програмування має оптимальний розв'язок, то вона має і оптимальний базисний розв'язок.*

Доведення. Нехай системою обмежень розглядуваної канонічної задачі є система рівнянь:

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m). \qquad (59)$$

Запишемо її у векторній формі:

$$x_1 p_1 + x_2 p_2 + \cdots + x_n p_n = b. \qquad (59')$$

Нехай $x^0 = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ — оптимальний розв'язок канонічної задачі. Цей розв'язок залежить від деякої системи векторів p_i.

Вважатимемо, що він залежить від системи $F = \{p_1, p_2, \ldots, p_r\}$. Якщо система векторів $F = \{p_1, p_2, \ldots, p_r\}$ лінійно незалежна, то

оптимальний розв'язок $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ є базисним розв'язком, і теорема в такому разі справедлива.

Припустимо, що система векторів $F = \{p_1, p_2, \ldots, p_r\}$ лінійно залежна. Оскільки розв'язок $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ залежить від системи $F = \{p_1, p_2, \ldots, p_r\}$, то $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_r^\circ, 0, 0, \ldots, 0)$, де $x_1^\circ, x_2^\circ, \ldots, x_r^\circ$ — числа, відмінні від нуля.

Нехай $y^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ — оптимальний розв'язок задачі, двоїстої для розглядуваної канонічної задачі (за теоремою 3 такий розв'язок існує).

Тоді, за теоремою 4,

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ = c_j \text{ для } j = 1, 2, \ldots, r; \quad (60)$$

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ > c_j \text{ для } j = r+1, r+2, \ldots, n. \quad (61)$$

Оскільки $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_r^\circ, 0, 0, \ldots, 0)$ — невід'ємний розв'язок рівняння (59′), то $\bar{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_r^\circ)$ є невід'ємним розв'язком рівняння

$$x_1p_1 + x_2p_2 + \cdots + x_rp_r = b. \quad (62)$$

За теоремою 5 рівняння (62) має базисний невід'ємний розв'язок $\bar{x}'$. Розв'язок $\bar{x}'$ залежить від деякої лінійно незалежної підсистеми F' системи векторів $F = \{p_1, p_2, \ldots, p_r\}$. Нехай $F' = \{p_1, p_2, \ldots, p_k\}$ $(k < r)$, тоді $\bar{x}' = (x_1', x_2', \ldots, x_k', \underbrace{0, 0, \ldots, 0}_{r-k \text{ нулів}})$, де $x_1', x_2', \ldots, x_k'$ — додатні числа.

Невід'ємний n-вимірний вектор $x' = (x_1', x_2', \ldots, x_k', 0, 0, \ldots, 0)$, очевидно, є розв'язком рівняння (59′), причому базисним, оскільки він залежить від лінійно незалежної системи векторів $F' = \{p_1, p_2, \ldots, p_k\}$. Розглянемо тепер допустимий розв'язок x' канонічної задачі й оптимальний розв'язок y° двоїстої задачі. Для розв'язку y° справедливі співвідношення (60) і (61), а в розв'язку x' маємо $x_{r+1}' = x_{r+2}' = \cdots = x_n' = 0$. Розв'язки x' і y°, таким чином, задовольняють умови теореми 4, і тому x' є оптимальним розв'язком канонічної задачі.

Отже, $x' = (x_1', x_2', \ldots, x_k', 0, 0, \ldots, 0)$ є оптимальний базисний розв'язок канонічної задачі. Теорему доведено.

З доведеної теореми випливає, що кожну канонічну задачу лінійного програмування можна розв'язати, виконавши скінчену кількість обчислень, а саме: треба знайти всі невід'ємні базисні розв'язки рівняння (59′) (а їх буде скінченне число), обчислити значення цільової функції $f = c_1x_1 + c_2x_2 + \ldots + c_nx_n$ для кожного з цих розв'язків і вибрати серед них найбільше (найменше). Невід'ємний базисний розв'язок, при якому значення цільової функції найбільше (найменше), і буде оптимальним розв'язком канонічної задачі максимізації (мінімізації). Зрозуміло, що при досить великій кількості невідомих цей метод розв'язання канонічної задачі непридатний, оскільки треба буде виконати надзвичайно велику кількість обчислень.

У попередніх параграфах цього розділу викладено теоретичні основи лінійного програмування, а тепер розглянемо питання, пов'язані з чисельним розв'язуванням конкретних задач лінійного програмування.

Реальні задачі лінійного програмування, як правило, містять велику кількість невідомих і обмежень.

У процесі розв'язування таких задач доводиться виконувати за допомогою швидкодіючих обчислювальних машин велику за обсягом обчислювальну роботу. В основі машинних програм, за якими провадяться необхідні обчислення, лежать певні алгоритми. Серед цих алгоритмів є спеціальні, придатні для розв'язання лише задач певного класу, — вони пов'язані з специфікою задач цього класу. Проте є й загальні методи, за допомогою яких можна розв'язати кожну задачу лінійного програмування. Одним з таких загальних методів є так званий симплекс-метод. Він є найбільш поширеним методом розв'язування задач лінійного програмування; його з успіхом застосовують також і при розв'язуванні інших задач. Зокрема симплекс-метод є ефективним методом розв'язання систем лінійних нерівностей і знаходження невід'ємних розв'язків систем лінійних рівнянь.

4.1. Заміщення вектора базису. В основі симплекс-методу лежить операція *заміщення вектора базису*. З'ясуємо, насамперед, що ж являє собою ця операція.

Нехай дано систему рівнянь

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad (i = 1, 2, \ldots, m). \quad (1)$$

Запишемо її у векторній формі:

$$x_1p_1 + x_2p_2 + \cdots + x_np_n = b. \quad (1')$$

Задачу розв'язання рівняння (1′), очевидно, можна сформулювати так: дано $(n+1)$ вектор простору V_m: $p_1, p_2, \ldots, p_n$ і b, треба вектор b подати, якщо це можливо, у вигляді лінійної комбінації векторів $p_1, p_2, \ldots, p_n$. З'ясуємо, як це можна зробити. Вектор b є лінійна комбінація одиничних векторів $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_m = (0, 0, \ldots, 1)$ простору V_m:

$$b = b_1e_1 + b_2e_2 + \cdots + b_me_m. \quad (2)$$

Кожен із векторів p_i також є лінійна комбінація векторів $e_1, e_2, \ldots, e_m$:

$$p_j = a_{1j}e_1 + a_{2j}e_2 + \cdots + a_{mj}e_m \quad (j = 1, 2, \ldots, n). \quad (3)$$

Якщо у рівності (2) один з векторів e_i замінимо, скориставшись співвідношеннями (3), на деякий із векторів p_j, наприклад, e_1 на p_1, то дістанемо запис вектора b у вигляді лінійної комбінації векторів $p_1, e_2, e_3, \ldots, e_m$:

$$b = b_1'p_1 + b_2'e_2 + b_3'e_3 + \cdots + b_m'e_m. \quad (4)$$

У цьому записі спробуємо замінити ще один одиничний вектор e_i деяким вектором p_j і так далі.

Якщо, діючи так, ми замінимо всі одиничні вектори $\boldsymbol{e}_i$ векторами $\boldsymbol{p}_j$, то дістанемо запис вектора $\boldsymbol{b}$ у вигляді лінійної комбінації векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$:

$$\boldsymbol{b} = \gamma_1\boldsymbol{p}_1 + \gamma_2\boldsymbol{p}_2 + \cdots + \gamma_n\boldsymbol{p}_n.$$

У такому разі вектор $\boldsymbol{g} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ буде розв'язком рівняння (1'). Цей метод розв'язування системи лінійних рівнянь (його іноді називають *методом заміщення*) на практиці виявляється досить ефективним, оскільки заміна вектора $\boldsymbol{e}_i$ вектором $\boldsymbol{p}_j$ і перехід від одного зображення вектора $\boldsymbol{b}$ до наступного (від зображення (2) до зображення (4), наприклад) здійснюється досить просто. Щоб формально описати цей перехід, розглянемо більш загальну задачу.

Припустимо, що дано лінійно незалежну систему векторів:

$$\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_m, \tag{5}$$

яку ми умовно називатимемо базисом, і систему векторів

$$\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n, \tag{6}$$

кожний з яких є лінійною комбінацією векторів системи (5).

Нехай

$$\boldsymbol{p}_j = \tau_{1j}\boldsymbol{d}_1 + \tau_{2j}\boldsymbol{d}_2 + \cdots + \tau_{mj}\boldsymbol{d}_m \quad (j = 1, 2, \ldots, n). \tag{7}$$

З коефіцієнтів τ_{ij} складемо матрицю

$$T = \begin{pmatrix} \tau_{11} & \tau_{12} & \ldots & \tau_{1j} & \ldots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \ldots & \tau_{2j} & \ldots & \tau_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \tau_{i1} & \tau_{i2} & \ldots & \tau_{ij} & \ldots & \tau_{in} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \tau_{m1} & \tau_{m2} & \ldots & \tau_{mj} & \ldots & \tau_{mn} \end{pmatrix}.$$

Підкреслимо, що елементами j-го стовпчика матриці T є коефіцієнти лінійного зображення вектора $\boldsymbol{p}_j$ через вектори $\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_m$.

Складемо тепер таку таблицю:

		$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\cdots$	$\boldsymbol{p}_j$	$\cdots$	$\boldsymbol{p}_n$
T	$\boldsymbol{d}_1$	τ_{11}	τ_{12}	$\cdots$	τ_{1j}	$\cdots$	τ_{1n}
	$\boldsymbol{d}_2$	τ_{21}	τ_{22}	$\cdots$	τ_{2j}	$\cdots$	τ_{2n}
	$\boldsymbol{d}_i$	τ_{i1}	τ_{i2}	$\cdots$	τ_{ij}	$\cdots$	τ_{in}
	$\boldsymbol{d}_m$	τ_{m1}	τ_{m2}	$\cdots$	τ_{mj}	$\cdots$	τ_{mn}

Цю таблицю будемо називати *таблицею векторів* $\boldsymbol{p}_i$ *відносно базису* $\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_m$.

Замінимо тепер у базисі (5) один з його векторів, наприклад вектор $\boldsymbol{d}_r$, деяким вектором $\boldsymbol{p}_s$, але таким, що $\tau_{rs} \neq 0$. Матимемо систему векторів

$$\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_{r-1}, \boldsymbol{p}_s, \boldsymbol{d}_{r+1}, \ldots, \boldsymbol{d}_m. \tag{8}$$

Системи векторів (5) і (8) — рівносильні. Справді, кожен вектор системи (8) є лінійною комбінацією векторів системи (5). А оскільки при $\tau_{rs} \neq 0$ з рівності

$$\boldsymbol{p}_s = \tau_{1s}\boldsymbol{d}_1 + \tau_{2s}\boldsymbol{d}_2 + \cdots + \tau_{rs}\boldsymbol{d}_r + \cdots + \tau_{ms}\boldsymbol{d}_m$$

випливає, що вектор $\boldsymbol{d}_r$ є лінійною комбінацією векторів системи (8), то й кожен вектор системи (5) є лінійною комбінацією векторів системи (8). Отже, система векторів (8) лінійно незалежна і кожен з векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$ є лінійною комбінацією векторів системи (8). Якою ж буде таблиця T_1 векторів $\boldsymbol{p}_i$ відносно базису (8)? Якщо коефіцієнти в лінійному зображенні вектора $\boldsymbol{p}_j$ через вектори базису (8) позначимо символами $\tau'_{1j}, \tau'_{2j}, \ldots, \tau'_{mj}$, то таблиця T_1 матиме вигляд

		$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\cdots$	$\boldsymbol{p}_s$	$\cdots$	$\boldsymbol{p}_n$
T_1	$\boldsymbol{d}_1$	τ'_{11}	τ'_{12}	$\cdots$	0	$\cdots$	τ'_{1n}
	$\boldsymbol{d}_2$	τ'_{21}	τ'_{22}	$\cdots$	0	$\cdots$	τ'_{2n}
	$\boldsymbol{p}_s$	τ'_{r1}	τ'_{r2}	$\cdots$	1	$\cdots$	τ'_{rn}
	$\boldsymbol{d}_m$	τ'_{m1}	τ'_{m2}	$\cdots$	0	$\cdots$	τ'_{mn}

Заміну вектора $\boldsymbol{d}_r$ вектором $\boldsymbol{p}_s$ і перехід від таблиці T до таблиці T_1 умовно називають *операцією заміщення вектора базису*. Знайдемо тепер формули, що виражають елементи τ'_{ij} через елементи τ_{ij}.

Оскільки за умовою $\tau_{rs} \neq 0$, то з рівності

$$\boldsymbol{p}_s = \tau_{1s}\boldsymbol{d}_1 + \tau_{2s}\boldsymbol{d}_2 + \cdots + \tau_{rs}\boldsymbol{d}_r + \cdots + \tau_{ms}\boldsymbol{d}_m$$

дістаємо:

$$\boldsymbol{d}_r = \frac{1}{\tau_{rs}}\boldsymbol{p}_s - \frac{\tau_{1s}}{\tau_{rs}}\boldsymbol{d}_1 - \frac{\tau_{2s}}{\tau_{rs}}\boldsymbol{d}_2 - \cdots - \frac{\tau_{r-1\,s}}{\tau_{rs}}\boldsymbol{d}_{r-1} - \frac{\tau_{r+1\,s}}{\tau_{rs}}\boldsymbol{d}_{r+1} - \cdots - \frac{\tau_{ms}}{\tau_{rs}}\boldsymbol{d}_m.$$

Підставимо у рівності (7) замість $\boldsymbol{d}_r$ його вираз через вектори $\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_{r-1}, \boldsymbol{d}_{r+1}, \ldots, \boldsymbol{d}_m, \boldsymbol{p}_s$, тоді матимемо:

$$\boldsymbol{p}_j = \left(\tau_{1j} - \frac{\tau_{1s}}{\tau_{rs}}\tau_{rj}\right)\boldsymbol{d}_1 + \left(\tau_{2j} - \frac{\tau_{2s}}{\tau_{rs}}\tau_{rj}\right)\boldsymbol{d}_2 + \cdots + \left(\tau_{r-1\,j} - \frac{\tau_{r-1\,s}}{\tau_{rs}}\tau_{rj}\right)\boldsymbol{d}_{r-1} + \frac{\tau_{rj}}{\tau_{rs}}\boldsymbol{p}_s + \left(\tau_{r+1\,j} - \frac{\tau_{r+1\,s}}{\tau_{rs}}\tau_{rj}\right)\boldsymbol{d}_{r+1} + \cdots + \left(\tau_{mj} - \frac{\tau_{ms}}{\tau_{rs}}\tau_{rj}\right)\boldsymbol{d}_m \quad (j = 1, 2, \ldots, n). \tag{9}$$

Оскільки система векторів (8) лінійно незалежна, то вектор p_j записується у вигляді лінійної комбінації векторів цієї системи єдиним способом. Тому з рівності (9) й таблиці T_1 випливає, що

$$\tau'_{ij} = \tau_{ij} - \frac{\tau_{is}}{\tau_{rs}}\tau_{rj} \text{ при } i \neq r \tag{10}$$

$$\tau'_{rj} = \frac{\tau_{rj}}{\tau_{rs}}. \tag{11}$$

Аналіз формул (10) і (11) показує, що матрицю T_1 можна дістати з матриці T за допомогою таких елементарних перетворень матриці T:

а) до i-го ($i = 1, 2, \ldots, r-1, r+1, \ldots, m$) рядка матриці T додати r-й рядок, помножений на число $-\frac{\tau_{is}}{\tau_{rs}}$;

б) r-й рядок матриці T помножити на число $\frac{1}{\tau_{rs}}$.

Інакше кажучи, треба виконати такі елементарні перетворення матриці T, щоб в s-му стовпці всі елементи, крім елемента, розміщеного в r-му рядку, дорівнювали нулю, а елемент, розміщений в r-му рядку і s-му стовпці, дорівнював 1.

Число τ_{rs} з формул (10) і (11) називають *розв'язувальним коефіцієнтом* даного заміщення. Вибір його обумовлюється вимогою $\tau_{rs} \neq 0$ і зручністю обчислень за формулами (10) і (11).

Якщо в таблиці T $\tau_{rs} = 0$, то на даному етапі процесу заміщення векторів базису вектор d_r не можна замінити вектором p_s: вектор d_r не можна подати у вигляді лінійної комбінації векторів $d_1, d_2, \ldots, d_{r-1}, p_s, d_{r+1}, \ldots, d_m$ і тому системи векторів (6) і (8) не будуть рівносильними. Всюди далі розв'язувальний коефіцієнт в таблицях векторів p_l відносно базисів будемо обводити прямокутником.

Приклад. Застосуємо операцію заміщення для розв'язання такої системи рівнянь:

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 5, \\ -2x_1 + x_2 + 2x_3 = 9, \\ x_1 + 4x_2 - 3x_3 = 5. \end{cases} \tag{12}$$

У векторній формі ця система запишеться так:

$$x_1p_1 + x_2p_2 + x_3p_3 = b,$$

де $p_1 = (3, -2, 1)$, $p_2 = (2, 1, 4)$, $p_3 = (1, 2, -3)$, $b = (5, 9, 5)$.

Записавши кожен із векторів p_1, p_2, p_3, b у вигляді лінійної комбінації одиничних векторів e_1, e_2, e_3, складемо початкову таблицю T:

		p_1	p_2	p_3	b
T	e_1	3	2	1	5
	e_2	−2	1	2	9
	e_3	1	4	−3	5

Візьмемо за розв'язувальний коефіцієнт першого кроку заміщення елемент $\tau_{13} = 1$. (За розв'язувальний коефіцієнт краще брати числа 1 і −1, якщо вони є в таблицях; тоді не треба буде виконувати дію ділення, що скоротить обчислення). При цьому вектор e_1 заміщується вектором p_3. Першу операцію заміщення виконуємо так: перший рядок матриці T помножимо на −2 і додамо до другого рядка, потім перший рядок помножимо на 3 і додамо до третього рядка. Перший рядок залишаємо без зміни. В результаті дістанемо таблицю T_1:

		p_1	p_2	p_3	b
T_1	p_3	3	2	1	5
	e_2	−8	−3	0	−1
	e_3	10	10	0	20

В таблиці T_1 за розв'язувальний коефіцієнт візьмемо число −3, тобто виконаємо заміщення вектора e_2 вектором p_2. До першого рядка матриці T_1 додамо другий рядок, помножений на $\frac{2}{3}$, а до третього рядка додамо другий, помножений на $\frac{10}{3}$. Потім другий рядок поділимо на −3. Дістанемо таблицю T_2:

		p_1	p_2	p_3	b
T_2	p_3	$-\frac{7}{3}$	0	1	$\frac{13}{3}$
	p_2	$\frac{8}{3}$	1	0	$\frac{1}{3}$
	e_3	$-\frac{50}{3}$	0	0	$\frac{50}{3}$

В таблиці T_2 за розв'язувальний коефіцієнт візьмемо число $-\frac{50}{3}$. До першого рядка матриці T_2 додамо третій, помножений на $-\frac{7}{50}$, а до другого — третій, помножений на $\frac{4}{25}$, і, крім того, третій рядок поділимо на $-\frac{50}{3}$. Дістанемо таблицю T_3:

		p_1	p_2	p_3	b
T_3	p_3	0	0	1	2
	p_2	0	1	0	3
	p_1	1	0	0	−1

З цієї таблиці випливає, що

$$b = (-1)p_1 + 3p_2 + 2p_3.$$

Отже, розв'язком системи рівнянь (12) є $x_1 = -1$, $x_2 = 3$, $x_3 = 2$, тобто вектор $x^0 = (-1, 3, 2)$.

Проаналізувавши уважно розв'язання системи рівнянь (12) методом заміщення, легко помітити, що по суті воно нічим не відрізняється від розв'язання її методом Гаусса. Перехід від таблиці T до таблиці T_1 — це не що інше, як виключення невідомого x_3 з другого і третього рівняння системи; перехід від таблиці T_1 до таблиці T_2 — це виключення невідомого x_2 з першого і третього рівняння, а перехід від таблиці T_2 до таблиці T_3 — це виключення невідомого x_1 з першого і другого рівнянь. Ми не розглядатимемо більше прикладів розв'язування систем лінійних рівнянь методом заміщення. Лише зауважимо, що при розв'язанні методом заміщення несумісної системи рівнянь ми обов'язково дійдемо до такого положення, коли дальша заміна одиничних векторів $\boldsymbol{e}_i$ векторами $\boldsymbol{p}_j$ буде неможливою: для всіх векторів $\boldsymbol{e}_i$, які ще треба заміщати, і всіх векторів $\boldsymbol{p}_j$, якими можна заміщати, елементи τ_{ij} виявляться рівними нулю. Це обов'язково станеться, бо в противному разі ми замістили б усі одиничні вектори $\boldsymbol{e}_i$ і система мала б розв'язки.

Однак, якщо при розв'язуванні заданої системи лінійних рівнянь методом заміщення ми дійдемо до такого етапу, на якому дальше заміщення одиничних векторів $\boldsymbol{e}_i$ векторами $\boldsymbol{p}_j$ уже неможливе, то це означатиме, що або ми вже знайшли розв'язок заданої системи лінійних рівнянь, або його немає зовсім.

4.2. Симплекс-метод у лінійному програмуванні. Розглянемо канонічну задачу мінімізації: знайти невід'ємний вектор $\boldsymbol{l} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, який задовольняє систему обмежень

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \ (i = 1, 2, \ldots, m) \qquad (13)$$

і мінімізує цільову функцію

$$f(x) = c_1x_1 + c_2x_2 + \cdots + c_nx_n.$$

Запишемо систему обмежень (13) у векторній формі:

$$x_1\boldsymbol{p}_1 + x_2\boldsymbol{p}_2 + \cdots + x_n\boldsymbol{p}_n = \boldsymbol{b}. \qquad (13')$$

Нагадаємо, що

$$\boldsymbol{p}_j = (a_{1j}, a_{2j}, \ldots, a_{mj}), \ (j = 1, 2, \ldots, n), \ \boldsymbol{b} = (b_1, b_2, \ldots, b_m).$$

Вважатимемо, що система (13) сумісна, в противному разі канонічна задача не мала б розв'язку.

Якщо система (13) сумісна, то, не втрачаючи загальності міркувань, завжди можна вважати, що ранг матриці A цієї системи дорівнює m. Справді, якщо матриця A має ранг $r < m$, то її рядки лінійно залежні і можна знайти такі r лінійно незалежних її рядків наприклад, $a_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $a_2 = (a_{21}, a_{22}, \ldots, a_{2n})$, $\ldots$, $a_r = (a_{r1}, a_{r2}, \ldots, a_{rn})$, що всі інші рядки будуть лінійними комбінаціями цих r рядків.

Система рівнянь

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \ (i = 1, 2, \ldots, r),$$

як відомо, рівносильна системі рівнянь (13).

Замінивши цією системою систему (13), матимемо систему обмежень, ранг матриці якої дорівнює числу її рівнянь. Крім того, вважатимемо, що система (13) задовольняє *умову невиродженості: вектор $\boldsymbol{b}$ не можна подати у вигляді лінійної комбінації меншого ніж m числа векторів $\boldsymbol{p}_i$.*

За цієї умови будь-який базисний розв'язок рівняння (13') не може залежати від лінійно незалежної системи векторів $\boldsymbol{p}_j$, яка складається з меншого ніж m числа векторів, і, отже, він залежить від деякого базису системи векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$.

Припустимо, що нам відомий деякий допустимий базисний розв'язок $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ розглядуваної канонічної задачі, тобто невід'ємний базисний розв'язок системи рівнянь (13).

Про те, як знайти невід'ємний базисний розв'язок системи лінійних рівнянь, мова йтиме далі. Він залежить від деякого базису $B\{\boldsymbol{p}_{j_1}, \boldsymbol{p}_{j_2}, \ldots, \boldsymbol{p}_{j_m}\}$. Оскільки нумерацію невідомих завжди можна змінити так, щоб вектори $\boldsymbol{p}_{j_1}, \boldsymbol{p}_{j_2}, \ldots, \boldsymbol{p}_{j_m}$ дістали індекси $1, 2, \ldots, m$, то вважатимемо, що $\boldsymbol{x}^\circ$ залежить від базису $B\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_m\}$, тобто що $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_m^\circ, 0, 0, \ldots, 0)$, де $x_i^\circ > 0$ $(i = 1, 2, \ldots, m)$.

Значення цільової функції $f(x)$ при $\boldsymbol{x} = \boldsymbol{x}^\circ$ позначимо символом ζ_0, тобто $\zeta_0 = f(\boldsymbol{x}^\circ) = c_1x_1^\circ + c_2x_2^\circ + \cdots + c_mx_m^\circ$.

Вектор $\boldsymbol{c} = (c_1, c_2, \ldots, c_m)$ умовно називатимемо *вектором вартості*, що відповідає базису $B\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_m\}$. Зауважимо, що із зміною базису вектор вартості також змінюється. Так, базису $B\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{r-1}, \boldsymbol{p}_s, \boldsymbol{p}_{r+1}, \ldots, \boldsymbol{p}_m\}$ відповідатиме вектор вартості

$$\boldsymbol{c} = (c_1, c_2, \ldots, c_{r-1}, c_s, c_{r+1}, \ldots, c_m).$$

Складемо таблицю T векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$, і $\boldsymbol{b}$ відносно базису B:

		$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\cdots$	$\boldsymbol{p}_m$	$\cdots$	$\boldsymbol{p}_s$	$\cdots$	$\boldsymbol{p}_n$	$\boldsymbol{b}$
T	$\boldsymbol{p}_1$	1	0	$\cdots$	0	$\cdots$	τ_{1s}	$\cdots$	τ_{1n}	x_1°
	$\boldsymbol{p}_2$	0	1	$\cdots$	0	$\cdots$	τ_{2s}	$\cdots$	τ_{2n}	x_2°
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	$\vdots$
	$\boldsymbol{p}_r$	0	0	$\cdots$	0	$\cdots$	τ_{rs}	$\cdots$	τ_{rn}	x_r°
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	$\vdots$
	$\boldsymbol{p}_m$	0	0	$\cdots$	1	$\cdots$	τ_{ms}	$\cdots$	τ_{mn}	x_m°

Знаходити оптимальний розв'язок $\boldsymbol{x}' = (x_1', x_2', \ldots, x_n')$ канонічної задачі мінімізації, як уже зазначалося вище, можна було б так: виходячи з базисного розв'язку $\boldsymbol{x}^\circ$ за допомогою операції заміщення знайти всі можливі допустимі базисні розв'язки і потім знайти

значення цільової функції для всіх цих розв'язків; якщо канонічна задача мінімізації має оптимальний розв'язок, то за теоремою 6, п. 13.5 оптимальним розв'язком буде той із допустимих базисних розв'язків, для якого значення цільової функції найменше. Однак, такий шлях знаходження оптимального розв'язку трудомісткий і малоефективний.

Звичайно цю задачу розв'язують так званим симплекс-методом. Суть симплекс-методу полягає в тому, що заміщення векторів базису провадять за певними правилами (правилами симплексного методу), завдяки чому після кожного наступного кроку заміщення ми дістаємо базисний розв'язок, який відрізняється від оптимального розв'язку менше ніж базисний розв'язок, знайдений на попередньому етапі. Ці правила будуть сформульовані нижче. При розв'язанні канонічної задачі симплекс-методом насамперед з'ясовують, чи не є вихідний допустимий базисний розв'язок x° оптимальним. Зробити це дають змогу теореми, які ми доведемо після виконання деякої підготовчої роботи.

Зафіксуємо у просторі V_m базис, що складається з одиничних векторів $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_m = (0, 0, \ldots, 0, 1)$. Компоненти вектора $a = (a_1, a_2, \ldots, a_m)$ будуть і координатами цього вектора в базисі $e_1, e_2, \ldots, e_m$. Скалярним добутком векторів $a = (a_1, a_2, \ldots, a_m)$ і $b = (b_1, b_2, \ldots, b_m)$ називатимемо суму добутків відповідних координат цих векторів у базисі $B\{e_1, e_2, \ldots, e_m\}$. Отже,

$$(a, b) = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

Простір V_m з визначеним таким способом скалярним добутком, тобто евклідів простір, позначатимемо символом $\widehat{V}_m$. Розглянемо тепер два вектори простору $\widehat{V}_m$: вектор $\bar{t}_j = (\tau_{1j}, \tau_{2j}, \ldots, \tau_{mj})$ і вектор вартості $c = (c_1, c_2, \ldots, c_m)$; компонентами вектора $\bar{t}_j$ є елементи таблиці T, що записані у стовпчик під вектором p_j, а вектора c — компоненти вектора $\bar{c} = (c_1, c_2, \ldots, c_m, \ldots, c_n)$, які мають ті самі індекси, що й вектори розглядуваного базису $B\{p_1, p_2, \ldots, p_m\}$ системи векторів $p_1, p_2, \ldots, p_n$.

Скалярний добуток векторів c і $\bar{t}_j$ позначимо символом ζ_j:

$$\zeta_j = (c, \bar{t}_j) = c_1\tau_{1j} + c_2\tau_{2j} + \cdots + c_m\tau_{mj} \quad (j = 1, 2, \ldots, n). \tag{14}$$

Визначимо тепер $(n+1)$-вимірні вектори z і c_0 за допомогою рівностей

$$z = (\zeta_1, \zeta_2, \ldots, \zeta_n, \zeta_0) \text{ і } c_0 = (c_1, c_2, \ldots, c_n, 0),$$

де ζ_j при $j > 0$ визначається рівністю (14), а $\zeta_0 = c_1x_1^\circ + c_2x_2^\circ + \cdots + c_mx_m^\circ$ значення цільової функції при $x = x^\circ$.

Приєднаємо до таблиці T ще один рядок, елементами якого є компоненти вектора $z - c_0 = (\zeta_1 - c_1, \zeta_2 - c_2, \ldots, \zeta_n - c_n, \zeta_0)$.

Розширена таблиця T матиме такий вигляд:

T		p_1	p_2	$\cdots$	p_m	$\cdots$	p_s	$\cdots$	p_n	b
	p_1	1	0	$\cdots$	0	$\cdots$	τ_{1s}	$\cdots$	τ_{1n}	x_1°
	p_2	0	1	$\cdots$	0	$\cdots$	τ_{2s}	$\cdots$	τ_{2n}	x_2°
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	$\vdots$
	p_r	0	0	$\cdots$	0	$\cdots$	$\boxed{\tau_{rs}}$	$\cdots$	τ_{rn}	x_r°
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	$\vdots$
	p_m	0	0	$\cdots$	1	$\cdots$	τ_{ms}	$\cdots$	τ_{mn}	x_m°
	$z - c_0$	0	0	$\cdots$	0		$\zeta_s - c_s$	$\cdots$	$\zeta_n - c_n$	ζ_0

Припустимо, що вектор p_r базису ми замінили вектором p_s (розв'язувальний коефіцієнт τ_{rs} обведено прямокутником) і перейшли до нової розширеної таблиці T_1. Ми знаємо, за яким правилом обчислюються елементи τ'_{ij} таблиці T_1. Постає питання, як за відомим додатковим рядком $z - c_0$ таблиці T знайти додатковий рядок $z' - c_0$ таблиці T_1. Позначимо i-й вектор-рядок таблиці T символом t_i, а таблиці T_1 — символом t'_i. Доведемо тепер, що *додатковий вектор-рядок $z' - c_0$ таблиці T_1 пов'язаний з додатковим вектором-рядком $z - c_0$ таблиці T співвідношенням:*

$$z' - c_0 = (z - c_0) - \frac{\zeta_s - c_s}{\tau_{rs}} t_r. \tag{15}$$

Справді, за означенням векторів z', t'_i, c і чисел ζ'_i в таблиці T_1 маємо:

$$z' = c_1t'_1 + c_2t'_2 + \cdots + c_{r-1}t'_{r-1} + c_st'_r + c_{r+1}t'_{r+1} + \cdots + c_mt'_m = \sum_{i=1\ (i\neq r)}^{m} c_it'_i + c_st'_r. \tag{16}$$

Але з формул (10) і (11) випливає[1], що

$$t'_i = t_i - \frac{\tau_{is}}{\tau_{rs}} t_r \text{ для } i \neq r \text{ й } t'_r = \frac{1}{\tau_{rs}} t_r.$$

Підставивши ці вирази у співвідношення (16), дістанемо:

$$z' = \sum_{i=1\ (i\neq r)}^{m} \left(t_i - \frac{\tau_{is}}{\tau_{rs}} t_r\right) c_i + \frac{c_s}{\tau_{rs}} t_r = \sum_{i=1\ (i\neq r)}^{m} c_it_i - \frac{1}{\tau_{rs}} \left(\sum_{i=1\ (i\neq r)}^{m} \tau_{is}c_i\right) t_r + \frac{c_s}{\tau_{rs}} t_r,$$

[1] Повторивши для вектора b міркування, які провадилися для вектора p_j при виведенні формул (10) і (11), читач пересвідчиться, що й для x'_i та x_i° справедливі співвідношення (10) і (11).

тобто

$$z' = \sum_{i=1\ (i\neq r)}^{m} c_i t_i - \frac{1}{\tau_{rs}}\Big(\sum_{i=1\ (i\neq r)}^{m} \tau_{is} c_i\Big) t_r + \frac{c_s}{\tau_{rs}} t_r.$$

Отже,

$$z' = \Big(\sum_{i=1\ (i\neq r)}^{m} c_i t_i + c_r t_r\Big) - \Big[c_r t_r + \frac{1}{\tau_{rs}}\Big(\sum_{i=1\ (i\neq r)}^{m} \tau_{is} c_i\Big) t_r - \frac{c_s}{\tau_{rs}} t_r\Big] =$$

$$= \sum_{i=1}^{m} c_i t_i - \frac{1}{\tau_{rs}}\Big(\tau_{rs} c_r + \sum_{i=1\ (i\neq r)}^{m} \tau_{is} c_i - c_s\Big) t_r = z - \frac{1}{\tau_{rs}}\Big(\sum_{i=1}^{m} \tau_{is} c_i -$$

$$- c_s\Big) t_r = z - \frac{\zeta_s - c_s}{\tau_{rs}} t_r, \text{ тобто } z' = z - \frac{\zeta_s - c_s}{\tau_{rs}} t_r.$$

Віднявши від обох частин цієї рівності вектор c_0, дістанемо

$$z' - c_0 = (z - c_0) - \frac{\zeta_s - c_s}{\tau_{rs}} t_r.$$

Твердження доведено.

Із співвідношення (15) випливає, що рядок $z' - c_0$ ми дістаємо так само, як і інші рядки таблиці T_1: з рядка $z - c_0$ ми віднімаємо рядок t_r, помножений на такий множник, щоб в s-му стовпчику з'явився нуль. Доведемо тепер теореми, про які говорилося вище.

Теорема 1. *Якщо принаймні для одного з чисел* j $(j = 1, 2, \ldots, n)$ $\zeta_j - c_j > 0$, *то існує такий допустимий розв'язок* x', *що* $\zeta_0' < \zeta_0$, *де* ζ_0' — *значення цільової функції при* $x = x'$.

Доведення. Нехай для деякого числа k $(1 \leqslant k \leqslant n)$ $\zeta_k - c_k > 0$. Для допустимого розв'язку $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_m^\circ, 0, \ldots, 0)$ маємо:

$$x_1^\circ p_1 + x_2^\circ p_2 + \cdots + x_m^\circ p_m = b, \tag{17}$$

$$c_1 x_1^\circ + c_2 x_2^\circ + \cdots + c_m x_m^\circ = \zeta_0, \tag{18}$$

де ζ_0 — значення цільової функції при $x = x^\circ$.

Вектор p_k запишемо у вигляді лінійної комбінації векторів базису $B\{p_1, p_2, \ldots, p_m\}$:

$$\tau_{1k} p_1 + \tau_{2k} p_2 + \cdots + \tau_{mk} p_m = p_k \tag{19}$$

і запишемо рівність:

$$\tau_{1k} c_1 + \tau_{2k} c_2 + \cdots + \tau_{mk} c_m = \zeta_k. \tag{20}$$

Нехай α — деяке дійсне число. Віднімемо почленно від рівностей (17) і (18) відповідно рівності (19) і (20), помножені на α. Дістанемо:

$$(x_1^\circ - \alpha\tau_{1k}) p_1 + (x_2^\circ - \alpha\tau_{2k}) p_2 + \cdots + (x_m^\circ - \alpha\tau_{mk}) p_m = b - \alpha p_k, \tag{21}$$

$$(x_1^\circ - \alpha\tau_{1k}) c_1 + (x_2^\circ - \alpha\tau_{2k}) c_2 + \cdots + (x_m^\circ - \alpha\tau_{mk}) c_m = \zeta_0 - \alpha\zeta_k. \tag{22}$$

В рівності (21) член αp_k перенесемо в ліву частину, а до обох частин рівності (22) додамо αc_k, тоді матимемо:

$$(x_1^\circ - \alpha\tau_{1k}) p_1 + (x_2^\circ - \alpha\tau_{2k}) p_2 + \cdots + (x_m^\circ - \alpha\tau_{mk}) p_m + \alpha p_k = b, \tag{23}$$

$$(x_1^\circ - \alpha\tau_{1k}) c_1 + (x_2^\circ - \alpha\tau_{2k}) c_2 + \cdots + (x_m^\circ - \alpha\tau_{mk}) c_m + \alpha c_k = \zeta_0 - \alpha(\zeta_k - c_k). \tag{24}$$

Виберемо тепер дійсне число α так, щоб усі числа $x_1^\circ - \alpha\tau_{1k}$, $x_2^\circ - \alpha\tau_{2k}$, $\ldots$, $x_m^\circ - \alpha\tau_{mk}$, α були додатними. Оскільки числа $x_1^\circ, x_2^\circ, \ldots, x_m^\circ$ додатні, то таке число α можна вибрати [1].

З означення числа ζ_k і умови $\zeta_k - c_k > 0$ випливає, що $m < k \leqslant n$. За цієї умови рівності (23) і (24) означають, що $w = (x_1^\circ - \alpha\tau_{1k}, x_2^\circ - \alpha\tau_{2k}, \ldots, x_m^\circ - \alpha\tau_{mk}, 0, \ldots, 0, \alpha, 0, \ldots, 0)$ є допустимий розв'язок канонічної задачі, а $\zeta_0' = \zeta_0 - \alpha(\zeta_k - c_k)$ — відповідне йому значення цільової функції. Оскільки $\zeta_k - c_k > 0$ і $\alpha > 0$, то з рівності $\zeta_0' = \zeta_0 - \alpha(\zeta - c_k)$ випливає, що $\zeta_0' < \zeta_0$. Теорему доведено.

Таким чином, якщо хоч одна з різниць $\zeta_1 - c_1, \zeta_2 - c_2, \ldots, \zeta_s - c_s, \ldots, \zeta_n - c_n$ є додатне число, то за доведеною теоремою допустимий розв'язок $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_m^\circ, 0, \ldots, 0)$ не є оптимальним.

Теорема 2. *Якщо* $\zeta_j - c_j \leqslant 0$ *при* $j = 1, 2, \ldots, n$, *то допустимий розв'язок* $x^\circ = (x_1^\circ, x_2^\circ, \ldots, x_m^\circ, 0, \ldots, 0)$ *є оптимальним розв'язком.*

Доведення. Нехай $\zeta_j - c_j \leqslant 0$ для будь-якого j, $1 \leqslant j \leqslant n$. Оскільки система векторів $p_1, p_2, \ldots, p_m$ лінійно незалежна, то

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{vmatrix} \neq 0$$

і тому система рівнянь

$$\{a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m = c_j \ (j = 1, 2, \ldots, m)$$

має розв'язок $y^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$. Таким чином, справедливі рівності

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ = c_j \ (j = 1, 2, \ldots, m). \tag{25}$$

А при $j > m$ маємо:

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ = (p_j, y^\circ) = (\tau_{1j}p_1 + \tau_{2j}p_2 + \cdots + \tau_{mj}p_m, y^\circ) = \tau_{1j}(p_1, y^\circ) + \tau_{2j}(p_2, y^\circ) + \cdots + \tau_{mj}(p_m, y^\circ) = \tau_{1j}c_1 + \tau_{2j}c_2 + \cdots + \tau_{mj}c_m = \zeta_j \leqslant c_j,$$

---

[1] Справді, нехай x_s° — найменше серед додатних чисел $x_1^\circ, x_2^\circ, \ldots, x_m^\circ$, а τ_{lk} — найбільше серед чисел $\tau_{1k}, \tau_{2k}, \ldots, \tau_{mk}$, виберемо α так, щоб справджувалися нерівності $\alpha > 0$ і $x_s^\circ - \alpha\tau_{lk} > 0$, тоді всі числа $x_i^\circ - \alpha\tau_{ik}$ $(i = 1, 2, \ldots, m)$ будуть додатні.

оскільки $\zeta_j - c_j \leqslant 0$, тобто

$$a_{1j}y_1^\circ + a_{2j}y_2^\circ + \cdots + a_{mj}y_m^\circ \leqslant c_j \quad (j = m+1,\ m+2,\ \ldots,\ n). \quad (26)$$

Співвідношення (25) і (26) означають, що вектор $\boldsymbol{y}^\circ = (y_1^\circ, y_2^\circ, \ldots, y_m^\circ)$ задовольняє систему обмежень.

$$\{a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \leqslant c_j \quad (j = 1,\ 2,\ \ldots,\ n),$$

причому виконується умова: якщо $a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m < c_j$, то компонента x_j° в невід'ємному розв'язку $\boldsymbol{x}^\circ = (x_1^\circ,\ x_2^\circ,\ \ldots,\ x_m^\circ,\ 0,\ 0,\ \ldots,\ 0)$ системи обмежень $\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ $(i = 1,\ 2,\ \ldots,\ m)$ розглядуваної канонічної задачі мінімізації дорівнює 0. Тому за теоремою 4, п. 3.4 $\boldsymbol{x}^\circ = (x_1^\circ,\ x_2^\circ,\ \ldots,\ x_m^\circ,\ 0,\ 0,\ \ldots,\ 0)$ є оптимальний розв'язок розглядуваної канонічної задачі мінімізації, а $\boldsymbol{y}^\circ = (y_1^\circ,\ y_2^\circ,\ \ldots,\ y_m^\circ)$ — оптимальний розв'язок двоїстої задачі максимізації[1]. Теорему доведено.

Об'єднавши теореми 1 і 2, дістаємо такий *критерій оптимальності: допустимий базисний розв'язок $\boldsymbol{x}^\circ$ канонічної задачі мінімізації є оптимальним тоді і тільки тоді, коли всі різниці $\zeta_j - c_j$ $(j = 1,\ 2,\ \ldots,\ n)$ недодатні.* Отже, для знаходження оптимального розв'язку заміщенням векторів базису треба виконувати так, щоб через кілька кроків дістати вектор $\boldsymbol{z}^{(k)} - \boldsymbol{c}_0$, у якого всі компоненти $\zeta_j^{(k)} - c_j$ $(j = 1,\ 2,\ \ldots,\ n)$ недодатні. Звідси *перше правило* симплекс-методу: *якщо для деякого s $\zeta_s - c_s > 0$, то вектор $\boldsymbol{p}_s$ вводимо у новий базис.*

Постає питання: *який же вектор базису слід замінити вектором $\boldsymbol{p}_s$?*

Припустимо, що ми вводимо у новий базис вектор $\boldsymbol{p}_s$. Можливі два випадки: 1) всі числа τ_{is} $(i = 1,\ 2,\ \ldots,\ m)$ недодатні; 2) серед чисел τ_{is} є додатні. У першому випадку справедлива така теорема.

Теорема 3. *Нехай $\boldsymbol{p}_s$ — вектор, вибраний за першим правилом. Якщо всі числа τ_{is} $(i = 1,\ 2,\ \ldots,\ m)$ недодатні, то цільова функція розглядуваної задачі не має мінімуму.*

Доведення. За таблицею T

$$\boldsymbol{b} = \sum_{i=1}^{m} x_i^\circ \boldsymbol{p}_i \quad (27)$$

і

$$\boldsymbol{p}_s = \sum_{i=1}^{m} \tau_{is}\boldsymbol{p}_i.$$

З останньої рівності дістаємо:

$$\boldsymbol{0} = \boldsymbol{p}_s - \sum_{i=1}^{m} \tau_{is}\boldsymbol{p}_i. \quad (28)$$

[1] Замінивши розглядувану канонічну задачу мінімізації лінійної форми $f = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ канонічною задачею максимізації форми $-f = (-c_1)x_1 + (-c_2)x_2 + \cdots + (-c_n)x_n$ й застосувавши потім теорему 4, п. 13.4, дістанемо результат, про який щойно говорилося.

До рівності (27) додамо почленно рівність (28), помножену на деяке число $k > 0$. Дістанемо:

$$\boldsymbol{b} = \sum_{i=1}^{m} x_i^\circ \boldsymbol{p}_i + k\boldsymbol{p}_s - k\sum_{i=1}^{m} \tau_{is}\boldsymbol{p}_i.$$

Звідси

$$k\boldsymbol{p}_s + \sum_{i=1}^{m} (x_i^\circ - k\tau_{is})\,\boldsymbol{p}_i = \boldsymbol{b}. \quad (29)$$

Оскільки $\tau_{is} \leqslant 0$ $(i = 1,\ 2,\ \ldots,\ m)$, $x_i^\circ > 0$ $(i = 1,\ 2,\ \ldots,\ m)$ і $k > 0$, то $\boldsymbol{w} = (k,\ x_1^\circ - k\tau_{1s},\ x_2^\circ - k\tau_{2s},\ \ldots,\ x_m^\circ - k\tau_{ms})$, за рівністю (29), є допустимим розв'язком розглядуваної канонічної задачі мінімізації. Цей розв'язок залежить від системи векторів $\boldsymbol{p}_s, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_m$. Тому при $\boldsymbol{x} = \boldsymbol{w}$ значенням цільової функції $f(\boldsymbol{x}) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ буде $f(\boldsymbol{w}) = c_sk + \sum_{i=1}^{m} c_i(x_i^\circ - k\tau_{is}) = \sum_{i=1}^{m} c_ix_i^\circ + c_sk - k\sum_{i=1}^{m} c_i\tau_{is} = \sum_{i=1}^{m} c_ix_i^\circ + c_sk - k\zeta_s = \zeta_0 - k(\zeta_s - c_s)$, тобто $f(\boldsymbol{w}) = \zeta_0 - k(\zeta_s - c_s)$.

Оскільки за умовою теореми $\zeta_s - c_s > 0$ і k — довільно вибране додатне число, то $f(\boldsymbol{w})$ може бути як завгодно малим (від'ємним) числом. Отже, цільова функція $f(\boldsymbol{x})$ не має мінімуму. Теорему доведено.

Нехай серед чисел τ_{is} є додатні. Для додатних чисел τ_{is} обчислимо відношення $\frac{x_i^\circ}{\tau_{is}}$. Оскільки таких відношень скінченне число, то серед них знайдеться найменше. Припустимо, що найменшим є відношення $\frac{x_r^\circ}{\tau_{rs}}$. У такому разі вектором $\boldsymbol{p}_s$ будемо заміщувати базисний вектор $\boldsymbol{p}_r$. Отже, вектор базису, який заміщується вектором $\boldsymbol{p}_s$, вибираємо за таким правилом: *вирішивши ввести в новий базис вектор $\boldsymbol{p}_s$, обчислюємо відношення $\frac{x_i^\circ}{\tau_{is}}$ для $\tau_{is} > 0$ і заміщуємо той вектор $\boldsymbol{p}_r$ старого базису, для якого це відношення мінімальне. Це — друге правило симплекс-методу.*

За першим правилом симплекс-методу ми визначаємо стовпчик, в якому розміщений розв'язувальний коефіцієнт τ_{rs}, за другим правилом — рядок, в якому розміщений цей коефіцієнт.

Доведемо теорему, яка характеризує заміщення вибраного нами базисного вектора $\boldsymbol{p}_r$ вектором $\boldsymbol{p}_s$.

Теорема 4 (про поліпшення розв'язку). *Якщо замістимо відповідно до правил симплекс-методу базисний вектор $\boldsymbol{p}_r$ вектором $\boldsymbol{p}_s$, то знайдений базисний розв'язок $\boldsymbol{x}'$ буде допустимим, а відповідне йому значення ζ_0' цільової функції, яку треба мінімізувати, буде менше, ніж попереднє її значення ζ_0.*

Доведення. Доведемо спочатку, що знайдений в результаті заміщення вектора $\boldsymbol{p}_r$ вектором $\boldsymbol{p}_s$ базисний розв'язок $\boldsymbol{x}' = (x_1', x_2', \ldots, x_{r-1}', x_s', x_{r+1}', \ldots, x_m', 0, \ldots, 0)$, допустимий, тобто що $x_1', x_2', \ldots, x_{r-1}', x_s', x_{r+1}', \ldots, x_m'$ є невід'ємні числа. Оскільки вектор $\boldsymbol{p}_r$ заміщено вектором $\boldsymbol{p}_s$, то за формулами (10) і (11)

$$x_i' = x_i^\circ - \frac{\tau_{is}}{\tau_{rs}} x_r^\circ \text{ при } i \neq r, \tag{30}$$

$$x_r' = \frac{x_r^\circ}{\tau_{rs}}, \tag{31}$$

де $x_i^\circ \geqslant 0$, $x_r^\circ \geqslant 0$, $\tau_{rs} > 0$.

Із співвідношення (31) випливає, що $x_r' \geqslant 0$. Якщо $\tau_{is} \leqslant 0$, то з співвідношення (30) випливає, що $x_i' \geqslant 0$. Якщо $\tau_{is} > 0$, то $\frac{x_i^\circ}{\tau_{is}} \geqslant \frac{x_r^\circ}{\tau_{rs}}$, оскільки вектор $\boldsymbol{p}_r$ вибрано за другим правилом заміщення. Звідси $\tau_{rs}x_i^\circ \geqslant \tau_{is}x_r^\circ$ і тому $\tau_{rs}x_i^\circ - \tau_{is}x_r^\circ \geqslant 0$. Отже, $x_i' = x_i^\circ - \frac{\tau_{is}}{\tau_{rs}} x_r^\circ = \frac{\tau_{rs}x_i^\circ - \tau_{is}x_r^\circ}{\tau_{rs}} \geqslant 0$.

Доведемо тепер, що $\zeta_0' < \zeta_0$. За співвідношенням (15)

$$\zeta_0' = \zeta_0 - \frac{\zeta_s - c_s}{\tau_{rs}} x_r^\circ. \tag{32}$$

Ми провадимо всі міркування в припущенні, що виконується умова невиродженості. Тому компоненти $x_1^\circ, x_2^\circ, \ldots, x_m^\circ$ допустимого базисного розв'язку $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_m^\circ, 0, 0, \ldots, 0)$ додатні, зокрема й $x_r^\circ > 0$. У противному разі базисний розв'язок залежав би від меншого ніж m числа векторів $\boldsymbol{p}_i$, чого не може бути.

Оскільки вектори $\boldsymbol{p}_s$ і $\boldsymbol{p}_r$ вибрано за правилами симплекс-методу, то $\zeta_s - c_s > 0$ і $\tau_{rs} > 0$.

Отже, $\zeta_0' - \zeta_0 = -\frac{\zeta_s - c_s}{\tau_{rs}} x_r^\circ < 0$ і тому $\zeta_0' < \zeta_0$. Теорему доведено.

Припустимо тепер, що, виходячи з таблиці T, яка відповідає початковому базисному розв'язку $\boldsymbol{x}^\circ = (x_1^\circ, x_2^\circ, \ldots, x_m^\circ, 0, 0, \ldots, 0)$, ми послідовно заміщуємо за правилами симплекс-методу вектори базису. Постає питання: *коли закінчиться цей процес послідовного заміщення?*

Якщо на певному етапі заміщень створиться ситуація, про яку говориться в теоремі 3, то цільова функція канонічної задачі мінімізації необмежена знизу і тому задача оптимального розв'язку не має.

Якщо ж у процесі послідовних заміщень така ситуація не створюється, то через скінченне число заміщень ми прийдемо до таблиці T_k, яка задовольняє критерій оптимальності[1] і, таким чином, знайдемо оптимальний розв'язок канонічної задачі. Справді, за теоремою 4, значення ζ_0 цільової функції при кожному черговому заміщенні зменшується. Тому в процесі послідовних заміщень ми не можемо одержати той самий базис більше одного разу і, отже, одержуємо різні базиси. Число послідовних заміщень, таким чином, не більше, ніж число всіх можливих базисів системи векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$, тобто не більше ніж A_n^m (число розміщень з n елементів по m елементів).

Отже, процес послідовного заміщення векторів базису через скінченне число кроків закінчиться. Але в розглядуваному випадку це можливо лише тоді, коли ми прийдемо до таблиці T_k, у додаткового вектора-рядка $\boldsymbol{z}^{(k)} - \boldsymbol{c}_0$ якої всі компоненти $\zeta_j^{(k)} - c_j$ недодатні. Базисний розв'язок, знайдений за таблицею T_k, буде оптимальним розв'язком канонічної задачі.

Таким чином, якщо канонічна задача мінімізації має оптимальний розв'язок, то в процесі розв'язування її симплекс-методом не може створитись ситуація, яка розглядається в теоремі 3, і, отже, її можна розв'язати за допомогою скінченного числа заміщень.

На закінчення зробимо кілька зауважень. Користуючись правилами симплексного методу для вибору векторів $\boldsymbol{p}_s$ і $\boldsymbol{p}_r$, слід мати на увазі таке. Може трапитися, що серед компонент $\zeta_j^{(k)} - c_j$ додаткового вектора-рядка $\boldsymbol{z}^{(k)} - \boldsymbol{c}_0$ таблиці T_k додатних буде кілька. Тоді кожен з векторів $\boldsymbol{p}_j$, що відповідає додатній різниці $\zeta_j^{(k)} - c_j$, можна ввести в новий базис.

Однак щоб скоротити процес обчислень, слід ввести у новий базис той із них, який відповідає найбільшій із різниць $\zeta_j^{(k)} - c_j$. Якщо є декілька однакових найбільших додатних різниць $\zeta_j^{(k)} - c_j$, то можна вводити в новий базис будь-який з векторів, що відповідають цим найбільшим різницям. Якщо при застосуванні другого правила симплекс-методу виявиться, що є декілька однакових найменших відношень $\frac{x_i^{(k)}}{t_{is}^{(k)}}$, то можна вивести з старого базису будь-який з векторів, що відповідають цим найменшим відношенням.

При викладі симплекс-методу ми вважали, що виконується умова невиродженості. Використали ми це лише при доведенні теореми 4. Якщо умова невиродженості не виконується, то теорема 4 не буде справедливою. Застосування другого правила симплекс-методу в цьому випадку може не дати поліпшення розв'язку[1].

4.3. Приклади розв'язування задач симплекс-методом. Розглянемо тепер кілька прикладів розв'язування задач лінійного програмування симплекс-методом.

1. Знайти невід'ємний вектор $\boldsymbol{l} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$, який задовольняє систему обмежень

$$\begin{cases} x_1 + 2x_2 + x_3 + 2x_4 + x_5 = 8, \\ -4x_1 + 7x_2 - x_3 - 5x_4 + 2x_5 = 7 \end{cases}$$

і мінімізує цільову функцію

$$f(x) = 4x_1 - 2x_2 + x_3 - 4x_4 + x_5.$$

[1] Таблиця T_k задовольняє критерій оптимальності — це означає, що всі елементи її додаткового рядка є недодатні числа.

[1] Про випадок виродженості див., наприклад, Д. Гейл, Теория линейных экономических моделей. М., Изд-во иностр. лит., 1963.

Розв'язання. Відповідно до прийнятих нами позначень маємо:

$$A=\begin{pmatrix}1 & 2 & 1 & 2 & 1\\ -4 & 7 & -1 & -5 & 2\end{pmatrix},\quad c=(4,\ -2,\ 1,\ -4,\ 1);\quad x=(x_1,\ x_2,\ x_3,\ x_4,\ x_5);$$

$$p_1=(1,\ -4),\quad p_2=(2,\ 7),\quad p_3=(1,\ -1),\quad p_4=(2,\ -5),\quad p_5=(1,\ 2).$$

$$b=(8,\ 7).$$

Легко показати, що вектори p_1 і p_5 утворюють базис системи векторів p_1, p_2, p_3, p_4, p_5. Невід'ємним базисним розв'язком, що відповідає базису $B\{p_1, p_5\}$, є вектор $x^0=\left(\frac{3}{2},\ 0,\ 0,\ 0,\ \frac{13}{2}\right)$. Виходячи з цього базисного розв'язку, відшукуватимемо тепер за симплекс-методом оптимальний розв'язок задачі. Щоб скласти таблицю T векторів p_1, p_2, p_3, p_4, p_5, b відносно базису $B\{p_1, p_5\}$, запишемо кожен з цих векторів у вигляді лінійної комбінації векторів p_1 і p_5. Матимемо:

$$p_1=p_1+0\cdot p_5,\quad p_2=-\frac{1}{2}p_1+\frac{5}{2}p_5,\quad p_3=\frac{1}{2}p_1+\frac{1}{2}p_5,\quad p_4=\frac{3}{2}p_1+$$

$$+\frac{1}{2}p_5,\quad p_5=0\cdot p_1+p_5,\quad b=\frac{3}{2}p_1+\frac{13}{2}p_5.$$

Отже таблиця T має вигляд:

		p_1	p_2	p_3	p_4	p_5	b	
T	p_1	1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	0	$\frac{3}{2}$	$c_1=4$
	p_5	0	$\frac{5}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{13}{2}$	$c_5=1$

Знайдемо тепер вектор $z=(\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_0)$. За формулою (14) $\zeta_j=c_1\tau_{1j}+$ $+c_5\tau_{5j}$ $(j=1, 2, 3, 4, 5)$; $\zeta_0=c_1x_1^0+c_5x_5^0$.

Отже, $\zeta_1=4$, $\zeta_2=\frac{1}{2}$, $\zeta_3=\frac{5}{2}$, $\zeta_4=\frac{13}{2}$, $\zeta_5=1$, $\zeta_0=\frac{25}{2}$,

$z=\left(4,\ \frac{1}{2},\ \frac{5}{2},\ \frac{13}{2},\ 1,\ \frac{25}{2}\right)$.

Оскільки вектор $c_0=(c_1, c_2, c_3, c_4, c_5, 0)=(4,\ -2,\ 1,\ -4,\ 1,\ 0)$, то $z-$ $-c_0=\left(0,\ \frac{5}{2},\ \frac{3}{2},\ \frac{21}{2},\ 0,\ \frac{25}{2}\right)$. Доповнимо таблицю T рядком $z-c_0$. Матимемо:

		p_1	p_2	p_3	p_4	p_5	b
T	p_1	1	$-\frac{1}{2}$	$\frac{1}{2}$	$\boxed{\frac{3}{2}}$	0	$\frac{3}{2}$
	p_5	0	$\frac{5}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{13}{2}$
	$z-c_0$	0	$\frac{5}{2}$	$\frac{3}{2}$	$\frac{21}{2}$	0	$\frac{25}{2}$

Застосуємо перше правило симплекс-методу. У рядку $z-c_0$ є три додатні елементи ζ_j-c_j: $\zeta_2-c_2=\frac{5}{2}$, $\zeta_3-c_3=\frac{3}{2}$, $\zeta_4-c_4=\frac{21}{2}$. Отже, в новий базис можна вводити будь-який з векторів p_2, p_3, p_4. Вводитимемо вектор p_4, який відповідає найбільшій додатній різниці ζ_j-c_j.

Тепер застосуємо друге правило симплекс-методу. У нашому випадку обидва елементи τ_{14} і τ_{54} — додатні. Тому обчислюємо відношення $\frac{x_1^0}{\tau_{14}}$ і $\frac{x_5^0}{\tau_{54}}:\frac{x_1^0}{\tau_{14}}=$ $=\frac{3}{2}:\frac{3}{2}=1$, $\frac{x_5^0}{\tau_{54}}=\frac{13}{2}:\frac{1}{2}=13$.

Отже, вектором p_4 треба заміщати в старому базисі вектор p_1.

Таким чином, розв'язувальним коефіцієнтом буде $\tau_{14}=\frac{3}{2}$. Виконавши заміщення вектора p_1 вектором p_4, дістаємо таблицю

		p_1	p_2	p_3	p_4	p_5	b
T_1	p_4	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	1	0	1
	p_5	$-\frac{1}{3}$	$\boxed{\frac{8}{3}}$	$\frac{1}{3}$	0	1	6
	$z'-c_0$	−7	6	−2	0	0	2

За правилами симплексного методу в новий базис треба ввести вектор p_2, а з старого базису слід вивести вектор p_5. Отже, розв'язувальним коефіцієнтом буде $\tau_{52}=\frac{8}{3}$. Виконавши чергове заміщення, дістанемо таблицю

		p_1	p_2	p_3	p_4	p_5	b
T_2	p_4	$\frac{5}{8}$	0	$\frac{3}{8}$	1	$\frac{1}{8}$	$\frac{7}{4}$
	p_2	$\frac{-1}{8}$	1	$\frac{1}{8}$	0	$\frac{3}{8}$	$\frac{9}{4}$
	$z''-c_0$	$\frac{-25}{4}$	0	$\frac{-11}{4}$	0	$\frac{-9}{4}$	$\frac{-23}{2}$

Оскільки в цій таблиці всі різниці ζ_j-c_j недодатні, то за критерієм оптимальності допустимий базисний розв'язок $l=\left(0,\ \frac{9}{4},\ 0,\ \frac{7}{4},\ 0\right)$ є оптимальним.

Значення $f\underset{\min}{(x)}$ уже також обчислено; воно дорівнює останній компоненті вектора $z''-c_0$, тобто $f\underset{\min}{(x)}=-\frac{23}{2}$.

2. Знайти невід'ємний вектор $l = (\lambda_1, \lambda_2, \lambda_3)$, який задовольняє систему обмежень

$$\begin{cases} -x_1 - x_2 - 2x_3 \leqslant 5, \\ 2x_1 - 3x_2 + x_3 \leqslant 3, \\ 2x_1 - 5x_2 + 6x_3 \leqslant 5 \end{cases} \tag{33}$$

і мінімізує цільову функцію

$$f(x) = -x_1 + 3x_2 - 2x_3.$$

Розв'язання. Запишемо систему обмежень (33) у вигляді системи рівнянь, ввівши до кожної нерівності додаткове невід'ємне невідоме. Матимемо:

$$\begin{cases} -x_1 - x_2 - 2x_3 + x_4 + 0 \cdot x_5 + 0 \cdot x_6 = 5, \\ 2x_1 - 3x_2 + x_3 + 0 \cdot x_4 + x_5 + 0 \cdot x_6 = 3, \\ 2x_1 - 5x_2 + 6x_3 + 0 \cdot x_4 + 0 \cdot x_5 + x_6 = 5. \end{cases} \tag{34}$$

Стандартна задача мінімізації зводиться, таким чином, до канонічної задачі мінімізації: знайти вектор $l = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$, який задовольняє систему рівнянь (34) і мінімізує цільову функцію

$$\varphi(x) = -x_1 + 3x_2 - 2x_3 + 0 \cdot x_4 + 0 \cdot x_5 + 0 \cdot x_6.$$

Розв'язуватимемо цю задачу симплекс-методом. Запишемо потрібні нам вектори: $\boldsymbol{p}_1 = (-1, 2, 2)$, $\boldsymbol{p}_2 = (-1, -3, -5)$, $\boldsymbol{p}_3 = (-2, 1, 6)$, $\boldsymbol{p}_4 = (1, 0, 0)$, $\boldsymbol{p}_5 = (0, 1, 0)$, $\boldsymbol{p}_6 = (0, 0, 1)$, $\boldsymbol{b} = (5, 3, 5)$, $\boldsymbol{c} = (-1, 3, -2, 0, 0, 0)$.

Вектори $\boldsymbol{p}_4$, $\boldsymbol{p}_5$, $\boldsymbol{p}_6$ є одиничними векторами простору $\hat{V}_3$ і утворюють базис цього простору, а отже, і системи векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4, \boldsymbol{p}_5, \boldsymbol{p}_6$. Вектор $x^0 = (0, 0, 0, 5, 3, 5)$, очевидно, є невід'ємним базисним розв'язком задачі, що відповідає базису $B\{\boldsymbol{p}_4, \boldsymbol{p}_5, \boldsymbol{p}_6\}$.

Таблиця T векторів $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$, $\boldsymbol{p}_4$, $\boldsymbol{p}_5$, $\boldsymbol{p}_6$, $\boldsymbol{b}$ відносно базису $B\{\boldsymbol{p}_4, \boldsymbol{p}_5, \boldsymbol{p}_6\}$ буде така:

		$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\boldsymbol{p}_3$	$\boldsymbol{p}_4$	$\boldsymbol{p}_5$	$\boldsymbol{p}_6$	$\boldsymbol{b}$	
T	$\boldsymbol{p}_4$	−1	−1	−2	1	0	0	5	$c_4 = 0$
	$\boldsymbol{p}_5$	2	−3	1	0	1	0	3	$c_5 = 0$
	$\boldsymbol{p}_6$	2	−5	6	0	0	1	5	$c_6 = 0$

Обчисливши компоненти $\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_6, \zeta_0$, знаходимо: $\zeta_1 = \zeta_2 = \zeta_3 = \zeta_4 = \zeta_5 = \zeta_6 = \zeta_0 = 0$. Отже, $\boldsymbol{z} = (0, 0, 0, 0, 0, 0, 0)$, $\boldsymbol{z} - \boldsymbol{c}_0 = (1, -3, 2, 0, 0, 0, 0)$.

Складаємо тепер розширену таблицю T і будемо провадити послідовне заміщення за симплекс-методом, виконуючи відповідні

обчислення. Результати цих обчислень запишемо так:

		$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\boldsymbol{p}_3$	$\boldsymbol{p}_4$	$\boldsymbol{p}_5$	$\boldsymbol{p}_6$	$\boldsymbol{b}$
T	$\boldsymbol{p}_4$	−1	−1	−2	1	0	0	5
	$\boldsymbol{p}_5$	2	−3	1	0	1	0	3
	$\boldsymbol{p}_6$	2	−5	[6]	0	0	1	5
	$\boldsymbol{z} - \boldsymbol{c}$	1	−3	2	0	0	0	0
T_1	$\boldsymbol{p}_4$	$-\frac{1}{3}$	$-\frac{8}{3}$	0	1	0	$\frac{1}{3}$	$\frac{20}{3}$
	$\boldsymbol{p}_5$	[$\frac{5}{3}$]	$-\frac{13}{6}$	0	0	1	$-\frac{1}{6}$	$\frac{13}{6}$
	$\boldsymbol{p}_3$	$\frac{1}{3}$	$-\frac{5}{6}$	1	0	0	$\frac{1}{6}$	$\frac{5}{6}$
	$\boldsymbol{z}' - \boldsymbol{c}$	$\frac{1}{3}$	$-\frac{4}{3}$	0	0	0	$-\frac{1}{3}$	$-\frac{5}{3}$
T_2	$\boldsymbol{p}_4$	0	$-\frac{31}{10}$	0	1	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{71}{10}$
	$\boldsymbol{p}_1$	1	$-\frac{13}{10}$	0	0	$\frac{3}{5}$	$-\frac{1}{10}$	$\frac{13}{10}$
	$\boldsymbol{p}_3$	0	$-\frac{2}{5}$	1	0	$-\frac{1}{5}$	$\frac{1}{5}$	$\frac{2}{5}$
	$\boldsymbol{z}'' - \boldsymbol{c}$	0	$-\frac{9}{10}$	0	0	$-\frac{1}{3}$	$-\frac{3}{10}$	$-\frac{21}{10}$

В останньому рядку таблиці T_2 всі елементи $\zeta_l'' - c_l$ недодатні. Отже, вектор $l' = \left(\frac{3}{10}; 0; \frac{2}{5}; \frac{71}{10}; 0; 0\right)$ є оптимальним розв'язком канонічної задачі мінімізації, до якої ми звели задану стандартну задачу. У такому разі вектор $l = \left(\frac{13}{10}; 0; \frac{2}{5}\right)$ є оптимальним розв'язком заданої стандартної задачі мінімізації. Значення стандартної задачі мінімізації $\min f(x) = -\frac{21}{10}$.

Примітка. Зауважимо, що в задачах лінійного програмування, в яких система обмежень має вигляд $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n + x_i = b_i$ $(i = 1, 2, \ldots, m)$ $(b_i > 0)$ проблеми відшукання вихідного базисного розв'язку не існує. Вектори $\boldsymbol{p}_{n+1}, \boldsymbol{p}_{n+2}, \ldots, \boldsymbol{p}_{n+m}$ є одиничні вектори простору $\hat{V}_m$ і, отже, вони утворюють базис системи векторів $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{n+m}$. Базису $B\{\boldsymbol{p}_{n+1}, \boldsymbol{p}_{n+2}, \ldots, \boldsymbol{p}_{n+m}\}$ відповідає невід'ємний базисний розв'язок $x^0 = (0, 0, \ldots, 0, b_1, b_2, \ldots, b_m)$. Тому, взявши базис $B\{\boldsymbol{p}_{n+1}, \boldsymbol{p}_{n+2}, \ldots, \boldsymbol{p}_{n+m}\}$ за вихідний, можна зразу приступати до заміщення векторів базису.

4.4. Знаходження невід'ємних розв'язків системи лінійних рівнянь. При вивченні симплекс-методу ми припускали, що нам відомий деякий допустимий базисний розв'язок канонічної задачі мінімізації, тобто невід'ємний базисний розв'язок системи обмежень цієї задачі. Проте розв'язування кожної конкретної канонічної задачі симплекс-методом розпочинається саме з відшукання невід'ємного базисного розв'язку її системи обмежень, тобто системи рівнянь вигляду

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \qquad (i = 1, 2, \ldots, m). \qquad (35)$$

Тому треба з'ясувати, як знайти невід'ємний базисний розв'язок системи (35). Звичайно можна було б спробувати відшукувати всі базисні розв'язки системи (35) з надією на те, що на певному етапі знайдемо й невід'ємний базисний розв'язок. Проте значно легше розв'язати цю задачу, якщо звести її до канонічної задачі мінімізації. Покажемо, як це зробити. Вважатимемо, що в системі рівнянь (35) всі вільні члени невід'ємні, тобто $b_i \geqslant 0$ при $i = 1, 2, \ldots, m$. Цього завжди можна досягти, помноживши кожне рівняння, у якого вільний член від'ємний, на -1.

Розглянемо таку канонічну задачу мінімізації: знайти невід'ємний вектор $\bar{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0, x_{n+1}^0, x_{n+2}^0, \ldots, x_{n+m}^0)$, який задовольняє систему обмежень

$$\{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + x_{n+i} = b_i \qquad (i = 1, 2, \ldots, m)$$

і мінімізує цільову функцію

$$f(\bar{x}) = x_{n+1} + x_{n+2} + \cdots + x_{n+m}. \qquad (36)$$

Називатимемо цю задачу задачею I.

Доведемо, що *система рівнянь* (35) *має невід'ємний розв'язок тоді і тільки тоді, коли значення задачі* I *дорівнює нулю.*

Справді, якщо система (35) має невід'ємний розв'язок $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$, то задача I має допустимий розв'язок $x^0 = (x_1^0, x_2^0, \ldots, x_n^0, 0, 0, \ldots, 0)$. Значення цільової функції (36) при $\bar{x} = \bar{x}^0$ дорівнює 0.

Для будь-якого іншого допустимого розв'язку задачі I значення цільової функції (36), очевидно, не менше, ніж нуль. Тому $\bar{x}^0$ є оптимальним розв'язком задачі I, і значення її $f(\bar{x}^0) = 0$.

Навпаки, якщо $\bar{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0, x_{n+1}^0, \ldots, x_{n+m}^0$ є оптимальним розв'язком задачі I і $f(\bar{x}^0) = 0$, то, за співвідношенням (36), $x_{n+1}^0 = x_{n+2}^0 = \ldots = x_{n+m}^0 = 0$ і, отже, $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ є невід'ємний розв'язок системи (35). Таким чином, якщо $\bar{x}^0 = (x_1^0, x_2^0, \ldots$ $\ldots, x_n^0, x_{n+1}^0, \ldots, x_{n+m}^0$ — оптимальний розв'язок задачі I і значення її $f(\bar{x}^0) = 0$, то $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ є невід'ємний розв'язок системи (35); якщо ж значення задачі I не дорівнює нулю, то система (35) невід'ємних розв'язків не має.

Приклад. Знайти невід'ємний розв'язок системи лінійних рівнянь

$$\begin{cases} x_1 + 4x_2 - x_4 = 3, \\ 2x_1 - x_3 = 3, \\ 3x_1 - x_2 - 2x_3 = 1. \end{cases}$$

Розв'язання. Відповідна канонічна задача мінімізації формулюється так: знайти вектор $\bar{x}^0 = (x_1^0, x_2^0, x_3^0, x_4^0, x_5^0, x_6^0, x_7^0)$, який задовольняє систему обмежень

$$\begin{cases} x_1 + 4x_2 - x_4 + x_5 = 3, \\ 2x_1 - x_3 + x_6 = 3, \\ 3x_1 - x_2 - 2x_3 + x_7 = 1 \end{cases}$$

і мінімізує цільову функцію

$$f(\bar{x}) = x_5 + x_6 + x_7.$$

Відповідно до прийнятих позначень маємо:

$$A = \begin{pmatrix} 1 & 4 & 0 & -1 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 0 & 1 & 0 \\ 3 & -1 & -2 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{array}{l} p_1 = (1, 2, 3), \ p_2 = (4, 0, -1), \\ p_3 = (0, -1, -2), \ p_4 = (-1, 0, 0), \\ p_5 = (1, 0, 0), \ p_6 = (0, 1, 0), \end{array}$$

$p_7 = (0, 0, 1)$, $c = (0, 0, 0, 0, 1, 1, 1)$, $b = (3, 3, 1)$, $\bar{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$.

Вектори p_5, p_6, p_7 є одиничними векторами простору $\hat{V}_3$. Позначимо їх відповідно $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Візьмемо за вихідний базис систему векторів e_1, e_2, e_3 і складемо таблицю T (з додатковим рядком) векторів $p_1, p_2, p_3, p_4, b, e_1, e_2, e_3$ відповідно базису $B\{e_1, e_2, e_3\}$.

Далі складаємо за симплекс-методом послідовність таблиць

		p_1	p_2	p_3	p_4	b	e_1	e_2	e_3
T	e_1	1	4	0	−1	3	1	0	0
	e_2	2	0	−1	0	3	0	1	0
	e_3	3	−1	−2	0	1	0	0	1
	$z - c_0$	6	3	−3	−1	7	0	0	0

		p_1	p_2	p_3	p_4	b	e_1	e_2	e_3
T_1	e_1	0	$\frac{13}{3}$	$\frac{2}{3}$	−1	$\frac{8}{3}$	1	0	$-\frac{1}{3}$
	e_2	0	$\frac{2}{3}$	$\frac{1}{3}$	0	$\frac{7}{3}$	0	1	$-\frac{2}{3}$
	p_1	1	$-\frac{1}{3}$	$-\frac{2}{3}$	0	$\frac{1}{3}$	0	0	$\frac{1}{3}$
	$z - c_0$	0	5	1	1	5	0	0	−2

ЦІЛІ ЧИСЛА Й ОСНОВИ ТЕОРІЇ ПОДІЛЬНОСТІ

§ 5. ОСНОВНІ ПОНЯТТЯ Й ТЕОРЕМИ ТЕОРІЇ ПОДІЛЬНОСТІ

5.1. Означення й основні властивості подільності. У множині натуральних чисел **N** визначені операції додавання і множення, але не завжди здійсненні обернені їм операції віднімання й ділення. Щоб операція віднімання завжди була здійсненна, у математиці вводять число 0 і цілі від'ємні числа —1, —2, —3, ... Натуральні числа, число 0 і цілі від'ємні числа утворюють *множину цілих чисел,* яку ми позначатимемо буквою **Z**.

У множині цілих чисел **Z** визначені операції додавання і множення й здійсненна операція віднімання. Операції додавання і множення цілих чисел, як відомо, асоціативні, комутативні й пов'язані дистрибутивним законом. Отже, множина цілих чисел **Z** є комутативне кільце; його називають *кільцем цілих чисел.* Ми не будемо торкатися питання побудови кільця цілих чисел, оскільки воно розглядається в курсі «Наукові основи курсу шкільної математики», а відразу перейдемо до вивчення деяких властивостей цілих чисел. Розглянемо насамперед питання про подільність цілих чисел.

Означення. *Якщо для цілих чисел a і b в кільці цілих чисел* **Z** *існує таке число q, що $a = bq$, то кажуть, що a ділиться на b або b ділить a і пишуть відповідно $a \vdots b$, b/a.* Число a при цьому називають *кратним* числа b, а b називають *дільником* a. Зрозуміло, якщо $a = bq$, то число q також є дільником числа a. Якщо в кільці **Z** не існує числа q, такого, що $a = bq$, то кажуть, що *a не ділиться* на b або *b не ділить a.*

Наведемо деякі властивості подільності цілих чисел, що випливають з означення.

1. $\forall\limits_{a\in Z} [0 \vdots a]$, оскільки $0 \cdot a = 0$.

2. $\forall\limits_{a\in Z} [a \vdots a \wedge a \vdots (-a) \wedge a \vdots 1 \wedge a \vdots (-1)]$, оскільки $a = a \cdot 1$ і $a = (-a) \cdot (-1)$.

3. $\forall\limits_{a\in Z} [a \vdots b] \Rightarrow a \vdots (-b) \wedge (-a) \vdots b \wedge (-a) \vdots (-b)]$.

Справді, $a = bq \Rightarrow a = (-b)(-q) \wedge (-a) = b(-q) \wedge (-a) = (q-b)q$.

4. $\forall\limits_{a,b,c\in Z} [a \vdots b \wedge b \vdots c \Rightarrow a \vdots c]$. Справді, $a = bq_1 \wedge b = cq_2 \Rightarrow a = c(q_2 \times q_1)$ і, отже, $a \vdots c$.

5. $\forall\limits_{a,b,c\in Z} [a \vdots c \wedge b \vdots c \Rightarrow (a+b) \vdots c \wedge (a-b) \vdots c]$.

Справді, $a = cq_1 \wedge b = cq_2 \Rightarrow (a+b) = c(q_1+q_2) \wedge (a-b) = c(q_1-q_2)$ і, отже, $(a+b) \vdots c$ і $(a-b) \vdots c$.

6. $\forall\limits_{a,b,c\in Z} [a \vdots b \Rightarrow ac \vdots b]$, оскільки $a = bq \Rightarrow ac = b(qc)$.

7. $\forall\limits_{a_1,b_1,a_2,b_2,\ldots a_n,b_n,c\in Z} [a_1 \vdots c \wedge a_2 \vdots c \wedge \cdots \wedge a_n \vdots c \Rightarrow (a_1b_1 + a_2b_2 + \cdots + a_nb_n) \vdots c]$.

Ця властивість є безпосереднім наслідком властивостей 6 і 5.

5.2. Ділення з остачею. Важливу роль у теорії подільності цілих чисел відіграє така теорема.

Теорема 1. (Про ділення з остачею). *Які б не були цілe число a і натуральне число b, завжди існує єдина пара цілих чисел q і r, така, що*

$$a = bq + r, \tag{1}$$

$$0 \leqslant r < b. \tag{2}$$

Число q називають *неповною часткою, r — остачею.*

Доведення. Доведемо спочатку, що числа q і r, які задовольняють умови (1) і (2), існують. Справді, якщо ціле число a ділиться на натуральне число b, тобто $a = bc$, де c — деяке ціле число, то, очевидно, числа $q = c$ і $r = 0$ задовольняють умови (1) і (2). Припустимо, що ціле число a не ділиться на b. Тоді можливі такі два випадки: 1) $|a| < b$; 2) $|a| > b$. Розглянемо кожен з цих випадків.

1) Нехай $|a| < b$. Якщо при цьому $a > 0$, то з рівності $a = b \cdot 0 + a$ випливає, що числа $q = 0$, $r = a$ задовольняють умови теореми. Якщо ж $a < 0$, то $0 < b + a < b$ і тому з рівності $a = b(-1) + (b + a)$ випливає, що умови теореми задовольняють числа $q = -1$, $r = a + b$. Отже, в першому випадку завжди існує пара чисел q і r, яка задовольняє умови (1) і (2).

2) Нехай тепер $|a| > b$. Якщо при цьому $a > 0$, то з рівності $a = b \cdot 1 + (a - b)$ випливає, що існують ціле число $q = 1$ і натуральне число $r = a - b$, такі, що

$$a = bq + r.$$

Якщо ж $a < 0$, то $b(-a) + a > 0$ і з рівності $a = b \cdot a + [b(-a) + a]$ випливає, що рівність (1) справджується для цілого числа $q = a$ і натурального числа $r = b(-a) + a$. Отже, у другому випадку завжди існують ціле число q і натуральне число r, які задовольняють умову (1). Нехай M — множина всіх тих і тільки тих натуральних чисел r, для кожного з яких при належному виборі цілого числа q справджується рівність (1). За принципом найменшого числа у множині натуральних чисел M є найменше число.

Нехай цим найменшим числом є r_1. Тоді $a = bq_1 + r_1$, де r_1 — деяке натуральне число. Покажемо, що $r_1 < b$. Справді, припустимо, що $r_1 > b$. Тоді $r_1 = b + r_2$, де r_2 — деяке натуральне число, і $a = bq_1 + r_1 = bq_1 + b + r_2 = b(q_1 + 1) + r_2$, тобто $a = b(q_1 + 1) + r_2$, причому $r_2 < r_1$. Отже, числа $q = q_1 + 1$ і $r = r_2$ задовольняють умову (1) і тому $r_2 \in M$, чого не може бути, оскільки r_1 — найменше з чисел множини M. Якщо припустити, що $r_1 = b$, то $a = bq_1 + b = b(q_1 + 1)$, тобто a ділиться на b, а це суперечить умові. Отже, і в другому випадку існує пара чисел q і r, яка задовольняє умову теореми. Таким чином, в усіх можливих випадках існує пара чисел q і r, така, що $a = bq + r$, $0 \leqslant r < b$.

Покажемо тепер, що ця пара єдино можлива. Нехай, крім пари чисел q і r, існує ще пара чисел q_1 і r_1, яка також задовольняє умову теореми. Тоді $a = bq + r$, $0 \leqslant r < b$ і $a = bq_1 + r_1$, $0 \leqslant r_1 < b$. Припустимо, що остачі r і r_1 різні. Для визначеності вважатимемо, що $r_1 > r$. З рівностей $a = bq + r$ і $a = bq_1 + r_1$ випливає, що

$$bq + r = bq_1 + r_1; \tag{3}$$

$$b(q - q_1) = r_1 - r. \tag{4}$$

Оскільки $b > r_1 > r$ і $r \geqslant 0$, то $0 < r_1 - r < b$.

Отже, права частина рівності (4) є натуральне число, а тому й ліва її частина також є натуральне число. Звідси випливає, що $q > q_1$, бо при $q = q_1$ ліва частина рівності дорівнює нулю, а при $q < q_1$ вона є цілим від'ємним числом.

Оскільки $q > q_1$, то $q - q_1 \geqslant 1$ і тому $b(q - q_1) \geqslant b$. Таким чином, ми прийшли до суперечності: число $r_1 - r$, яке менше ніж b, дорівнює числу $b(q - q_1)$, не меншому ніж b. Отже, остачі r і r_1 не можуть бути різними. Оскільки $r = r_1$, то з рівності (3) випливає, що й $q = q_1$. Цим єдиність пари чисел q і r, а отже, і теорему доведено.

Зауважимо, що теорему про ділення з остачею можна поширити й на випадок цілого від'ємного числа b. Справді, якщо $b < 0$, то $b = -|b|$. За щойно доведеною теоремою, існує єдина пара цілих чисел q і r, така, що $a = |b| q_1 + r$, $0 \leqslant r < |b|$. Звідси $a = bq + r$, де

$$q = -q_1, \quad 0 \leqslant r < |b|. \tag{5}$$

Пара чисел q і r, що задовольняє умови (5), очевидно, існує тільки одна. Отже, справедлива така теорема.

Теорема 2. *Для будь-яких цілих чисел a і b, де $b \neq 0$, існує одна і тільки одна пара чисел q і r, така, що*

$$a = bq + r, \quad 0 \leqslant r < |b|.$$

Приклади. **1.** Нехай $a = 347$, $b = 18$. Маємо

$$347 = 18 \cdot 19 + 5; \quad 0 < 5 < 18.$$

2. Нехай $a = -347$, $b = 18$. Маємо

$$-347 = 18 \cdot (-20) + 13; \quad 0 < 13 < 18.$$

3. Нехай $a = -334$, $b = -17$. Маємо

$$-334 = (-17) \cdot 20 + 6, \quad 0 < 6 < |17|.$$

З теореми 2 випливає такий наслідок.

Наслідок. *Ціле число a тоді і тільки тоді кратне цілому числу $b \neq 0$, коли остача від ділення a на b дорівнює нулю.*

Справді, якщо остача r від ділення a на b дорівнює нулю, то $a = b \cdot q$ і, отже, a кратне b. Навпаки, якщо a кратне b, то $a = b \cdot c$, де c — деяке ціле число. Оскільки частка q і остача r від ділення a на b визначені однозначно, то з останньої рівності випливає, що $q = c$, а $r = 0$.

5.3. Найбільший спільний дільник двох чисел і алгоритм Евкліда. За властивістю 3 подільності цілих чисел з подільності цілого числа a на ціле число b випливає подільність $\pm a$ на $\pm b$. Тому при вивченні

питання про подільність цілих чисел можна обмежитися розглядом лише цілих додатних чисел. Зокрема, всюди далі ми розглядатимемо лише додатні дільники цілих чисел.

Означення. *Ціле число δ називається спільним дільником цілих чисел a і b, якщо кожне з цих чисел ділиться на δ. Найбільший із спільних дільників чисел a і b називають найбільшим спільним дільником цих чисел* (скорочено НСД) *і позначають символом* (a, b).

Цілі числа a і b, найбільший спільний дільник яких дорівнює 1, називають *взаємно простими*.

Безпосередньо з означень спільного дільника і найбільшого спільного дільника двох чисел випливає правильність таких тверджень:

Теорема 3. *Якщо ціле число a ділиться на натуральне число b, то множина спільних дільників чисел a і b збігається з множиною дільників числа b. Зокрема, $(a, b) = b$.*

Доведення. Справді, будь-який спільний дільник чисел a і b є дільником числа b. Але оскільки a ділиться на b, то й, навпаки, кожний дільник числа b буде спільним дільником чисел a і b. Отже, множина спільних дільників чисел a і b збігається з множиною дільників числа b. А оскільки найбільшим дільником числа b є саме це число, то $(a, b) = b$.

Теорема 4. *Якщо цілі числа a, b, q, r пов'язані співвідношенням $a = bq + r$, то множина спільних дільників чисел a і b збігається з множиною спільних дільників чисел b і r. Зокрема, $(a, b) = (b, r)$.*

Доведення. Справді, кожний спільний дільник чисел a і b за сьомою властивістю подільності є дільником числа r, а кожний спільний дільник чисел b і r за цією самою властивістю є дільником і числа a. Таким чином, множина спільних дільників чисел a і b збігається з множиною спільних дільників чисел b і r і, отже, $(a, b) = (b, r)$.

Перейдемо тепер до питання про знаходження НСД двох чисел.

Ще Евклід у книзі VII своїх «Начал» виклав спосіб знаходження НСД двох чисел, який відомий тепер як спосіб послідовного ділення, або алгоритм Евкліда. Полягає він ось у чому. Нехай a і b — натуральні числа. Якщо a не ділиться на b, то за теоремою 1 $a = bq_1 + r_1$, $0 < r_1 < b$. Якщо b не ділиться на r_1, то за цією самою теоремою $b = r_1 q_2 + r_2$, $0 < r_2 < r_1$. Якщо r_1 не ділиться на r_2, то $r_1 = r_2 q_3 + r_3$, $0 < r_3 < r_2$ і т. д. Цей процес послідовного ділення не може продовжуватись нескінченно, бо в противному разі множина натуральних чисел $r_1 > r_2 > r_3 > \ldots > r_{n-1} > r_n \ldots$ не матиме найменшого числа, а це суперечить принципу найменшого числа. Отже, існує таке n, що r_{n-1} ділиться на r_n. Процес послідовного ділення закінчиться через $n + 1$ кроків, і ми дістанемо таку систему рівностей:

$$\begin{aligned} a &= bq_1 + r_1, \\ b &= r_1 q_2 + r_2, \\ r_1 &= r_2 q_3 + r_3, \\ &\cdots\cdots\cdots \\ r_{n-2} &= r_{n-1} q_n + r_n, \\ r_{n-1} &= r_n q_{n+1}. \end{aligned} \tag{6}$$

Розглядаючи ці рівності зверху вниз, на підставі теореми 4 приходимо до висновку, що множина спільних дільників числа a і b збігається з множиною спільних дільників чисел b і r_1, з множиною спільних дільників чисел r_1 і r_2, r_2 і r_3 і т. д. Отже, множина спільних дільників чисел a і b збігається з множиною спільних дільників чисел r_{n-1} і r_n, тобто за теоремою 3 вона збігається з множиною дільників числа r_n.

Зокрема, $(a, b) = (b, r_1) = (r_1, r_2) = \ldots = (r_{n-1}, r_n) = r_n$. Отже, ми довели таку теорему:

Теорема 5. НСД *чисел a і b дорівнює останній відмінній від нуля остачі r_n в алгоритмі Евкліда.*

З викладеного вище випливає також правильність такого твердження:

Теорема 6. *Множина спільних дільників чисел a і b збігається з множиною дільників* НСД (a, b) *цих чисел.*

З цієї теореми випливає такий наслідок: НСД *чисел a і b ділиться на будь-який їх спільний дільник.*

Очевидно також, що додатний спільний дільник чисел a і b, який ділиться на будь-який їх спільний дільник, є найбільшим серед спільних дільників цих чисел.

Виходячи з щойно викладеного, поняття найбільшого спільного дільника двох чисел можна означити так: *найбільшим спільним дільником чисел a і b називається такий додатний їх спільний дільник d, який ділиться на будь-який спільний дільник δ цих чисел.*

Приклад. Застосуємо алгоритм Евкліда для знаходження найбільшого спільного дільника чисел 816 і 187. Маємо:

$$\begin{array}{r|l} 816 & 187 \\ \hline -748 & 4 \\ \hline 68 & \end{array} \qquad \begin{array}{r|l} 187 & 68 \\ \hline -136 & 2 \\ \hline 51 & \end{array} \qquad \begin{array}{r|l} 68 & 51 \\ \hline -51 & 1 \\ \hline 17 & \end{array} \qquad \begin{array}{r|l} 51 & 17 \\ \hline -51 & 3 \\ \hline 0 & \end{array}$$

Отже,

$$816 = 187 \cdot 4 + 68, \quad 187 = 68 \cdot 2 + 51, \quad 68 = 51 \cdot 1 + 17, \quad 51 = 17 \cdot 3.$$

Остання відмінна від 0 остача дорівнює 17. Отже, $(816, 187) = 17$.

Теорема 7. *Якщо натуральні числа a і b помножимо на натуральне число m, то їх* НСД *також помножиться на m, тобто*

$$(am, bm) = (a, b)\, m.$$

Доведення. Помноживши обидві частини кожної з рівностей (6) на число m, дістанемо рівності:

$$am = bmq_1 + r_1 m,$$
$$bm = r_1 mq_2 + r_2 m,$$

$$r_1 m = r_2 mq_3 + r_3 m,$$
$$\ldots\ldots\ldots\ldots$$
$$r_{n-2} m = r_{n-1} mq_n + r_n m,$$
$$r_{n-1} m = r_n mq_{n+1},$$

тобто матимемо алгоритм Евкліда для чисел am і bm. Оскільки остання відмінна від 0 остача тут дорівнює $r_n\ m$, то

$$(am, bm) = r_n m = (a, b)\, m.$$

Теорема 8. *Якщо натуральні числа a і b поділимо на який-небудь їх спільний дільник δ, то* НСД *цих чисел також поділиться на δ, тобто*

$$\left(\frac{a}{\delta}, \frac{b}{\delta}\right) = \frac{(a, b)}{\delta}.$$

Доведення. Справді, $\left(\frac{a}{\delta}\,\delta, \frac{b}{\delta}\,\delta\right) = (a, b)$. З другого боку, за теоремою 7 $\left(\frac{a}{b}\,\delta, \frac{b}{\delta}\,\delta\right) = \left(\frac{a}{\delta}, \frac{b}{\delta}\right) \cdot \delta$. Отже, $(a, b) = \left(\frac{a}{\delta}, \frac{b}{\delta}\right)\delta$. Звідси $\left(\frac{a}{\delta}, \frac{b}{\delta}\right) = \frac{(a, b)}{\delta}$.

З теорем 7 і 8 випливають такі очевидні наслідки:

Наслідок 1. *Частки $\frac{a}{d}$ і $\frac{b}{d}$ від ділення чисел a і b на їх найбільший спільний дільник d взаємно прості.*

Наслідок 2. *Якщо частки $\frac{a}{d}$ і $\frac{b}{d}$ від ділення чисел a і b на їх спільний дільник d взаємно прості, то d є найбільший спільний дільник цих чисел.*

5.4. Найбільший спільний дільник кількох чисел. Якщо кожне з цілих чисел $a_1, a_2, \ldots, a_n$ ділиться на ціле число δ, то число δ називають *спільним дільником чисел $a_1, a_2, \ldots, a_n$*. Найбільший із спільних дільників чисел $a_1, a_2, \ldots, a_n$ називають *найбільшим спільним дільником* цих чисел і позначають символом $(a_1, a_2, \ldots, a_n)$.

Якщо найбільший спільний дільник чисел $a_1, a_2, \ldots, a_n$ дорівнює 1, то ці числа називають *взаємно простими*.

Якщо кожне з чисел $a_1, a_2, \ldots, a_n$ взаємно просте з будь-яким іншим з них, то числа $a_1, a_2, \ldots, a_n$ називають *попарно взаємно простими*.

З викладених вище означень безпосередньо випливає, що коли числа $a_1, a_2, \ldots, a_n$ попарно взаємно прості, то вони і взаємно прості, бо якби $(a_1, a_2, \ldots, a_n) = d > 1$, то, наприклад, числа a_1 і a_2 ділилися б на число $d > 1$, тобто не були б взаємно простими.

Обернене твердження неправильне. Це видно, зокрема, з такого прикладу: числа 15, 10, 13 — взаємно прості, але не попарно взаємно прості, бо $(15, 10) = 5$.

У випадку двох чисел поняття «попарно взаємно прості» збігається з поняттям «взаємно прості».

Нехай $a_1, a_2, a_3, \ldots, a_{n-1}, a_n$ — будь-які цілі числа, серед яких принаймні одне відмінне від 0. Введемо такі позначення:

$(a_1, a_2) = d_2$, $(d_2, a_3) = d_3$, $\ldots$, $(d_{n-2}, a_{n-1}) = d_{n-1}$, $(d_{n-1}, a_n) = d_n$.

Теорема 9. *Множина спільних дільників чисел* $a_1, a_2, a_3, \ldots, \ldots, a_{n-1}, a_n$ *збігається з множиною дільників числа* d_n.

Доведення. При $n = 2$ теорема правильна. Припустимо, що вона правильна для $n - 1$, $(n \geqslant 3)$, тобто, що множина спільних дільників чисел $a_1, a_2, \ldots, a_{n-1}$ збігається з множиною усіх дільників числа d_{n-1}. В такому разі множина спільних дільників чисел $a_1, a_2, a_3, \ldots, a_{n-1}, a_n$ збігається з множиною спільних дільників чисел d_{n-1} і a_n. Але оскільки множина спільних дільників чисел d_{n-1} і a_n збігається з множиною дільників числа d_n, то множина спільних дільників чисел $a_1, a_2, \ldots, a_{n-1}, a_n$ збігається з множиною дільників числа d_n.

Таким чином, з припущення, що твердження правильне для $n - 1$ цілих чисел, випливає правильність його і для n цілих чисел, і, отже, за принципом математичної індукції, воно правильне для будь-якого числа цілих чисел. Цим теорему доведено.

Через те що найбільшим серед дільників числа d_n є саме число d_n, то з теореми 8 випливає правильність такого твердження:

Теорема 10. *Число* d_n *є найбільшим спільним дільником чисел* $a_1, a_2, \ldots, a_n$.

Як видно з викладеного вище, задача знаходження найбільшого спільного дільника чисел $a_1, a_2, \ldots, a_n$ зводиться до знаходження чисел $d_2, d_3, \ldots, d_{n-1}, d_n$, тобто до знаходження найбільшого спільного дільника двох чисел.

Оскільки множина спільних дільників чисел $a_1, a_2, \ldots, a_n$ збігається з множиною дільників їх найбільшого спільного дільника d, то *найбільший спільний дільник* d *чисел* $a_1, a_2, \ldots, a_n$ *ділиться на будь-який спільний дільник* δ *цих чисел.*

З другого боку, додатний спільний дільник чисел $a_1, a_2, \ldots, a_n$, який ділиться на будь-який їх спільний дільник, є найбільшим серед спільних дільників цих чисел. Отже, *найбільшим спільним дільником чисел* $a_1, a_2, \ldots, a_n$ *називається додатний спільний дільник цих чисел, який ділиться на будь-який їх спільний дільник.*

Зауважимо, що при множенні всіх чисел $a_1, a_2, \ldots, a_n$ на будь-яке натуральне число m кожне з чисел $d_2, d_3, \ldots, d_{n-1}, d_n$ також помножиться на число m. При діленні всіх чисел $a_1, a_2, \ldots, a_n$ на будь-який їх спільний дільник δ на δ поділяться також і всі числа $d_2, d_3, \ldots, \ldots, d_{n-1}, d_n$. Звідси випливає, що для найбільшого спільного дільника $d = (a_1, a_2, \ldots, a_n)$ правильні твердження, аналогічні теоремам 7 і 8.

5.5. Взаємно прості числа. Доведемо тепер кілька простих, але важливих теорем про взаємно прості числа.

Теорема 11.

$$\forall_{a,b,c \in Z} [(a, b) = 1 \Rightarrow (ac, b) = (c, b)].$$

Доведення. Справді, оскільки числа ac і bc діляться на (ac, b), то за наслідком теореми 6 і їх найбільший спільний дільник (ac, bc) ділиться на (ac, b). Але за теоремою 7 $(ac, bc) = c\,(a, b) = c$. Отже, число c ділиться на (ac, b). Оскільки на (ac, b) ділиться також і число b, то за наслідком теореми 6 на (ac, b) ділиться і (c, b). Навпаки, (ac, b) ділиться на (c, b), бо кожне з чисел ac і b ділиться на (c, b).

Таким чином, (c, b) ділиться на (ac, b), а (ac, b) ділиться на (c, b) і тому $(ac, b) = (c, b)$.

Теорема 12. *Якщо число* a *взаємно просте з кожним з чисел* b *і* c, *то* a *взаємно просте й з добутком* bc.

Доведення. Справді, $(b, a) = 1$, $(c, a) = 1$, тоді за попередньою теоремою $(bc, a) = (c, a) = 1$, тобто bc і a взаємно прості.

Цю теорему можна узагальнити так: *якщо число* a *взаємно просте з кожним з чисел* $b_1, b_2, \ldots, b_n$, *то* a *взаємно просте й з добутком* $b_1, b_2, \ldots, b_n$.

Справді, за теоремою 12 a взаємно просте з b_1 b_2, b_1 b_2 b_3 і т. д., взаємно просте з $b_1, b_2, \ldots, b_n$.

Теорема 13. *Якщо добуток* ab *ділиться на* c, *причому* b *і* c *взаємно прості, то* a *ділиться на* c.

Доведення. Оскільки $(b, c) = 1$, то за теоремою 7 $(ab, ac) = a\,(b, c) = a$. За умовою ab ділиться на c, тому c є спільним дільником ab і ac і, отже, згідно з наслідком теореми 6 є також дільником їх НСД $(ab, ac) = a$, тобто a ділиться на c.

Теорема 14. *Якщо* a *ділиться на кожне з чисел* b *і* c, *причому* b *і* c *взаємно прості, то* a *ділиться і на добуток* bc.

Доведення. Справді, $(b, c) = 1$, тому за теоремою 7 $(ab, ac) = a$.

Оскільки a ділиться на b, то $a = bm$ і, отже, $ac = bc \cdot m$, тобто ac ділиться на bc. Але a ділиться і на c, тому $a = ck$, отже, $ab = bc \cdot k$, тобто ab ділиться на bc. Таким чином, bc є спільний дільник чисел ab і ac, звідки за теоремою 6 найбільший спільний дільник цих чисел $(ab, ac) = a$ ділиться на bc.

5.6. Найменше спільне кратне. Нехай $a_1, a_2, \ldots, a_n$ — будь-які цілі числа. Числа $a_1a_2 \ldots a_n$, $2a_1a_2 \ldots a_n$, $3a_1a_2 \ldots a_n$, $\ldots$, $ma_1a_2 \ldots a_n$ діляться на кожне з чисел $a_1, a_2, \ldots, a_n$. Отже, існує нескінченна множина цілих чисел, які діляться на $a_1, a_2, \ldots, a_n$.

Означення. *Ціле число, яке ділиться на кожне з чисел* $a_1, a_2, \ldots, a_n$, *називають спільним кратним цих чисел.*

Найменше з додатних спільних кратних чисел $a_1, a_2, \ldots, a_n$ називають *найменшим спільним кратним* цих чисел.

Найменше спільне кратне чисел $a_1, a_2, \ldots, a_n$ (скорочено НСК) позначають символом $[a_1, a_2, \ldots, a_n]$.

З'ясуємо насамперед як знайти найменше спільне кратне двох натуральних чисел a і b.

Нехай M — будь-яке спільне кратне натуральних чисел a і b. Оскільки M ділиться на a, то $M = a \cdot k$, де k — деяке ціле число. Але M ділиться також і на b. Тому $\frac{M}{b} = \frac{ak}{b}$ також є деяке ціле число.

Позначимо найбільший спільний дільник (a, b) чисел a і b через d. Тоді $a = a_1d$, $b = b_1d$,

$$\frac{M}{b} = \frac{ak}{b} = \frac{a_1dk}{b_1d} = \frac{a_1k}{b_1}.$$

Оскільки $\frac{a_1 k}{b_1}$ є ціле число і за наслідком 1 з теорем 7 і 8 $(a_1, b_1) = 1$, то за теоремою 13 k ділиться на b_1, тобто $k = b_1 \cdot t = \frac{b}{d} \cdot t$, де t — деяке ціле число.

З викладеного вище випливає, що

$$M = \frac{ab}{d} t. \qquad (7)$$

Таким чином, ми довели, що будь-яке спільне кратне M чисел a і b можна записати у формі (7), тобто його можна дістати з формули (7) при певному значенні t.

Очевидно, що й, навпаки, кожне число M виду (7) є спільним кратним чисел a і b. Отже, формула (7) дає загальний вигляд усіх спільних кратних чисел a і b.

Найменше спільне кратне $[a, b]$ чисел a і b дістанемо, очевидно, при $t = 1$. Отже, ми довели таку теорему:

Теорема 15. *Найменше спільне кратне натуральних чисел a і b дорівнює добутку цих чисел, поділеному на їх найбільший спільний дільник, тобто* $[a, b] = \frac{a \cdot b}{(a, b)}$.

Найменше спільне кратне чисел a і b позначимо через m. Тоді формулу (7) запишемо так: $M = m \cdot t$.

Звідси випливає правильність такої теореми:

Теорема 16. *Будь-яке спільне кратне M чисел a і b ділиться на їх найменше спільне кратне $[a, b]$.*

Оскільки будь-яке кратне найменшого спільного кратного $[a, b]$ є спільним кратним чисел a і b, то очевидно, що множина спільних кратних чисел a і b збігається з множиною кратних найменшого спільного кратного цих чисел.

Розглянемо питання про знаходження найменшого спільного кратного натуральних чисел $a_1, a_2, \ldots, a_n$ при $n > 2$.

Введемо такі позначення:

$$[a_1, a_2] = m_2, \ [m_2, a_3] = m_3, \ \ldots, \ [m_{n-2}, a_{n-1}] = m_{n-1}, \ [m_{n-1}, a_n] = m_n.$$

Теорема 17. *Множина спільних кратних чисел $a_1, a_2, \ldots, a_n$ збігається з множиною кратних числа m_n.*

Доведення. Для $n = 2$ теорема правильна. Припустимо, що теорема правильна для $n - 1$ $(n \geqslant 3)$, і доведемо, що вона правильна і для n. Справді, за припущенням, множина спільних кратних чисел $a_1, a_2, \ldots, a_{n-1}$ збігається з множиною кратних числа m_{n-1}. Отже, множина спільних кратних чисел $a_1, a_2, \ldots, a_{n-1}, a_n$ збігається з множиною спільних кратних чисел m_{n-1} і a_n. Але множина спільних кратних чисел m_{n-1} і a_n збігається з множиною кратних числа m_n. Звідси випливає, що множина спільних кратних чисел $a_1, a_2, \ldots, a_n$ збігається з множиною кратних числа m_n.

Отже, за принципом математичної індукції, теорема правильна для будь-якого натурального n. Оскільки серед кратних числа m_n найменшим є саме число m_n, то безпосередньо з теореми 16 випливає правильність такого твердження.

Теорема 18. *Число m_n є найменшим спільним кратним чисел $a_1, a_2, \ldots, a_n$.*

Таким чином, знаходження найменшого спільного кратного кількох чисел зводиться до знаходження найменшого спільного кратного двох чисел.

Приклад. Знайти [245, 147, 84]. Знаходимо спочатку [247, 147]. Як відомо, $[245, 147] = \frac{245 \cdot 147}{(245, 147)}$. Застосовуючи алгоритм Евкліда, знаходимо, що $(245, 147) = 49$. Отже, $[245, 147] = \frac{245 \cdot 147}{49} = 245 \cdot 3 = 735$. Знайдемо тепер [245, 147, 84]. За теоремою 18, $[245, 147, 84] = [735, 84] = \frac{735 \cdot 84}{(735, 84)} = \frac{735 \cdot 84}{21} = 2940$.

§ 6. ЦІЛІ СИСТЕМНІ ЧИСЛА

6.1. Системи числення. Щоб оперувати з цілими числами й вивчати їх властивості, треба, насамперед, уміти називати й записувати їх. У різні часи в різних народів були різні способи найменування й запису чисел. Усякий спосіб найменування й запису чисел називають *системою числення* або *нумерацією*.

У кожній системі числення числа записують за допомогою певних знаків (символів), які тепер називають *цифрами*. В одних системах кожна цифра завжди позначає одне й те саме число незалежно від її місця (позиції) в записі числа. Такі системи називають *непозиційними*. В інших — значення кожної цифри визначається не лише самою цифрою, а й позицією, яку вона займає у записі числа. Такі системи називають *позиційними*.

Добре відомим прикладом непозиційної системи числення є *римська система*, яка дійшла до нас із стародавнього Риму. В ній для запису чисел використовують сім цифр: цифра I завжди означає одиницю, цифра V — п'ять, цифра X — десять, L — п'ятдесят, C — сто, D — п'ятсот, M — тисячу. За допомогою цих цифр можна записати будь-яке число, застосовуючи принцип додавання і віднімання. Віднімати можна, як правило, не більше одного знака, а додавати не більше трьох однакових знаків. Наприклад, число сім римськими цифрами — це VII, тобто 5 + 2, число дев'ять — це IX, тобто 10 — 1, а наприклад, у записі числа 14 — XIV використано обидва принципи: від п'яти віднято один і результат додано до десяти.

Проте записи навіть не дуже великих чисел у римській нумерації довгі, множення і ділення на письмі виконувати неможливо, тому ці дії доводиться виконувати усно. Навіть для того щоб прочитати число, потрібно спочатку виконати усно дії додавання і віднімання.

Саме тому римська система числення в математичній практиці не застосовується. Римські цифри тепер використовують дуже рідко, в основному для нумерації століть у хронології, розділів у книжках і т. п.

З'ясуємо тепер суть позиційного принципу запису чисел. Нехай g — деяке зафіксоване натуральне число, більше від одиниці. Назвемо це число *основою системи числення.*

Теорема 1. *Кожне натуральне число m можна записати і притому єдиним способом у вигляді*

$$m = a_n g^n + a_{n-1} g^{n-1} + \cdots + a_1 g + a_0, \qquad (1)$$

де a_i ($i = 0, 1, \ldots, n$) — певні цілі невід'ємні числа, менші від g, причому $a_n \neq 0$.

Доведення. Застосувавши метод математичної індукції, доведемо можливість подання будь-якого натурального числа m у вигляді (1). Для 1 запис виду (1), очевидно, можливий; в цьому випадку $n = 0$, $a_n = a_0 = 1$ і, отже, $a_0 < g$. Припустимо тепер, що кожне натуральне число k, яке менше від m, можна записати у вигляді (1) і доведемо, що тоді запис виду (1) можливий і для числа m. Для цього розглянемо такий нескінченний ряд чисел

$$1, g, g^2, g^3, \ldots, g^n, g^{n+1}, \ldots. \qquad (2)$$

Число 1 менше від m, а число g^m, очевидно, більше за m. Отже, серед чисел ряду (2) є менші і більші від m. Серед чисел ряду (2), більших від m, за принципом найменшого числа є найменше. Нехай ним є число g^{n+1}. Тоді $g^n \leqslant m < g^{n+1}$.

За теоремою 1, § 5 маємо

$$m = a_n g^n + r, \qquad (3)$$

де $0 \leqslant r < g^n$.

Коефіцієнт a_n задовольняє нерівність $0 < a_n < g$. Справді, з рівності (3) випливає, що

$$a_n g^n = m - r > g^n - g^n = 0, \quad a_n > 0;$$

$$a_n g^n \leqslant m < g^{n+1}, \quad a_n < g.$$

Якщо $r = 0$, то для числа m можливий запис виду (1), бо тоді $m = 0 + 0 \cdot g + 0 \cdot g^2 + \ldots + 0 \cdot g^{n-1} + a_n g^n$. Якщо ж $r > 0$, то оскільки $r < g^n \leqslant m$, за індуктивним припущенням для r можливий запис виду (1). Отже,

$$r = a_s g^s + a_{s-1} g^{s-1} + \cdots + a_1 g + a_0,$$

причому $0 \leqslant s < n$, бо $r < g^n$. Але тоді

$$m = a_n g^n + a_{n-1} g^{n-1} + \cdots + a_2 g^2 + a_1 g + a_0,$$

і, отже, для числа m можливий запис виду (1). Тому за принципом математичної індукції запис виду (1) можливий для будь-якого натурального числа m.

Доведемо тепер єдиність запису (1). Покажемо спочатку, що при будь-якому відмінному від 1 натуральному g і при цілих невід'ємних b_i ($i = 0, 1, 2, \ldots, s-1$), менших від g, справджується нерівність

$$g^s > b_{s-1} g^{s-1} + b_{s-2} g^{s-2} + \cdots + b_2 g^2 + b_1 g + b_0. \qquad (4)$$

Справді, оскільки кожен з коефіцієнтів b_i ($i = 0, 1, 2, \ldots, s-1$) менший від g, то кожне з чисел $b_0, b_1, \ldots, b_{s-2}, b_{s-1}$ не більше

від $g - 1$. Отже,

$$(g-1) g^{s-1} + (g-1) g^{s-2} + \cdots + (g-1) g + (g-1) \geqslant b_{s-1} g^{s-1} + b_{s-2} g^{s-2} + \cdots + b_1 g + b_0. \qquad (5)$$

Після розкриття дужок і зведення подібних членів у лівій частині нерівності (5) дістанемо нерівність

$$g^s - 1 \geqslant b_{s-1} g^{s-1} + b_{s-2} g^{s-2} + \cdots + b_1 g + b_0. \qquad (6)$$

А з нерівності (6) випливає нерівність (4).

Припустимо тепер, що для числа m існує два записи виду (1):

$$m = a_n g^n + a_{n-1} g^{n-1} + \cdots + a_1 g + a_0, \qquad (7)$$

$$m = b_n g^n + b_{n-1} g^{n-1} + \cdots + b_1 g + b_0. \qquad (8)$$

Для зручності міркувань вважатимемо, що показники вищих степенів g в обох рівностях дорівнюють n. У противному разі ми могли б дописати відсутні додатки з нульовими коефіцієнтами і, таким чином, добитися, щоб вищі показники числа g в обох рівностях стали однаковими.

З рівностей (7) і (8) випливає, що

$$a_n g^n + a_{n-1} g^{n-1} + \cdots + a_1 g + a_0 = b_n g^n + b_{n-1} g^{n-1} + \cdots + b_1 g + b_0. \qquad (9)$$

Доведемо, що $a_n = b_n$, $a_{n-1} = b_{n-1}$, ..., $a_1 = b, a_0 = b_0$. Припустимо, що це не так, тобто припустимо, що $a_n = b_n$, $a_{n-1} = b_{n-1}$, ..., $a_{s+1} = b_{s+1}$, але $a_s \neq b_s$. Нехай $a_s > b_s$. Тоді $a_s - 1 \geqslant b_s$ і тому

$$(a_s - 1) g^s \geqslant b_s g^s. \qquad (10)$$

Крім того, очевидно, здійснюється така нерівність:

$$a_{s-1} g^{s-1} + a_{s-2} g^{s-2} + \cdots + a_1 g + a_0 \geqslant 0. \qquad (11)$$

Додавши почленно нерівності (4), (10), (11), дістанемо нерівність $a_s g^s + a_{s-1} g^{s-1} + \ldots + a_1 g + a_0 > b_s g^s + b_{s-1} g^{s-1} + \ldots + b_1 g + b_0$.

До обох частин цієї нерівності додамо тепер порівну, а саме:

$$a_n g^n + a_{n-1} g^{n-1} + \cdots + a_{s+1} g^{s+1}$$

— до лівої частини і

$$b_n g^n + b_{n-1} g^{n-1} + \cdots + b_{s+1} g^{s+1}$$

— до правої частини. Дістанемо нерівність

$$a_n g^n + a_{n-1} g^{n-1} + \cdots + a_1 g + a > b_n g^n + b_{n-1} g^{n-1} + \cdots + b,$$

яка суперечить рівності (9).

Таким чином, наше припущення приводить до суперечності і тому воно неправильне. Отже,

$$a_n = b_n, \ a_{n-1} = b_{n-1}, \ \ldots, \ a_1 = b_1, \ a_0 = b_0.$$

Цим теорему доведено.

Означення. *Вираз (1) називають записом числа m у системі числення з основою g. Символи, що позначають числа $a_n, a_{n-1}, \ldots, a_1, a_0$, називають цифрами числа m в системі числення з основою g.*

Права частина виразу (1) є сума степенів числа g з цілими невід'ємними коефіцієнтами, меншими від g.

Отже, *запис кожного натурального числа m в системі числення з основою g — це зображення числа у вигляді суми степенів основи g з цілими невід'ємними коефіцієнтами, меншими ніж основа g.*

Вираз

$$m = a_n g^n + a_{n-1} g^{n-1} + \ \cdots \ + a_1 g + a_0$$

скорочено записують так:

$$m = (a_n a_{n-1} \ \ldots \ a_2 a_1 a_0)_g . \qquad (12)$$

У цьому записі значення кожної цифри визначається як самою цифрою, так і місцем (позицією), яку вона займає у записі. Так, якщо у записі (12), наприклад, $a_2 = a_1 = a_0 = 2$ ($g \geqslant 2$), то a_0 означає число 2, a_1 — число $2g$, а a_2 — число $2g^2$. Отже, система числення, в якій числа записуються щойно викладеним способом, є *позиційною системою*. Зауважимо, що для запису натуральних чисел в позиційній системі числення з основою g потрібно рівно g цифр: (g — 1) цифр потрібно для позначення натуральних чисел, менших ніж g, оскільки всі ці числа в даній системі є однозначними і, отже, повинні позначитися різними символами, і, крім того, потрібна ще цифра нуль. У системі числення з основою g число один вважають *одиницею* першого розряду, g одиниць першого розряду об'єднуються в одну одиницю другого розряду, об'єднання g одиниць другого розряду становить одиницю третього розряду і т. д.

Таким чином, запис (12) означає, що число m складається з a_0 одиниць першого розряду, a_1 одиниць другого розряду, a_2 одиниць третього розряду і т. д.

До цього часу мова йшла про запис натуральних чисел у системі числення з основою g. Аналогічно записують і додатні дробові числа. Запис додатного дробового числа у системі числення з основою g являє собою зображення цього числа у вигляді суми степенів основи g з цілими невід'ємними коефіцієнтами, меншими від основи, але не лише додатних і нульового, а й від'ємних степенів. Для запису додатного дробового числа, крім цифр, використовують ще один знак — кому. Дробове число

$$1 \cdot 8^3 + 2 \cdot 8^2 + 5 \cdot 8^1 + 4 \cdot 8^0 + 7 \cdot 8^{-1} + 6 \cdot 8^{-2} + 3 \cdot 8^{-3} + 2 \cdot 8^{-4} ,$$

наприклад, записують так: $1254,\ 7632_8$.

Для кожного додатного дробового числа так само, як і для кожного натурального числа, в системі числення з будь-якою основою g існує тільки один запис.

Введення знака мінус дає змогу записувати у системі числення з основою g також і від'ємні числа.

У десятковій системі числення, як відомо, одні раціональні числа записуються у вигляді скінченного дробу, інші — у вигляді нескінченного. Це ж саме спостерігається і в інших позиційних системах числення. При цьому те саме число може в одній системі числення записуватися скінченним дробом, а в іншій — нескінченним.

Наприклад, число $\frac{1}{5}$ в десятковій системі числення записують скінченним дробом, а в дванадцятковій — нескінченним $\frac{1}{5} = 0{,}2_{10}$, $\frac{1}{5} = 0{,}249724972497\ldots_{12}$; число $\frac{1}{5}$, навпаки, в дванадцятковій системі числення записують скінченним дробом, а в десятковій — нескінченним: $\frac{1}{6} = 0{,}2_{12}$, $\frac{1}{6} = 0{,}1666\ldots_{10}$. В обох цих системах число $\frac{1}{4}$ записують скінченним дробом, а число $\frac{1}{7}$ — нескінченним: $\frac{1}{4} = 0{,}25_{10} = 0{,}3_{12}$, $\frac{1}{7} = 0{,}142857142857\ldots_{10} = 0{,}186035186035\ldots_{12}$.

Загальновживаною тепер є позиційна система числення, основа якої $g = 10$. Її називають *десятковою позиційною системою*.

Десяткова позиційна система була винайдена в Індії; там її запозичили араби потім перенесли в Європу. Звичайно, можливі системи числення й з будь-якою іншою основою. Так, за 2—3 тисячі років до нашої ери у Вавілоні успішно використовувалася *шістдесяткова система числення*, в основу якої було покладено число ш і с т д е с я т. Сліди цієї системи числення збереглися до наших днів. І в наш час годину ділять на шістдесят хвилин, хвилину — на шістдесят секунд; коло ділять на триста шістдесят градусів, градус — на шістдесят мінут, а мінуту — на шістдесят секунд.

Досить поширеною була колись *дванадцяткова система числення*, основою якої є число д в а н а д ц я т ь. Про це свідчить те, що рік ділиться на дванадцять місяців, доба — на (12×2) годин. Ще й тепер деякі предмети (тарілки, ножі, ложки, стільці тощо) лічать не десятками, а *дюжинами*.

Найстародавнішою є *двійкова система числення* з основою два, якою, як гадають, користувались стародавні єгиптяни. Вважають, що її було винайдено у IV тисячолітті до нашої ери.

Можливо, що двійкову систему числення як найпростішу використовували на певному етапі всі народи. Тепер двійкову позиційну систему числення широко застосовують в електронних обчислювальних машинах.

В електронних обчислювальних машинах використовують також *вісімкову позиційну систему*.

6.2. Арифметичні операції над системними числами. Число, записане в певній системі числення, називають *системним числом*. Щоб розрізняти, в якій системі числення записане дане число, ми будемо вказувати основу системи числення, записуючи її (в десятковій системі числення!) справа внизу від числа у вигляді індекса.

Наприклад, 372_{10} — це запис числа у звичайній десятковій системі числення, а 372_8 — у вісімковій системі.

При виконанні арифметичних операцій над числами, записаними в десятковій системі числення, ми користуємося правилами додавання, віднімання і множення чисел «стовпцем» і ділення — «кутом». За цими ж правилами виконують операції й над числами, записаними в будь-якій іншій позиційній системі числення. Грунтуються ці правила на п'яти основних законах додавання і множення цілих чисел: асоціативності й комутативності додавання, асоціативності й комутативності множення, дистрибутивності множення відносно додавання.

Щоб з'ясувати це, розглянемо операції додавання і множення чисел у системі числення з основою g.

Д о д а в а н н я. Нехай $a = (a_k a_{k-1} \ \ldots \ a_2 a_1 a_0)_g$ і $b = (b_s b_{s-1} \ \ldots \ b_2 b_1 b_0)_g$. Не втрачаючи загальності міркувань,

вважатимемо, що $k \geqslant s$, знайдемо суму $a + b$:

$$a + b = (a_k a_{k+1} \ldots a_2 a_1 a_0)_g + (b_s b_{s-1} \ldots b_2 b_1 b_0)_g =$$

$$= (a_k g^k + a_{k-1} g^{k-1} + \cdots + a_{s+1} g^{s+1} + a_s g^s + a_{s-1} g^{s-1} +$$

$$+ \cdots + a_1 g + a_0) + (b_s g^s + b_{s-1} g^{s-1} + \cdots + b_1 g + b_0) =$$

$$= a_k g^k + a_{k-1} g^{k-1} + \cdots + a_{s+1} g^{s+1} + (a_s g^s + b_s g^s) +$$

$$+ (a_{s-1} g^{s-1} + b_{s-1} g^{s-1}) + \cdots + (a_1 g + b_1 g) + (a_0 + b_0) =$$

$=$ (асоціативність і комутативність додавання) $= a_k g^k + a_{k-1} g^{k-1} + \ldots$ $\ldots + a_{s+1} g^{s+1} + (a_s + b_s)\ g^s + (a_{s-1} + b_{s-1})\ g^{s-1} + \ldots +$ $+ (a_1 + b_1)\ g + (a_0 + b_0)$ (дистрибутивність множення відносно додавання).

Отже,

$$a + b = a_k g^k + a_{k-1} g^{k-1} + \cdots + a_{s+1} g^{s+1} + (a_s + b_s)\, g^s +$$

$$+ (a_{s-1} + b_{s-1})\, g^{s-1} + (a_1 + b_1)\, g + (a_0 + b_0). \qquad (13)$$

У цьому записі деякі з чисел $a_0 + b_0$, $a_1 + b_1$, ..., $a_s + b_s$ можуть виявитися більшими або дорівнювати основі числення g. Якщо $a_m + b_m \geqslant g$ $(0 \leqslant m \leqslant s)$, то $a_m + b_m = g + r_m$, де $0 \leqslant r_m < g$. Тому, за асоціативним законом додавання та дистрибутивним законом множення і додавання, $(a_{m+1} + b_{m+1})\, g^{m+1} + (a_m + b_m)\, g^m = (a_{m+1} + b_{m+1} + 1)\, g^{m+1} + r_m g^m$. Замінивши у записі (13) кожен вираз $(a_m + b_m) \geqslant g$, починаючи $a_0 + b_0$ і закінчуючи $a_s + b_s$, рівнозначним йому виразом $g + r_m$ і перенісши там, де це потрібно, одиницю в наступний розряд, дістанемо запис суми $a + b$ в системі числення з основою g:

$$a + b = c_k g^k + c_{k-1} g^{k-1} + \cdots + c_{s+1} g^{s+1} + c_s g^s + c_{s-1} g^{s-1} +$$

$$+ \cdots + c_1 g + c_0,$$

або скорочено

$$a + b = (c_k c_{k-1} + \cdots + c_1 c_0)_g.$$

Таким чином, щоб додати два цілі додатні числа, записані в системі числення з основою g, потрібно додати їхні цифри першого розряду, потім другого розряду, третього розряду і т. д. При цьому кожен раз, коли при додаванні цифр даного розряду дістанемо суму більшу або рівну основі числення g, потрібно зробити перенесення одиниці в наступний розряд. Для зручності при виконанні додавання чисел доданки доцільно підписувати один під одним (у стовпчик) так, щоб цифри, які відповідають однаковим розрядам, були одна під одною.

Як бачимо, додавання багатоцифрових чисел зводиться до додавання чисел одноцифрових. Основою додавання системних чисел є, таким чином, таблиця додавання одноцифрових чисел, що визначає суму двох чисел, менших, ніж основа числення g. У вісімковій системі числення,

наприклад, ця таблиця має такий вигляд:

	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7	10
2	2	3	4	5	6	7	10	11
3	3	4	5	6	7	10	11	12
4	4	5	6	7	10	11	12	13
5	5	6	7	10	11	12	13	14
6	6	7	10	11	12	13	14	15
7	7	10	11	12	13	14	15	16

(14)

У кожній клітині цієї таблиці записано у вісімковій системі числення суму чисел, що є номерами рядка і стовпчика, на перетині яких стоїть дана клітина. Користуючись цією таблицею, знайдемо за встановленим вище правилом додавання системних чисел суму $23451_8 + 15\,254_8$

$$+\begin{array}{r} 23\,451_8 \\ 15\,254_8 \\ \hline 40\,725_8 \end{array}$$

М н о ж е н н я. Виходячи з означення операції множення, складаємо таблицю множення одноцифрових чисел у системі числення з основою g. У вісімковій системі числення ця таблиця така:

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	10	12	14	16
3	0	3	6	11	14	17	22	25
4	0	4	10	14	20	24	30	34
5	0	5	12	17	24	31	36	43
6	0	6	14	22	30	36	44	52
7	0	7	16	25	34	43	52	61

У кожній клітині цієї таблиці записано у вісімковій системі числення добуток чисел, що є номерами рядка і стовпця, на перетині яких стоїть дана клітина.

Від множення одноцифрових чисел переходимо до множення чисел, позначених однією значущою цифрою і нулями, тобто чисел виду $a_m \underbrace{000\ \ldots\ 0}_{m \text{ нулів}}$, або $a_m g^m$, на одноцифрові числа:

$$a_m \underbrace{0\,0\,0\ \ldots\ 0}_{m \text{ нулів}} \cdot b_0 = a_m \cdot g^m b_0 = (a_m \cdot b_0) \cdot g^m$$

(асоціативний і комутативний закони множення) $= (a_m \cdot b_0)_g \underbrace{000\ \ldots\ 0}_{m \text{ нулів}}$.

Отже, *щоб помножити число виду* $a_m \underbrace{000\ \ldots\ 0}_{m \text{ нулів}}$ *на одноцифрове число* b_0, *треба* a_m *помножити на* b_0 *і до знайденого результату дописати справа стільки нулів, скільки їх є в першому множнику.* Зауважимо, що за цим самим правилом, в силу комутативності дії множення, виконується і множення одноцифрового числа b_0 на число виду $a_m \underbrace{000\ \ldots\ 0}_{m \text{ нулів}}$.

Числа, позначені однією значущою цифрою і кількома нулями, помножають так:

$$a_m \underbrace{000\ \ldots\ 0}_{m \text{ нулів}} \cdot b_n \underbrace{000\ \ldots\ 0}_{n \text{ нулів}} = a_m \cdot g^m \cdot b_n \cdot g^n = (a_m \cdot b_n)\, g^{m+n}$$

(асоціативний і комутативний закони множення) $= (a_m \cdot b_n)_g \underbrace{000\ \ldots\ 0}_{m+n \text{ нулів}}$,

тобто за правилом: *щоб помножити число виду* $a_m \underbrace{000\ \ldots\ 0}_{m \text{ нулів}}$ *на число виду* $b_n \underbrace{000\ \ldots\ 0}_{n \text{ нулів}}$, *треба* a_m помножити на b_n *і до знайденого результату дописати справа стільки нулів, скільки їх є в обох множниках.*

Множення багатоцифрового числа на одноцифрове зводиться, в силу дистрибутивного закону множення і додавання, до множення чисел виду $a_m \underbrace{000\ \ldots\ 0}_{m \text{ нулів}}$ і одноцифрового числа a_0 на одноцифрове число b_0, тобто до таблиці множення одноцифрових чисел. Справді,

$$a_k a_{k-1} \ldots a_1 a_0 \cdot b_0 = (a_k \cdot g^k + a_{k-1} \cdot g^{k-1} + \ldots + a_1 \cdot g + a_0) \cdot b_0 =$$
$$= (a_k b_0) g^k + (a_{k-1} b_0) g^{k-1} + \ldots + (a_1 b_0) g + a_0 b_0.$$

Щоб помножити багатоцифрове число $a_k a_{k-1} \ldots a_1 a_0$ *на число виду* $b_s \underbrace{000\ \ldots\ 0}_{s \text{ нулів}}$, *треба число* $a_k a_{k-1} \ldots a_1 a_0$ *помножити на число* b_s *і до знайденого результату дописати справа стільки нулів, скільки їх є в другому множнику.* Справді,

$$(a_k a_{k-1} \ldots a_1 a_0 \cdot b_s) \underbrace{000\ \ldots\ 0}_{s \text{ нулів}} = a_k a_{k-1} \ldots a_1 a_0 \cdot (b_s \cdot g^s) =$$
$$= (a_k a_{k-1} \ldots a_1 a_0 \cdot b_s)_g \underbrace{000\ \ldots\ 0}_{s \text{ нулів}}.$$

Встановивши ці правила, можна обгрунтувати правило множення багатоцифрового числа на багатоцифрове. Справді,

$$a_k a_{k-1} \ldots a_1 a_0 \cdot b_s b_{s-1} \ldots b_1 b_0 = a_k a_{k-1} \ldots a_1 a_0 \cdot (b_s \cdot g^s + b_{s-1} g^{s-1} +$$
$$+ \ldots + b_1 g + b_0) = a_k a_{k-1} \ldots a_1 a_0 \cdot (b_0 + b_1 g + \ldots + b_{s-1} g^{s-1} + b^s g^s) =$$
$$= (\text{комутативний закон додавання}) = (a_k a_{k-1} \ldots a_1 a_0 \cdot b_0)_g +$$
$$+ (a_k a_{k-1} \ldots a_1 a_0 \cdot b_1)_g 0 + \cdots + (a_k a_{k-1} \ldots a_1 a_0 \cdot b_{s-1})_g \underbrace{000\ \ldots\ 0}_{s-1 \text{ нулів}} +$$
$$+ (a_k a_{k-1} \ldots a_0 \cdot b_s)_g \underbrace{000\ \ldots\ 0}_{s \text{ нулів}}.$$

У практиці обчислень доданки останньої суми записують не в рядок, а в стовпчик, один під одним, так, щоб цифри того самого розряду стояли в одному стовпчику. При цьому нулі, що стоять у кінці доданків, починаючи з другого опускають, оскільки при додаванні цифр даного розряду нулі на суму цифр не впливають.

Проілюструємо викладене на конкретному прикладі: $532_8 \cdot 425_8 = 532_8\ (4 \cdot 8^2 + 2 \cdot 8 + 5) = 532_8\ (5 + 2 \cdot 8 + 4 \cdot 8^2) = 532_8 \cdot 5 + 532_8 \cdot 20_8 + 532_8 \cdot 400_8 = 3302_8 + 12\,640_8 + 255\,000_8 = 273\,142_8$, або скорочено:

так:
$$\begin{array}{r} 532_8 \\ \times\ 425_8 \\ \hline 3302_8 \\ 1264_8 \\ 2550_8 \\ \hline 273\,142_8 \end{array}$$

Правило множення «стовпцем» формулюється порівняно складно. Тому ми і не будемо формулювати його, вважаючи, що для чисел, записаних у десятковій системі числення, воно читачеві добре відоме. Зауважимо тільки, що множення багатоцифрових чисел «стовпцем» у будь-якій позиційній системі числення зводиться до кількаразового множення одноцифрових чисел і наступного додавання.

6.3. Переведення цілих чисел з однієї позиційної системи числення в іншу. У процесі розв'язування задач доводиться переводити цілі числа з однієї позиційної системи числення в іншу. Як же перевести число a, записане в системі числення з основою p, в систему числення з основою g? Як відомо, записати число a в системі числення з основою g — це означає зобразити його у вигляді суми

$$a = a_k g^k + a_{k-1} g^{k-1} + \ldots + a_1 g + a_0.$$

Отже, щоб записати число a в системі числення з основою g, треба знайти коефіцієнти $a_0, a_1, a_2, \ldots, a_k$. Ці коефіцієнти знаходимо так. Поділимо в системі числення з основою p число a на g, дістанемо $a = b_0 g + a_0$. Далі поділимо b_0 на g, дістанемо $b_0 = b_1 g + a_1$. Звідси

$$a = b_0 g + a_0 = (b_1 g + a_1)\, g + a_0 = b_1 g^2 + a_1 g + a_0.$$

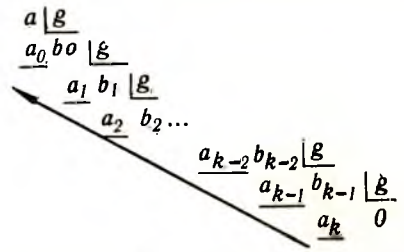
Потім поділимо b_1 на g і т. д. Цей процес продовжуватимемо доти, поки не дістанемо частку, яка дорівнює нулю. Внаслідок цього матимемо:

$$a = a_k g^k + a_{k-1} g^{k-1} + \ldots + a_1 g + a_0.$$

Оскільки за теоремою 1 число a можна подати в такому вигляді єдиним способом, то a_0, a_1, a_2, ..., a_k є цифри числа a в системі числення з основою g.

Таким чином, *цифрами a_0, a_1, ..., a_k числа a в системі числення з основою g є остачі, що утворюються при послідовному діленні a на g.*

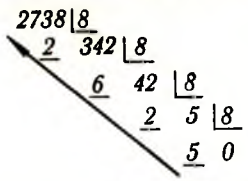
Хід послідовного ділення числа a на g скорочено записують так:



Стрілка показує напрям від вищих до нижчих розрядів числа, записаного в системі числення з основою g, цифри числа a у цій системі підкреслено.

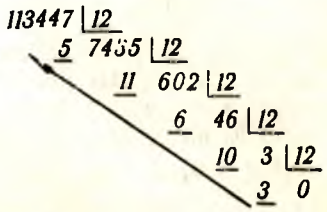
Послідовне ділення числа a на g провадиться в системі числення з основою p. Якщо $g < p$, то дільник g, а отже, і всі остачі a_0, a_1, ..., a_k є одноцифрові числа і ми відразу дістаємо потрібні нам цифри числа a. Якщо ж $g > p$, то в системі числення з основою p дільник g і, можливо, деякі остачі міститимуть більш як одну цифру. У системі числення з основою g ці остачі слід записати новими цифрами, яких у системі числення з основою p немає.

Приклади. 1. Запишемо число 2738_{10} у вісімковій системі числення. Для цього поділимо послідовно у десятковій системі числення число 2738_{10} на 8_{10}:



Звідси випливає, що число 2738_{10} у вісімковій системі числення записується так: 5262_8.

2. Запишемо число 113447_8 у десятковій системі числення. Нова основа числення у вісімковій системі числення дорівнює 12, тобто $10_{10} = 12_8$. Поділимо послідовно у вісімковій системі числення число 113447_8 на 12_8:



Результати послідовного ділення записані у вісімковій системі числення. Але ми знаємо, що $10_8 = 8_{10}$, $11_8 = 9_{10}$. Отже, $113447_8 = 38695_{10}$.

7.1. Прості числа. Повернемося знову до вивчення властивостей цілих додатних чисел.

Число 1 має тільки один дільник, а саме 1, а кожне натуральне число a, відмінне від 1, має принаймні два дільники: 1 і a. (Тут і далі мова йде тільки про додатні дільники).

Означення. *Відмінне від 1 натуральне число a називають простим, якщо воно не має дільників, відмінних від 1 і a. Його називають складеним, якщо воно має дільники, відмінні від 1 і a.*

Простими є, наприклад, числа 2, 7, 13; числа 4, 9, 15 — складені. Число 1 не належить ні до простих, ні до складених чисел.

Доведемо кілька важливих теорем про прості числа.

Теорема 1. *Всяке натуральне число a або ділиться на дане просте число p, або взаємно просте з p.*

Доведення. Справді, найбільший спільний дільник (a, p) як дільник числа p може дорівнювати або p, або 1. У першому випадку a ділиться на p, у другому a і p — взаємно прості числа.

Теорема 2. *Якщо добуток кількох натуральних чисел ділиться на просте число p, то принаймні один із співмножників ділиться на p.*

Доведення. Справді, внаслідок попередньої теореми, кожен із співмножників або взаємно простий з p, або ділиться на p. Але якби всі множники були взаємно прості з p, то за теоремою 12, § 5, і їх добуток був би взаємно простий з p. Тому хоч один з множників ділиться на p.

Теорема 3. *Найменший відмінний від 1 дільник більшого від 1 натурального числа a є число просте.*

Доведення. Нехай q — найменший дільник натурального числа $a > 1$. Якби q було числом складеним, то воно мало б дільник q_1, такий, що $1 < q_1 < q$. Але тоді a ділилося б на q_1 і q не було б найменшим дільником числа a.

Теорема 4. *Найменший відмінний від одиниці дільник складеного числа a не більший за $\sqrt{a}$.*

Доведення. Нехай найменшим відмінним від 1 дільником числа a є q. Тоді $a = q \cdot a_1$, де a_1 — деяке натуральне число, причому $a_1 \geqslant q$, бо в противному разі q не було б найменшим додатним дільником числа a. Отже, $a \cdot a_1 \geqslant q a_1 \cdot q$ і тому $a \geqslant q^2$, $q \leqslant \sqrt{a}$.

7.2. Нескінченість множини простих чисел, Решето Ератосфена.

Теорема 5. (Евкліда). *Множина простих чисел нескінченна.*

Доведення. Припустимо, що множина простих чисел скінченна. Нехай вона складається з простих чисел p_1, p_2, ..., p_n. Отже, ми припускаємо, що простих чисел, відмінних від p_1, p_2, ..., p_n, немає. Розглянемо тепер ціле число $p = p_1 p_2 \ldots p_n + 1$. Очевидно, що це число відмінне від кожного з чисел p_1, p_2, ..., p_n. Оскільки число 1 не має дільників, відмінних від самого себе, то жодне з простих чисел p_1, p_2, ..., p_n не може бути дільником числа p. Ціле число $p > 1$. Тому воно або само просте, або за теоремою 3 ділиться на просте число q, відмінне від кожного з чисел p_1, p_2, ..., p_n. Звідси випливає, що існує принаймні одне просте число, відмінне від чисел p_1, p_2, ..., p_n, а

це суперечить нашому припущенню. Отже, наше припущення неправильне. Цим теорему доведено.

Природно постає запитання: як у ряду натуральних чисел виділити всі прості числа?

Таблицю всіх простих чисел, що не перевищують даного натурального числа N, можна скласти так. Випишемо підряд усі натуральні числа від 2 до N:

$$2,\ 3,\ 4,\ 5,\ \ldots,\ N. \qquad (1)$$

Потім закреслимо в ряду (1) всі числа, кратні 2, крім самого числа 2. Першим числом у ряду (1), яке залишилося після цього, є число 3. Число 3 не ділиться на 2, бо в противному разі ми закреслили б його: отже, число 3 ділиться лише на 1 і на самого себе, тому воно просте. Закреслимо тепер у ряду (1) всі числа, кратні 3, крім самого числа 3. Першим числом, яке залишилося після цього в ряду (1), є число 5; воно не ділиться ні на 2, ні на 3, бо в противному разі воно виявилося б закресленим; отже, 5 ділиться тільки на 1 і на самого себе, тому воно просте число. Потім у ряду (1) закреслимо всі числа, кратні 5, крім самого числа 5 і т. д. Закресливши в ряду (1) всі числа, кратні простим числам, не більшим ніж $\sqrt{N}$, дістанемо за теоремою 4 таблицю всіх простих чисел, які не перевищують числа N.

Уперше для складання таблиць простих чисел описаний щойно метод застосував грецький математик Ератосфен. Ератосфен писав числа на папірусі, натягнутому на рамку; числа він не закреслював, а проколював. Внаслідок цього він діставав дещо схоже на решето: складені числа «просіювалися» крізь це решето, а прості числа залишалися. Тому цей метод називають *решетом Ератосфена*.

Метод Ератосфена поступово удосконалювався, завдяки чому складання таблиць простих чисел спрощувалося. Це, в свою чергу, дало можливість скласти таблиці простих чисел, що містять порівняно велику кількість чисел. Тепер складено таблиці простих чисел приблизно до 10 мільйонів.

7.3. Основна теорема арифметики. Доведемо тепер теорему, яка відіграє фундаментальну роль як у теорії подільності, так і в усій теорії чисел. Її називають основною теоремою арифметики.

Теорема 6. (Основна теорема арифметики). *Кожне відмінне від 1 натуральне число n можна записати у вигляді добутку простих чисел і притому єдиним способом, якщо не брати до уваги порядку розміщення співмножників.*

Доведення. Доведемо спочатку можливість запису натурального числа $q \neq 1$ у вигляді добутку простих чисел. Для натурального числа 2 це можливо, бо число 2 просте, тобто його можна вважати добутком простих чисел з числом співмножників, що дорівнює одиниці. Припустимо, що це можливо для всякого натурального числа k, такого, при якому $2 \leqslant k < n$, і доведемо, що в такому разі і натуральне число n можна записати у вигляді добутку простих чисел. Справді, якщо натуральне число n просте, то воно є добутком простих чисел з числом співмножників, що дорівнює одиниці. Якщо ж число n складене, то воно має дільник k_1, відмінний від 1 і від n. Отже, $n = k_1 \cdot k_2$, де k_2 — натуральне число, відмінне від 1. Тоді $2 \leqslant k_1 < n$, $2 \leqslant k_2 < n$.

Через те що за припущенням кожне з чисел k_1 і k_2 записується у

вигляді добутку простих чисел, то й число $n = k_1 \cdot k_2$ записується у вигляді добутку простих чисел. Отже, внаслідок принципу математичної індукції будь-яке натуральне число n можна записати у вигляді добутку простих чисел.

Доведемо тепер єдиність запису числа n у вигляді добутку простих чисел. Нехай число n двома способами записано у вигляді добутку простих чисел, тобто $n = p_1 p_2 \ldots p_r$ і $n = q_1 q_2 \ldots q_s$, де $r \geqslant 2$, $s \geqslant 2$ і всі числа p_l і q_k прості. Доведемо, що ці записи можуть відрізнятися лише порядком співмножників, тобто що $r = s$ і при належному виборі нумерації співмножників $p_i = q_i$, $i = 1, 2, \ldots, s$. Доводитимемо це індукцією по n. Для числа 2 правильність цього твердження очевидна. Число 2 є добуток простих чисел з числом співмножників, що дорівнює одиниці. Нехай твердження правильне для всякого числа k, такого, що $2 \leqslant k < n$. Доведемо, що в такому разі твердження правильне і для числа n. Справді, оскільки $n = p_1 p_2 \ldots p_r$ і $n = q_1 q_2 \ldots q_s$, то

$$p_1 p_2 \ldots p_r = q_1 q_2 \ldots q_s. \qquad (2)$$

Ліва частина цієї рівності ділиться на просте число p_1. Отже, і права частина її ділиться на просте число p_1. Звідси за теоремою 2 принаймні один із співмножників $q_1, q_2, \ldots, q_s$ ділиться на просте число p_1. Змінивши, якщо потрібно, нумерацію множників $q_1, q_2, \ldots, q_s$, ми вважатимемо, що на p_1 ділиться співмножник q_1. Оскільки q_1 є просте число і ділиться на відмінне від 1 просте число p_1, то $q_1 = p_1$.

Поділивши обидві частини рівності (2) на число $p_1 = q_1$, дістанемо рівність $p_2 p_3 \ldots p_r = q_2 q_3 \ldots q_s$.

Число $n_1 = p_2 p_3 \ldots p_r = q_2 q_3 \ldots q_s$ задовольняє умову $2 \leqslant n_1 < n$. Тому за припущенням для нього твердження правильне, тобто $r - 1 = s - 1$, і при відповідній нумерації $p_2 = q_2$, $p_3 = q_3$, ..., $p_s = q_s$. Звідси випливає, що $r = s$ і при відповідній нумерації $p_i = q_i$, $i = 1, 2, \ldots, s$.

Отже, внаслідок принципу математичної індукції, всяке відмінне від 1 натуральне число n єдиним способом записується у вигляді добутку простих чисел. Цим теорему доведено.

Зауважимо, що запис $n = p_1 p_2 \ldots p_s$ натурального числа n у вигляді добутку простих чисел $p_1, p_2, \ldots, p_s$ називають також *розкладом числа n у добуток простих множників*, або *розкладом на прості множники*.

Основна теорема арифметики показує, що всі натуральні числа дістають з простих чисел за допомогою операції множення: кожне натуральне число (складене) є деякий добуток простих чисел, причому різні добутки дають різні числа. Тепер зрозуміло, чому одиницю не слід вважати простим числом: віднісши 1 до простих чисел, ми порушили б єдиність розкладу числа в добуток простих чисел, оскільки до будь-якого добутку можна приєднати множником 1.

7.4. Канонічний розклад складеного числа. У розкладі

$$n = p_1 p_2 \ldots p_s \qquad (3)$$

натурального числа n на прості множники $p_1, p_2, \ldots, p_s$ деякі з цих множників можуть повторюватись. Якщо простий множник p_i повто-

рюється в розкладі (3) k раз, то його називають *k-кратним множником числа n*, або кажуть, що множник p_i має кратність k.

Позначимо символами $p_1, p_2, \ldots, p_m$ $(m \leqslant s)$ різні множники в розкладі (3). Нехай множник p_i $(i = 1, 2, \ldots, m)$ має кратність k_i. Тоді розклад числа n у добуток простих множників можна записати так: $n = p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$.

Цей запис називають *канонічним розкладом числа n на прості множники, або канонічним зображенням числа n.*

Оскільки число n єдиним способом записується у вигляді добутку простих чисел, то й канонічне зображення числа n існує тільки одне.

Приклад 1. Знайти канонічний розклад числа 12600. Маємо: $12\,600 = 2 \cdot 2 \cdot 2 \cdot 5 \cdot 5 \cdot 3 \cdot 3 \cdot 7 = 2^3 \cdot 3^2 \cdot 5^2 \cdot 7$.

З єдності канонічного зображення числа випливає правильність такого твердження.

Теорема 7. *Якщо $n = p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$ — канонічний розклад числа n, то всі дільники цього числа збігаються з числами вигляду*

$$d = p_1^{s_1} p_2^{s_2} \ldots p_m^{s_m}, \qquad (4)$$

де $0 \leqslant s_i \leqslant k_i$, $i = 1, 2, \ldots, m$.

Доведення. Справді, очевидно, що всяке число d виду (4) є дільником числа n. Навпаки, якщо d є дільником числа n, то $n = dq$, де q — деяке натуральне число. Через те що для n існує тільки один канонічний розклад, то з рівності $n = dq$ випливає, що в канонічний розклад числа d можуть входити тільки прості числа $p_1, p_2, \ldots, p_m$, причому їх степені відповідно не вищі від $k_1, k_2, \ldots, k_m$. Тому канонічне зображення d має вигляд (4).

Візьмемо тепер довільні два натуральні числа a і b. Припустимо, що вони мають такі канонічні розклади:

$$a = r_1^{l_1} r_2^{l_2} \ldots r_m^{l_m}, \quad b = q_1^{s_1} q_2^{s_2} \ldots q_t^{s_t}.$$

Позначимо символами $p_1 p_2 \ldots p_s$ всі різні множники, кожен з яких входить до розкладу принаймні одного з чисел a і b. Якщо простий множник p_i не зустрічається в розкладі якого-небудь з чисел a і b, то вважатимемо, що він входить до цього розкладу в нульовому степені. За цієї умови канонічні розклади чисел a і b можна записати так:

$$a = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s}, \quad b = p_1^{m_1} p_2^{m_2} \ldots p_s^{m_s},$$

де кожен з показників k_i і m_i, $i = 1, 2, \ldots, s$, є ціле невід'ємне число. Для чисел a і b правильні такі твердження.

Теорема 8. *Найбільшим спільним дільником чисел a і b є число*

$$(a, b) = p_1^{l_1} p_2^{l_2} \ldots p_s^{l_s}.$$

де $l_i = \min(k_i, m_i)$, $i = 1, 2, \ldots, s$.

Доведення. Справді, за теоремою 7 всі спільні дільники чисел a і b збігаються з числами вигляду $d = p_1^{n_1} p_2^{n_2} \ldots p_s^{n_s}$, де $0 \leqslant n_i \leqslant l_i$, $i = 1, 2, \ldots, s$. Серед усіх чисел такого виду найбільшим, очевидно, буде число, у якого показники n_i $(i = 1, 2, \ldots, s)$ найбільші, тобто число, у якому $n_i = l_i$. Отже,

$$(a, b) = p_1^{l_1} p_2^{l_2} \ldots p_s^{l_s}.$$

Теорема 9. *Найменшим спільним кратним чисел a і b є число*

$$[a, b] = p_1^{n_1} p_2^{n_2} \ldots p_s^{n_s},$$

де $n_i = \max(k_i, m_i)$, $i = 1, 2, \ldots, s$.

Доведення. Нехай M — будь-яке кратне чисел a і b. Число M ділиться на кожне з чисел a і b і тому воно ділиться й на кожне з чисел $p_i^{n_i}$, де $n_i = \max(k_i, m_i)$, $i = 1, 2, \ldots, s$.

Оскільки для числа M існує тільки один канонічний розклад, то з викладеного вище випливає, що він має такий вигляд:

$$M = p_1^{t_1} p_2^{t_2} \ldots p_s^{t_s} q_1^{s_1} q_2^{s_2} \ldots q_r^{s_r},$$

де $t_i \geqslant n_i$, $s_j \geqslant 0$, $i = 1, 2, \ldots, s$, $j = 1, 2, \ldots, r$.

Серед усіх чисел такого виду найменшим, очевидно, буде те, у якого $t_i = n_i$, $s_j = 0$, $i = 1, 2, \ldots, s$, $j = 1, 2, \ldots, r$. Отже,

$$[a, b] = p_1^{n_1} p_2^{n_2} \ldots p_s^{n_s}.$$

Теореми 8 і 9 поширюються на будь-яке скінченне число натуральних чисел. Зауважимо, що з цих теорем випливають відомі читачеві з шкільного курсу арифметики правила знаходження найбільшого спільного дільника і найменшого спільного кратного кількох чисел:

Щоб знайти НСД *даних чисел, треба кожне з цих чисел розкласти в добуток простих множників і взяти добуток спільних простих множників з найменшими показниками, з якими вони входять у всі розклади.*

Щоб знайти НСК *даних чисел, потрібно кожне з них розкласти в добуток простих множників і взяти добуток усіх простих множників з найбільшими показниками, з якими вони входять у всі розклади.*

7.5. Числові функції. Число і сума натуральних дільників.

Означення. *Функція $f(x)$ називається числовою, якщо вона визначена при всіх натуральних значеннях аргументу x.*

За цим означенням x^n, $\sin x$, $\ln x$, e^x — числові функції. Вивчаючи властивості цілих чисел, доводиться розглядати числові функції, визначені лише при натуральних значеннях аргументу x. Прикладами таких функцій є число $\tau(n)$ додатних дільників натурального n і сума $\sigma(n)$ додатних дільників натурального n. Розглянемо числові функції $\tau(n)$, $\sigma(n)$ $\varnothing$; $\tau(1) = 1$, $\tau(4) = 3$, оскільки число 4 має три додатні дільники: 1, 2, 4; $\tau(12) = 6$, бо число 12 має шість додатних дільників: 1, 2, 3, 4, 6, 12. $\sigma(1) = 1$, $\sigma(4) = 7$, $\sigma(12) = 28$. Якщо p — просте число, то $\tau(p) = 2$, а $\sigma(p) = p + 1$.

Теорема 10. *Якщо $n = p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$ — канонічний розклад натурального числа n, то*

$$\tau(p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}) = (k_1 + 1)(k_2 + 1) \ldots (k_m + 1).$$

Доведення. За теоремою 7 всі додатні дільники числа n збігаються з числами вигляду $d = p_1^{s_1} p_2^{s_2} \dots p_m^{s_m}$, де $0 \leqslant s_i \leqslant k_i$, $i = 1, 2, \dots, m$. Оскільки показник s_1 може набувати $k_1 + 1$ значень від 0 до k_1, s_2 може набувати $k_2 + 1$ значень від 0 до k_2, ..., s_m може набувати $k_m + 1$ значень від 0 до k_m, то $p_1^{s_1}$ може набувати $k_1 + 1$ різних значень, $p_2^{s_2}$ може набувати $k_2 + 1$ різних значень ..., $p_m^{s_m}$ може набувати $k_m + 1$ різних значень. Тому за узагальненим правилом добутку d може набувати $(k_1 + 1)(k_2 + 1) \dots (k_m + 1)$ різних значень. Отже, число n має $(k_1 + 1)(k_2 + 1) \dots (k_m + 1)$ додатних дільників, тобто

$$\tau(p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}) = (k_1 + 1)(k_2 + 1) \dots (k_m + 1).$$

Теорему доведено.

Приклади. 1. $\tau(100\,000) = \tau(2^5 \cdot 5^5) = 6 \cdot 6 = 36$.
2. $\tau(720) = \tau(2^4 \cdot 3^2 \cdot 5) = 5 \cdot 3 \cdot 2 = 30$.

Теорема 11. *Якщо* $n = p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}$ *— канонічний розклад натурального числа* n, *то*

$$\sigma(p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}) = \frac{p_1^{k_1+1} - 1}{p_1 - 1} \cdot \frac{p_2^{k_2+1} - 1}{p_2 - 1} \cdots \frac{p_m^{k_m+1} - 1}{p_m - 1}.$$

Доведення. За теоремою 7 множина всіх додатних дільників числа n збігається з множиною чисел

$$p_1^{s_1} p_2^{s_2} \dots p_m^{s_m} \qquad (s_i = 0, 1, 2, \dots, k_i, \quad i = 1, 2, \dots, m).$$

Тому

$$\sigma(p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}) = \Sigma p_1^{s_1} p_2^{s_2} \dots p_m^{s_m}$$
$$(s_i = 0, 1, 2, \dots, k_i, \ i = 1, 2, \dots, m).$$

Але

$$\Sigma p_1^{s_1} p_2^{s_2} \dots p_m^{s_m} = (1 + p_1 + p_1^2 + \dots + p_1^{k_1})(1 + p_2 + p_2^2 + \dots + p_2^{k_2}) \dots (1 + p_m + p_m^2 + \dots + p_m^{k_m}) \qquad (5)$$
$$(s_i = 0, 1, 2, \dots, k_i, \quad i = 1, 2, \dots, m).$$

Справді, виконавши множення у правій частині рівності (5), дістанемо доданки вигляду $p_1^{s_1} p_2^{s_2} \dots p_m^{s_m}$, де s_i набуває значень від 0 до k_i ($i = 1, 2, \dots, m$), тобто всі доданки лівої частини рівності (5), причому кожен з цих доданків одержиться лише один раз.

Оскільки $(1 + p_i + p_i^2 + \dots + p_i^{k_i}) = \dfrac{p_i^{k_i+1} - 1}{p_i - 1}$ $(i = 1, 2, \dots, m)$, то

$$\Sigma p_1^{s_1} p_2^{s_2} \dots p_m^{s_m} = \frac{p_1^{k_1+1} - 1}{p_1 - 1} \cdot \frac{p_2^{k_2+1} - 1}{p_2 - 1} \cdots \frac{p_m^{k_m+1} - 1}{p_m - 1}$$
$$(s_i = 1, 2, \dots, k_i, \quad i = 1, 2, \dots, m).$$

Отже,

$$\sigma(p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}) = \frac{p_1^{k_1+1} - 1}{p_1 - 1} \cdot \frac{p_2^{k_2+1} - 1}{p_2 - 1} \cdots \frac{p_m^{k_m+1} - 1}{p_m - 1}.$$

Приклади. 3. $\sigma(100\,000) = \sigma(2^5 \cdot 5^5) = \dfrac{2^6 - 1}{2 - 1} = \dfrac{5^6 - 1}{5 - 1} = 63 \cdot 3906 = 246\,078$.

4. $\sigma(720) = \sigma(2^4 \cdot 3^2 \cdot 5) = \dfrac{2^5 - 1}{2 - 1} \cdot \dfrac{3^3 - 1}{3 - 1} \cdot \dfrac{5^2 - 1}{5 - 1} = 2418$.

5. $\sigma(1001) = \sigma(7 \cdot 11 \cdot 13) = \dfrac{7^2 - 1}{7 - 1} \cdot \dfrac{11^2 - 1}{11 - 1} \cdot \dfrac{13^2 - 1}{14 - 1} = 1344$.

6. $\sigma(7429) = \sigma(17 \cdot 19 \cdot 23) = \dfrac{17^2 - 1}{17 - 1} \cdot \dfrac{19^2 - 1}{19 - 2} \cdot \dfrac{23^2 - 1}{23 - 1} = 8640$.

§ 8. РОЗПОДІЛ ПРОСТИХ ЧИСЕЛ

8.1. Формули, які задають прості числа. Аналіз таблиць простих чисел дав змогу вченим зробити деякі висновки про розподіл цих чисел у натуральному ряді й про їх властивості. Зокрема, грунтуючись на табличних підрахунках простих чисел, були зроблені спроби знайти прості алгебраїчні формули, які давали б прості числа. Так, Ферма[1] висловив припущення, що всі числа вигляду

$$F(n) = 2^{2^n} + 1 \qquad (1)$$

при цілому невід'ємному n прості. При $n = 0, 1, 2, 3, 4$ за формулою (1) маємо прості числа: $F(0) = 3$, $F(1) = 5$, $F(2) = 17$, $F(3) = 257$, $F(4) = 65537$.
Але $F(5)$, як встановив у 1732 р. Ейлер,— число складене: $F(5) = 641 \cdot 6700417$. Пізніше було доведено, що й деякі інші з чисел, заданих формулою (1), також складені. Отже, серед чисел, заданих формулою (1), є як числа прості, так і складені. Проте ще й досі невідомо, скінченною чи нескінченною є множина простих чисел, що задаються формулою (1).

Ейлер вказав на дві формули, які дають багато простих чисел:

$$F(n) = n^2 - n + 41, \qquad (2)$$

$$F(n) = n^2 - 79n + 1601. \qquad (3)$$

Формула (2) при $n = 0, 1, 2, 3, \dots, 39$ дає прості числа, а при $n = 40$ — складене число.

Формула (3) дає прості числа при $n = 0, 1, 2, \dots, 79$, а при $n = 80$ вона дає складене число. Пошуки алгебраїчних формул, які давали б лише прості числа, виявилися безуспішними. Ще більш безнадійною справою, очевидно, слід вважати відшукання такої формули, яка давала б усі прості числа.

8.2. Функція $\pi(x)$. Нерівності Чебишова. Ейлера, а згодом й інших математиків, цікавило питання, як часто зустрічаються прості числа в натуральному ряді. Як відомо, в натуральному ряді простих чисел безліч. А скільки їх міститься в тому чи іншому відрізку натурального ряду? Інакше кажучи, якщо позначити число простих чисел, що не перевищують числа x, символом $\pi(x)$, то яка природа функції $\pi(x)$? Аналізуючи таблиці простих чисел, які на той час досягали мільйона, відомий французький математик Лежандр у 1798—1808 рр. дістав емпіричну формулу для набли-

[1] П'єр Ферма (1601—1665) — французький юрист і математик.

женого обчислення функції $\pi(x)$:

$$\pi(x) \approx \frac{x}{\ln x - 1{,}08366}.$$

Звичайно, залишалося невідомим, чи придатна ця формула для чисел x, більших від мільйона. На основі табличних підрахунків простих чисел Гаусс у 1849 р. прийшов до висновку, що відношення $\frac{\Delta\pi(x)}{\Delta x}$ для великих x наближено виражається функцією $\frac{1}{\ln x}$ і $\pi(x) \approx \int_2^x \frac{dx}{\ln x}$.

Результат Гаусса було опубліковано тільки в 1863 р. Між тим у 1848 і 1850 рр. з'явилися дві роботи видатного російського математика П. Л. Чебишова, присвячені розподілу простих чисел. У статті «Про визначення простих чисел, що не перевищують даної величини» (1848 р.) П. Л. Чебишов довів теорему: *при* $2 \leqslant x < \infty$ *функція* $\pi(x)$ *задовольняє безліч разів і нерівність* $\pi(x) > \int_2^x \frac{dt}{\ln t} - \frac{\alpha x}{\ln^n x}$ *і нерівність* $\pi(x) < \int_2^x \frac{dt}{\ln t} + \frac{\alpha x}{\ln^n x}$, *де* α *— як завгодно мале додатне число і* n *— як завгодно велике натуральне число.*

На основі цієї теореми Чебишов довів, що *вираз* $\frac{x}{\pi(x)} - \ln(x)$ при $x \to \infty$ *не може мати границі, відмінної від* -1.

З формули Лежандра $\pi(x) \approx \frac{x}{\ln x - 1{,}08366}$ випливає, що $\lim_{x\to\infty}\left[\frac{x}{\pi(x)} - \ln x\right] = -1{,}08366$, а не -1. Формула Лежандра, таким чином, суперечить твердженню, доведеному Чебишовим, і, отже, вона помилкова. Далі у названій роботі Чебишов показав, що точніше, ніж формула Лежандра, функцію $\pi(x)$ виражає функція $\frac{x}{\ln x - 1}$, а ще точніше при великих значеннях x її виражає функція $\int_2^x \frac{dt}{\ln t}$.

З теорем, доведених Чебишовим, випливає також, що *кожне з відношень* $\pi(x) : \frac{x}{\ln x}$ *і* $\pi(x) : \int_2^x \frac{dt}{\ln t}$ *при* $x \to \infty$ *не може мати границі, відмінної від 1.*

У роботі «Мемуар про прості числа» (1850 р.) Чебишов встановив межі, між якими лежить функція $\pi(x)$. Він довів, що при $2 \leqslant x < \infty$ виконуються нерівності

$$0{,}92129\,\frac{x}{\ln x} < \pi(x) < 1{,}10555\,\frac{x}{\ln x}, \qquad (4)$$

тобто $0{,}92129 < \pi(x) : \frac{x}{\ln x} < 1{,}10555$.

Ці нерівності називають *нерівностями Чебишова.* З нерівностей Чебишова безпосередньо випливає теорема, яку вперше сформулював і довів Ейлер.

Теорема Ейлера. *При зростанні* x *до* ∞ *відношення* $\frac{\pi(x)}{x}$ *прямує до нуля.*

Справді, з нерівностей (4) дістаємо

$$0{,}92129\,\frac{1}{\ln x} < \frac{\pi(x)}{x} < 1{,}10555\,\frac{1}{\ln x}.$$

А звідси, оскільки $\lim_{x\to\infty}\frac{1}{\ln x} = 0$, $\lim_{x\to\infty}\frac{\pi(x)}{x} = 0$.

У першій сотні натуральних чисел міститься 25 простих чисел, тобто прості числа становлять 25%. У першій тисячі натуральних чисел є 169 простих чисел (16,9%); серед перших десяти тисяч натуральних чисел є 1229 простих чисел (12,29%); серед перших ста тисяч натуральних чисел 9522 простих числа (9,592%). Таблиці простих чисел показують, що при зростанні x процент простих чисел на відрізку $[1, x]$ зменшується. Проте цей факт сам по собі не виключає того, що при як завгодно великому x прості числа, що належать відрізку $[1, x]$, все-таки становлять не менш ніж, наприклад, 0,0000001% від числа всіх натуральних чисел відрізка $[1, x]$. Теорема Ейлера свідчить про те, що цього не може бути: при зростанні x до ∞ «середня щільність» $\frac{\pi(x)}{x}$ простих чисел на відрізок $[1, x]$ стає меншою від будь-якого числа $\varepsilon > 0$.

8.3. Асимптотичний закон розподілу простих чисел. Нерівності Чебишова вперше дали змогу судити про характер зростання функції $\pi(x)$ при зростанні x.

Результати Чебишова в дослідженні функції $\pi(x)$ привели математиків до висновку, що при $x \to \infty$ відношення $\pi(x) : \frac{x}{\ln x}$ прямує до 1. Як зазначалось вище, Чебишов довів, що коли при $x \to \infty$ границя відношення $\pi(x) : \frac{x}{\ln x}$ існує, то вона дорівнює 1. Проте довести існування цієї границі Чебишов не зміг. У 1881 р. англійському математикові Сільвестру вдалося вмістити відношення $\pi(x) : \frac{x}{\ln x}$ у більш тісні межі:

$$0{,}95695 < \pi(x) : \frac{x}{\ln x} < 1{,}04423. \qquad (5)$$

А в 1896 р. французький математик Адамар і бельгійський математик Валле Пуссен за допомогою апарата теорії функцій комплексної змінної незалежно один від одного довели існування границі відношення $\pi(x) : \frac{x}{\ln x}$. Таким чином, було доведено, що

$$\lim_{x\to\infty}\left[\pi(x) : \frac{x}{\ln x}\right] = 1.$$

Якщо при $x \to \infty$ *відношення* $\frac{f(x)}{g(x)}$ *додатних функцій* $f(x)$ *і* $g(x)$, *визначених для дій ших додатних вначень* x, *прямув до* 1, *тобто* $\lim_{x\to\infty}\frac{f(x)}{g(x)} = 1$, *то функції* $f(x)$ *і* $g(x)$ *називають асимптотично рівними і записують* $f(x) \sim g(x)$.
Отже,

$$\pi(x) \sim \frac{x}{\ln x}. \qquad (6)$$

Теорему про те, що $\lim_{x\to\infty}\left[\pi(x) : \frac{x}{\ln x}\right] = 1$, *називають асимптотичним законом розподілу простих чисел.*

Адамар і Валле Пуссен, власне, довели, що

$$\pi(x) \sim \int_2^x \frac{dt}{\ln t}. \qquad (7)$$

Формула (6) є безпосереднім наслідком з формули (7), оскільки, як легко довести,

$$\int_2^x \frac{dt}{\ln t} \sim \frac{x}{\ln x}.$$

Дослідження Рімана, Адамаре і Валле Пуссена показали, що за величиною абсолютної похибки функція $\int_2^x \frac{dt}{\ln t}$ дає більш точне наближення до $\pi(x)$, ніж $\frac{x}{\ln x}$. Із встановленням асимптотичного закону розподілу простих чисел дослідження характеру зростання функції $\pi(x)$ не припинилися. Дальші зусилля математиків були спрямовані на якомога точнішу оцінку модуля різниці між $\pi(x)$ і $\int_2^x \frac{dt}{\ln t}$. Істотних результатів в цьому напрямку досягли радянські математики М. Г. Чудаков, І. М. Виноградов, М. М. Коробов.

Вище ми зазначали, що при доведенні асимптотичного закону розподілу простих чисел Адамар і Валле Пуссен використовували зовнішні для теорії чисел методи теорії функцій комплексної змінної. Сам же асимптотичний закон належить виключно до натуральних чисел. Тому математикам здавалося природним шукати таке доведення цього закону, в якому не використовувалися б сторонні для теорії чисел ідеї. Пошуки такого доведення успішно закінчилися порівняно недавно: у 1949 р. норвезький математик А. Сальберг і угорський математик П. Ердеш опублікували доведення асимптотичного закону, в якому вони оперували тільки з натуральними числами.

8.4. Прості числа в арифметичних прогресіях. Після того як було доведено, що в натуральному ряді міститься нескінченна множина простих чисел, цілком природно постає питання: а скільки є простих чисел того або іншого спеціального вигляду? Інакше кажучи, скільки простих чисел міститься у тій або іншій підпослідовності натурального ряду?

Найпростішими підпослідовностями натурального ряду, звичайно, є арифметичні прогресії. Математиків, насамперед, цікавило питання про кількість простих чисел саме в арифметичних прогресіях. Наприклад, на початку арифметичної прогресії

3, 11, 19, 27, 35, *43,* 51, *59, 67,* 75, *83,* 91, 99, *107, ...,* (8)

різниця якої дорівнює 8, простих чисел порівняно багато (вони виділені курсивом). Але невідомо, скінченну чи нескінченну множину утворюватимуть прості числа, що містяться в цій прогресії.

Виявляється, не тільки в прогресії (8), а й у будь-якій іншій арифметичній прогресії натуральних чисел, перший член якої взаємно простий з її різницею, міститься нескінченна множина простих чисел. Це стверджується теоремою, яку довів Діріхле у 1837 р.

Теорема 1 (Діріхле). *У кожній арифметичній прогресії, перший член і різниця якої є взаємно прості натуральні числа, міститься нескінченна множина простих чисел.*

Доведення цієї теореми, яке дав Діріхле, грунтується на використанні теорії функцій комплексної змінної; воно досить складне, і тому ми не наводимо його. У 1948 р. А. Сельберг уперше довів цю теорему, не використовуючи сторонні для теорії чисел ідеї. Зауважимо, що умова теореми про взаємну простоту першого члена a і різниці b прогресії істотна: якщо $(a, b) = d \neq 1$, то кожен член прогресії $a, a+b, a+2b, \ldots$ ділиться на d і, отже, не є простим (хіба що за винятком першого члена a) числом. Для деяких окремих прогресій теорему Діріхле можна довести тим самим методом, яким ми доводили теорему Евкліда. Розглянемо два приклади таких прогресій. Множина всіх непарних чисел, очевидно, вичерпується такими двома арифметичними прогресіями:

1, 5, 9, 13, 17, 21, 25, 29, ..., (9)

3, 7, 11, 15, 19, 23, 27, 31, ... (10)

Ці прогресії складаються відповідно з чисел вигляду $4n+1$ і $4n+3$ (або $4n-1$, що те ж саме). Оскільки будь-яке просте число $p > 2$ непарне, то кожне просте число, більше від 2, належить до однієї з прогресій (8) та (9) і, таким чином, є число або вигляду $4n+1$, або вигляду $4n-1$.

Теорема 2. *Множина простих чисел вигляду* $4n-1$ *нескінченна.*

Доведення. Припустимо, що множина простих чисел вигляду $4n-1$ скінченна. Нехай вона складається з чисел $p_1, p_2, \ldots, p_s$. Розглянемо число $N = 4(p_1 p_2 \ldots p_s) - 1$ вигляду $4n-1$.

При діленні N на будь-яке з чисел $p_1, p_2, \ldots, p_s$ буде остача -1. Отже, число N відмінне від кожного з чисел $p_1, p_2, \ldots, p_s$ і жодне з цих простих чисел не є дільником N. З другого боку, не можуть усі прості дільники числа N бути числами вигляду $4n+1$, оскільки добуток чисел вигляду $4n+1$ є число такого самого вигляду:

$$(4m+1)(4n+1) = 16mn + 4m + 4n + 1 = 4(mn + m + n) + 1.$$

Таким чином, є дві можливості: або число N просте, або воно розкладається в добуток простих множників, серед яких є множник вигляду $4n-1$, відмінний від чисел $p_1, p_2, \ldots, p_s$. В обох випадках існує просте число вигляду $4n-1$, відмінне від кожного з чисел $p_1, p_2, \ldots, p_s$. А це суперечить припущенню, що множина простих чисел вичерпується числами $p_1, p_2, \ldots, p_s$. Тому це припущення неправильне. Теорему доведено.

Множина всіх чисел, які не діляться на 2 або на 3, очевидно, вичерпується такими двома арифметичними прогресіями:

1, 7, 13, 19, 25, 31, 37, 43, ..., (11)

5, 11, 17, 23, 29, 35, 41, 47, ... (12)

Прогресія (11) складається з чисел вигляду $6n+1$, а прогресія (12) — з чисел вигляду $6n-1$ (або вигляду $6n+5$). Тому кожне просте число $p > 3$ належить до однієї з прогресій (11) та (12) і, отже, є число або вигляду $6n+1$, або $6n-1$. Міркуваннями, аналогічними щойно викладеним, доводиться така теорема.

Теорема 3. *Множина простих чисел вигляду* $6n-1$ *нескінченна.*

Складніше, але все ж порівняно просто, доводяться також теореми про нескінченність множин простих чисел вигляду $4n+1$ і $6n+1$. Про множини простих чисел, що задаються не загальними членами арифметичних прогресій, а більш складними формулами, відомо мало. Наприклад, вважається, що множина простих чисел вигляду n^2+1 нескінченна; ось кілька перших чисел з цієї множини: 2, 5, 17, 37, 101, 197, 257, ... Проте всі спроби з'ясувати це питання не мали успіхів і до цього часу не знайдено способів його розв'язання.

§ 9. СКІНЧЕННІ ЛАНЦЮГОВІ ДРОБИ

9.1. Ланцюгові дроби. Запис раціональних чисел у вигляді скінченних ланцюгових дробів. У теорії чисел, математичному аналізі, теорії ймовірностей і в обчислювальній математиці широко використовують так звані ланцюгові дроби. У цьому параграфі ми розглянемо питання про скінченні ланцюгові дроби та деякі застосування їх.

Нехай α — деяке дійсне число і нехай q_0 — найбільше з цілих чисел, не більших ніж α. Тоді

$$\alpha = q_0 + x_0 = q_0 + \frac{1}{\alpha_1}, \text{ де } \alpha_1 > 1.$$

Аналогічно можна записати

$$\alpha_1 = q_1 + x_1 = q_1 + \frac{1}{\alpha_2}, \quad \alpha_2 > 1,$$

$$\alpha_2 = q_2 + x_2 = q_2 + \frac{1}{\alpha_3}, \quad \alpha_3 > 1,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\alpha_{n-1} = q_{n-1} + x_{n-1} = q_{n-1} + \frac{1}{\alpha_n}, \quad \alpha_n > 1,$$

$$\alpha_n = q_n + x_n = q_n + \frac{1}{\alpha_{n+1}}, \quad \alpha_{n+1} > 1$$

і т. д.

Якщо число α раціональне, то, як буде доведено нижче, для деякого натурального n матимемо $x_n = 0$, отже, записи, про які щойно йшла мова, обірвуться; якщо ж число α ірраціональне, то їх, очевидно, можна буде продовжувати нескінченно.

У першому випадку матимемо:

$$\alpha = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}}},$$

у другому випадку

$$\alpha = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n + \cdots}}}}$$

Означення 1. *Вираз вигляду*

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdots}}},$$

де q_0, q_1, q_2, ... — деякі цілі числа — називають елементарним ланцюговим, або елементарним неперервним, дробом.

Числа q_0 q_1, q_2, q_3, ... називають *елементами* даного ланцюгового дробу, а правильні дроби $\frac{1}{q_1}$, $\frac{1}{q_2}$, $\frac{1}{q_3}$, ... — відповідно першою, другою, третьою і т. д. *ланкою ланцюгового дробу*. Число ланок у ланцюговому дробі може бути як скінченним, так і нескінченним. Ланцюговий дріб, у якого число ланок скінченне, записують у вигляді

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}} \qquad (1)$$

і називають *скінченним,* точніше, *n-членним* ланцюговим дробом, а дріб, у якого число ланок нескінченне, записують у вигляді

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots}}$$

і називають *нескінченним* ланцюговим дробом.

Ми обмежимося розглядом скінченних елементарних ланцюгових дробів і називатимемо їх просто скінченними ланцюговими дробами. Крім того, вважатимемо, що в ланцюговому дробі (1) $q_1 \geqslant 1$, $q_2 \geqslant 1$, ..., $q_{n-1} \geqslant 1$, $q_n > 1$, а q_0 може бути будь-яким цілим числом. Такі дроби становлять найбільш важливий і разом з тим найбільш вивчений клас ланцюгових дробів; вони лежать в основі майже всіх арифметичних і багатьох аналітичних застосувань теорії ланцюгових дробів.

Умовимося, для зручності, n-членний ланцюговий дріб

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots + \cfrac{1}{q_n}}}$$

позначати символом $[q_0; q_1, q_2, \ldots, q_n]$. Значення кожного скінченного ланцюгового дробу $[q_0; q_1, q_2, \ldots, q_n]$ знаходимо в результаті виконання скінченної кількості разів раціональних операцій над елементами цього дробу. Отже, за нашими припущеннями відносно елементів ланцюгового дробу кожен скінченний ланцюговий дріб, очевидно, виражає собою деяке раціональне число $\frac{a}{b}$. Покажемо, що справедливе також і обернене твердження.

Теорема 1. *Кожне раціональне число можна подати у вигляді деякого скінченного ланцюгового дробу.*

Доведення. Нехай z — довільно вибране раціональне число. Тоді $z = \frac{a}{b}$, де a і b — деякі цілі числа, причому $b \geqslant 1$.

Застосувавши до чисел a і b алгоритм Евкліда, дістанемо рівності:

$$\left.\begin{aligned} a &= bq_0 + r_1; \\ b &= r_1q_1 + r_2; \\ r_1 &= r_2q_2 + r_3; \\ &\cdots\cdots\cdots \\ r_{n-2} &= r_{n-1}q_{n-1} + r_n; \\ r_{n-1} &= r_nq_n, \end{aligned}\right\} \qquad (2)$$

де $b > r_1 > r_2 > \cdots > r_{n-1} > r_n > 0$.

З рівностей (2) випливають відповідно рівності

$$\left.\begin{aligned} \frac{a}{b} &= q_0 + \cfrac{1}{\left(\frac{b}{r_1}\right)}; \\ \frac{b}{r_1} &= q_0 + \cfrac{1}{\left(\frac{r_1}{r_2}\right)}; \\ \frac{r_1}{r_2} &= q_2 + \cfrac{1}{\left(\frac{r_2}{r_3}\right)}; \\ &\cdots\cdots\cdots \\ &\cdots\cdots\cdots \\ \frac{r_{n-2}}{r_{n-1}} &= q_{n-1} + \cfrac{1}{\left(\frac{r_{n-1}}{r_n}\right)}; \\ \frac{r_{n-1}}{r_n} &= q_n. \end{aligned}\right\}$$

Звідси

$$\frac{a}{b} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{\left(\frac{r_1}{r_2}\right)}} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\left(\frac{r_2}{r_3}\right)}}} =$$

$$= \cdots = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2} + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}$$

або скорочено

$$\frac{a}{b} = [q_0;\ q_1, q_2, \ldots, q_n]. \tag{3}$$

Теорему доведено.

У тому разі, коли z — ціле число, тобто коли $b = 1$, у рівностях (2) матимемо лише одну рівність: $a = 1 \cdot a + 0$ і ланцюговий дріб обірветься на $q_0 = a$.

Запис (3) *називають зображенням раціонального числа* $\frac{a}{b}$ *скінченним ланцюговим дробом* або *розкладом числа* $\frac{a}{b}$ *у скінченний ланцюговий дріб.*

Теорема 2. *Кожне раціональне число зображується тільки одним скінченним ланцюговим дробом.*

Доведення. Справді, якщо

$$\frac{a}{b} = [q_0;\ q_1, q_2, \ldots, q_n], \tag{4}$$

$$\frac{a}{b} = [q_0';\ q_1', q_2', \ldots, q_s'], \tag{5}$$

то

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}} = q_0' + \cfrac{1}{q_1' + \cfrac{1}{q_2' + \cdots + \cfrac{1}{q_{s-1}' + \cfrac{1}{q_s'}}}}. \tag{6}$$

Не втрачаючи загальності міркувань, вважатимемо, що $s \geqslant n$. Оскільки дроби

$$\cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}} \quad \text{і} \quad \cfrac{1}{q_1' + \cfrac{1}{q_2' + \cdots + \cfrac{1}{q_{s-1}' + \cfrac{1}{q_s'}}}}$$

менші від 1, то кожне з чисел q_0 і q_0' дорівнює цілій частині числа $\frac{a}{b}$ і тому $q_0 = q_0'$. Віднявши почастинно від рівності (6) рівність $q_0 = q_0'$, дістанемо

$$\cfrac{1}{q_1 + \cfrac{1}{q_2} + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}} = \cfrac{1}{q_1' + \cfrac{1}{q_2' + \cdots + \cfrac{1}{q_{s-1}' + \cfrac{1}{q_s'}}}}. \tag{7}$$

Дроби, що стоять у лівій і правій частинах рівності (7), мають однакові чисельники: кожен з чисельників дорівнює 1. Тому знаменники цих дробів також рівні між собою, тобто

$$q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3} + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}} = q_1' + \cfrac{1}{q_2' + \cfrac{1}{q_3'} + \cdots + \cfrac{1}{q_{s-1}' + \cfrac{1}{q_s'}}}.$$

Міркуваннями, аналогічними викладеним вище, доведемо, що $q_1 = q_1'$, $q_2 = q_2'$, $q_3 = q_3'$ і т. д.

Через n кроків ми прийдемо до рівності

$$q_n = q_n' + \cfrac{1}{q_{n+1}' + \cfrac{1}{q_{n+2}'} + \cdots + \cfrac{1}{q_{s-1}' + \cfrac{1}{q_s'}}},$$

де $s > n$. Звідси випливає, що $s = n$ і, отже, $q_n = q_n$, бо при $s > n$ ціле число q_n мало б дорівнювати дробовому числу

$$q_n' + \cfrac{1}{q_{n+1}' + \cfrac{1}{q_{n+2}'} + \cdots + \cfrac{1}{q_{s-1}' + \cfrac{1}{q_s'}}},$$

чого не може бути.

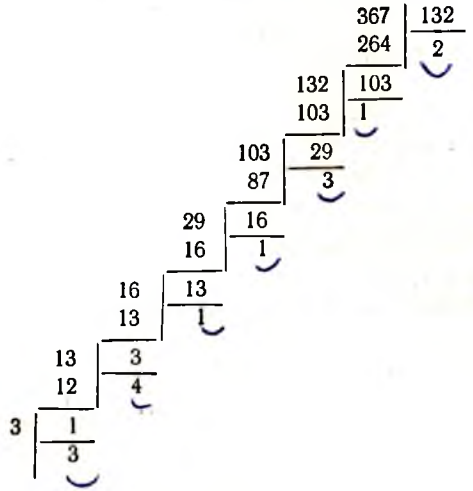
Отже, ми довели, що в зображеннях (4) і (5) $n = s$ і $q_0 = q_0'$, $q_1 = q_1'$, $q_2 = q_2'$, ..., $q_n = q_n'$, тобто ці зображення нічим не відрізняються одне від одного. Теорему доведено.

Примітка. Теорему 2 доведено в припущенні, що останній елемент скінченного ланцюгового дробу $q_n > 1$ ($n > 0$). Якщо ж не вимагати цього, то для раціонального числа $\frac{a}{b}$ існуватиме два розклади в скінченний ланцюговий дріб:

$$\frac{a}{b} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2} + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}} \quad (q_n > 1)$$

$$\frac{a}{b} = q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2}} + \cdots + \cfrac{1}{q_{n-1} + \cfrac{1}{(q_n - 1) + \cfrac{1}{1}}}.$$

Приклад. Розкласти у неперервний дріб число $-\frac{602}{367}$. Маємо $-\frac{602}{367} = -2 + \frac{132}{367}$. Виконаємо послідовне ділення



Отже, $-\frac{602}{367} = [-2;\ 2,\ 1,\ 3,\ 1,\ 1,\ 4,\ 3]$.

9.2. Підхідні дроби. Нехай

$$[q_0;\ q_1,\ q_2,\ \ldots,\ q_n] \tag{8}$$

є деякий ланцюговий дріб. Значенням цього ланцюгового дробу є деякий звичайний дріб $\frac{R}{S}$. Отже, ланцюговий дріб (8) зображується звичайним дробом $\frac{R}{S}$. Проте таке зображення не є єдиним, бо якщо $\frac{R}{S} = [q_0;\ q_1,\ q_2,\ \ldots,\ q_k]$, то й $\frac{Rl}{Sl} = [q_0, q_1, q_2,\ \ldots,\ q_n]$, де l — будь-яке відмінне від нуля ціле число. Всюди далі нам потрібно буде мати деяке цілком певне зображення скінченного ланцюгового дробу у вигляді звичайного дробу — зображення, яке ми називатимемо *канонічним*. Це зображення ми визначимо індуктивно [1].

Канонічним зображенням нуль-членного ланцюгового дробу $[q_0] = q_0$ вважатимемо дріб $\frac{q_0}{1}$. Припустимо тепер, що канонічне зображення визначене для кожного ланцюгового дробу, у якого число ланок менше ніж n, і визначимо його для n-членного ланцюгового дробу.

Розглянемо n-членний ланцюговий дріб $[q_0;\ q_1,\ q_2,\ \ldots,\ q_n]$. Як випливає з означення ланцюгового дробу, справедливе співвідношення

$$[q_0;\ q_1,\ q_2,\ \ldots,\ q_n] = q_0 + \frac{1}{[q_1;\ q_2,\ q_3,\ \ldots,\ q_n]}.$$

Дріб $[q_1;\ q_2,\ q_3,\ \ldots,\ q_n]$ — $(n-1)$-членний, отже, для нього канонічне зображення за припущенням уже визначене. Нехай цим зображенням є звичайний дріб $\frac{P'}{Q'}$, тоді

$$[q_0;\ q_1,\ q_2,\ \ldots,\ q_n] = q_0 + \frac{Q'}{P'} = \frac{q_0P' + Q'}{P'}.$$

Дріб $\frac{q_0P' + Q'}{P'}$ ми і вважатимемо канонічним зображенням ланцюгового дробу $[q_0;\ q_1,\ q_2,\ \ldots,\ q_n]$. Таким чином, тепер канонічне зображення однозначно визначене для будь-якого скінченного ланцюгового дробу. Позначимо канонічне зображення дробу $[q_0;\ q_1,\ q_2,\ \ldots,\ q_n]$ символом $\frac{P}{Q}$. Тоді для чисельників і знаменників канонічних зображень $\frac{P}{Q}$ і $\frac{P'}{Q'}$ ланцюгових дробів $[q_0;\ q_1,\ q_2,\ \ldots,\ q_n]$ і $[q_1;\ q_2,\ q_3,\ \ldots,\ q_n]$ матимемо співвідношення

$$P = q_0P' + Q', \quad Q = P'. \tag{9}$$

Умовимося називати ланцюговий дріб

$$[q_0;\ q_1,\ q_2,\ \ldots,\ q_s], \tag{10}$$

де $0 \leqslant s \leqslant n$, відрізком ланцюгового дробу (8).

Канонічне зображення відрізка (10) *називають s-м підхідним дробом або підхідним дробом порядку s ланцюгового дробу* (8).

Ланцюговий дріб (8) має $n + 1$ підхідних дробів:

$$\frac{P_0}{Q_0},\ \frac{P_1}{Q_1},\ \frac{P_2}{Q_2},\ \ldots,\ \frac{P_n}{Q_n}. \tag{11}$$

За означеннями підхідного дробу й канонічного зображення нуль-членного ланцюгового дробу

$$\frac{P_0}{Q_0} = \frac{q_0}{1}.$$

За формулами (9) $P_1 = q_0q_1 + 1$, $Q_1 = q_1$, отже, $\frac{P_1}{Q_1} = \frac{q_0q_1 + 1}{q_1}$,

$$P_2 = q_0(q_1q_2 + 1) + q_2 = q_2(q_0q_1 + 1) + q_0 = q_2P_1 + P_0,\ Q_2 = q_2q_1 + 1 = q_2Q_1 + Q_0 \tag{12}$$

і

$$\frac{P_2}{Q_2} = \frac{q_2P_1 + P_0}{q_2Q_1 + Q_0}.$$

Теорема 3. (Правило утворення підхідних дробів). *Для будь-якого* $s \geqslant 2$

$$P_s = q_sP_{s-1} + P_{s-2},\quad Q_s = q_sQ_{s-1} + Q_{s-2}. \tag{13}$$

[1] Див.: Хинчин А. Я. Цепные дроби. Изд. 3. М., Физматгиз, 1961.

Доведення. При $s = 2$, як показують співвідношення (12), формули (13) правильні. Припустимо, що вони правильні для $s = m - 1$ $(m > 2)$, і доведемо, що тоді вони правильні й для $s = m$. Розглянемо відрізок

$$[q_1;\ q_2,\ q_3,\ \ldots,\ q_m] \qquad (2 < m \leqslant n) \tag{14}$$

ланцюгового дробу (8). Підхідний дріб порядку r дробу (14) позначатимемо символом $\frac{P'_r}{Q'_r}$.

За формулами (9)

$$P_m = q_0 P'_{m-1} + Q'_{m-1} \text{ і } Q_m = P'_{m-1}. \tag{15}$$

Але оскільки за припущенням формули (13) правильні для $s = m - 1$, то, застосувавши їх до дробу $[q_1; q_2, q_3, \ldots, q_m]$, дістаємо:

$$P'_{m-1} = q_m P'_{m-2} + P'_{m-3},\quad Q'_{m-1} = q_m Q'_{m-2} + Q'_{m-3} \tag{16}$$

(тут стоїть q_m, а не q_{m-1}, оскільки дріб $[q_1; q_2, q_3, \ldots, q_m]$ починається з q_1, а не з q_0).

З співвідношень (15), (16) і за формулами (9)

$$P_m = q_0 (q_m P'_{m-2} + P'_{m-3}) + (q_m Q'_{m-2} + Q_{m-3}) = q_m (q_0 P'_{m-2} +$$
$$+ Q'_{m-2}) + (q_0 P'_{m-3} + Q'_{m-3}) = q_m P_{m-1} + P_{m-2};$$
$$Q_m = q_m P'_{m-2} + P'_{m-3} = q_m Q_{m-1} + Q_{m-2},$$

тобто $\quad P_m = q_m P_{m-1} + P_{m-2}$ і $Q_m = q_m Q_{m-1} + Q_{m-2}$.

Цим теорему доведено.

Формула (13) виражає чисельник P_s і знаменник Q_s підхідного дробу порядку s через елемент q_s і через чисельники і знаменники двох попередніх підхідних дробів й, отже, дає можливість за відомими підхідними дробами порядку $s - 2$ і $s - 1$ знайти підхідний дріб порядку s. Зауважимо, що на формулі (13) грунтується вся теорія ланцюгових дробів.

Обчислення чисельників і знаменників підхідних дробів за допомогою формули (13) зручно провадити за такою схемою:

	q_0	q_1	q_2	...	q_{n-1}	q_n
P_s	$P_0 = q_0$	$P_1 = q_0 q_1 + 1$	$P_2 = q_2 P_1 + P_0$	...	$P_{n-1} = q_{n-1} P_{n-2} + P_{n-3}$	$P_n = q_n P_{n-1} + P_{n-2}$
Q_s	$Q_0 = 1$	$Q_1 = q_1$	$Q_2 = q_2 Q_1 + Q_0$	...	$Q_{n-1} = q_{n-1} Q_{n-2} + Q_{n-3}$	$Q_n = q_n Q_{n-1} + Q_{n-2}$

Щоб обчислити P_s $(s = 2, 3, \ldots, n)$ за цією схемою, потрібно число q_s, що стоїть над P_s, помножити на число P_{s-1}, яке передує P_s, і до одержаного добутку додати число P_{s-2}, що передує P_{s-1}. За аналогічним правилом обчислюють Q_s.

Приклад 1. Знайти підхідні дроби ланцюгового дробу

$$[-2;\ 2,\ 1,\ 3,\ 1,\ 1,\ 4,\ 3].$$

За наведеною вище схемою:

q_s	−2	2	1	3	1	1	4	3
P_s	−2	−3	−5	−18	−23	−41	−187	−602
Q_s	1	2	3	11	14	25	114	367

$(S = 0, 1, 2, \ldots, 8)$.

Розглянемо деякі властивості підхідних дробів.

Теорема 4. *При $s = 1, 2, 3, \ldots, n$ справджується співвідношення*

$$P_s Q_{s-1} - P_{s-1} Q_s = (-1)^{s-1}. \tag{17}$$

Доведення. При $s = 1$ рівність (17) справедлива, бо $P_1 = q_0 q_1 + 1$, $Q_0 = 1$, $P_0 = q_0$, $Q_1 = q_1$, і тому $P_1 Q_0 - P_0 Q_1 = 1$. Припустимо, що рівність (17) правильна при $s = m$ $(1 \leqslant m \leqslant n - 1)$, і доведемо, що тоді вона правильна й при $s = m + 1$. Це справді так:

$$P_{m+1} Q_m - P_m Q_{m+1} = (P_m q_{m+1} + P_{m-1}) Q_m - P_m (Q_m q_{m+1} + Q_{m-1}) =$$
$$= -(P_m Q_{m-1} - P_{m-1} Q_m) = -(-1)^{m-1} = (-1)^m,$$

тобто

$$P_{m+1} Q_m - P_m Q_{m+1} = (-1)^m.$$

Отже, за принципом математичної індукції, рівність (17) правильна при будь-якому s $(1 \leqslant s \leqslant n)$. Теорему доведено. З теореми 4 випливає справедливість такого твердження.

Наслідок. *Кожний підхідний дріб — нескоротний.*

Доведення. Дріб $\frac{P_0}{Q_0}$ нескоротний, оскільки $Q_0 = 1$.

Нескоротний також і кожен з дробів $\frac{P_s}{Q_s}$ $(s = 1, 2, \ldots, n)$. Справді, припустимо, що деякий дріб $\frac{P_m}{Q_m}$ скоротний, тобто $(P_m, Q_m) = d > 1$ $(1 \leqslant m \leqslant n)$. Тоді ліва частина рівності

$$P_m Q_{m-1} - P_{m-1} Q_m = (-1)^{m-1}$$

ділитиметься на число d, а тому і права її частина $(-1)^{m-1}$ також має ділитися на d, що неможливо. Отже, наше припущення неправильне. Твердження доведено.

Цей висновок дає змогу застосувати розклад раціональних чисел у ланцюгові дроби для скорочення звичайних дробів. Справді, якщо звичайний дріб $\frac{P}{Q}$ розкласти у ланцюговий дріб, то останній підхідний дріб $\frac{P_n}{Q_n}$ цього ланцюгового дробу буде нескоротним дробом і дорівнюватиме $\frac{P}{Q}$.

Приклад 2. Скоротити дріб $\frac{P}{Q} = \frac{2329}{9911}$. Розклавши цей дріб у скінченний ланцюговий дріб, матимемо: $\frac{2329}{9911} = [0;\ 4,\ 3,\ 1,\ 10,\ 1,\ 2]$.

Знаходимо підхідні дроби

	0	4	3	1	10	1	2
P_s	0	1	3	4	43	47	137
Q_s	1	4	13	17	183	200	583

Як відомо, $\frac{2329}{9911} = \frac{P_6}{Q_6} = \frac{137}{583}$, де $\frac{137}{583}$ — нескоротний дріб.

Теорема 5. *При $s \geqslant 2$ справджується співвідношення*

$$P_sQ_{s-2} - P_{s-2}Q_s = (-1)^s q_s. \tag{18}$$

Доведення. Справді, за формулою (13)

$$P_s = q_sP_{s-1} + P_{s-2}, \quad Q_s = q_sQ_{s-1} + Q_{s-2}.$$

Тому

$$P_sQ_{s-2} - P_{s-2}Q_s = (q_sP_{s-1} + P_{s-2})Q_{s-2} - P_{s-2}(q_sQ_{s-1} + Q_{s-2}) =$$
$$= q_s(P_{s-1}Q_{s-2} - P_{s-2}Q_{s-1}).$$

Але оскільки за теоремою 4 $P_{s-1}Q_{s-2} - P_{s-2}Q_{s-1} = (-1)^{s-2} = (-1)^s$, то $P_sQ_{s-2} - P_{s-2}Q_s = (-1)^s q_s$. Цим теорему доведено.

Теорема 6. *Підхідні дроби парного порядку даного ланцюгового дробу утворюють зростаючу, а підхідні дроби непарного порядку — спадну послідовність.*

Доведення. Поділимо обидві частини співвідношення (18) на $Q_s \cdot Q_{s-2}$. Тоді матимемо:

$$\frac{P_s}{Q_s} - \frac{P_{s-2}}{Q_{s-2}} = \frac{(-1)^s q_s}{Q_sQ_{s-2}}.$$

Звідси випливає, що при s парному справджується нерівність

$$\frac{P_s}{Q_s} > \frac{P_{s-2}}{Q_{s-2}},$$

а при s непарному — нерівність

$$\frac{P_s}{Q_s} < \frac{P_{s-2}}{Q_{s-2}}.$$

Цим теорему доведено.

Теорема 7. *З двох підхідних дробів $\frac{P_{s-1}}{Q_{s-1}}$ і $\frac{P_s}{Q_s}$ даного ланцюгового дробу дріб парного порядку завжди менший від дробу непарного порядку.*

Доведення. Поділимо обидві частини співвідношення (17) на $Q_s \cdot Q_{s-1}$; тоді матимемо:

$$\frac{P_s}{Q_s} - \frac{P_{s-1}}{Q_{s-1}} = \frac{(-1)^{s-1}}{Q_sQ_{s-1}}.$$

Звідси випливає, що при парному s справджується нерівність

$$\frac{P_s}{Q_s} < \frac{P_{s-1}}{Q_{s-1}},$$

а при непарному s — нерівність

$$\frac{P_s}{Q_s} > \frac{P_{s-1}}{Q_{s-1}}.$$

Отже, з двох дробів $\frac{P_{s-1}}{Q_{s-1}}$ і $\frac{P_s}{Q_s}$ менший той, порядок якого парний. Теорему доведено.

З цієї теореми випливає справедливість такого твердження.

Наслідок. *Кожен підхідний дріб парного порядку даного ланцюгового дробу менший від будь-якого підхідного дробу непарного порядку цього ланцюгового дробу.*

Справді, якщо б принаймні один підхідний дріб $\frac{P_{2m}}{Q_{2m}}$ парного порядку був не менший від деякого підхідного дробу $\frac{P_{2k+1}}{Q_{2k+1}}$ непарного порядку, то за теоремою 6 останній підхідний дріб парного порядку був би більший від останнього підхідного дробу непарного порядку, а це суперечило б теоремі 7.

Нехай $\frac{P}{Q}$ — деяке раціональне число, задане у вигляді скінченного ланцюгового дробу:

$$\frac{P}{Q} = [q_0; q_1, q_2, \ldots, q_n],$$

а $\frac{P_0}{Q_0}, \frac{P_1}{Q_1}, \ldots, \frac{P_n}{Q_n}$ — підхідні дроби цього ланцюгового дробу. Тоді на основі щойно доведених теорем, враховуючи, що $\frac{P_n}{Q_n} = \frac{P}{Q}$, можна записати

$$\frac{P_0}{Q_0} < \frac{P_2}{Q_2} < \frac{P_4}{Q_4} < \cdots \leqslant \frac{P}{Q} \leqslant \cdots < \frac{P_5}{Q_5} < \frac{P_3}{Q_3} < \frac{P_1}{Q_1}.$$

Таким чином, підхідні дроби парного порядку є наближеними значеннями $\frac{P}{Q}$ з недостачею, а непарного порядку — з надлишком. Оцінка похибки при цьому визначається нерівністю

$$\left|\frac{P}{Q} - \frac{P_s}{Q_s}\right| < \frac{1}{Q_s^2}.$$

Справді,

$$\left|\frac{P}{Q} - \frac{P_s}{Q_s}\right| \leqslant \left|\frac{P_{s+1}}{Q_{s+1}} - \frac{P_s}{Q_s}\right| = (\text{за теоремою 4}) = \frac{1}{Q_{s+1}Q_s} =$$
$$= (\text{за теоремою 3}) = \frac{1}{(q_{s+1}Q_s + Q_{s-1})Q_s} < \frac{1}{Q_s^2}.$$

9.3. Розв'язування в цілих числах лінійного рівняння з двома невідомими. Розглянемо, як застосовують ланцюгові дроби для знаходження цілих розв'язків лінійного рівняння з двома невідомими, коефіцієнти і вільний член якого — цілі числа. Нехай

$$ax + by = c \tag{19}$$

є довільне рівняння. Якщо $a \neq 0$ і $b \neq 0$, то рівняння (19) неозначене: воно має безліч розв'язків. Загальним розв'язком цього рівняння є $x = \frac{c - by}{a}$ або $y = \frac{c - ax}{b}$.

Припустимо, що в рівнянні (19) a, b і c — цілі числа і що потрібно знайти цілі розв'язки цього рівняння, тобто розв'язки, які складаються з цілих чисел. У цьому разі при будь-якому цілому значенні y число $c - by$ буде цілим, а число $x = \frac{c - by}{a}$ при цьому, взагалі кажучи, не буде цілим, оскільки ціле число $c - by$ може і не ділитися на ціле число a. Про те, як знайти цілі розв'язки рівняння (19), у якого a, b і c — цілі числа, і йтиме далі мова. Якщо вільний член c рівняння (19) не ділиться на найбільший спільний дільник $d = (a, b)$ його коефіцієнтів a і b, то воно не має цілих розв'язків, бо в противному разі c повинно було б ділитися на $d = (a, b)$. Якщо ж c ділиться на $d = (a, b)$, то, поділивши обидві частини рівняння (19) на $d = (a, b)$, дістанемо рівняння, рівносильне даному, коефіцієнти якого є взаємно прості числа. Має місце така теорема.

Теорема 8. *Якщо пара цілих чисел x_0, y_0 задовольняє рівняння*

$$ax + by = c, \tag{20}$$

де a, b, c — цілі числа й $(a, b) = 1$, то

$$x = x_0 + bt, \quad y = y_0 - at, \tag{21}$$

де t — будь-яке ціле число, є загальним розв'язком цього рівняння в цілих числах.

Доведення. За умовою теореми

$$ax_0 + by_0 = c. \tag{22}$$

Віднявши почастинно від рівняння (20) рівність (22), дістанемо рівняння

$$a(x - x_0) + b(y - y_0) = 0, \tag{23}$$

рівносильне рівнянню (20). Покажемо, що формули (21) задають множину всіх цілих розв'язків рівняння (23), а отже, і рівняння (20). Очевидно, що кожна пара цілих чисел, задана формулами (21), задовольняє рівняння (23). Навпаки, якщо пара цілих чисел x_1, y_1 задовольняє рівняння (23), тобто $a(x_1 - x_0) + b(y_1 - y_0) = 0$, то $a(x_1 - x_0) = -b(y_1 - y_0)$. Звідси, оскільки $(a, b) = 1$, випливає, що $x_1 - x_0$ ділиться на b, тобто $x_1 - x_0 = bt$, де t — деяке ціле число. Тому $abt = -b(y_1 - y_0) \Rightarrow y_1 - y_0 = -at$.

З співвідношень $x_1 - x_0 = bt$, $y_1 - y_0 = -at$, де t — деяке ціле число, дістаємо $x_1 = x_0 + bt$, $y_1 = y_0 - at$, де t — деяке ціле число. Отже, кожна пара цілих чисел x_1, y_1, що задовольняє рівняння (23),

задається формулами (21). Цим теорему доведено. Таким чином, щоб розв'язати рівняння (20) в цілих числах, потрібно знайти який-небудь окремий цілий розв'язок (x_0, y_0) цього рівняння.

Зробити це можна, скориставшись розкладом числа $\frac{a}{b}$ у ланцюговий дріб. Справді, нехай $\frac{a}{b} = [q_0; q_1, q_2, \ldots, q_n]$ розклад числа $\frac{a}{b}$ у ланцюговий дріб, а $\frac{P_s}{Q_s}$ $(s = 0, 1, 2, \ldots, n)$ є підхідні дроби цього розкладу. Тоді $\frac{P_n}{Q_n} = \frac{a}{b}$. За умовою дріб $\frac{a}{b}$ — нескоротний і дріб $\frac{P_n}{Q_n}$ за висновком з теореми 4 також нескоротний; тому $P_n = a$, $Q_n = b$. За теоремою 4

$$P_nQ_{n-1} - P_{n-1}Q_n = (-1)^{n-1}, \text{ тобто } aQ_{n-1} - bP_{n-1} = (-1)^{n-1}.$$

Помноживши обидві частини останньої рівності на $(-1)^{n-1}c$, дістанемо рівність

$$a[(-1)^{n-1} \cdot c \cdot Q_{n-1}] + b[(-1)^n \cdot c \cdot P_{n-1}] = c.$$

Ця рівність означає, що пара чисел $x_0 = (-1)^{n-1} \cdot c \cdot Q_{n-1}$, $y_0 = (-1)^n \cdot c \cdot P_{n-1}$ є цілий розв'язок рівняння (20).

Теорема 9. *Загальний розв'язок у цілих числах рівняння $ax + by = c$, де a, b, c — цілі числа й $(a, b) = 1$, можна подати у вигляді*

$$x = (-1)^{n-1} \cdot c \cdot Q_{n-1} + bt, \quad y = (-1)^n \cdot c \cdot P_{n-1} - at, \tag{24}$$

де t — довільне ціле число, а P_{n-1} і Q_{n-1} — чисельник і знаменник передостаннього підхідного дробу розкладу числа $\frac{a}{b}$ у ланцюговий дріб.

Приклад. Розв'язати в цілих числах рівняння $61x + 48y = 3$.

Розклавши $\frac{61}{48}$ у ланцюговий дріб, матимемо:

$$\frac{61}{48} = [1; 3, 1, 2, 4].$$

Підхідними дробами для ланцюгового дробу $[1; 3, 1, 2, 4]$ є

$$\frac{1}{1}, \frac{4}{3}, \frac{5}{4}, \frac{14}{11}, \frac{61}{48}.$$

Передостаннім підхідним дробом є $\frac{P_3}{Q_3} = \frac{14}{11}$.

Отже, за формулами (24) загальним розв'язком у цілих числах заданого рівняння є

$$\begin{aligned} x &= (-1)^3 \cdot 3 \cdot 11 + 48t = -33 + 48t; \\ y &= (-1)^4 \cdot 3 \cdot 14 - 61t = 42 - 61t. \end{aligned} \tag{25}$$

У цьому загальному розв'язку $x_0 = -33$, $y_0 = 42$. Узявши у формулах (25) $t = 1$, дістанемо частинний розв'язок $x_1 = 15$, $y_1 = -19$, і загальний розв'язок заданого рівняння за теоремою 8 можна записати так:

$$x = 15 + 48t, \quad y = -19 - 61t.$$

ГРУПИ І КІЛЬЦЯ

Вивчаючи алгебру на першому курсі, читач ознайомився з основними алгебраїчними структурами — групами, кільцями, полями — і їх найпростішими властивостями.

Групи і кільця є найважливішими алгебраїчними структурами. Вони мають досить широку область застосувань і є предметом великих самостійних алгебраїчних наук — теорії груп і теорії кілець.

Цей розділ можна розглядати як вступ у теорію груп і теорію кілець. У ньому буде викладено елементарні відомості про групи і кільця, з якими повинен бути ознайомлений кожен учитель математики середньої школи. Значна частина цих відомостей знайде також застосування в наступних розділах цієї книги, а також в інших математичних курсах.

§ 10. ГРУПИ І ПІДГРУПИ

10.1. Групи. Як відомо, алгебраїчну операцію, визначену в групі, називають множенням або додаванням. Групу відносно операції множення називають *мультиплікативною,* а відносно операції додавання — *адитивною.*

Умовимося, як це прийнято в загальній теорії груп, розглядати далі в основному мультиплікативні групи.

Нагадаємо означення мультиплікативної групи й її найпростіші властивості.

Нехай G — непорожня множина, в якій визначена операція множення.

Означення 1. *Непорожня множина G, в якій визначена операція множення, називається групою, якщо виконуються такі умови:*

1. *Операція множення асоціативна.*
2. *Для операції множення в множині G здійсненна обернена операція — ділення, тобто для будь-яких елементів a і b множини G кожне з рівнянь $ax = b$ і $ya = b$ має у множині G розв'язок і притому тільки один.*

Якщо операція множення, визначена в групі G, комутативна, то група G називається *комутативною* або *абельовою.* Група G називається *скінченною,* якщо множина її елементів скінченна; вона називається *нескінченною,* якщо множина її елементів нескінченна. Число елементів скінченної групи називають *порядком групи.*

З означення групи (1, § 11) випливають такі наслідки.

1. *У кожній групі G можна виконувати лівосторонні і правосторонні скорочення: якщо $ab_1 = ab_2$ або $b_1a = b_2a$, то $b_1 = b_2$.*
2. *У кожній групі G існує і притому тільки один елемент e, такий, що* $\forall\limits_{a \in G} [ae = ea = a]$. **Елемент e називають одиничним елементом або одиницею групи G;** іноді одиничний елемент позначають також символом 1.
3. *У кожній групі G для будь-якого її елемента a існує єдиний обернений йому елемент a^{-1}, тобто такий, що $a^{-1}a = aa^{-1} = e$.*
4. *Які б не були цілі числа m і n, для кожного елемента a групи G справджуються рівності*

$$a^m \cdot a^n = a^{m+n}, \quad (a^m)^n = a^{mn}.$$

Скориставшись тим, що для кожного елемента $a \in G$ в групі G існує обернений елемент a^{-1}, розв'язки рівнянь $ax = b$ і $ya = b$, про які йшлося вище, можна записати в явному вигляді:

$$x = a^{-1}b, \quad y = ba^{-1}.$$

Рівносильним означенню 1 є таке означення групи (1, § 11).

Означення 2. *Непорожня множина G, в якій визначена операція множення, називається групою, якщо виконуються такі умови:*

1) *Операція множення асоціативна.*
2) *У множині G існує одиничний елемент.*
3) *Для кожного елемента $a \in G$ у множині G існує обернений елемент a^{-1}.*

Користуючись означенням 2, іноді легше буває перевірити, що дана множина є мультиплікативна група.

Прикладами мультиплікативних груп є множина всіх додатних раціональних чисел, всіх відмінних від нуля раціональних чисел, множина всіх додатних дійсних чисел, всіх відмінних від нуля дійсних чисел, множина всіх відмінних від нуля комплексних чисел. Усі ці групи — нескінченні, абельові. Прикладом мультиплікативної нескінченної некомутативної групи є множина неособливих матриць n-го порядку над полем комплексних чисел **C**. Множина всіх комплексних коренів n-го степеня з 1 є мультиплікативною абельовою групою порядку n.

10.2. Підстановки. Важливими прикладами скінченних некомутативних груп є групи підстановок. Спочатку пригадаємо деякі відомості про перестановки та підстановки з n елементів.

Нехай дано деяку множину M, що складається з n елементів. Елементи цієї множини можна перенумерувати за допомогою чисел 1, 2, 3, ..., n. Індивідуальні властивості елементів множини M далі не відіграватимуть ніякої ролі, тому ми просто вважатимемо, що множина M складається з чисел 1, 2, 3, ..., n.

Всяке розташування чисел 1, 2, 3, ..., n в деякому певному порядку називається перестановкою з n чисел або з n елементів.

Число різних перестановок з n елементів дорівнює $n! = 1 \cdot 2 \times \times 3 \ldots n$ (1, § 9). Прийнято вважати, що в перестановці $i_1, i_2, \ldots, i_k, \ldots, i_s, \ldots, i_n$ елементи i_k й i_s утворюють інверсію, якщо $i_k > i_s$, але i_k стоїть у перестановці лівіше від i_s.

Перестановку, елементи якої утворюють парне число інверсій, називають *парною,* а перестановку, елементи якої утворюють непарне число інверсій, називають *непарною.*

Перетворення перестановки, при якому деякі два її елементи міняються місцями, а решта елементів залишаються нерухомими, називають *транспозицією*.

Як відомо (1, § 25), *кожна транспозиція змінює парність перестановки*.

Нехай A_1 і A_2 — дві різні перестановки з символів 1, 2, 3, ..., n.

Якщо ми виконаємо в обох цих перестановках транспозицію будь-яких, але тих самих символів, то дістанемо дві різні перестановки A_1' і A_2'. Справді, якби перестановки A_1' і A_2' були однакові, то були б однакові й перестановки A_1 і A_2, оскільки їх дістають з A_1' і A_2' за допомогою зворотної транспозиції двох тих самих символів.

Теорема 1. *При $n \geqslant 2$ число парних перестановок з n елементів дорівнює числу непарних, тобто дорівнює $\frac{1}{2}\,n!$*

Доведення. Справді, виконаємо у всіх $n!$ перестановках з n елементів транспозицію двох тих самих елементів. У результаті дістанемо $n!$ різних перестановок, тобто всі $n!$ перестановок з n елементів. Але при цьому всі парні перестановки перейдуть у непарні, а непарні — в парні. Отже, число парних перестановок дорівнює числу непарних, тобто дорівнює $\frac{1}{2}\,n!$. Цим теорему доведено.

Відомо також (1, § 25), що *від кожної перестановки з n елементів можна перейти до будь-якої іншої перестановки з цих самих елементів за допомогою кількох транспозицій*.

Означення 1. *Всяке взаємно однозначне відображення множини $M = \{1, 2, 3, \ldots, n\}$ самої на себе називають підстановкою з n елементів або підстановкою n-го степеня.*

Підстановки позначатимемо великими буквами латинського алфавіту: A, B, C і ін.

Якщо при підстановці A число i ($i = 1, 2, \ldots, n$) відображається в число a_i, то записують

$$A = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}, \qquad (1)$$

тобто під кожним з чисел 1, 2, 3 ... n підписують те число, в яке воно відображається, і одержані два рядки беруть у дужки.

Запис (1) слід читати так: при підстановці A 1 переходить в a_1, 2 переходить в a_2, ..., n переходить в a_n.

Оскільки підстановка є *взаємно однозначним відображенням*, то всі числа $a_1, a_2, \ldots, a_n$ різні і, отже, другий рядок у записі (1) являє собою деяку перестановку з елементів 1, 2, 3, ..., n.

Слід зауважити, що стовпчики в записі (1) можна поміняти місцями, тобто у верхньому рядку замість перестановки 1, 2, 3, ..., n можна записати будь-яку іншу перестановку $b_1, b_2, b_3, \ldots, b_n$ з елементів 1, 2, 3, ..., n і потім у нижньому рядку числа $a_1, a_2, a_3, \ldots, a_n$ переставити так, щоб під числом i ($i = 1, 2, 3, \ldots, n$) стояло число a_i. В результаті дістанемо запис тієї самої підстановки, але вже іншого

вигляду. Наприклад,

$$\begin{pmatrix} 3 & 1 & 2 & 4 & 5 & 6 & \ldots & n \\ a_3 & a_1 & a_2 & a_4 & a_5 & a_6 & \ldots & a_n \end{pmatrix}, \quad \begin{pmatrix} 3 & 4 & 5 & \ldots & n & 1 & 2 \\ a_3 & a_4 & a_5 & \ldots & a_n & a_1 & a_2 \end{pmatrix},$$

$$\begin{pmatrix} n & n-1 & n-2 & \ldots & 3 & 2 & 1 \\ a_n & a_{n-1} & a_{n-2} & \ldots & a_3 & a_2 & a_1 \end{pmatrix}$$

є різні записи тієї самої підстановки $A = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}$, оскільки в кожному з них число i ($i = 1, 2, 3, \ldots, n$) переходить у число a_i.

Отже, будь-яку підстановку n-го степеня можна записати за допомогою двох перестановок з чисел 1, 2, 3, ..., n, підписаних одна під одною:

$$\begin{pmatrix} a_1 & a_2 & a_3 & \ldots & a_n \\ b_1 & b_2 & b_3 & \ldots & b_n \end{pmatrix},$$

причому верхню перестановку $a_1, a_2, a_3, \ldots, a_n$ завжди можна вибрати довільно.

Навпаки, якщо під деякою перестановкою $c_1, c_2, c_3, \ldots, c_n$ з чисел 1, 2, 3, ..., n ми підпишемо будь-яку іншу перестановку $d_1, d_2, d_3, \ldots, d_n$ з цих самих чисел, то дістанемо запис

$$\begin{pmatrix} c_1 & c_2 & c_3 & \ldots & c_n \\ d_1 & d_2 & d_3 & \ldots & d_n \end{pmatrix}$$

деякої підстановки n-го степеня — підстановки, при якій число c_i ($i = 1, 2, 3, \ldots, n$) переходить у число d_i.

Кожна підстановка n-го степеня звичайно може бути записана у вигляді (1). При такому записі підстановок різні підстановки n-го степеня відрізнятимуться одна від одної нижніми перестановками. Звідси випливає справедливість такого твердження.

Теорема 2. *Число підстановок n-го степеня дорівнює $n!$*

Доведення. Справді, підпишемо почергово під перестановкою 1, 2, 3, ..., n кожну перестановку з чисел 1, 2, 3, ..., n; тоді дістанемо всі можливі різні підстановки n-го степеня. Оскільки число всіх перестановок з n символів дорівнює $n!$, то й число всіх підстановок n-го степеня дорівнює $n!$. Цим теорему доведено.

Візьмемо деяку підстановку n-го степеня

$$A = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ a_{i_1} & a_{i_2} & \ldots & a_{i_n} \end{pmatrix}. \qquad (2)$$

Верхня і нижня перестановки в записі (2) можуть бути або однакової, або протилежної парності. Припустимо, що вони мають однакову парність. Нехай

$$A = \begin{pmatrix} j_1 & j_2 & \ldots & j_n \\ a_{j_1} & a_{j_2} & \ldots & a_{j_n} \end{pmatrix} \qquad (3)$$

є інший, довільно вибраний, запис підстановки A.

Покажемо, що в записі (3) верхня і нижня перестановки також мають однакову парність. Справді, перестановку $j_1, j_2, \ldots, j_n$, як відомо, можна одержати з перестановки $i_1, i_2, \ldots, i_n$ послідовним виконанням кількох транспозицій. Якщо одночасно з транспозиціями, що переводять перестановку $i_1, i_2, \ldots, i_n$ в перестановку $j_1, j_2, \ldots, j_n$, ми виконаємо і транспозиції відповідних символів у нижній перестановці, то, очевидно, від запису (2) ми перейдемо до запису (3). Проте одночасне виконання однієї транспозиції у верхній і нижній перестановках одночасно змінює парності цих перестановок на протилежні і, отже, зберігає збіг їх парностей. Тому верхня й нижня перестановки у записі (3) мають однакову парність. Якщо ж верхня і нижня перестановки у записі (2) мають протилежні парності, то й у записі (3) верхня і нижня перестановки також мають протилежні парності.

Оскільки запис (3) підстановки A вибраний довільно, то з щойно викладеного випливає, що або у всіх записах підстановки A парності верхньої й нижньої перестановок збігаються, або ж у всіх записах вони протилежні. Отже, збіг чи протилежність парностей верхньої й нижньої перестановок у записі даної підстановки є особливістю самої підстановки, а не того чи іншого запису її. Тому правомірно ввести таке означення парності й непарності підстановки.

Означення 2. *Підстановка A називається парною, якщо парності верхньої й нижньої перестановок довільного запису її збігаються; вона називається непарною, якщо парності цих перестановок протилежні.*

Рівносильним цьому означенню є таке означення.

Означення 3. *Підстановка A називається парною, якщо сумарне число інверсій у верхній і нижній перестановках довільного запису її парне, в протилежному разі вона називається непарною.*

Рівносильність означень 2 і 3 очевидна: загальне число інверсій у верхній і нижній перестановках запису підстановки буде парним тоді і тільки тоді, коли парності цих підстановок збігаються.

Якщо підстановка A записана у вигляді (1), тобто

$$A = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix},$$

то, оскільки верхня перестановка цього запису парна, парність підстановки A визначається парністю нижньої перестановки $a_1, a_2, \ldots, a_n$. Звідси випливає справедливість такого твердження.

Теорема 3. *При $n \geqslant 2$ число парних підстановок n-го степеня дорівнює числу непарних, тобто дорівнює $\frac{1}{2}\, n!$*

Справді, підписавши почергово під перестановкою 1, 2, 3, ..., n кожну з $\frac{1}{2}\, n!$ парних перестановок, дістанемо $\frac{1}{2}\, n!$ парних підстановок, а підписавши під перестановкою 1, 2, 3, ..., n кожну непарну перестановку, дістанемо $\frac{1}{2}\, n!$ непарних підстановок.

Приклад 1. Визначимо парність підстановки 6-го степеня.

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

У верхній перестановці цього запису 5 інверсій, а в нижній їх 11. Загальне число інверсій в обох перестановках — 16. Отже, підстановка A є парна.

Запишемо тепер розглядувану підстановку так:

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

У верхній перестановці цього запису 0 інверсій, а в нижній — 8. Загальне число інверсій — 8. Цей приклад показує, що при різних записах даної підстановки парність загального числа інверсій в обох перестановках запису її зберігається, а саме число інверсій, взагалі кажучи, змінюється.

10.3. Групи підстановок. Візьмемо дві довільні підстановки n-го степеня

$$A = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} \text{ і } B = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ a_{i_1} & a_{i_2} & \ldots & a_{i_n} \end{pmatrix}.$$

Виконаємо послідовно підстановки A і B. В результаті цього дістанемо підстановку

$$C = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_{i_1} & a_{i_2} & \ldots & a_{i_n} \end{pmatrix}.$$

Означення 1. *Підстановку C, що є результатом послідовного виконання підстановок A і B, називають добутком підстановки A на підстановку B і записують: $C = AB$.*

Наприклад, якщо

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \text{ і } B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}, \text{ то } AB = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

Справді, підстановка A відображає 1 в 2, а підстановка B відображає 2 в 4. Отже, AB відображає 1 в 4. Аналогічно відображаються й інші символи.

Операція множення підстановок n-го степеня при $n \geqslant 3$ некомутативна. Справді, для підстановок

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \ldots & n \\ 2 & 3 & 1 & i_4 & i_5 & \ldots & i_n \end{pmatrix} \text{ і } B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \ldots & n \\ 3 & 2 & 1 & j_4 & j_5 & \ldots & j_n \end{pmatrix}$$

$AB \neq BA$, оскільки підстановка AB відображає елемент 1 в 2, а підстановка BA відображає —1 в 1.

Операція множення підстановок асоціативна. Справді, нехай дано підстановки n-го степеня A, B і C. Припустимо, що елемент k ($1 \leqslant k \leqslant n$) при підстановці A переходить в a_k, a_k при підстановці B переходить у b_k, а b_k при підстановці C переходить у c_k. Тоді при підстановці AB елемент k переходить у b_k, а при підстановці $(AB)\,C$ елемент k перейде в c_k. Розглянемо тепер підстановку $A(BC)$. При підстановці A елемент k переходить в a_k, при підстановці BC елемент a_k перейде в c_k, тому при підстановці $A\,(BC)$ елемент k також перейде в c_k. Отже, $(AB)\,C = A\,(BC)$.

Теорема 4. *Множина всіх підстановок n-го степеня є група за множенням .*

Доведення. Справді, операція множення підстановок асоціативна. Серед підстановок n-го степеня є підстановка $E = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$, при якій кожен елемент відображається в самого себе; цю підстановку називають *тотожною*. Очевидно, що добуток будь-якої підстановки A на тотожну підстановку E, а також добуток E на A дорівнює A:

$$AE = EA = A,$$

тобто підстановка E відіграє роль одиничного елемента. Для кожної підстановки $A = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$ в множині підстановок n-го степеня існує підстановка

$$A^{-1} = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix},$$

така, що $AA^{-1} = A^{-1}A = E$.

Підстановку A^{-1} називають *оберненою* для підстановки A.

Отже, множина підстановок n-го степеня, за означенням 2 п. 10.1, є група. Цим теорему доведено. Групу всіх підстановок n-го степеня називають *симетричною* групою n-го степеня і позначають S_n. Порядок групи S_n дорівнює $n!$

Розглянемо тепер підстановку виду

$$\begin{pmatrix} 1 & 2 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & 2 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}.$$

Елементи, що залишаються нерухомими, замінимо крапками й запишемо цю підстановку скорочено

$$\begin{pmatrix} \dots & i & \dots & j & \dots \\ \dots & j & \dots & i & \dots \end{pmatrix}. \qquad (4)$$

Підстановку (4), очевидно, можна дістати з тотожної підстановки $E = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$ за допомогою виконання транспозиції i й j у нижній перестановці. На цій підставі підстановки виду (4) також називають *транспозиціями*. Умовимося позначати транспозицію (4) символом (ij).

Теорема 5. *Кожну підстановку* n-го *степеня можна подати у вигляді добутку кількох транспозицій.*

Доведення. Нехай дано деяку підстановку

$$A = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ i_1 & i_2 & i_3 & \dots & i_n \end{pmatrix}.$$

Відомо, що від кожної перестановки з n елементів можна перейти до будь-якої іншої перестановки з цих самих елементів за допомогою кількох транспозицій. Зокрема перестановку $i_1, i_2, \dots, i_n$ можна дістати з перестановки $1, 2, 3, \dots, n$, виконавши послідовно кілька транспозицій.

Припустимо, що $i_1, i_2, \dots, i_n$ можна дістати з $1, 2, 3, \dots, n$ послідовним виконанням транспозицій елементів

$$j_1 \text{ і } s_1,\ j_2 \text{ й } s_2,\ \dots,\ j_m \text{ й } s_m. \qquad (5)$$

Тоді підстановку A, очевидно, можна дістати з підстановки $E = \begin{pmatrix} 1 & 2 & 3 & \dots \\ 1 & 2 & 3 & \dots \end{pmatrix}$, виконавши послідовно в її нижній перестановці транспозиції (5). Але виконання транспозиції елементів j_k й j_l у нижній перестановці будь-якої підстановки

$$B = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$$

рівносильне множенню підстановки B справа на підстановку

$$\begin{pmatrix} \dots & j_k & \dots & j_l & \dots \\ \dots & j_l & \dots & j_k & \dots \end{pmatrix}, \text{ тобто на } (j_k, j_l).$$

Отже,

$$A = E(j_1, s_1)(j_2, s_2) \dots (j_m, s_m).$$

Опустивши в правій частині цієї рівності множник E, матимемо:

$$A = (j_1, s_1)(j_2, s_2) \dots (j_m, s_m).$$

Цим теорему доведено.

Кожну підстановку можна різними способами записати у вигляді добутку транспозицій, бо завжди до даного добутку транспозицій можна дописати, наприклад, дві транспозиції виду (i, j) (i, j), добуток яких дорівнює тотожній підстановці E. Проте справедлива така теорема.

Теорема 6. *У всіх записах даної підстановки у вигляді добутку транспозицій парність числа транспозицій буде та сама: вона збігається з парністю підстановки.*

Доведення. Для доведення теореми достатньо показати, що добуток будь-яких k транспозицій є підстановка, парність якої збігається з парністю числа k. При $k = 1$ це справді так, оскільки будь-яка транспозиція (i, j) є непарна підстановка. Припустимо, що це твердження правильне для $k - 1$ $(k > 2)$ множників. Тоді воно правильне й для k множників, оскільки числа $k - 1$ й k мають протилежні парності, а множення підстановки (добутку перших $(k - 1)$ транспозицій) на транспозицію рівносильне виконанню транспозиції в нижній перестановці підстановки і, отже, змінює її парність на протилежну. Цим теорему доведено.

Розглянемо тепер множину всіх парних підстановок n-го степеня. Як відомо, число її елементів дорівнює $\frac{1}{2}\, n!$

Теорема 7. *Множина всіх парних підстановок* n-го *степеня є група по множенню.*

Доведення. Нехай A і B — довільно вибрані парні підстановки n-го степеня. Оскільки розклад підстановки AB у добуток транспозицій можна дістати, записавши у вигляді добутку транспозицій підстановки A і B, то за теоремою 6 підстановка AB є парна. Отже, у

множині парних підстановок здійсненна операція множення. Множення підстановок n-го степеня, як відомо, асоціативне. А тому і множення парних підстановок асоціативне. Тотожна підстановка, що відіграє роль одиничного елемента, парна і, отже, належить до множини парних підстановок. Нарешті, якщо підстановка $A = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$ є парна, то й обернена їй підстановка $A^{-1} = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix}$ також парна і, отже, належить до множини парних підстановок. Цим теорему доведено.

Групу парних підстановок називають *знакозмінною групою* n-го *степеня*. Порядок цієї групи дорівнює $\frac{1}{2}\,n!$

Легко перевірити, що знакозмінна група n-го степеня при $n = 3$ комутативна, а при $n \geqslant 4$ — некомутативна.

Зауважимо, що симетричні й знакозмінні групи відіграють надзвичайно важливу роль у теорії скінченних груп і теорії Галуа.

Це обумовлюється тим, що, як можна довести, будь-яка група порядку n ізоморфна деякій підгрупі симетричної групи S_n. Внаслідок цього вивчення скінченних груп зводиться до вивчення груп підстановок. Слід також зазначити, що саме дослідження підстановок у зв'язку з проблемою розв'язування алгебраїчних рівнянь у радикалах (у кінці XVIII і в першій половині XIX ст.) було відправним пунктом розвитку загальної теорії груп.

10.4. Підгрупи. Нехай дано групу G і деяку підмножину H цієї групи. Підмножину H називають *підгрупою* групи G, якщо вона є групою відносно бінарної операції, визначеної в G. Справедлива теорема: *для того щоб підмножина H групи G була підгрупою цієї групи, необхідно й достатньо, щоб вона разом з будь-якими своїми елементами a і b містила й їх добуток ab і разом з кожним своїм елементом a містила також і обернений йому елемент a^{-1}* (1, § 11).

Кожна мультиплікативна група G, очевидно, має такі тривіальні підгрупи: саму групу G і так звану одиничну підгрупу, яка складається лише з одиничного елемента 1. Але, звичайно, в групі можуть бути й інші підгрупи. Так, група за множенням, що складається з 1 і -1, і мультиплікативна група додатних раціональних чисел Q^+ є підгрупами мультиплікативної групи всіх відмінних від нуля раціональних чисел. Мультиплікативна група відмінних від нуля раціональних чисел є підгрупою групи всіх відмінних від нуля дійсних чисел. Знакозмінна група n-го степеня є підгрупою симетричної групи n-го степеня.

Множина всіх матриць n-го порядку над числовим полем P, детермінант кожної з яких дорівнює 1, як легко довести, є мультиплікативна група: її називають *унімодулярною групою матриць*. Унімодулярна група матриць є підгрупою мультиплікативної групи всіх невироджених матриць n-го порядку над полем P.

Важливим прикладом підгруп є так звані ц и к л і ч н і п і д г р у п и. Нехай G — деяка група і a — довільний елемент цієї групи. Позначимо символом $\{a\}$ підмножину групи G, що складається з усіх степенів елемента a. Покажемо, що підмножина $\{a\}$ *є підгрупою групи G.*

Справді, добуток будь-яких двох елементів a^m і a^n з $\{a\}$ міститься в $\{a\}$, оскільки $a^m \cdot a^n = a^n \cdot a^m = a^{m+n}$. В $\{a\}$ міститься також елемент $1 = a^0$. Разом з усяким своїм елементом a^n підмножина $\{a\}$ містить і обернений йому елемент a^{-n}.

Означення. *Підгрупа $\{a\}$, що складається з усіх степенів елемента a, називається циклічною підгрупою групи G, породженою елементом a.*

Зауважимо, що можуть бути такі два випадки: 1) усі степені елемента a є різні елементи групи G; в цьому разі a називають *елементом нескінченного порядку;* 2) серед степенів елемента a є рівні між собою, наприклад $a^l = a^s$, де $s \neq l$. Це завжди буде так, якщо група G скінченна, але може трапитися й у нескінченній групі. Розглянемо другий випадок докладніше. Отже, припустимо, що $a^l = a^s$, де $s > l$.

Тоді $a^{s-l} = 1$, тобто існують додатні степені елемента a, які дорівнюють 1.

Нехай серед усіх додатних степенів елемента a, які дорівнюють 1, n є найменший, тобто

1) $a^n = 1$, $n > 0$;

2) якщо $a^l = 1$, $l > 0$, то $l \geqslant n$.

У цьому разі елемент a називають *елементом* скінченного порядку, а саме *порядку* n.

Якщо a є елемент n-го порядку, то породжена ним циклічна підгрупа $\{a\}$ складається з таких елементів:

$$1,\ a,\ a^2,\ a^3,\ \dots,\ a^{n-2},\ a^{n-1}. \tag{6}$$

Справді, всі ці елементи різні, бо якби $a^l = a^s$, $0 \leqslant l < s \leqslant n-1$, то $a^{s-l} = 1$ і, отже, a було б елементом порядку $s - l < n$.

З другого боку, будь-який інший степінь елемента a, додатний чи від'ємний, дорівнює одному з елементів (6). Справді, якщо k — деяке ціле число, то

$$k = nq + r,\ 0 \leqslant r < n \tag{7}$$

і тому $a^k = a^{nq+r} = (a^n)^q a = a^r$.

Звідси випливає, що коли $a^k = 1$, то в рівності (7) $r = 0$, тобто k ділиться на n, бо в противному разі a було б елементом порядку $r < n$. Множина (6) складається з n елементів; отже, порядок циклічної підгрупи $\{a\}$ дорівнює порядку елемента a, що породжує цю підгрупу.

Зауважимо, що в кожній групі G є єдиний елемент першого порядку — це 1. Циклічна підгрупа $\{1\}$ збігається з одиничною підгрупою.

Доведемо тепер одну теорему, що стосується підгруп даної групи.

Теорема 8. *Якщо H і F є підгрупи групи G, то їх перетин $H \cap F$ також є підгрупа цієї групи.*

Д о в е д е н н я. Справді, якщо елементи a і b належать перетину $H \cap F$, то вони містяться в кожній з підгруп H і F. Отже, елементи ab і a^{-1} також містяться в кожній з підгруп H і F, а тому ab та a^{-1} міститься і в перетині $H \cap F$. Отже, за сформульованою вище теоремою $H \cap F$ є підгрупа групи G. Цим теорему доведено.

Доведена нами теорема поширюється на будь-яке число (скінченне чи нескінченне) підгруп групи G.

10.5. Циклічні групи. *Означення. Група G називається циклічною, якщо вона складається з степенів одного з своїх елементів a, тобто збігається з однією з своїх циклічних підгруп $\{a\}$.*

Елемент a називають *твірним елементом* циклічної групи $\{a\}$. Кожна циклічна група абельова, бо $a^m a^n = a^n a^m = a^{m+n}$.

Приклади циклічних груп. 1. Адитивна група цілих чисел $\mathbf{Z}$ є нескінченна циклічна група. Її твірним елементом є число 1. За твірний елемент цієї групи, очевидно, можна взяти також число -1.

2. Мультиплікативна група коренів n-го степеня з 1 є циклічною групою порядку n. Справді, множини всіх коренів n-го степеня з 1 є група по множенню (1, § 16). Як відомо, корені n-го степеня з 1 знаходять за формулою:

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

За формулою Муавра

$$\left(\cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}\right)^k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n},$$

тобто $(\varepsilon_1)^k = \varepsilon_k$.

Таким чином, кожен корінь n-го степеня з 1 є певним степенем кореня ε_1 і, отже, група коренів n-го степеня з 1 є циклічною групою $\{\varepsilon_1\}$, твірним елементом якої є корінь

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

Теорема 9. *Кожна нескінченна циклічна група ізоморфна адитивній групі цілих чисел $\mathbf{Z}$.*

Доведення. Нехай $G = \{a\}$ — довільна нескінченна циклічна група з твірним елементом a. Кожному елементу a^k групи G поставимо у відповідність елемент k групи $\mathbf{Z}$. Цим, очевидно, буде задано взаємно однозначне відображення групи G на групу $\mathbf{Z}$. Це відображення є ізоморфним, оскільки з $a^k \to k$ і $a^s \to s$ випливає, що $a^k \cdot a^s = a^{k+s} \to k + s$. Теорему доведено.

Теорема 10. *Кожна циклічна група порядку n ізоморфна мультиплікативній групі коренів n-го степеня з 1.*

Доведення. Нехай $G = \{a\}$ — довільна циклічна група з твірним елементом a порядку n. Вона складається з таких елементів: $a^0 = 1, a, a^2, a^3, \ldots, a^{n-1}$. Мультиплікативна група коренів n-го степеня з 1 складається з коренів $\varepsilon_0 = \varepsilon_1^0 = 1, \varepsilon_1, \varepsilon_2 = \varepsilon_1^2, \varepsilon_3 = \varepsilon_1^3, \ldots, \varepsilon_{n-1} = \varepsilon_1^{n-1}$.

Кожному елементу a^k групи G поставимо у відповідність елемент ε_1^k групи $\{\varepsilon_1\}$. Цим, очевидно, буде задано ізоморфне відображення групи циклічної групи G на групу коренів n-го степеня з 1, оскільки в $a^k \to \varepsilon_1^k$ і $a^s \to \varepsilon_1^s$ випливає, що $a^k a^s = a^{k+s} \to \varepsilon_1^{k+s} = \varepsilon_1^k \cdot \varepsilon_1^s$. Теорему доведено.

З теорем 9 і 10 випливає, що адитивною групою цілих чисел $\mathbf{Z}$ і мультиплікативною групою коренів n-го степеня з 1 по суті вичерпуються всі циклічні групи.

Доведемо тепер теорему про підгрупи циклічної групи.

Теорема 11. *Кожна підгрупа циклічної групи сама циклічна.*

Доведення. Нехай $G = \{a\}$ — довільна циклічна група з твірним елементом a і Q — деяка її підгрупа. Вважатимемо, що підгрупа Q відмінна від одиничної підгрупи E: в противному разі не треба було б доводити, що вона циклічна. Серед додатних степенів елемента a, що містяться в підгрупі Q, існує найменший, оскільки в будь-якій множині натуральних чисел, за принципом найменшого числа, існує найменше число. Нехай цим найменшим додатним степенем є a^k. Покажемо, що коли $a^l \in Q$, то l ділиться на k. Справді, за теоремою 1 § 5 $l = kq + r$, $0 \leqslant r < k$. Якщо $r > 0$, то в підгрупі Q міститься елемент $a^l (a^k)^{-q} = a^{kq+r} a^{-kq} = a^r$, тобто міститься додатний степінь елемента a, менший ніж a^k, що суперечить нашому припущенню. Отже, $r = 0$ і l ділиться на k. Теорему доведено.

10.6. Розклад групи за підгрупою. Нехай дано групу G та підмножини M і N цієї групи. *Сукупність усіх елементів групи G, кожен з яких можна записати у вигляді добутку деякого елемента з множини M на деякий елемент з множини N, називатимемо добутком множини M на множину N і позначатимемо його символом MN.*

Звичайно, одна з множин M і N може складатися лише з одного елемента. Якщо, наприклад, множина M складається з елемента a, то мова йтиме про добуток aN елемента a на множину N.

З асоціативності множення в групі G, як легко бачити, випливає асоціативність множення підмножин цієї групи:

$$(MN)P = M(NP).$$

Зауважимо, що *коли Q є підгрупа групи G, то $Q \cdot Q = Q$*. Справді, добуток ab будь-яких двох елементів a і b з підгрупи Q міститься в Q і тому $Q \cdot Q \subset Q$. З другого боку, $Q \subset Q \cdot Q$, оскільки $Q = Q \cdot l$. Отже, $Q \cdot Q = Q$.

Нехай H — довільно вибрана підгрупа групи G. Використавши підгрупу H, введемо на множині елементів $a, b, c, \ldots$ групи G бінарне відношення ρ (1, § 5), вважаючи, що $a\rho b \iff a^{-1}b \in H$, або, що те саме, $a\rho b \iff b = ah$, де h — деякий елемент підгрупи H.

Покажемо, що ρ є відношення еквівалентності (1, § 6). Справді,

1) $\forall\limits_{a \in G} [a\rho a]$, оскільки $a = a \cdot 1$, $1 \in H$;

2) $\forall\limits_{a,b \in G} [a\rho b \Rightarrow b\rho a]$, оскільки з $b = ah$, $h \in H$ випливає, що $a = bh^{-1}$;

3) $\forall\limits_{a,b,c \in G} [a\rho b \wedge b\rho c \Rightarrow a\rho c]$, бо з $b = ah_1$ і $c = bh_2$; $h_1, h_2 \in H$ випливає, що $c = a(h_1 h_2)$, $h_1 h_2 \in H$.

Ми знаємо, що будь-яке відношення еквівалентності, задане на множині M, визначає розбиття цієї множини на класи еквівалентних елементів, які не перетинаються. Отже, і відношення еквівалентності ρ визначає розбиття групи G на класи еквівалентних елементів. З'ясуємо, що являють собою ці класи розбиття. Якщо $H = G$, то розбиття складається лише з одного класу, бо $\forall\limits_{a,b \in G} [b = a(a^{-1}b)]$, $a^{-1}b \in H$ і, отже, $a\rho b$. Якщо $H = E = \{e\}$, то відношення еквівалентності ρ, очевидно, є звичайна рівність, і тому кожен елемент групи G становить

клас розбиття. Припустимо тепер, що H — підгрупа, відмінна від E і від G. Нехай B_l — один з класів розбиття групи G, яке матимемо при цьому, і нехай g — довільний елемент класу B_l. Тоді кожен елемент $b = gh$, де h — будь-який елемент з H, належить до B_l, оскільки $g\rho b$, і, навпаки, якщо $b \in B_l$, то $g\rho b$, тому $b = gh$, $h \in H$. Отже, $B_l = gH$. Таким чином, ми довели, що кожен клас розбиття групи G, що визначається відношенням еквівалентності ρ, коли $E \subset H \subset G$ є добуток gH довільно вибраного елемента g цього класу на підгрупу H. Ці класи розбиття називають *лівими суміжними класами групи G за підгрупою H*, а саме розбиття називають *лівостороннім розкладом* групи G за підгрупою H. Про суміжний клас $B_l = gH$ говорять, що він породжується елементом g.

Елемент g довільно вибраний в суміжному класі B_l. Отже, суміжний клас $B_l = gH$ породжується будь-яким з своїх елементів і тому будь-який з елементів класу gH можна взяти за представника цього класу. Зауважимо, що одним з лівих суміжних класів є сама підгрупа H. Цей суміжний клас породжується одиничним елементом e, а також будь-яким іншим елементом h з H, оскільки $hH = H$. Його позначають не eH або hH, а просто H.

Якщо група G скінченна, то лівосторонній розклад групи G за підгрупою H записують так:

$$G = H \dot{+} g_1H \dot{+} g_2H \dot{+} \cdots \dot{+} g_{s-1}H = \dot{\sum_{i=1}^{s-1}} q_iH_i \dot{+} H,$$

де знаки $\dot{+}$ і $\dot{\Sigma}$ означають об'єднання множин, що не перетинаються,— лівих суміжних класів. На множині елементів групи G можна було б ввести відношення еквівалентності ρ': $a\rho'b \Leftrightarrow ba^{-1} \in H$, тобто $a\rho'b \Leftrightarrow b = ha$, $h \in H$.

У цьому разі ми прийшли б до поняття *правого суміжного класу* Hg групи G за підгрупою H, породженого елементом g, і до *правостороннього розкладу* групи G за підгрупою H. Правосторонній розклад скінченної групи G за підгрупою H записують так:

$$G = H \dot{+} Hg_1 \dot{+} Hg_2 \dot{+} \cdots \dot{+} Hg_{s-1} = \dot{\sum_{i=1}^{s-1}} Hg_i \dot{+} H.$$

Цілком природно постає питання: лівосторонній і правосторонній розклади групи G за підгрупою H — це різні розбиття групи G на підмножини чи те саме розбиття? Інакше кажучи, відношення еквівалентності ρ і ρ' — це різні відношення чи те саме бінарне відношення?

Для абельової групи G лівосторонній і правосторонній розклади за будь-якою підгрупою H збігаються, оскільки $gH = Hg$ для будь-якого $g \in G$. Для неабельової ж групи розклади за однією підгрупою можуть збігатися, а за іншою — можуть виявитися різними.

Приклади розкладу групи за підгрупою. 1. Нехай G — адитивна група цілих чисел, а H підгрупа, що складається з усіх чисел, кратних натуральному числу k. Група G — абельова. Лівостороиній і правосторонній розклади цієї групи за підгрупою H збігаються. Кожен з цих розкладів складається з k різних суміжних класів, що породжуються відповідно числами 0, 1, 2, 3, ..., k — 1. Лівосторонній суміжний клас, що породжується числом l, $0 \leqslant l \leqslant k-1$, має вигляд $l + H$, а правосторонній $-H + l$.

2. Нехай G_n — група неособливих матриць n-го порядку над полем дійсних чисел R, а Q — підгрупа, що складається з матриць, детермінант кожної з яких дорівнює 1. Множина всіх матриць з рівними детермінантами становитиме лівий (а також і правий) суміжний клас.

Справді, якщо $B \in AQ_n$, тобто $B = AU$, $U \in Q_n$, то $|B| = |AU| = |A||U| = |A| \cdot 1 = |A|$, тобто $|B| = |A|$. Навпаки, якщо $|C| = |A|$, то $C = A(A^{-1}C) \in AQ_n$, оскільки $|A^{-1}C| = |A^{-1}||C| = |A|^{-1}|C| = 1$, тому $A^{-1}C \in Q_n$.

Отже, згрупувавши в один суміжний клас (лівий чи правий) всі матриці з рівними детермінантами, дістанемо розклад (відповідно лівосторонній чи правосторонній) групи G_n за підгрупою Q_n.

Цей приклад показує, що й у некомутативних групах можуть бути підгрупи, за якими лівосторонні розклади збігаються з правосторонніми.

3. Розглянемо тепер симетричну групу третього степеня S_3. Вона складається з підстановок

$$E = \begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad A = \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\quad B = \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\quad AB = \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},$$

$$B^2 = \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\quad AB^2 = \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}.$$

Підмножина групи S_3, що складається з підстановок E і A, є підгрупа цієї групи, оскільки, як легко перевірити, $A^2 = E$. Позначимо цю підгрупу символом H. Ліві суміжні класи групи S_3 за підгрупою H такі: $H = \{E, A\}$, $BH = \{B, AB^2\}$, $B^2H = \{B^2, AB\}$. Праві суміжні класи S_3 за H: $H = \{E, A\}$, $HB = \{B, AB\}$, $HB^2 = \{B^2, AB^2\}$.

Як бачимо, ліві й праві суміжні класи різні. Отже, лівосторонній і правосторонній розклади групи S_3 за підгрупою H різні. Цей приклад показує, що некомутативні групи можуть мати підгрупи, лівосторонні й правосторонні розклади за якими різні.

Для скінченних груп справедливе таке твердження.

Теорема Лагранжа. *У кожній скінченній групі порядок будь-якої її підгрупи є дільником порядку групи.*

Доведення. Нехай G — скінченна група порядку n, H — деяка її підгрупа порядку k. Розглянемо лівосторонній розклад групи G за підгрупою H. Припустимо, що він складається з s суміжних класів. (Число суміжних класів s називається *індексом* підгрупи H в групі G).

$$G = H \dot{+} g_1H \dot{+} g_2H \dot{+} \cdots \dot{+} g_{s-1}H. \qquad (8)$$

Підгрупа H складається з k елементів, а тому і кожен суміжний клас g_iH ($i = 1, 2, \ldots, s-1$) також складається з k елементів, бо якщо $g_ih_1 = g_ih_2$, де h_1 і h_2 — елементи з H, то $h_1 = h_2$. Отже, з розкладу (8) випливає, що

$$n = k \cdot s. \qquad (9)$$

Цим теорему доведено.

З теореми Лагранжа випливають такі наслідки:

Наслідок 1. *Порядок кожного елемента a скінченної групи G є дільником порядку групи.*

Справді, порядок елемента a дорівнює порядку породжуваної ним циклічної підгрупи $\{a\}$ і, отже, за теоремою Лагранжа, є дільником порядку групи G.

Наслідок 2. *Кожна скінченна група G, порядок якої є просте число p, є циклічна група.*

Справді, кожен відмінний від 1 елемент групи G є елементом порядку p і, отже, породжена ним циклічна підгрупа збігається з групою G.

§ 11. НОРМАЛЬНІ ДІЛЬНИКИ. ФАКТОР-ГРУПИ. ГОМОМОРФІЗМИ

11.1. Нормальні дільники. З викладеного вище ми знаємо, що групи можуть мати підгрупи, лівосторонні й правосторонні розклади, за якими істотно відрізняються, а також і підгрупи, за якими ці розклади збігаються. Підгрупи, лівосторонні й правосторонні розклади за якими збігаються, відіграють надзвичайно важливу роль в теорії груп. До вивчення таких підгруп ми і перейдемо.

Означення 1. *Підгрупа H групи G називається нормальним дільником цієї групи або інваріантною підгрупою, якщо лівосторонній і правосторонній розклади групи G за підгрупою H збігаються.*

Лівосторонній і правосторонній розклади групи G за підгрупою H збігатимуться, очевидно, тоді і тільки тоді, коли лівий суміжний клас gH групи G за підгрупою H, породжений будь-яким елементом $g \in G$, збігатиметься з її правим суміжним класом Hg, що містить елемент g. Тому поняття нормального дільника можна означити так.

Означення 2. *Підгрупа H групи G називається нормальним дільником цієї групи, якщо*

$$\forall_{g \in G} [gH = Hg].$$

Умова $\forall_{g \in G} [gH = Hg]$, *очевидно, означає, що*

$$\forall_{g \in G} \forall_{h \in H} \exists_{h', h'' \in H} [gh = h'g \wedge hg = gh'']. \tag{1}$$

Приклади нормальних дільників груп. 1. У будь-якій групі G сама група G і одинична підгрупа E є її нормальними дільниками: лівосторонній і правосторонній розклади групи G за підгрупою G складаються з одного суміжного класу G, а лівосторонній і правосторонній розклади групи за підгрупою E складаються з усіх елементів групи G.

2. У кожній абельовій групі G будь-яка її підгрупа H є нормальним дільником, оскільки для будь-якого елемента g групи G $gH = Hg$. Зокрема, мультиплікативна група додатних дійсних чисел є нормальним дільником мультиплікативної групи всіх відмінних від нуля дійсних чисел; мультиплікативна група відмінних від нуля раціональних чисел є нормальним дільником мультиплікативної групи відмінних від нуля дійсних чисел.

3. У мультиплікативній групі G_n невироджених матриць n-го порядку з елементами поля дійсних чисел $\mathbf{R}$ підгрупа Q_n матриць, детермінант кожної з яких дорівнює 1, є нормальним дільником цієї групи, оскільки лівосторонній і правосторонній розклади групи G_n за підгрупою Q_n, як показано вище, збігаються.

4. У симетричній групі n-го степеня S_n знакозмінна група n-го степеня A_n є нормальний дільник. Справді, оскільки група A_n складається з $\frac{1}{2} n!$ елементів, то і кожен суміжний (лівий і правий) клас групи S_n за підгрупою A_n також складається з $\frac{1}{2} n!$ елементів. Тому як лівосторонній, так і правосторонній розклади групи S_n за підгрупою A_n складаються з двох класів. Одним з цих суміжних класів є підгрупа A_n, а другим — сукупність непарних підстановок. Отже, лівосторонній і правосторонній розклади групи S_n за підгрупою A_n збігаються.

Теорема 1. *Підгрупа H групи G є її нормальним дільником тоді і тільки тоді, коли*

$$\forall_{g \in G} [h \in H \Rightarrow g^{-1}hg \in H]. \tag{2}$$

Доведення. Доведемо спочатку необхідність умови. Нехай H є нормальний дільник групи G. Тоді, за співвідношенням (1), $\forall_{g \in G} \forall_{h \in H} \exists_{h' \in H} [hg = gh']$. Отже, $\forall_{g \in G} \forall_{h \in H} \exists_{h' \in H} [g^{-1}hg = h']$, тобто $\forall_{g \in G} [h \in H \Rightarrow g^{-1}hg \in H]$. Цим необхідність умови доведено. Доведемо тепер достатність умови. Припустимо, що $\forall_{g \in G} [h \in H \Rightarrow g^{-1}hg \in H]$. Тоді $\forall_{g \in G} \forall_{h \in H} \exists_{h', h'' \in H} [g^{-1}hg = h' \wedge (g^{-1})^{-1}hg^{-1} = h'']$, тобто $\forall_{g \in G} \forall_{h \in H} \exists_{h', h'' \in H} [hg = gh' \wedge gh = h''g]$ і, отже, за означенням 2 H є нормальний дільник групи G. Цим достатність умови, а отже, і теорему доведено.

Елементи a і b групи G називають *спряженими* в цій групі, якщо в G існує принаймні один такий елемент g, що $b = g^{-1}ag$. Умова (2), таким чином, означає, що підгрупа H разом з кожним своїм елементом h містить і всі елементи, спряжені з ним в групі G.

Умову (2) часто беруть за означення нормального дільника.

Означення 3. *Підгрупа H групи G називається нормальним дільником цієї групи, якщо вона разом з кожним своїм елементом h містить і всі елементи, спряжені з ним в G.*

Користуючись цим означенням, легко довести таку теорему.

Теорема 2. *Перетин будь-якої множини нормальних дільників групи G є нормальним дільником цієї групи.*

Доведення. Нехай D — перетин деякої множини нормальних дільників групи G. Як перетин підгруп групи G D є підгрупа групи G. Якщо $a \in D$, то a міститься у всіх нормальних дільниках, перетином яких є D. Тому кожен елемент, спряжений з елементом a в G, також міститься у всіх цих нормальних дільниках, а отже, він міститься і в їх перетині D. Теорему доведено.

11.2. Фактор-групи. Нехай H — довільний нормальний дільник групи G. Оскільки кожен лівий суміжний клас gH групи G за нормальним дільником H є одночасно і правим суміжним класом Hg і навпаки, то далі ми говоритимемо просто про суміжні класи групи G за нормальним дільником H. Суміжний клас gH, породжений елементом g, позначатимемо $\bar{g}$. Виходячи з поняття добутку підмножин групи, означимо в множині суміжних класів групи G за нормальним дільником H операцію множення.

Нехай $\bar{g}_1 = g_1H$ і $\bar{g}_2 = g_2H$ — два довільні суміжні класи групи G за нормальним дільником H. Розглянемо добуток $\bar{g}_1 \cdot \bar{g}_2 = g_1H \cdot g_2H$ цих суміжних класів як підмножин групи G. Оскільки множення підмножин асоціативне й $H \cdot H = H$, то

$$\bar{g}_1 \cdot \bar{g}_2 = g_1H \cdot g_2H = (g_1Hg_2) \cdot H = [g_1 \cdot (g_2H)] \cdot H =$$
$$= (g_1 \cdot g_2)(H \cdot H) = g_1g_2H = \overline{g_1 \cdot g_2},$$

тобто

$$\bar{g}_1 \cdot \bar{g}_2 = \overline{g_1 \cdot g_2}. \qquad (3)$$

Отже, *добуток двох суміжних класів групи G за нормальним дільником H як підмножин групи G є суміжним класом G за H.* Цим у множині суміжних класів групи G за нормальним дільником H визначена операція множення.

Рівність (3) показує, що для відшукання добутку двох даних суміжних класів групи G за нормальним дільником H потрібно в кожному з цих класів вибрати по одному представнику і потім взяти той суміжний клас, до якого належить добуток вибраних представників.

Теорема 3. *Множина суміжних класів групи G за нормальним дільником H з визначеною в ній операцією множення є група. Вона називається фактор-групою групи G за нормальним дільником H і позначається символом G/H.*

Доведення. Справді, операція множення суміжних класів асоціативна — це випливає з асоціативності множення підмножин групи. Суміжний клас $\bar{e} = H$ відіграє роль одиничного елемента: для будь-якого суміжного класу $\bar{g} = gH$ справджуються рівності $\bar{g} \cdot \bar{e} = gH \cdot H = g(H \cdot H) = gH = \bar{g}$, тобто $\bar{g} \cdot \bar{e} = \bar{g}$; $\bar{e} \cdot \bar{g} = H \cdot gH = (Hg) \cdot H = (g \cdot H)H = gH = \bar{g}$, тобто $\bar{e} \cdot \bar{g} = \bar{g}$.

Для кожного суміжного класу $\bar{g} = gH$ існує обернений суміжний клас $\bar{g}^{-1} = \overline{g^{-1}} = g^{-1}H : \bar{g} \cdot \bar{g}^{-1} = \bar{g} \cdot \overline{g^{-1}} = gH \cdot g^{-1}H = (gHg^{-1}) \times H = 1 \cdot H \cdot H = H = \bar{e}$. Так само $g^{-1} \cdot g = e$. Цим теорему доведено.

Приклади фактор-груп. **1.** Нехай G адитивна група цілих чисел, а $H_k = \{k\}$ — підгрупа цілих чисел, кратних цілому числу k. Фактор-група G/H_k складається з суміжних класів $\bar{0} = H_k$, $\bar{1} = 1 + H_k$, $\bar{2} = 2 + H_k$, ..., $\overline{(k-1)} = (k-1) + H_k$.

2. Нехай S_n — симетрична група n-го степеня, а A_n — знакозмінна група n-го степеня. Як відомо, A_n є нормальний дільник S_n. Фактор-група S_n/A_n складається з двох суміжних класів: множини парних підстановок A_n і множини непарних підстановок B_n.

3. Нехай G_n — група неособливих матриць над полем дійсних чисел $\mathbf{R}$, а Q_n — нормальний дільник, що складається з матриць, детермінант кожної з яких дорівнює 1. Фактор-група G_n/Q_n складається з суміжних класів, кожен з яких містить усі матриці, детермінанти яких дорівнюють даному числу a.

Встановимо деякі найпростіші властивості фактор-груп.

Теорема 4. *Кожна фактор-група G/H абелевої групи G також абелева.*

Доведення. Справді, оскільки $\forall\limits_{a,b\in G} [ab = ba]$, то

$$\forall_{\bar{a},\bar{b}\in G/H} [\bar{a}\bar{b} = aH \cdot bH = abH = baH = bH \cdot aH = \bar{b} \cdot \bar{a}].$$

Теорема 5. *Кожна фактор-група G/H циклічної групи G також циклічна.*

Доведення. Нехай G — циклічна група, породжена елементом g, тобто $G = \{g\}$, H — деяка підгрупа групи G і aH — довільно вибраний елемент фактор-групи G/H. Тоді існує таке ціле число m, що

$a = g^m$, і тому $aH = g^m H = (gH)^m$. Отже, $G/H = \{gH\}$. Цим теорему доведено.

Теорема 6. *Порядок будь-якої фактор-групи G/H скінченної групи G є дільником порядку цієї групи.*

Доведення. Справді, порядок s фактор-групи G/H дорівнює індексу нормального дільника H в групі G і тому, за рівністю (9) п. 10.5, s є дільник порядку групи G.

11.3. Гомоморфізми груп. Природним узагальненням поняття ізоморфізму груп є гомоморфізм груп. З поняттям гомоморфізму груп, як можна пересвідчитися далі, тісно пов'язані поняття нормального дільника групи і фактор-групи.

Ізоморфізм φ групи G на групу G' визначається як взаємно однозначне відображення G на G', що не порушує множення:

$$\forall_{a,b\in G} [\varphi(ab) = \varphi(a) \cdot \varphi(b)].$$

Якщо не вимагати, щоб відображення було взаємно однозначним, а лише зберігало операцію множення, то ми приходимо до поняття гомоморфного відображення групи G в (або на) групу G'.

Означення 1. *Гомоморфізмом, або гомоморфним відображенням, групи G в групу G' називають відображення φ групи G в групу G', яке задовольняє умову:*

$$\forall_{a,b\in G} [\varphi(ab) = \varphi(a) \cdot \varphi(b)]. \qquad (4)$$

Якщо групи G і G' — адитивні, то умову гомоморфізму (4) можна записати так:

$$\forall_{a,b\in G} [\varphi(a + b) = \varphi(a) + \varphi(b)].$$

Якщо гомоморфне відображення $\varphi : G \to G'$ є відображенням групи G на групу G', то його називають *гомоморфізмом групи G на групу G'* або *епіморфізмом групи G*. В цьому разі говорять, що група G' є *гомоморфним образом* групи G, і пишуть $G \sim G'$.

Щоб зазначити, що φ є гомоморфізм групи G на групу G', пишуть $\varphi : G \sim G'$.

Приклади гомоморфного відображення групи на групу. **1.** Нехай S_n — симетрична група n-го степеня і G — мультиплікативна група, що складається з чисел 1 і — 1. Розглянемо відображення φ групи S_n на групу G, що задається таким способом: для кожної підстановки $A \in S_n$:

$\varphi(A) = 1$, якщо A — парна підстановка;

$\varphi(A) = -1$, якщо A — непарна підстановка.

Очевидно, що φ — відображення групи S_n на групу G, причому виконується умова $\varphi(AB) = \varphi(A) \cdot \varphi(B)$. Отже, φ — гомоморфізм групи S_n на групу G.

2. Нехай G_n — група неособливих матриць порядку n над полем дійсних чисел $\mathbf{R}$, $\mathbf{R}^*$ — мультиплікативна група відмінних від нуля дійсних чисел. Розглянемо відображення ψ групи G_n в групу $\mathbf{R}^*$, яке кожній матриці A ставить у відповідність її детермінант $|A|$. Для будь-якого числа $d \in \mathbf{R}^*$ існує така матриця $D = \begin{pmatrix} d & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$.

що $|D| = d$. Отже, ψ — однозначне відображення групи G_n на групу $\mathbf{R}^*$. А оскільки $|A \cdot B| = |A| \cdot |B|$, то ψ — гомоморфізм групи G_n на групу $\mathbf{R}^*$.

3. Нехай G — деяка група і H — будь-який її нормальний дільник. Нехай $\varkappa$ — відображення, яке кожному елементу $g \in G$ ставить у відповідність той суміжний клас gH групи G за нормальним дільником H, в якому міститься цей елемент. Очевидно, що $\varkappa$ є відображенням групи G на фактор-групу G/H. З означення, множення в фактор-групі G/H : $g_1H \cdot g_2H = g_1g_2H$ випливає, що

$$\forall_{g_1, g_2 \in G} [\varkappa(g_1g_2) = g_1g_2H = g_1H \cdot g_2H = \varkappa(g_1) \cdot \varkappa(g_2)].$$

Отже, $\varkappa$ — гомоморфне відображення групи G на G/H.

Гомоморфізм $\varkappa$ (приклад 3) називається *природним або канонічним* гомоморфізмом групи G на фактор-групу G/H.

Доведемо тепер кілька теорем, що характеризують гомоморфні відображення груп.

Теорема 7. *При гомоморфному відображенні φ групи G в групу G' одиничний елемент e групи G відображається в одиничний елемент e' групи G'.*

Доведення. Справді, з $e \cdot e = e$ випливає $\varphi(e) \cdot \varphi(e) = \varphi(e)$. З другого боку $e' \cdot \varphi(e) = \varphi(e)$. Отже, $\varphi(e) \cdot \varphi(e) = e' \cdot \varphi(e)$, а звідси $\varphi(e) = e'$.

Теорему доведено.

Теорема 8. *Якщо φ — гомоморфізм групи G в групу G', то*

$$\forall_{g \in G} [\varphi(g^{-1}) = [\varphi(g)]^{-1}].$$

Доведення. Справді, нехай $\varphi(g^{-1}) = g'$. Тоді

$$e' = \varphi(e) = \varphi(g \cdot g^{-1}) = \varphi(g) \cdot \varphi(g^{-1}) = \varphi(g) \cdot g',$$

$$e' = \varphi(e) = \varphi(g^{-1} \cdot g) = \varphi(g^{-1}) \cdot \varphi(g) = g'\varphi(g).$$

Звідси

$$\varphi(g^{-1}) = g' = [\varphi(g)]^{-1}.$$

Теорему доведено.

Теорема 9. *Якщо φ є гомоморфізм групи G в групу G', то $\varphi(G)$ є підгрупа групи G'.*

Доведення. Нехай a' і b' — два будь-які елементи з множини $\varphi(G)$. Тоді $a' = \varphi(a)$ і $b' = \varphi(b)$, де $a, b \in G$ і $a'b' = \varphi(a) \cdot \varphi(b) = \varphi(ab) \in \varphi(G)$, $a'^{-1} = [\varphi(a)]^{-1} = \varphi(a^{-1}) \in \varphi(G)$. Отже, $\varphi(G)$ є підгрупа групи G.

Означення 2. *Нехай φ є гомоморфне відображення групи G в групу G'. Сукупність K всіх елементів групи G, які при гомоморфізмі φ відображаються в одиницю e' групи G', називають ядром гомоморфізму φ і записують $K = \operatorname{Ker}\varphi$.*

Теорема 10. *Ядро будь-якого гомоморфізму φ групи G є нормальним дільником групи G.*

Доведення. Справді, якщо елементи a і b групи G містяться в $\operatorname{Ker}\varphi$, то й $ab \in \operatorname{Ker}\varphi$, бо $\varphi(ab) = \varphi(a) \cdot \varphi(b) = e' \cdot e' = e'$, якщо

$a \in \operatorname{Ker}\varphi$, тобто $\varphi(a) = e'$, то й $a^{-1} \in \operatorname{Ker}\varphi$, бо $\varphi(a^{-1}) = [\varphi(a)]^{-1} = e'$. Отже, $\operatorname{Ker}\varphi$ є підгрупа групи G. Нехай тепер a — довільний елемент ядра $\operatorname{Ker}\varphi$, а g — будь-який елемент групи G. Тоді

$$\varphi(g^{-1}ag) = \varphi(g^{-1}) \cdot \varphi(a) \cdot \varphi(g) = [\varphi(g)]^{-1} \cdot e' \cdot \varphi(g) = e'.$$

Таким чином, підгрупа $\operatorname{Ker}\varphi$ разом з будь-яким своїм елементом a містить і всі елементи, спряжені з ним у групі G, і тому $\operatorname{Ker}\varphi$ є нормальний дільник групи G. Теорему доведено.

Зауважимо, що ядром природного гомоморфізма групи G на фактор-групу G/H є, очевидно, нормальний дільник H.

Теорема про гомоморфізми груп. *Нехай φ є гомоморфізм групи G на групу G' і $H = \operatorname{Ker}\varphi$. Тоді група G' ізоморфна фактор-групі G/H, причому існує такий ізоморфізм ψ фактор-групи G/H на групу G', що добуток $\varkappa\psi$ природного гомоморфізму $\varkappa : G \sim G/H$ на ізоморфізм ψ є гомоморфізм φ.*

Доведення. Нехай g' — довільно вибраний елемент групи G', а g — такий елемент групи G, що $g' = \varphi(g)$. Оскільки H — ядро гомоморфізму φ, то $\forall_{h \in H} [\varphi(h) = e']$, тому $\forall_{h \in H} [\varphi(gh) = \varphi(g) \cdot \varphi(h) = g' \cdot e' = g']$, тобто кожен елемент суміжного класу $\bar{g} = gH$ при гомоморфізмі φ відображається в елемент g'. З другого боку, якщо елемент $q \in G$ при гомоморфізмі φ відображається в елемент $g' \in G'$, тобто $\varphi(q) = g'$, то $\varphi(g^{-1}q) = \varphi(g^{-1}) \cdot \varphi(q) = [\varphi(g)]^{-1} \cdot \varphi(q) = g'^{-1} \cdot g' = e'$, тому $g^{-1} \cdot q \in H$, тобто $g^{-1} \cdot q = h$, де $h \in H$. Звідси $q = gh \in gH = \bar{g}$. Таким чином, множина всіх елементів групи G, які при гомоморфізмі φ відображаються в елемент $g' \in G'$, становить суміжний клас $\bar{g} = gH$ групи G за нормальним дільником H. Позначимо символом ψ відображення, яке кожному суміжному класу $\bar{g} = gH$ ставить у відповідність елемент $g' \in G'$, у який при гомоморфізмі φ відображаються елементи класу $\bar{g} = gH$, тобто $\psi(\bar{g}) = \varphi(g)$. Очевидно, що ψ є відображення фактор-групи G/H на групу G'. Покажемо, що відображення ψ є ізоморфне. Справді, нехай $\bar{g}_1 = g_1H$ і $\bar{g}_2 = g_2H$ — будь-які елементи фактор-групи G/H. Оскільки $\bar{g}_1\bar{g}_2 = (g_1H) \cdot (g_2H) = (g_1g_2)H$, то $\psi(\bar{g}_1 \cdot \bar{g}_2) = \varphi(g_1 \cdot g_2) = \varphi(g_1) \cdot \varphi(g_2) = \psi(\bar{g}_1) \times \psi(\bar{g}_2)$. Отже,

$$\forall_{\bar{g}_1, \bar{g}_2 \in G/H} [\psi(\bar{g}_1 \cdot \bar{g}_2) = \psi(\bar{g}_1) \cdot \psi(\bar{g}_2)].$$

Крім того, відображення ψ взаємно однозначне, тобто

$$\forall_{\bar{g}_1, \bar{g}_2 \in G/H} [\bar{g}_1 \neq \bar{g}_2 \Rightarrow \psi(\bar{g}_1) \neq \psi(\bar{g}_2)],$$

оскільки

$$\psi(\bar{g}_1) = \psi(\bar{g}_2) \Rightarrow \varphi(g_1) = \varphi(g_2) \Rightarrow g_1H = g_2H \Rightarrow \bar{g}_1 = \bar{g}_2.$$

Отже, ми довели, що ψ є ізоморфне відображення фактор-групи G/H на групу G'. Розглянемо тепер відображення $\varkappa\psi$. Оскільки $\varkappa$ — при-

родний гомоморфізм групи G на фактор-групу G/H, а ψ — ізоморфізм фактор-групи G/H на групу G', то $\varkappa\psi$, очевидно, є відображення групи G на групу G'. Доведемо, що $\varkappa\psi = \varphi$. Нехай g — довільний елемент групи G. За означенням природного гомоморфізму $\varkappa$, $\varkappa(g) = \overline{g} = gH$ і, за означенням ізоморфізму ψ, $\psi(\overline{g}) = \varphi(g)$. Отже, $\varkappa\psi(g) = \psi[\varkappa(g)] = \psi(\overline{g}) = \varphi(g)$, тобто $\varkappa\psi(g) = \varphi(g)$. Таким чином, $\forall_{g\in G}[\varkappa\psi(g) = \varphi(g)]$. А це й означає, що $\varkappa\psi = \varphi$. Теорему доведено.

Ізоморфні групи з алгебраїчної точки зору, тобто з точки зору їх властивостей, які є наслідком визначених у них алгебраїчних операцій, а не індивідуальних властивостей їхніх елементів, нерозрізненні. Отже, теорема про гомоморфізми показує, що всі групи, на які може гомоморфно відображатися група G, по суті вичерпуються її фактор-групами, а всі гомоморфізми групи G вичерпуються природними гомоморфізмами G на її фактор-групи.

§ 12. КІЛЬЦЕ. ОБЛАСТЬ ЦІЛІСНОСТІ. ПОЛЕ ЧАСТОК

12.1. Елементарні відомості про кільця. Нагадаємо елементарні відомості про кільця, про які докладно йшла мова у першій частині підручника.

Кільцем називається непорожня множина K, в якій визначені дві бінарні операції — додавання і множення, причому за додаванням K є абельова група — адитивна група кільця K, а операція множення — асоціативна і пов'язана дистрибутивними законами з операцією додавання. Якщо операція множення в кільці K комутативна, то кільце K називають *комутативним*. Прикладами комутативних кілець є множина цілих чисел $\mathbf{Z}$, множина цілих чисел кратних деякому відмінному від 1 натуральному числу m (зокрема множина парних чисел), множина раціональних чисел $\mathbf{Q}$, множина дійсних чисел $\mathbf{R}$, множина комплексних чисел $\mathbf{C}$, множина всіх чисел вигляду $a + b\sqrt{2}$, де a і b — будь-які раціональні числа, множина всіх дійсних неперервних функцій від дійсного змінного x, заданих на відрізку $[0, 1]$, кільце Z_m класів, конгруентних за модулем m цілих чисел (1, § 12). Некомутативними кільцями є кільце Q_n квадратних матриць n-го порядку над полем раціональних чисел $\mathbf{Q}$, кільце R_n матриць n-го порядку над полем дійсних чисел $\mathbf{R}$, кільце C_n матриць n-го порядку над полем комплексних чисел $\mathbf{C}$. Скалярною матрицею над полем P називають матрицю, яка має на головній діагоналі той самий елемент a, а поза головною діагоналлю — нулі. Множина R_n^* всіх скалярних матриць n-го порядку над полем дійсних чисел $\mathbf{R}$ є комутативне кільце.

Справді, нехай

$$Q_a = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & a & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a \end{pmatrix} \text{ і } Q_b = \begin{pmatrix} b & 0 & \dots & 0 \\ 0 & b & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b \end{pmatrix}$$

є довільно вибрані матриці множини R_n^*. Тоді їх сума $Q_a + Q_b =$

$$= \begin{pmatrix} a+b & 0 & \dots & 0 \\ 0 & a+b & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a+b \end{pmatrix}, \text{ різниця } Q_a - Q_b =$$

$$= \begin{pmatrix} a-b & 0 & \dots & 0 \\ 0 & a-b & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a-b \end{pmatrix} \text{ і добуток}$$

$$Q_a \cdot Q_b = \begin{pmatrix} ab & 0 & \dots & 0 \\ 0 & ab & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & ab \end{pmatrix}$$

також є скалярні матриці і, отже, належать до множини R_n^*.

Операції додавання і множення матриць в R_n^*, як легко перевірити, асоціативні, комутативні й пов'язані дистрибутивним законом. Отже, R_n^* — комутативне кільце.

Поширюючи термінологію, яка вживається для цілих чисел, на елементи будь-якого кільця K, приймають таке означення.

Означення. *Елемент $b \in K$ називають лівим* (відповідно правим) *дільником елемента $a \in K$, якщо існує елемент $c \in K$ такий, що $a = bc$* (відповідно $a = cb$); *при цьому говорять також, що a є правим* (відповідно лівим) *кратним елемента b.*

Якщо кільце K — комутативне, то, оскільки порядок слідування множників у добутку можна змінити, поняття лівого дільника (кратного) збігається з поняттям правого дільника (кратного). Тому в цьому випадку говорять просто «дільник» і «кратне».

Зауважимо, що коли в кільці K немає одиниці, тобто такого елемента e, що $\forall_{a\in K}[ae = ea = a]$, елемент $a \in K$ може не бути дільником (лівим чи правим) самого себе. Так, у кільці парних чисел жодне з відмінних від нуля чисел не є дільником самого себе. Так само, якщо в кільці K немає одиниці e, то елемент na, де $a \in K$, а n — деяке ціле число, не буде, взагалі кажучи, кратним елемента a у смислі наведеного вище означення. Так, у кільці цілих чисел, кратних 3, елемент $5(3) = 15$ не є кратним елемента 3, оскільки число 5 не є елементом кільця.

Якщо ж у кільці K є одиничний елемент e, то для будь-якого $a \in K$

$$na = n \cdot (ea) = \underbrace{ea + ea + \dots + ea}_{n \text{ доданків}} = \underbrace{(e + e + \dots + e)}_{n \text{ доданків}} a = ne \cdot a.$$

Отже, na є кратним елемента a.

Підмножина K' кільця K називається *підкільцем* кільця K, якщо K' є кільце відносно операцій додавання і множення, визначених у кільці K.

Так, кільце парних чисел є підкільце кільця цілих чисел $\mathbf{Z}$, а останнє, в свою чергу, є підкільцем кільця раціональних чисел $\mathbf{Q}$. Кільце раціональних чисел і кільце чисел вигляду $a + b\sqrt{2}$, де a і b — будь-які раціональні числа, є підкільцями кільця дійсних чисел $\mathbf{R}$.

Кільце Q_n матриць n-го порядку над полем раціональних чисел $\mathbf{Q}$ є підкільце кільця $\mathbf{R}_n$ матриць n-го порядку над полем дійсних чисел $\mathbf{R}$, а останнє, в свою чергу, є підкільцем кільця C_n матриць порядку n над полем комплексних чисел $\mathbf{C}$. Кільце R_n^* скалярних матриць є підкільце кільця R_n.

У кожному кільці K, очевидно, є такі підкільця: само кільце K і нульове підкільце, що складається лише з нуля кільця K. Ці підкільця називають *тривіальними*.

З'ясовуючи, чи є дана підмножина K' кільця K підкільцем цього кільця, звичайно користуються такою теоремою (1, § 12).

Теорема 1. *Для того щоб непорожня підмножина K' кільця K була його підкільцем, необхідно й достатньо, щоб сума $a + b$, різниця $a - b$ й добуток ab будь-яких двох елементів a і b підмножини K' належали до K'.*

Нехай K і K_1 — два кільця. Кільця K й K_1 називаються *ізоморфними*, якщо існує таке взаємно однозначне відображення φ кільця K на кільце K_1, що

$$\forall_{a,b\in K}\,[\varphi(a+b)=\varphi(a)+\varphi(b)\wedge\varphi(ab)=\varphi(a)\cdot\varphi(b)].$$

Саме відображення φ з цими властивостями називають *ізоморфним відображенням*.

Кільце скалярних матриць R_n^* ізоморфне кільцю дійсних чисел $\mathbf{R}$, бо, як легко бачити, відображення φ, яке матриці $Q_a = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & a & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a \end{pmatrix}$ ставить у відповідність число a, є *ізоморфним відображенням*.

Справедливе таке твердження (1, § 14).

Теорема 2. *Якщо множина F, в якій визначені дві бінарні операції — додавання і множення, ізоморфно відображається на деяке кільце K, то множина F також є кільце відносно визначених у ній операцій.*

12.2. Кільця з одиницею. З означення кільця не випливає наявність чи відсутність в даному кільці K одиниці e. Але якщо в кільці K одиничний елемент існує, то тільки один (1, § 12). В нульовому кільці, тобто в кільці, що складається тільки з одного нуля, елемент 0 є одночасно й одиницею, оскільки $0 \cdot 0 = 0$.

Означення 1. *Ненульове кільце K, в якому є одиничний елемент e, називають кільцем з одиницею.*

Одиницю e кільця K ми часто позначатимемо також символом 1, але при цьому цей символ не слід ототожнювати з числом 1. В кільці цілих чисел, кратних довільно вибраному натуральному числу $m > 1$, одиниці немає. Зокрема, немає одиниці в кільці парних чисел. Кільце цілих чисел $\mathbf{Z}$, кільце раціональних чисел $\mathbf{Q}$, кільце дійсних чисел $\mathbf{R}$ —

кільця з одиницею. Кільцем з 1 є також кільце матриць n-го порядку над полем раціональних чисел $\mathbf{Q}$, над полем дійсних чисел $\mathbf{R}$ і над полем комплексних чисел $\mathbf{C}$. Одиницею цього кільця є одинична матриця

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Нехай K — довільне кільце з одиницею. Для будь-якого відмінного від нуля елемента $a \in K$ справедливі рівності $a \cdot 0 = 0 \cdot a = 0$ і $a \cdot e = e \cdot a = a$. Звідси випливає, що e і 0 є різні елементи кільця K, тобто $e \neq 0$. Якщо для елемента $a \in K$ в кільці K існує обернений елемент a^{-1}, то тільки один (1, § 10). Елемент e є оберненим для самого себе. З рівності $(-e) \times (-e) = e$ випливає, що елемент $-e$ також є оберненим для самого себе. Елемент 0 не має оберненого елемента, оскільки $a \cdot 0 = 0 \cdot a = 0 \neq e$ для будь-якого елемента $a \in K$. Якщо для елемента $a \in K$ в кільці K існує обернений елемент a^{-1}, то a, за означенням дільників елемента кільця, є дільником одиниці e.

Тому приймають таке означення.

Означення 2. *Елемент a, для якого в кільці K існує обернений елемент a^{-1}, називають оборотним або дільником одиниці.*

Кільце цілих чисел є найпростішим прикладом комутативного кільця, в якому тільки 1 і -1 є дільниками одиниці.

Теорема 3. *Множина K^* всіх дільників одиниці кільця K є група за множенням.*

Доведення. Справді, якщо a і b містяться в множині K^*, тобто є дільники 1, то $a^{-1}a = aa^{-1} = 1$ і $(b^{-1}a^{-1})\ ab = ab\ (b^{-1}a^{-1}) = 1$. А це означає, що a^{-1} і ab також є дільниками 1 і, отже, містяться в множині K^*. Тому, за теоремою п. 10.3, K^* є група за множенням. Теорему доведено.

Групу K^* називають *групою дільників одиничного елемента*, або, більш коротко, *групою одиниць кільця K*.

12.3. Дільники нуля. Область цілісності. Нехай K — довільне кільце. Для будь-якого елемента $a \in K$ справджується рівність $a \cdot 0 = 0 \cdot a$. Отже, за означенням дільників елементів кільця, кожен елемент a є дільником нуля. Проте в теорії кілець приймають таке означення дільників нуля:

Означення 1. *Елементи a і b кільця K називаються дільниками нуля, якщо $a \neq 0$, $b \neq 0$, а $ab = 0$; при цьому a називають лівим, а b — правим дільником нуля.*

В комутативних кільцях поняття лівого і правого дільника нуля, очевидно, збігаються. Так, якщо m є складене натуральне число, наприклад $m = p \cdot q$, то в кільці $\mathbf{Z}_m$ класів цілих чисел, конгруентних за модулем m, класи C_p і C_q відмінні від нульового класу C_0, а їх добуток дорівнює нульовому класу: $C_p \cdot C_q = C_0$. Отже, класи C_p і C_q

є дільниками нуля в кільці Z_m. Якщо $n > 2$, то матриці n-го порядку

$$A = \begin{pmatrix} 100 \ldots 00 \\ 000 \ldots 00 \\ \ldots\ldots \\ 000 \ldots 00 \end{pmatrix} \text{ і } B = \begin{pmatrix} 000 \ldots 00 \\ 000 \ldots 00 \\ \ldots\ldots \\ 000 \ldots 01 \end{pmatrix}$$

є дільниками нуля в кільці R_n.

Означення 2. *Комутативне кільце, в якому немає дільників нуля, називають областю цілісності.*

Кожне числове кільце є областю цілісності. Областю цілісності є також будь-яке поле P, оскільки

$$\underset{a,b\in P}{\forall} [a \neq 0 \wedge ab = 0 \Rightarrow a^{-1}(ab) = 0 \Rightarrow b = 0].$$

12.4. Поле часток. Областю цілісності, з якою найчастіше доводиться зустрічатися, є кільце цілих чисел **Z**. Кільце цілих чисел **Z** є підкільцем поля раціональних чисел **Q**. Цілком природно постає запитання: чи кожна область цілісності є підкільцем деякого поля. Відповідь на це запитання дає така теорема.

Теорема 4. *Для кожної області цілісності R існує поле Q, що містить, як підкільце, область цілісності R.*

Доведення. Нехай R — деяка область цілісності. Ми виключимо з розгляду той випадок, коли R складається тільки з одного нуля. Вважатимемо, що R складається з елементів $a, b, c, \ldots$. Розглянемо множину всіх можливих пар (a, b), де $a \in R$, $b \in R \setminus \{0\}$, тобто декартів добуток $R \times (R \setminus \{0\})$. У множині цих пар визначимо відношення ω:

$$(a, b)\,\omega\,(a', b') \underset{\text{df}}{\Leftrightarrow} ab' = a'b.$$

Покажемо, що ω — відношення еквівалентності. Рефлективність і симетричність відношення ω очевидна. Доведемо, що відношення ω має властивість транзитивності. Нехай $(a, b)\ \omega\ (a', b')$ і $(a', b')\ \omega\ (a'', b'')$. Тоді $ab' = a'b$ і $a'b'' = a''b'$. Помноживши обидві частини першої з цих рівностей на b'', дістаємо рівність $ab'b'' = a'bb''$. Далі підставимо у правій частині останньої рівності замість $a'b''$ добуток $a''b'$, дістанемо рівність $ab'b'' = a''b'b$. Оскільки $b' \neq 0$ і не є дільником нуля, то з останньої рівності випливає, що $ab'' = a''b$, тому $(a, b)\ \omega\ (a'', b'')$. Отже, ω — відношення еквівалентності. Як відомо, відношення еквівалентності ω визначає розбиття розглядуваної множини пар на класи еквівалентних між собою елементів, які ми називатимемо *класами еквівалентності* ω. Позначимо множину всіх класів еквівалентності ω, буквою Q, а класи еквівалентності позначатимемо малими буквами $\alpha, \beta, \gamma, \rho, \sigma, \tau$ і т. д. Кожну пару, що входить до даного класу еквівалентності ρ, називатимемо представником цього класу.

Визначимо тепер у множині Q операції додавання й множення.

Нехай (a, b) і (c, d) — довільно вибрані представники відповідно класів ρ і σ.

Сумою $\rho + \sigma$ класів ρ і σ називатимемо клас еквівалентності, що містить пару $(ad + bc, bd)$, а добутком $\rho\sigma$ — клас еквівалентності, що містить пару (ac, bd).

Операції додавання і множення класів еквівалентності ми визначили через представників цих класів. Покажемо, що визначена так сума $\rho + \sigma$ та добуток $\rho\sigma$ класів ρ і σ не залежать від вибору представників цих класів і, отже, визначаються однозначно. Для цього доведемо, що коли $(a, b)\ \omega\ (a', b')$ і $(c, d)\ \omega\ (c', d')$, то $(ad + bc, bd)\ \omega\ (a'd' + b'c', b'd')$ і $(ac, bd)\ \omega\ (a'c', b'd')$.

Справді,

$$(a, b)\,\omega\,(a', b') \wedge (c, d)\,\omega\,(c', d') \Leftrightarrow ab' = a'b \wedge cd' = c'd.$$

Але $ab' = a'b \wedge (c\ d) = (c'\ d') \Rightarrow ab' \cdot dd' = a'b \cdot dd' \wedge cd' \cdot bb' = c'd \cdot bb'$.

Додавши почастинно останні дві рівності, матимемо

$$ab'dd' + cd'bb' = ba'dd' + dc'bb',$$

тобто $(ad + bc)\,b'd' = (a'd' + b'c')\,bd$.

Отже, $(ad + bc, bd)\ \omega\ (a'd' + b'c', b'd')$, і тому $\rho + \sigma$ не залежить від вибору представників класів.

Перемноживши почастинно рівності $ab' = a'b$ і $cd' = c'd$, дістанемо $ab' \cdot cd' = a'b \cdot c'd$. Тому $(ac, bd)\ \omega\ (a'c', b'd')$ і, отже, $\rho \cdot \sigma$ не залежить від вибору представників класів.

Покажемо тепер, що множина класів еквівалентності Q з визначеними в ній операціями додавання і множення є поле.

Операції додавання і множення, визначені в множині Q, асоціативні, комутативні й пов'язані дистрибутивним законом. Доводять це безпосередньою перевіркою. Доведемо, наприклад, що операція множення дистрибутивна відносно операції додавання. Нехай ρ, σ, τ — довільно вибрані класи еквівалентності з множини Q, а (a, b), (c, d) і (k, l) — відповідно деякі представники цих класів. Тоді представником класу $(\rho + \sigma)\,\tau$ буде пара $(adk + bck, bdl)$, а представником класу $\rho\tau + \sigma\tau$ — пара $(ak \cdot dl + bl \cdot ck,\ bl \cdot dl)$. Оскільки $(adk + bck)\ bl \cdot dl = (ak \cdot dl + bl \cdot ck)\ bdl$, то, за означенням еквівалентності ω,

$$(adk + bck,\ bdl)\,\omega\,(ak \cdot dl + bl \cdot ck,\ bl \cdot dl).$$

Отже, пари $(adk + bck, bdl)$ і $(ak \cdot dl + bl \cdot ck, bl \cdot dl)$ належать до того самого класу еквівалентності і тому

$$(\rho + \sigma)\,\tau = \rho\tau + \sigma\tau.$$

Серед класів множини Q є, очевидно, клас, до якого входить пара $(0, c)$, де c — деякий відмінний від нуля елемент області цілісності R. Легко довести, що він складається з пар виду $(0, b)$, $b \neq 0$ і тільки з них. Позначимо цей клас символом o. Клас o є нульовим елементом у множині Q. Справді, нехай ρ — довільний клас з множини Q і нехай його представником є пара (a, b). Тоді представником класу $\rho + o$ буде $(ac + b0, bc)$, тобто (ac, bc). Але оскільки $(ac, bc)\ \omega\ (a, b)$, то клас $\rho + o$ збігається з класом ρ, тобто $\rho + o = \rho$. Отже, клас o є нульовим елементом множини Q.

Для будь-якого класу $\rho \in Q$ в множині Q є протилежний клас $-\rho$: якщо представником класу ρ є пара (a, b), то протилежним буде клас $-\rho$, представником якого є пара $(-a, b)$. Справді, представником класу $\rho + (-\rho)$ є пара $(ab - ab, bb)$ ω $(0, b)$ і, отже, $\rho + (-\rho) = o$. Таким чином, ми довели, що множина класів еквівалентності Q з визначеними в ній операціями додавання і множення є комутативне кільце. У кільці Q є відмінні від нуля елементи: відмінним від нульового елемента є кожен клас еквівалентності ρ, представником якого є пара (a, b), де $a \neq 0$.

Доведемо тепер, що в кільці Q здійсненна операція ділення, крім ділення на нуль. Справді, нехай ρ — довільний, а σ — будь-який відмінний від нуля елемент кільця Q і нехай представниками класів ρ і σ є відповідно пари (a, b) і (c, d). Оскільки $\sigma \neq o$, то $c \neq 0$. Клас τ, представником якого є пара (ad, bc), є часткою від ділення класу ρ на клас σ, бо (c,d) $(ad, bc) = (acd, bcd)$ ω (a, b), тому $\sigma\tau = \rho$. Отже, множина Q з визначеними в ній операціями додавання і множення є поле.

Покажемо тепер, що в полі є підкільце R', ізоморфне області цілісності R відносно операцій додавання і множення, визначених у полі Q і області цілісності R. Нехай α — клас з поля Q, що містить деяку пару (ac, c) з декартового добутку $R \times (R \setminus \{0\})$. З'ясуємо, з яких пар складається клас α. $\forall\limits_{c_1 \in R \setminus \{0\}} [(ac_1, c_1)$ ω $(ac, c)]$, оскільки $ac \cdot c = ac \cdot c_1$. Навпаки, якщо (b, c_2) ω (ac, c), то $bc = ac \cdot c_2$, тому $b = ac_2$. Звідси випливає, що клас α складається з усіх пар виду (ac, c), де a — заданий, а c — будь-який, відмінний від нуля, елемент з R. Позначимо символом R' множину всіх тих і тільки тих класів поля Q, кожен з яких складається з пар виду (ac, c), де a — фіксований, а c — будь-який, відмінний від нуля, елемент області цілісності R. Кожному класу α множини R' поставимо у відповідність елемент області цілісності R за таким правилом: якщо клас α складається з пар виду (ac, c), то в R йому відповідає елемент a. Цим визначається взаємно однозначне відображення f множини R' на область цілісності R.

Справді, оскільки клас α складається з усіх пар виду (ac, c) і тільки з них, то йому відповідає один і тільки один елемент $a \in R$: двом різним класам α і β відображення f ставить у відповідність два різні елементи a і b з R, бо якби $a = b$, то і $\alpha = \beta$; кожен елемент $d \in R$ є відповідним для деякого класу $\delta \in R'$, а саме класу, що складається з усіх пар виду (dc, c).

Покажемо тепер, що відображення $f : R \to R$ ізоморфне відносно операцій додавання і множення, визначених у полі Q і в області цілісності R. Справді, нехай α і β — будь-які класи з R' і нехай $f(\alpha) = a$, $f(\beta) = b$. Тоді клас α складається з усіх пар виду (ac, c), а клас β — з усіх пар виду (bc, c). За означенням суми і добутку класів еквівалентності $\alpha + \beta$ є клас, що містить пару виду $(ac^2 + bc^2, c^2)$, а $\alpha\beta$ — клас, що містить пару виду (abc^2, c^2). Але оскільки $(ac^2 + bc^2, c^2)$ ω $((a + b)c, c)$, а (abc^2, c^2) ω (abc, c), то клас $\alpha + \beta$ складається з усіх пар виду $((a + b)c, c)$, а клас $\alpha\beta$ — з усіх пар виду (abc, c). Тому $f(\alpha + \beta) = a + b = f(\alpha) + f(\beta)$, $f(\alpha + \beta) = f(\alpha) + f(\beta)$ і $f(\alpha\beta) = ab = f(\alpha) \cdot f(\beta)$, $f(\alpha\beta) = f(\alpha) \cdot f(\beta)$.

Таким чином, множина R' ізоморфна області цілісності R і тому, за теоремою 2, R' є підкільце поля Q.

Оскільки R — область цілісності, то й ізоморфне їй кільце R', як легко довести, також є область цілісності. Отже, у полі Q міститься, як підкільце, область цілісності R', ізоморфна області цілісності R. Але будь-які дві ізоморфні області цілісності, з точки зору визначених у них бінарних операцій, нерозрізнені. Тому елементи області цілісності R' ми ототожнимо з відповідними їм елементами області цілісності R, тобто клас α, що складається з пар виду (ac, c), ототожнимо з елементом a. Після цього можна вважати, що поле Q містить, як підкільце, область цілісності R. Теорему доведено.

Далі у тому випадку, коли в полі P міститиметься, як підкільце, область цілісності R', ізоморфна області цілісності R відносно операцій додавання і множення, визначених у полі P і в області цілісності R, ми вважатимемо, що поле P містить область цілісності R. При цьому елементи області цілісності R' ми ототожнюватимемо з відповідними елементами області цілісності R. Продовжимо вивчення побудованого нами поля Q.

Теорема 5. *Кожен елемент поля Q дорівнює частці деяких двох елементів області цілісності R.*

Доведення. Справді, нехай ρ — клас, що містить пару (a, b), а α і β — класи, що містять відповідно пари (ac, c) і (bc, c). За означенням добутку класів $\rho\beta$ є клас еквівалентності, що містить пару (abc, bc), а отже, і пару (ac, c), оскільки (abc, bc) ω (ac, c). Тому $\rho\beta = \alpha$. Звідси $\rho = \frac{\alpha}{\beta}$. Але класи α і β ми ототожнюємо відповідно з елементами a і b. Отже, $\rho = \frac{a}{b}$. Зокрема, для кожного елемента $a \in R$ маємо: $a = \frac{am}{m}$, де m — будь-який відмінний від нуля елемент з R. Теорему доведено.

Нехай P — поле, що містить деяку область цілісності R. Поле P, очевидно, містить кожну частку $\frac{a}{b}$, де a — довільний, а b — будь-який відмінний від нуля елемент області цілісності R.

Означення. *Поле P, що містить область цілісності R, кожен елемент якого може бути записаний у вигляді частки деяких двох елементів області цілісності R, називають полем часток або полем відношень області цілісності R.*

Побудоване нами поле Q, за теоремою 5, є полем часток розглядуваної області цілісності R, а теорема 4 встановлює існування поля часток для будь-якої області цілісності.

Покажемо тепер, *що поле часток Q області цілісності R з точністю до ізоморфізму однозначно визначається областю цілісності R, тобто, що будь-які два поля часток однієї й тієї області цілісності R ізоморфні.*

Ми доведемо більш сильне твердження.

Теорема 6. *Нехай R — деяка область цілісності, а Q і Q' — її поля часток. Тоді між Q і Q' існує ізоморфна відповідність φ, яка тотожно відображає область цілісності R саму на себе.*

Доведення. Насамперед нагадаємо, що для елементів будь-якого поля P (1, §13) справедливі такі твердження:

1) $\underset{\alpha,\beta,\gamma,\delta\in P}{\forall}\left[\beta\neq 0\wedge\delta\neq 0\Rightarrow\frac{\alpha}{\beta}=\frac{\gamma}{\delta}\Leftrightarrow\alpha\delta=\beta\gamma\right]$,

2) $\underset{\alpha,\beta,\gamma,\delta\in P}{\forall}\left[\beta\neq 0\wedge\delta\neq 0\Rightarrow\frac{\alpha}{\beta}+\frac{\gamma}{\delta}=\frac{\alpha\delta+\beta\gamma}{\beta\delta}\right]$,

3) $\underset{\alpha,\beta,\gamma,\delta\in P}{\forall}\left[\beta\neq 0\wedge\delta\neq 0\Rightarrow\frac{\alpha}{\beta}\cdot\frac{\gamma}{\delta}=\frac{\alpha\gamma}{\beta\delta}\right]$.

Ці твердження правильні, зокрема, й для елементів області цілісності R, розглядуваних як елементи поля Q, так і поля Q'.

Умовимося елементи $a, b, c, d, \ldots$ області цілісності R, розглядувані як елементи поля Q', позначати символами $a', b', c', d', \ldots$.

Оскільки Q і Q' є поля часток області цілісності R, то частка будь-яких двох елементів $a, b, c, d, \ldots$, якщо вона існує, належить полю Q, а частка будь-яких двох елементів $a', b', c', d', \ldots$ — полю Q' і, навпаки, кожен елемент поля Q дорівнює частці деяких двох елементів $a, b, c, d, \ldots$, а кожен елемент поля Q' — частці деяких двох елементів $a', b', c', d', \ldots$.

Перейдемо тепер до доведення теореми. Поставимо у відповідність елементу $\rho=\frac{a}{b}$ поля Q елемент $\rho'=\frac{a'}{b'}$ поля Q', причому вважатимемо, що $\rho'=\varphi(\rho)$. Цим задається деяке відображення $\varphi: Q\to Q'$. Елемент ρ' не залежить від вибору запису елемента ρ у вигляді частки елементів з R. Справді, якщо $\rho=\frac{a}{b}$ і $\rho=\frac{c}{d}$, то $\frac{a}{b}=\frac{c}{d}$. Тому, за твердженням 1) $ad=bc$. Але тоді і в полі Q' справджується рівність $a'd'=b'c'$, а отже, і рівність $\frac{a'}{b'}=\frac{c'}{d'}$. Тому елементу ρ поля Q як частці $\frac{a}{b}$ і як частці $\frac{c}{d}$ ставиться у відповідність той самий елемент ρ' поля Q'. Таким чином, кожному елементу ρ поля Q ставиться у відповідність один і тільки один елемент $\rho'=\varphi(\rho)$ поля Q'.

Аналогічно можна довести, що кожен елемент ρ' поля Q' є відповідним для одного і тільки одного елемента ρ поля Q. Отже, φ є взаємно однозначне відображення поля Q на поле Q'. Доведемо, що це відображення ізоморфне. Нехай $\rho=\frac{a}{b}$ і $\tau=\frac{c}{d}$ — будь-які елементи поля Q. Тоді $(\rho)\varphi=\rho'=\frac{a'}{b'}$, $(\tau)\varphi=\tau'=\frac{c'}{d'}$. Беручи до уваги твердження 2) і 3), матимемо:

$$(\rho+\tau)\varphi=\left(\frac{a}{b}+\frac{c}{d}\right)\varphi=\left(\frac{ad+bc}{bd}\right)\varphi=\frac{a'd'+b'c'}{b'd'}=$$
$$=\frac{a'}{b'}+\frac{c'}{d'}=\rho'+\tau'=(\rho)\varphi+(\tau)\varphi,$$

тобто $(\rho+\tau)\varphi=(\rho)\varphi+(\tau)\varphi$,

$$(\rho\tau)\varphi=\left(\frac{a}{b}\cdot\frac{c}{d}\right)\varphi=\left(\frac{ac}{bd}\right)\varphi=\frac{a'c'}{b'd'}=\frac{a'}{b'}\cdot\frac{c'}{d'}=\rho'\tau'=$$
$$=(\rho)\varphi\cdot(\tau)\varphi,$$

тобто $(\rho\cdot\tau)\varphi=(\rho)\varphi\cdot(\tau)\varphi$. А це й означає, що відображення φ є ізоморфне. Відображення φ тотожне на області цілісності R. Справді, для будь-якого $a\in R$

$$(a)\varphi=\left(\frac{am}{m}\right)\varphi=\frac{a'm'}{m'}=a',$$

тобто $(a)\varphi=a$, бо a' — це позначення елемента $a\in R$ як елемента поля Q'. Теорему доведено.

Зауважимо, що спосіб побудови поля часток, яким ми скористалися при доведенні теореми 4, загальний; його можна застосувати до будь-якої області цілісності, зокрема й в тому випадку, коли R є кільце цілих чисел. Якщо розглядувана нами область цілісності R є кільце цілих чисел, то побудоване нами поле часток Q буде поле раціональних чисел.

§ 13. ІДЕАЛИ КІЛЬЦЯ. ФАКТОР-КІЛЬЦЯ. ГОМОМОРФІЗМИ КІЛЕЦЬ

13.1. Ідеали кільця. Операції над ідеалами. В теорії кілець особливу роль, аналогічну ролі нормальних дільників у теорії груп, відіграють підкільця, що дістали назву ідеалів.

Означення 1. *Непорожня підмножина* $\mathfrak{a}$ *кільця* K *називається лівим* (відповідно правим) *ідеалом цього кільця, якщо* $\mathfrak{a}$ *є підгрупа адитивної групи кільця* K *і якщо для будь-яких елементів* $a\in\mathfrak{a}$ *і* $x\in K$ *добуток* xa (відповідно ax) *міститься в* $\mathfrak{a}$.

Підмножина $\mathfrak{a}$ *кільця* K, *яка одночасно є лівим і правим ідеалом, називається двостороннім ідеалом, або просто ідеалом кільця* K.

У комутативному кільці кожен лівий і кожен правий ідеал, очевидно, є двостороннім ідеалом.

З означення 1 випливає, що кожен лівий, правий, а отже, і двосторонній ідеал є підкільцем кільця K. Зауважимо, що оскільки ідеал — це деяка підмножина кільця K, то можна говорити про відношення включення між ідеалами даного кільця K.

Приклади ідеалів. 1. Кожне кільце K, очевидно, є своїм двостороннім ідеалом. Цей ідеал називають *одиничним*. У кожному кільці K нульове підкільце $\{0\}$ є ідеалом, його називають *нульовим ідеалом* і позначають символом $\mathfrak{o}$.

Одиничний ідеал K кільця K, очевидно, містить будь-який ідеал $\mathfrak{a}$ цього кільця, а нульовий ідеал $\mathfrak{o}$ міститься в кожному ідеалі $\mathfrak{o}$ кільця K. Отже, в смислі відношення включення одиничний ідеал — найбільший, а нульовий — найменший серед ідеалів кільця K.

2. Нехай K — деяке кільце і a — будь-який елемент цього кільця. Множина Ka всіх елементів виду xa, де x — будь-який елемент кільця K, є лівий ідеал, множина aK всіх елементів виду ax — правий ідеал, а множина $\mathfrak{m}(a)$ всіх елементів виду

$$x_1ay_1+x_2ay_2+\cdots+x_nay_n,$$

де n — будь-яке натуральне число, x_i й y_i — будь-які елементи кільця K, є двосторонній ідеал кільця K.

Покажемо, що множина Ka є лівий ідеал кільця K. Сума x_1a+x_2a будь-яких двох елементів x_1a і x_2a множини Ka належить до цієї самої множини, оскільки $x_1+x_2\in K$, і елемент $-xa=(-x)a$, протилежний будь-якому елементу $xa\in Ka$, також належить до множини Ka, оскільки $-x\in K$; тому множина Ka є

підгрупа адитивної групи кільця K. Для будь-яких елементів $xa \in Ka$ і $x' \in K$ добуток $x'(xa) = (x'x)\,a \in Ka$. Отже, Ka — лівий ідеал кільця K. Аналогічними міркуваннями доводять, що aK — правий, а $\mathfrak{m}(a)$ — двосторонній ідеал кільця K. Якщо кільце K — комутативне, то, очевидно, $Ka = aK = \mathfrak{m}(a)$. Зауважимо, що коли в кільці K немає одиниці, то кожен з ідеалів Ka, aK, $\mathfrak{m}(a)$ може не містити елемента a.

~~8.~~ Нехай K — деяке комутативне кільце і a — будь-який елемент цього кільця. Множина елементів виду $xa + na$, де x — будь-який елемент кільця K, а n — будь-яке ціле число, є ідеал кільця K. В тому, що це справді так, легко пересвідчитися за допомогою міркувань, аналогічних тим, які ми проводили в прикладі 2.

Цей ідеал називають *головним ідеалом, породженим елементом* a, і позначають символом (a). Серед ідеалів, що містять елемент a, головний ідеал (a) є найменшим (в смислі відношення включення).

Справді, кожен ідеал, який містить елемент a, містить всі кратні xa і всі суми $\pm \Sigma a = na$, а отже, і всі суми $xa + na$, тобто містить ідеал (a).

Якщо в кільці K є одиниця e, то $(a) = Ka$. Справді, з означення ідеалу (a) випливає, що $Ka \subseteq (a)$. З другого боку, $xa + na = xa + nea = (x + ne)\,a = x'a \in Ka$, тому $(a) \subseteq Ka$. Отже, $(a) = Ka$. Наприклад, головний ідеал (m) кільця цілих чисел $\mathbf{Z}$ складається з усіх цілих чисел, кратних числу m : $(m) = \mathbf{Z}m$.

Зауважимо, що нульовий ідеал $\mathfrak{o}$ кільця K є головний ідеал (0). Якщо в кільці K є одиниця e, то одиничний ідеал K також є головним ідеалом, а саме: $K = (e)$.

4. Аналогічно тому, як ми визначили поняття головного ідеалу (a) комутативного кільця K, можна визначити поняття ідеалу, породженого кількома елементами $a_1, a_2, \ldots, a_s$. Нехай K — деяке комутативне кільце і нехай $a_1, a_2, \ldots, a_s$ — будь-які елементи цього кільця. Множина елементів виду $\sum_{i=1}^{s} x_i a_i + \sum_{j=1}^{s} n_j a_j$, де x_i — будь-який елемент з кільця K, n_j — будь-яке ціле число, є ідеал кільця K. Цей ідеал позначають символом $(a_1, a_2, \ldots, a)_s$; множину елементів $a_1, a_2, \ldots, a_s$ називають *базисом ідеалу* $(a_1, a_2, \ldots, a_s)$. Звичайно, для ідеалу $(a_1, a_2, \ldots, a_s)$, крім базису $a_1, a_2, \ldots, a_s$, можуть існувати й інші базиси, причому деякі з них можуть складатися з меншого ніж s числа елементів. Перейдемо тепер до розгляду деяких операцій над ідеалами кільця K. Першою операцією, яку ми розглянемо, є операція теоретико-множинного перетину. Нехай $\mathfrak{a}$ і $\mathfrak{b}$ — будь-які ідеали кільця K.

Теорема 1. *Перетин* $\mathfrak{a} \cap \mathfrak{b}$ *ідеалів* $\mathfrak{a}$ *і* $\mathfrak{b}$ *кільця* K *є ідеал цього кільця.*

Доведення. За теоремою 8 § 10, перетин $\mathfrak{a} \cap \mathfrak{b}$ є підгрупа адитивної групи кільця K. Крім того, для будь-яких елементів $a \in \mathfrak{a} \cap \mathfrak{b}$ і $x \in K$ добутки xa і ax містяться в ідеалах $\mathfrak{a}$ і $\mathfrak{b}$, а тому містяться і в їх перетині $\mathfrak{a} \cap \mathfrak{b}$. Отже, перетин $\mathfrak{a} \cap \mathfrak{b}$ є ідеал кільця K.

Легко перевірити, що операція перетину ідеалів асоціативна і комутативна. Зауважимо, що теорема 1 легко поширюється на будь-яке скінченне чи нескінченне число ідеалів.

Нехай A і B — деякі непорожні підмножини кільця K.

Означення 2. *Множину всіх елементів виду* $a + b$, *де* $a \in A$, $b \in B$, *називають сумою підмножин* A *і* B *й позначають символом* $A + B$.

Якщо підмножина A складається тільки з одного елемента a, то суму $A + B$ позначають символом $a + B$. Оскільки операція додавання елементів кільця K асоціативна й комутативна, то й операція додавання підмножин кільця K, як легко перевірити, також асоціативна і комутативна.

Означення 3. *Добутком* AB *підмножин* A *і* B *називають множину всіх елементів виду* $\sum_{i=1}^{n} a_i b_i$, *де* n — *деяке натуральне число,* $a_i \in A$, $b_i \in B$.

Якщо підмножина A складається лише з одного елемента a, то добуток AB позначають символом aB. Цей добуток, очевидно, складається з усіх елементів виду ab, $b \in B$.

Читач самостійно легко доведе, що визначена так операція множення підмножин кільця K асоціативна. Якщо кільце K комутативне, то операція множення підмножин, очевидно, також буде комутативною.

Якщо $A_1, A_2, \ldots, A_s$ — підмножини кільця K, то добуток $A_1 \cdot A_2 \cdot \ldots \cdot A_s$ (його позначають символом $\prod_{i=1}^{s} A_i$) складається з усіх сум добутків виду $a_1 a_2 \ldots a_s$, де $a_i \in A_i$, $i = 1, 2, \ldots, s$.

Операції додавання й множення підмножин кільця можна, звичайно, застосувати до ідеалів.

Нехай $\mathfrak{a}$ і $\mathfrak{b}$ — довільні ідеали кільця K.

Теорема 2. *Сума* $\mathfrak{a} + \mathfrak{b}$ *ідеалів* $\mathfrak{a}$ *і* $\mathfrak{b}$ *кільця* K *є ідеал цього кільця.*

Доведення. Справді, сума $(a_1 + b_1) + (a_2 + b_2)$ будь-яких двох елементів $a_1 + b_1$ і $a_2 + b_2$ множини $\mathfrak{a} + \mathfrak{b}$ належить до $\mathfrak{a} + \mathfrak{b}$, оскільки $a_1 + a_2 \in \mathfrak{a}$, $b_1 + b_2 \in \mathfrak{b}$, і елемент $-(a + b) = (-a) + (-b)$, протилежний довільно вибраному елементу $a + b \in \mathfrak{a} + \mathfrak{b}$, також належить до $\mathfrak{a} + \mathfrak{b}$, бо $-a \in \mathfrak{a}$, $-b \in \mathfrak{b}$. Отже, $\mathfrak{a} + \mathfrak{b}$ є підгрупа адитивної групи кільця K. Крім того, для будь-яких елементів $a + b \in \mathfrak{a} + \mathfrak{b}$ і $x \in K$ $x(a + b) = xa + xb \in \mathfrak{a} + \mathfrak{b}$ і $(a + b)\,x = ax + bx \in \mathfrak{a} + \mathfrak{b}$. Цим теорему доведено.

Теорема 3. *Добуток* $\mathfrak{a}\,\mathfrak{b}$ *ідеалів* $\mathfrak{a}$ *і* $\mathfrak{b}$ *кільця* K *також є ідеал кільця* K.

Доведення. Справді, сума $\sum_{i=1}^{n} a_i b_i + \sum_{j=1}^{m} a_j' b_j'$ будь-яких двох елементів $\sum_{i=1}^{n} a_i b_i$ й $\sum_{j=1}^{m} a_j' b_j'$ множини $\mathfrak{a}\mathfrak{b}$ є, очевидно, елемент цієї самої множини, і елемент $-\sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n} (-a_i)\,b_i$, протилежний довільно вибраному елементу $\sum_{i=1}^{n} a_i b_i \in \mathfrak{a}\mathfrak{b}$, належить до $\mathfrak{a}\mathfrak{b}$. Крім того, для будь-яких

$$\sum_{i=1}^{n} a_i b_i \in \mathfrak{a}\mathfrak{b} \text{ і } x \in K \quad x \sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n} (x a_i)\,b_i \in \mathfrak{a}\mathfrak{b} \text{ й } \left(\sum_{i=1}^{n} a_i b_i\right) x =$$

$$= \sum_{i=1}^{n} a_i\,(b_i x) \in \mathfrak{a}\mathfrak{b}.$$

Цим теорему доведено.

Таким чином, у множині ідеалів кільця K здійсненні операції додавання й множення. Операція додавання ідеалів — асоціативна і комутативна, а операція множення — асоціативна. Якщо кільце K — комутативне, то операція множення ідеалів також комутативна.

Теорема 4. *Операції множення і додавання ідеалів кільця K пов'язані дистрибутивними законами:*

$$\forall_{a,b,c\in K}\,[(a+b)c = ac + bc \wedge c(a+b) = [a+cb].$$

Справедливість цього твердження очевидна.

13.2. Конгруенції і класи лишків за ідеалом. Фактор-кільце.

Продовжимо вивчення загальних питань теорії кілець. Нехай K — деяке кільце, а $\mathfrak{m}$ — довільний ідеал цього кільця. Ми знаємо, що K є адитивна абельова група, а ідеал $\mathfrak{m}$ — підгрупа цієї групи. Оскільки в абельовій групі всі її підгрупи є нормальними дільниками, то ідеал $\mathfrak{m}$ є нормальний дільник групи K. Отже, існує фактор-група $K/\mathfrak{m}$ групи K за нормальним дільником $\mathfrak{m}$. Вона складається з суміжних класів групи K за нормальним дільником $\mathfrak{m}$:

$$0+\mathfrak{m},\ a+\mathfrak{m},\ b+\mathfrak{m},\ c+\mathfrak{m},\ \ldots .$$

Нагадаємо, що елементи $x, y \in K$ належать до того самого суміжного класу адитивної групи K за підгрупою $\mathfrak{m}$ тоді і тільки тоді, коли $x - y \in \mathfrak{m}$. Оскільки група K — абельова, то й $K/\mathfrak{m}$ — адитивна абельова група. Ми покажемо, що в групі $K/\mathfrak{m}$ можна так означити операцію множення, що вона буде кільцем відносно визначених у ній операцій додавання і множення.

Але спочатку введемо одне досить важливе поняття — поняття *конгруенції* й вивчимо деякі його властивості.

Означення. *Вважають, що елемент $x \in K$ конгруентний елементу $y \in K$ за ідеалом $\mathfrak{m}$ або за модулем $\mathfrak{m}$, якщо $x - y \in \mathfrak{m}$, тобто якщо x і y належать до того самого суміжного класу адитивної групи K за підгрупою $\mathfrak{m}$.*

Висловлення «x конгруентно y за модулем $\mathfrak{m}$» скорочено записують так:

$$x \equiv y \pmod{\mathfrak{m}}.$$

Отже, $x \equiv y \pmod{\mathfrak{m}} \Leftrightarrow x - y \in \mathfrak{m}$.

Якщо $\mathfrak{m}$ є головний ідеал (m), то замість $x \equiv y \pmod{\mathfrak{m}}$ можна було б писати $x \equiv y \pmod{(m)}$. Проте в цьому випадку пишуть просто $x \equiv y \pmod{m}$ і говорять: x конгруентне y за модулем m. У випадку, коли x не конгруентне y за модулем $\mathfrak{m}$, пишуть $x \not\equiv y \pmod{\mathfrak{m}}$. Сформульоване вище означення визначає в множині K бінарне відношення; його називають *відношенням конгруентності*, або просто *конгруентністю*. Відношення конгруентності, як випливає з його означення, задається розбиттям адитивної групи K на суміжні класи за підгрупою $\mathfrak{m}$ і, отже, є відношенням е к в і в а л е н т н о с т і на множині K, тобто воно рефлексивне, симетричне і транзитивне:

$$\forall_{x\in K}\,[x \equiv x \pmod{\mathfrak{m}}],\quad \forall_{x,y\in K}\,[x \equiv y \pmod{\mathfrak{m}} \Rightarrow y \equiv x \pmod{\mathfrak{m}}],$$

$$\forall_{x,y,z\in K}\,[x \equiv y \pmod{\mathfrak{m}} \wedge y \equiv z \pmod{\mathfrak{m}} \Rightarrow x \equiv z \pmod{\mathfrak{m}}].$$

Класи еквівалентності відношення конгруентності в кільці K є, таким чином, суміжними класами групи K за підгрупою $\mathfrak{m}$; їх називають *класами лишків кільця K за ідеалом* $\mathfrak{m}$, *або за модулем* $\mathfrak{m}$.

Ми позначатимемо їх символами $\bar{a}, \bar{b}, \bar{c}, \ldots$.

Співвідношення вигляду $x \equiv y \pmod{\mathfrak{m}}$ називають *конгруенціями*. Спинимося на деяких властивостях конгруенцій.

1. *Обидві частини конгруенції можна помножити на будь-яке ціле число n:*

$$\forall_{x,y\in K}\ \forall_{n\in Z}\,[x \equiv y \pmod{\mathfrak{m}} \Rightarrow nx \equiv ny \pmod{\mathfrak{m}}].$$

Справді, $x \equiv y \pmod{\mathfrak{m}} \Leftrightarrow (x - y) \in \mathfrak{m}$. Тому $nx - ny = n(x - y) \in \mathfrak{m}$ і, отже, $nx \equiv ny \pmod{\mathfrak{m}}$. Зокрема, при $n = -1$ матимемо: $x \equiv y \pmod{\mathfrak{m}} \Rightarrow -x \equiv -y \pmod{\mathfrak{m}}$.

2. *До обох частин конгруенції можна додати будь-який елемент $z \in K$:*

$$\forall_{x,y,z\in K}\,[x \equiv y \pmod{\mathfrak{m}} \Rightarrow x + z \equiv y + z \pmod{\mathfrak{m}}].$$

Справді, $x \equiv y \pmod{\mathfrak{m}} \Leftrightarrow x - y \in \mathfrak{m}$. Тому $(x + z) - (y + z) = x - y \in \mathfrak{m}$ і, отже, $x + z \equiv y + z \pmod{\mathfrak{m}}$.

3. *Обидві частини конгруенції можна помножити на будь-який елемент $z \in K$:*

$$\forall_{x,y,z\in K}\,[x \equiv y \pmod{\mathfrak{m}} \Rightarrow zx \equiv zy \pmod{\mathfrak{m}} \wedge xz \equiv yz \pmod{\mathfrak{m}}].$$

Справді, $x \equiv y \pmod{\mathfrak{m}} \Leftrightarrow x - y \in m$. Тому $zx - zy = z(x - y) \in \mathfrak{m}$ і $xz - yz = (x - y)z \in \mathfrak{m}$, отже, $zx \equiv zy \pmod{\mathfrak{m}}$ і $xz \equiv yz \pmod{\mathfrak{m}}$.

4. *Конгруенції можна почленно додавати і віднімати:*

$$\forall_{x,y,x',y'\in K}\,[x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow x \pm x' \equiv y \pm y' \pmod{\mathfrak{m}}].$$

Справді, $x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow x + x' \equiv y + x' \pmod{\mathfrak{m}} \wedge y + x' \equiv y + y' \pmod{\mathfrak{m}} \Rightarrow x + x' \equiv y + y' \pmod{\mathfrak{m}}$, тобто $x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow x + x' \equiv y + y' \pmod{\mathfrak{m}}$.

Далі $x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow x \equiv y \pmod{\mathfrak{m}} \wedge -x' \equiv -y' \pmod{\mathfrak{m}} \Rightarrow x + (-x') \equiv y + (-y') \pmod{\mathfrak{m}} \Rightarrow x - x' \equiv y - y' \pmod{\mathfrak{m}}$, тобто $x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow x - x' \equiv y - y' \pmod{\mathfrak{m}}$.

5. *Конгруенції можна почленно перемножати:*

$$\forall_{x,y,x',y'\in K}\,[x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow xx' \equiv yy' \pmod{\mathfrak{m}}].$$

Справді, $x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow xx' \equiv yx' \pmod{\mathfrak{m}} \wedge yx' \equiv yy' \pmod{\mathfrak{m}} \Rightarrow xx' \equiv yy' \pmod{\mathfrak{m}}$, тобто $x \equiv y \pmod{\mathfrak{m}} \wedge x' \equiv y' \pmod{\mathfrak{m}} \Rightarrow xx' \equiv yy' \pmod{\mathfrak{m}}$.

З викладеного видно, що над конгруенціями можна виконувати всі ті операції, які виконують над рівностями, за винятком скорочення обох частин конгруенції на їх спільний дільник. Скорочувати конгруенції, взагалі кажучи, не можна. Наприклад, у кільці цілих чисел $\mathbf{Z}$ справедлива конгруенція $16 \equiv 4 \pmod{6}$, проте $4 \not\equiv 1 \pmod{6}$, бо $4 - 1 \equiv 3 \not\equiv (6)$. Отже, скорочувати обидві частини конгруенції $16 \equiv 4 \pmod{6}$ на їх спільний дільник 4 не можна.

Примітка. Співвідношення $x \equiv y \pmod{\mathfrak{m}}$ ми назвали конгруенцією. Проте символом $x \equiv y \pmod{\mathfrak{m}}$ часто позначають також і відношення конгруентності. З контексту, звичайно, буває зрозуміло, що саме означає $x \equiv y \pmod{\mathfrak{m}}$.

Повернемося тепер знову до фактор-групи $K/\mathfrak{m}$. Нагадаємо, що $K/\mathfrak{m}$ складається з класів лишків $\bar{a}, \bar{b}, \bar{c}, \ldots$. Як відомо, кожен клас a породжується будь-яким з своїх елементів: якщо $a \in \bar{a}$, то $\bar{a} = a + \mathfrak{m}$; тому будь-який з елементів класу a можна вважати представником цього класу (п. 10.6).

Додавання класів лишків (суміжних класів), як відомо, означається так: якщо $a \in \bar{a}$ і $b \in \bar{b}$, то $\bar{a} + \bar{b}$ — це той клас лишків, який містить елемент $a + b$. Інакше кажучи,

$$\bar{a} + \bar{b} = (a + \mathfrak{m}) + (b + \mathfrak{m}) = (a + b) + \mathfrak{m}.$$

Означимо тепер в $K/\mathfrak{m}$ операцію множення.

Нехай a — будь-який елемент класу $\bar{a}$, b — класу $\bar{b}$. Умовимося вважати, що $\overline{ab}$ — це клас, який містить елемент ab, тобто що $\bar{a} \cdot \bar{b} = (a + \mathfrak{m})(b + \mathfrak{m}) = ab + \mathfrak{m}$.

Покажемо, що означений так добуток класів не залежить від вибору представників цих класів. Справді, якщо a, a' — елементи з класу $\bar{a}$ і b, b' — з класу $\bar{b}$, то $a \equiv a' \pmod{\mathfrak{m}}$, $b \equiv b' \pmod{\mathfrak{m}}$ і, за властивістю 5, $ab \equiv a'b' \pmod{\mathfrak{m}}$, тобто ab і $a'b'$ належить до того самого класу. Тому $ab + \mathfrak{m} = a'b' + \mathfrak{m}$ і, отже, добуток $\overline{ab}$ не залежить від вибору представників у класах $\bar{a}$ і $\bar{b}$.

Теорема 5. *Множина $K/\mathfrak{m}$ класів лишків кільця K за ідеалом $\mathfrak{m}$ з означеними у ній операціями додавання і множення є кільце. Це кільце називають фактор-кільцем кільця K за ідеалом $\mathfrak{m}$ або за модулем $\mathfrak{m}$.*

Доведення. Множина $K/\mathfrak{m}$ є адитивна абельова група. Означена в цій групі операція множення є асоціативна і пов'язана дистрибутивними законами з операцією додавання. Справді,

$$\forall_{\bar{a},\bar{b},\bar{c} \in K/\mathfrak{m}} \{(\bar{a}\,\bar{b})\,\bar{c} = [(a + \mathfrak{m})(b + \mathfrak{m})](c + \mathfrak{m}) = abc + \mathfrak{m} =$$

$$= (a + \mathfrak{m})(bc + \mathfrak{m}) = (a + \mathfrak{m})[(b + \mathfrak{m})(c + \mathfrak{m})] = \bar{a}\,(\bar{b}\,\bar{c})\},$$

тобто

$$\forall_{\bar{a},\bar{b},\bar{c} \in K/\mathfrak{m}} [(\bar{a}\,\bar{b})\,\bar{c} = \bar{a}\,(\bar{b}\,\bar{c})]; \quad \forall_{\bar{a},\bar{b},\bar{c} \in K/\mathfrak{m}} \{(\bar{a} + \bar{b})\,\bar{c} =$$

$$= [(a + \mathfrak{m}) + (b + \mathfrak{m})](c + \mathfrak{m})\,[(a + b) + \mathfrak{m}](c + \mathfrak{m}) =$$

$$= (a + b)c + \mathfrak{m} = (ac + bc) + \mathfrak{m} = (ac + \mathfrak{m}) + (bc + \mathfrak{m}) =$$

$$= \bar{a}\,\bar{c} + \bar{b}\,\bar{c}\},$$

тобто $\forall_{\bar{a},\bar{b},\bar{c} \in K/\mathfrak{m}} [(\bar{a} + \bar{b})\,\bar{c} = \bar{a}\,\bar{c} + \bar{b}\,\bar{c}]$. Аналогічно доводять, що

$$\forall_{\bar{a},\bar{b},\bar{c} \in K/\mathfrak{m}} [\bar{c}\,(\bar{a} + \bar{b}) = \bar{c}\,\bar{a} + \bar{c}\,\bar{b}].$$

Отже, $K/\mathfrak{m}$ — кільце. Теорему доведено. Зауважимо, що фактор-кільце $K/\mathfrak{m}$ називають також *кільцем класів лишків*. Всюди далі символ $K/\mathfrak{m}$ означатиме фактор-кільце кільця K за модулем $\mathfrak{m}$.

Приклади. **1.** У кожному кільці K є одиничний ідеал K і нульовий ідеал (0). Фактор-кільце K/K є нульове кільце $\{0\}$, а фактор-кільце $K/(0)$ ізоморфне кільцю K.

2. В кільці цілих чисел $\mathbf{Z}$ візьмемо головний ідеал $\mathfrak{m} = (m)$ (m — деяке відмінне від 1 натуральне число). Цей ідеал складається з усіх цілих чисел, кратних числу m. Ідеал (m) є нульовим класом лишків: $\bar{0} = 0 + (m) = (m)$. Всі цілі числа, конгруентні за модулем m числу 1, утворюють клас лишків $\bar{1} = 1 + (m)$, конгруентні числу 2 — клас лишків $\bar{2} = 2 + (m)$, конгруентні числу $\bar{3}$ — клас лишків $\bar{3} = 3 + (m)$ і т. д.; всі числа, конгруентні за модулем m числу $m - 1$, утворюють клас лишків $\overline{m-1} = m - 1 + (m)$. Інших класів лишків бути не може, оскільки кожне ціле число належить до одного з перелічених нами класів.

Отже, фактор-кільце $\mathbf{Z}/(m)$ складається з класів лишків $\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{m-1}$. Операції додавання і множення в кільці $Z/(m)$ виконують за такими правилами: щоб додати класи $\bar{k} = k + (m)$ і $\bar{l} = l + (m)$, треба знайти суму $k + l$ представників цих класів і потім знайти лишок від ділення $k + l$ на m; якщо цим лишком є число r, то $\bar{k} + \bar{l} = \bar{r}$. Аналогічно, щоб перемножити класи $\bar{k}$ і $\bar{l}$, потрібно знайти добуток kl представників цих класів і потім знайти лишок від ділення kl на m; якщо цим лишком є число s, то $\bar{k}\bar{l} = \bar{s}$. В теорії чисел кільце $\mathbf{Z}/(m)$ називають *кільцем класів конгруентних чисел за модулем m або кільцем лишків за модулем m.*

13.3. Гомоморфізми кілець. Теорема про гомоморфізми. З викладеного в п. 11.3 ми знаємо, що між фактор-групами групи G та її гомоморфізмами існує тісний зв'язок. Виявляється, що аналогічний зв'язок існує й між фактор-кільцями даного кільця та його гомоморфізмами. До з'ясування цього зв'язку ми і перейдемо. Нехай K і K' — деякі кільця.

Означення 1. *Відображення $\varphi : K \to K'$ кільця K в кільце K' називається гомоморфним відображенням K в K', або гомоморфізмом K в K', якщо виконуються такі умови:*

1. $\forall_{a,b \in K} [\varphi(a + b) = \varphi(a) + \varphi(b)]$;
2. $\forall_{a,b \in K} [\varphi(ab) = \varphi(a)\,\varphi(b)]$.

Якщо гомоморфне відображення φ є відображенням кільця K на кільце K', то його називають *гомоморфізмом кільця K на кільце K'*, або *епіморфізмом кільця K*. У цьому випадку говорять, що кільце K' є *гомоморфним образом кільця K'*, і пишуть

$$K \sim K'.$$

Щоб зазначити, що φ є гомоморфізм кільця K на кільце K', пишуть $\varphi : K \sim K'$.

Приклади гомоморфізмів кілець. **1.** Нехай $\mathbf{C}$ — кільце всіх комплексних чисел і C_2 — кільце матриць 2-го порядку над полем комплексних чисел. Розглянемо відображення Ψ, яке визначають так:

$$\forall_{a+bi \in \mathbf{C}} \left[\psi(a + bi) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\right].$$

Очевидно, що ψ є відображення кільця C в кільце C_2. Відображення ψ задовольняє умови 1 і 2 означення гомоморфізму. Справді,

$$\underset{a+bi,c+di\in C}{\forall}\{\psi[(a+bi)+(c+di)]=\psi[(a+c)+(b+d)i]=$$

$$=\begin{pmatrix}a+c & b+d\\ -(b+d) & a+c\end{pmatrix}=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}+\begin{pmatrix}c & d\\ -d & c\end{pmatrix}=\psi(a+bi)+\psi(c+di)\},$$

тобто $\underset{a+bi,c+di\in C}{\forall}\{\psi[(a+bi)+(c+di)]=\psi(a+bi)+\psi(c+di)\}$,

$$\underset{a+bi,c+di\in C}{\forall}\Big\{\psi[(a+bi)\cdot(c+di)[=\psi[(ac-bd)+(ad+bc)i]=$$

$$=\begin{pmatrix}ac-bd & ad+bc\\ -(ad+bc) & ac-bd\end{pmatrix}=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}\begin{pmatrix}c & d\\ -d & c\end{pmatrix}=\psi(a+bi)\cdot\psi(c+di)\Big\},$$

тобто $\underset{a+bi,c+di\in C}{\forall}\{\psi[(a+bi)(c+di)]=\psi(a+bi)\psi(c+di)\}$.

Отже, ψ є гомоморфізм кільця C в кільце C_2. Цей гомоморфізм, як легко бачити, є ізоморфізмом.

2. Нехай K — множина всіх матриць виду $\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}$, де a і b — деякі дійсні числа. Читач самостійно легко доведе, що відносно операцій додавання і множення матриць множина K є кільце. Задамо відображення φ таким способом:

$$\underset{a,b\in R}{\forall}\left|\varphi\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}=a+b\sqrt{2}\right|.$$

Очевидно, що φ є відображення кільця K в кільце дійсних чисел R. Покажемо, що відображення φ задовольняє вимоги означення гомоморфізму. Справді,

$$\underset{a,b,a_1,b_1\in R}{\forall}\left\{\varphi\left[\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}+\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right]=\varphi\begin{pmatrix}a+a_1 & b+b_1\\ 2(b+b_1) & a+a_1\end{pmatrix}=(a+a_1)+\right.$$

$$\left.+(b+b_1)\sqrt{2}=(a+b\sqrt{2})+(a_1+b_1\sqrt{2})=\varphi\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}+\varphi\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right\},\ \text{тобто}$$

$$\underset{a,b,a_1,b_1\in R}{\forall}\left\{\varphi\left[\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}+\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right]=\varphi\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}+\varphi\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right\};$$

$$\underset{a,b,a_1,b_1\in R}{\forall}\left\{\varphi\left[\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right]=\varphi\begin{pmatrix}aa_1+2bb_1 & ab_1+a_1b\\ 2(ab_1+a_1b) & aa_1+2bb_1\end{pmatrix}=\right.$$

$$=(aa_1+2bb_1)+(ab_1+a_1b)\sqrt{2}=(a+b\sqrt{2})(a_1+b_1\sqrt{2})=$$

$$\left.=\varphi\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}\varphi\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right\},\ \text{тобто}$$

$$\underset{a,b,a_1,b_1\in R}{\forall}\left\{\varphi\left[\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right]=\varphi\begin{pmatrix}a & b\\ 2b & a\end{pmatrix}\varphi\begin{pmatrix}a_1 & b_1\\ 2b_1 & a_1\end{pmatrix}\right\}.$$

Таким чином, φ є гомоморфізм кільця K в кільце R.

3. Нехай K — деяке кільце і $\mathfrak{m}$ — довільний ідеал цього кільця. Розглянемо відображення $\varkappa$ кільця K в фактор-кільце $K/\mathfrak{m}$, яке задається так: кожному елементу $a\in K$ відповідає той клас лишків за модулем $\mathfrak{m}$, до якого належить елемент a, тобто клас $a+\mathfrak{m}$. Очевидно, що $\varkappa$ є відображення кільця K на кільце $K/\mathfrak{m}$. Покажемо, що $\varkappa$ задовольняє вимоги означення гомоморфізму. Справді,

$$\underset{a,b\in K}{\forall}[\varkappa(a+b)=(a+b)+\mathfrak{m}=(a+\mathfrak{m})+(b+\mathfrak{m})=\varkappa(a)+\varkappa(b)],$$

тобто $\underset{a,b\in K}{\forall}[\varkappa(a+b)=\varkappa(a)+\varkappa(b)]$;

$$\underset{a,b\in K}{\forall}[\varkappa(ab)=ab+\mathfrak{m}=(a+\mathfrak{m})(b+\mathfrak{m})=\varkappa(a)\varkappa(b)],$$

тобто $\underset{a,b\in K}{\forall}[\varkappa(ab)=\varkappa(a)\varkappa(b)]$.

Отже, $\varkappa$ є гомоморфне відображення кільця K на фактор-кільце $K/\mathfrak{m}$.

Відображення $\varkappa: K\backsim K/\mathfrak{m}$ називають *природним,* або *канонічним, гомоморфізмом.* Доведемо кілька теорем, що характеризують властивості гомоморфних відображень кілець.

Теорема 6. *Якщо φ є гомоморфізм кільця K в кільце K', то:*

1. $\varphi(0)=0'$.
2. $\underset{a\in K}{\forall}[\varphi(-a)=-\varphi(a)]$.
3. *(K) φ є підкільце кільця K'.*
4. *Якщо в кільці K є одиничний елемент e, то $\varphi(e)$ — одиничний елемент кільця $\varphi(K)$ й якщо для елемента $a\in K$ в кільці K існує обернений елемент a^{-1}, то $\varphi(a^{-1})$ — обернений елемент елементу $\varphi(a)$ кільця $\varphi(K)$.*

Доведення. Справедливість тверджень 1 і 2 випливає відповідно з теорем 7 і 8 п. 11.3, застосованих до адитивної групи кільця K. Доведемо справедливість твердження 3. За теоремою 9 п. 11.3, застосованою до адитивної групи кільця K, $\varphi(K)$ є підгрупа адитивної групи кільця K. Покажемо, що множина $\varphi(K)$ замкнена відносно операції множення, визначеної в кільці K'. Справді, якщо a', $b'\in\varphi(K)$, то $a'=\varphi(a)$, $b'=\varphi(b)$, де a, $b\in K$, і $a'b'=\varphi(a)\varphi(b)=\varphi(ab)\in$ $\in\varphi(K)$. Отже, за теоремою 1 п. 12.1 $\varphi(K)$ є підкільце кільця K'. Справедливість твердження 4 доводиться дослівним повторенням міркувань, які ми проводили при доведенні теорем 7 і 8 п. 11.3. Теорему доведено.

Означення 2. *Нехай φ є гомоморфізм кільця K в кільце K'. Множину $\mathfrak{A}$ всіх елементів кільця K, які гомоморфізмом φ відображаються в $0'$ кільця K', називають ядром гомоморфізму φ і записують $\mathfrak{A}=\operatorname{Ker}\varphi$.*

Теорема 7. *Ядро $\mathfrak{A}=\operatorname{Ker}\varphi$ будь-якого гомоморфізму φ кільця K в кільце K' є ідеал кільця K.*

Доведення. Справді, за теоремою 10 п. 11.3, застосованою до адитивної групи кільця K, $\mathfrak{A}=\operatorname{Ker}\varphi$ є підгрупа адитивної групи кільця K. Крім того, для будь-якого елемента $x\in K$, $\mathfrak{A}x\in\mathfrak{A}$ і $x\mathfrak{A}\in$ $\in\mathfrak{A}$, оскільки для кожного $a\in\mathfrak{A}$ $\varphi(ax)=\varphi(a)\varphi(x)=0'$ і $\varphi(xa)=$ $=\varphi(x)\varphi(a)=0'$. Отже, $\mathfrak{A}=\operatorname{Ker}\varphi$ — ідеал кільця K. Теорему доведено.

Теорема про гомоморфізми кілець. *Якщо φ є гомоморфізм кільця K на кільце K' і $\mathfrak{A}=\operatorname{Ker}\varphi$, то кільце K' ізоморфне фактор-кільцю $K/\mathfrak{A}$, причому існує такий ізоморфізм ψ кільця $K/\mathfrak{A}$ на кільце K', що добуток $\varkappa\psi$ природного гомоморфізму $\varkappa$: $K\backsim K/\mathfrak{A}$ на ізоморфізм ψ є гомоморфізм φ.*

Доведення. Нехай r' — довільно вибраний елемент кільця K' і r — деякий елемент кільця K такий, що $\varphi(r)=r'$. Тоді $\underset{b\in K}{\forall}[b\equiv$

$\equiv r \pmod{\mathfrak{A}} \to \varphi(b) = \varphi(r) + \varphi(b - r) = \varphi(r) = r'$], оскільки $b - r \in \mathfrak{A}$ і тому $\varphi(b - r) = 0'$. З другого боку, якщо елемент $c \in K$ при гомоморфізмі φ відображається в елемент r', тобто $\varphi(c) = r'$, то $\varphi(c - r) = \varphi(c) - \varphi(r) = 0'$ і тому $c - r \in \mathfrak{A}$, тобто $c \equiv r \pmod{\mathfrak{A}}$. Отже, множина всіх елементів кільця K, які при гомоморфізмі φ відображаються в елемент $r' \in K'$, є клас лишків кільця K за модулем $\mathfrak{A}$, до якого належить елемент r, тобто клас $\bar{r} = r + \mathfrak{A}$. Позначимо символом ψ відображення, яке кожному класу лишків $\bar{r} = r + \mathfrak{A}$ ставить у відповідність елемент $r' \in K'$, у який при гомоморфізмі φ відображаються елементи класу $\bar{r}$, тобто $\psi(\bar{r}) = \varphi(r)$, де r — будь-який елемент з класу лишків $\bar{r}$. Очевидно, що ψ є відображення фактор-кільця $K/\mathfrak{A}$ на кільце K'. Відображення ψ є гомоморфне. Справді, нехай $\bar{r}_1 = r_1 + \mathfrak{A}$, $\bar{r}_2 = r_2 + \mathfrak{A}$ — довільно вибрані елементи кільця $K/\mathfrak{A}$.

Тоді $\bar{r}_1 + \bar{r}_2 = (r_1 + r_2) + \mathfrak{A}$, $\bar{r}_1\bar{r}_2 = r_1r_2 + \mathfrak{A}$ і тому

$$\psi(\bar{r}_1 + \bar{r}_2) = \varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2) = \psi(\bar{r}_1) + \psi(\bar{r}_2),$$

$$\psi(\bar{r}_1 \cdot \bar{r}_2) = \varphi(r_1r_2) = \varphi(r_1) \cdot \varphi(r_2) = \psi(\bar{r}_1)\,\psi(\bar{r}_2),$$

тобто

$$\psi(\bar{r}_1 + \bar{r}_2) = \psi(\bar{r}_1) + \psi(\bar{r}_2),\ \psi(\bar{r}_1\bar{r}_2) = \psi(\bar{r}_1)\,\psi(\bar{r}_2).$$

Доведемо тепер, що відображення ψ взаємно однозначне, тобто що коли $\bar{r}_1 \neq \bar{r}_2$, то $\psi(\bar{r}_1) \neq \psi(\bar{r}_2)$. Це справді так. Якщо $\psi(\bar{r}_1) = \psi(\bar{r}_2)$, то за означенням відображення ψ, $\varphi(r_1) = \varphi(r_2)$, тобто $\varphi(\bar{r}_1) - \varphi(\bar{r}_2) = 0'$. Тому $\varphi(r_1 - r_2) = \varphi(r_1) - \varphi(r_2) = 0'$ і $r_1 - r_2 \in \mathfrak{A}$. Звідси $r_1 \equiv r_2 \pmod{\mathfrak{A}}$ і $\bar{r}_1 = \bar{r}_2$. Отже, ψ є ізоморфне відображення $K/\mathfrak{A}$ на K'.

Розглянемо тепер відображення $\varkappa\psi$. Оскільки $\varkappa$ — природний гомоморфізм кільця K на фактор-кільце $K/\mathfrak{A}$, а ψ — ізоморфізм кільця $K/\mathfrak{A}$ на кільце K', то $\varkappa\psi$, очевидно, є відображення кільця K на кільце K'. Доведемо тепер, що $\varkappa\psi = \varphi$. Нехай r — довільний елемент кільця K. За означенням природного гомоморфізму $\varkappa$, $x(r) = \bar{r} = r + \mathfrak{A}$ і за означенням ізоморфізму ψ, $\psi(\bar{r}) = \varphi(r)$.

Отже, $\varkappa\psi(r) = \psi[\varkappa(r)] = \psi(\bar{r}) = \varphi(r)$, тобто $\varkappa\psi(r) = \varphi(r)$. Таким чином, $\forall\limits_{r \in K}[\varkappa\psi(r) = \varphi(r)]$. А це й означає, що $\varkappa\psi = \varphi$. Теорему доведено. Теорема про гомоморфізми кілець показує, що природними гомоморфізмами кільця K на його фактор-кільця по суті вичерпуються всі його гомоморфізми.

13.4. Характеристика кільця з одиницею. На закінчення цього параграфа коротко спинимося на понятті характеристики кільця з одиницею. З'ясуємо насамперед, які ідеали є в найпростішому з кілець — кільці цілих чисел. Як відомо, кожне ціле число n породжує головний ідеал $(n) = \mathbf{Z}n$. Такими ідеалами вичерпується множина всіх ідеалів кільця $\mathbf{Z}$, оскільки справедлива така теорема.

Теорема 8. *Кожен ідеал кільця цілих чисел є головним ідеалом.*

Доведення. Нехай $\mathfrak{A}$ — деякий ідеал кільця $\mathbf{Z}$. Якщо $\mathfrak{A}$ — нульовий ідеал, то $\mathfrak{A} = (0)$. Якщо ж в ідеалі $\mathfrak{A}$ міститься число $c \neq 0$, то в ньому міститься також і число $-c$. Одне з чисел c і $-c$ додатне, тому в ідеалі $\mathfrak{A}$ містяться натуральні числа. Нехай a — найменше з натуральних чисел, що містяться в $\mathfrak{A}$. Тоді $\forall\limits_{n \in \mathbf{Z}}[na \in \mathfrak{A}]$ і, отже, $(a) = \mathbf{Z}a \subset \mathfrak{A}$. Покажемо, що і, навпаки, $\mathfrak{A} \subset (a)$.

Справді, нехай b — довільне число з ідеалу $\mathfrak{A}$. Поділивши b на $\mathfrak{A}$, дістанемо $b = aq + r$, $0 \leqslant r < a$. Оскільки $a \in \mathfrak{A}$, $b \in \mathfrak{A}$, то і $r = b - aq \in \mathfrak{A}$. Звідси й з умови $0 \leqslant r < a$ випливає, що $r = 0$, бо в противному разі a не було б найменшим серед натуральних чисел, що містяться в ідеалі $\mathfrak{A}$. Таким чином, $b = aq$; тому $b \in (a)$, а отже, і $\mathfrak{A} \subset (a)$. Оскільки $(a) \subset \mathfrak{A}$ і $\mathfrak{A} \subset (a)$, то $\mathfrak{A} = (a)$. Теорему доведено.

Нехай тепер A — довільне кільце з одиницею e і $\mathbf{Z}$ — кільце цілих чисел. Розглянемо відображення $\varphi : \mathbf{Z} \to A$, яке задається так: $\forall\limits_{n \in \mathbf{Z}}[\varphi(n) = ne]$.

Очевидно, що φ є гомоморфізм кільця $\mathbf{Z}$ в кільце A. За теоремою 6 множина $E = \varphi(\mathbf{Z})$ є підкільце кільця A. Підкільце E складається з усіх цілих кратних ne одиничного елемента e; будемо називати його *підкільцем, породженим одиницею* e. Як легко бачити, φ є гомоморфізм кільця $\mathbf{Z}$ на підкільце E. Тому, за теоремою про гомоморфізми кілець, підкільце E ізоморфне фактор-кільцю $\mathbf{Z}/\mathfrak{A}$, де $\mathfrak{A}$ — ядро гомоморфізму φ: $\mathbf{Z} \sim E$. Оскільки в кільці цілих чисел $\mathbf{Z}$ кожен ідеал головний, то $\mathfrak{A} = (p)$, де p — деяке невід'ємне число.

Можливі два випадки:

1. $p = 0$. Тоді $E \cong \mathbf{Z}/(0) \cong \mathbf{Z}$, тобто підкільце E ізоморфне кільцю цілих чисел $\mathbf{Z}$.

2. $p > 0$. Тоді підкільце E ізоморфне кільцю класів лишків $\mathbf{Z}/(p)$.

Отже, в будь-якому кільці A з одиницею e підкільце E, породжене елементом e, ізоморфне або кільцю цілих чисел $\mathbf{Z}$, або кільцю класів лишків $\mathbf{Z}/(p)$, де p — деяке натуральне число. Виходячи з цього, ми введемо таке означення.

Означення 1. *Нехай A — деяке кільце з одиницею e. Ми будемо говорити, що кільце A має характеристику* 0, *якщо його підкільце E, породжене одиничним елементом e, ізоморфне кільцю цілих чисел $\mathbf{Z}$; ми говоритимемо, що кільце A має характеристику $p > 0$, якщо підкільце E ізоморфне кільцю класів лишків $\mathbf{Z}/(p)$.*

Примітка. Зауважимо, що поряд з висловом «кільце A має характеристику 0 (або p)» вживають також вирази «характеристика кільця A дорівнює 0 (або p)», «A є кільце характеристики 0 (або p)».

За означенням 1, кільце цілих чисел має характеристику 0, а кільце $\mathbf{Z}/(p)$ — характеристику p.

Теорема 9. *Якщо кільце A має характеристику* 0, *то $ne = 0$ лише при $n = 0$; якщо ж A має характеристику $p > 0$, то $pe = 0$ і немає такого натурального числа $m < p$, що $me = 0$.*

Доведення. Нехай A — кільце характеристики 0. Тоді існує ізоморфне відображення φ підкільця $E \subset A$ на кільце $\mathbf{Z}$. Оскільки за теоремою 6, при гомоморфізмі (зокрема, ізоморфізмі) кільця E на

кільце $\mathbf{Z}$ $e \to 1$, то $\forall_{n\in\mathbf{Z}} [\varphi(ne) = n]$. Тому $ne = 0$ лише при $n = 0$, бо якщо $ne = 0$, то $\varphi(ne) = n$, за теоремою 6, дорівнює 0, тобто $n = 0$.

Припустимо тепер, що A — кільце характеристики p. Тоді існує ізоморфне відображення ψ кільця $\mathbf{Z}/(p)$ на підкільце E. За теоремою 6, $\psi(\bar{1}) = e$, де $\bar{1} = 1 + (p)$, і тому $\psi(n \cdot \bar{1}) = ne$. Оскільки $p \cdot \bar{1} = \bar{0}$ і, за теоремою 6, $\varphi(\bar{0}) = 0$, то $pe = 0$. Якщо $m < p$, то $m \cdot \bar{1} = \bar{m} \neq \bar{0}$, а тому і $me = \psi(m \cdot \bar{1}) \neq 0$, бо в противному разі відображення ψ було б гомоморфізмом, а не ізоморфізмом. Цим теорему доведено. Справедлива також і обернена теорема.

Теорема 10. *Якщо в кільці A з одиницею e рівність $ne = 0$ справджується лише при $n = 0$, то A має характеристику 0; якщо в кільці A справджується рівність $pe = 0$ і немає такого натурального $m < p$, що $me = 0$, то A має характеристику p.*

Доведення. Розглянемо гомоморфізм φ: $\mathbf{Z} \sim E$, для якого $\forall_{n\in\mathbf{Z}} [\varphi(n) = ne]$ за теоремою про гомоморфізми кілець, $E \simeq \mathbf{Z}/\mathfrak{A}$, де $\mathfrak{A}$ — ядро гомоморфізму φ: $\mathbf{Z} \sim E$. Якщо в A рівність $ne = 0$ справджується лише при $n = 0$, то при гомоморфізмі φ: $\mathbf{Z} \sim E$ в нуль кільця E відображається лише 0 кільця $\mathbf{Z}$; тому $\mathfrak{A} = (0)$ і, отже, $E \simeq \mathbf{Z}/(0) \simeq \mathbf{Z}$, тобто $E \simeq \mathbf{Z}$. Якщо в кільці A справджується рівність $pe = 0$ і немає такого натурального $m < p$, що $me = 0$, то при гомоморфізмі φ: $\mathbf{Z} \sim E$ в нуль кільця E відображаються всі цілі кратні np числа p і тільки вони; тому $\mathfrak{A} = (p)$ і, отже, $E \simeq \mathbf{Z}/(p)$. Теорему доведено.

З теорем 9 і 10 випливає таке означення.

Означення 2. *Характеристикою кільця A з одиницею e називають число 0, якщо $ne = 0$ лише при $n = 0$; характеристикою кільця A називають натуральне число p, якщо $pe = 0$ і немає такого натурального числа $m < p$, що $me = 0$.*

Всі числові кільця з одиницею, очевидно, мають характеристику 0. Кожне скінченне кільце A з одиницею e є кільце ненульової характеристики. Справді, якщо кільце A скінченне, то серед усіх цілих додатних кратних одиничного елемента e обов'язково будуть кратні, рівні між собою, бо в противному разі кільце A було б нескінченним. Нехай $ke = me$, де k і m — деякі натуральні числа, причому $m > k$. Тоді $(m - k)e = 0$ і, отже, A є кільце ненульової характеристики.

Кожне натуральне число n є характеристикою деякого кільця з одиницею: n є характеристикою кільця $\mathbf{Z}/(n)$. Доведемо тепер дві теореми, які характеризують властивості кілець характеристики 0 і характеристики p.

Теорема 11. *Якщо R є область цілісності характеристики 0, то*

$$\forall_{a\in R} \forall_{n\in\mathbf{Z}} [a \neq 0 \wedge n \neq 0 \to na \neq 0].$$

Доведення. Нехай a — довільно вибраний, відмінний від 0, елемент з R і n — будь-яке натуральне число. Тоді $na = \underbrace{a + a + \ldots + a}_{n \text{ доданків}} = a\underbrace{(e + e + \ldots + e)}_{n \text{ доданків}} = a(ne)$.

Припустимо, що $na = 0$, тоді й $a(ne) = 0$. Оскільки в R немає дільників нуля і за умовою теореми $a \neq 0$, то з рівності $a(ne) = 0$ випливає, що $ne = 0$, чого не може бути. Отже, припущення, що $na = 0$, неправильне. Таким чином, для будь-якого натурального n маємо $na \neq 0$. При будь-якому цілому від'ємному n також $na \neq 0$, бо якби елемент na $(n < 0)$ кільця R дорівнював нулю, то й протилежний йому елемент $(-n)a$ також дорівнював би нулю, чого за доведеним вище не може бути. Теорему доведено.

Теорема 12. *Якщо A — кільце характеристики p, то* $\forall_{a\in A} [pa = 0]$.

Справді,

$$\forall_{a\in A} [pa = \underbrace{a + a + \ldots + a}_{p \text{ доданків}} = a\underbrace{(e + e + \ldots + e)}_{p \text{ доданків}} = a(pe) = a \cdot 0 = 0].$$

§ 14. КІЛЬЦЯ ГОЛОВНИХ ІДЕАЛІВ ТА ЕВКЛІДОВІ КІЛЬЦЯ

14.1. Подільність в області цілісності. В теорії кілець особливої уваги заслуговують кільця, які за своїми властивостями досить близькі до кільця цілих чисел. Зокрема, для цих кілець можна розвинути теорію подільності, аналогічну теорії подільності цілих чисел. Ці кільця дістали назву *кілець головних ідеалів*. Вивченням їх ми і будемо займатись. Але спочатку викладемо деякі загальні відомості, що стосуються подільності в області цілісності з одиницею.

Нехай R — область цілісності з одиницею. Оскільки область цілісності — комутативне кільце, то в ній поняття правого і лівого дільника елемента збігаються і тому означення подільності формулюється так:

Означення 1. *Якщо для елементів a і b області цілісності R в R існує такий елемент c, що $a = bc$, то говорять, що a ділиться на b або b ділить a і пишуть відповідно $a \vdots b$; b/a* або $a \equiv 0 \pmod{b}$.

Як бачимо, означення 1 є поширенням на область цілісності означення подільності в кільці цілих чисел, яке є конкретним прикладом області цілісності.

З означення 1 випливають такі властивості подільності в області цілісності:

1. $\forall_{a,b,c\in R} [a \vdots b \wedge b \vdots c \Rightarrow a \vdots c]$.
2. $\forall_{a,b,c\in R} [a \vdots c \wedge b \vdots c \Rightarrow (a + b) \vdots c \wedge (a - b) \vdots c]$.
3. $\forall_{a,b,c\in R} [a \vdots b \Rightarrow ac \vdots b]$.
4. $\forall_{a_1,b_1,a_2,b_2,\ldots,a_n,b_n c\in R} [a_1 \vdots c \wedge a_2 \vdots c \wedge \ldots \wedge a_n \vdots c \Rightarrow$

$\Rightarrow (a_1b_1 + a_2b_2 + \ldots + a_nb_n) \vdots c]$.

Ці властивості, як легко бачити, є поширенням на область цілісності відповідних властивостей подільності в кільці цілих чисел.

5. *Кожен елемент $a \in R$ ділиться на будь-який дільник ε одиниці e.* Справді, $a = \varepsilon(\varepsilon^{-1}a)$ і, отже, ε/a.

6. *Якщо $a \in R$ ділиться на $b \in R$, то a ділиться і на $b\varepsilon$, де ε — будь-який дільник одиниці.*

Справді, з рівності $a = bc$ випливає рівність $a = b\varepsilon(\varepsilon^{-1}c)$ і, отже, $b\varepsilon/a$.

7. *Кожен з дільників одного з елементів $a \in R$ і $a\varepsilon \in R$, де ε — будь-який дільник одиниці, є дільником і іншого.*

Справді, з рівності $a = cg$ випливає рівність $a\varepsilon = c(\varepsilon g)$, а з рівності $a\varepsilon = cq$ — рівність $a = c(\varepsilon^{-1}q)$. Отже, якщо c/a, то $c/a\varepsilon$, і навпаки.

Всюди далі будемо розглядати елементи області цілісності R, відмінні від нуля.

Означення 2. *Елементи a і b області цілісності R називаються асоційованими, якщо кожен з них є дільником іншого:*

$$a = bc, \quad b = ad. \tag{1}$$

З рівностей (1) випливає, що $a = a(cd)$. Звідси, скоротивши обидві частини рівності на $a \neq 0$, дістаємо $cd = 1$. Отже, c і d є дільники одиниці. Таким чином, якщо a і b — асоційовані елементи, то $b = a\varepsilon$, де ε — деякий дільник одиниці. З другого боку, який би ми не взяли дільник одиниці ε, елементи a і $a\varepsilon$ асоційовані між собою, оскільки $a = (a\varepsilon)\varepsilon^{-1}$.

Означення 2'. *Елементи a і b області цілісності R називаються асоційованими, якщо $b = a\varepsilon$, де ε — деякий дільник одиниці.*

В кільці цілих чисел, наприклад, асоційованими є кожні два числа m і $-m$.

Якщо a і b — асоційовані елементи, тобто $a = bc$ і $b = ad$, то $(a) \subseteq (b)$ і $(b) \subseteq (a)$ і, отже, $(a) = (b)$.

Таким чином, *два асоційовані елементи a і b породжують той самий головний ідеал.*

Нехай a і b — довільні елементи області цілісності R.

Означення 3. *Елемент $c \in R$ називається спільним дільником елементів a і b, якщо кожен з цих елементів ділиться на c.* За властивістю 5, всі дільники одиниці e області цілісності R є спільними дільниками елементів a і b. Але в елементів a і b можуть бути й інші спільні дільники. Ми хочемо ввести поняття найбільшого спільного дільника цих елементів. Означення НСД двох цілих чисел, за яким найбільшим спільним дільником називають найбільший із спільних дільників, поширити на область цілісності не можна, оскільки в довільній області цілісності R немає відношення порядку. Проте ми знаємо й інше означення НСД двох чисел, а саме: НСД двох чисел називають такий спільний дільник цих чисел, який ділиться на будь-який інший їхній спільний дільник. Саме це означення ми й поширимо на область цілісності.

Означення 4. *Найбільшим спільним дільником елементів a і b області цілісності R називається такий спільний дільник цих елементів, який ділиться на будь-який інший їхній спільний дільник.*

Щоб зазначити, що d є найбільший спільний дільник елементів a і b, пишуть $d = (a, b)$.

Якщо також $d' = (a, b)$, то елементи d і d' діляться один на одного і, отже, вони асоційовані. З другого боку, якщо $d = (a, b)$ і ε — будь-який дільник одиниці, то, очевидно, $d\varepsilon = (a, b)$. Як бачимо, найбільший спільний дільник елементів a і b визначається з точністю до множника ε, що є дільником одиниці.

Означення 5. *Елементи a, $b \in R$ називаються взаємно простими, якщо вони не мають спільних дільників, відмінних від дільників одиниці, тобто якщо $(a, b) = 1$.*

Нехай ε — будь-який дільник одиниці і a — довільний елемент області цілісності R. Тоді $a = a\varepsilon \cdot \varepsilon^{-1}$. З цієї рівності випливає, що всі елементи, асоційовані з елементом a, і всі дільники одиниці є дільниками елемента a. Їх називають *тривіальними*, або *невласними*, *дільниками* елемента a. Всі інші дільники елемента a, тобто дільники, відмінні від $a\varepsilon$ і ε, якщо такі існують, називають *нетривіальними*, або *власними*. Так, в кільці цілих чисел $\mathbf{Z}$ тривіальними дільниками числа 10 є числа ± 1, ± 10 і нетривіальними — числа ± 2, ± 5.

Означення 6. *Елемент $a \in R$ називається нерозкладним, або простим, якщо він не є дільником одиниці й не має нетривіальних дільників; елемент $a \in R$ називається розкладним, або складеним, якщо він має нетривіальні дільники.*

Інакше кажучи, елемент $a \in R$ називається розкладним, якщо його можна записати у вигляді добутку $a = bc$ двох нетривіальних множників b і c; він називається нерозкладним, якщо його не можна записати у вигляді добутку двох нетривіальних дільників, тобто якщо з $a = bc$ завжди випливає, що один з множників b і c є дільник одиниці, а інший — асоційований з a. Так, у кільці цілих чисел $\mathbf{Z}$ нерозкладними є числа $\pm 2, \pm 3, \pm 5, \ldots$ (тобто числа прості й протилежні простим); всі інші числа, відмінні від ± 1,— розкладні.

Наведемо такі дві властивості нерозкладних елементів.

1. *Якщо елемент $p \in R$ нерозкладний, то і будь-який асоційований з ним елемент $p\varepsilon$ також нерозкладний.* Ця властивість випливає з властивості 7 подільності елементів області цілісності R.

2. *Якщо a — будь-який, а p — нерозкладний елемент з R, то або a ділиться на p, або a і p — взаємно прості.*

Справді, якщо $(a, p) = d$, то d, як дільник нерозкладного елемента p, або є деякий дільник ε одиниці, або елемент вигляду $p\varepsilon$. У першому випадку a і p взаємно прості, в другому — a ділиться на p.

14.2. Кільце головних ідеалів. Перейдемо тепер до вивчення кілець головних ідеалів.

Означення. *Кільцем головних ідеалів називається область цілісності з одиницею, в якій кожен ідеал є головний.*

Найпростішим прикладом кілець головних ідеалів є кільце цілих чисел $\mathbf{Z}$: кільце $\mathbf{Z}$, як відомо, є область цілісності з 1 і, за теоремою 8 § 13, кожен його ідеал головний.

Кожне поле P є кільце головних ідеалів. Справді, поле P є областю цілісності з одиницею; якщо $\mathfrak{A}$ є ненульовий ідеал поля P, то разом з будь-яким своїм елементом $a \neq 0$ він містить і елемент $aa^{-1} = 1$ і,

отже, $\mathfrak{A} = (1)$. Кільцем головних ідеалів є також кільце многочленів від змінної x з коефіцієнтами з поля P, яке буде об'єктом вивчення в розділі V.

Звичайно, не кожна область цілісності з одиницею є кільцем головних ідеалів. Нижче ми наведемо приклади таких областей цілісності. А тепер займемося вивченням властивостей кілець головних ідеалів. Всюди далі вважатимемо, що R — кільце головних ідеалів.

Теорема 1. *Будь-які два елементи a і b кільця головних ідеалів R мають найбільший спільний дільник d, причому $d = ra + sb$, де r і s — деякі елементи кільця R.*

Доведення. Якщо один з елементів a і b дорівнює нулю, то справедливість теореми очевидна. Нехай a і b — будь-які відмінні від нуля елементи кільця R. Вони породжують ідеал (a, b), який складається з усіх елементів вигляду $xa + yb$, де x і y — будь-які елементи кільця R. Оскільки R — кільце головних ідеалів, то ідеал (a, b) є головний, тобто породжується деяким елементом $d \in R$: $(a, b) = (d)$.

Тому

$$d = ra + sb \quad (r, s \in R), \tag{2}$$

$$a = gd, \; b = hd \quad (g, h \in R). \tag{3}$$

З рівностей (3) випливає, що d є спільний дільник елементів a і b; з рівності ж (2) випливає, що d ділиться на будь-який спільний дільник елементів a і b. Отже, $d = (a, b)$. Теорему доведено.

Спираючись на теорему 1, доведемо твердження, яке є критерієм взаємної простоти двох елементів кільця головних ідеалів.

Теорема 2. *Елементи a і b кільця головних ідеалів R взаємно прості тоді і тільки тоді, коли в кільці R є такі елементи r і s, що $ra + sb = 1$.*

Доведення. Необхідність умови очевидна: якщо a і b — взаємно прості, тобто $(a, b) = 1$, то, за теоремою 1, в кільці R існують такі елементи r і s, що $ra + sb = 1$. Доведемо достатність умови. Припустимо, що в кільці R існують такі елементи r і s, що $ra + sb = 1$.

З цієї рівності випливає, що спільними дільниками елементів a і b можуть бути лише дільники одиниці і, отже, елементи a і b взаємопрості. Теорему доведено.

Теорема 3. *Якщо елемент $a \in R$ взаємно простий з кожним із елементів $b \in R$ і $c \in R$, то він взаємно простий і з добутком цих елементів.*

Доведення. Оскільки a і b — взаємно прості, то, за теоремою 2, існують такі $r, s \in R$, що

$$ra + sb = 1.$$

Помноживши цю рівність на c, дістаємо: $a(rc) + (bc)s = c$. З цієї рівності випливає, що кожен спільний дільник елементів a і bc буде дільником і елемента c. Але за умовою теореми спільними дільниками елементів a і c є лише дільники одиниці, тому і спільними дільниками a і bc будуть лише дільники одиниці й, отже, a і bc взаємно прості.

Теорема 4. *Якщо добуток елементів $a \in R$ і $b \in R$ ділиться на елемент $c \in R$, але a і c взаємно прості, то b ділиться на c.*

Доведення. Оскільки a і c — взаємно прості, то в кільці R існують такі r і s, що

$$ra + sc = 1.$$

Помноживши цю рівність на b, дістаємо:

$$(ab)r + c(bs) = b.$$

Обидва доданки лівої частини останньої рівності діляться на c, а тому і права її частина b ділиться на c.

Теорема 5. *Якщо елемент $a \in R$ ділиться на кожен з елементів $b \in R$ і $c \in R$, які між собою взаємно прості, то a ділиться і на добуток bc.*

Доведення. Справді, за умовою теореми, $a \vdots b$, тобто $a = bg$. Оскільки $a \vdots c$, то $bg \vdots c$. Але b і c взаємно прості, тому, за теоремою 4, $g \vdots c$, тобто $g = cq$.

Отже, $a = (bc)q$, тобто $a \vdots bc$.

Теорема 6. *Якщо R — кільце головних ідеалів і p — простий елемент цього кільця, то фактор-кільце $R/(p)$ є поле.*

Доведення. Одиничний елемент $\bar{1} = 1 + (p)$ кільця $R/(p)$ відмінний від $\bar{0} = (p)$. Справді, якби $\bar{1} = \bar{0}$, то елемент 1 містився б в ідеалі (p) і тому $p/1$. Але елемент p не може бути дільником одиниці, оскільки він нерозкладний. Отже, в кільці $R/(p)$ є принаймні один відмінний від нуля елемент.

Покажемо, що в кільці $R/(p)$ здійсненна операція ділення, крім ділення на нуль, тобто що для будь-яких елементів $\bar{a} = a + (p) \neq \bar{0}$ і $\bar{b} = b + (p)$ кільця $R/(p)$ рівняння $\bar{a} \cdot \bar{x} = \bar{b}$ має в цьому кільці розв'язок. Справді, оскільки $\bar{a} \neq \bar{0}$, то $a \not\equiv 0 \pmod p$, тобто a не ділиться на p. Отже, за другою властивістю нерозкладних елементів, елементи a і p — взаємно прості, тобто $(a, p) = 1$. Тому, за теоремою 2, в кільці R існують такі елементи r і s, що $ar + ps = 1$.

Звідси

$$arb + psb = b, \; arb \equiv b \pmod p,$$

і, отже, $\bar{a} \cdot \overline{rb} = \bar{b}$. Таким чином, $\bar{x} = \overline{rb}$ є розв'язком рівняння $\bar{a}\bar{x} = \bar{b}$. Теорему доведено.

Наслідок. *Якщо добуток кількох елементів кільця головних ідеалів R ділиться на простий елемент $p \in R$, то принаймні один із співмножників ділиться на p.*

Доведення. Припустимо, що добуток $a_1 \cdot a_2 \cdot \ldots \cdot a_s$ ($a_i \in R$) ділиться на нерозкладний елемент $p \in R$, тобто що $a_1 a_2 \ldots a_s \in (p)$. Розглянемо елементи $\bar{a}_i = a_i + (p)$ $(i = 1, 2, \ldots, s)$ і $\overline{a_1 a_2 \ldots a_s} = \overline{a_1 a_2 \ldots a_s + (p)}$. За означенням операції множення в кільці $R/(p)$, $\overline{a_1 a_2 \ldots a_s} = \bar{a}_1 \bar{a}_2 \ldots \bar{a}_s$. Оскільки $a_1 a_2 \ldots a_s \in (p)$, то $\overline{a_1 a_2 \ldots a_s} = \bar{0}$ і, отже, $\bar{a}_1 \bar{a}_2 \ldots \bar{a}_s = \bar{0}$. Звідси, оскільки, за теоремою 6, $R/(p)$ є поле, випливає, що для деякого m $(1 \leqslant m \leqslant s)$ $\bar{a}_m = \bar{0}$. Але $\bar{a}_m = \bar{0}$ означає, що $a_m \in (p)$, тобто що $a_m \vdots p$. Цим справедливість наслідку доведено.

Нашою метою буде тепер доведення твердження про можливість розкладу кожного елемента кільця головних ідеалів у добуток простих (нерозкладних) множників. Воно грунтується на такій лемі.

Лема. *В кільці головних ідеалів R не існує нескінченної строго зростаючої послідовності ідеалів*

$$\mathfrak{A}_0 \subset \mathfrak{A}_1 \subset \mathfrak{A}_2 \subset \cdots \subset \mathfrak{A}_n \subset \cdots . \tag{4}$$

Доведення. Припустимо, що нескінченна строго зростаюча послідовність (4) існує. Позначимо символом $\mathfrak{b}$ об'єднання всіх ідеалів послідовності (4). Множина $\mathfrak{b}$ є ідеал кільця R. Справді, якщо $a \in \mathfrak{b}$ і $b \in \mathfrak{b}$, то a є елемент деякого ідеала $\mathfrak{A}_s$ і b — деякого ідеала $\mathfrak{A}_l$. Тому a і b є елементи ідеала $\mathfrak{A}_m$, де m — більший з індексів s і l. Отже, $(a+b) \in \mathfrak{A}_m \subset \mathfrak{b}$, $(a-b) \in \mathfrak{A}_m \subset \mathfrak{b}$ і для будь-якого $r \in R$ $ar \in \mathfrak{A}_m \subset \mathfrak{b}$. Оскільки R — кільце головних ідеалів, то ідеал $\mathfrak{b}$ головний. Нехай $\mathfrak{b} = (b)$. Елемент b, як елемент об'єднання ідеалів послідовності (4), належить до деякого ідеалу $\mathfrak{A}_k$, а отже, і до кожного ідеалу $\mathfrak{A}_i$ при $i \geqslant k$.

Тому $(b) = \mathfrak{A}_k = \mathfrak{A}_{k+1} = \mathfrak{A}_{k+2} = \ldots$. А це суперечить нашому припущенню. Лему доведено.

Теорема 7. *В кільці головних ідеалів R кожен відмінний від нуля елемент, що не є дільником одиниці, розкладається в добуток простих множників.*

Доведення. Для кожного простого елемента кільця R теорема справедлива: для простого елемента добуток, про який говориться в теоремі, складається з одного множника. Припустимо, що в кільці R є відмінний від нуля елемент a, який не можна розкласти в добуток простих множників. Елемент a не простий і, отже, $a = a_1 \cdot a_2$, де a_1 і a_2 — нетривіальні дільники елемента a.

Принаймні один з елементів a_1 і a_2 не можна розкласти в добуток простих множників, бо в противному разі і елемент a розкладався б у добуток простих множників. Не втрачаючи загальності міркувань, припустимо, що a_1 не можна розкласти в добуток простих множників. Тоді $a_1 = a_{11} \cdot a_{12}$, де a_{11} і a_{12} — нетривіальні дільники. Принаймні один з елементів a_{11} і a_{12} також не можна розкласти в добуток простих множників. Нехай цим елементом є a_{11}. Для елемента a_{11} міркування повторимо і т. д. Цей процес послідовного розкладу, очевидно, не може обірватися. Таким чином, ми дістанемо нескінченну послідовність елементів

$$a, a_1, a_{11}, a_{111}, \ldots , \tag{5}$$

у якій кожен наступний член є власним дільником попереднього.

Якщо a_{i+1} є власним дільником a_i, то $(a_{i+1}) \supset (a_i)$, оскільки $a_i = a_{i+1} \cdot r$, де r — деякий елемент R. Тому головні ідеали, породжені елементами послідовності (5), утворюють нескінченну строго зростаючу послідовність ідеалів

$$(a) \subset (a_1) \subset (a_{11}) \subset (a_{111}) \subset \cdots ,$$

а це суперечить доведеній вище лемі. Отже, наше припущення неправильне. Теорему доведено.

Покажемо тепер, що розклад, про який іде мова в теоремі 7, однозначний з точністю до порядку співмножників і до дільників одиниці.

Теорема 8. *Якщо*

$$a = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s$$

є два розклади елемента a кільця головних ідеалів R в добуток простих множників, то $r = s$ і, при відповідній нумерації співмножників, справджуються рівності $q_i = \varepsilon_i p_i$ $(i = 1, 2, \ldots, r)$, де ε_i — деякий дільник одиниці кільця R.

Доведення. Доводитимемо індукцією по r. При $r = 1$ справедливість твердження очевидна. Справді, оскільки елемент $a = p_1$ простий, то добуток $q_1 q_2 \ldots q_s$ може містити лише один множник $q_1 = p_1$. Припустимо, що теорема правильна для $r - 1$ $(2 \leqslant r)$, і доведемо, що в такому разі теорема справедлива й для r. Справді, оскільки $a = p_1 p_2 \ldots p_r$ і $a = q_1 q_2 \ldots q_s$, то

$$p_1 p_2 p_3 \ldots p_r = q_1 q_2 q_3 \cdots q_s. \tag{6}$$

З рівності (6) випливає, що $q_1 q_2 q_3 \ldots q_s$ ділиться на p_1. Тому, за наслідком з теореми 6, принаймні один із співмножників $q_1, q_2, q_3, \ldots, q_s$ ділиться на p_1. Ми вважатимемо, що на p_1 ділиться множник q_1: цього завжди можна досягти зміною нумерації множників $q_1, q_2, \ldots, q_s$. Оскільки q_1 — простий елемент і ділиться на простий елемент p_1, то $q_1 = \varepsilon_1 p_1$, де ε_1 — деякий дільник одиниці кільця R. Підставивши в рівність (6) $\varepsilon_1 p_1$ замість q_1 і скоротивши обидві частини одержаної рівності на p_1, дістанемо:

$$p_2 p_3 \ldots p_r = (\varepsilon_1 q_2) q_3 \cdots q_s.$$

Але, за індуктивним припущенням, $r - 1 = s - 1$ і при відповідній нумерації множників $q_2, q_3, \ldots, q_r$:

$$q_2' = \varepsilon_1 q_2 = \varepsilon_2' p_2, \quad q_3 = \varepsilon_3 p_3, \quad \ldots , \quad q_r = \varepsilon_r p_r,$$

де $\varepsilon_2', \varepsilon_3, \ldots, \varepsilon_r$ — деякі дільники одиниці кільця R. Тому $r = s$ і при відповідній нумерації множників $q_1, q_2, \ldots, q_r$:

$$q_1 = \varepsilon_1 p_1, \quad q_2 = \varepsilon_1^{-1} \varepsilon_2' p_2 = \varepsilon_2 p_2, \quad q_3 = \varepsilon_3 p_3, \quad \ldots , \quad q_r = \varepsilon_r p_r.$$

Теорему доведено.

Зауважимо, що теореми 7 і 8 справедливі, зокрема, для кільця цілих чисел, яке є кільцем головних ідеалів.

Постає запитання: чи не можна теореми 7 і 8 поширити на клас областей цілісності більш широкий, ніж кільце головних ідеалів? Відповідь на це запитання в загальному випадку негативна. Є області цілісності, в яких не справджується теорема про розклад елементів області цілісності в добутки простих множників, а також області цілісності, в яких розклад елементів на прості множники хоч і можливий, але не однозначний. Наведемо приклади таких областей цілісності, не вивчаючи її докладно.

Нехай K — множина всіх дійсних чисел виду

$$c = a_1 2^{r_1} + a_2 2^{r_2} + \cdots + a_n 2^{r_n},$$

де n — будь-яке натуральне число, $a_1, a_2, \ldots, a_n$ — будь-які цілі числа й $r_1, r_2, \ldots, r_n$ — будь-які числа виду $\frac{m}{2^k}$ (m, k — цілі невід'ємні числа). Сума, різниця й добуток чисел такого виду — числа такого самого виду. Отже, K — кільце. При $n = 1$ і $r_1 = 0$ дістанемо $c = a_1$; тому K містить усі цілі числа, зокрема 1. Легко бачити, що кільце K є область цілісності. У цій області цілісності число 2 розкладається на множники так:

$$2 = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{2}} = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{4}} \cdot 2^{\frac{1}{4}} = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{4}} \cdot 2^{\frac{1}{8}} \cdot 2^{\frac{1}{8}} = \ldots.$$

Можна довести, що числа виду $2^{\frac{1}{2^k}}$, де k — ціле невід'ємне число, не є дільниками одиниці в кільці K. Таким чином, число 2 не можна розкласти на прості множники в кільці K.

Нехай тепер Q — множина всіх комплексних чисел виду $z = a + bi\sqrt{5}$, де a і b — будь-які цілі числа. Сума, різниця й добуток чисел такого виду є, очевидно, числа такого самого виду. Отже, Q — кільце. При $b = 0$, $z = a$, а тому в Q містяться всі цілі числа. Отже, кільце Q є область цілісності. Можна довести, що в цій області цілісності кожне число розкладається на прості множники. Проте не можна стверджувати, що для цього кільця характерна однозначність розкладу на прості множники. Для числа 6, наприклад, у цьому кільці існують такі два розклади: $6 = 2 \cdot 3$ і $6 = (1 + i\sqrt{5})(1 - i\sqrt{5})$.

Поряд з цим існують області цілісності, які не є кільцями головних ідеалів, проте в них справджуються теореми 7 і 8.

14.3. Евклідові кільця. Порівняно з кільцями головних ідеалів більш близькими до кільця цілих чисел за своїми властивостями є кільця, в яких справедлива теорема, що є аналогом теореми про ділення з остачею в кільці цілих чисел. Ці кільця називають *евклідовими*. Вони означаються так:

Означення. *Область цілісності R з одиницею називається евклідовим кільцем, якщо існує відображення φ множини відмінних від 0 елементів цієї області цілісності в множину цілих невід'ємних чисел N^0, тобто φ: $R \setminus \{0\} \to N^0$, яке задовольняє таку вимогу: для будь-яких елементів a, $b \in R$, $b \neq 0$ в R існують такі елементи q і r, що $a = bq + r$, причому або $r = 0$, або $\varphi(r) < \varphi(b)$.* Кільце цілих чисел $\mathbf{Z}$ — евклідове; відображення φ, про яке йде мова в означенні, задається так: $\forall[\varphi(a) = |a|]$. Евклідовим також є кільце многочленів від невідомого x з коефіцієнтами з поля P. $a \in \mathbf{Z}, a \neq 0$

Теорема 9. *Кожне евклідове кільце R є кільцем головних ідеалів.*

Доведення. Нехай $\mathfrak{A}$ — довільний ідеал евклідового кільця R. Якщо $\mathfrak{A}$ — нульовий ідеал, то $\mathfrak{A} = (0)$. Припустимо, що ідеал $\mathfrak{A}$ — відмінний від нульового. Тоді в $\mathfrak{A}$ є елементи, відмінні від нуля. Серед відмінних від нуля елементів ідеалу $\mathfrak{A}$, очевидно, є такий елемент a_0, що $\varphi(a_0) \leqslant \varphi(a)$ для будь-якого ненульового елемента $a \in \mathfrak{A}$. За означенням евклідового кільця, для будь-якого елемента $a \in \mathfrak{A}$ в кільці R існують такі елементи q і r, що $a = a_0 q + r$, причому, якщо $r \neq 0$, то $\varphi(r) < \varphi(a_0)$. Але оскільки $r = a - a_0 q \in \mathfrak{A}$, то можливість $r \neq 0$ виключається і тому $r = 0$. Таким чином, $a = a_0 q$ і, отже, $\mathfrak{A}$ є головний ідеал, породжений елементом a_0. Теорему доведено.

Оскільки кожне евклідове кільце є кільцем головних ідеалів, то для елементів будь-якого евклідового кільця справедливі теореми 7 і 8. Зауважимо, що твердження, обернене твердженню 9, неправильне: існують кільця головних ідеалів, які не є евклідовими.

В п. 14.2 було доведено існування найбільшого спільного дільника для будь-яких двох елементів a і b кільця головних ідеалів R. Але там нічого не говориться про те, як же відшукати цей найбільший спільний дільник. Методу, який би давав змогу відшукати найбільший спільний дільник будь-яких двох елементів a і b довільного кільця головних ідеалів R, не існує. В евклідових же кільцях його можна відшукати за допомогою відомого вже читачеві алгоритму Евкліда. Справді, нехай a_0 і a_1 — будь-які відмінні від нуля елементи евклідового кільця R і нехай $\varphi(a_0) \geqslant \varphi(a_1)$. Тоді, за означенням евклідового кільця, в R існують такі елементи q_1 і a_2, що $a_0 = a_1 q_1 + a_2$, причому або $a_2 = 0$, або $\varphi(a_1) > \varphi(a_2)$. Якщо $a_2 \neq 0$, то в R існують такі елементи q_2 і a_3, що $a_1 = a_2 q_2 + a_3$, причому або $a_3 = 0$, або $\varphi(a_2) > \varphi(a_3)$. Якщо $a_3 \neq 0$, то в R існують такі елементи q_3 і a_4, що $a_2 = a_3 q_3 + a_4$ і т. д.

Оскільки $\varphi(a_1) > \varphi(a_2) > \varphi(a_3) > \ldots > \varphi(a_{s-1}) > \varphi(a_s) > \ldots$, то цей процес послідовного ділення не може продовжуватись нескінченно: в противному разі множина цілих невід'ємних чисел $\varphi(a_1) > \varphi(a_2) > \ldots > \varphi(a_s) > \ldots$ не мала б найменшого числа. Отже, через кілька кроків ми дійдемо до ділення з остачею нуль: $a_{m-1} = a_m q_m$. Таким чином, ми матимемо рівності

$$\begin{aligned}
a_0 &= a_1 q_1 + a_2,\\
a_1 &= a_2 q_2 + a_3,\\
a_2 &= a_3 q_3 + a_4,\\
&\cdots\cdots\cdots\\
a_{m-3} &= a_{m-2} q_{m-2} + a_{m-1},\\
a_{m-2} &= a_{m-1} q_{m-1} + a_m,\\
a_{m-1} &= a_m q_{m+1}.
\end{aligned}$$

Остання рівність означає, що a_m є дільником a_{m-1}. Оскільки кожен з доданків правої частини передостанньої рівності ділиться на a_m, то і її ліва частина ділиться на a_m, тобто a_m є дільником a_{m-2}. Аналогічними міркуваннями ми доведемо, що a_m є дільником $a_{m-3}, a_{m-4}, \ldots, a_4, a_3, a_2, a_1, a_0$. Отже, a_m є спільним дільником елементів a_0 і a_1. Покажемо тепер, що a_m ділиться на будь-який спільний дільник елементів a_0 і a_1. Нехай b — довільно вибраний спільний дільник a_0 і a_1. Тоді з рівності $a_0 = a_1 q_1 + a_2$ випливає, що a_2 ділиться на b, з рівності $a_1 = a_2 q_2 + a_3$ випливає, що a_3 ділиться на b і т. д. Нарешті, з рівності $a_{m-2} = a_{m-1} q_{m-1} + a_m$ випливає, що a_m ділиться на b. Таким чином, елемент a_m є спільним дільником елементів a_0 і a_1 і ділиться на будь-який спільний дільник цих елементів, тобто a_m *є найбільшим спільним дільником елементів a_0 і a_1*.

Розділ IV

ТЕОРІЯ КОНГРУЕНЦІЙ І СТЕПЕНЕВІ ЛИШКИ

§ 15. КОНГРУЕНЦІЇ В КІЛЬЦІ ЦІЛИХ ЧИСЕЛ

15.1. Властивості конгруенцій за даним модулем. В п. 13.2 вже було введено відношення конгруентності в будь-якому кільці K за ідеалом $\mathfrak{m}$, або за модулем $\mathfrak{m}$, і розглянуто деякі його властивості. Розглянемо тепер це поняття в кільці $\mathbf{Z}$ цілих чисел — комутативному кільці з одиницею.

Спочатку сформулюємо деякі означення і властивості стосовно цілих чисел.

Означення 1. *Числа a і b називаються конгруентними за модулем m, якщо остачі при діленні їх на число m рівні між собою, тобто $a = mq + r$, $b = mq_1 + r$ і $0 \leqslant r < m$.*

Записують це, як було домовлено в п. 13.2, так:

$$a \equiv b \pmod m.$$

Якщо розглядається кілька чисел, конгруентних між собою за тим самим модулем m, то роблять такий запис:

$$a \equiv b \equiv c \equiv d \pmod m,$$

наприклад, $2 \equiv 5 \equiv 8 \pmod 3$.

Теорема 1. *Для того щоб числа a і b були конгруентні за модулем m, необхідно і достатньо, щоб різниця $a - b$ ділилася на m, або що те ж саме, $a = b + mt$, де t — довільне ціле число.*

Доведення. *Необхідність.*

$$[a \equiv b \pmod m] \Rightarrow [a = mq + r \wedge b = mq_1 + r] \Rightarrow [a - b = m(q - q_1)],$$

тобто $(a - b) \vdots m$, а позначаючи $q - q_1 = t$, дістанемо $a = b + mt$.

Достатність.

$$[(a - b) \vdots m \wedge a = mq + r] \Rightarrow [a = b + mt \wedge a = mq + r] \Rightarrow$$

$$\Rightarrow [b = m(q - t) + r] \Rightarrow [b = mq_1 + r].$$

Отже, число b, як і число a, при діленні на m має остачу, рівну r, тобто $a \equiv b \pmod m$. У зв'язку з тим, що конгруенції за теоремою 1 тісно пов'язані з рівностями, вони мають багато властивостей, аналогічних властивостям рівностей. Властивості конгруенцій можна умовно поділити на дві групи: властивості при незмінному модулі і властивості при змінному модулі.

Розглянемо спочатку конгруенції при незмінному модулі. В п. 13.2 властивості 1, 2, 3, 6, 7 вже доводилися в загальному вигляді. Проте, щоб зберегти елементарність викладу, наведемо доведення цих властивостей конгруенцій і в кільці цілих чисел.

1. *Конгруенції за тим самим модулем можна почленно додавати.*

Справді, $[a_1 \equiv b_1 \pmod m \wedge \ldots \wedge a_k \equiv b_k \pmod m] \Rightarrow [a_1 = b_1 + mt_1 \wedge \ldots \wedge a_k = b_k + mt_k] \Rightarrow [a_1 + \ldots + a_k = (b_1 + \ldots + b_k) + m(t_1 + \ldots + t_k)] \Rightarrow [a_1 + \ldots + a_k \equiv b_1 + \ldots + b_k \pmod m]$.

2. *Конгруенції за тим самим модулем можна почленно віднімати.*

Справді, $[a_1 \equiv b_1 \pmod m \wedge a_2 \equiv b_2 \pmod m] \Rightarrow [a_1 = b_1 + mt_1 \wedge a_2 = b_2 + mt_2] \Rightarrow [a_1 - a_2 = b_1 - b_2 + m(t_1 - t_2)] \Rightarrow [a_1 - a_2 \equiv b_1 - b_2 \pmod m]$.

3. *До обох частин конгруенції можна додати будь-яке ціле число, тобто з конгруенції $a \equiv b \pmod m$ випливає $a + c \equiv b + c \pmod m$, де c — довільне ціле число.*

Справді, на основі рефлексивності конгруентності число c конгруентне з самим собою за будь-яким модулем, у тому числі і за модулем m. Тому, використовуючи властивість 1, маємо

$$[a \equiv b \pmod m \wedge c \equiv c \pmod m] \Rightarrow [a + c \equiv b + c \pmod m].$$

4. *З однієї частини конгруенції до другої її частини можна переносити доданок з протилежним знаком, тобто з $a + b \equiv c \pmod m$ випливає $a \equiv c - b \pmod m$.*

Використовуючи рефлексивність конгруентності і властивість 1, дістаємо

$$[a + b \equiv c \pmod m \wedge -b \equiv -b \pmod m] \Rightarrow [a \equiv c - b \pmod m].$$

5. *До будь-якої частини конгруенції можна додати або відняти довільне ціле число, кратне модулю, тобто з конгруенції $a \equiv b \pmod m$ випливає $a + km \equiv b \pmod m$ або $a \equiv b + km \pmod m$.* Оскільки $km \equiv 0 \pmod m$, то

$$[a \equiv b \pmod m \wedge km \equiv 0 \pmod m] \Rightarrow [a + km \equiv b \pmod m],$$

або

$$[a \equiv b \pmod m \wedge 0 \equiv km \pmod m] \Rightarrow [a \equiv b + km \pmod m],$$

що й треба було довести.

6. *Конгруенції за одним модулем можна почленно перемножати.*

Справді, $[a_1 \equiv b_1 \pmod m \wedge \ldots \wedge a_k \equiv b_k \pmod m] \Rightarrow [a_1 = b_1 + mt_1 \wedge \ldots \wedge a_k = b_k + mt_k] \Rightarrow [a_1 a_2 \ldots a_k = b_1 b_2 \ldots b_k + mt] \Rightarrow [a_1 a_2 \ldots a_k \equiv b_1 b_2 \ldots b_k \pmod m]$, де записом mt позначено суму всіх доданків із спільним множником m в добутку $(b_1 + mt_1) \cdot \ldots \cdot (b_k + mt_k)$. Властивість доведено.

Наслідок. *Конгруенцію можна піднести до будь-якого натурального степеня n.*

Справді, з властивості 6 при умові $a_1 = a_2 = \ldots = a_k = a$, $b_1 = b_2 = \ldots = b_k = b$ випливає, що $a^n \equiv b^n \pmod m$.

7. *Обидві частини конгруенції можна помножити на те саме ціле число, тобто при $a \equiv b \pmod m$ і k цілому справедлива конгруенція $ak \equiv bk \pmod m$.*

Справді, на основі 6 і рефлексивності конгруентності маємо

$$[a \equiv b \pmod m) \wedge k \equiv k \pmod m)] \Rightarrow [ak \equiv bk \pmod m)].$$

8. *Обидві частини конгруенції можна поділити на їх спільний дільник d, якщо він взаємно простий з модулем m.*

Справді, якщо $d = (a, b)$, тобто $a = a_1 d$, $b = b_1 d$, то

$$[a \equiv b \pmod m)] \Rightarrow [a = b + mt] \Rightarrow [a_1 d = b_1 d + mt] \Rightarrow [(a_1 - b_1) d = mt].$$

Права частина рівності ділиться на m. Тому на m ділиться і ліва частина. Оскільки $(m, d) = 1$, то $(a_1 - b_1) \vdots m$, тобто $a_1 - b_1 = ms$ і $a_1 = b_1 + ms$. Отже, $a_1 \equiv b_1 \pmod m$, що й треба було довести.

9. *Якщо у виразі*

$$f(a_1, a_2, \ldots, a_k) = \Sigma A a_1^{n_1} a_2^{n_2} \ldots a_k^{n_k}$$

всі коефіцієнти A і числа $a_1, a_2, \ldots, a_k$ замінити на конгруентні їм за модулем m коефіцієнти B і числа $b_1, b_2, \ldots, b_k$ відповідно, то новий вираз

$$g(b_1, b_2, \ldots, b_k) = \Sigma B b_1^{n_1} b_2^{n_2} \ldots b_n^{n_k}$$

буде конгруентний за модулем m до заданого

$$f(a_1, a_2, \ldots, a_k) \equiv g(b_1, b_2, \ldots, b_k) \pmod m.$$

Справді, з конгруентності $a_i \equiv b_i \pmod m$ при всіх $i = 1, 2, \ldots, k$ випливає $a_i^{n_i} \equiv b_i^{n_i} \pmod m$. Враховуючи також, що $A \equiv B \pmod m$, маємо за властивістю 6

$$A a_1^{n_1} a_2^{n_2} \ldots a_k^{n_k} \equiv B b_1^{n_1} b_2^{n_2} \ldots b_k^{n_k} \pmod m$$

і далі за 1, додаючи всі аналогічні конгруенції, дістанемо

$$\Sigma A a_1^{n_1} a_2^{n_2} \ldots a_k^{n_k} \equiv \Sigma B b_1^{n_1} b_2^{n_2} \ldots b_k^{n_k} \pmod m,$$

або

$$f(a_1, a_2, \ldots, a_k) \equiv g(b_1, b_2, \ldots, b_k) \pmod m,$$

що й треба було довести.

Наслідок 1. *Якщо $a_i \equiv b_i \pmod m$, $i = 1, 2, \ldots, k$, то*

$$f(a_1, a_2, \ldots, a_k) \equiv f(b_1, b_2, \ldots, b_k) \pmod m.$$

Це окремий випадок теореми і доведення не потребує.

Наслідок 2. *Якщо в многочлені з цілими коефіцієнтами*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0, \qquad (1)$$

заданому на множині цілих чисел $\mathbf{Z}$, всі коефіцієнти a_i замінити коефіцієнтами b_i, конгруентними з a_i за модулем m, то дістанемо многочлен

$$g(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots + b_1 x + b_0,$$

конгруентний з многочленом $f(x)$, тобто

$$\forall_{x \in \mathbf{Z}} [f(x) \equiv g(x) \pmod m].$$

Справді, $\forall_k [a_k \equiv b_k \pmod m) \wedge x \equiv x \pmod m)] \Rightarrow \forall_k [a_k x^k \equiv b_k x^k \pmod m)] \Rightarrow$

$$\Rightarrow \left[\sum_{k=0}^{n} a_k x^k \equiv \sum_{k=0}^{n} b_k x^k \pmod m\right], \text{ тобто } f(x) \equiv g(x) \pmod m.$$

Наслідок 3. *Якщо $x \equiv y \pmod m$, то для многочлена (1) справедлива конгруенція*

$$f(x) \equiv f(y) \pmod m.$$

Це безпосередньо випливає з наслідку 1.

15.2. Властивості конгруенцій за різними модулями.

10. *Обидві частини конгруенції і модуль можна множити на те саме ціле число.*

Справді, $[a \equiv b \pmod m)] \Rightarrow [a = b + mt] \Rightarrow [ak = bk + (mk)\, t] \Rightarrow$
$\Rightarrow [ak \equiv bk \pmod{mk}]$.

11. *Обидві частини конгруенції і модуль можна скоротити на спільний дільник.*

Справді, $[ak \equiv bk \pmod{mk})] \Rightarrow [ak = bk + (mk)\, t] \Rightarrow [a = b + mt] \Rightarrow [a \equiv b \pmod m)]$.

12. *Якщо конгруенція $a \equiv b$ має місце за кількома модулями, то вона має місце і за модулем, який дорівнює спільному найменшому кратному цих модулів.*

Припустимо, що

$$\begin{aligned} &a \equiv b \pmod{m_1}, \\ &a \equiv b \pmod{m_2}, \\ &\ldots\ldots\ldots \\ &a \equiv b \pmod{m_k} \end{aligned} \qquad (2)$$

і $m \equiv [m_1, m_2, \ldots, m_k]$ — найменше спільне кратне чисел $m_1, m_2, \ldots, m_k$. З конгруенцій (2) випливає, що різниця чисел $a - b$ ділиться на числа $m_1, m_2, \ldots, m_k$. Але в цьому випадку, як відомо з другого розділу, вона повинна ділитися і на їх найменше спільне кратне m. Отже, $a - b = mt$, тобто $a \equiv b \pmod m)$, що й треба було довести.

13. *Якщо конгруенція має місце за модулем m, число d — дільник m, то вона має місце і за модулем d.*

Справді, якщо $m = dm_1$, то $[a \equiv b \pmod m) \wedge m = dm_1] \Rightarrow [a = b + mt \wedge m = dm_1] \Rightarrow [a = b + dm_1 t] \Rightarrow [a = b + d\,(m_1 t)] \Rightarrow$
$\Rightarrow [a \equiv b \pmod d)]$.

14. *Якщо одна частина конгруенції і модуль діляться на число d, то й друга частина конгруенції ділиться на це число.*

Справді, з конгруенції $a \equiv b \pmod m)$ випливає рівність

$$a = b + mt.$$

Нехай $b \vdots d$ і $m \vdots d$. Тоді права частина рівності ділиться на d і, отже, $a \vdots d$. Якщо $a \vdots d$ і $m \vdots d$, то робимо аналогічний висновок, що $b \vdots d$.

15. *Якщо $a \equiv b \pmod m)$, то НСД чисел a, m і b, m рівні між собою: $(a, m) = (b, m)$.*

Справедливість цього твердження випливає з рівності $a = b + mt$. Справді, a ділиться на будь-який спільний дільник d чисел b і m і тому d є спільним дільником чисел a і m. Навпаки, число b ділиться на

будь-який спільний дільник d чисел a і m, тому d є спільним дільником чисел b і m. Таким чином, спільні дільники чисел a і m є ті ж самі, що й чисел b і m. Зокрема, повинні збігатися і найбільші спільні дільники, тобто $(a, m) = (b, m)$.

15.3. Класи чисел за даним модулем. В п. 13.2 вже зазначалося, що відношення конгруентності ділить кільце K на класи чисел, конгруентних між собою за даним модулем — класи лишків. *Лишком класу за модулем m називається будь-яке число цього класу.* Кільце цілих чисел $\mathbf{Z}$ за модулем m розпадається на m класів, лишків, кожний з яких породжується будь-яким числом цього класу. До класу лишків, який містить число a, за теоремою 1 п. 15.1 належать усі цілі числа x виду $x = a + mt$, де t — будь-яке ціле число. Цей клас ми позначали символом $\bar{a}$. Позначатимемо далі цей клас символом $K_a^{(m)}$. Очевидно, що його можна позначити і символом $K_{a+mt}^{(m)}$, бо число $a + mt \in K_a^{(m)}$. Іншими словами, справедлива така рівність

$$K_a^{(m)} = K_{a+mt}^{(m)} \tag{3}$$

при будь-якому цілому значенні t.

Оскільки відношення конгруентності є відношенням еквівалентності, то різні класи лишків за модулем m не мають спільних елементів. Тим самим можна стверджувати, що *два класи лишків збігаються, якщо вони мають хоч один спільний елемент.* Доведемо таку теорему.

Теорема 2. *Кожний клас лишків $K_a^{(m)}$ за модулем m розпадається на d ($d \geqslant 1$) класів лишків за модулем dm а саме:*

$$\{K_a^{(dm)}, K_{a+m}^{(dm)}, K_{a+2m}^{(dm)}, \ldots, K_{a+(d-1)m}^{(dm)}\}.$$

Доведення. У класі $K_a^{(m)}$ містяться всі числа x, конгруентні з a за модулем m, тобто числа виду $x = a + mt$. Зокрема, він містить d таких чисел:

$$x_0 = a,\ x_1 = a + m,\ x_2 = a + 2m,\ \ldots,\ x_p = a + pm,\ \ldots,\ x_{d-1} = a + (d-1)m. \tag{4}$$

Зауважимо, що за модулем dm ці числа не конгруентні. Справді, різниця будь-яких з них не ділиться на dm, бо є меншою за число dm. Навіть найбільша з можливих різниць цих чисел — різниця крайніх чисел $x_{d-1} - x_0 = (d-1)m$ — менша за dm. Отже, за модулем dm числа (4) належать різним класам. З другого боку, легко встановити, що будь-яке число $x = a + mt$ з класу $K_a^{(m)}$ конгруентне з одним із чисел (4).

Нехай число t конгруентне з деяким числом p за модулем d ($0 \leqslant p < d$), тобто $t = d\tau + p$. Покажемо тоді, що число $x = a + mt \in K_a^{(m)}$ і число $x_p = a + pm$ із (4) конгруентні між собою за модулем dm. Це випливає з того, що їх різниця

$$x - x_p = (a + mt) - (a + pm) = m(t - p) = md\tau = (md)\tau,$$

і тому $(x - x_p) \vdots dm$. Отже, $x \equiv x_p \pmod{dm}$.

Отже, клас $K_a^{(m)}$ лишків за модулем m розпадається на такі класи лишків за модулем dm:

$$\{K_a^{(dm)}, K_{a+m}^{(dm)}, K_{a+2m}^{(dm)}, \ldots, K_{a+(d-1)m}^{(dm)}\}.$$

Теорему доведено.

15.4. Фактор-кільце класів лишків за даним модулем. Відповідно до п. 13.3, у фактор-множині $\mathbf{Z}/m$ класів лишків за даним модулем m вводяться операції додавання й множення, погоджені з операціями додавання й множення в кільці цілих чисел, а саме:

Сумою класів $K_a^{(m)}$ і $K_b^{(m)}$ називається такий клас $K_{a+b}^{(m)}$, який містить у собі число $a + b$.

Добутком класів $K_a^{(m)}$ і $K_b^{(m)}$ називається такий клас $K_{ab}^{(m)}$, який містить у собі число ab.

Звідси дістаємо:

$$K_a^{(m)} + K_b^{(m)} = K_{a+b}^{(m)};\quad K_a^{(m)} \cdot K_b^{(m)} = K_{ab}^{(m)}. \tag{5}$$

Приклад. За модулем $m = 6$ кільце $\mathbf{Z}$ цілих чисел утворює фактор-множину класів лишків:

$$\{K_0^{(6)}, K_1^{(6)}, K_2^{(6)}, K_3^{(6)}, K_4^{(6)}, K_5^{(6)}\}.$$

Очевидно, $K_3^{(6)} + K_4^{(6)} = K_7^{(6)} = K_1^{(6)}$, а $K_2^{(6)} \cdot K_5^{6} = K_{10}^{6} = K_4^{6}$.

Теорема 3. *Фактор-множина класів лишків за даним модулем є комутативним кільцем з одиницею.*

Цю теорему доведено в загальному випадку в п. 13.2 (теорема 5).

Відповідно до цього фактор-множину класів лишків за модулем m називають фактор-кільцем. Позначатимемо його $\mathbf{Z}/m$.

Теорема 4. *Якщо m — складене число, то $\mathbf{Z}/m$ є комутативне кільце з дільниками нуля. Якщо ж m — просте число, то $\mathbf{Z}/m$ — поле.*

Доведення: Нехай m — складене, тобто, наприклад, $m = pq$. Числа p і q менші за m і більші від одиниці. Оскільки $K_p^{(m)} \neq K_0^{(m)}$ і $K_q^{(m)} \neq K_0^{(m)}$, а $K_p^{(m)} \cdot K_q^{(m)} = K_{pq}^{(m)} = K_m^{(m)} = K_0^{(m)}$ і, то класи $K_p^{(m)}$ і $K_q^{(m)}$ є дільниками нуля.

Нехай тепер m — просте число. Тоді не існує таких чисел p і q, щоб добуток їх дорівнював m. Тому в кільці $\mathbf{Z}/m$ не існує й дільників нуля. Покажемо, що в кільці

$$\mathbf{Z}/m = \{K_0^{(m)}, K_1^{(m)}, K_2^{(m)}, \ldots, K_r^{(m)}, \ldots, K_{m-1}^{(m)}\}$$

для кожного елемента $K_p^{(m)} \neq K_0^{(m)}$ існує обернений йому $(K_r^{(m)})^{-1}$, такий, що $K_p^{(m)} \cdot (K_p^{(m)})^{-1} = K_1^{(m)}$. У зв'язку з тим, що m — просте число, числа m і p — взаємно прості, тобто $(p, m) = 1$. Але тоді існують такі цілі числа x і y, що справджується рівність $px + my = 1$, тобто $px = 1 - my$. Але права частина — число $1 - my \equiv 1 \pmod{m}$ і тому $px \equiv 1 \pmod{m}$. Це означає, що $K_p^{(m)} \cdot K_x^{(m)} = K_1^{(m)}$, тобто для елемента $K_p^{(m)} \neq K_0^{(m)}$ з кільця $\mathbf{Z}/m$ обернений йому елемент $K_x^{(m)}$. Отже, при m простому $\mathbf{Z}/m$ є полем. Теорему доведено.

Приклад. Розглянемо кільце $Z/7$:

$$Z/7 = \{K_0^{(7)}, K_1^{(7)}, K_2^{(7)}, K_3^{(7)}, K_4^{(7)}, K_5^{(7)}, K_6^{(7)}\}.$$

Для ненульових елементів знайдемо обернений.

Елемент $K_p^{(7)} \neq K_0^{(7)}$	$K_1^{(7)}$	$K_2^{(7)}$	$K_3^{(7)}$	$K_4^{(7)}$	$K_5^{(7)}$	$K_6^{(7)}$
Обернений до $K_p^{(7)}$ елемент	$K_1^{(7)}$	$K_4^{(7)}$	$K_5^{(7)}$	$K_2^{(7)}$	$K_3^{(7)}$	$K_6^{(7)}$
Добуток	$K_1^{(7)}$	$K_8^{(7)}=K_1^{(7)}$	$K_{15}^{(7)}=K_1^{(7)}$	$K_8^{(7)}=K_1^{(7)}$	$K_{15}^{(7)}=K_1^{(7)}$	$K_{36}^{(7)}=K_1^{(7)}$

§ 16. ПОВНА І ЗВЕДЕНА СИСТЕМИ ЛИШКІВ. ФУНКЦІЯ ЕЙЛЕРА

16.1. Повна система лишків. У кожному класі $K_a^{(m)}$ лишків за модулем m можна знайти *найменший невід'ємний лишок* — остачу r від ділення a на m: $a = mq + r$, де $0 \leqslant r < m$, і *абсолютно найменший лишок* — лишок, який за абсолютною величиною менший від всіх абсолютних величин лишків класу $K_a^{(m)}$.

Наприклад, у класі $K_5^{(7)}$ за модулем $m = 7$

$$K_5^{(7)} = \{\ldots, -16, -9, -2, 5, 12, 19, 26, \ldots\}$$

найменшим невід'ємним лишком є число $r = 5$ (яке задовольняє нерівності $0 \leqslant r < m$), а абсолютно найменшим лишком є число -2, бо його абсолютна величина менша від абсолютних величин всіх інших лишків даного класу. Звичайно, може бути, що найменший невід'ємний лишок і абсолютно найменший лишок даного класу $K_a^{(m)}$ збігаються. Наприклад, у класі $K_3^{(7)}$

$$K_3^{(7)} = \{\ldots, -11, -4, 3, 10, 17, 24, \ldots\}$$

число 3 і є найменшим невід'ємним і абсолютно найменшим лишком.

Означення 1. *Система лишків, утворена з m чисел, узятих по одному з кожного класу, називається повною системою лишків за модулем m.*

Скорочено позначатимемо її буквами ПСЛ. При $m = 7$ повними системами лишків є, наприклад, такі системи чисел:

$$p_1 = \{0, 1, 2, 3, 4, 5, 6\},$$
$$p_2 = \{-3, -2, -1, 0, 1, 2, 3\},$$
$$p_3 = \{-24, -9, -1, 56, 1, 16, 73\}$$

і т. д.

Система p_1 складається з найменших невід'ємних лишків усіх класів, система p_2 — з абсолютно найменших лишків, а система p_3 — з довільних 7 чисел, узятих по одному з кожного класу:

$$-24 \in K_4^{(7)},\ -9 \in K_5^{(7)},\ -1 \in K_6^{(7)},\ 56 \in K_0^{(7)},\ 1 \in K_1^{(7)},\ 16 \in K_2^{(7)},\ 73 \in K_3^{(7)}.$$

Теорема 1. *Якщо $(a, m) = 1$, b — довільне число, а x пробігає ПСЛ за модулем m, то й лінійна форма $ax + b$ також пробігає ПСЛ за модулем m.*

Доведення. Нехай система чисел

$$x_0, x_1, x_2, \ldots, x_i, \ldots, x_j, \ldots, x_{m-1} \qquad (1)$$

утворює ПСЛ за модулем m. Числа попарно неконгруенті. Але тоді й числа

$$ax_0 + b,\ ax_1 + b,\ \ldots,\ ax_i + b,\ \ldots,\ ax_j + b,\ \ldots,\ ax_{m-1} + b$$

утворюють ПСЛ. Справді, цих чисел m, і вони попарно неконгруентні між собою. Якби

$$ax_i + b \equiv ax_j + b \pmod m,$$

то на підставі властивостей 3 і 8 (п. 15.1) звідси випливало б, що $x_i \equiv x_j \pmod m$, а це неправильно. Теорему доведено.

16.2. Зведена система лишків. Уведемо тепер поняття найбільшого спільного дільника класу лишків за модулем m. Згідно з властивістю 15 (п. 15.2), усі числа того самого класу $K_a^{(m)}$ мають однаковий НСД з модулем m: якщо $a, b \in K_a^{(m)}$, тобто $a \equiv b \pmod m$, то $(a, m) = (b, m)$.

Означення 2. *Найбільшим спільним дільником класу $K_a^{(m)}$ називається найбільший спільний дільник чисел a і m. Якщо $(a, m) = 1$, то клас $K_a^{(m)}$ називається взаємно простим з модулем m.*

Означення 3. *Система лишків, узятих по одному з кожного класу, взаємно простого з модулем, називається зведеною системою лишків.*

Скорочено позначатимемо цю систему буквами ЗСЛ.

Приклад. Якщо $m = 8$, то ПСЛ є 0, 1, 2, 3, 4, 5, 6, 7. Члени 1, 3, 5, 7 взаємно прості з числом 8. Вони утворюють ЗСЛ. Отже,

$$\text{ПСЛ} = \{0, 1, 2, 3, 4, 5, 6, 7\},\ \text{ЗСЛ} = \{1, 3, 5, 7\}.$$

Відповідно до цього класи $K_1^{(8)}, K_3^{(8)}, K_5^{(8)}, K_7^{(8)}$ взаємно прості з модулем m.

Означення 4. *Функцією Ейлера $\varphi(m)$ називається функція, визначена на множині натуральних чисел; значення $\varphi(m)$ є кількість невід'ємних чисел, менших за m і взаємно простих з m.*

Наприклад, $\varphi(8) = 4$, бо існує 4 невід'ємних числа, менших за 8 і взаємно простих з 8, а саме числа 1, 3, 5, 7.

Аналогічно легко встановити, що

$$\varphi(2) = 1,\ \varphi(3) = 2,\ \varphi(4) = 2,\ \varphi(5) = 4,\ \varphi(6) = 2,\ \varphi(7) = 6 \text{ і т. д.}$$

Зазначимо ще, що $\varphi(1) = 1$, бо існує одне невід'ємне число — нуль,— для якого $(0, 1) = 1$.

Очевидно, число $\varphi(m)$ дорівнює кількості чисел, які утворюють ЗСЛ за модулем m.

Теорема 2. *Якщо $(a, m) = 1$, x пробігає ЗСЛ за модулем m, то лінійна форма $y = ax$ також пробігає ЗСЛ за модулем m.*

Доведення. Нехай числа

$$\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_i, \ldots, \tilde{x}_{\varphi(m)} \qquad (2)$$

утворюють зведену систему лишків за модулем m. Всі $\bar{x}_i$ взаємно прості з числом m. Крім того, всі ці числа попарно не конгруентні між собою за модулем m:

$$\bar{x}_i \not\equiv \bar{x}_j \pmod m, \quad i \neq j, \quad i, j = 1, 2, \ldots, \varphi(m).$$

Але тоді і числа

$$a\bar{x}_1, a\bar{x}_2, \ldots, a\bar{x}_i, \ldots, a\bar{x}_{\varphi(m)}, \tag{3}$$

де $(a, m) = 1$, утворюють ЗСЛ за модулем m. Справді, перш за все, всіх чисел тут $\varphi(m)$ і вони взаємно прості з числом m: якщо $(a, m) = 1$ і $(\bar{x}_i, m) = 1$, то $(a\bar{x}_i, m) = 1$. Крім того, числа $a\bar{x}_i$, $a\bar{x}_j$ не конгруентні між собою за модулем m. Якби $a\bar{x}_i \equiv a\bar{x}_j \pmod m$, то на підставі властивості 8 п. 15.1 звідси випливало б, що $\bar{x}_i \equiv \bar{x}_j \pmod m$, а це неправильно. Отже, (3) являє собою ЗСЛ за модулем m. Теорему доведено.

З а у в а ж е н н я. Якщо число $\bar{x}_i$ є лишок ЗСЛ за модулем m і належить класу $K_i^{(m)}$, то добуток $a\bar{x}_i$ хоч і належатиме до ЗСЛ за модулем m (відповідно до теореми 2), але може належати зовсім іншому класу. Тільки в тому випадку, коли $a = km + 1$, добуток

$$a\bar{x}_i = (km + 1)\bar{x}_i = km\bar{x}_i + \bar{x}_i \equiv \bar{x}_i \pmod m$$

і тому знову належатиме класу $K_i^{(m)}$.

Наприклад, при модулі $m = 8$ ЗСЛ з найменших невід'ємних лишків складається з чисел 1, 3, 5, 7, які відповідно належать класам $K_1^{(8)}$, $K_3^{(8)}$, $K_5^{(8)}$, $K_7^{(8)}$. Якщо $a = 11$, то добутки $a\bar{x}_i$ є числами 11, 33, 55, 77, які належать відповідно до класів $K_3^{(8)}$, $K_1^{(8)}$, $K_7^{(8)}$, $K_5^{(8)}$.

16.3. Властивості функції Ейлера. В п. 7.5 було введено поняття ч и с л о в и х функцій, які визначені для всіх натуральних значень аргументу; було розглянуто деякі числові функції.

У теорії чисел використовують ряд спеціальних функцій, які дають важливу арифметичну характеристику цілих чисел. Одним з найпростіших прикладів є функція $f(x) = [x]$ (читається «ант'є від x»), яка задана на множині всіх дійсних чисел; $[x]$ — це найбільше ціле число, *менше за* x. Так, $[8, 6] = 8$, а $[-9, 3] = -10$. За допомогою функції $[x]$ можна, наприклад, вказати степінь, з яким у канонічний розклад числа $n!$ входить простий множник p. Очевидно, що цей степінь дорівнюватиме такому натуральному числу:

$$\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \cdots + \left[\frac{n}{p^k}\right] \quad (p^k < n < p^{k+1}).$$

У справедливості цього можна впевнитися на прикладі. Якщо $n = 9$, то просте число 2 в канонічний розклад числа $9! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \times \times 6 \cdot 7 \cdot 8 \cdot 9$ входить в степені 7, бо:

$$\left[\frac{9}{2}\right] + \left[\frac{9}{2^2}\right] + \left[\frac{9}{2^3}\right] = 4 + 2 + 1 = 7.$$

Число 3 входить у степені

$$\left[\frac{9}{3}\right] + \left[\frac{9}{3^2}\right] = 3 + 1 = 4,$$

а 5 і 7 — у степені 1, бо $\left[\frac{9}{5}\right] = 1$, $\left[\frac{9}{7}\right] = 1$. Отже, число $9!$ в канонічному вигляді можна записати так:

$$9! = 2^7 \cdot 3^4 \cdot 5 \cdot 7.$$

Серед числових функцій особливу роль відіграють так звані м у л ь т и п л і к а т и в н і функції.

Означення 5. *Числова функція $f(n)$, визначена на множині натуральних чисел, називається мультиплікативною, якщо для кожного n $f(n) \neq 0$ і для будь-яких взаємно простих натуральних чисел n і m*

$$f(nm) = f(n) \cdot f(m). \tag{4}$$

Мультиплікативні функції мають такі властивості:

1) *Якщо $f(n)$ — мультиплікативна функція, то* $f(1) = 1$. Справді,

$$f(n) = f(1 \cdot n) = f(1) \cdot f(n),$$

звідки $f(1) = 1$.

2) *Якщо $f(n)$ — мультиплікативна функція і числа $n_1, n_2, \ldots, n_k$ попарно взаємно прості, то*

$$f(n_1 \cdot n_2 \cdot \ldots \cdot n_k) = f(n_1) f(n_2) \cdot \ldots \cdot f(n_k). \tag{5}$$

Справді, якщо $(n_1, n_2) = 1$, $(n_1, n_3) = 1$, ..., $(n_1, n_k) = 1$, то й $(n_1, n_2 \cdot n_3 \cdot \ldots \cdot n_k) = 1$, а тому

$$f(n_1 \cdot n_2 \cdot n_3 \cdot \ldots \cdot n_k) = f(n_1 \cdot (n_2 \cdot n_3 \cdot \ldots \cdot n_k)) = \\ = f(n_1) \cdot f(n_2 \cdot n_3 \cdot \ldots \cdot n_k).$$

Аналогічно міркуючи відносно $f(n_2 \cdot n_3 \cdot \ldots \cdot n_k)$ і наступних значень функції $f(n)$, дістанемо (5).

3) *Добуток мультиплікативних функцій є мультиплікативна функція.*

Справді, якщо $f_1(n)$ і $f_2(n)$ є мультиплікативні функції, то функція $F(n) = f_1(n) \cdot f_2(n)$ має таку особливість: $F(n) \neq 0$ ні для яких n, бо $f_1(n) \neq 0$ і $f_2(n) \neq 0$. Крім того, для будь-яких n_1 і n_2

$$F(n_1 \cdot n_2) = f_1(n_1 \cdot n_2) \cdot f_2(n_1 \cdot n_2) = f_1(n_1) \cdot f_1(n_2) \cdot f_2(n_1) \times \\ \times f_2(n_2) = F(n_1) \cdot F(n_2).$$

Отже, $F(n)$ — мультиплікативна функція.

Розглянемо тепер властивості функції Ейлера $\varphi(m)$.

Теорема 3. *Функція Ейлера $\varphi(m)$ мультиплікативна.*

Д о в е д е н н я. Нагадаємо, перш за все, що $\varphi(1) = 1$. Для всіх $n > 1$ $\varphi(n) \neq 0$, бо хоч одне натуральне число, менше за n і взаємно просте з n, — одиниця, завжди існує.

Нехай тепер маємо числа m і n, причому $(m, n) = 1$. Підрахуємо кількість натуральних чисел, менших за mn і взаємно простих з mn.

Для цього всі числа від 1 до mn розмістимо у вигляді такої таблиці:

$$\begin{array}{lllll}
1, & 2, & 3, \ldots, & r, \ldots, & m; \\
m+1, & m+2, & m+3, \ldots, & m+r, \ldots, & 2m; \\
2m+1, & 2m+2, & 2m+3, \ldots, & 2m+r, \ldots, & 3m; \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
sm+1, & sm+2, & sm+3, \ldots, & sm+r, \ldots, & (s+1)m; \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
(n-1)m+1, & (n-1)m+2, & \ldots, & (n-1)\times m+r, \ldots, & nm. \qquad (6)
\end{array}$$

Взаємно простими з добутком mn будуть, очевидно, ті і тільки ті числа, які взаємно прості і з m, і з n. Тому відберемо спочатку ті числа, які взаємно прості з m, а потім з них відберемо ті, які ще взаємно прості з числом n. З числом m взаємно простих чисел у першому рядку є $\varphi(m)$ чисел. Вони утворюють ЗСЛ за модулем m і є представниками $\varphi(m)$ класів чисел, в яких всі числа взаємно прості з m. Нехай для конкретності число r з таблиці (6) задовольняє умову $(r, m) = 1$. Тоді всі числа класу $K_r^{(m)}$ будуть взаємно прості з m. У стовпчику таблиці (6), який починається числом r, міститься n чисел класу $K_r^{(m)}$. Отже, у таблиці (6) існує $\varphi(m)$ стовпчиків, кожний з яких містить n чисел, взаємно простих з m. Усього таких чисел $\varphi(m) \cdot n$.

Розглянемо тепер усі n елементів будь-якого r-го стовпчика таблиці (6):

$$r,\ m+r,\ 2m+r,\ \ldots,\ sm+r,\ \ldots,\ (n-1)m+r. \qquad (7)$$

Не важко впевнитися, що ці числа утворюють ПСЛ за модулем n. Справді, числа (7) є значеннями лінійної форми $f(x) = mx + r$, де x пробігає ПСЛ за модулем n: 0, 1, 2, ..., $n-1$, причому $(m, n) = 1$. Але тоді за теоремою 1 п. 16.1 і лінійна форма $f(x) = mx + r$ пробігає ПСЛ за модулем n. Отже, числа r-го стовпчика утворюють ПСЛ за модулем n, серед яких є $\varphi(n)$ чисел, взаємно простих з n.

У всіх $\varphi(m)$ таких стовпчиках є, очевидно, всього $\varphi(n) \cdot \varphi(m)$ чисел, взаємно простих з n. Тому при $(m, n) = 1$ маємо $\varphi(mn) = \varphi(m) \cdot \varphi(n)$. Теорему доведено.

Теорема 4. *Якщо p — просте число, а k — натуральне число, то*

$$\varphi(p^k) = p^k\left(1 - \frac{1}{p}\right). \qquad (8)$$

Доведення. Перш за все зазначимо, що для простого числа p

$$\varphi(p) = p - 1, \qquad (9)$$

що безпосередньо випливає з означення функції Ейлера.

Випишемо тепер всі натуральні числа від 1 до p^k:

$$1, 2, 3, \ldots, p-1, p, p+1, p+2, \ldots, 2p, 2p+1, \ldots$$
$$\ldots, 3p, 3p+1, \ldots, p^2, \ldots, p^k. \qquad (10)$$

Числа (10) можна поділити на групи по p послідовних натуральних чисел:

$$\{1, 2, 3, \ldots, p-1, p\}, \{p+1, p+2, \ldots, 2p\}, \{2p+1, \ldots, 3p\}, \ldots$$
$$\ldots, \{(p^{k-1}-1)p+1, \ldots, p^k\}.$$

У кожній з цих груп p^{k-1} останнє число ділиться на p, а решта на p не діляться, і, отже, взаємно прості з p. Тому вони взаємно прості і з p^k. Таким чином, усіх чисел, взаємно простих з p^k, є $p^k - p^{k-1} = p^k \times \times \left(1 - \frac{1}{p}\right)$, тобто $\varphi(p^k) = p^k\left(1 - \frac{1}{p}\right)$, що й треба було довести.

Теорема 5. *Якщо канонічний розклад числа m має вигляд*

$$m = p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s},$$

то

$$\varphi(m) = m\prod_{i=1}^{s}\left(1 - \frac{1}{p_i}\right). \qquad (11)$$

Доведення. Оскільки $\varphi(m)$ — мультиплікативна функція, а $p_1, p_2, \ldots, p_s$ — прості числа, то на основі (8) маємо

$$\varphi(m) = \varphi(p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s}) = \varphi(p_1^{k_1}) \cdot \varphi(p_2^{k_2}) \cdot \ldots \cdot \varphi(p_s^{k_s}) =$$
$$= p_1^{k_1}\left(1 - \frac{1}{p_1}\right) \cdot p_2^{k_2}\left(1 - \frac{1}{p_2}\right) \cdot \ldots \cdot p_s^{k_s}\left(1 - \frac{1}{p_s}\right) =$$
$$= p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s}\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdot \ldots \cdot \left(1 - \frac{1}{p_s}\right)$$

і тому

$$\varphi(m) = m\prod_{i=1}^{s}\left(1 - \frac{1}{p_i}\right).$$

Теорему доведено.

Приклад. Для $m = 360 = 2^3 \cdot 3^2 \cdot 5$

$$\varphi(360) = 360\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right) = 96.$$

Теорема 6. *Сума значень функції Ейлера для всіх дільників d_j числа m дорівнює m:*

$$\sum_j \varphi(d_j) = m. \qquad (12)$$

Доведення. Нехай канонічний розклад числа m є

$$m = p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s}.$$

Будь-який дільник d числа m має вигляд

$$d_l = p_1^{l_1} \cdot p_2^{l_2} \cdot \ldots \cdot p_s^{l_s}.$$

Для того щоб скласти $\sum_j \varphi(d_j)$, розглянемо добуток

$$[1+\varphi(p_1)+\varphi(p_1^2)+\ \cdots\ +\varphi(p_1^{k_1})]\cdot[1+\varphi(p_2)+\varphi(p_2^2)+\ \cdots$$

$$\cdots\ +\varphi(p_2^{k_2})]\cdot\ \ldots\ \cdot[1+\varphi(p_s)+\varphi(p_s^2)+\ \cdots\ +\varphi(p_s^{k_s})]. \quad (13)$$

Якщо тут розкрити дужки, то кожний член суми являє собою значення функції Ейлера для деякого дільника d_j числа m. Отже,

$$\sum_j \varphi(d_j) = [1+\varphi(p_1)+\ \cdots\ +\varphi(p_1^{k_1})]\cdot[1+\varphi(p_2)+\ \cdots$$

$$\cdots\ +\varphi(p_2^{k_2})]\cdot\ \cdots\ \cdot[1+\varphi(p_s)+\ \cdots\ +\varphi(p_s^{k_s})].$$

Але вираз у перших квадратних дужках можна перетворити так:

$$1+\varphi(p_1)+\varphi(p_1^2)+\cdots+\varphi(p_1^{k_1}) = 1+(p_1-1)+p_1(p_1-1)+\cdots+$$

$$+p_1^{k_1-1}(p_1-1) = 1+p_1-1+p_1^2-p_1+\ \cdots\ +p_1^{k_1}-p_1^{k_1-1} = p_1^{k_1}.$$

Аналогічний результат дістанемо для виразів в інших квадратних дужках. Тому

$$\sum_j \varphi(d_j) = p_1^{k_1}\cdot p_2^{k_2}\cdot\ \ldots\ \cdot p_s^{k_s} = m,$$

що й треба було довести.

Формула (12) називається формулою Гаусса.

16.4. Теореми Ейлера і Ферма. З властивостей систем лишків безпосередньо випливають відомі теореми Ейлера і Ферма.

Теорема 7 (Ейлера). *Якщо m — натуральне число і $m>1$, $(a, m)=1$, то*

$$a^{\varphi(m)} \equiv 1 \pmod m. \quad (14)$$

Доведення. Нехай числа

$$\bar{x}_1, \bar{x}_2, \ \ldots\ , \bar{x}_{\varphi(m)} \quad (15)$$

утворюють ЗСЛ найменших невід'ємних лишків за модулем m. Тоді при $(a, m)=1$ числа $a\bar{x}_1, a\bar{x}_2, \ldots, a\bar{x}_{\varphi(m)}$ утворюють теж ЗСЛ за модулем m. Ці числа належатимуть якимсь класам, взаємно простим з модулем m. Числа $a\bar{x}_1, a\bar{x}_2, \ldots, a\bar{x}_{\varphi(m)}$ можна замінити системою найменших невід'ємних лишків

$$\bar{x}_1, \bar{x}_2, \ \ldots\ , \bar{x}_{\varphi(m)} \quad (16)$$

з тих же класів, до яких вони належать. Вважаючи, що

$$a\bar{x}_1 \equiv \bar{x}_1 \pmod m,$$

$$a\bar{x}_2 \equiv \bar{x}_2 \pmod m,$$

$$\ldots\ldots\ldots$$

$$a\bar{x}_{\varphi(m)} \equiv \bar{x}_{\varphi(m)} \pmod m,$$

перемножимо ці $\varphi(m)$ конгруенцій. Дістанемо

$$a^{\varphi(m)}\cdot\bar{x}_1\cdot\bar{x}_2\cdot\ \ldots\ \cdot\bar{x}_{\varphi(m)} \equiv \bar{x}_1\cdot\bar{x}_2\cdot\ \ldots\ \cdot\bar{x}_{\varphi(m)} \pmod m.$$

Оскільки (15) і (16) — це ті самі числа, взаємно прості з m, то, скорочуючи на них обидві частини конгруенції за властивістю 8 (п. 15.1), дістанемо

$$a^{\varphi(m)} \equiv 1 \pmod m,$$

що й треба було довести.

Теорема 8 (Ферма). *Якщо число p просте і $(a, p)=1$, то*

$$a^{p-1} \equiv 1 \pmod p. \quad (17)$$

Доведення. Теорема Ферма є наслідком теореми Ейлера. Згідно з формулою (9), $\varphi(p)=p-1$, тому формула (14) матиме вигляд (17). Теорему доведено.

Наслідок. *Якщо число p просте, то для будь-якого цілого числа a має місце конгруенція*

$$a^p \equiv a \pmod p. \quad (18)$$

Справді, множачи обидві частини конгруенції (17) на a, дістанемо

$$[a^{p-1} \equiv 1 \pmod p] \Rightarrow [a^p = a \pmod p].$$

При цьому результат має місце і тоді, коли числа a і p не взаємно прості. Якщо $(a, p) \neq 1$ при p простому, то $(a, p)=p$, тобто $a \vdots p$. Але тоді й $a^p - a = a(a^{p-1}-1)$ також ділиться на p. Отже, $a^p - a \equiv 0 \pmod p$, або $a^p \equiv a \pmod p$, що й треба було довести.

Приклад. Знайти остачу від ділення числа 42^{50} на 17. Оскільки 17 — просте число і $(42, 17)=1$, то за (17) $42^{16} \equiv 1 \pmod{17}$. Підносячи до куба обидві частини конгруенції, далі маємо $42^{48} \equiv 1 \pmod{17}$. Крім того, $42 \equiv 8 \pmod{17}$, а в квадраті це дає $42^2 \equiv 13 \pmod{17}$. Потім дістаємо

$$[42^{48} \equiv 1 \pmod{17} \wedge 42^2 \equiv 13 \pmod{17}] \Rightarrow [42^{50} \equiv 13 \pmod{17}].$$

Отже, остача дорівнює 13.

§ 17. ЛІНІЙНІ КОНГРУЕНЦІЇ З ОДНИМ НЕВІДОМИМ

17.1. Загальні означення. *Конгруенціями з одним невідомим за модулем m називаються конгруенції виду*

$$f(x) \underset{df}{=} a_n x^n + a_{n-1}x^{n-1} + \ \cdots\ + a_1 x + a_0 \equiv 0 \pmod m, \quad (1)$$

де в лівій частині міститься многочлен з цілими коефіцієнтами. Якщо a_n не ділиться на число m, то n називається *степенем конгруенції;* при $a_n \vdots m$ старший член $a_n x^n \equiv 0 \pmod m$ і його можна відкинути.

Розв'язком конгруенції $f(x) \equiv 0 \pmod m$ є будь-яке ціле число a, яке задовольняє конгруенцію, тобто $f(a) \equiv 0 \pmod m$. Легко зрозуміти, що в цьому випадку разом з числом a конгруенцію задовольняють і всі числа класу $K_a^{(m)}$, конгруентні з a за модулем m. Саме це стверджує наслідок 3 властивості 9 п. 15.1. Тому введемо таке означення.

Означення 1. *Розв'язком конгруенції $f(x) \equiv 0 \pmod m$ називається клас лишків за модулем m, кожне число якого задовольняє цю конгруенцію.*

Оскільки класів чисел за даним модулем m є m, то конгруенція може мати лише скінченну кількість розв'язків або може не мати їх зов-

сім. Наприклад, конгруенція $2x \equiv 3 \pmod 4$ не має розв'язків, бо з неї випливає рівність $2x = 3 + 4t$. Така рівність неможлива, бо при будь-яких x і t ліва частина рівності — парне число, а права — непарне число.

Щоб знайти розв'язки, досить підставити в конгруенцію замість невідомого x числа з різних класів за модулем m. Для цього можна перебрати ПСЛ з найменших невід'ємних лишків, а ще краще — повну систему абсолютно найменших лишків.

Приклад. Щоб знайти розв'язки конгруенції

$$2x^2 - 5x - 2 \equiv 0 \pmod 5, \tag{2}$$

знайдемо ПСЛ з найменших невід'ємних лишків за модулем 5:

$$0, 1, 2, 3, 4. \tag{3}$$

В результаті підстановки в конгруенцію впевнюємося, що числа 1 і 4 задовольняють її. Отже, розв'язком конгруенції (2) є класи лишків $K_1^{(5)}$ і $K_4^{(5)}$.

Конгруенції розв'язують за допомогою побудови більш простих конгруенцій, рівносильних заданим.

Означення 2. *Конгруенції називаються рівносильними, якщо множини їх розв'язків збігаються.*

Щоб побудувати рівносильні конгруенції, над заданою конгруенцією проводять операції, які грунтуються на властивостях, розглянутих у § 15. До операцій, які не порушують множину розв'язків конгруенцій, належать такі:

а) Додавання до обох частин конгруенції будь-якого многочлена $g(x)$ з цілими коефіцієнтами.

б) Додавання до однієї з частин конгруенції многочлена з коефіцієнтами, кратними модулю.

в) Множення обох частин конгруенції на число, взаємно просте з модулем.

г) Множення обох частин конгруенції і модуля на те саме додатне число.

Наприклад, конгруенцію (2) можна спростити так. Відкинемо спочатку член $-5x$, коефіцієнт якого кратний модулю. Дістанемо конгруенцію $2x^2 - 2 \equiv 0 \pmod 5$. Далі, скорочуючи на число 2 ліву (і праву) частину, дістанемо конгруенцію $x^2 - 1 \equiv 0 \pmod 5$, рівносильну конгруенції (2). Підставляючи в неї числа з ПСЛ (3), встановлюємо, що розв'язками конгруенції (2) є класи лишків $K_1^{(5)}$ і $K_4^{(5)}$.

Конгруенції 1-го степеня мають вигляд $a_1x + a_0 \equiv 0 \pmod m$. Переносячи вільний член у праву частину конгруенції і змінюючи позначення коефіцієнтів, дістанемо

$$ax \equiv b \pmod m. \tag{4}$$

При розв'язуванні таких конгруенцій розглядають два випадки: $(a, m) = 1$ і $(a, m) = d > 1$.

Теорема 1. *Якщо* $(a, m) = 1$, *то конгруенція* (4) *має єдиний розв'язок.*

Доведення. Конгруенція може мати не більш як m розв'язків відповідно до кількості чисел в ПСЛ. Якщо x пробігає ПСЛ, то й лінійна форма $ax - b$ також пробігає ПСЛ. При цьому один раз

лінійна форма $ax - b$ прийме числове значення, яке конгруентне нулю, бо нуль є один з лишків ПСЛ. Нехай при $x = j$ матимемо $aj \equiv b \pmod m$. Тоді клас лишків $K_j^{(m)}$ і є єдиним розв'язком конгруенції (4). Теорему доведено.

Теорема 2. *Якщо* $(a, m) = d > 0$ *і число* b *не ділиться на* d, *то конгруенція* $ax \equiv b \pmod m$ *не має розв'язків.*

Доведення. Припустимо супротивне. Нехай при деякому x_0 конгруенція (4) задовольняється. Тоді

$$[ax_0 \equiv b \pmod m] \Rightarrow [ax_0 = b + mt].$$

Але така рівність неможлива, якщо $a \vdots d$, $m \vdots d$, а b не ділиться на d. Отже, наше припущення неправильне, і теорему доведено.

Теорема 3. *Якщо* $(a, m) = d > 1$ *і* $b \vdots d$, *то конгруенція* (4) *має* d *розв'язків.*

Доведення. Нехай $a = a_1d$, $b = b_1d$, $m = m_1d$. Тоді, скорочуючи (4) на d, дістанемо конгруенцію $a_1x \equiv b_1 \pmod{m_1}$, рівносильну конгруенції (4). Оскільки $(a_1, m_1) = 1$, то остання конгруенція за теоремою 1 має єдиний розв'язок — клас лишків $K_j^{(m_1)}$. Проте цей клас лишків розпадається на d класів лишків за модулем $m = dm_1$ (теорема 2 п. 15.3), які й утворюють d розв'язків заданої конгруенції (4):

$$\{K_j^{(m_1)}, K_{j+m_1}^{(m_1)}, \ldots, K_{j+(d-1)m_1}^{(m_1)}.$$

Теорему доведено.

Приклад. Розв'яжемо конгруенцію

$$6x \equiv 15 \pmod{21}. \tag{5}$$

Оскільки $(6, 21) = 3$ і 15 ділиться на число 3, то конгруенція має три розв'язки. Скорочуючи всі члени конгруенції і модуль на 3, матимемо

$$2x \equiv 5 \pmod 7. \tag{6}$$

Конгруенція (6) рівносильна заданій і має єдиний розв'язок, бо $(2, 7) = 1$.

Випробовуючи числа, які входять в ПСЛ за модулем 7:

$$0, 1, 2, 3, 4, 5, 6,$$

знаходимо, що тільки число $j = 6$ є розв'язком (6). Отже, клас $K_6^{(7)}$ є розв'язком конгруенції (6). А за модулем $m = 21$ він розпадається на класи лишків

$$K_6^{(21)}, K_{13}^{(21)}, K_{20}^{(21)},$$

які є розв'язками конгруенції (5).

17.2. Способи розв'язування конгруенцій 1-го степеня. Розглянемо деякі, найбільш поширені способи розв'язування конгруенцій 1-го степеня:

а) *Підстановка в конгруенцію чисел ПСЛ.* Цей спосіб використовують при невеликих модулях. При великих модулях підстановку лишків ПСЛ проводять на заключному етапі побудови рівносильних конгруенцій.

б) *Зведення конгруенцій 1-го степеня до рівносильної їй конгруенції з коефіцієнтом при* x, *рівному одиниці.* Цей спосіб полягає в проведенні ряду рівносильних перетворень заданої конгруенції за допомогою операцій, розглянутих у п. 17.1.

Приклад 1. Конгруенція $22x \equiv 9 \pmod{29}$ має єдиний розв'язок, бо $(22,29) = 1$. Його можна знайти так. Додамо до лівої частини конгруенції $-29x$. Дістанемо $-7x \equiv 9 \pmod{29}$. Помноживши обидві частини конгруенції на 4, а потім, додавши до неї $+29x$, дістанемо $x \equiv 36 \pmod{29}$, або $x \equiv 7 \pmod{29}$. Отже, розв'язком заданої конгруенції є клас $K_7^{(29)}$.

в) *Спосіб Ейлера.*

Нехай задано конгруенцію

$$ax \equiv b \pmod{m}, \tag{7}$$

де $(a, m) = 1$. Конгруенція має єдиний розв'язок. За теоремою Ейлера

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

Справедливою є, очевидно, і така конгруенція:

$$a^{\varphi(m)} \cdot b \equiv b \pmod{m},$$

або

$$a(a^{\varphi(m)-1} \cdot b) \equiv b \pmod{m}. \tag{8}$$

Порівнюючи конгруенції (7) і (8), бачимо, що

$$x \equiv a^{\varphi(m)-1} b \pmod{m}. \tag{9}$$

Приклад 2. Розв'язком конгруенції

$$3x \equiv 2 \pmod{8}$$

за формулою (9) є клас чисел x,

$$x \equiv 3^{\varphi(8)-1} \cdot 2 \pmod{8} \equiv 3^3 \cdot 2 \pmod{8}$$

або

$$x \equiv 6 \pmod{8}.$$

Отже, $K_6^{(8)}$ є розв'язком заданої конгруенції.

г) *Розв'язування конгруенцій* 1-го *степеня за допомогою неперервних дробів.*

Нехай знову дано конгруенцію

$$ax \equiv b \pmod{m},$$

де $(a, m) = 1$. Розкладемо $\frac{m}{a}$ в ланцюговий дріб (див. п. 9. 1). Якщо $\frac{P_{n-1}}{Q_{n-1}}$ і $\frac{P_n}{Q_n} = \frac{m}{a}$ є останні підхідні дроби, то на основі теореми 4.§ 9 маємо

$$P_n Q_{n-1} - P_{n-1} Q_n = (-1)^{n-1},$$

тобто

$$mQ_{n-1} - P_{n-1} a = (-1)^{n-1}.$$

Враховуючи, що перший член кратний модулю m, дістанемо далі

$$-P_{n-1} a \equiv (-1)^{n-1} \pmod{m},$$

або

$$a(-1)^n P_{n-1} \equiv 1 \pmod{m}$$

і, нарешті,

$$a(-1)^n P_{n-1} b = b \pmod{m}.$$

Порівнюючи цей результат з конгруенцією (7), дістаємо

$$x \equiv (-1)^n P_{n-1} b \pmod{m}, \tag{10}$$

що й є розв'язком конгруенції (7).

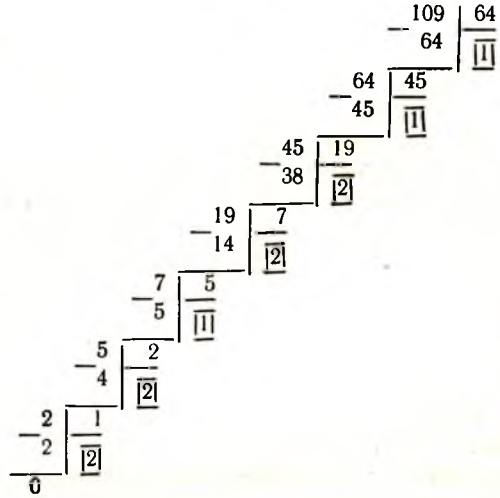
Приклад 3. Конгруенція

$$192x \equiv 9 \pmod{327} \tag{11}$$

має три розв'язки, бо $(192, 327) = 3$ і $9 : 3$. Скорочуючи всі члени і модуль конгруенції на 3, дістанемо конгруенцію

$$64x \equiv 3 \pmod{109}, \tag{12}$$

рівносильну заданій. Розкладемо число $\frac{109}{64}$ в ланцюговий дріб. За алгоритмом Евкліда маємо



Неповними частками є числа 1, 1, 2, 2, 1, 2, 2. Очевидно, тут $n = 6$. Обчислимо тепер число P_5 за схемою.

q_i	1	1	2	2	1	2	2
P_i	1	2	5	12	17	46	109

Отже, чисельник P_5 передостаннього підхідного дробу дорівнює 46. Тому за формулою (10) маємо

$$x \equiv (-1)^6 \cdot 46 \cdot 3 \pmod{109},$$

або

$$x \equiv 29 \pmod{109}.$$

Таким чином, клас чисел $K_{29}^{(109)}$ є розв'язком конгруенції (12) і заданої конгруенції (11). Але при модулі $m = 327$ цей клас лишків розпадається на три класи

$$K_{29}^{(327)}, K_{138}^{(327)}, K_{247}^{(327)},$$

які й є розв'язками конгруенції (11).

У цьому параграфі ми розглянули питання про розв'язування однієї лінійної конгруенції з одним невідомим. Цікавий і важливий розділ теорії чисел становить вивчення систем таких конгруенцій. Читачеві, який зацікавиться цими питаннями, рекомендуємо книгу Б о р о д і н а О. І. Теорія чисел. Вид. 2. К., «Вища школа», 1970.

§ 18. КОНГРУЕНЦІЇ n-ГО СТЕПЕНЯ

18.1. Побудова рівносильних конгруенцій. Розглянемо методику розв'язування конгруенцій n-го степеня виду

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \equiv 0 \pmod p, \qquad (1)$$

де p — просте число. За допомогою операцій, описаних у п. 17.1, можна побудувати рівносильну до (1) конгруенцію степеня, не вище $p-1$, з коефіцієнтами, які є найменшими невід'ємними або абсолютно найменшими лишками ПСЛ за модулем p.

Побудову такої конгруенції можна провести в такому порядку.

а) *Замінити всі коефіцієнти a_i многочлена $f(x)$ відповідними їм найменшими невід'ємними або абсолютно найменшими лишками з* ПСЛ *за модулем p.* Наприклад, конгруенція $12x^8 - 17x^3 + 12x^2 - 8 \equiv 0 \pmod 7$ рівносильна конгруенції $5x^8 - 3x^3 + 5x^2 - 1 \equiv 0 \pmod 7$, яку ми дістали, відкинувши з першої конгруенції многочлен $7x^8 - 14x^3 + 7x^2 - 7$, що має коефіцієнти, кратні модулю 7, тому при будь-якому x

$$7x^8 - 14x^3 + 7x^2 - 7 \equiv 0 \pmod 7.$$

б) *Зробити коефіцієнт при старшому члені конгруенції рівним одиниці.* Можна вважати, що $(a_n, p) = 1$. Справді, якщо a_n і просте число p не взаємно прості, то $(a_n, p) = p$, тобто $a_n \vdots p$. Але тоді в конгруенції (1) можна відкинути старший член і вона вже не буде конгруенцією n-го степеня.

За теоремою 1 п. 17.1, при умові $(a_n, p) = 1$, конгруенція $a_n x \equiv 1 \pmod p$ має єдиний розв'язок. Нехай найменший невід'ємний лишок цього класу - розв'язку є x_0. Отже, $a_n x_0 \equiv 1 \pmod p$ і $a_n x_0 x^n \equiv x^n \pmod p$. Тому, перемножаючи конгруенцію (1) на число x_0, дістанемо

$$a_n x_0 x^n + a_{n-1}x_0 x^{n-1} + \cdots + a_1 x_0 x + a_0 x_0 \equiv 0 \pmod p,$$

або

$$x^n + a_{n-1}x_0 x^{n-1} + \cdots + a_1 x_0 x + a_0 x_0 \equiv 0 \pmod p.$$

Приклад 1. Оскільки конгруенція $5x \equiv 1 \pmod 7$ має своїм розв'язком клас лишків

$$K_3^{(7)} = \{\ldots, -11, -4, 3, 10, 17, \ldots\},$$

то обидві частини конгруенції

$$5x^8 - 3x^3 + 5x^2 - 1 \equiv 0 \pmod 7$$

можна помножити на будь-яке число 3 цього класу лишків. Помножимо їх на число $x_0 = 3$. Дістанемо

$$15x^8 - 9x^3 + 15x^2 - 3 \equiv 0 \pmod 7,$$

або конгруенцію

$$x^8 - 2x^3 + x^2 - 3 \equiv 0 \pmod 7,$$

в якій коефіцієнт при старшому члені дорівнює одиниці.

в) *Понизити степінь конгруенції.*

Якщо степінь $n \geqslant p$, то конгруенцію (1) можна замінити рівносильною їй конгруенцією степеня, не вищого $p-1$. За наслідком теореми Ферма (п. 16.4) для будь-якого цілого числа x і простого p

$$x^p \equiv x \pmod p,$$

або

$$x^p - x \equiv 0 \pmod p. \qquad (2)$$

З другого боку, ділячи многочлен $f(x)$ на $x^p - x$, дістанемо

$$f(x) = (x^p - x)\,g(x) + r(x).$$

Враховуючи це, на основі конгруенції (2) матимемо

$$f(x) = (x^p - x)\,g(x) + r(x) \equiv r(x) \pmod p,$$

тобто

$$f(x) \equiv r(x) \pmod p. \qquad (3)$$

Звідси випливає, що конгруенції $f(x) \equiv 0 \pmod p$ і $r(x) \equiv 0 \pmod p$ рівносильні. Справді, якщо x_0 задовольняє конгруенцію $f(x) \equiv 0 \pmod p$, то на основі (3) маємо $r(x_0) \equiv 0 \pmod p$. Справедливим є і обернене твердження: будь-який розв'язок конгруенції $r(x) \equiv 0 \pmod p$ задовольняє і конгруенцію $f(x) \equiv 0 \pmod p$.

При практичному перетворенні конгруенції ділення многочлена $f(x)$ на $x^p - x$ не виконують, а користуються деякими простими конгруенціями. Поділимо n на число $p-1$, поставивши умову, що остача m може дорівнювати лишкам $1, 2, \ldots, p-1$ за модулем p. Тоді $n = (p-1)k + m$ $(1 \leqslant m \leqslant p-1)$.

На основі теореми Ейлера

$$x^{p-1} \equiv 1 \pmod p.$$

Тому

$$x^n = x^{(p-1)k+m} = x^{(p-1)k} \cdot x^m \equiv 1 \cdot x^m \pmod p,$$

тобто $x^n \equiv x^m \pmod p$.

Приклад 2. Конгруенція

$$x^{16} - 3x^{13} + 5x^8 - 4x^7 + 3x^4 + x + 1 \equiv 0 \pmod 7$$

рівносильна конгруенції

$$x^4 - 3x + 5x^2 - 4x + 3x^4 + x + 1 \equiv 0 \pmod 7,$$

тобто конгруенції

$$4x^4 + 5x^2 - 6x + 1 \equiv 0 \pmod 7.$$

18.2. Число розв'язків конгруенції n-го степеня.

Теорема 1. *Конгруенція n-го степеня за простим модулем може мати не більш як n розв'язків.*

Доведення. Використаємо метод математичної індукції. При $n = 1$ конгруенція (1) має вигляд

$$a_1x + a_0 \equiv 0 \pmod p, \text{ де } (a, p) = 1.$$

За теоремою 1 п. 17.1 така конгруенція має єдиний розв'язок. Отже, для $n = 1$ теорема правильна. Нехай тепер теорема правильна для $n = k - 1$. Доведемо, що вона правильна і для $n = k$. При $n = k$ маємо конгруенцію виду

$$f(x) = a_kx^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0 \equiv 0 \pmod p, \qquad (4)$$

де a_k не ділиться на p, тобто $(a_k, p) = 1$. Якщо конгруенція (4) не має розв'язку, то в цьому разі теорема правильна: конгруенція (4) має не більш як k розв'язків. Нехай тепер вона має розв'язки і, зокрема, число x_1 задовольняє конгруенцію (4), тобто

$$f(x_1) \equiv 0 \pmod p. \qquad (5)$$

Оскільки за теоремою Безу[1]

$$f(x) = (x - x_1)\,g(x) + f(x_1),$$

де $g(x)$ — многочлен степеня $(k - 1)$, то конгруенція (4) запишеться так:

$$(x - x_1)\,g(x) + f(x_1) \equiv 0 \pmod p \text{ і}$$

з урахуванням (5) матиме вигляд:

$$(x - x_1)\,g(x) \equiv 0 \pmod p.$$

Множина розв'язків цієї конгруенції складається з розв'язків конгруенції $x - x_1 \equiv 0 \pmod p$ і конгруенції $g(x) \equiv 0 \pmod p$. Перша з них має один розв'язок, а друга за припущенням індукції — не більш як $k - 1$. Разом конгруенція (4) має не більш як k розв'язків. Теорему доведено.

Доведене твердження має практичне застосування. Наприклад, конгруенція $3x^2 \equiv 5 \pmod 7$ задовольняється числом $x = 2$ і, очевидно, $x = -2$. Оскільки $-2 \equiv 5 \pmod 7$, то $K_2^{(7)}$, $K_5^{(7)}$ є розв'язками цієї конгруенції і більше розв'язків шукати не треба — їх не існує.

Введемо в розгляд поняття тотожної конгруенції.

Означення. *Якщо всі коефіцієнти многочлена $f(x)$ є кратними модуля m, то конгруенція*

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \equiv \pmod m$$

називається тотожною конгруенцією.

Така конгруенція задовольняється будь-яким числовим значенням x; про степінь такої конгруенції можна говорити лише умовно, бо і старший член, і будь-які інші члени можна відкинути, як кратні модулю m.

Встановимо достатню ознаку того, що дана конгруенція є тотожною. Для цього спочатку звернемо увагу на те, що будь-який многочлен n-го степеня $f(x) = a_nx_n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ завжди можна подати у вигляді

$$\begin{aligned} f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = \\ = a(x - x_1)(x - x_2) \ldots (x - x_{n-1})(x - x_n) + \\ + b(x - x_1)(x - x_2) \ldots (x - x_{n-1}) + \\ \cdots\cdots\cdots\cdots \\ + k(x - x_1)(x - x_2) + \\ + l(x - x_1) + \\ + m, \end{aligned} \qquad (6)$$

де $x_1, x_2, \ldots, x_{n-1}, x_n$ — будь-які попарно різні числа. Прирівнюючи коефіцієнти при однакових степенях x в лівій і правій частинах цієї тотожності, можна знайти коефіцієнти $a, b, \ldots, k, l, m$ за коефіцієнтами $a_n, a_{n-1}, \ldots, a_1, a_0$, або навпаки.

Приклад 3. Подати многочлен $f(x) = x^3 - 2x^2 + 3x - 7$ у формі (6), коли $x_1 = 2$, $x_2 = 3$, $x_3 = 1$. Маємо

$$\begin{aligned} f(x) = x^3 - 2x^2 + 3x - 7 = \\ = a(x - 2)(x - 3)(x - 1) + b(x - 2)(x - 3) + c(x - 2) + d. \end{aligned}$$

Зрівнювання коефіцієнтів дає результати: $a = 1$, $b = 4$, $c = 12$, $d = -1$, тобто

$$x^3 - 2x^2 + 3x - 7 = (x - 2)(x - 3)(x - 1) + 4(x - 2)(x - 3) + 12(x - 2) - 1.$$

Теорема 2. *Якщо конгруенція n-го степеня $f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \equiv 0 \pmod p$ має $n + 1$ розв'язків, то всі її коефіцієнти кратні модулю, тобто ця конгруенція тотожна.*

Доведення. Припустимо, що $x_1, x_2, \ldots, x_n, x_{n+1}$ — лишки $n + 1$ розв'язків конгруенції

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \equiv 0 \pmod p.$$

Користуючись формулою (6), цю конгруенцію можна записати так:

$$\begin{aligned} f(x) = a(x - x_1)(x - x_2) \ldots (x - x_{n-1})(x - x_n) + b(x - x_1)(x - x_2) \ldots \\ \ldots (x - x_{n-1}) + \cdots + k(x - x_1)(x - x_2) + l(x - x_1) + m. \end{aligned} \qquad (7)$$

Якщо тут узяти $x = x_1$, то за умовою теореми $f(x_1) \equiv 0 \pmod p$. Але це означає, що $m \equiv 0 \pmod p$, тобто $m \vdots p$. Відкидаючи з конгруенції (7) число m і виносячи за дужки $x - x_1$, дістанемо

$$\begin{aligned} (x - x_1)\,[a(x - x_2) \ldots (x - x_{n-1})(x - x_n) + \cdots + k(x - x_2) + l] \equiv \\ \equiv 0 \pmod p. \end{aligned}$$

Узявши тут знову $x = x_2$ і враховуючи, що $x_2 - x_1$ не ділиться на p як різниця лишків неконгруентних класів, бачимо, що це можливо тільки при умові $l \vdots p$. Аналогічно впевнюємося, що й інші коефіцієнти $a, b, \ldots, k$ є кратними p. Але тоді кратними p будуть і всі коефіцієнти $a_n, a_{n-1}, \ldots, a_1, a_0$ многочлена $f(x)$, бо вони є алгебраїчними сумами чисел $a, b, \ldots, k, l, m$, кратних модулю p. Теорему доведено.

[1] Етьєн Безу (1730—1783) — французький математик.

Теорема 3 (***Вільсона*** [1]). *Якщо p — просте число, то*

$$(p-1)! + 1 \equiv 0 \pmod{p}. \tag{8}$$

Доведення. Якщо $p = 2$, то теорема очевидна. Нехай p — непарне просте число. За теоремою Ферма для будь-якого числа x, взаємно простого з p, справедлива конгруенція

$$x^{p-1} \equiv 1 \pmod{p}. \tag{9}$$

Отже, конгруенція (9) має $p - 1$ розв'язків, бо задовольняється лишками 1, 2, ..., $p - 1$. Конгруенція $(p - 1)$-го степеня

$$g(x) = (x-1)(x-2)\ \ldots\ [x-(p-1)] \equiv 0 \pmod{p} \tag{10}$$

також має ті самі розв'язки, і її задовольняють числа 1, 2, ..., $p - 1$. З (9) і (10) побудуємо конгруенцію

$$[x^{p-1} - 1] - (x-1)(x-2)\ \ldots\ [x-(p-1)] \equiv 0 \pmod{p}. \tag{11}$$

Очевидно, її задовольняють також числа 1, 2, ..., $p - 1$, тобто конгруенція (11) має $p - 1$ розв'язків. Але в лівій частині конгруенції (11) міститься многочлен степеня $p - 2$. Таким чином, (11) є конгруенцією $(p - 2)$-го степеня, яка має $p - 1$ розв'язків і тому вона тотожна. Беручи в ній $x = 0$, дістанемо

$$1 + (-1)(-2)\ \ldots\ [-(p-1)] \equiv 0 \pmod{p},$$

або

$$(-1)^{p-1}(p-1)! + 1 \equiv 0 \pmod{p}.$$

При непарному p маємо

$$(p-1)! + 1 \equiv 0 \pmod{p}.$$

Отже, теорема правильна для будь-якого простого числа p.

18.3. Квадратичні лишки і нелишки. Розглянемо, як найпростіші, двочленні конгруенції степеня $n > 1$ виду

$$x^n \equiv a \pmod{p}, \tag{12}$$

де p — просте число і $(a, p) = 1$.

Якщо така конгруенція має розв'язок, то число a називається *лишком степеня n за простим модулем p*, якщо ні,— то *нелишком степеня n за модулем p*. При $n = 2$ лишки і нелишки називаються *квадратичними*, при $n = 3$ — *кубічними*, при $n = 4$ — *біквадратними*. Зробимо кілька зауважень відносно конгруенції (12). Перш за все зауважимо, що конгруенція (12) розглядається при умові $(a, p) = 1$, тому число нуль не належить ні до лишків, ні до нелишків. Це пояснюється тим, що конгруенція виду $x^n \equiv 0 \pmod{p}$ має єдиний розв'язок $x \equiv 0 \pmod{p}$ і внаслідок тривіальності цього факту такий випадок виключається при розгляді конгруенції (12).

У конгруенції (12) достатньо вважати, що число a задовольняє умову $1 \leqslant a < p$, бо будь-яка конгруенція $x^n \equiv b \pmod{p}$, де $b > p$, завжди зводиться до конгруенції з невід'ємними коефіцієнтами, меншими за p.

[1] Джон Вільсон (1741—1793) — англійський математик.

Якщо в конгруенції (12) модуль $p = 2$, то фактично розглядається також тривіальний випадок: $a = 0$, або $a = 1$. Отже конгруенцію $x^n \equiv a \pmod{2}$ задовольняє або клас непарних чисел $K_1^{(2)}$, якщо a — непарне число, або клас всіх парних чисел $K_0^{(2)}$, якщо a — парне число. Тому вважатимемо далі, що $p > 2$, і перейдемо до розгляду найпростішої квадратної двочленної конгруенції

$$x^2 \equiv a \pmod{p}, \quad (a, p) = 1, \quad p > 2. \tag{13}$$

Розглянемо спочатку приклад. Знайдемо всі квадратичні лишки і нелишки за модулем $p = 17$. Множина чисел

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 \tag{14}$$

являє собою ЗСЛ за модулем $p = 17$. Якщо ж цю множину чисел розглянути з точки зору можливості існування розв'язків квадратної конгруенції (13), то частина чисел множини (14) є квадратичними лишками, а частина їх — квадратичними нелишками.

Враховуючи, що

$1^2 \equiv 1 \pmod{17}$, $5^2 \equiv 8 \pmod{17}$, $9^2 \equiv 13 \pmod{17}$, $13^2 \equiv 16 \pmod{17}$,
$2^2 \equiv 4 \pmod{17}$, $6^2 \equiv 2 \pmod{17}$, $10^2 \equiv 15 \pmod{17}$, $14^2 \equiv 9 \pmod{17}$,
$3^2 \equiv 9 \pmod{17}$, $7^2 \equiv 15 \pmod{17}$, $11^2 \equiv 2 \pmod{17}$, $15^2 \equiv 4 \pmod{17}$,
$4^2 \equiv 16 \pmod{17}$, $8^2 \equiv 13 \pmod{17}$, $12^2 \equiv 8 \pmod{17}$, $16^2 \equiv 1 \pmod{17}$,

бачимо, що квадратичними лишками за модулем 17 є окантовані рамкою числа в множині

$$\boxed{1}, \boxed{2}, 3, \boxed{4}, 5, 6, 7, \boxed{8}, \boxed{9}, 10, 11, 12, \boxed{13}, 14, \boxed{15}, \boxed{16}.$$

Неокантовані рамкою числа є квадратичними нелишками за модулем 17.

Відповідно до сказаного конгруенцію

$$x^2 \equiv 2 \pmod{17}$$

задовольняють числа $x = 6$ і $x = 11$, і тому вона має два розв'язки: $K_6^{(17)}$ і $K_{11}^{(17)}$. Мають по два розв'язки і конгруенції $x^2 \equiv 15 \pmod{17}$, $x^2 \equiv 8 \pmod{17}$ та інші. А конгруенції $x^2 \equiv 3 \pmod{17}$, $x^2 \equiv 7 \pmod{17}$, $x^2 \equiv 11 \pmod{17}$ та інші, в правій частині яких стоять квадратичні нелишки за модулем 17, розв'язків не мають.

Проведений аналіз свідчить про те, що квадратні конгруенції або мають два розв'язки, або не мають їх зовсім. І половина лишків ЗСЛ за модулем $p = 17$ є квадратичними лишками, а половина — квадратичними нелишками.

Доведемо це в загальному випадку.

Теорема 4. *Якщо a — квадратичний лишок за модулем p, $(a, p) = 1$, $p > 2$, то квадратна конгруенція*

$$x^2 \equiv a \pmod{p} \tag{15}$$

має два розв'язки.

Доведення. Оскільки a — квадратичний лишок, то конгруенція має розв'язок. Нехай це клас чисел $K_{x_0}^{(p)}$, де x_0 — один з лишків

ЗСЛ за модулем p, тому $(x_0, p) = 1$. Число x_0 задовольняє (15): $x_0^2 \equiv a \pmod p$. Але тоді і $-x_0$ задовольняє (15), бо $(-x_0)^2 \equiv a \pmod p$. При цьому x_0 не конгруентне з $-x_0$ за модулем p. Справді, якби $x_0 \equiv -x_0 \pmod p$, то $2x_0 \equiv 0 \pmod p$, що неможливо, бо і $(2, p) = 1$, і $(x_0, p) = 1$, тому $2x_0$ не може бути кратним p. Таким чином, $K_{x_0}^{(p)}$ і $K_{-x_0}^{(p)}$ — два різні розв'язки квадратної конгруенції (15). Оскільки квадратна конгруенція не може мати більше двох розв'язків, то інших розв'язків не існує. Теорему доведено.

Теорема 5. *Для будь-якого простого числа $p > 2$ половина лишків ЗСЛ є квадратичними лишками, а половина — квадратичними нелишками.*

Доведення. Просте число $p > 2$ непарне і тому $\frac{p-1}{2}$ — ціле число. ЗСЛ за модулем p, складена з абсолютно найменших лишків, є

$$\pm 1,\ \pm 2,\ \pm 3,\ \ldots,\ \pm \frac{p-1}{2}. \tag{16}$$

Ці числа є представниками $p-1$ класів лишків $K_{-\frac{p-1}{2}}^{(p)}, \ldots, K_{-2}^{(p)}, K_{-1}^{(p)}, K_1^{(p)}, K_2^{(p)}, \ldots, K_{\frac{p-1}{2}}^{(p)}$. Одночасно деякі з чисел множини (16) є і квадратичними лишками за модулем p.

Для того щоб дослідити, якими саме числами множини (16) задовольняється конгруенція

$$x^2 \equiv a \pmod p, \quad (a, p) = 1, \quad p > 2,$$

досить підставити їх у конгруенцію. При піднесенні до квадрата $p-1$ чисел (16) перейдуть у $\frac{p-1}{2}$ квадратів чисел

$$1^2,\ 2^2,\ 3^2,\ \ldots,\ \left(\frac{p-1}{2}\right)^2. \tag{17}$$

Якщо задане число a — квадратичний лишок за модулем p, то одне з чисел (17) мусить бути конгруентним з a. Покажемо, що числа (17) неконгруентні між собою. Справді, нехай серед чисел (17) є числа s^2 і t^2, де $1 \leqslant s < t \leqslant \frac{p-1}{2}$. Тоді, якщо

$$t^2 \equiv s^2 \pmod p, \tag{18}$$

то $t^2 - s^2 \equiv 0 \pmod p$ і, отже, $(t-s)(t+s) \equiv 0 \pmod p$. Остання конгруенція вимагає, щоб або $t - s$, або $t + s$ були кратними p. Але $t - s < \frac{p-1}{2}$, а $t + s < p - 1$, тому ні $t - s$, ні $t + s$ не можуть ділитися на більше за них число p. Отже, конгруенція (18) неможлива і тому $\frac{p-1}{2}$ чисел множини (17) належать різним класам лишків за модулем p. Всі вони є квадратичними лишками, їх усього $\frac{p-1}{2}$. Зрозуміло, що вони конгруентні певним $\frac{p-1}{2}$ числам ЗСЛ. Тому серед

ЗСЛ є $\frac{p-1}{2}$ квадратичних лишків і $\frac{p-1}{2}$ квадратичних нелишків. Теорему доведено.

18.4. Критерій Ейлера. При досить великих модулях p множина ЗСЛ (16) налічує велику кількість членів, і тоді процес підстановки цих чисел у конгруенцію (15) для знаходження її розв'язків стає громіздким. Тому перед розв'язуванням конгруенції (15) важливо наперед встановити, чи є число a квадратичним лишком. Відповідь на це запитання дають *критерій Ейлера* і *символ Лежандра*.

Теорема 6. (Критерій Ейлера). *При простому $p > 2$ число a є квадратичним лишком за модулем p тоді і тільки тоді, коли*

$$a^{\frac{p-1}{2}} \equiv 1 \pmod p,$$

і квадратичним нелишком тоді і тільки тоді, коли

$$a^{\frac{p-1}{2}} \equiv -1 \pmod p.$$

Доведення. Якщо $(a, p) = 1$, то за теоремою Ферма

$$a^{p-1} \equiv 1 \pmod p,$$

або

$$a^{p-1} - 1 \equiv 0 \pmod p.$$

Оскільки число $p - 1$ парне, то, розглядаючи вираз $a^{p-1} - 1$ як різницю квадратів, маємо

$$(a^{\frac{p-1}{2}} - 1)(a^{\frac{p-1}{2}} + 1) \equiv 0 \pmod p. \tag{19}$$

Звідси випливає, що якийсь із множників у лівій частині конгруенції є кратним модулю p. Обидва ці множники одночасно ділитися на p не можуть, бо тоді б і їх різниця

$$(a^{\frac{p-1}{2}} + 1) - (a^{\frac{p-1}{2}} - 1) = 2$$

теж повинна була б ділитися на p, що неможливо, бо $p > 2$.

Нехай a — квадратичний лишок. Тоді конгруенція (15) має розв'язок $K_{x_0}^{(p)}$. Зокрема,

$$x_0^2 \equiv a \pmod p$$

і далі

$$x_0^{2\left(\frac{p-1}{2}\right)} \equiv a^{\frac{p-1}{2}} \pmod p,$$

тобто

$$x_0^{p-1} \equiv a^{\frac{p-1}{2}} \pmod p.$$

Оскільки $x_0 \in$ ЗСЛ за модулем p і тому $(x_0, p) = 1$, то за теоремою Ферма

$$x_0^{p-1} \equiv 1 \pmod p.$$

Порівнюючи дві останні конгруенції, маємо

$$a^{\frac{p-1}{2}} \equiv 1 \pmod{p}. \qquad (20)$$

Отже, будь-який квадратичний лишок a за модулем p задовольняє конгруенцію (20). Оскільки всіх квадратичних лишків за модулем p є $\frac{p-1}{2}$, а степінь конгруенції (20) також дорівнює $\frac{p-1}{2}$, то конгруенцію (20) задовольняють тільки квадратичні лишки за модулем p — ніяких інших розв'язків вона більше не має. Очевидно, і навпаки — якщо справджується конгруенція (20), то a є квадратичним лишком за модулем p.

Звідси відразу ж випливає і друге твердження критерію Ейлера. З конгруенції (19) видно, що коли a є квадратичним нелишком, то на p ділиться другий множник $a^{\frac{p-1}{2}} + 1$, тобто

$$a^{\frac{p-1}{2}} \equiv -1 \pmod{p}. \qquad (21)$$

Конгруенція (21) справедлива для всіх $\frac{p-1}{2}$ квадратичних нелишків і ніяких інших розв'язків вона не має. Тому і навпаки: якщо має місце (21), то a — квадратичний нелишок за модулем p. Теорему доведено.

Приклад. Встановимо, чи має розв'язки квадратна конгруенція

$$x^2 \equiv 13 \pmod{17}.$$

За критерієм Ейлера знаходимо число $13^{\frac{17-1}{2}}$ і встановлюємо, що за модулем 17

$$13^{\frac{17-1}{2}} = 13^8 \equiv (-4)^8 \equiv 16^4 \equiv (-1)^4 \equiv 1 \pmod{17}.$$

Отже, виконується конгруенція (20) і тому задана конгруенція має розв'язки.

18.5. Символ Лежандра[1]. При великих p і a користуватися критерієм Ейлера практично майже неможливо. Наприклад, для доведення того, що конгруенція

$$x^2 \equiv 579 \pmod{821}$$

не має розв'язків, треба встановити, що справедлива конгруенція

$$579^{\frac{821-1}{2}} \equiv -1 \pmod{821},$$

а це, звичайно, справа дуже складна.

Значно ефективнішим є спосіб, який грунтується на так званому символі Лежандра $\left(\frac{a}{p}\right)$. Читається символ $\left(\frac{a}{p}\right)$ так: «a відносно p».

[1] Андрієн Марі Лежандр (1752—1833) — видатний французький математик.

Означення. *Символ Лежандра* $\left(\frac{a}{p}\right)$ *визначається для всіх цілих чисел* a, *які не діляться на просте число* $p > 2$ *рівністю*

$$\left(\frac{a}{p}\right) = \begin{cases} +1, & \text{якщо } a \text{ в квадратичним лишком за модулем } p, \\ -1, & \text{якщо } a \text{ в квадратичним нелишком за модулем } p. \end{cases}$$

Використовуючи критерій Ейлера, очевидно, маємо

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}. \qquad (22)$$

Властивості символу Лежандра:

1. *Якщо* $a \equiv a_1 \pmod{p}$, *то* $\left(\frac{a}{p}\right) = \left(\frac{a_1}{p}\right)$.

Справді, всі числа того самого класу $K_a^{(p)}$ є одночасно або квадратичними лишками, або квадратичними нелишками за модулем p. Вважаючи, що $a_1 = a + pt$, цю властивість можна записати ще так:

$$\left(\frac{a+pt}{p}\right) = \left(\frac{a}{p}\right).$$

Наприклад,

$$\left(\frac{400}{151}\right) = \left(\frac{2 \cdot 151 + 98}{151}\right) = \left(\frac{98}{151}\right).$$

2. $\left(\frac{a^2}{p}\right) = 1$.

Справді, на підставі (22) і теореми Ферма дістанемо

$$\left(\frac{a^2}{p}\right) = (a^2)^{\frac{p-1}{2}} \pmod{p} \equiv a^{p-1} \pmod{p} \equiv 1 \pmod{p} \Rightarrow \left(\frac{a^2}{p}\right) = 1.$$

Нарешті рівність $\left(\frac{a^2}{p}\right) = 1$ очевидна, бо будь-яке число a^2 є квадратичним лишком за модулем p, тобто конгруенція $x^2 \equiv a^2 \pmod{p}$ завжди має розв'язок.

3. $\left(\frac{1}{p}\right) = 1$.

Це наслідок 2.

4. $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

Це безпосередньо випливає з конгруенції (22). Якщо в ній вважати $a = -1$, то матимемо

$$\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \pmod{p}.$$

Але і ліва, і права частини конгруенції є ± 1. При $p > 2$ числа -1 і $+1$ конгруентними бути не можуть. Тому в обох частинах конгруенції міститься або 1, або -1, тобто маємо рівність

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}.$$

Властивість 4 можна конкретизувати. Справді, просте число $p > 2$ є непарним числом. А будь-яке непарне число можна подати у вигляді $p = 4m + 1$ або $p = 4m + 3$.

Якщо $p = 4m + 1$, то на основі 4

$$\left(\frac{-1}{p}\right) = \left(\frac{-1}{4m+1}\right) = (-1)^{\frac{(4m+1)-1}{2}} = 1.$$

Якщо $p = 4m + 3$, то

$$\left(\frac{-1}{p}\right) = \left(\frac{-1}{4m+3}\right) = (-1)^{\frac{(4m+3)-1}{2}} = -1.$$

Отже, справедливим є таке твердження: *Число -1 є квадратичним лишком за модулем p, якщо p можна подати у вигляді $p = 4m + 1$, і квадратичним нелишком за модулем p, якщо p має вигляд $p = 4m + 3$.*

Наприклад, число $p = 17$ має вигляд $17 = 4 \cdot 4 + 1$, тому -1 є квадратичним лишком за модулем 17. А число $p = 23$ має вигляд $23 = 5 \cdot 4 + 3$, тому -1 є квадратичним нелишком за модулем 23.

5. $\left(\frac{a_1 \cdot a_2 \cdot \ldots \cdot a_k}{p}\right) = \left(\frac{a_1}{p}\right) \cdot \left(\frac{a_2}{p}\right) \cdot \ldots \cdot \left(\frac{a_k}{p}\right)$.

Справді, за (22) маємо

$$\left(\frac{a_1}{p}\right) \equiv a_1^{\frac{p-1}{2}} \pmod p,$$

$$\left(\frac{a_2}{p}\right) \equiv a_2^{\frac{p-1}{2}} \pmod p,$$

$$\ldots\ldots\ldots\ldots$$

$$\left(\frac{a_k}{p}\right) \equiv a_k^{\frac{p-1}{2}} \pmod p.$$

Перемножаючи всі ці конгруенції, дістанемо

$$\left(\frac{a_1}{p}\right) \cdot \left(\frac{a_2}{p}\right) \cdot \ldots \cdot \left(\frac{a_k}{p}\right) \equiv a_1^{\frac{p-1}{2}} \cdot a_2^{\frac{p-1}{2}} \cdot \ldots \cdot a_k^{\frac{p-1}{2}} \pmod p \equiv$$

$$\equiv (a_1 \cdot a_2 \cdot \ldots \cdot a_k)^{\frac{p-1}{2}} \pmod p \equiv \left(\frac{a_1 \cdot a_2 \cdot \ldots \cdot a_k}{p}\right),$$

що й треба було довести.

Наслідок 1. $\left(\frac{ab^2}{p}\right) = \left(\frac{a}{p}\right)$.

Справді, на основі 5 і 2 маємо $\left(\frac{ab^2}{p}\right) = \left(\frac{a}{p}\right) \cdot \left(\frac{b^2}{p}\right) = \left(\frac{a}{p}\right)$.

Наслідок 2. $\left(\frac{a^n}{p}\right) = \left(\frac{a}{p}\right)^n$.

Цю властивість дістаємо з 5 при рівності $a_1 = a_2 = \ldots = a_n$.

Наслідок 3. *Якщо число a не ділиться на просте число $p > 2$ і в канонічному розкладі на добуток простих множників має вигляд*

$$a = q_1^{l_1} \cdot q_2^{l_2} \cdot \ldots \cdot q_k^{l_k},$$

то

$$\left(\frac{q_1^{l_1} \cdot q_2^{l_2} \cdot \ldots \cdot q_k^{l_k}}{p}\right) = \left(\frac{q_1}{p}\right)^{l_1} \cdot \left(\frac{q_2}{p}\right)^{l_2} \cdot \ldots \cdot \left(\frac{q_k}{p}\right)^{l_k}. \qquad (23)$$

Останній наслідок зводить питання обчислення символу Лежандра $\left(\frac{a}{p}\right)$ до обчислення символів Лежандра виду $\frac{q}{p}$, де обидва числа і p, і q є простими числами, більшими за 2.

6. $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

За модулем 8 всі непарні числа можна подати як числа виду

$$8k + 1, \quad 8k + 3, \quad 8k + 5, \quad 8k + 7,$$

або

$$8k \pm 1, \quad 8k \pm 3.$$

Оскільки при $p = 8k \pm 1$ число $\frac{p^2-1}{8}$ є парним, а при $p = 8k \pm 3$ є непарним, то звідси випливає таке твердження: *Число 2 є квадратичним лишком за простим модулем p, якщо p подано у вигляді суми $p = 8k \pm 1$, і квадратичним нелишком за модулем p, якщо p подано у вигляді $p = 8k \pm 3$.*

7. *Закон взаємності квадратичних лишків:*

$$\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} \cdot \left(\frac{q}{p}\right).$$

Доведення властивостей 6 і 7 досить громіздкі, тому їх приводити не будемо. Зазначимо лише, що ці властивості важливі і широко застосовуються при знаходженні числового значення символу Лежандра $\left(\frac{p}{q}\right)$.

Приклад. Знайти числове значення символу Лежандра. $\left(\frac{342}{677}\right)$. Оскільки $342 = 2 \cdot 3^2 \cdot 19$, то

$$\left(\frac{342}{677}\right) = \left(\frac{2 \cdot 3^2 \cdot 19}{677}\right) = \left(\frac{2}{677}\right) \cdot \left(\frac{3^2}{677}\right) \cdot \left(\frac{19}{677}\right) = \left(\frac{2}{677}\right) \cdot \left(\frac{19}{677}\right).$$

Далі

$$\left(\frac{2}{677}\right) = (-1)^{\frac{677^2-1}{8}} = (-1)^{57\,291} = -1;$$

$$\left(\frac{19}{677}\right) = (-1)^{\frac{677-1}{2} \cdot \frac{19-1}{2}} \left(\frac{677}{19}\right) = \left(\frac{677}{19}\right) = \left(\frac{12}{19}\right) = \left(\frac{3 \cdot 2^2}{19}\right) = \left(\frac{3}{19}\right) =$$

$$= (-1)^{\frac{3-1}{2} \cdot \frac{19-1}{2}} \left(\frac{19}{3}\right) = -\left(\frac{1}{3}\right) = -1.$$

Отже,

$$\left(\frac{342}{677}\right)=\left(\frac{2}{677}\right)\cdot\left(\frac{19}{677}\right)=(-1)\cdot(-1)=1.$$

Звідси робимо висновок, що 342 є квадратичним лишком за модулем 677, або, що те ж саме, конгруенція $x^2 \equiv 342 \pmod{677}$ має розв'язки.

18.6. Розв'язування квадратичних конгруенцій. Критерій Ейлера і символ Лежандра $\left(\frac{a}{p}\right)$ дають можливість встановити, чи є число a квадратичним лишком за модулем p, тобто встановити, чи має розв'язки квадратна конгруенція

$$x^2 \equiv a \pmod{p}. \qquad (25)$$

Розв'язують цю конгруенцію *простими пробами,* підставляючи в (25) замість x числа ЗСЛ за модулем p, або *за допомогою спеціальних таблиць.*

При великих p метод розв'язування за допомогою проб внаслідок своєї громіздкості стає явно непрактичним. Тільки в окремих випадках можна знайти розв'язки конгруенції (25) в загальному вигляді. Нехай, наприклад, a — квадратичний лишок за модулем p. Тоді за критерієм Ейлера маємо

$$a^{\frac{p-1}{2}} \equiv 1 \pmod{p}.$$

Якщо число p має вигляд $p = 4k + 3$, то дістанемо

$$a^{2k+1} \equiv 1 \pmod{p}.$$

Помноживши на a обидві частини конгруенції, матимемо далі

$$a^{2k+2} \equiv a \pmod{p},$$

або

$$(a^{k+1})^2 \equiv a \pmod{p}.$$

Очевидно, $x \equiv \pm a^{k+1} \pmod{p}$ дає нам класи чисел $K^{(p)}_{a^{k+1}}$ і $K^{(p)}_{-a^{k+1}}$, які є розв'язками конгруенції (15). Отже, у даному випадку розв'язки знайдені в загальному вигляді.

Приклад. Знайдемо розв'язки квадратної конгруенції

$$x^2 \equiv 10 \pmod{41}.$$

Насамперед, встановимо, чи має розв'язки ця конгруенція. Знаходимо

$$\left(\frac{10}{41}\right)=\left(\frac{2}{41}\right)\cdot\left(\frac{5}{41}\right)=(-1)^{\frac{41^2-1}{8}}\cdot\left(\frac{5}{41}\right)=(-1)^{210}\cdot\left(\frac{5}{41}\right)=\left(\frac{5}{41}\right)=$$

$$=(-1)^{\frac{5-1}{2}\cdot\frac{41-1}{2}}\left(\frac{41}{5}\right)=\left(\frac{1}{5}\right)=1.$$

Отже, $\left(\frac{10}{41}\right)=1$ і тому задана конгруенція має розв'язки. Підставляючи в конгруенцію послідовно числа ЗСЛ, дістанемо що її задовольняють числа $x = \pm 16$, або, що те ж саме, $x = 16$ і $x = 25$. Отже, розв'язками заданої конгруенції є класи чисел $K^{(41)}_{16}$ і $K^{(41)}_{25}$.

19.1. Показники за модулем. Розглянемо множину цілих додатних степенів числа a:

$$a,\ a^2,\ a^3,\ a^4,\ \ldots \qquad (1)$$

з точки зору їх конгруентності за деяким модулем $m > 1$, вважаючи, що $(a, m) = 1$. За теоремою Ейлера маємо:

$$a^{\varphi(m)} \equiv 1 \pmod{m},$$

і, отже,

$$a^{\varphi(m)k} \equiv 1 \pmod{m}$$

при будь-якому цілому додатному k. Таким чином, серед степенів (1) числа a знайдеться нескінченна кількість чисел, конгруентних з 1 за модулем m.

Найменше натуральне число δ, для якого справедлива конгруенція $a^\delta \equiv 1 \pmod{m}$, називається показником, до якого належить число a, за модулем m. Інакше число δ позначають ще символом $P_m(a)$. Число $\delta \leqslant \varphi(m)$, бо може бути і меншим за значення $\varphi(m)$ функції Ейлера. У випадку, коли $\delta = \varphi(m)$, число a називається первісним коренем за модулем m.

Приклади. 1. Знайти $P_m(a)$, якщо $a = 3$, $m = 8$. Для модуля $m = 8$ ЗСЛ $= \{1, 3, 5, 7\}$, тому $\varphi(8) = 4$. Отже, $3^4 \equiv 1 \pmod{8}$. Проте, як виявляється, $\delta = P_8(3) < \varphi(8)$. Справді, аналізуючи числа ПСЛ, бачимо, що вже

$$3^2 \equiv 1 \pmod{8}.$$

Отже, $P_8(3) = 2$, тобто число 3 належить до показника 2 за модулем 8.

2. Знайти $P_m(a)$, якщо $a = 3$, $m = 10$. Для $m = 10$ число $\varphi(10) = 4$. Отже,

$$3^4 \equiv 1 \pmod{10}.$$

Аналізуючи степені чисел 3^1, 3^2, 3^3, бачимо, що вони не конгруентні 1 за модулем 10. Тому $\delta = P_{10}(3) = 4$, тобто число 3 належить показнику 4 за модулем 10. Отже, число 3 є первісним коренем за модулем 10.

Зауважимо, що вимога $(a, m) = 1$ є істотною. У випадку, коли числа a і m не є взаємно простими, конгруенція

$$a^\delta \equiv 1 \pmod{m}$$

не виконується ні при яких δ (властивість 14 п. 15.2).

Теорема 1. *Якщо $a_1 \equiv a_2 \pmod{m}$, то числа a_1 і a_2 належать до того самого показника за цим модулем.*

Доведення. Припустимо супротивне, а саме: $P_m(a_1) = \delta_1$, $P_m(a_2) = \delta_2$ і $\delta_1 \neq \delta_2$. Тоді

$$[a_1 \equiv a_2 \pmod{m} \wedge a_1^{\delta_1} \equiv 1 \pmod{m}] \Rightarrow [a_2^{\delta_1} \equiv 1 \pmod{m}].$$

Тому показник δ_2, до якого належить число a_2, або дорівнює δ_1, або може бути меншим за δ_1: $\delta_2 \leqslant \delta_1$. Аналогічно

$$[a_1 \equiv a_2 \pmod{m} \wedge a_2^{\delta_2} \equiv 1 \pmod{m}] \Rightarrow [a_1^{\delta_2} \equiv 1 \pmod{m}],$$

тому $\delta_1 \leqslant \delta_2$. З нерівностей $\delta_1 \leqslant \delta_2$ і $\delta_2 \leqslant \delta_1$ випливає єдиний висновок: $\delta_1 = \delta_2$, що й треба було довести.

Наслідок. *Усі числа одного класу $K_a^{(m)}$ належать тому самому показнику δ за модулем m.*

Таким чином, говорять не про числа, а *про класи чисел, які належать до даного показника за модулем m*. У прикладі 1 до показника $\delta = 2$ належить клас $K_3^{(3)}$, а в прикладі 2 до показника $\delta = 4$ належить клас $K_3^{(10)}$. При цьому клас $K_3^{(10)}$ є, очевидно, *класом первісних коренів* за модулем m.

19.2. Властивості показників за модулем.

1. *Якщо $(a, m) = 1$ і δ — показник, до якого належить число a за модулем m, то серед степенів*

$$1 = a^0,\ a,\ a^2,\ a^3,\ \ldots,\ a^{\delta-1} \tag{2}$$

немає чисел, конгруентних між собою за модулем m.

Доведення. Припустимо супротивне. Нехай при деяких натуральних k і l, причому $0 \leqslant k < l < \delta$, маємо

$$a^l = a^k \pmod m.$$

Враховуючи, що $(a, m) = 1$, маємо далі

$$a^{l-k} \equiv 1 \pmod m.$$

При цьому $l - k < \delta$, що неможливо, оскільки δ є найменше з натуральних чисел, для яких справджується конгруенція $a^\delta \equiv 1 \pmod m$. Властивість доведено.

Наслідок. *Якщо a — первісний корінь за модулем m, тобто $\delta = \varphi(m)$, то множина степенів*

$$1 = a^0,\ a,\ a^2,\ a^3,\ \ldots,\ a^{\varphi(m)-1} \tag{3}$$

є ЗСЛ за модулем m.

Справді, у цій множині є $\varphi(m)$ чисел. Усі вони взаємно прості з числом m і всі вони неконгруентні між собою за модулем m. Тим самим вони належать до різних класів за модулем m і утворюють ЗСЛ за модулем m. Якщо $m = p$ (просте число), то (3) перетворюється в таку сукупність чисел:

$$1,\ a,\ a^2,\ a^3,\ \ldots,\ a^{p-2},$$

яка являє собою зведену систему лишків за модулем p.

2. *Якщо $\delta = P_m(a)$, то*

$$[a^{k_1} \equiv a^{k_2} \pmod m] \Leftrightarrow [k_1 \equiv k_2 \pmod \delta].$$

Доведення. *Необхідність.* За теоремою про ділення чисел з остачею маємо при діленні на δ $k_1 = \delta q_1 + r_1$ $k_2 = \delta q_2 + r_2$, причому $0 \leqslant r_1 < \delta$, $0 \leqslant r_2 < \delta$. Враховуючи, що

$$a^\delta \equiv 1 \pmod m,$$

маємо

$$a^{k_1} \equiv a^{\delta q_1 + r_1} \pmod m \equiv a^{r_1} \pmod m,$$
$$a^{k_2} \equiv a^{\delta q_2 + r_2} \pmod m \equiv a^{r_2} \pmod m.$$

Отже,

$$[a^{k_1} \equiv a^{k_2} \pmod m] \Rightarrow [a^{r_1} \equiv a^{r_2} \pmod m].$$

Але при $r_1, r_2 < \delta$ це може бути лише у випадку $r_1 = r_2$. Таким чином, при діленні на δ чисел k_1 і k_2 дістаємо рівні остачі. Це означає, що $k_1 \equiv k_2 \pmod \delta$. Необхідність доведено.

Достатність. Оскільки $k_1 \equiv k_2 \pmod \delta$, то $k_1 = k_2 + \delta t$. Враховуючи, що $a^\delta \equiv 1 \pmod m$, дістаємо $a^{k_1} \equiv a^{k_2 + \delta t} \pmod m \equiv a^{k_2} \pmod m$, що й треба було довести.

Наслідок 1. *Якщо число a належить до показника δ за модулем m, і $a^k \equiv 1 \pmod m$, то $k \vdots \delta$.*

Справді, на основі 2

$$[a^\delta \equiv 1 \pmod m \wedge a^k \equiv 1 \pmod m] \Rightarrow [a^\delta \equiv a^k \pmod m] \Rightarrow$$
$$\Rightarrow [k \equiv \delta \pmod \delta] \Rightarrow [k \equiv 0 \pmod \delta],$$

тобто $k \vdots \delta$.

Наслідок 2. *Показник δ, до якого належить число a за модулем m, є дільником числа $\varphi(m)$.*

Справді, $[a^{\varphi(m)} \equiv 1 \pmod m \wedge a^\delta \equiv 1 \pmod m] \Rightarrow [a^{\varphi(m)} \equiv a^\delta \pmod m] \Rightarrow$ $\Rightarrow [\varphi(m) \equiv \delta \pmod \delta] \Rightarrow [\varphi(m) \equiv 0 \pmod \delta]$, тобто $\varphi(m) \vdots \delta$, що й треба було довести. Цей наслідок дає можливість спростити знаходження показника δ, до якого належить число a за модулем m.

Приклад. Знайти $P_{20}(7)$.

Оскільки $\varphi(20) = 8$, то для знаходження δ треба дослідити тільки степені $7^1, 7^2, 7^4, 7^8$, показники яких є дільниками числа $8 = \varphi(20)$. Встановлюємо, що $\delta = P_{20}(7) = 4$.

19.3. Добуток показників. Як буде показано далі, первісні корені за модулем m існують обов'язково при умові, коли $m = p$ є простим числом; при цьому первісних коренів буде $\varphi(p-1)$ класів. Доведемо спочатку таку теорему.

Теорема 2. *Якщо число a_1 належить до показника δ_1 за модулем m, а число a_2 — до показника δ_2 за модулем m і $(\delta_1, \delta_2) = 1$, то добуток чисел $a_1 \cdot a_2$ належить до добутку показників $\delta_1 \cdot \delta_2$ за модулем m.*

Доведення. Оскільки $P_m(a_1) = \delta_1$ і $P_m(a_2) = \delta_2$, то

$$a_1^{\delta_1} \equiv 1 \pmod m, \quad a_2^{\delta_2} \equiv 1 \pmod m. \tag{4}$$

Нехай тепер $\delta = P_m(a_1 a_2)$, тобто δ є найменшим натуральним числом, при якому виконується конгруенція

$$(a_1 a_2)^\delta \equiv 1 \pmod m. \tag{5}$$

Треба довести, що $\delta = \delta_1 \cdot \delta_2$. Цей результат ми дістанемо, якщо покажемо, що одночасно справедливі два співвідношення: $\delta \vdots \delta_1\delta_2$ і $\delta_1\delta_2 \vdots \delta$. З (5) і (4) випливає справедливість таких конгруенцій

$$[(a_1a_2)^{\delta\delta_1} \equiv 1 \pmod m \Rightarrow [a_1^{\delta\delta_2} \cdot a_2^{\delta\delta_2} \equiv 1 \pmod m] \Rightarrow [a_1^{\delta\delta_2} \equiv 1 \pmod m]. \tag{6}$$

Порівнюючи останню конгруенцію з першою конгруенцією (4), дістаємо, що $\delta\delta_2 \vdots \delta_1$. Враховуючи, що $(\delta_1, \delta_2) = 1$, робимо далі висновок, що $\delta \vdots \delta_1$.

Розглядаючи аналогічно до (6) наслідок з конгруенції

$$(a_1a_2)^{\delta\delta_1} \equiv 1 \pmod m,$$

дістанемо, що $\delta : \delta_2$. Оскільки $(\delta_1, \delta_2) = 1$, робимо висновок, що *число* δ *ділиться на добуток* $\delta_1 \cdot \delta_2$.

Розглянемо тепер очевидну конгруенцію

$$(a_1a_2)^{\delta_1\delta_2} \equiv 1 \pmod m.$$

З неї і конгруенції (5) випливає

$$[(a_1a_2)^{\delta_1\delta_2} \equiv 1 \pmod m \wedge (a_1a_2)^{\delta} \equiv 1 \pmod m] \Rightarrow [\delta_1\delta_2 \equiv \delta \pmod \delta],$$

тобто $\delta_1\delta_2 : \delta$.

З двох співвідношень $\delta : \delta_1\delta_2$ і $\delta_1\delta_2 : \delta$ випливає $\delta = \delta_1\delta_2$. Теорему доведено.

Твердження теореми можна поширити на добуток n чисел.

Наслідок. *Якщо числа* $a_1, a_2, \ldots, a_n$ *належать за модулем* m *відповідно до показників* $\delta_1, \delta_2, \ldots, \delta_n$, *які попарно взаємно прості між собою, то*

$$P_m(a_1 \cdot a_2 \cdot \ldots \cdot a_n) = P_m(a_1) \cdot P_m(a_2) \cdot \ldots \cdot P_m(a_n) = \delta_1\delta_2 \ldots \delta_n, \quad (7)$$

тобто показник, до якого належить добуток чисел $a_1 \cdot a_2 \cdot \ldots \cdot a_n$ *за модулем* m, *дорівнює добутку показників, до яких належать числа* $a_1, a_2, \ldots, a_n$ *за модулем* m.

Доведення. Доведення проведемо методом математичної індукції.

Припустимо, що твердження справедливе для $n = k \leqslant 2$

$$P_m(a_1 \cdot a_2 \cdot \ldots \cdot a_k) = P_m(a_1) \cdot P_m(a_2) \cdot \ldots \cdot P_m(a_k), \quad (8)$$

і доведемо його для $n = k + 1$. За доведеною теоремою при умові, що всі $P_m(a_i)$ попарно взаємно прості між собою, маємо

$$P_m(a_1a_2 \ldots a_ka_{k+1}) = P_m((a_1a_2 \ldots a_k)a_{k+1}) =$$

$$= P_m(a_1a_2 \ldots a_k) \cdot P_m(a_{k+1}) = P_m(a_1) \cdot P_m(a_2) \cdot \ldots \cdot P_m(a_k) \cdot P_m(a_{k+1}),$$

що й треба було довести.

19.4. Існування первісних коренів. В попередньому параграфі вже зазначалося, що конгруенції

$$a^{\delta} \equiv 1 \pmod m$$

можуть не мати розв'язків. Це має місце, коли $(a, m) = d > 1$. Якщо ж $(a, m) = 1$, то існує таке натуральне число $\delta \leqslant \varphi(m)$, при якому ця конгруенція справедлива. При цьому постають питання про числові значення показника δ для різних чисел a за модулем m. Чи може, зокрема, бути випадок, коли для всіх a, взаємно простих з m, показник $\delta < \varphi(m)$? Іншими словами, чи може бути випадок, коли за даним модулем m не існує первісних коренів? Чи для всіх m існують первісні корені? Скільки класів первісних коренів існує? Як їх знаходити?

Виявляється, що коли m — складене число, то первісні корені можуть і не існувати за цим модулем. Наприклад, якщо $m = 15$ (15 не є простим числом), то

$$\text{ПСЛ} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}.$$

Очевидно, тут $\varphi(m) = \varphi(15) = 8$, бо серед ПСЛ є 8 чисел, взаємно простих з m: 1, 2, 4, 7, 8, 11, 13, 14. Числа 3, 5, 6, 9, 10, 12 не взаємно прості з m і тому конгруенцію $a^{\delta} \equiv 1 \pmod m$ задовольняти не можуть. Легко перевірити, що

$1^1 \equiv 1 \pmod{15}$, тобто $\delta = 1$; $\quad 8^4 \equiv 1 \pmod{15}$, тобто $\delta = 4$;

$2^4 \equiv 1 \pmod{15}$, тобто $\delta = 4$; $\quad 11^2 \equiv 1 \pmod{15}$, тобто $\delta = 2$;

$4^2 \equiv 1 \pmod{15}$, тобто $\delta = 2$; $\quad 13^4 \equiv 1 \pmod{15}$, тобто $\delta = 4$;

$7^4 \equiv 1 \pmod{15}$, тобто $\delta = 4$; $\quad 14^2 \equiv 1 \pmod{15}$, тобто $\delta = 2$.

Таким чином, число $a = 1$ належить показнику $\delta = 1$ за модулем 15, числа $a = 4, 11, 14$ належать показнику $\delta = 2$, а числа $a = 2, 7, 8, 13$ — показнику $\delta = 4$ за модулем 15. Отже, всі показники $\delta < \varphi(m)$ і тому *не існує первісних коренів за модулем* 15.

Проте виявляється, що коли $m = p$ є простим числом, то первісні корені за цим модулем завжди існують. Доведемо це.

Теорема 3. *За будь-яким простим модулем* p *існує хоча б один первісний корінь.*

Доведення. Для $p = 2$ теорема правильна. Справді $\varphi(2) = 1$ і конгруенцію $a^{\varphi(2)} \equiv 1 \pmod 2$, тобто $a \equiv 1 \pmod 2$ задовольняє число $a = 1$; отже, число $a = 1$, а отже, клас $K_1^{(2)}$ є класом первісних коренів за модулем 2.

Нехай тепер p — непарне просте число $\varphi(p) = p - 1$. ЗСЛ за модулем p є множиною таких чисел:

$$1, 2, 3, \ldots, p - 1.$$

Для кожного з них за теоремою Ейлера справедлива конгруенція

$$a^{\varphi(p)} \equiv 1 \pmod p,$$

або

$$a^{p-1} \equiv 1 \pmod p. \quad (9)$$

Покажемо, що серед чисел ЗСЛ існує таке число a, що конгруенція

$$a^{\delta} \equiv 1 \pmod p \quad (10)$$

не може справджуватися при $\delta < p - 1$, тобто що $\delta = p - 1$ — це *найменший натуральний показник* степеня, при якому справедлива конгруенція (10). Саме ця умова і визначатиме те, що число a є первісним коренем за модулем p.

Нехай парне число $p - 1$ має такий канонічний розклад:

$$p - 1 = q_1^{k_1} \cdot q_2^{k_2} \cdot \ldots \cdot q_s^{k_s}. \quad (11)$$

Розглянемо конгруенцію

$$x^{\frac{p-1}{q_i}} \equiv 1 \pmod p, \quad (12)$$

де $i = 1, 2, \ldots, s$. Вона має, очевидно, не більше ніж $\frac{p-1}{q_i}$ розв'язків, які слід шукати серед чисел множини

$$\text{ЗСЛ} = \left\{1, 2, 3, \ldots, \frac{p-1}{q_i}, \ldots, p-1\right\}.$$

Оскільки $\frac{p-1}{q_i} < p-1$, то серед чисел ЗСЛ знайдеться хоча б одне число a_i, яке *не є розв'язком* конгруенції (12), тобто

$$a_i^{\frac{p-1}{q_i}} \not\equiv 1 \pmod p. \tag{13}$$

Побудуємо тепер число

$$b_i = a_i^{\frac{p-1}{q_i^{k_i}}} \tag{14}$$

і покажемо, що воно належить показнику $q_i^{k_i}$. Справді, перш за все

$$b_i^{q_i^{k_i}} = a_i^{p-1} \equiv 1 \pmod p.$$

Якщо позначити через δ показник, до якого належить число b_i, то $q_i^{k_i} \vdots \delta$. Оскільки q_i — просте число, то, очевидно, $\delta = q_i^{l_i}$, де $l_i \leqslant k_i$. Легко впевнитися, що $l_i = k_i$. Справді, якщо $l_i < k_i$, то конгруенція $b_i^{\delta} \equiv 1 \pmod p$ матиме вигляд

$$b_i^{\delta} = b_i^{q_i^{l_i}} = \left(a_i^{\frac{p-1}{q_i^{k_i}}}\right)^{q_i^{l_i}} = a_i^{\frac{p-1}{q_i^{k_i-l_i}}} \equiv 1 \pmod p,$$

і, підносячи обидві частини цієї конгруенції до степеня $q_i^{k_i-l_i-1}$, дістанемо

$$a_i^{\frac{p-1}{q_i}} \equiv 1 \pmod p,$$

що суперечить (13). Отже, випадок $l_i < k_i$ неможливий і $l_i = k_i$, а $\delta = q_i^{k_i}$:

$$P_p(b_i) = q_i^{k_i}. \tag{15}$$

Побудуємо тепер число $a = b_1 \cdot b_2 \cdot \ldots \cdot b_s$, де всі b_i визначаються рівностями (14). Оскільки за (15)

$$P_p(b_1) = q_1^{k_1},\ P_p(b_2) = q_2^{k_2},\ \ldots,\ P_p(b_s) = q_s^{k_s},$$

і всі числа $q_1^{k_1}, q_2^{k_2}, \ldots, q_s^{k_s}$ попарно взаємно прості, то на основі теореми 2 (п. 19.3) та рівностей (15) і (11) дістанемо

$$P_p(a) = P_p(b_1 \cdot b_2 \cdot \ldots \cdot b_s) = P_p(b_1) \cdot P_p(b_2) \cdot \ldots \cdot P_p(b_s) =$$

$$= q_1^{k_1} \cdot q_2^{k_2} \cdot \ldots \cdot q_s^{k_s} = p-1.$$

Таким чином, число a належить показнику $p-1$ за модулем p і тому є первісним коренем за модулем p. Теорему доведено.

19.5. Число класів первісних коренів.

Теорема 4. *Якщо існує хоч одне число, яке належить до показника δ за модулем p, то всього класів таких чисел є $\varphi(\delta)$.*

Доведення. Нехай $P_p(a) = \delta$. Тим самим передбачається, що $(a, p) = 1$ і справедливою є конгруенція

$$a^{\delta} \equiv 1 \pmod p.$$

Числа

$$1 = a^0,\ a,\ a^2,\ \ldots,\ a^{\delta-1} \tag{16}$$

є представниками δ класів, неконгруентних між собою чисел за модулем p. Всі вони, очевидно, є розв'язками конгруенції

$$x^{\delta} \equiv 1 \pmod p. \tag{17}$$

Оскільки остання не може мати більш як δ розв'язків, то множина (16) є множиною всіх розв'язків конгруенції (17).

Встановимо тепер кількість розв'язків з множини (16), які належать показнику δ. Припустимо, що деяке число a^k з множини (16) належить показнику β. Тоді β буде найменшим числом, при якому є справедливою конгруенція

$$(a^k)^{\beta} \equiv 1 \pmod p. \tag{18}$$

Іншими словами,

$$a^{k\beta} \equiv 1 \pmod p.$$

Оскільки $P_p(a) = \delta$, то $k\beta \vdots \delta$. При цьому можуть бути два випадки:

а) Нехай $(k, \delta) = 1$, тоді $\beta \vdots \delta$. З другого боку, поряд з конгруенцією (18) справедливою є конгруенція

$$(a^k)^{\delta} \equiv 1 \pmod p.$$

Враховуючи, що β — *найменше* натуральне число, при якому справедлива конгруенція (18), дістаємо, що $\delta \vdots \beta$. Оскільки числа δ і β діляться одне на одне, то $\beta = \delta$, тобто a^k належить до показника δ.

б) Нехай $(k, \delta) = d > 1$. Покажемо тепер, що в цьому випадку число a^k не належить до показника δ. Якщо $k = k_1 d$, $\delta = \delta_1 d$, то a^k належить до показника $\delta_1 < \delta$, бо

$$(a^k)^{\delta_1} = (a^{k_1 d})^{\frac{\delta}{d}} = a^{k_1 \delta} \equiv 1 \pmod p.$$

Тим самим *ми встановили, що числа множини* (16), *які належать до показника δ, мають вигляд a^k, де $(k, \delta) = 1$ і $k \leqslant \delta - 1$*. Але таких k, як відомо, існує $\varphi(\delta)$. Тим самим існує $\varphi(\delta)$ класів чисел за модулем p, які належать до показника δ. Теорему доведено.

Наслідок 1. *Первісних коренів за модулем p існує $\varphi(p-1)$.* За теоремою 3 первісні корені існують для будь-якого простого модуля p, тобто існує таке хоч одне натуральне число a, яке належить показнику $\delta = \varphi(p) = p-1$.

За теоремою 4, всіх класів таких чисел, тобто класів первісних коренів, є $\varphi(\delta) = \varphi(p-1)$.

Наслідок 2. *Якщо a — первісний корінь за модулем p, то інші первісні корені слід шукати серед степенів $a^2, a^3, \ldots, a^{p-1}$ — вони мають вигляд a^k, де $(k, p-1) = 1$ і $k \leqslant p-1$.* Це твердження безпосередньо випливає з самого ходу доведення теореми 4, де воно сформульоване для будь-якого показника δ.

Зручного способу для знаходження первісних коренів практично не існує. Їх знаходять за допомогою звичайних проб. Щоб дещо полегшити процес обчислень, можна використати таку теорему.

Теорема 5. *Якщо $p-1 = q_1^{k_1} \cdot q_2^{k_2} \cdot \ldots \cdot q_s^{k_s}$ — канонічний розклад числа $p-1$, то для того щоб число a було первісним коренем за модулем p, достатньо, щоб виконувались умови:*

$$a^{\frac{p-1}{q_i}} \not\equiv 1 \pmod{p}$$

для всіх $i = 1, 2, \ldots, s$.

Доведення. Якщо d — якийсь дільник числа $p-1$ і $d < p-1$, то на число d ділиться хоч одне з чисел

$$\frac{p-1}{q_1}, \frac{p-1}{q_2}, \ldots, \frac{p-1}{q_r}, \ldots, \frac{p-1}{q_s}. \tag{19}$$

Нехай, наприклад, число $\frac{p-1}{q_r}$ ділиться на d, тобто

$$\frac{p-1}{q_r} = dl.$$

Легко зрозуміти, що число a не може належати показнику d, бо тоді з конгруенції

$$a^d \equiv 1 \pmod{p}$$

відразу випливало б

$$(a^d)^l \equiv a^{\frac{p-1}{q_r}} \equiv 1 \pmod{p}, \text{ що}$$

за умовою теореми неможливо. Отже, a не може належати жодному числу d — дільнику числа $p-1$. Таким чином, a — первісний корінь за модулем p.

Приклад. Знайти первісні корені за модулем $p = 13$. За доведенням первісних коренів є $\varphi(p-1) = \varphi(12) = 4$. Шукати їх слід серед чисел ЗСЛ за модулем 13:

$$\text{ЗСЛ} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

Обчислення показують, що

$2^1 \equiv 2 \pmod{13}$,	$2^2 \equiv 4 \pmod{13}$,	$2^3 \equiv 8 \pmod{13}$,	$2^4 \equiv 3 \pmod{13}$,
$2^5 \equiv 6 \pmod{13}$	$2^6 \equiv -1 \pmod{13}$,	$2^7 \equiv -2 \pmod{13}$,	$2^8 \equiv -4 \pmod{13}$,
$2^9 \equiv -8 \pmod{13}$,	$2^{10} \equiv -3 \pmod{13}$,	$2^{11} \equiv -6 \pmod{13}$,	$2^{12} \equiv 1 \pmod{13}$.

Отже, 2 належить показнику $12 = \varphi(13)$, тобто 2 є первісним коренем за модулем 13. Проте для встановлення цього факту краще скористатися попередньою теоремою 5. Число $p-1 = 12$ в канонічному розкладі має вигляд $p-1 = 2^2 \cdot 3$. Побудуємо

числа виду (19):

$$\frac{p-1}{2} = \frac{12}{2} = 6, \quad \frac{p-1}{3} = \frac{12}{3} = 4.$$

Досліджуючи конгруентність степенів 2^6 і 2^4 за модулем 13, бачимо, що

$$2^6 = 64 \equiv -1 \pmod{13}, \quad 2^4 = 16 \equiv 3 \pmod{13}.$$

Отже, в обох випадках

$$2^6 \not\equiv 1 \pmod{13} \text{ і } 2^4 \not\equiv 1 \pmod{13},$$

і тому за теоремою 5 число 2 є первісним коренем за модулем 13.

За наслідком 2 з теореми 4, інші первісні корені є числами виду 2^k, де $(k, p-1) = (k, 12) = 1$ і $0 < k \leqslant 12$. Ці умови задовольняють такі значення: $k = 5, 7, 11$. Отже, іншими первісними коренями за модулем 13 є числа $2^5, 2^7, 2^{11}$. Враховуючи, що

$$2^5 \equiv 6 \pmod{13}, \quad 2^7 \equiv 11 \pmod{13}, \quad 2^{11} \equiv 7 \pmod{13},$$

встановлюємо, що первісними коренями за модулем 13 є числа 2, 6, 7, 11.

19.6. Індекси за простим модулем. ЗСЛ за модулем p можна подати сукупністю чисел — найменших невід'ємних лишків

$$1, 2, 3, 4, \ldots, p-1. \tag{20}$$

Разом з тим, ЗСЛ може бути поданою й інакше — за допомогою степенів якогось первісного кореня за модулем p. Якщо q є первісний корінь за модулем p (нагадуємо, що p і q взаємно прості), то степені

$$q, q^2, q^3, \ldots, q^{p-1} \tag{21}$$

є також сукупністю $p-1$ чисел, неконгруентних між собою за модулем p (властивість 1, п. 19.2). Тому ці числа є також представниками різних класів лишків за модулем p і утворюють ЗСЛ за модулем p. Кожне число сукупності (20) конгруентне деякому числу сукупності (21). Звідси кожний клас лишків ЗСЛ за модулем p можна подати якимсь числом виду q^γ. Тим самим кожному класу лишків $K_a^{(p)}$ ЗСЛ можна поставити у відповідність показник степеня γ числа q^γ, який і називається індексом класу $K_a^{(p)}$ при основі q.

Означення. *Індексом числа a за модулем p* (або класу $K_a^{(p)}$) *при основі q називається таке ціле невід'ємне число γ, що*

$$q^\gamma \equiv a \pmod{p}.$$

Позначають індекс $\gamma = \operatorname{ind}_q a$ за модулем p.

Зауважимо, що для класу $K_0^{(p)}$ чисел, кратних модулю p, поняття індексу не вводять, бо при умові $(p, q) = 1$ конгруенція виду $kp \equiv q^\gamma \pmod{p}$ неможлива.

Приклад. Нехай $p = 7$. Можна встановити, що первісними коренями за модулем 7 є числа 3 і 5. Візьмемо за q менший первісний корінь 3. Очевидно, $3^1 \equiv 3 \pmod{7}$, $3^2 \equiv 2 \pmod{7}$, $3^3 \equiv 6 \pmod{7}$, $3^4 \equiv 4 \pmod{7}$, $3^5 \equiv 5 \pmod{7}$, $3^6 \equiv 1 \pmod{7}$. У правих частинах конгруенцій стоять числа ЗСЛ виду (20). Перепишемо ці конгруенції інакше: $1 \equiv 3^6 \pmod{7}$, $2 \equiv 3^2 \pmod{7}$, $3 \equiv 3^1 \pmod{7}$, $4 \equiv 3^4 \pmod{7}$, $5 \equiv 3^5 \pmod{7}$, $6 \equiv 3^3 \pmod{7}$. Кожному з чисел 1, 2, 3, 4, 5, 6 або, що те саме, кожному з класів лишків за модулем 7

$$K_1^{(7)}, K_2^{(7)}, K_3^{(7)}, K_4^{(7)}, K_5^{(7)}, K_6^{(7)}$$

поставимо у відповідність індекс γ — показник степеня первісного кореня 3. Цю відповідність можна подати у вигляді такої таблиці:

Число	1	2	3	4	5	6
Індекс	6	2	1	4	5	3

(22)

Таким чином, поняття індексів у теорії конгруенцій аналогічне поняттю логарифмів чисел. Ми далі побачимо, що й роль їх аналогічна: операції з числами в конгруенціях можна замінити певними операціями над індексами. На практиці користуються таблицями індексів, аналогічними побудованій. Для зручності користування будують і другу таблицю, де упорядкована в порядку зростання множина індексів:

Індекс	1	2	3	4	5	6
Число	3	2	6	4	5	1

(22′)

Основні властивості індексів.

1. *Числа $\gamma' \geqslant 0$ є індексами числа a за модулем p при основі q тоді і тільки тоді, коли*

$$\gamma' \equiv \gamma \pmod{p-1}, \text{ де } \gamma = \operatorname{ind}_q a \text{ за модулем } p.$$

Доведення. *Необхідність.*

Нехай $\gamma = \operatorname{ind}_q a$ і $\gamma' = \operatorname{ind}_q a$ за модулем p. Тоді

$$q^{\gamma} \equiv a \pmod{p}, \quad q^{\gamma'} \equiv a \pmod{p},$$

звідки

$$q^{\gamma} \equiv q^{\gamma'} \pmod{p}. \tag{23}$$

За властивістю 2 п. 19.2, де $m = p$, а δ — показник, до якого належить первісний корінь q, тобто $\delta = p - 1$, з конгруенції (23) маємо

$$\gamma' \equiv \gamma \pmod{p-1}. \tag{24}$$

Достатність. Якщо виконується конгруенція (24), то за властивістю 2 п. 19.2 дістанемо конгруенцію (23), де $\gamma = \operatorname{ind}_q a$ за модулем p, тобто

$$q^{\gamma} \equiv a \pmod{p}.$$

Але тоді й γ' є також індексом числа a за модулем p при основі q, тобто

$$q^{\gamma'} \equiv a \pmod{p}.$$

2. *Для того щоб*

$$a \equiv b \pmod{p},$$

необхідно і досить, щоб виконувалася конгруенція

$$\operatorname{ind}_q a \equiv \operatorname{ind}_q b \pmod{p-1}.$$

Доведення. *Необхідність.* Нехай

$$a \equiv b \pmod{p}, \quad \gamma = \operatorname{ind}_q a, \quad \gamma' = \operatorname{ind}_q b.$$

Тоді

$$a \equiv q^{\gamma} \pmod{p}, \quad b \equiv q^{\gamma'} \pmod{p},$$

і тому

$$q^{\gamma} \equiv q^{\gamma'} \pmod{p}.$$

З останньої конгруенції випливає

$$\gamma \equiv \gamma' \pmod{p-1}.$$

Достатність. Якщо $\gamma \equiv \gamma' \pmod{p-1}$ і $\gamma = \operatorname{ind}_q a$, $\gamma' = \operatorname{ind}_q b$, то $q^{\gamma} \equiv q^{\gamma'} \pmod{p}$, $a \equiv q^{\gamma} \pmod{p}$, $b \equiv q^{\gamma'} \pmod{p}$ і тому $a \equiv b \pmod{p}$, що й треба було довести.

Доведеними властивостями можна встановити, що кожному класу лишків $K_a^{(p)}$, де a входить до ЗСЛ за модулем p, взаємно однозначно відповідає клас лишків $K_{\gamma}^{(p-1)}$, де ν входить до ПСЛ за модулем $p-1$.

У зв'язку з цим повніше таблицю (22) можна записати так:

Класи за модулем 7	$K_1^{(7)}$	$K_2^{(7)}$	$K_3^{(7)}$	$K_4^{(7)}$	$K_5^{(7)}$	$K_6^{(7)}$
Класи індексів (класи за модулем 6)	$K_6^{(6)}$	$K_2^{(6)}$	$K_1^{(6)}$	$K_4^{(6)}$	$K_5^{(6)}$	$K_3^{(6)}$

(25)

3. $\operatorname{ind}_q 1 \equiv 0 \pmod{p-1}$.

4. $\operatorname{ind}_q q \equiv 1 \pmod{p-1}$.

Обидві властивості очевидні, бо

$$1 \equiv q^0 \pmod{p-1}, \quad q \equiv q^1 \pmod{p-1}.$$

5. *Індекс добутку чисел за модулем p при основі q конгруентний за модулем $p-1$ сумі індексів окремих множників при основі q, тобто*

$$\operatorname{ind}_q (a_1 a_2 \dots a_n) \equiv \operatorname{ind}_q a_1 + \operatorname{ind}_q a_2 + \dots + \operatorname{ind}_q a_n \pmod{p-1}.$$

Доведення. Якщо

$$\operatorname{ind}_q a_1 = \gamma_1, \quad \operatorname{ind}_q a_2 \equiv \gamma_2, \dots, \operatorname{ind}_q a_n = \gamma_n,$$

то

$$a_1 \equiv q^{\gamma_1} \pmod{p}, \quad a_2 \equiv q^{\gamma_2} \pmod{p}, \dots, \quad a_n \equiv q^{\gamma_n} \pmod{p}.$$

Перемножаючи почленно ці конгруенції за властивістю 6 п. 15.1, далі дістаємо

$$a_1 a_2 \dots a_n \equiv q^{\gamma_1 + \gamma_2 + \dots + \gamma_n} \pmod{p}. \tag{26}$$

Якщо позначити тепер $\gamma = \operatorname{ind}_q (a_1 a_2 \dots a_n)$, тобто

$$a_1 a_2 \dots a_n \equiv q^{\gamma} \pmod{p}, \tag{27}$$

то з конгруенцій (26) і (27) за властивістю 2 п. 19.2 маємо

$$\gamma \equiv \gamma_1 + \gamma_2 + \dots + \gamma_n \pmod{p-1},$$

тобто

$$\operatorname{ind}_q (a_1 a_2 \dots a_n) \equiv \operatorname{ind}_q a_1 + \operatorname{ind}_q a_2 + \dots + \operatorname{ind}_q a_n \pmod{p-1},$$

що й треба було довести.

6. $\operatorname{ind}_g a^n \equiv n \operatorname{ind}_g a \pmod{p-1}$.

7. Якщо $a : b$, то $\operatorname{ind}_g \frac{a}{b} \equiv \operatorname{ind}_g a - \operatorname{ind}_g b \pmod{p-1}$.

Останні властивості є наслідком властивості 5.

Зауважимо, що перехід від конгруенції між числами до конгруенції їх індексів називається *індексацією*, а зворотний перехід — *потенціюванням*.

19.7. Розв'язування двочленних конгруенцій n-го степеня за допомогою індексів. З двочленними конгруенціями n-го степеня за простим модулем ми вже зустрічалися в п. 18.3. В загальному вигляді двочленні конгруенції можна записати так:

$$ax^n \equiv b \pmod{p}, \tag{28}$$

де $(a, p) = 1$ і n — натуральне число.

Якщо провести індексацію цієї конгруенції при однаковій основі, то дістанемо конгруенцію $\operatorname{ind}(ax^n) \equiv \operatorname{ind} b \pmod{p-1}$, або, що те ж саме,

$$\operatorname{ind} a + n \operatorname{ind} x \equiv \operatorname{ind} b \pmod{p-1}. \tag{29}$$

Конгруенції (28) і (29) еквівалентні. Якщо позначити

$$\operatorname{ind} b - \operatorname{ind} a = c, \quad \operatorname{ind} x = y,$$

то конгруенція (29) має вигляд

$$ny \equiv c \pmod{p-1}. \tag{30}$$

Тим самим від конгруенції n-го степеня (28) за допомогою індексації ми прийшли до конгруенції (29) першого степеня. Розв'язавши її і взявши величину $y = \operatorname{ind} x$, знайдемо x за відповідними таблицями.

Приклад. Розв'язати конгруенцію $3x^3 \equiv 4 \pmod 7$.
Проводячи індексацію при деякій основі q (як правило, це найменший первісний корінь за модулем p), дістанемо

$$\operatorname{ind} 3 + 3 \operatorname{ind} x \equiv \operatorname{ind} 4 \pmod 6.$$

За таблицею (22) маємо $\operatorname{ind} 3 = 1$, $\operatorname{ind} 4 = 4$ і тому маємо

$$1 + 3 \operatorname{ind} x \equiv 4 \pmod 6, \quad 3 \operatorname{ind} x \equiv 3 \pmod 6.$$

Розв'язками цієї лінійної конгруенції є числа $\operatorname{ind} x = 1, 3, 5$ з ПСЛ за модулем 6. З таблиці (22′) дістаємо відповідні три значення невідомого x:

$$x \equiv 3,\ 6,\ 5 \pmod 7.$$

Отже, задана конгруенція має три розв'язки.
За допомогою індексів дуже легко знайти показники за модулем. Справді, нехай треба знайти P_7 (6). Маємо конгруенцію

$$6^x \equiv 1 \pmod 7.$$

Виконавши індексацію, дістанемо конгруенцію 1-го степеня

$$x \operatorname{ind} 6 \equiv \operatorname{ind} 1 \pmod 6.$$

За таблицею (22) далі маємо $x \cdot 3 \equiv 0 \pmod 6$, або, скоротивши на 3 обидві частини і модуль конгруенції, маємо $x \equiv 0 \pmod 2$. Вибираючи найменше натуральне число з класу $K_0^{(2)}$, знаходимо, що $x = 2$, тобто $P_7(6) = 2$.

20.1. Застосування конгруенцій до встановлення ознак подільності. Як відомо, в кільці $\mathbf{Z}$ цілих чисел визначені операції додавання, віднімання і множення, а дія ділення не завжди можлива. Тому виникає потреба визначити, при яких умовах цілі числа діляться одно на одне.

Подільність чисел — це певне відношення між числами, яке в $\mathbf{Z}_+$ має такі властивості: *рефлексивність* $(a : a)$, *транзитивність* $([a : b \wedge b : c] \Rightarrow [a : c])$ і *антисиметричність* $([a : b \wedge b : a \Rightarrow a = b])$. Будь-яке відношення, яке має властивості рефлексивності, транзитивності і антисиметричності, називається *відношенням не строгого порядку*. Отже, подільність чисел в $\mathbf{Z}_+$ є відношенням не строгого порядку. Аналогічним відношенням частинної упорядкованості є, наприклад, відношення «$\geqslant$» в кільці $\mathbf{Z}$. Воно рефлексивне $(a \geqslant a)$, транзитивне $[a \geqslant b \wedge b \geqslant c] \Rightarrow [a \geqslant b]$, антисиметричне (якщо $a \geqslant b$ і $b \geqslant a$, то $b \geqslant a$).

Між відношеннями подільності і $\geqslant$ в кільці $\mathbf{Z}$ можна встановити і таку аналогію. Відношення $a : b$ означає, що існує таке число c, при якому виконується рівність $a = bc$. Відношення $a \geqslant b$, або $a - b \geqslant 0$, означає, що існує таке число $c \geqslant 0$, при якому $a = b + c$. Рівності $a = bc$ і $a = b + c$, як бачимо, аналогічні.

Факт подільності двох чисел можна, звичайно, встановити за допомогою алгоритму ділення чисел з остачею. Проте для великих чисел це завдання досить складне. Тому бажано знайти зручні ознаки, за якими можна було б судити про подільність чисел, не виконуючи самого ділення. В цілому суть ознак подільності зводиться до того, що розгляд подільності деякого натурального числа a на натуральне число m замінюється розглядом подільності на число m іншого, меншого за a натурального числа b, яке можна знайти за деяким правилом, що визначається числовою функцією $f(a)$, тобто $b = f(a)$. При цьому числа a і $b = f(a)$ є, як кажуть, рівноподільними на число m, тобто такі, які одночасно діляться або одночасно не діляться на число m. Часто вимагають, щоб вони були конгруентними за модулем m.

Одним із способів знаходження ознак подільності, основаних на конгруентності чисел, є так званий спосіб Паскаля [1]. Нехай деяке натуральне число a при основі числення g має вигляд

$$a = p_n g^n + p_{n-1} g^{n-1} + \cdots + p_1 g + p_0, \tag{1}$$

де коефіцієнти p_k є натуральні числа, які задовольняють нерівності $0 \leqslant p_k < g$. Позначимо через r_k остачу від ділення числа g^k на m, тобто $g^k \equiv r_k \pmod m$, і побудуємо число $b = f(a)$ за таким правилом:

$$b = f(a) = p_n r_n + p_{n-1} r_{n-1} + \cdots + p_1 r_1 + p_0. \tag{2}$$

На основі властивості 9 п. 15.1 $a \equiv b \pmod m$. Оскільки $b < a$, то дістаємо таку ознаку Паскаля подільності чисел:

[1] Блез Паскаль (1623—1662) — французький математик і фізик.

Якщо число $b = p_n r_n + p_{n-1} r_{n-1} + \dots + p_1 r_1 + p_0$ *ділиться на число* m, *то ділиться на нього і число* $a = p_n g^n + p_{n-1} g^{n-1} + \dots + p_1 g + p_0$.

Якщо ж b *на число* m *не ділиться, то не ділиться на* m *і число* a.

За допомогою цієї загальної ознаки можна встановити зручні конкретні ознаки подільності чисел, записаних у звичайній для нас десятковій системі числення. У цій системі $g = 10$ і число a має вигляд:

$$a = p_n \cdot 10^n + p_{n-1} \cdot 10^{n-1} + \dots + p_1 \cdot 10 + p_0. \qquad (3)$$

Коротко це можна записати так: $a = \overline{p_n p_{n-1} \dots p_1 p_0}$.

а) *Ознака подільності на* 2 *і на* 5.

Оскільки $10^k : 2$ і $10^k : 5$, то всі остачі r_k від ділення 10^k на числа 2 і 5 дорівнюють нулю. Тому за формулою (2) число $b = f(a) = p_0$. Отже, маємо таку ознаку:

Число a *ділиться на* 2 *і на* 5 *тоді і тільки тоді, коли на них ділиться цифра одиниць числа* a.

Приклад 1. Число $a = 8574 : 2$, бо $4 : 2$. Число 8127 не ділиться на 5, бо 7 не ділиться на 5.

б) *Ознака подільності на* 3 *і на* 9.

Оскільки всі остачі r_k від ділення 10^k на число 3 або 9 дорівнюють 1, то за (2)

$$b = f(a) = p_n + p_{n-1} + \dots + p_1 + p_0.$$

Отже, маємо таку ознаку:

Число a *ділиться на* 3 (або на 9) *тоді і тільки тоді, коли сума цифр, які його зображують, ділиться на* 3 (або відповідно на 9).

Приклад 2. Число $a = 5742 : 3$, бо сума цифр $b = 5 + 7 + 4 + 2 = 18 : 3$.

в) *Ознака подільності на* 11.

За модулем 11 маємо

$$10^{2k} \equiv 1 \pmod{11}, \quad 10^{2k-1} \equiv -1 \pmod{11}.$$

Тому $r_{2k} = 1$, $r_{2k-1} = -1$, і, отже, за рівністю (2)

$$b = p_0 - p_1 + p_2 - p_3 + p_4 - p_5 + \dots = (p_0 + p_2 + p_4 + \dots) - (p_1 + p_3 + p_5 + \dots).$$

Враховуючи, що цифри p_{2k} з парними індексами в числі a стоять на непарних місцях, можна сформулювати таку ознаку:

Число a *ділиться на* 11 *тоді і тільки тоді, коли різниця між сумою цифр, які стоять на непарних місцях, і сумою цифр, які стоять на парних місцях, ділиться на* 11.

Приклад 3. Число $a = 53746 : 11$, бо число

$$b = (6 + 7 + 5) - (4 + 3) = 18 - 7 = 11$$

ділиться на 11.

У системі числення з основою $g = 10^2$ можна знайти зручні ознаки подільності на числа 4, 25, 50. Число a в цій системі можна записати так:

$$a = c_s \cdot 100^s + c_{s-1} \cdot 100^{s-1} + \dots + c_1 \cdot 100 + c_0.$$

Порівнюючи це з (3), бачимо, що $c_0 = p_1 \cdot 10 + p_0 = \overline{p_1 p_0}$, тобто є двоцифровим числом, яке зображується двома останніми цифрами числа a в десятковій системі числення.

Враховуючи, що $a \equiv c_0 \pmod{100}$ і числа 100^k діляться на числа 4, 25, 50, дістаємо такі ознаки подільності:

Число $a = \overline{p_n p_{n-1} \dots p_1 p_0}$ *ділиться на* 4 (або на 25 чи 50), *якщо на* 4 (або відповідно на 25 чи 50) *ділиться двоцифрове число* $c_0 = \overline{p_1 p_0}$, *утворене двома останніми цифрами числа* a, *записаного в десятковій системі числення.*

Ознаки подільності є цінними, якщо вони прості, зручні для користування. Проте більшість ознак, які можна вивести з ознаки Паскаля, є складними. Існує ряд зручних ознак подільності, які не випливають з загальної ознаки Паскаля, а знайдені іншими способами. Наприклад, одну з ознак подільності на 7 можна сформулювати так:

Число $a = 10a_1 + a_0$ *ділиться на* 7 *тоді і тільки тоді, коли ділиться на* 7 *число* $b = a_1 - 2a_0$.

Зазначимо, що на відміну від усіх попередніх ознак числа a і b тут рівноподільні на 7, а не конгруентні між собою за модулем $m = 7$.

Приклади. 4. $a = 285 = 28 \cdot 10 + 5$, $b = 28 - 2 \cdot 5 = 18$.

Оскільки b не ділиться на 7, то не ділиться на 7 і число 285. Зазначимо, що при діленні на 7 числа 285 дістаємо остачу 5, а при діленні на 7 числа 18 остача дорівнює 4 і тому $285 \not\equiv 18 \pmod 7$.

5. Встановити, чи ділиться на 7 число $a = 63\,364$.

Приклад можна розв'язати так. Перш за все $a = 63\,000 + 364$ і тому $63\,364 \equiv 364 \pmod 7$. А $364 = 36 \cdot 10 + 4$. Тому $b = 36 - 2 \cdot 4 = 28$. Оскільки $28 : 7$, то й число $63\,364 : 7$.

20.2. Перетворення звичайного дробу в систематичний і визначення довжини періоду систематичного дробу. Розглянемо питання про перетворення звичайного дробу в десятковий. Як відомо з арифметики звичайні дроби перетворюються або в скінченні, або в нескінченні періодичні десяткові дроби. При цьому звичайний дріб $\frac{a}{b}$ перетворюється в скінченний десятковий дріб тоді і тільки тоді, коли канонічний розклад знаменника має вигляд $2^\alpha 5^\beta$, тобто не містить ніяких простих множників, крім 2 і 5. Для спрощення вважатимемо $\frac{a}{b}$ нескоротним правильним дробом. (Якщо він неправильний, то можна спочатку виділити цілу частину). Звичайні нескоротні правильні дроби виду $\frac{a}{2^\alpha \cdot 5^\beta}$ перетворюються в скінченні десяткові дроби з числом десяткових знаків, яке дорівнює найбільшому з чисел α або β. Справді, якщо $\alpha = \beta$, то $\frac{a}{2^\alpha \cdot 5^\beta} = \frac{a}{2^\alpha \cdot 5^\alpha} = \frac{a}{10^\alpha}$ — скінченний десятковий дріб. Якщо $\alpha <$

$<\beta$, то $\frac{a}{2^\alpha \cdot 5^\beta} = \frac{a \cdot 2^{\beta-\alpha}}{2^\beta \cdot 5^\beta} = \frac{a \cdot 2^{\beta-\alpha}}{10^\beta}$ — скінченний десятковий дріб.

Якщо $\alpha > \beta$, то $\frac{a}{2^\alpha \cdot 5^\beta} = \frac{a \cdot 5^{\alpha-\beta}}{2^\alpha \cdot 5^\alpha} = \frac{a \cdot 5^{\alpha-\beta}}{10^\alpha}$ — скінченний десятковий дріб.

Легко зрозуміти, що *нескоротний дріб виду* $\frac{a}{c \cdot 2^\alpha \cdot 5^\beta}$, *де* c *відмінне від* 2 *і* 5, *в скінченний десятковий дріб не перетворюється.*

Справді, припускаючи супротивне, маємо

$$\frac{a}{c \cdot 2^\alpha \cdot 5^\beta} = \frac{d}{10^m},$$

звідки $a \cdot 10^m = d \cdot c \cdot 2^\alpha 5^\beta$, де c — дільник числа $a \cdot 10^m$, що неможливо, бо c відмінне від 2 і 5 за умовою і $(a, c) = 1$. Ця суперечність доводить справедливість твердження.

Теорема 1. *Якщо канонічний розклад знаменника* b *нескоротного дробу* $\frac{a}{b}$ *не містить у собі множників* 2 *і* 5, *то цей дріб перетворюється у чистий періодичний десятковий дріб; при цьому число цифр у періоді дорівнює показнику* δ, *до якого належить число* 10, *за модулем* b.

Доведення. Для спрощення дріб $\frac{a}{b}$ вважатимемо правильним. Процес ділення числа a на число b при умові $a < b$ можна схематично зобразити так:

$$\begin{array}{l|l} a \cdot 10 & b \\ \hline b\ \ q_1 & 0,\ q_1 q_2 \ldots q_m q_{m+1} \ldots \\ \hline r_1 \cdot 10 & \\ \hline b\ \ q_2 & \\ \hline r_2 & \\ \hline \ldots\ldots & \\ \hline r_{m-1} \cdot 10 & \\ \hline b\ \ \ q_m & \\ \hline r_m \cdot 10 & \\ \hline b\ \ q_{m+1} & \\ \hline r_{m+1} & \\ \hline \ldots\ldots & \end{array} \qquad (4)$$

Цю схему в свою чергу можна подати у вигляді системи рівностей:

$$\begin{aligned} 10a &= bq_1 + r_1, \\ 10r_1 &= bq_2 + r_2, \\ &\ldots\ldots\ldots \\ 10r_{m-1} &= bq_m + r_m, \\ 10r_m &= bq_{m+1} + r_{m+1}, \\ &\ldots\ldots\ldots\ldots \end{aligned} \qquad (5)$$

де $r_1, r_2, \ldots, r_m, r_{m+1}, \ldots$ — остачі, а $q_1, q_2, \ldots, q_m, q_{m+1}, \ldots$ — частки проміжних обчислень. Будь-яка остача r_i, очевидно, задовольняє нерівність

$$0 < r_i < b, \qquad (6)$$

а будь-які числа q_i задовольняють нерівність $0 \leqslant q_i < 10$, тобто є цифрами, з яких складається частка $0, q_1 q_2 \ldots q_m q_{m+1} \ldots$ в схемі (4).

Проаналізуємо властивості чисел r_i і q_i докладніше. Насамперед нагадаємо, що дріб $\frac{a}{b}$ є нескоротним і правильним. Це означає, що $(a, b) = 1$ і $a < b$. Таким чином, число a є один з найменших додатних лишків ЗСЛ за модулем b.

Покажемо тепер, що й усі r_i також є найменшими додатними лишками ЗСЛ за модулем b. Справді

$$[(a, b) = 1 \wedge (10, b) = 1)] \Rightarrow [(10a, b) = 1]. \qquad (7)$$

Оскільки числа $10a$ і b взаємно прості, то з першої рівності (5) випливає, що $(r_1, b) = 1$. Справді з умови $(r_1, b) = d > 1$ випливало б, що вся права частина, а отже, і ліва частина ділилась би на d. Тому числа $10a$ і b не були б взаємно простими, що суперечить (7). З умов $0 < r_1 < b$ і $(r_1, b) = 1$ випливає, що остача r_1 є одним з найменших додатних лишків ЗСЛ за модулем b. Аналогічно можна показати, що й числа $r_2, r_3, \ldots, r_m, r_{m+1}, \ldots$ є найменшими додатними лишками ЗСЛ за модулем b. Але ЗСЛ за модулем b може мати не більше $\varphi(b)$ найменших додатних лишків. Тому в системі рівностей (5) настане момент, коли одна з остач дорівнюватиме a. Нехай $r_m = a$. Тоді $m + 1$ рівність (5) збіжиться з першою рівністю цієї системи. І тому $q_1 = q_{m+1}$, $r_1 = r_{m+1}$. Далі, $m + 2$-а рівність збіжиться з другою рівністю (5) і тому $q_2 = q_{m+2}$, $r_2 = r_{m+2}$. Таким чином, остачі r_i і частки q_i проміжних обчислень повторюватимуться. Тим самим частка в схемі (4) буде чистим періодичним десятковим дробом виду

$$0, (q_1 q_2 \ldots q_m). \qquad (8)$$

Для доведення теореми залишається показати, що перше повторення настане після δ кроків проміжних обчислень, де δ — показник, до якого належить 10 за модулем b. Справді, якщо δ — найменший показник, при якому здійснюється конгруенція

$$10^\delta \equiv 1 \pmod b,$$

то при $r < b$ і $(r, b) = 1$ рівносильною їй є і конгруенція

$$10^\delta \cdot r \equiv r \pmod b.$$

Остання конгруенція якраз і показує, що, приписавши до r δ нулів, що відповідає визначенню δ послідовних цифр частки, дістанемо при діленні $10^\delta \cdot r$ на b остачу r. При діленні a на b при $(a, b) = 1$ і $a < b$ аналогічно дістанемо через δ ділень остачу, яка дорівнює числу a. Отже, частка (8) має вигляд $0, (q_1 q_2 \ldots q_\delta)$, що й треба було довести.

Зауваження. З конгруенції $10^\delta \equiv 1 \pmod b$ випливає, що $10^\delta - 1 \equiv 0 \pmod b$, або $\underbrace{999\ldots9}_{\delta \text{ цифр}} \equiv 0 \pmod b$.

Іншими словами, число 999...9, що складається з δ дев'яток,— найменше з можливих чисел такої структури, яке ділиться на b. Це дає можливість досить легко знаходити число δ. Для цього треба послідовно ділити на b числа 9, 99, 999, 9999, ... і т. д., аж поки таке ділення не відбудеться. Кількість дев'яток у такому числі і дорівнює числу δ.

Теорема 2. *Якщо канонічний розклад знаменника b нескоротного дробу $\frac{a}{b}$ має вигляд $b = 2^\alpha \cdot 5^\beta \cdot c$, де $(c, 10) = 1$, то цей дріб перетворюється у мішаний періодичний дріб; число цифр до періоду дорівнює γ, де γ — найбільше з чисел α і β; число цифр періоду дорівнює δ, де δ — показник, якому належить число 10 за модулем c.*

Доведення. Дріб

$$\frac{a}{b} = \frac{a}{2^\alpha \cdot 5^\beta \cdot c}$$

помножимо на 10^γ, де $\gamma = \max\{\alpha, \beta\}$. Матимемо

$$\frac{10^\gamma \cdot a}{b} = \frac{a \cdot 2^{\gamma-\alpha} \cdot 5^{\gamma-\beta}}{c} = \frac{a_1}{c}$$

і далі

$$[(a, c) = 1 \wedge (2, c) = 1 \wedge (5, c) = 1] \Rightarrow$$
$$\Rightarrow [(a \cdot 2^{\gamma-\alpha} \cdot 5^{\gamma-\beta}, c) = 1] \Rightarrow [(a_1, c) = 1].$$

За теоремою 1, дріб $\frac{a_1}{c}$ перетворюється в чистий періодичний дріб з числом цифр у періоді, яке дорівнює δ, де δ — показник, до якого належить 10 за модулем c. Щоб з нього дістати початковий дріб $\frac{a}{b}$, треба розділити його на 10^γ, або інакше, перенести кому в знайденому періодичному дробі на γ знаків ліворуч; у результаті дістанемо мішаний періодичний дріб з числом γ цифр до періоду. Теорему доведено.

Приклади. **1.** Знайти число цифр періоду десяткового періодичного дробу, в який перетворюється дріб $\frac{7}{39}$.

Ділимо на 39 послідовно числа 9, 99, 999, 9999, 99999. Нарешті з'ясовується, що тільки число 999999 націло ділиться на 39. Кількість дев'яток у цьому числі визначає довжину періоду: $\delta = 6$.

2. Знайти число цифр, яке міститься до періоду, і довжину періоду періодичного дробу, в який перетворюється дріб $\frac{13}{550}$.

Знаменник цього дробу в канонічному розкладі має вигляд $550 = 2 \cdot 5^2 \cdot 11$. Тому $\gamma = \max(\alpha, \beta) = 2$ є найбільший з показників степеня цифр 2 і 5. Це означає, що періодичний десятковий дріб має дві цифри до періоду. Найменше з чисел, складених з дев'яток, яке ділиться на 11, є число 99. Воно складається з двох дев'яток. Це означає, що довжина δ періоду періодичного дробу дорівнює 2. І справді, як неважко перевірити, розглядуваний дріб перетворюється в такий періодичний дріб:

$$\frac{13}{550} = 0{,}02\,(36).$$

МНОГОЧЛЕНИ ВІД ОДНІЄЇ ЗМІННОЇ

§ 21. КІЛЬЦЕ МНОГОЧЛЕНІВ НАД ОБЛАСТЮ ЦІЛІСНОСТІ

21.1. Попередні зауваження. Поняття многочлена [1], безперечно, не є новим для читача. З цим поняттям зустрічалися як у середній школі, так і при вивченні математичного аналізу та алгебри на І курсі.

З курсу математичного аналізу відомо, що многочленом від однієї змінної є ціла раціональна функція виду

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \qquad (1)$$

задана на всій дійсній осі, де коефіцієнти a_n, a_{n-1}, ..., a_1, a_0 — довільні задані дійсні числа. В аналізі вивчались деякі властивості многочленів від дійсної змінної, зокрема їх неперервність, вирази для похідних різного порядку тощо. Найпростіші типи многочленів (зокрема лінійну функцію $f(x) = ax + b$ та квадратний тричлен $f(x) = ax^2 + bx + c$) досить детально вивчають ще в середній школі.

З другого боку, в алгебрі многочлени зустрічались у зв'язку з розв'язуванням алгебраїчних рівнянь вищих степенів з одним невідомим, тобто рівнянь виду

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0, \qquad (2)$$

ліва частина яких є многочлен від однієї змінної. На відміну від аналізу, в алгебрі многочлени вважалися цілими раціональними функціями комплексної змінної, тобто виразами виду (1), в яких коефіцієнти a_n, a_{n-1}, ..., a_1, a_0 є комплексні числа, а змінна x може набувати довільних комплексних значень. Доцільність розгляду многочленів в елементарній алгебрі як функцій комплексної змінної була пов'язана з тим, що вже в найпростішому випадку квадратного рівняння розв'язування його приводить до комплексних коренів навіть тоді, коли коефіцієнти цього рівняння — дійсні числа. Тим більше це справедливо для рівнянь 3-го, 4-го і вищих степенів. Якщо ж розглядати многочлени з довільними комплексними коефіцієнтами, то всі їхні корені виявляються також комплексними (в окремому випадку — дійсними) числами, тобто розв'язування алгебраїчних рівнянь будь-якого степеня не виводить нас за межі поля **C** комплексних чисел (якщо коефіцієнти належали цьому полю).

Хоч трактування многочленів як функцій комплексної змінної достатнє для потреб дослідження і розв'язування алгебраїчних рівнянь з числовими коефіцієнтами, ми все ж таки покладемо в основу дальшого викладу ще більш загальний погляд на многочлен. Адже ми тепер знайомі не лише з числовими, а й абстрактними алгебраїчними структурами (кільцями, полями, лінійними просторами, алгеб-

[1] У математиці вживають також термін поліном, який походить від грецьких слів πολις (багато) і νομη (член).

рами та ін.), і природно ставити питання про існування, кількість та способи знаходження розв'язків рівнянь виду (2), в яких коефіцієнти та невідоме є елементи деякої абстрактної алгебраїчної структури, причому дії додавання і множення, за допомогою яких утворено многочлен, є операції цієї структури, а 0 — її нульовий елемент.

Звичайно, абстрактне означення многочленів повинно бути таким, щоб їх властивості узагальнювали важливі з теоретичного і практичного погляду властивості многочленів з числовими коефіцієнтами.

21.2. Означення многочлена. Вираз виду

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \qquad (3)$$

повністю визначається коефіцієнтами a_n, a_{n-1}, ..., a_1, a_0. Але ці коефіцієнти не можуть бути цілком довільними об'єктами. Якщо ми хочемо, щоб над многочленами можна було виконувати операції додавання та множення і, крім того, за тими самими правилами, що й в елементарній математиці, то потрібно подбати про те, щоб для коефіцієнтів многочленів мали сенс дії додавання та множення і щоб ці дії мали достатньо «добрі» властивості (асоціативність, дистрибутивність, комутативність).

Отже, природно вважати, що коефіцієнти розглядуваних многочленів повинні належати деякому комутативному кільцю R. Ми також вимагатимемо, щоб кільце R *не мало дільників нуля,* тобто щоб

$$ab = 0 \Leftrightarrow a = 0 \vee b = 0$$

для довільних a, $b \in R$ (0 — нульовий елемент кільця).

Як відомо (§ 12), комутативне кільце, в якому не існує дільників нуля, називається областю цілісності. Отже, вважатимемо, що коефіцієнти многочленів, які ми розглядаємо, належать деякій області цілісності R.

Означення 1. *Многочленом* (поліномом) *від однієї змінної над областю цілісності R називається вираз виду* (3), *де n — довільне ціле невід'ємне число, a_n, a_{n-1}, ..., a_1, a_0* — елементи R, *а x* (або x^1), x^2, ..., x^{n-1}, x^n — деякі символи; *x^k називається k-м степенем змінної x* (або невідомого x), *а a_k — k-м коефіцієнтом многочлена* (3) *або коефіцієнтом при x^k* ($k = 0, 1, \ldots, n$).

У читача не повинно виникати непорозуміння в зв'язку з тим, що символам x^k ми не дали ніякого реального тлумачення. Многочлен повністю визначається своїми коефіцієнтами a_n, a_{n-1}, ..., a_1,a_0, а символи x^n, x^{n-1}, ..., x відіграють поки що, так би мовити, роль «розділових знаків», що відокремлюють коефіцієнти один від одного та упорядковують їх. Зауважимо також, що знак «+» у символічному запису (3) поки що не є позначенням якоїсь операції, вираз $a_k x^k$ не є добутком a_k на x^k, а x^k не є добутком k множників x. По суті, многочлен (3) можна було б записати просто — як упорядковану сукупність коефіцієнтів або як $(n+1)$-вимірний вектор $(a_n, a_{n-1}, \ldots, a_1, a_0)$. Проте, як виявиться далі, запис у формі (3) має переваги над векторним записом.

Многочлени від змінної x позначатимемо маленькими латинськими буквами: $f(x)$, $g(x)$, $q(x)$, $s(x)$ і т. п., сукупність усіх многочленів від x над областю цілісності R — символом $R[x]$.

Означення 2. *Вираз $a_k x^k$ ($k = 1, \ldots, n-1, n$) називається k-м членом або членом k-го степеня многочлена*

$$f(x) \underset{\text{df}}{=} a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \qquad (4)$$

a_0 — нульовим або вільним членом, причому записи a_0 і $a_0 x^0$ рівнозначні. Якщо $a_k = 0$ (тобто є нульовим елементом області цілісності R), *то кажуть, що k-й член многочлена $f(x)$ дорівнює нулю або його немає.*

Відповідно до означень 1 і 2 різні члени многочлена від однієї змінної є завжди членами різних степенів. Зауважимо, що k-й коефіцієнт a_k многочлена іноді називають коефіцієнтом його k-го члена.

У виразі для многочлена (4) члени, які дорівнюють нулю, можна не писати. Так, многочлен

$$f(x) \underset{\text{df}}{=} 0x^5 + 2x^4 + 0x^3 + 0x^2 + 0x + 4, \qquad (5)$$

розглядуваний як многочлен над кільцем усіх парних цілих чисел (це кільце, як відомо, є областю цілісності), можна записати коротше:

$$f(x) \underset{\text{df}}{=} 2x^4 + 4. \qquad (6)$$

Обидва записи (5) і (6) містять ту саму інформацію про многочлен, а саме: нульовий коефіцієнт його дорівнює 4, четвертий — 2, решта коефіцієнтів дорівнюють нулю. Роль «розділових символів» x^k у записі многочлена, як бачимо, саме в тому і полягає, щоб цю інформацію зробити незалежною від способу запису. Єдина незручність, яка при цьому залишається, полягає у тому, що ми не можемо однозначно сказати, що розуміємо під «рештою коефіцієнтів». Зрозуміло, що в цю решту напевно входять a_1, a_2, a_3. Але оскільки многочлен (5) можна записати й у вигляді

$$f(x) = 0x^n + 0x^{n-1} + \cdots + 0x^5 + 2x^4 + 0x^3 + 0x^2 + 0x + 4,$$

де n — довільне натуральне число, більше за 4, то «решта» коефіцієнтів визначається неоднозначно (залежно від обраного n).

Означення 3. *Відмінний від нуля член многочлена $f(x)$, степінь якого більший за степінь усіх інших відмінних від нуля членів цього многочлена, називається старшим членом, його коефіцієнт — старшим коефіцієнтом, а його степінь — степенем многочлена $f(x)$.*

Степінь многочлена $f(x)$ позначають [1] $\deg f$.

Домовимось тепер, що будь-який многочлен $f(x)$ записуватимемо так, щоб запис починався з старшого члена, тобто не включати у запис рівних нулю членів, степінь яких більший за $\deg f$. Так, многочлен (6) подаватимемо в будь-якому з виглядів

$$f(x) = 2x^4 + 4; \quad f(x) = 2x^4 + 0x^3 + 4; \quad f(x) = 2x^4 + 0x^2 + 4;$$
$$f(x) = 2x^4 + 0x^3 + 0x^2 + 4$$

тощо, але не у вигляді (5).

[1] Позначення походить від слова degree (англ.), що означає «степінь».

Згідно з цією домовленістю будь-який многочлен n-го степеня подаватимемо, як правило, у вигляді

$$f(x) \underset{\text{df}}{=} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \qquad (7)$$

причому $a_n \neq 0$, а з решти коефіцієнтів частина або всі можуть дорівнювати нулю. Запис (7) характеризується тим, що члени упорядковано за спаданням степеня x^k; таку форму запису називають *канонічною*. Вживають також назву «*многочлен стандартного виду*». Ми переважно користуватимемося канонічною формою запису многочленів, хоч в окремих випадках будуть зручні інші форми (наприклад, розміщення членів у порядку зростання степенів). Зокрема, довільний многочлен першого степеня над областю цілісності R можна записати так: $f(x) \underset{\text{df}}{=} a_1 x + a_0$ ($a_1,\ a_0 \in R,\ a_1 \neq 0$); такі многочлени називають також *лінійними двочленами*. Будь-якому многочлену нульового степеня можна надати вигляду $f(x) \underset{\text{df}}{=} a_0$ ($a_0 \in R,\ a_0 \neq 0$). Многочлени нульового степеня називають також константами. Очевидно, будь-який елемент $a \in R$, відмінний від нульового, можна розглядати як многочлен нульового степеня над R. Елемент $0 \in R$ ми також вважатимемо константою і многочленом над R; цей многочлен називатимемо *нуль-многочленом* і позначатимемо $\theta(x)$, тобто $\theta(x) = 0$. Зрозуміло, що означення 3 незастосовне до $\theta(x)$, так що нуль-многочлену не приписують ніякого степеня. Все ж нам буде зручно вважати, що канонічна форма (6) охоплює і випадок нуль-многочлена, тобто допускати у цій формі при $n = 0$ випадок $a_n = 0$. Слід, отже, мати на увазі, що завжди істинна імплікація $[\deg f = n] \Rightarrow [a_n \neq 0]$, але якщо в канонічній формі $n = 0$, то a_n може бути будь-яким елементом області цілісності R (хоч при $a_0 = 0$ не вважаємо $n = 0$ степенем многочлена). У дальшому викладі ми іноді припускатимемо деяку мовну вільність і тлумачитимемо висловлення «степінь многочлена $f(x)$ менший (не більший) за n» так: «$f(x) = \theta(x)$ або $\deg f < n$ ($\deg f \leqslant n$)». Зокрема, усі константи можна вважати многочленами, степінь яких не більший від нуля (або, як кажуть, многочленами н е в и щ е від нульового степеня).

21.3. Дії над многочленами. Многочлени, як і вектори, становлять інтерес не самі по собі, а як об'єкти деяких алгебраїчних операцій, до означення яких ми й переходимо.

Нехай дано два многочлени над областю цілісності R.

$$f(x) \underset{\text{df}}{=} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \qquad (a_i \in R,\ i = 0, 1, \ldots, n), \qquad (8)$$

$$g(x) \underset{\text{df}}{=} b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 \qquad (b_j \in R,\ j = 0, 1, \ldots, m). \qquad (9)$$

Означимо спочатку поняття р і в н о с т і многочленів (8) і (9).

Означення 1. *Многочлени $f(x)$ і $g(x)$ називають рівними між собою і записують $f(x) = g(x)$, якщо канонічні форми цих многочленів збігаються, тобто*

$$[f(x) = g(x)] \underset{\text{df}}{\equiv} [n = m] \wedge \bigwedge_{k=0}^{n} [a_k = b_k]. \qquad (10)$$

Якщо $f(x)$ і $g(x)$ відмінні від $\theta(x)$, то (10) можна переписати у вигляді

$$[f(x) = g(x)] \underset{\text{df}}{\equiv} [\deg f = \deg g] \wedge \bigwedge_{k=0}^{n} [a_k = b_k], \qquad (11)$$

а для випадку нуль-многочлена можна записати

$$[f(x) = \theta(x)] \underset{\text{df}}{\equiv} [\deg f \leqslant 0] \wedge [a_0 = 0].$$

Нерівність $\deg f \leqslant 0$ тут означає, що $f(x)$ не має членів ненульового степеня.

Читачеві, який звик дивитись на многочлен як на функції, а їхню рівність розуміти як рівність усіх можливих значень цих функцій, нагадаємо, що для нас тепер многочлен — нове абстрактне поняття і що вираз «значення многочлена» поки що позбавлений для нас сенсу. До обговорення цього питання ми повернемось у п. 21.5.

З означення 1 безпосередньо випливає, що рівність многочленів має властивості рефлексивності, симетричності та транзитивності, тобто є відношенням еквівалентності на множині $R[x]$. Природно рівні між собою многочлени вважати тим самим многочленом. Так само природно вживати означення $f(x) \neq g(x)$ для висловлення $\overline{f(x) = g(x)}$, тобто заперечення рівності многочленів $f(x)$ і $g(x)$.

Означимо тепер с у м у многочленів (8) і (9). Без обмеження загальності можна вважати, що $n \geqslant m \geqslant 0$.

Означення 2. *Сумою многочленів $f(x)$ і $g(x)$ називається многочлен*

$$s(x) \underset{\text{df}}{=} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_{m+1}x^{m+1} + (a_m + b_m)x^m + (a_{m-1} + b_{m-1})x^{m-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0). \qquad (12)$$

Те, що $s(x)$ є сумою многочленів $f(x)$ і $g(x)$, записують так: $s(x) = f(x) + g(x)$.

З цього означення, яке є природним узагальненням «шкільної» практики, випливають важливі наслідки:

Наслідок 1. *Якщо $f(x) \in R[x]$, $g(x) \in R[x]$, то й $f(x) + g(x) \in R[x]$.* Справді, для коефіцієнтів d_k многочлена $f(x) + g(x)$ маємо з (12):

$$d_k = \begin{cases} a_k, & k = n, n-1, \ldots, m+1, \\ a_k + b_k, & k = m, m-1, \ldots, 1, 0. \end{cases} \qquad (13)$$

Оскільки R є кільце, то $a_i \in R \wedge b_j \in R \Rightarrow d_k \in R$, тобто $f(x) + g(x) \in R[x]$. Зауважимо, що вираз для коефіцієнтів суми даних многочленів можна записати також у більш простій і симетричній формі, якщо ввести до розгляду символи $b_{m+1}, b_{m+2}, \ldots, b_{n-1}, b_n$, вважаючи їх рівними нулю кільця R. Тоді (13) можна записати так:

$$d_k = a_k + b_k, \quad k = 0, 1, \ldots, n; \quad b_k = 0 \text{ при } k > m. \qquad (14)$$

Наслідок 2. *Степінь суми двох многочленів не перевищує більшого з степенів даних многочленів:* $\deg(f+g) \leqslant \max\{\deg f, \deg g\}$.

З (12) безпосередньо видно, що коли $\deg f = n > m = \deg g$, то $\deg(f+g) = n = \max\{\deg f, \deg g\}$. Але у випадку додавання многочленів однакового степеня ($m = n$) може статись, що $\deg(f+g) < \max\{\deg f, \deg g\}$. Це трапиться тоді, коли старші коефіцієнти многочленів, які додаються, є протилежні елементи кільця R:

$$a_n + b_n = 0 \Rightarrow \deg(f+g) < n.$$

Наслідок 3. *Для довільного многочлена* $f(x) \in R[x]$ *і* $\theta(x) \in R[x]$

$$f(x) + \theta(x) = f(x). \tag{15}$$

Цей наслідок випливає безпосередньо з (12) при $m = 0$, $b_0 = 0$.

Перейдемо тепер до означення множення многочленів.

Означення 3. *Добутком многочленів* $f(x)$ *і* $g(x)$ (див. (8) і (9)) *називається многочлен*

$$p(x) \underset{\text{df}}{=} c_{n+m}x^{n+m} + c_{n+m-1}x^{n+m-1} + \cdots + c_1x + c_0, \tag{16}$$

де

$$c_k \underset{\text{df}}{=} \sum_{i+j=k} a_{k-j}b_j = a_kb_0 + a_{k-1}b_1 + \cdots + a_1b_{k-1} + a_0b_k$$

$(k = 0, 1, \ldots, n+m)$, $a_{k-j} = 0$ при $k - j > n$, $b_j = 0$ при $j > m$. (17)

Те, що $p(x)$ *є добуток многочленів* $f(x)$ *і* $g(x)$, *записують так:* $p(x) = f(x)\,g(x)$ *або* $p(x) = f(x) \cdot g(x)$.

Наведене означення є також безпосереднім узагальненням «шкільного» правила множення многочленів. Щоб у цьому переконатись, досить обчислити за формулою (17) кілька коефіцієнтів многочлена $p(x)$

$$c_{n+m} = a_{n+m}b_0 + a_{n+m-1}b_1 + \cdots + a_{n+1}b_{m-1} + a_nb_m + a_{n-1}b_{m+1} + \cdots + a_0b_{m+n} = a_nb_m, \tag{18}$$

бо у цій сумі всі члени, крім підкресленого, дорівнюють нулю: адже згідно з (17) $a_{n+m} = \ldots = a_{n+1} = b_{m+1} = \ldots = b_{m+n} = 0$. Аналогічно

$$c_{n+m-1} = a_{n+m-1}b_0 + a_{n+m-2}b_1 + \cdots + a_{n+1}b_{m-2} + a_nb_{m-1} + a_{n-1}b_m + a_{n-2}b_{m+1} + \cdots + a_0b_{n+m-1} = a_nb_{m-1} + a_{n-1}b_m;$$

$$\ldots;\ c_1 = a_1b_0 + a_0b_1;\quad c_0 = a_0b_0.$$

Як бачимо, коефіцієнт c_k при x^k в многочлені $p(x) = f(x)\,g(x)$ є сумою всіх можливих попарних добутків коефіцієнтів даних многочленів таких, що сума індексів співмножників (або, що те саме, сума степенів відповідних членів) дорівнює k. Саме так і утворюються коефіцієнти многочлена-добутку за «шкільним» правилом, коли всі члени даних многочленів попарно перемножають, а потім «зводять подібні члени», тобто додають коефіцієнти при однакових степенях x.

Зауважимо, що вираз (17) для коефіцієнтів добутку многочленів можна подати в більш зручній формі, а саме

$$c_k = \sum_{i+j=k} a_ib_j, \tag{19}$$

зрозуміло, що суму беруть по всіх індексах i та j, для яких $i + j = k$. Тут, як і раніше, $a_i = 0$ при $i > n$, $b_j = 0$ при $j > m$.

Щоб не робити щоразу подібні застереження, домовимося про таке. Якщо $f(x)$ — будь-який многочлен над R, a_i — його коефіцієнти, $n = \deg f$, то символ a_i при $i > n$ означає нульовий елемент кільця R.

Наслідок 4. *Якщо* $f(x) \in R[x]$, $g(x) \in R[x]$, *то й* $f(x)\,g(x) \in R[x]$. Це випливає з формули (19) для коефіцієнтів многочлена $f(x)\,g(x)$ з урахуванням того, що всі a_i та b_j ($i, j = 1, 2, \ldots, n+m$) належать R (бо є або коефіцієнти даних многочленів $f(x)$, $g(x) \in R[x]$, або нулі).

Наслідок 5. *Якщо* $f(x)$ *і* $g(x)$ *не є нуль-многочлени, то*

$$\deg(fg) = \deg f + \deg g. \tag{20}$$

Справді, нехай $\deg f = n$, $\deg g = m$. Оскільки $f(x) \neq \theta(x)$ і $g(x) \neq \theta(x)$, то $a_n \neq 0$, $b_m \neq 0$. Але тоді згідно з (18) $c_{n+m} = a_nb_m \neq 0$, бо R є область цілісності і тому в R не існує дільників нуля:

$$a_nb_m = 0 \Rightarrow a_n = 0 \vee b_m = 0,$$

що неможливо. Звідси випливає (20).

Зауважимо, що саме при доведенні наслідку 5 ми вперше використали те, що R — не просто кільце, а область цілісності.

Наслідок 6.

$$[f(x) = \theta(x)] \vee [g(x) = \theta(x)] \Rightarrow [f(x)\,g(x) = \theta(x)], \tag{21}$$

тобто при множенні двох многочленів, з яких хоч один є нуль-многочленом, дістаємо нуль-многочлен.

Цей наслідок безпосередньо випливає з (17).

В п. 21.2 ми домовились розглядати елементи $a \in R$ як многочлени не вище нульового степеня над R. Після того як введено означення рівності між многочленами та дій над ними, потрібно переконатись, що в застосуванні до елементів кільця R як до многочленів не вище нульового степеня, ці означення збігаються з означеннями рівності та відповідних операцій у кільці R.

Але це справді так, бо якщо $f(x) = a_0$, $g(x) = b_0$ — многочлени не вище нульового степеня над R (зокрема, це можуть бути і нуль-многочлени), то

$$[f(x) = g(x)] \equiv [a_0 = b_0],\quad f(x) + g(x) = a_0 + b_0,$$
$$f(x)\,g(x) = a_0b_0,$$

тобто рівність цих многочленів рівносильна рівності відповідних елементів кільця R, а сума (добуток) цих многочленів є сумою (добутком) відповідних елементів R (у розумінні операцій в кільці R).

Тим самим перевірено коректність розгляду елементів області цілісності R як многочленів над R не вище нульового степеня. Точніше,

показано ізоморфізм між сукупністю усіх многочленів не вище нульового степеня над R і самим кільцем R. Саме це і дає право ототожнювати многочлен не вище нульового степеня з відповідним елементом кільця R.

Зокрема, нуль-многочлен $\theta(x)$ можна розглядати як нульовий елемент кільця R, в зв'язку з чим ми у більшості випадків замість $\theta(x)$ писатимемо 0.

21.4. Кільце многочленів над областю цілісності. Нехай, як і раніше, R — якась область цілісності, $R[x]$ — сукупність усіх многочленів над R. Звичайно, властивості $R[x]$ визначаються властивостями R і можуть бути істотно різними для різних областей цілісності R. Наприклад, многочлен $f(x) \underset{\text{df}}{=} 4x^2 + 1$ з числовими коефіцієнтами $a_2 = 4$, $a_1 = 0$, $a_0 = 1$ можна розглядати і як многочлен над кільцем $\mathbf{Z}$ усіх цілих чисел, і як многочлен над полями $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ усіх раціональних, дійсних та комплексних чисел відповідно (всі ці структури є областями цілісності). Але розглядуваний як елемент різних із цих сукупностей, $f(x)$ має р і з н і властивості. Так, у $\mathbf{C}[x]$ многочлен $f(x)$ розкладається на лінійні множники, тобто може бути поданий як добуток двох лінійних двочленів з $\mathbf{C}[x]$:

$$f(x) \underset{\text{df}}{=} 4x^2 + 1 = (2x + i)(2x - i) = g(x)\,h(x),$$

де $g(x) \underset{\text{df}}{=} 2x + i \in \mathbf{C}[x]$, $h(x) \underset{\text{df}}{=} 2x - i \in \mathbf{C}[x]$. В той же час у множинах $\mathbf{Z}[x]$, $\mathbf{Q}[x]$, $\mathbf{R}[x]$ таких двочленів $g(x)$ і $h(x)$ не існує, і вказаний розклад неможливий. Отже, властивість $f(x)$ розкладатись на лінійні множники залежить від того, над якою областю цілісності розглядаємо цей многочлен.

Проте існують властивості сукупності $R[x]$ многочленів над областю цілісності R, які не залежать від специфічних особливостей R, а лише від того, що R є областю цілісності. Саме такі спільні для усіх $R[x]$ властивості ми тут і розглянемо.

Теорема 1. *Сукупність $R[x]$ усіх многочленів над областю цілісності R є область цілісності відносно операцій додавання та множення многочленів.*

Д о в е д е н н я. З наслідків 1 та 4 п. 21.3 випливає, що сукупність $R[x]$ є *комутативне кільце* відносно згаданих операцій. Справді, нехай $f(x)$, $g(x)$ — будь-які многочлени з $R[x]$. Згідно з формулами (14) і (19) для коефіцієнтів суми і добутку многочленів відповідно, помічається, що $f(x) + g(x) = g(x) + f(x)$ і $f(x)\,g(x) = g(x)\,f(x)$, бо область цілісності R є комутативне кільце і

$$a_k + b_k = b_k + a_k, \quad \sum_{i+j=k} a_i b_j = \sum_{i+j=k} b_i a_j.$$

Розглянемо тепер довільні три многочлени з $R[x]$:

$$f(x) \underset{\text{df}}{=} \sum_{k=1}^{n} a_k x^k; \quad g(x) \underset{\text{df}}{=} \sum_{k=0}^{m} b_k x^k; \quad h(x) \underset{\text{df}}{=} \sum_{k=0}^{l} c_k x^k.$$

Асоціативність додавання многочленів

$$[f(x) + g(x)] + h(x) = f(x) + [g(x) + h(x)] \tag{22}$$

випливає безпосередньо з асоціативності додавання елементів у кільці R; адже за означенням рівності многочленів співвідношення (22) рівносильне системі рівностей

$$(a_k + b_k) + c_k = a_k + (b_k + c_k), \qquad k = 1, 2, \ldots, \max\{n, m, l\}.$$

Нульовим елементом в $R[x]$ є, очевидно, нуль-многочлен $\theta(x)$ (див. наслідок 3 п. 21.3). Для довільного многочлена $f(x) \underset{\text{df}}{=} \sum_{k=0}^{n} a_k x^k \in R[x]$ існує в $R[x]$ протилежний елемент, а саме — многочлен $\bar{f}(x) = \sum_{k=0}^{n} (-a_k)\,x^k$. Для перевірки *асоціативності множення* многочленів

$$[f(x)\,g(x)]\,h(x) = f(x)\,[g(x)\,h(x)]$$

доведеться виконати деякі обчислення. Якщо через p_k позначити k-й коефіцієнт многочлена $f(x)\,g(x)$, то згідно з (19)

$$p_k = \sum_{i+j=k} a_i b_j.$$

Але тоді s-й коефіцієнт многочлена $[f(x)\,g(x)]\,h(x)$ є

$$\sum_{k+l=s} p_k c_l = \sum_{k+l=s} \Big(\sum_{i+j=k} a_i b_j \Big) c_l = \sum_{i+j+l=s} a_i b_j c_l. \tag{23}$$

З другого боку, позначаючи через q_r r-й коефіцієнт многочлена $g(x) \times h(x)$, дістаємо для s-го коефіцієнта многочлена $f(x)\,[g(x)\,h(x)]$ такий вираз:

$$\sum_{i+r=s} a_i q_r = \sum_{i+r=s} a_i \Big(\sum_{j+l=r} b_j c_l \Big) = \sum_{i+j+l=s} a_i b_j c_l. \tag{24}$$

Зіставляючи (23) і (24), пересвідчуємось у справедливості асоціативного закону для множення многочленів.

Дистрибутивний закон

$$[f(x) + g(x)]\,h(x) = f(x)\,h(x) + g(x)\,h(x)$$

випливає з такої очевидної рівності:

$$\sum_{i+j=k} (a_i + b_i)\,c_j = \sum_{i+j=k} a_i c_j + \sum_{i+j=k} b_i c_j.$$

Тим самим показано, що $R[x]$ — комутативне кільце. Залишається встановити, що це кільце є *область цілісності*, тобто що в $R[x]$ не існує дільників нуля. Але це справді так. Нехай $f(x)$ і $g(x)$ — многочлени з $R[x]$, які не є нуль-многочленами, і їх старші коефіцієнти є a_n та b_m відповідно; отже, $a_n \neq 0$, $b_m \neq 0$. Многочлен $f(x)\,g(x)$ має тоді, згідно з (18), старший коефіцієнт $a_n b_m$, відмінний від нуля (адже в R не має дільників нуля), і тому $f(x)\,g(x)$ не є нуль-многочлен.

Отже,

$$[f(x) \neq \theta(x)] \wedge [g(x) \neq \theta(x)] \Rightarrow [f(x)\,g(x) \neq \theta(x)],$$

звідки з урахуванням (21) випливає, що *добуток двох многочленів є нуль-многочлен тоді і тільки тоді, коли хоч один з цих многочленів є нуль-многочлен.*

Теорему доведено.

Зауважимо, що відсутність дільників нуля в R використано лише при доведенні відсутності дільників нуля в $R[x]$ і пов'язаної з цим формули (20).

Можна було б розглядати сукупність многочленів $C[x]$ над комутативним кільцем C, що не є областю цілісності; при цьому $C[x]$ також була б комутативним кільцем і не була б областю цілісності. Йдучи далі по шляху узагальнень, можна було б відмовитись і від вимоги комутативності кільця C, але це зробило б властивості дій в $C[x]$ надто складними.

В усіх застосуваннях теорії многочленів, з якими нам доведеться зустрічатись, цілком досить розглядати сукупності многочленів над областями цілісності; як показує теорема 1, ці сукупності також є областями цілісності, тобто мають досить «добрі» алгебраїчні властивості, які можна покласти в основу стрункої теорії.

Можна вказати ще одну «добру» властивість кільця R, яка переноситься на $R[x]$. Справедливе таке твердження:

$R[x]$ є кільце з одиницею тоді і тільки тоді, коли R є кільце з одиницею.

Справді, якщо в R існує одиниця 1, то вона, розглядувана як многочлен нульового степеня, є одиницею і в кільці $R[x]$, бо $f(x) \cdot 1 = f(x)$. Навпаки, нехай $e(x)$ — одиниця кільця многочленів, тобто для довільного $f(x) \in R[x]$

$$f(x)\, e(x) = f(x).$$

Зрозуміло, що $e(x)$ відмінний від нуль-многочлена і є многочлен нульового степеня (бо $\deg f + \deg e = \deg f \Rightarrow \deg e = 0$), тобто ненульовий елемент кільця R. Далі, з рівності $f(x)\, e = f(x)$ випливає $a_0 e = a_0$, де $a_0 \in R$ — вільний член многочлена $f(x)$. Оскільки $f(x)$ — довільний многочлен з $R[x]$, то a_0 — довільний елемент з R. Тому $a_0 e = a_0 \Rightarrow e = 1$, тобто e — одиниця кільця R. Цим наше твердження доведено.

Оскільки $R[x]$ є кільце, можна розглядати *різницю* будь-яких многочленів $f(x) = \sum\limits_{k=0}^{n} a_k x^k$, $g(x) = \sum\limits_{k=0}^{m} b_k x^k$, узявши

$$f(x) - g(x) \underset{\text{df}}{=} f(x) + \bar{g}(x) = \sum_{k=0}^{l} [a_k + (-b_k)]\, x^k =$$

$$= \sum_{k=0}^{l} (a_k - b_k)\, x^k; \quad l = \max\{m, n\}.$$

В зв'язку з тим, що всі елементи $a \in R$ можна розглядати як многочлени над R, то в $R[x]$ означено не тільки дії додавання і множення, а й операцію *множення на елементи з R*. Ураховуючи це, можна показати, що $R[x]$ є лінійний простір і алгебра над R.

Беручи до уваги введені означення дій над многочленами та їх властивості, можемо по-новому подивитись на вираз (4) многочлена $f(x)$. Досі ми вважали (4) просто деяким символічним записом. Але будь-який член $a_k x^k$ многочлена $f(x)$ можна розглядати як многочлен k-го степеня (при $a_k \neq 0$) або нуль-многочлен (при $a_k = 0$) над R. Якщо взяти

$$u_k(x) \underset{\text{df}}{=} a_k x^k \qquad (k = 0,\ 1,\ \ldots,\ n-1,\ n)$$

і урахувати, що внаслідок асоціативності додавання многочленів набуває цілком певного змісту поняття суми довільного скінченного числа многочленів, то знайдемо, що (побіжно використовуємо і комутативність додавання)

$$f(x) = \sum_{k=0}^{n} u_k(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Це показує, що на знак «+» в записі (4) можна дивитись не просто як на умовний символ, а як на позначення операції додавання многочленів.

Якщо R — *кільце з одиницею*, то в $R[x]$ є многочлен

$$p(x) \underset{\text{df}}{=} 1x = x.$$

Але тоді для довільного натурального k за означенням добутку многочленів

$$[p(x)]^k = \underbrace{p(x)\, p(x)\ \ldots\ p(x)}_{k} = x^k,$$

а k-й член $u_k(x)$ є добутком многочлена a_k, не вище нульового степеня, на многочлен x^k. Отже, остаточно:

$$a_k x^k = a_k \underbrace{\cdot\, x \cdot x \cdot\ \ldots\ \cdot x}_{k}, \tag{25}$$

тобто k-й член многочлена $f(x)$ можна розглядати як добуток многочленів, а вираз (4) многочлена $f(x)$ в цілому — як суму добутків многочленів:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0. \tag{26}$$

Правила (12) і (16)—(17) додавання та множення многочленів тепер можна розглядати як запис результату дій над сумами добутків многочленів виду (26) з застосуванням комутативних, асоціативних та дистрибутивного законів для додавання і множення. Зауважимо, що так зване зведення подібних членів, тобто заміна виразу $ax^k + bx^k$ виразом $(a + b)\, x^k$, спирається на справедливість дистрибутивного закону.

Разом з цим важливо розуміти, що означень 2 та 3 дій над многочленами ми цим не замінюємо і відкинути ці означення не можна. Адже саме на основі згаданих означень ми з'ясували властивості додавання і множення та дістали право дивитись на формальний запис многочлена $a_n x^n + \ldots + a_1 x + a_0$ як на суму добутків більш простих многочленів.

Крім того, не слід забувати, що останні міркування справедливі при тій д о д а т к о в і й умові, що в області цілісності R існує одиничний елемент. Як відомо, це не завжди так (наприклад, кільце усіх парних цілих чисел є область цілісності без одиниці).

21.5. Функціональне тлумачення многочлена. Нехай $f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ — многочлен над областю цілісності R, а C — деяке комутативне кільце, що є розширенням R. Якщо α — будь-який елемент з C, то має сенс такий вираз:

$$a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_1 \alpha + a_0, \tag{27}$$

бо в C визначено дії множення і додавання над елементами α, a_n, a_{n-1}, ..., a_1, a_0.

Вираз (27) утворений з $f(x)$ заміною символа x елементом α. У зв'язку з цим його позначають $f(\alpha)$ і називають *значенням многочлена* $f(x)$ *при* $x = \alpha$ (або «в точці α»). Очевидно, $f(\alpha)$ є елементом C: $f(\alpha) \in C$. Кожному $\alpha \in C$ відповідає за цим правилом єдиний цілком певний елемент $f(\alpha) \in C$. Пригадуючи загальне означення функції (відображення), можемо висловити таке твердження:

Теорема 2. *Якщо $f(x)$ — будь-який многочлен над областю цілісності R, а C — деяке комутативне кільце, яке є розширенням R, то, поставивши кожному елементу $\alpha \in C$ у відповідність елемент $f(\alpha) \in C$, дістаємо функцію φ_f: $C \to C$; $\varphi_f(\alpha) = f(\alpha)$.*

Іншими словами, *многочлен $f(x) \in R[x]$ визначає відображення φ_f будь-якого комутативного розширення C кільця R у себе.*

Приклади. 1. Якщо $R = \mathbf{Z}$, $C = \mathbf{R}$, то φ_f є многочлен з цілими коефіцієнтами, розглядуваний як функція дійсної змінної.

2. Якщо $R = \mathbf{C}$, $C = \mathbf{C}$, то многочлен $f(x)$ з комплексними коефіцієнтами визначає функцію φ_f комплексної змінної (тобто є відображенням множини $\mathbf{C}$ усіх комплексних чисел у себе).

3. Якщо $f(x)$ — многочлен нульового степеня над R, тобто $f(x) = a_0$, і C — будь-яке розширення R, то $f(\alpha)$ не залежить від α і є сталим: $\forall\limits_{\alpha \in C} [f(\alpha) = a_0]$. Саме тому многочлени нульового степеня називають також константами.

Наведені міркування пояснюють, чому в школі, у курсі аналізу та у нашому попередньому викладі (частина 1) можна було розглядати многочлен як деяку функцію дійсної чи комплексної змінної.

Але у цьому питанні слід проявити певну обережність. Адже ми розглядаємо многочлени не самі по собі, а як об'єкти певних дій. У п. 21.3 було введено означення операцій додавання і множення многочленів, а також означення рівності многочленів. З другого боку, ми знаємо загальні означення додавання, множення та рівності двох функцій. Функціональне тлумачення введених у цьому розділі многочленів стане коректним лише після того, як ми пересвідчимося, що сума, добуток і рівність двох многочленів, розглядуваних як функції, є те саме, що й сума, добуток і рівність двох многочленів за означенням п. 21.3.

Як відомо, сумою (добутком) двох функцій φ: $X \to Y$ і ψ: $X \to Y$ називається функція χ: $X \to Y$ така, що

$$\forall_{x \in X} [\chi(x) = \varphi(x) + \psi(x)] \qquad \left(\forall_{x \in X} [\chi(x) = \varphi(x)\,\psi(x)]\right).$$

Звичайно, це означення передбачає, що у множині Y означено операцію додавання (множення) елементів.

Нехай тепер $f(x)$, $g(x)$ — многочлени над R і $s(x) = f(x) + g(x)$; $p(x) \underset{\text{df}}{=} f(x)\,g(x)$ — сума і добуток цих многочленів у розумінні п. 21.3. Як відомо, $s(x) \in R[x]$ і $p(x) \in R[x]$.

Якщо C — деяке комутативне розширення кільця R, то можна розглянути функції f: $C \to C$; g: $C \to C$; s: $C \to C$; p: $C \to C$. Щоб функцію s можна було розглядати як суму, а функцію p — як добуток функцій f і g, слід довести таке твердження:

Теорема 3. *Нехай* $f(x) \underset{\text{df}}{=} \sum\limits_{k=0}^{n} a_k x^k$, $g(x) \underset{\text{df}}{=} \sum\limits_{k=0}^{m} b_k x^k$ *— многочлени над областю цілісності R,* $s(x) = f(x) + g(x)$, $p(x) = f(x)\,g(x)$, *а C — будь-яке комутативне розширення кільця R. Тоді*

$$\forall_{\alpha \in C} [s(\alpha) = f(\alpha) + g(\alpha)], \qquad \forall_{\alpha \in C} [p(\alpha) = f(\alpha)\,g(\alpha)], \tag{28}$$

тобто

$$\varphi_s = \varphi_f + \varphi_g; \quad \varphi_p = \varphi_f \cdot \varphi_g. \tag{29}$$

Доведення. Згідно з означеннями 2 і 3 п. 21.3

$$s(x) = \sum_{k=0}^{l} (a_k + b_k)\,x^k, \text{ де } l = \max\{m, n\};$$

$$p(x) = \sum_{k=0}^{m+n} c_k x^k, \text{ де } c_k = \sum_{i+j=k} a_i b_j.$$

Для будь-якого $\alpha \in C$ маємо

$$s(\alpha) = \sum_{k=0}^{l} (a_k + b_k)\,\alpha^k, \qquad p(\alpha) = \sum_{k=0}^{m+n} c_k \alpha^k.$$

З другого боку,

$$f(\alpha) = \sum_{k=0}^{n} a_k \alpha^k, \qquad g(\alpha) = \sum_{k=0}^{m} b_k \alpha^k,$$

звідки, ураховуючи властивості комутативного кільця C,

$$f(\alpha) + g(\alpha) = \sum_{k=0}^{n} a_k \alpha^k + \sum_{k=0}^{m} b_k \alpha^k = \sum_{k=0}^{l} (a_k + b_k)\,\alpha^k = s(\alpha);$$

$$f(\alpha)\,g(\alpha) = \sum_{i=0}^{n} a_i \alpha^i \cdot \sum_{j=0}^{m} b_j \alpha^j = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j\right) \alpha^k = \sum_{k=0}^{m+n} c_k \alpha^k = p(\alpha)$$

(нагадаємо, що, як завжди, $a_i = 0$ при $i > n$, $b_j = 0$ при $j > m$). Цим доведено (28), а тому й теорему.

Отже, введені в п. 21.3 означення суми і добутку многочленів узгоджені з загальними поняттями суми і добутку функцій.

Зауваження. Ми вимагали, щоб C було **комутативним** розширенням кільця R. Проте для того, щоб сказане про функціональний сенс многочлена було справедливим, досить, щоб C було кільцем, *усі елементи якого комутують з будь-яким елементом комутативного підкільця R.* Справді, при множенні значень $f(\alpha)$ і $g(\alpha)$ ми користуємось лише тим, що

$$(a_i \alpha^i)\,(b_j \alpha^j) = a_i b_j \alpha^{i+j},$$

тобто можливістю переставляти α з коефіцієнтами a_i, b_j, які є елементами R.

Спираючись на це зауваження, можна, зокрема, розглядати многочлени як функції від **матриць**. Точніше, нехай R — будь-яка область цілісності, а M_n — кільце квадратних матриць n-го порядку, елементи яких належать R. Як відомо (1, § 30), M_n можна розглядати як розширення кільця R, ототожнюючи кожний елемент $a \in R$ з діагональною матрицею

$$\alpha = \begin{pmatrix} a & 0 & 0 & \dots & 0 \\ 0 & a & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & & a \end{pmatrix} \in M_n,$$

оскільки $a \to \alpha$ є ізоморфізм. При цьому будь-який елемент $\beta \in M_n$ переставний з будь-яким елементом $a \in R$ (адже діагональні матриці переставні з довільними матрицями з M_n). Саме ж кільце M_n, як відомо, не є комутативним.

Отже, якщо $f(x) = \sum_{k=0}^{n} a_k x^k$ — многочлен над R, а A — будь-яка матриця з M_n, то

$$f(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0$$

також є матрицею з M_n і $f(x)$ можна розглядати як функцію

$$\varphi_f : M_n \to M_n.$$

Функції (зокрема, многочлени) від матриць широко використовують у математиці. Все ж у дальшому викладі, якщо не зазначене супротивне, ми завжди вважатимемо C комутативним розширенням R.

Перейдемо тепер до розгляду питання про р і в н і с т ь многочленів. Згідно з означенням 1 п. 21.3, для многочленів $f(x) = \sum_{k=0}^{n} a_k x^k$, $g(x) = \sum_{k=0}^{m} b_k x^k$

$$[f(x) = g(x)] \underset{\text{df}}{\equiv} [m = n] \wedge \bigwedge_{k=0}^{n} [a_k = b_k].$$

З другого боку, дві функції $\varphi\colon X \to Y$; $\psi\colon X \to Y$ вважають рівними між собою, якщо

$$\forall_{x \in X} [\varphi(x) = \psi(x)],$$

тобто, якщо значення цих функцій рівні між собою при всіх $x \in X$ (або, як кажуть, якщо ці функції т о т о ж н о р і в н і на X).

Отже, рівносильність алгебраїчного і функціонального означень рівності многочленів $f(x)$ і $g(x)$ означає справедливість еквіваленції

$$\forall_{\alpha \in C} [f(\alpha) = g(\alpha)] \Leftrightarrow [m = n] \wedge \bigwedge_{k=0}^{n} [a_k = b_k], \tag{30}$$

де C — будь-яке комутативне розширення кільця R. Зрозуміло, що (30) буде істинне для довільного C, якщо буде істинне для $C = R$; тому замість (30) досить написати

$$\forall_{\alpha \in R} [f(\alpha) = g(\alpha)] \Leftrightarrow [m = n] \wedge \bigwedge_{k=0}^{n} [a_k = b_k]. \tag{31}$$

Слід відразу зауважити, що у випадку д о в і л ь н о ї області цілісності R справедливість (31) стверджувати н е м о ж н а. Нехай, наприклад, R є поле $\mathbf{Z}/(2) = \{\bar{0}, \bar{1}\}$ класів лишків цілих чисел за модулем 2 (див. пп. 13.2, 14.2). Для многочленів $f(x) = x^2 + x$ (тобто $f(x) \underset{\text{df}}{=} \bar{1}x^2 + \bar{1}x + \bar{0}$) і $g(x) = \bar{0}$ з $\mathbf{Z}/_{(2)}[x]$ маємо: $f(\bar{0}) = \bar{0} = g(\bar{0})$, $f(\bar{1}) = \bar{1} + \bar{1} = \bar{0} = g(\bar{1})$, тобто $\forall_{\alpha \in \mathbf{Z}/(2)} [f(\alpha) = g(\alpha)]$. Але, звичайно, многочлени $f(x)$ і $g(x)$ не дорівнюють один одному в алгебраїчному розумінні, бо права частина еквіваленції (31) в цьому випадку хибна.

Проте можна навести умови, які досить накласти на область цілісності R, щоб справджувалась еквіваленція (31).

Звичайно, імплікація

$$[m = n] \wedge \bigwedge_{k=0}^{n} [a_k = b_k] \Rightarrow \forall_{\alpha \in R} [f(\alpha) = g(\alpha)]$$

очевидна для будь-якого R. Питання полягає в тому, щоб знайти умови, при яких має місце обернене твердження

$$\forall_{\alpha \in R} [f(\alpha) = g(\alpha)] \Rightarrow [m = n] \wedge \bigwedge_{k=0}^{n} [a_k = b_k].$$

Зазначимо, що рівність многочленів $f(x) = \sum_{k=0}^{n} a_k x^k$, $g(x) = \sum_{k=0}^{m} b_k x^k$, очевидно, рівносильна тому, що їх різниця

$$f(x) - g(x) = \sum_{k=0}^{l} d_k x^k \underset{\text{df}}{=} q(x) \qquad (l = \max\{m, n\})$$

є нуль-многочлен. Отже,

$$[f(x) = g(x)] \equiv [q(x) = \theta(x)] \equiv [\deg q \leqslant 0] \wedge [d_0 = 0]. \tag{32}$$

З другого боку,

$$\forall_{\alpha \in R} [f(\alpha) = g(\alpha)] \equiv \forall_{\alpha \in R} [q(\alpha) = 0]. \tag{33}$$

Ураховуючи (32) і (33), можна замість еквіваленції (31) доводити еквіваленцію

$$\forall_{\alpha \in R} [q(\alpha) = 0] \Leftrightarrow [\deg q \leqslant 0] \wedge [d_0 = 0]. \tag{34}$$

Отже, достатні умови справедливості (34) є також достатніми умовами справедливості (31).

Теорема 4. *Якщо R — область цілісності характеристики 0, то многочлен $q(x) \in R[x]$ є нуль-многочленом тоді і тільки тоді, коли його значення в усіх точках області R дорівнюють нулю* (або $q(x)$ дорівнює нулю тотожно на R).

Д о в е д е н н я. Як випливає з попереднього зауваження, для доведення (34) досить довести імплікацію

$$\forall_{\alpha \in R} [q(\alpha) = 0] \Rightarrow [\deg q \leqslant 0] \wedge [d_0 = 0]. \tag{35}$$

Покажемо спочатку, що

$$\forall_{\alpha \in R} [q(\alpha) = 0] \wedge [\deg q \leqslant 0] \Rightarrow [d_0 = 0].$$

Справді, якщо $\deg q \leqslant 0$, то $\forall_{\alpha \in R} [q(\alpha) = d_0]$ $(d_0 \in R)$ і тому

$$\forall_{\alpha \in R} [q(\alpha) = 0] \wedge [\deg q \leqslant 0] \Rightarrow \forall_{\alpha \in R} [q(\alpha) = 0] \wedge$$
$$\wedge \forall_{\alpha \in R} [q(\alpha) = d_0] \Rightarrow [d_0 = 0].$$

Отже, будь-який многочлен не вище нульового степеня, який тотожно дорівнює нулю на R, є нуль-многочлен. Для доведення імплікації (35) і всієї теореми встановимо правильність такого твердження: *якщо многочлен* $q(x) \in R[x]$ *тотожно дорівнює нулю, то він є многочленом, не вище нульового степеня.* Якщо l — будь-яке натуральне число ($l \geqslant 1$), то слід показати, що $\deg q \neq l$. Це останнє твердження доведемо індукцією по l. (Принцип математичної індукції використовується тут у формі теореми 12 з § 9, ч. 1).

Нехай $l = 1$ і припустимо, що $\deg q = 1$, тобто $q(x) = d_1x + d_0$, де $d_1 \neq 0$. Оскільки за умовою теореми $q(0) = 0$, то $d_0 = 0$. Нехай тепер $\alpha \in R$ і $\alpha \neq 0$. Тоді

$$q(\alpha) = d_1\alpha = 0.$$

Оскільки в R немає дільників нуля і $\alpha \neq 0$, то

$$d_1\alpha = 0 \Rightarrow d_1 = 0.$$

Ми дістали суперечливість з припущенням ($\deg q = 1$), звідки випливає справедливість нашого твердження при $l = 1$.

Нехай це твердження справедливе при усіх цілих l, для яких $1 \leqslant l < s$, і доведемо його справедливість при $l = s$. Припустимо, що $q(x) = \sum\limits_{k=0}^{s} d_kx^k$ має степінь s, тобто $d_s \neq 0$. За умовою теореми для довільного $\alpha \in R$

$$q(\alpha) = d_s\alpha^s + d_{s-1}\alpha^{s-1} + \ldots + d_1\alpha + d_0 = 0.$$

Розглянемо значення $q(x)$ у точці $\beta = 2\alpha$ (тобто $\beta = \alpha + \alpha$). Враховуючи, що $(2\alpha)^k = 2^k\alpha^k$ ($k = 1, 2, \ldots$) і що 2α — також елемент R, маємо

$$q(2\alpha) = 2^sd_s\alpha^s + 2^{s-1}d_{s-1}\alpha^{s-1} + \ldots + 2d_1\alpha + d_0 = 0.$$

З другого боку, усі кратні елемента $q(\alpha)$ дорівнюють нулю, тому

$$2^sq(\alpha) = 2^sd_s\alpha^s + 2^sd_{s-1}\alpha^{s-1} + \ldots + 2^sd_1\alpha + 2^sd_0 = 0.$$

Але тоді й $2^sq(\alpha) - q(2\alpha) = 0$ для довільного $\alpha \in R$, тобто

$$\forall_{\alpha\in R}\,[(2^s - 2^{s-1})d_{s-1}\alpha^{s-1} + (2^s - 2^{s-2})d_{s-2}\alpha^{s-2} + \ldots + (2^s - 2)d_1\alpha + (2^s - 1)d_0 = 0]. \qquad (36)$$

Розглянемо многочлен

$$u(x) \underset{\text{df}}{=} h_{s-1}x^{s-1} + h_{s-2}x^{s-2} + \ldots + h_1x + h_0,$$

де $h_k \underset{\text{df}}{=} (2^s - 2^k)d_k$, $k = 0,1, 2, \ldots, s-1$.

Очевидно, $u \in R[x]$ і $\deg u \leqslant s - 1$, тому за припущенням індукції наше твердження правильне для многочлена $u(x)$. Оскільки (36) означає, що $\forall\limits_{\alpha\in R} u(\alpha) = 0$, то звідси випливає, що $\deg u \leqslant 0$, тобто

$$h_k = (2^s - 2^k)d_k = 0, \quad k = 0,1, 2, \ldots, s-1. \qquad (37)$$

Тепер нам потрібно використати те, що R має характеристику 0, тобто якщо $\alpha \in R$ і $\alpha \neq 0$, то будь-яке його кратне $n\alpha \neq 0$ при $n \neq 0$. Оскільки $n_k = 2^s - 2^k$ при $k = 0,1, 2, \ldots, s-1$ — відмінні від нуля цілі числа, то

$$(2^s - 2^k)d_k = 0 \Rightarrow d_k = 0 \quad (k = 0,1, 2, \ldots, s-1).$$

Отже, $q(x)$ має вигляд

$$q(x) = d_sx^s, \qquad d_s \neq 0.$$

При $\alpha \neq 0$

$$q(\alpha) = 0 \Leftrightarrow d_s\alpha^s = 0 \Rightarrow d_s = 0.$$

Ми дістали суперечність з припущенням $d_s \neq 0$ і цим довели наше твердження, а отже, і теорему.

Наслідок. *Якщо* R *— область цілісності характеристики* 0, *то многочлени* $f(x), g(x) \in R[x]$ *рівні між собою тоді і тільки тоді, коли їх значення в довільній точці області* R *рівні між собою.*

На основі викладеного можна зробити такий висновок.

Нехай R — область цілісності характеристики 0. З теореми 2 видно, що кожному многочлену $f(x) \in R[x]$ відповідає певна функція $\varphi_f\colon R \to R$. Відповідність $f \to \varphi_f$ *взаємно однозначна,* бо з наслідку теореми 4 випливає, що для будь-яких $f(x), g(x) \in R[x]$

$$\varphi_f = \varphi_g \Leftrightarrow \forall_{\alpha\in R}\,[f(\alpha) = g(\alpha)] \Rightarrow f(x) = g(x).$$

Далі, теорема 3 свідчить про те, що ця відповідність є *ізоморфізм,* бо внаслідок (29) $\varphi_{f+g} = \varphi_f + \varphi_g$, $\varphi_{fg} = \varphi_f \cdot \varphi_g$. Це й означає, що *у випадку, коли* R *є область цілісності* (зокрема, поле) *характеристики* 0, *має місце повна рівноправність алгебраїчного та функціонального тлумачень многочленів.* Для довільної ж області цілісності теорема 4, а тому й щойно наведене твердження неправильні.

Зокрема, алгебраїчне і функціональне тлумачення многочленів цілком рівнозначні, коли многочлени розглядати над ч и с л о в и м и полями (1, § 16), бо будь-яке числове поле має характеристику 0. З цього видно, що всі відомості, одержані з курсів алгебри та математичного аналізу про многочлени з дійсними чи комплексними коефіцієнтами, розглядувані як цілі раціональні функції однієї змінної, зберігають силу і при новому підході до поняття многочлена, викладеному в п. п. 21.2—21.3.

§ 22. ТЕОРІЯ ПОДІЛЬНОСТІ МНОГОЧЛЕНІВ

22.1. Многочлени над полем. Досі ми розглядали кільце многочленів над областю цілісності R. Хоч означення кільця многочленів $R[x]$ було дано для д о в і л ь н о ї області цілісності R, для деяких змістовних результатів, як ми бачили (п. 21.5), слід було додатково вима-

гати, щоб в R існувала одиниця і щоб характеристика кільця R дорівнювала нулю.

Тепер ми вимагатимемо, щоб область цілісності R була п о л е м, тобто щоб в R для довільного елемента $a \neq 0$ існував обернений елемент a^{-1}, або щоб

$$\underset{a,b\in R}{\forall}\left[a\neq 0 \Rightarrow \underset{c\in R}{\exists}\left[c = a^{-1}b \underset{\text{df}}{=} \frac{b}{a}\right]\right].$$

Отже, в дальшому викладі розглядатимемо *многочлени над полем* P. Оскільки будь-яке поле є областю цілісності з одиницею, то всі результати § 21 залишаються в силі для цих многочленів. Зокрема, *сукупність усіх многочленів над полем P є область цілісності з одиницею $P[x]$ відносно додавання і множення многочленів.*

Природно, що при цьому ми дещо звужуємо клас розглядуваних многочленів, але в усіх питаннях, які вивчатимуться в цьому курсі (зокрема, у застосуванні теорії многочленів до розв'язування алгебраїчних рівнянь та їх систем), поняття многочлена над полем буде цілком достатнім.

Звичайно, можна було б відразу означити многочлени над полем P, а не над областю цілісності R (так і роблять автори багатьох посібників з алгебри). Проте прийнятий у цьому курсі порядок викладу сприяє чіткішому усвідомленню того, які властивості кільця многочленів $P[x]$ залежать від того, що P — к і л ь ц е, а які від того, що P — саме п о л е.

Особливо важливу роль в елементарній алгебрі, в аналізі та теорії функцій, а також у практичних застосуваннях математики відіграють многочлени над числовими полями, тобто многочлени над полем **C** комплексних чисел або його підполями (**R**, **Q** та ін.). Як зазначалося у попередньому параграфі, оскільки числові поля мають характеристику нуль, *для многочленів над числовими полями алгебраїчне означення многочлена рівносильне функціональному.* Це полегшує вивчення їх і дає змогу встановити ряд важливих спеціальних властивостей.

Було б помилкою думати, що у випадку, коли кільце, над яким розглядаються многочлени, є полем P, то й кільце многочленів $P[x]$ виявиться полем. Більше того: *для жодного многочлена ненульового степеня з $P[x]$ не існує оберненого елемента.* Справді, для довільного $f(x) \in P[x]$ такого, що $\deg f \geqslant 1$, рівність $f(x)\,g(x) = 1$ неможлива ні при якому $g(x) \in P[x]$; адже $g(x)$ не може бути нуль-многочленом і тому $\deg fg = \deg f + \deg g \geqslant 1$, звідки $f(x)\,g(x) \neq 1$.

Що ж до многочленів нульового степеня, які є елементами поля P, то для кожного з них обернений елемент в $P[x]$ існує і є також многочленом нульового степеня. Іншими словами, *дільниками одиниці в області цілісності $P[x]$ є многочлени нульового степеня* (відмінні від нуля константи) *і тільки вони.*

Як бачимо, не тільки все кільце $P[x]$ не є полем, а й будь-яке підкільце, яке містить хоч один многочлен ненульового степеня, не є полем.

Із сказаного зрозуміло, що два різні многочлени з $P[x]$, як правило, не діляться один на одного. Все ж для $P[x]$ може бути побудована

теорія подільності, цілком аналогічна теорії подільності цілих чисел, якщо операцію ділення многочленів (обернену до операції множення в $P[x]$) замінити більш загальною операцією *ділення з остачею.*

Аналогія між теорією подільності у кільці цілих чисел і теорією подільності у кільці многочленів є проявом глибокої спільності властивостей згаданих кілець. Суть цієї спільності полягає в тому, що обидва згадані кільця є моделями тієї самої алгебраїчної структури — е в к л і д о в о г о к і л ь ц я. Саме абстрактне поняття евклідового кільця виникло при порівнянні та узагальненні властивостей подільності цілих чисел і многочленів, які в історії математики досить довго вивчалися незалежно в арифметиці і в алгебрі відповідно. Але в сучасному викладі, зокрема, коли читач вже обізнаний з основними властивостями абстрактних евклідових кілець (п. 14.3), ми маємо змогу дістати всі основні факти теорії подільності многочленів за допомогою простої конкретизації теорії подільності в евклідових кільцях. Іншими словами, якщо буде доведено, що кільце многочленів над полем є евклідовим кільцем, можна буде перенести всі результати, встановлені в п. 14.3 для довільних евклідових кілець, на кільце многочленів без спеціальних доведень.

Отже, перш за все нам потрібно встановити, що кільце многочленів над полем є евклідове кільце.

22.2. **Кільце многочленів як евклідове кільце.** Щоб довести, що кільце $P[x]$ многочленів над полем P є евклідовим, потрібно, згідно з означенням евклідового кільця (п. 14.3), показати, що

1. *$P[x]$ є область цілісності.*
2. *Існує відображення* $\varphi\colon P[x] \setminus \{0\} \to N^0 \underset{\text{df}}{=} \mathbf{N} \cup \{0\}$ (1)

таке, що має місце ділення з остачею, тобто

$$\underset{f(x),g(x)\in P[x]}{\forall}\ \underset{\substack{s(x),r(x)\in P[x]\\ g(x)\neq 0}}{\exists}\ [f(x) = g(x)\,s(x) + r(x)],$$

причому або $r(x) = 0$, *або* $\varphi(r(x)) < \varphi(g(x))$. Той факт, що $P[x]$ — область цілісності, було доведено в § 21 (п. 21.4). Відображення $\varphi: P[x] \setminus \{0\} \to N^0$ побудуємо в такий спосіб. Кожному многочлену $f(x) \in P[x]$, відмінному від нуля, поставимо у відповідність його степінь, тобто

$$\varphi(f(x)) \underset{\text{df}}{=} \deg f \in N^0.$$

Залишається довести властивість 3, тобто здійсненність ділення з остачею в кільці $\mathbf{P}[x]$ при тому означенні відображення φ, яке було щойно наведено. Іншими словами, слід показати, що *для довільних многочленів $f(x)$ і $g(x) \neq 0$ з кільця $P[x]$ існують многочлени $s(x)$ і $r(x)$ такі, що $f(x)$ можна подати у вигляді*

$$f(x) = g(x)\,s(x) + r(x), \tag{2}$$

де $r(x) = 0$ *або* $\deg r < \deg g$. Ураховуючи зауваження в кінці п. 21.2, це можна записати коротше $\deg r < \deg g$.

При діленні з остачею многочленів вживають ту саму термінологію, що й для цілих чисел: $f(x)$ — *ділене*, $g(x)$ — *дільник*, $s(x)$ — *частка*, $r(x)$ — *остача*. Умова для чисел, щоб остача була менша за модуль дільника, у випадку многочленів замінюється умовою, щоб степінь остачі був менший від степеня дільника.

Приклад. Нехай $f(x) = x^3 - 3x + 1$, $g(x) = x^2 + 1$. Рівність $x^3 - 3x + 1 = (x^2 + 1)x + (-4x + 1)$ свідчить про те, що $f(x)$ ділиться на $g(x)$ з остачею, причому частка $s(x) = x$, остача $r(x) = -4x + 1$. Так само $g(x)$ ділиться з остачею на $f(x)$; у цьому випадку частка $s_1(x) = 0$, остача $r_1(x) = x^2 + 1$, адже $x^2 + 1 = (x^3 - 3x + 1) \cdot 0 + x^2 + 1$.

Теорема 1. *Довільний многочлен $f(x)$ з кільця $P[x]$ ділиться з остачею на будь-який многочлен $g(x)$ з цього кільця, відмінний від нульмногочлена; при цьому частка й остача також належать до $P[x]$ і визначаються однозначно.*

Доведення. Встановимо спочатку саму можливість знайти серед многочленів з кільця $P[x]$ частку $s(x)$ і остачу $r(x)$ для будь-яких многочленів $f(x)$, $g(x) \in P[x]$ при $g(x) \neq 0$. Нехай

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0.$$

Якщо $f(x) = 0$, то зрозуміло, що $s(x) = 0$, $r(x) = 0$. Нехай тепер $n = \deg f < \deg g = m$; тоді можна вважати, що $s(x) = 0$, $r(x) = f(x)$. Залишається розглянути випадок, коли $n \geqslant m$. Виконаємо доведення методом індукції по n, починаючи з $n = 0$. При $n = 0$ маємо $m = 0$, $f(x) = a_0$, $g(x) = b_0 \neq 0$, тому $s(x) = \frac{a_0}{b_0}$, а $r(x) = 0$. Очевидно, $s(x) \in P[x]$, бо $\frac{a_0}{b_0} \in P$.

Припустимо, що теорема справедлива для всіх многочленів таких, що $\deg f < n$, і доведемо її для многочленів степеня n.

Розглянемо многочлен

$$p(x) = f(x) - \frac{a_n}{b_m} x^{n-m} g(x)$$

(нагадаємо, що a_n і b_m відмінні від нуля). Старший член многочлена $\frac{a_n}{b_m} x^{n-m} g(x)$ (який, очевидно, належить $P[x]$) дорівнює $a_n x^n$, тобто старшому члену многочлена $f(x)$. Тому $\deg p < n$, і за припущенням індукції $p(x)$ можна поділити з остачею на $g(x)$:

$$p(x) = g(x)\, s_1(x) + r_1(x); \quad s_1(x),\ r_1(x) \in P[x];$$

$$r_1(x) = 0 \text{ або } \deg r_1 < \deg g.$$

Отже,

$$f(x) - \frac{a_n}{b_m} x^{n-m} g(x) = g(x)\, s_1(x) + r_1(x),$$

звідки $f(x) = g(x)\, s(x) + r(x)$, де

$$r(x) = r_1(x), \quad s(x) = s_1(x) + \frac{a_n}{b_m} x^{n-m}.$$

Зрозуміло, що $s(x)$ і $r(x)$ належать $P[x]$ і що $r(x) = 0$ або $\deg r = \deg r_1 < \deg g$. Цим показано можливість ділення $f(x)$ на $g(x)$ з остачею.

Щоб закінчити доведення теореми, слід встановити єдиність частки $s(x)$ і остачі $r(x)$. Припустимо, що можливі два записи:

$$f(x) = g(x)\, s(x) + r(x), \quad \deg r < \deg g;$$

$$f(x) = g(x)\, \hat{s}(x) + \hat{r}(x), \quad \deg \hat{r} < \deg g.$$

Віднімаючи другу рівність від першої, дістанемо

$$g(x)\,[s(x) - \hat{s}(x)] = r(x) - \hat{r}(x). \tag{3}$$

За умовою $g(x) \neq 0$. Якщо припустити, що $r(x) \neq \hat{r}(x)$, то й $s(x) \neq \hat{s}(x)$ (адже кільце $P[x]$ не має дільників нуля). Але тоді ми приходимо до суперечності. Справді, права частина (3) є многочлен, степінь якого менший від $\deg g$ і, отже, менший від степеня лівої частини рівності (2). Отже, рівність (3) можлива лише при $r(x) = \hat{r}(x)$ і $s(x) = \hat{s}(x)$, що й свідчить про єдиність частки і остачі. Теорему доведено.

Як наслідок з попереднього викладу дістаємо таке важливе твердження:

Теорема 2. *Кільце $P[x]$ многочленів над полем P є евклідове кільце.*

Зауважимо, що у кільці $R[x]$, де R — область цілісності, але не поле, ділення з остачею, взагалі кажучи, нездійсненне. Так, у кільці $\mathbf{Z}[x]$ для многочленів $f(x) = x^2 + 1$, $g(x) = 3x - 1$ не можна знайти многочленів $s(x)$ і $r(\deg r < 1)$ таких, щоб $x^2 + 1 = (3x - 1)\, s(x) + r$. В доведенні теореми 1 було істотно використано те, що многочлени розглядаються саме над полем (при доведенні того, що $s(x)$ і $r(x)$ належать $P[x]$). Однією з істотних причин того, що далі ми розглядатимемо кільця многочлена над полем, а не над довільними областями цілісності, є можливість здійснювати в таких кільцях ділення з остачею.

22.3. Техніка ділення з остачею. Схема Горнера. Практичне здійснення ділення з остачею для двох заданих многочленів (тобто знаходження частки та остачі) грунтується на методі, використаному при доведенні теореми 1. А саме, спочатку від діленого $f(x)$ віднімають $\frac{a_n}{b_m} x^{n-m} g(x)$ (де $g(x)$ — дільник, a_n і b_m — старші коефіцієнти $f(x)$ і $g(x)$ відповідно), а з многочленом-різницею діють так само, як з $f(x)$, а саме: якщо старший член цього многочлена є $c_l x^l$, $l \geqslant m$, то від нього віднімають $\frac{c_l}{b_m} x^{l-m}\, g(x)$ і т. д. Цей процес продовжують доти,

поки не дістануть многочлен, степінь якого менший від m. Такий момент обов'язково настане, бо степінь діленого щоразу зменшується щонайменше на одиницю. Знайдений таким способом многочлен і буде $r(x)$, а $s(x)$ буде сумою множників $\frac{a_n}{b_m}x^{n-m}$, $\frac{c_l}{b_m}x^{l-m}$, ... при $g(x)$, які будувалися в процесі цього віднімання.

Очевидно, що саме до цього способу зводиться метод ділення многочлена на многочлен, відомий з курсу елементарної математики.

Приклад 1. Нехай $f(x) = x^4 - 2x^3 + x - 1$, а $g(x) = x^2 - 2$.

Знайдемо $g_1(x) = f(x) - \frac{a_n}{b_m}x^{n-m}g(x)$. У цьому разі

$$g_1(x) = (x^4 - 2x^3 + x - 1) - x^2(x^2 - 2) = -2x^3 + 2x^2 + x - 1.$$

Далі, з $g_1(x)$ діємо так само, як і з $f(x)$:

$$g_2(x) = (-2x^3 + 2x^2 + x - 1) - (-2x)(x^2 - 2) = 2x^2 - 3x - 1.$$

Аналогічно

$$g_3(x) = 2x^2 - 3x - 1 - 2(x^2 - 2) = -3x + 3.$$

Як бачимо, степінь $g_3(x)$ вже менший від степеня $g(x)$. Отже, $r(x) = g_3(x) = -3x + 3$, $s(x) = x^2 + (-2x) + 2 = x^2 - 2x + 2$, і тому

$$x^4 - 2x^3 + x - 1 = (x^2 - 2)(x^2 - 2x + 2) + (-3x + 3).$$

Це саме ділення з остачею можна подати у звичайній формі:

$$\begin{array}{r|l} x^4 - 2x^3 + x - 1 & x^2 - 2 \\ \underline{-x^4 \quad\quad - 2x^2} & \\ -2x^3 + 2x^2 + x - 1 & x^2 - 2x + 2 \\ \underline{-2x^3 \quad\quad + 4x} & \\ 2x^2 - 3x - 1 & \\ \underline{-2x^2 \quad\quad - 4} & \\ -3x + 3 & \end{array}$$

Оскільки за теоремою 1 частка $s(x)$ і остача $r(x)$ визначаються однозначно, для знаходження їх можна користуватися і методом невизначених коефіцієнтів.

Пояснимо цей метод на тому самому прикладі. Нам відомо, що існують такі многочлени $s(x)$ і $r(x)$, для яких справджується рівність

$$x^4 - 2x^3 + x - 1 = (x^2 - 2)s(x) + r(x), \qquad (4)$$

причому степінь $s(x)$ не може перевищувати $n - m$, тобто в цьому разі двох, а степінь $r(x)$ менший від m, тобто в цьому разі не перевищує одиницю.

Це означає, що $s(x)$ і $r(x)$ можна подати в канонічній формі так:

$$s(x) = A_2x^2 + A_1x + A_0; \quad r(x) = B_1x + B_0,$$

де A_0, A_1, A_2, B_0, B_1 — поки що невідомі коефіцієнти.

Підставляючі ці вирази у рівність (4), маємо

$$x^4 - 2x^3 + x - 1 = (x^2 - 2)(A_2x^2 + A_1x + A_0) + (B_1x + B_0).$$

На підставі означення рівності многочленів коефіцієнти при однакових степенях x рівні між собою. Звідси дістаємо систему рівнянь:

$$A_2 = 1, \quad A_1 = -2, \quad -2A_2 + A_0 = 0, \quad -2A_1 + B_1 = 1, \quad -2A_0 + B_0 = -1.$$

Розв'язавши цю систему, матимемо

$$A_2 = 1, \quad A_1 = -2, \quad A_0 = 2, \quad B_1 = -3, \quad B_0 = 3,$$

тобто $s(x) = x^2 - 2x + 2$, $r(x) = -3x + 3$, як і слід було чекати.

У загальному випадку $s(x)$ шукають у вигляді многочлена з невизначеними коефіцієнтами степеня $n - m$, а $r(x)$ — степеня $m - 1$.

Застосуємо такий перехід до окремого, але важливого для дальшого викладу випадку, коли $g(x) = x - \alpha$, тобто коли многочлен-дільник є лінійний двочлен. Використовуючи метод невизначених коефіцієнтів, візьмемо

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = (x - \alpha)(A_{n-1}x^{n-1} + A_{n-2}x^{n-2} + \cdots + A_1x + A_0) + r. \qquad (5)$$

Звернемо увагу на те, що остача в цьому разі є многочлен, не вище нульового степеня, тобто константа (елемент поля P).

Прирівнюючи коефіцієнти у рівності (5), дістанемо:

$$\begin{aligned} a_n &= A_{n-1}, \\ a_{n-1} &= A_{n-2} - \alpha A_{n-1}, \\ &\cdots\cdots\cdots \\ a_1 &= A_0 - \alpha A_1, \\ a_0 &= r - \alpha A_0, \end{aligned}$$

звідки

$$\begin{aligned} A_{n-1} &= a_n, \\ A_{n-2} &= a_{n-1} + \alpha A_{n-1}, \\ A_{n-3} &= a_{n-2} + \alpha A_{n-2}, \\ &\cdots\cdots\cdots \\ A_0 &= a_1 + \alpha A_1, \\ r &= a_0 + \alpha A_0. \end{aligned} \qquad (6)$$

Формули (6) показують, що поділити многочлен на $x - \alpha$ можна за такою схемою, яка називається схемою Горнера[1].

	a_n	a_{n-1}	a_{n-2}	a_{n-3}	...	a_1	a_0
α	a_n	$\alpha A_{n-1} + a_{n-1}$	$\alpha A_{n-2} + a_{n-2}$	$\alpha A_{n-3} + a_{n-3}$	...	$\alpha A_1 + a_1$	$\alpha A_0 + a_0$
	A_{n-1}	A_{n-2}	A_{n-3}	A_{n-4}		A_0	r

Виконуючи ділення за цією схемою, кожний наступний коефіцієнт A_{k-1} частки й остачу r дістають множенням щойно обчисленого кое-

[1] У. Горнер (1768—1837) — англійський математик.

фіцієнта A_k на α і додаванням до знайденого добутку відповідного коефіцієнта a_k даного многочлена.

Зауважимо, що ділення з остачею многочлена $f(x)$ на лінійний двочлен виду $x - \alpha$ здійснене і в кільці многочленів над будь-якою областю цілісності R (що не є полем), хоч, як було зазначено в попередньому пункті, для довільних многочленів $f(x) \in R[x]$, $g(x) \in R[x]$ частка $s(x)$ і остача $r(x)$ можуть в $R[x]$ не існувати. У випадку ж $g(x) = x - \alpha$ коефіцієнти частки й остача, як показують формули (6), належать області цілісності R.

Приклад 2. Поділимо за схемою Горнера многочлен $x^4 - 3x^2 + 2x - 1$ на $x - 2$. Тут $a_4 = 1$, $a_3 = 0$, $a_2 = -3$, $a_1 = 2$, $a_0 = -1$, $\alpha = 2$. Складемо схему:

	1	0	−3	2	−1
2	1	1·2 + 0	2·2 + (−3)	1·2 + 2	4·2 + (−1)
	1	2	1	4	7

Отже, частка дорівнює $s(x) = x^3 + 2x^2 + x + 4$, а остача $r = 7$.

Схему Горнера особливо доцільно використовувати тоді, коли знайдену частку треба знову ділити на який-небудь лінійний множник. При застосуванні схеми Горнера для послідовного ділення не потрібно щоразу наново виписувати коефіцієнти часток: закінчення одного з процесів ділення є в той же час початком наступного.

Приклад 3.. Якщо потрібно не тільки многочлен $x^4 - 3x^2 + 2x - 1$ поділити на $x - 2$, але й знайдену частку $s(x) = x^3 + 2x^2 + x + 4$ знову поділити на $x - 2$, то обчислювати за схемою Горнера можна так:

	1	0	−3	2	−1
2	1	2	1	4	7
2	1	4	9	22	—

Другий рядок цієї таблиці показує, що при діленні $x^4 - 3x^2 + 2x - 1$ на $x - 2$ ми дістаємо частку $x^3 + 2x^2 + x + 4$ і остачу 7. Третій рядок показує, що при діленні $x^3 + 2x^2 + x + 4$ на $x - 2$ ми дістаємо частку $x^2 + 4x + 9$ і остачу 22.

Легко перевірити, що в наведених прикладах остача при діленні $f(x)$ на $x - \alpha$ дорівнює значенню многочлена при $x = \alpha$, тобто $f(\alpha)$. Цей результат у загальному випадку вже відомий читачеві з курсу алгебри середньої школи під назвою теореми Безу. Нагадаємо тут формулювання і доведення цієї теореми.

Теорема 2 (Безу). Для будь-якого елемента α з поля P остача при діленні многочлена $f(x) \in P[x]$ на $x - \alpha$ дорівнює $f(\alpha)$.

Доведення. За формулою (2) ділення з остачею маємо:

$$f(x) = (x - \alpha)\, s(x) + r, \tag{7}$$

де многочлен r є константа, бо має степінь, нижчий від степеня $x - \alpha$. Оскільки многочлени, що стоять у лівій і правій частинах (7), рівні між собою, то рівні між собою і їх значення при будь-якому $x \in P$. Тому, взявши $x = \alpha$, дістаємо $f(\alpha) = r$, що й треба було довести. Зауважимо, що ми не вимагали, щоб поле P мало характеристику 0, і не твердили, що алгебраїчне і функціональне означення многочлена у нашому випадку рівносильні. Але рівність двох многочленів $f(x)$, $g(x) \in P[x]$ імплікує $\forall_{\beta \in P} [f(\beta) = g(\beta)]$ для довільного поля P.

Як було зазначено, за схемою Горнера зручно багаторазово ділити многочлен на лінійний двочлен. За допомогою такого ділення легко дістати розклад довільного многочлена $f(x)$ за степенями $x - \alpha$, який широко використовується в алгебрі та аналізі.

Справді, нехай $f(x)$ — многочлен n-го степеня над полем P, α — елемент цього поля. Ділення на $x - \alpha$ дає:

$$f(x) = (x - \alpha)\, f_1(x) + c_0, \tag{8}$$

де $f_1(x)$ — многочлен $(n-1)$-го степеня з $P[x]$, а c_0 — елемент поля P. Якщо $n > 1$, аналогічно маємо:

$$\begin{aligned} f_1(x) &= (x - \alpha)\, f_2(x) + c_1, \\ f_2(x) &= (x - \alpha)\, f_3(x) + c_2, \\ &\cdots\cdots\cdots\cdots \\ f_{n-1}(x) &= (x - \alpha)\, f_n(x) + c_{n-1}. \end{aligned} \tag{9}$$

Очевидно, $f_n(x)$ є многочленом нульового степеня; візьмемо $f_n(x) = c_n$. Виключаючи з (9) і (8) $f_{n-1}(x)$, $f_{n-2}(x)$, ..., $f_2(x)$, $f_1(x)$, дістаємо

$$f(x) = c_n (x - \alpha)^n + c_{n-1} (x - \alpha)^{n-1} + \cdots + c_1 (x - \alpha) + c_0. \tag{10}$$

Отже, многочлен $f(x)$ над полем P ми подали як многочлен того самого степеня над тим самим полем, але вже від змінного $y = x - \alpha$. При цьому коефіцієнти $c_0, c_1, \ldots, c_n$ однозначно визначаються через α і коефіцієнти $a_0, a_1, \ldots, a_n$ многочлена $f(x)$. А саме, c_0 є остача від ділення $f(x)$ на $x - \alpha$; c_1 — остача від ділення першої частки $f_1(x)$ на $x - \alpha$ і т. д. Що ж до c_n, то це є остання частка в процесі послідовного ділення.

Приклад 4. Знайдемо розклад многочлена $f(x) = x^5 - 3x^3 + x^2 - 2x + 1$ за степенями двочлена $x - 1$.

Складемо таку таблицю:

	1	0	−3	1	−2	1
1	1	1	−2	−1	−3	[−2]
1	1	2	0	−1	[−4]	
1	1	3	3	[2]		
1	1	4	[7]			
1	1	[5]				
1	[1]					

У першому рядку цієї таблиці стоять коефіцієнти даного многочлена $f(x)$; у другому рядку — коефіцієнти частки $f_1(x)$ і остача $c_0 = -2$; у третьому рядку маємо вже результат ділення $f_1(x)$ на $x - 1$ і остачу $c_1 = -4$, що утворилась при цьому діленні, і т. д.

Отже, обведені у таблиці числа є шукані коефіцієнти $c_5 = 1$, $c_4 = 5$, $c_3 = 7$, $c_2 = 2$, $c_1 = -4$, $c_0 = -2$, а розклад многочлена $f(x)$ за степенями $x - 1$ матиме вигляд:

$$f(x) = (x-1)^5 + 5(x-1)^4 + 7(x-1)^3 + 2(x-1)^2 - 4(x-1) - 2.$$

22.4. Подільність многочленів. Ідеали кільця $P[x]$. Ми пересвідчились у тому, що будь-який многочлен $f(x) \in P[x]$ ділиться на довільний ненульовий многочлен $g(x) \in P[x]$ з остачею. Проте нас особливо цікавитиме той окремий випадок, коли це ділення відбувається «без остачі», або, як кажуть, «націло». У цьому випадку говорять просто, що $f(x)$ д і л и т ь с я на $g(x)$.

Вважатимемо, що многочлен $f(x) \in P[x]$ ділиться на многочлен $g(x) \in P[x]$, і записуватимемо $f(x) \vdots g(x)$, якщо остача $r(x)$ при діленні $f(x)$ на $g(x)$ дорівнює нулю, тобто якщо існує многочлен $s(x) \in P[x]$ такий, що

$$f(x) = g(x)\,s(x). \tag{11}$$

Якщо $f(x)$ ділиться на $g(x)$, то кажуть також, що $g(x)$ ділить $f(x)$ або є дільником $f(x)$, і записують $g(x) \mid f(x)$.

На підставі теореми 1 частка $s(x)$ визначається однозначно. Зауважимо, що нуль-многочлен ділиться на довільний многочлен, відмінний від нуля; при цьому частка також є нуль-многочлен. Це дає змогу далі в цьому пункті вважати всі розглядувані многочлени відмінними від нуля.

Оскільки $P[x]$ не є поле, то многочлени з $P[x]$, взагалі кажучи, не діляться один на одного. Але це не виключає того, що в окремих випадках таке ділення «націло» можливе (як і в кільці цілих чисел). Ці випадки становлять значний інтерес у зв'язку з тим, що подання многочлена як добутку кількох многочленів часто дає змогу спростити дослідження цього многочлена та розв'язування асоційованих з ним задач. Зокрема, ця обставина використовується при розв'язуванні алгебраїчних рівнянь.

Властивості подільності многочленів у кільці $P[x]$ дістаємо безпосередньо з властивостей подільності в довільній області цілісності (п. 14.1). При цьому слід урахувати, що дільниками одиниці в кільці $P[x]$ є всі відмінні від нуля константи і тільки вони (п. 22.1).

1. $\forall\limits_{f(x),g(x),h(x)\in P[x]} [f(x) \vdots g(x) \wedge g(x) \vdots h(x) \Rightarrow f(x) \vdots h(x)]$.

2. $\forall\limits_{h(x),f(x),g(x)\in P[x]} [f(x) \vdots h(x) \wedge g(x) \vdots h(x) \Rightarrow (f(x) + g(x)) \vdots h(x) \wedge$

$\wedge (f(x) - g(x)) \vdots h(x)]$.

3. $\forall\limits_{f(x),h(x)\in P[x]} [f(x) \vdots h(x) \Rightarrow \forall\limits_{g(x)\in P[x]} f(x)\,g(x) \vdots h(x)]$.

4. $\forall\limits_{h(x),f_1(x),\ldots,f_m(x)\in P[x]} [f_1(x) \vdots h(x) \wedge f_2(x) \vdots h(x) \wedge \cdots \wedge f_m(x) \vdots h(x) \Rightarrow$

$\Rightarrow \forall\limits_{g_1(x),g_2(x),\ldots,g_m(x)\in P[x]} [f_1(x)\,g_1(x) + \cdots + f_m(x)\,g_m(x) \vdots h(x)]$.

5. $\forall\limits_{f(x),g(x)\in P[x]} \forall\limits_{\substack{c\in P\\ c\neq 0}} [f(x) \vdots c]$.

6. $\forall\limits_{f(x),g(x)\in P[x]} \forall\limits_{\substack{c\in P\\ c\neq 0}} [f(x) \vdots g(x) \Rightarrow f(x) \vdots cg(x)]$.

Як відомо (п. 14.1), два елементи області цілісності з одиницею називаються а с о ц і й о в а н и м и, якщо вони діляться один на одного, або, що те саме, відрізняються лише множником, що є дільником одиниці. Зокрема, у кільці $P[x]$ многочлени $f(x)$, $g(x)$ асоційовані, якщо вони відрізняються лише множником, який є відмінною від нуля константою:

$$f(x) = cg(x) \text{ або } g(x) = \frac{1}{c} f(x) = c' f(x).$$

Легко бачити, що відношення між многочленами «*бути асоційованими*» рефлексивне, симетричне і транзитивне, тобто є відношенням е к в і в а л е н т н о с т і.

Якщо не вважати різними асоційовані многочлени (тобто розглядати класи асоційованих многочленів), то відношення подільності $f(x) \vdots g(x)$ можна тлумачити як відношення порядку (нестрогого), адже це відношення рефлексивне ($f(x) \vdots cf(x)$), транзитивне (властивість 1), антисиметричне, бо

$$f(x) \vdots c_1 g(x) \wedge g(x) \vdots c_2 f(x) \Rightarrow f(x) = cg(x),$$

тобто *якщо $f(x)$ і $g(x)$ або асоційовані з ними многочлени взаємно діляться один на одного, то $f(x)$ і $g(x)$ збігаються з точністю до асоційованості*.

Розглянемо тепер питання про будову і д е а л і в кільця $P[x]$.

Згідно з загальним означенням ідеалу комутативного кільця (п. 13.1), непорожня сукупність I многочленів $f(x) \in P[x]$ є ідеалом кільця $P[x]$, якщо вона є групою відносно додавання і якщо

$$\forall_{f(x)\in I} \forall_{g(x)\in P[x]} [f(x)\,g(x) \in I]. \tag{12}$$

Серед ідеалів кільця особливу роль відіграють г о л о в н і ідеали, кожний з яких породжений деяким елементом кільця. У кільці з одиницею головний ідеал, породжений елементом f, складається з усіх елементів кільця, кратних f, і позначається (f). Якщо в області цілісності з одиницею усі ідеали — головні, то її називають *кільцем головних ідеалів* (див. п. 14.2). Такі кільця мають ряд спільних властивостей, на які спирається теорія подільності.

Покажемо, що $P[x]$ є кільце головних ідеалів. Зауважимо, що це, твердження є наслідком теореми 2 і наведеного у п. 14.3 твердження, що *кожне евклідове кільце є кільце головних ідеалів*. Проте враховуючи особливе значення висловленої теореми для теорії подільності многочленів, наведемо її доведення, яке спирається безпосередньо на теорему 1.

Теорема 4. *Кільце $P[x]$ многочленів над довільним полем P є кільце головних ідеалів.*

Доведення. Зауважимо спочатку, що будь-який головний ідеал кільця $P[x]$ має вигляд $(f) = \{f(x)\, g(x)\}$, де $f(x)$ — фіксований, а $g(x)$ — довільний многочлен з $P[x]$ (оскільки $P[x]$ — кільце з одиницею). Іншими словами, (f) (ідеал, породжений елементом $f(x)$) складається з усіх многочленів кільця $P(x)$, які діляться на $f(x)$. Нам потрібно довести, що для довільного ідеалу I кільця $P[x]$ знайдеться многочлен $f(x)$ такий, що $I = (f)$.

У випадку, коли $I = \{0\}$, тобто складається лише з нуль-многочлена, твердження теореми правильне, бо тоді $I = (0)$ — головний ідеал, породжений нуль-многочленом. Нехай тепер $I \neq \{0\}$, тобто в I існують многочлени, відмінні від нуля. Позначимо через $f(x)$ якийсь многочлен найменшого степеня з I. Такий многочлен існує, бо степені многочленів — невід'ємні цілі числа, отже, в будь-якій сукупності многочленів є многочлени найменшого степеня. Оскільки $f(x) \in I$ і I — ідеал, то для довільного $g(x) \in P[x]$ маємо $f(x)\, g(x) \in I$, тобто $(f) \subset I$.

Залишається показати, що $(f) = I$, тобто що будь-який многочлен $q(x) \in I$ можна подати у вигляді $q(x) = f(x)\, s(x)$, де $s(x) \in P[x]$. Поділимо $q(x)$ на $f(x)$ з остачею; маємо $q(x) = f(x)\, s(x) + r(x)$, де $\deg r < \deg f$ або

$$r(x) = 0. \quad (13)$$

Через те що $q(x) \in I$ і $f(x)\, s(x) \in I$, то й $r(x) = q(x) - f(x)\, s(x) \in I$. Але тоді зрозуміло, що $r(x) = 0$, бо серед відмінних від нуля многочленів з I немає многочленів, степінь яких менший від $\deg f$.

Отже, $r(x) = 0$ і з (13) випливає, що $q(x) = f(x)\, s(x)$.

22.5. Найбільший спільний дільник. Алгоритм Евкліда. Конкретизуючи у випадку кільця многочленів $P[x]$ над полем P загальні поняття спільного дільника та найбільшого спільного дільника (НСД) многочленів, введемо такі означення.

Означення 1. *Якщо многочлен $d(x)$ є дільником многочлена $f(x)$ і многочлена $g(x)$, то він називається спільним дільником многочленів $f(x)$ і $g(x)$.*

Означення 2. *Спільний дільник многочленів $f(x)$ і $g(x)$, який ділиться на кожний інший спільний дільник цих многочленів, називається найбільшим спільним дільником многочленів $f(x)$ і $g(x)$ і позначається через (f, g).*

Ці означення природно узагальнюються на випадок m ($m > 2$) многочленів.

Звичайно, будь-які два многочлени мають тривіальні спільні дільники, а саме — дільники одиниці кільця $P[x]$; такими дільниками є усі відмінні від нуля константи (елементи поля P) (див. властивість 5 подільності, п. 22.4). Очевидно також, що НСД двох многочленів не визначається цілком однозначно. Якщо $d(x)$ — найбільший спільний дільник, то й кожний многочлен $cd(x)$, де c — елемент поля P, відмінний від нуля, також є НСД цих многочленів. Проте *з точністю до сталого множника найбільший спільний дільник визначається однозначно.* Справді, якщо $d(x)$ і $d_1(x)$ — найбільші спільні дільники даних двох

многочленів, то $d(x)$ повинен ділитися на $d_1(x)$ (бо $d(x)$ — найбільший спільний дільник), а $d_1(x)$ повинен ділитися на $d(x)$ (бо $d_1(x)$ — найбільший спільний дільник). Тоді $d(x)$ і $d_1(x)$ асоційовані, тобто $d_1(x) = cd(x)$, де c — константа, відмінна від нуля. Розглядаючи спільні дільники двох многочленів, ми не братимемо до уваги тривіальні дільники і вважатимемо, що многочлени взаємно прості, якщо в них немає спільних дільників, відмінних від многочленів нульового степеня. Точніше, приймемо таке означення:

Означення 3. *Многочлени $f(x)$, $g(x) \in P[x]$ називаються взаємно простими, якщо кожний їхній спільний дільник є многочленом нульового степеня* (відмінною від нуля константою).

Зрозуміло, що многочлени $f(x)$, $g(x)$ взаємно прості тоді і тільки тоді, коли $(f, g) = 1$ (адже ця умова рівнозначна тому, що кожний спільний дільник многочленів є дільником одиниці).

Основне твердження відносно НСД многочленів, яким ми постійно користуватимемось, можна сформулювати так:

Теорема 5. *Для будь-яких двох многочленів $f(x)$, $g(x) \in P[x]$ існує найбільший спільний дільник $d(x)$, причому $d(x)$ можна подати у вигляді*

$$d(x) = f(x)\, u(x) + g(x)\, v(x), \quad (14)$$

де $u(x)$, $v(x)$ — деякі многочлени з $P[x]$.

Ця теорема не потребує доведення, оскільки вона безпосередньо випливає з теореми 1 § 14, справедливої для будь-якого кільця головних ідеалів.

Зауважимо, що многочлен $d(x)$, поданий у формі (14), ділиться на підставі властивості 4 п. 22.4 на будь-який спільний дільник многочленів $f(x)$ і $g(x)$.

Наслідок. *Многочлени $f(x)$, $g(x) \in P[x]$ взаємно прості тоді і тільки тоді, коли існують многочлени $u(x)$, $v(x) \in P[x]$ такі, що*

$$f(x)\, u(x) + g(x)\, v(x) = 1. \quad (15)$$

З цього наслідку випливає ряд простих, але важливих властивостей взаємно простих многочленів. Сформулюємо деякі з них:

1. $\forall_{f(x),g(x),h(x)\in P[x]} [(f, g) = 1 \wedge (f, h) = 1 \Rightarrow (f, gh) = 1]$.
2. $\forall_{f(x),g(x),h(x)\in P[x]} [f(x)\, g(x) \vdots h(x) \wedge (f, h) = 1 \Rightarrow g(x) \vdots h(x)]$.
3. $\forall_{f(x),g(x),h(x)\in P[x]} [f(x) \vdots g(x) \wedge f(x) \vdots h(x) \wedge (g, h) = 1 \Rightarrow f(x) \vdots g(x) \times h(x)]$.

Ці властивості є конкретизацією для випадку кільця многочленів загальних теорем 3, 4 і 5 § 14.

Розглянемо тепер спосіб знаходження найбільшого спільного дільника двох многочленів.

Оскільки $P[x]$ є евклідове кільце, у ньому застосовна процедура знаходження НСД за допомогою послідовного ділення з остачею або алгоритму Евкліда. Цей алгоритм у загальному вигляді було описано й обґрунтовано в п. 14.3. Тут ми розглянемо алгоритм Евкліда стосовно до знаходження НСД двох многочленів.

Нехай дано два многочлени $f(x)$ і $g(x)$, причому степінь $f(x)$ не менший від степеня $g(x)$. Виконаємо послідовне ділення з остачею, яке можна записати за допомогою такої системи рівностей:

$$\begin{aligned}
&f(x) = g(x)\,s_1(x) + r_1(x),\\
&g(x) = r_1(x)\,s_2(x) + r_2(x),\\
&r_1(x) = r_2(x)\,s_3(x) + r_3(x),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&r_{n-2}(x) = r_{n-1}(x)\,s_n(x) + r_n(x),\\
&r_{n-1}(x) = r_n(x)\,s_{n+1}(x).
\end{aligned} \tag{16}$$

Ми тут виходимо з того, що після скінченного числа ділень остача $r_{n+1}(x)$ дорівнюватиме 0. Справді, з самого означення остачі зрозуміло, що степінь многочлена $r_1(x)$ менший від степеня $g(x)$; степінь $g_2(x)$ менший від степеня $g_1(x)$ і взагалі степінь $r_k(x)$ менший від степеня $r_{k-1}(x)$. Але це означає, що або якась з остач $r_k(x)$ дорівнюватиме нулю, або степінь остачі, зменшуючись при кожному діленні принаймні на одиницю, дорівнюватиме нулю. Якщо $\deg r_n = 0$, то $r_{n+1} = 0$, бо будь-який многочлен ділиться на многочлен нульового степеня. У всякому разі алгоритм Евкліда для многочленів зводиться до скінченного числа ділень з остачею. Оскільки степінь $r_1(x)$ не більший за $m-1$, де m — степінь $g(x)$, то число кроків у схемі (16) не може перевищувати m.

Приклад 1. Нехай $f(x) = x^3 - 3x^2 + 3x - 1$, $g(x) = x^3 - 1$. Застосовуючи алгоритм Евкліда до цих многочленів, дістанемо такі рівності:

$$x^3 - 3x^2 + 3x - 1 = (x^3 - 1)\cdot 1 + (-3x^2 + 3x);$$

$$x^3 - 1 = (-3x^2 + 3x)\left(-\frac{1}{3}x - \frac{1}{3}\right) + (x - 1);$$

$$-3x^2 + 3x = (x-1)(-3x).$$

$$\begin{aligned}
&s_1(x) = 1,\\
&r_1(x) = -3x^2 + 3x;\\
&s_2(x) = -\frac{1}{3}x - \frac{1}{3};\\
&r_2(x) = x - 1;\\
&s_3(x) = -3;\\
&r_3(x) = 0.
\end{aligned}$$

Відповідно до загальної теорії (п. 14.3) остання відмінна від нуля остача $r_n(x)$ у системі рівностей (16) і є НСД многочленів $f(x)$ і $g(x)$.

Приклади. 2. Для многочленів $f(x) = x^3 - 3x^2 + 3x - 1$, $g(x) = x^3 - 1$, які було розглянуто у прикладі 1, найбільший спільний дільник $(f, g) = x - 1$.

3. Знайдемо найбільший спільний дільник многочленів $f(x) = x^4 + x^3 + x^2 + 1$, $g(x) = 4x^3 + 3x^2 + 2x + 1$. На цей раз виконаємо обчислення докладно.

Ділимо $x^4 + x^3 + x^2 + x + 1$ на $4x^3 + 3x^2 + 2x + 1$.

При цьому, щоб уникнути дробових коефіцієнтів, перший з цих многочленів множимо на 4. Зрозуміло, що при цьому частка й остача також помножаються на 4, що не має істотного значення, бо всі многочлени ми визначаємо з точністю до сталого множника. Маємо:

$$\begin{array}{rl|l}
 & x^4 + x^3 + x^2 + x + 1 & 4x^3 + 3x^2 + 2x + 1\\
\text{(помножаємо на 4)} & 4x^4 + 4x^3 + 4x^2 + 4x + 4 & x\\
\hline
 & 4x^4 + 3x^3 + 2x^2 + x & \\
\hline
 & x^3 + 2x^2 + 3x + 4 &
\end{array}$$

Перш ніж ділити далі, помножимо знайдену різницю знову на 4. При цьому частку дістанемо неправильну, бо її перший коефіцієнт у 4 рази, а другий — у 16 раз більший за той, який повинен бути. Що ж до остачі, то вона збільшиться в 16 раз. Оскільки нас цікавить не частка, а остача і оскільки цю остачу можна визначити з точністю до сталого множника, то такий процес «порушеного ділення» веде нас до мети.

Отже, маємо далі:

$$\begin{array}{rl|l}
 & x^3 + 2x^2 + 3x + 4 & 4x^3 + 3x^2 + 2x + 1\\
\text{(помножаємо на 4)} & 4x^3 + 8x^2 + 12x + 16 & 1\\
\hline
 & 4x^3 + 3x^2 + 2x + 1 & \\
\hline
\text{(ділимо на 5)} & 5x^2 + 10x + 15 & \\
 & x^2 + 2x + 3 &
\end{array}$$

Таким чином, $r(x) = x^2 + 2x + 3$. Ділимо далі:

$$\begin{array}{rl|l}
 & 4x^3 + 3x^2 + 2x + 1 & x^2 + 2x + 3\\
\hline
 & 4x^3 + 8x^2 + 12x & 4x - 5\\
\hline
 & -5x^2 - 10x + 1 & \\
 & - & \\
\hline
 & -5x^2 - 10x - 15 & \\
\hline
\text{(ділимо на 16)} & 16 & \\
 & 1 &
\end{array}$$

Отже, $r_2(x) = 1$. Далі ділити не потрібно, бо видно, що $r_3(x) = 0$ і тому $(f, g) = r_2(x) = 1$. Найбільший спільний дільник дорівнює 1, тобто многочлени $x^4 + x^3 + x^2 + x + 1$ і $4x^3 + 3x^2 + 2x + 1$ взаємно прості.

Іноді доводиться шукати НСД не двох, а більшого числа многочленів, скажімо, $f_1(x), f_2(x), \ldots, f_n(x)$. У цьому випадку спочатку знаходимо $d_1(x) = (f_1, f_2)$; потім шукаємо $d_2(x) = (d_1, f_3)$; $d_3(x) = (d_2, f_4)$; …; $d_{k-1}(x) = (d_{k-2}, f_k)$; $d_k(x) = (d_{k-1}, f_{k+1})$; …; $d_{n-1}(x) = (d_{n-2}, f_n)$.

$d_{n-1}(x)$ і буде НСД усіх многочленів. Справді, $f_k(x)$ ділиться на $d_{n-1}(x)$, бо $f_k(x)$ ділиться на $d_{k-1}(x)$; $d_{k-1}(x)$ ділиться на $d_k(x)$; $d_k(x)$ — на $d_{k+1}(x)$ і т. д., нарешті, $d_{n-2}(x)$ ділиться на $d_{n-1}(x)$. Якщо тепер $d(x)$ — будь-який спільний дільник для $f_1(x), f_2(x), \ldots, f_n(x)$, то він є також дільником для $d_1(x), d_2(x), \ldots, d_{n-1}(x)$, як це випливає з означень цих многочленів. Отже, $d_{n-1}(x)$ — найбільший спільний дільник многочленів $f_1(x), f_2(x), \ldots, f_n(x)$.

Зрозуміло, що коли які-небудь два з многочленів $f_1(x), f_2(x), \ldots, f_n(x)$ взаємно прості, то НСД усіх многочленів дорівнює одиниці.

За допомогою алгоритму Евкліда можна також знайти для заданих многочленів $f(x), g(x) \in P$ многочлени $u(x)$ і $v(x)$ з $P[x]$, які задовольняють рівність (14).

Оскільки $d(x) = r_n(x)$, то з рівностей (16) дістаємо (для спрощення записів позначатимемо многочлени скорочено, опускаючи букву x):

$$d = r_{n-2} - r_{n-1}s_n. \tag{17}$$

У свою чергу,

$$r_{n-1} = r_{n-3} - r_{n-2}s_{n-1},$$
$$r_{n-2} = r_{n-4} - r_{n-3}s_{n-2} \qquad (18)$$

і т. д. В загальному випадку

$$r_{n-k} = r_{n-k-2} - r_{n-k-1}s_{n-k}, \qquad (19)$$

де $k = 1, 2, \ldots, n-1$, причому під r_{-1} слід розуміти многочлен f, а під r_0 — многочлен g.

Підставимо тепер у (17) вираз для r_{n-1}. Дістанемо:

$$d = r_{n-2} - [r_{n-3} - r_{n-2}s_{n-1}]\, s_n = -r_{n-3}s_n + r_{n-2}(1 + s_n s_{n-1}).$$

Отже,

$$d = -r_{n-3}s_n + r_{n-2}(1 + s_n s_{n-1}). \qquad (20)$$

З цього виразу можна аналогічно виключити r_{n-2}, підставивши замість нього вираз з рівності (18):

$$d = -r_{n-3}s_n + [r_{n-4} - r_{n-3}s_{n-2}](1 + s_n s_{n-1}) =$$
$$= r_{n-4}(1 + s_n s_{n-1}) - r_{n-3}[s_n + s_{n-2}(1 + s_n s_{n-1})]. \qquad (21)$$

Виключаючи далі r_{n-3}, r_{n-4}, ..., ми щоразу діставатимемо вираз d через r_{k-1} і r_k, в якому множники при них утворені з часток s_i за допомогою додавання і множення. Процес поступового виключення r_k припиниться тоді, коли в правій частині рівності з'являться $r_{-1} = f$ і $r_0 = g$.

Отже, ми дістанемо, що

$$d = fu + gv, \qquad (22)$$

де u і v — деякі многочлени, утворені з часток $s_n, s_{n-1}, \ldots, s_1$. З самого способу побудови $u(x)$ і $v(x)$ бачимо, що ці многочлени належать до того самого кільця $P[x]$, що й $f(x)$, $g(x)$.

У кільці $P[x]$ многочленів над полем P можна означити і *найменше спільне кратне* елементів.

Означення. *Спільним кратним многочленів* $f(x), g(x) \in P[x]$ *називається будь-який многочлен* $s(x) \in P[x]$ *такий, що* $s(x) \vdots f(x) \wedge \wedge s(x) \vdots g(x)$. *Найменшим спільним кратним* (НСК) *многочленів* $f(x)$, $g(x)$ *називається спільне кратне* $f(x)$ *і* $g(x)$, *яке ділить будь-яке інше спільне кратне цих многочленів;* НСК *многочленів* $f(x)$ *і* $g(x)$ *позначають* $[f, g]$.

Теорема 6. *Для будь-яких відмінних від нуля многочленів* $f(x)$, $g(x)$ *найменше спільне кратне існує і визначається однозначно з точністю до сталого множника.*

Доведення. Розглянемо многочлен $q(x) = \frac{f(x)\,g(x)}{(f, g)}$, де (f, g), як завжди, НСД многочленів $f(x)$ та $g(x)$ і тому ділить кожний з них. Зрозуміло, що $q(x)$ є спільне кратне $f(x)$ і $g(x)$, бо

$$q(x) = \frac{f(x)}{(f, g)} \cdot g(x) = \frac{g(x)}{(f, g)} \cdot f(x).$$

Якщо тепер $s(x)$ — будь-яке інше спільне кратне многочленів $f(x)$ і $g(x)$, то $s(x) \vdots f(x)$ і $s(x) \vdots g(x)$ і тому $s(x) = s_1(x) f(x)$, причому $\frac{s_1(x) f(x)}{g(x)} = p(x)$ — многочлен з $P[x]$.

Подамо тепер $f(x)$ і $g(x)$ у вигляді $f(x) = (f, g) f_1(x)$, $g(x) = (f, g) g_1(x)$, де $f_1(x)$ і $g_1(x)$ — многочлени з $P[x]$; при цьому [1] $(f_1, g_1) = 1$. Тоді

$$p(x) = \frac{s_1(x) f_1(x) (f, g)}{g_1(x) (f, g)} = \frac{s_1(x) f_1(x)}{g_1(x)}.$$

Оскільки $(f_1, g_1) = 1$, то $s_1(x) \vdots g_1(x)$. Вводячи означення $\frac{s_1(x)}{g_1(x)} = t(x) \in P[x]$, дістанемо: $s_1(x) = g_1(x)\, t(x)$, звідки

$$s(x) = s_1(x) f(x) = f(x) g_1(x) t(x) = \frac{f(x)\, g(x)}{(f, g)} t(x) = q(x) t(x),$$

тобто $s(x) \vdots q(x)$.

Отже, $q(x)$ справді є НСК многочленів $f(x)$, $g(x)$: $q(x) = [f, g]$. Якщо $q_1(x)$ — будь-яке інше НСК цих многочленів, то $q(x) \vdots q_1(x)$ і $q_1(x) \vdots q(x)$, звідки ясно, що $q(x)$ і $q_1(x)$ відрізняються лише сталим множником. Теорему доведено.

Можна, аналогічно до НСД, розглядати НСК довільного числа многочленів. Ми на цьому спинятися не будемо.

22.6. Незвідні многочлени. Розглянемо тепер, які з елементів області цілісності $P[x]$ є неперозкладними або простими. Згідно з загальним означенням (п. 14.1), елемент області цілісності нерозкладний або простий, якщо він *не є дільником одиниці і не має нетривіальних дільників*. Переформулюємо це означення стосовно кільця многочленів над полем P, увівши для простого многочлена спеціальний термін — *незвідний*.

Означення 1. *Многочлен* $f(x) \in P[x]$ *називається незвідним у полі* P, *якщо він не є константа і не має дільників, відмінних від константи і від многочленів виду* $cf(x)$, *де* c — *константа.*

Іншими словами, $f(x) \in P[x]$ — незвідний у полі P, якщо $\deg f \geqslant 1$ і якщо з рівності $f(x) = g(x) s(x)$, $g(x), s(x) \in P[x]$ випливає $\deg g = 0 \vee \deg s = 0$.

Складені елементи області цілісності $P[x]$ називатимемо *звідними многочленами у полі*.

Означення 2. *Многочлен* $f(x) \in P[x]$ *називається звідним у полі* P, *коли* $\deg f \geqslant 1$ *і коли існують такі многочлени* $g(x), s(x) \in P[x]$, *що* $f(x) = g(x) s(x)$, *причому* $\deg g \geqslant 1$ *і* $\deg s \geqslant 1$.

Останню умову можна записати й у формі

$$f(x) = g(x) s(x), \text{ причому } \deg g < \deg f \text{ і } \deg s < \deg f, \qquad (23)$$

бо $\deg f = \deg g + \deg s$.

Отже, будь-який многочлен, вищий від нульового степеня, є або звідним, або незвідним у даному полі.

[1] Те, що многочлени $f_1(x)$ і $g_1(x)$ взаємно прості, випливає з співвідношення $f(x)\, u(x) + g(x)\, v(x) = (f, g) \Rightarrow f_1(x)\, u(x) + g_1(x)\, v(x) = 1$ (див. теорему 5 п. 22.5 і її наслідок).

Підкреслимо, що звідність чи незвідність многочлена є поняття в і д н о с н е і залежить від поля P, над яким многочлен розглядається. Як відомо, будь-який многочлен $f(x) \in P[x]$ можна вважати також многочленом над полем Δ, де Δ — довільне розширення поля P. Якщо $f(x)$ звідний у полі P, то він звідний і в будь-якому розширенні цього поля. Але цілком можливо, що многочлен $f(x)$, незвідний у полі P, виявиться звідним у деякому розширенні Δ поля P.

П р и к л а д и. **1.** Многочлен x^2+1 незвідний у полі раціональних чисел і в полі дійсних чисел. Цей многочлен звідний у полі комплексних чисел. Для доведення незвідності у полі дійсних чисел припустимо супротивне, тобто що $x^2+1=(ax+b)(cx+d)$, де a, b, c, d — дійсні числа, $a \neq 0$, $c \neq 0$. Нехай $x=-\frac{b}{a}$; тоді $\left(-\frac{b}{a}\right)^2+1=0$, тобто $a^2+b^2=0$, що неможливо, бо $a \neq 0$.

2. Многочлен x^2-2 незвідний у полі раціональних чисел. У цьому можна переконатися, міркуючи аналогічно попередньому. Він звідний у полі чисел виду $a+b\sqrt{2}$ ($a, b \in \mathbf{Q}$), а також у всіх розширеннях цього поля (зокрема, у полях $\mathbf{R}$ і $\mathbf{C}$), бо справедлива рівність $x^2-2=(x-\sqrt{2})(x+\sqrt{2})$.

До многочленів нульового степеня поняття звідності і незвідності не застосовується. Вони відіграють у теорії подільності многочленів ту саму роль, яку числа ± 1 відіграють у теорії подільності цілих чисел.

Що ж до многочленів першого степеня, то справедлива така теорема.

Теорема 7. *Многочлен першого степеня над довільним полем P незвідний у цьому полі.*

Це твердження очевидне, коли врахувати, що степінь добутку многочленів дорівнює сумі степенів співмножників.

Зауважимо, що кожний многочлен $f(x)$ першого степеня з раціональними коефіцієнтами незвідний у будь-якому числовому полі P, бо це поле можна розглядати як розширення поля раціональних чисел, і тому $f(x) \in P[x]$.

Сформулюємо деякі властивості незвідних многочленів, які є конкретизацією для випадку кільця $P[x]$ загальних властивостей простих елементів у будь-якій області цілісності з одиницею.

1) *Якщо $p(x)$ — многочлен, незвідний у даному полі, то і многочлен $cp(x)$, де c — довільна відмінна від нуля константа, незвідний у цьому полі.*

2) *Якщо $p(x)$ — незвідний у даному полі многочлен, а $f(x)$ — довільний многочлен над цим полем, то або $f(x)$ ділиться на $p(x)$, або ці многочлени взаємно прості.*

3) *Якщо незвідний у даному полі многочлен $p(x)$ ділиться на інший незвідний у цьому полі многочлен $q(x)$, то ці многочлени збігаються з точністю до сталого множника.*

Перші дві з цих властивостей не потребують доведення, бо відтворюють для кільця $P[x]$ властивості 1—2 простих елементів, доведені в п. 14.1 в загальному випадку. Властивість 3 випливає з таких простих міркувань. За умовою $p(x)$ і $q(x)$ мають спільний дільник $q(x)$ ненульового степеня і тому не взаємно прості. Через те що $p(x)$ — незвідний многочлен, то $q(x)$, на підставі властивості 2, повинен

ділитися на нього. Отже, многочлени $p(x)$ і $q(x)$ діляться один на одного і тому є асоційованими (тобто відрізняються лише множником нульового степеня).

Ми бачили, що звідність чи незвідність многочлена істотно залежить від того, над яким полем цей многочлен розглядається. На відміну від цього, в з а є м н а п р о с т о т а двох многочленів і, більше того, їх НСД повністю визначається даними многочленами незалежно від того, до якого кільця ми їх відносимо. Наприклад, многочлени $f(x)=x^2+1$ і $g(x)=x^4-2$ взаємно прості і над полем раціональних, і над полем дійсних або й комплексних чисел. Інакше кажучи, найбільший спільний дільник в усіх цих випадках однаковий (а саме 1), хоч дільники зовсім різні у кільцях над різними полями. Це пояснюється тим, що найбільший спільний дільник визначається за допомогою раціональних дій над даними многочленами і, отже, його коефіцієнти залежать тільки від коефіцієнтів даних многочленів і належать до того самого поля. З цих самих міркувань зрозуміло, що подільність чи неподільність многочлена $f(x)$ на многочлен $g(x)$ також не залежить від того, над яким полем вони розглядаються.

22.7. Канонічний розклад многочлена. Фундаментальну роль у теорії подільності цілих чисел відіграє теорема про можливість і єдиність розкладу довільного цілого числа (відмінного від 0, 1 і —1) у добуток простих множників. Це твердження навіть називають *основною теоремою арифметики* (див. п. 7.3.) Виявляється, що для многочленів справедливе цілком аналогічне твердження про можливість і однозначність розкладу довільного многочлена над полем P у добуток незвідних у цьому полі многочленів.

Точніше, справедлива така важлива теорема.

Теорема 8. *Кожний многочлен $f(x)$ ненульового степеня над полем P можна подати у вигляді*

$$f(x)=p_1(x)\,p_2(x)\,\ldots\,p_l(x), \tag{24}$$

де всі $p_k(x)$ є незвідними многочленами у полі P. Зображення (24) єдине з точністю до сталих множників і до порядку нумерації многочленів $p_k(x)$.

Зображення (24) називають *розкладом многочлена $f(x)$ на незвідні множники* (або у добуток незвідних множників) *у полі P*.

Д о в е д е н н я. Оскільки $P[x]$ є кільце головних ідеалів, то у ньому справджується теорема 7 п. 14.2, згідно з якою будь-який елемент $f(x) \in P[x]$, який не є нулем і дільником одиниці (тобто $\deg f \geqslant 1$), розкладається у добуток простих множників (тобто незвідних многочленів) з $P[x]$. Єдиність цього розкладу з точністю до сталих множників і до нумерації співмножників випливає з теореми 8 п. 14.2, згідно з якою два розклади елемента кільця головних ідеалів у добуток простих множників можуть відрізнятися лише порядком співмножників і множниками, що є дільниками одиниці; у нашому випадку дільники одиниці — це відмінні від нуля константи (сталі). Теорему доведено.

Наслідок. *Довільний многочлен ненульового степеня над полем P можна подати у вигляді*

$$f(x)=[p_1(x)]^{k_1}[p_2(x)]^{k_2}\,\ldots\,[p_m(x)]^{k_m}, \tag{25}$$

де $p_1(x)$, $p_2(x)$, ..., $p_m(x)$ — *попарно різні* (неасоційовані) *многочлени, незвідні у полі* P. *Це зображення єдине з точністю до сталих множників* (і нумерації співмножників).

Зображення (25) називатимемо *канонічним розкладом многочлена* $f(x)$ *у полі* P.

Розклад (25) відразу випливає з зображення (24), коли врахувати те, що деякі з незвідних множників $p_1(x)$, $p_2(x)$, ..., $p_l(x)$ у формулі (24) можуть бути однакові.

Означення. *Якщо многочлен* $p_j(x)$ *входить у канонічний розклад* (25) *у степені з показником* k_j, *кажуть, що* $p_j(x)$ *є множником кратності* k_j *многочлена* $f(x)$. *Множники, кратність яких більша за одиницю, називаються кратними множниками многочлена.*

Іншими словами, *незвідний многочлен* $p_j(x)$ *є множником* k_j-*ї кратності многочлена* $f(x)$, *якщо* $f(x)$ *ділиться на* $[p_j(x)]^{k_j}$, *але не ділиться на* $[p_j(x)]^{k_j+1}$.

Це означення, очевидно, рівносильне попередньому. Справді, якщо множник $p_j(x)$ входить у розклад (25) у степені k_j, то $f(x)$ ділиться на $[p_j(x)]^{k_j}$, але не ділиться на $[p_j(x)]^{k_j+1}$, бо всі множники $p_i(x)$ $(i \neq j)$ і їх добуток взаємно прості з $p_j(x)$. Якщо ж $f(x)$ ділиться на $[p_j(x)]^{k_j}$, але не ділиться на $p_j[(x)]^{k_j+1}$, то $f(x) = [p_j(x)]^{k_j} f_1(x)$, де $f_1(x)$ не ділиться на $p_j(x)$; тому розклад $f(x)$ на незвідні множники у формі (25) містить $[p_j(x)]^{k_j}$.

Приклади. 1. Розкладемо $f(x) = x^3 + x^2 - 5x + 3$ на незвідні множники в полі раціональних чисел. Очевидно,

$$f(x) = x^3 + x^2 - 5x + 3 = (x-1)(x-1)(x+3) = (x-1)^2(x+3).$$

Отже, $f(x)$ має два різні незвідні множники: $p_1(x) = x - 1$ кратності 2 і $p_2(x) = x + 3$ кратності 1. Цей самий розклад матимемо і в полях дійсних та комплексних чисел, бо $p_1(x)$ і $p_2(x)$ є многочлени першого степеня і тому незвідні в усякому числовому полі.

2. Нехай $f(x) = x^4 - 4x^2 + 4$. Очевидно, $f(x) = (x^2 - 2)^2$. Це зображення і є розкладом на незвідні множники у полі раціональних чисел. У полі дійсних чисел маємо інший розклад на незвідні множники: $f(x) = (x - \sqrt{2})^2 (x + \sqrt{2})^2$. Обидва множники мають кратність 2.

Зауваження. Доведення єдиності (на відміну від доведення можливості) розкладу (24) істотно спиралося на той факт, що $P[x]$ — кільце головних ідеалів. Коли замість $P[x]$ розглянути якусь сукупність многочленів над P, що не є кільцем головних ідеалів, то ця властивість, взагалі кажучи, не матиме місця. Так, сукупність $\hat{P}[x]$ усіх многочленів над полем P, які не містять членів першого степеня, є область цілісності з одиницею, але не є кільцем головних ідеалів. Щоб переконатись, що не всі ідеали в $\hat{P}[x]$ головні, досить розглянути ідеал I, який складається з усіх многочленів з $\hat{P}[x]$ без вільного члена, тобто усі многочлени виду $a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2$. Такий ідеал не може бути породжений многочленом $s(x)$ при $\deg s = 0$ (бо $I \neq \hat{P}[x]$), при $\deg s = 2$ (бо $x^3 \in I$ не можна подати як $s(x) f(x)$, $f(x) \in \hat{P}[x]$) і при $\deg s > 2$ (бо $x^2 \in I$ не можна подати як $s(x) f(x)$). Отже, I не є головний ідеал. В кільці $\hat{P}[x]$ многочлен $f(x) = x^6$ має два істотно різні розклади на незвідні множники: $x^6 = x^2 \cdot x^2 \cdot x^2$ і $x^6 = x^3 \cdot x^3$ (незвідність многочленів x^2 і x^3 у кільці $\hat{P}[x]$ очевидна, бо в $\hat{P}[x]$ немає множників степеня 1).

Покажемо тепер, що до многочленів можна застосувати метод знаходження НСД, подібний до методу розкладання на прості множники в арифметиці.

При цьому будемо користуватись тим очевидним фактом, що *кожний спільний дільник двох многочленів* $f(x)$ *і* $g(x)$ *над полем* P *може мати тільки такі незвідні множники в цьому полі, які є незвідними множниками як многочлена* $f(x)$, *так і многочлена* $g(x)$.

Теорема 9. *Якщо многочлени* $f(x)$ *і* $g(x)$ *розкладені на незвідні множники у довільному полі* P, *то найбільший спільний дільник* (f, g) *дорівнює добутку всіх незвідних множників, які входять у розклад*[1] *як* $f(x)$, *так і* $g(x)$. *Якщо таких спільних незвідних множників немає, то* $(f, g) = 1$.

Доведення. Припустимо спочатку, що розклади $f(x)$ і $g(x)$ мають спільні незвідні множники, які дорівнюють $d_1(x)$, $d_2(x)$, ..., $d_r(x)$. Тоді розклади $f(x)$ і $g(x)$ на незвідні множники можна записати у вигляді:

$$f(x) = d_1(x)\, d_2(x) \ \ldots\ d_r(x) \cdot p_{r+1}(x) \ \ldots\ p_l(x),$$

$$g(x) = d_1(x)\, d_2(x) \ \ldots\ d_r(x) \cdot q_{r+1}(x) \ \ldots\ q_m(x).$$

Многочлен $d(x) = d_1(x)\, d_2(x) \ \ldots\ d_r(x)$, очевидно, є спільним дільником многочленів $f(x)$ і $g(x)$. Але це є і найбільший спільний дільник. Справді, якщо $d'(x)$ — довільний спільний дільник $f(x)$ і $g(x)$, то легко зрозуміти, що його розклад на незвідні у даному полі множники має вигляд:

$$d'(x) = d_{i1}(x)\, d_{i2}(x) \ \ldots\ d_{is}(x),$$

де $d_{i1}(x)$, ..., $d_{is}(x)$ — якісь з многочленів $d_1(x)$, $d_2(x)$, ..., $d_r(x)$. Отже, $d(x)$ ділиться на $d'(x)$ і тому є найбільшим спільним дільником.

Якщо спільних незвідних множників у розкладах $f(x)$ і $g(x)$ немає, то $f(x)$ і $g(x)$ взаємно прості. Справді, якби ці многочлени мали найбільший спільний дільник ненульового степеня $d(x)$, то на підставі основної теореми 8 вони мали б хоч один незвідний спільний дільник, що суперечить умові. Теорему доведено. Звичайно, її легко поширити на випадок більшого числа заданих многочленів.

Приклад 3. Нехай $f(x) = x^3 + x^2 - 5x + 3$, $g(x) = x^3 - 3x^2 + 3x - 1$. Розкладемо ці многочлени на незвідні множники у полі $\mathbf{Q}$:

$$f(x) = (x-1)^2(x+3), \quad g(x) = (x-1)^3.$$

Відповідно до теореми 9, $(f, g) = (x-1)^2$, тобто $(f, g) = x^2 - 2x + 1$.

§ 23. КОРЕНІ МНОГОЧЛЕНІВ

23.1 Поняття кореня многочлена. Кратні корені. Нехай $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ — многочлен над полем P, а Δ — будь-яке розширення поля P (зокрема, може бути $\Delta = P$). Як ві-

[1] Це слід розуміти так, що в розкладі найбільшого спільного дільника кожний спільний незвідний множник має кратність, яка є меншою з кратностей цього множника в розкладах $f(x)$ і $g(x)$.

домо (п. 21.5), для довільного $\alpha \in \Delta$ можна обчислити *значення многочлена* $f(x)$ *при* $x = \alpha$ (або *в точці* α), тобто елемент

$$f(\alpha) = a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 \in \Delta.$$

Нас особливо цікавитимуть такі елементи $\alpha \in \Delta$, для яких значення $f(\alpha)$ є нулем поля P (або, що те саме, поля Δ).

Означення 1. *Коренем многочлена* $f(x) \in P[x]$ *називається елемент* α *будь-якого розширення* Δ *поля* P *такий, що* $f(\alpha) = 0$.

Корінь многочлена $f(x)$ називають також нулем многочлена $f(x)$.

Користуючись функціональною термінологією, можна сказати, що корені многочлена $f(x)$ — це *прообрази нульового елемента* при відображенні $f\colon \Delta \to \Delta$.

Поняття кореня многочлена має велике теоретичне і практичне значення. Адже розв'язування алгебраїчних рівнянь вищих степенів, тобто рівнянь виду

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0,$$

полягає в знаходженні всіх коренів многочлена, який утворює ліву частину рівняння.

Теорема 1. *Елемент* $\alpha \in P$ *є коренем многочлена* $f(x) \in P[x]$ *тоді і тільки тоді, коли лінійний двочлен* $x - \alpha$ *є дільником многочлена* $f(x)$.

Доведення. За теоремою Безу (п. 22.3) остача від ділення $f(x)$ на $x - \alpha$ дорівнює $f(\alpha)$. Отже, $f(x) \vdots (x - \alpha)$ тоді і тільки тоді, коли $f(\alpha) = 0$, тобто коли α — корінь многочлена $f(x)$. Цим теорему доведено.

Теорема 1 дає необхідну і достатню умову того, що $\alpha \in P$ є коренем многочлена $f(x)$, або нове означення поняття кореня рівносильне означенню 1 (при $\Delta = P$).

Означення 2. *Елемент* $\alpha \in P$ *називається коренем многочлена* $f(x) \in P[x]$, *якщо* $f(x)$ *ділиться на* $x - \alpha$.

Ця форма означення має певні переваги над попередньою. Вона носить чисто алгебраїчний характер (бо спирається лише на поняття дільника многочлена), тоді як перша форма використовує поняття значення многочлена $f(x)$ при $x = \alpha$, яке має функціональну природу. Далі, що найбільш важливо, означення 2, на відміну від означення 1, може бути узагальнене на випадок так званих кратних коренів многочлена.

Означення 3. *Елемент* $\alpha \in P$ *називається* k*-кратним коренем* (або коренем k-ї кратності) *многочлена* $f(x) \in P[x]$, *якщо* $f(x)$ *ділиться на* $(x - \alpha)^k$, *але не ділиться на* $(x - \alpha)^{k+1}$.

Іншими словами, кратність кореня α многочлена $f(x)$ є найбільше з натуральних чисел m таких, що $(x - \alpha)^m$ є дільником $f(x)$ у кільці $P[x]$.

Корені кратності 1 називаються простими, корені кратності 2 і більше — кратними, причому двократні та трикратні корені іноді називають також *подвійними* та *потрійними* відповідно. Елементи поля P, які не є коренями $f(x) \in P[x]$, іноді називають нульовими кратними коренями $f(x)$; зрозуміло, що це не суперечить означенню 3.

Якщо $f(x)$ — нуль-многочлен, то будь-який елемент $\alpha \in P$ є його коренем, причому кратність цього кореня не можна визначити, бо нуль-многочлен ділиться на $(x - \alpha)^m$ при довільному натуральному m. Якщо ж $f(x) \neq 0$, то будь-який корінь $\alpha \in P$ має певну кратність $k \leqslant \deg f$: адже $f(x)$ ділиться на $x - \alpha$ і не ділиться на $(x - \alpha)^m$ при $m > \deg f$.

Очевидно, що $\alpha \in P$ є k-кратним коренем многочлена $f(x) \in P[x]$ тоді і тільки тоді, коли

$$f(x) = (x - \alpha)^k g(x), \tag{1}$$

де $g(x)$ — многочлен над полем P, для якого α не є коренем: адже для того, щоб $f(x)$ ділився на $(x - \alpha)^k$ і не ділився на $(x - \alpha)^{k+1}$, необхідно і достатньо, щоб $g(x) \in P[x]$ і не ділився на $x - \alpha$.

Зрозуміло, що $\deg g = \deg f - k$ при $k = \deg f$ многочлен $g(x)$ є константа.

Приклади. 1. Многочлен $f(x) = x^n$ над полем Q має n-кратний корінь $\alpha = 0$, бо $f(x)$ ділиться на $(x - 0)^n$ і не ділиться на $(x - 0)^{n+1}$.

2. Многочлен $f(x) = x^5 - 8x^4 + 25x^3 - 38x^2 + 28x - 8$ можна подати у вигляді добутку $f(x) = (x - 1)^2(x - 2)^3$. Звідси видно, що 1 є подвійним, а 2 — потрійним коренем цього многочлена.

23.2. Число коренів многочлена. Інтерполяційний многочлен. Нехай $f(x)$ — многочлен n-го степеня над полем P, а Δ — будь-яке розширення поля P. Припустимо, що $\alpha_1 \in \Delta$ є коренем $f(x)$ кратності k_1, $\alpha_2 \in \Delta$ — коренем $f(x)$ кратності k_2, ..., $\alpha_m \in \Delta$ — коренем $f(x)$ кратності k_m, причому $\alpha_i \neq \alpha_j$ при $i \neq j$. Тоді, згідно з (1), можна записати

$$f(x) = (x - \alpha_1)^{k_1} g_1(x), \tag{2}$$

де $g_1(x)$ не ділиться на $x - \alpha_1$. Оскільки $f(x)$ має ділитись на $(x - \alpha_2)^{k_2}$ (але не на $(x - \alpha_2)^{k_2+1}$), а $(x - \alpha_1)^{k_1}$ взаємно простий з $(x - \alpha_2)^{k_2}$, то з (2), видно, що $g_1(x)$ ділиться на $(x - \alpha_2)^{k_2}$ (але не на $(x - \alpha_2)^{k_2+1}$), тобто

$$f(x) = (x - \alpha_1)^{k_1}(x - \alpha_2)^{k_2} g_2(x),$$

де $g_2(x)$ — многочлен, який не має своїми коренями α_1 та α_2.

Продовжуючи міркувати в такий же спосіб (тобто, по суті, застосовуючи метод математичної індукції), дістанемо

$$f(x) = (x - \alpha_1)^{k_1}(x - \alpha_2)^{k_2} \ldots (x - \alpha_m)^{k_m} g_m(x), \tag{3}$$

де $g_m(x)$ — многочлен, для якого жодний з елементів $\alpha_1, \alpha_2, \ldots, \alpha_m$ не є коренем.

З (3) видно, що

$$\deg f = n = k_1 + k_2 + \cdots + k_m + \deg g_m,$$

тобто

$$k_1 + k_2 + \cdots + k_m \leqslant n. \qquad (4)$$

Отже, число коренів многочлена $f(x)$ у полі Δ не може перевищувати степеня цього многочлена, коли навіть кожний корінь ураховувати стільки разів, яка його кратність. Оскільки Δ — довільне розширення поля P, а будь-який корінь многочлена $f(x) \in P[x]$ лежить у якомусь розширенні поля P, то ми дістали такий результат:

Теорема 2. *Число усіх можливих коренів многочлена $f(x)$ над полем P не перевищує його степеня.*

Згідно з попереднім викладом, це твердження справедливе, якщо при підрахунку числа коренів кожний з них рахуємо стільки разів, яка його кратність. Надалі постійно дотримуватимемося цієї домовленості.

Зауважимо, що теорема 2 справедлива і для многочленів нульового степеня (тобто для відмінних від нуля констант, які не мають жодного кореня).

Наслідок. *Якщо многочлен $f(x) \in P[x]$, степінь якого не перевищує n, має $n+1$, різних коренів, то $f(x)$ є нуль-многочлен.*

Це твердження уточнює теорему 4 п. 21.5, згідно з якою многочлен $f(x)$ над областю цілісності R характеристики 0 є нуль-многочленом, якщо $\forall_{\alpha \in R}\,[f(\alpha) = 0]$. Кільце характеристики 0 має нескінченне число різних елементів (зокрема, усі кратні *та* ненульового елемента a попарно різні). Умова $\forall_{\alpha \in R}\,[f(\alpha) = 0]$ означає, що всі елементи $\alpha \in R$ є корені $f(x)$, тобто що $f(x)$ має безліч коренів. Тепер зрозуміло, що теорема 4 п. 21.5 безпосередньо випливає з щойно наведеного наслідку.

Отже, вимога $f(\alpha) = 0$ для всіх $\alpha \in R$ є надмірною для того, щоб було $f(x) = 0$; досить, щоб $f(\alpha_i) = 0$, $\alpha_i \in R$, $j = 1, 2, \ldots, (n+1)$, якщо тільки $\alpha_i \neq \alpha_j$ при $i \neq j$.

Можна дати інше формулювання наведеному наслідку: *два многочлени $f(x)$, $g(x) \in P[x]$, степені яких не перевищують n і які приймають однакові значення в $n+1$ різних точках з P, рівні між собою.* Для доведення цього твердження досить розглянути многочлен $f(x) - g(x)$.

Це означає, що серед многочленів не вище n-го степеня існує *не більше одного многочлена,* який приймає наперед задані значення $\beta_j \in P$ в $n+1$ різних точках $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_{n+1}$ поля P. Можна показати, що такий многочлен існує, отже, він єдиний.

Теорема 3. *Існує один і тільки один многочлен $f(x) \in P[x]$ не вище n-го степеня, який приймає в $n+1$ різних точках $\alpha_j \in P$ задані значення $\beta_j \in P$ $(j = 1, 2, \ldots, n+1)$.*

Доведення. З попереднього зрозуміло, що досить переконатися в існуванні хоч одного многочлена з потрібними властивостями. Дамо конструктивне доведення цього факту, тобто покажемо, як конкретно побудувати шуканий многочлен. Розглянемо многочлен

$$f(x) = \frac{(x-\alpha_2)(x-\alpha_3)\ldots(x-\alpha_{n+1})}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)\ldots(\alpha_1-\alpha_{n+1})}\beta_1 +$$

$$+ \frac{(x-\alpha_1)(x-\alpha_3)\ldots(x-\alpha_{n+1})}{(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)\ldots(\alpha_2-\alpha_{n+1})}\beta_2 + \cdots +$$

$$+ \frac{(x-\alpha_1)\ldots(x-\alpha_{j-1})(x-\alpha_{j+1})\ldots(x-\alpha_{n+1})}{(\alpha_j-\alpha_1)\ldots(\alpha_j-\alpha_{j-1})(\alpha_j-\alpha_{j+1})\ldots(\alpha_j-\alpha_{n+1})}\beta_j + \cdots +$$

$$+ \frac{(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n)}{(\alpha_{n+1}-\alpha_1)(\alpha_{n+1}-\alpha_2)\ldots(\alpha_{n+1}-\alpha_n)}\beta_{n+1}. \qquad (5)$$

Усі члени суми (5) є многочлени n-го степеня. При цьому j-й член побудований так, що при $x = \alpha_j$ він перетворюється в β_j, а при $x = \alpha_i$ $(i \neq j)$ — перетворюється в 0. Отже, $f(x)$ є многочлен не вище n-го степеня і такий, що $f(\alpha_j) = \beta_j$ $(j = 1, 2, \ldots, n+1)$, тобто — шуканий многочлен. Теорему доведено.

Многочлен (5) називають *інтерполяційним многочленом Лагранжа* [1]. Він розв'язує задачу інтерполяції [2] многочлена, яка полягає в тому, щоб за якимись $n+1$ значеннями многочлена не вище n-го степеня знайти всі його значення.

Приклад. Побудуємо многочлен не вище третього степеня над полем $\mathbf{Q}$ який при $x = -1$, $x = 0$, $x = 1$, $x = 2$ приймає значення 0, 1, 0, −1 відповідно.

За формулою (5)

$$f(x) = \frac{(x-0)(x-1)(x-2)}{(-1-0)(-1-1)(-1-2)}\cdot 0 + \frac{(x+1)(x-1)(x-2)}{(0+1)(0-1)(0-2)}\cdot 1 +$$

$$+ \frac{(x+1)(x-0)(x-2)}{(1+1)(1-0)(1-2)}\cdot 0 + \frac{(x+1)(x-0)(x-1)}{(2+1)(2-0)(2-1)}\cdot(-1) =$$

$$= \frac{1}{3}x^3 - x^2 - \frac{1}{3}x + 1.$$

Зауважимо, що теорема 2 та її наслідки справедливі і у випадку, коли P є область цілісності з одиницею. Проте теорема 2 неправильна для многочленів над кільцями, що мають дільники нуля. Так, якщо розглянути многочлен степеня 2 $f(x) = \bar{1}x^2$ над кільцем $\mathbf{Z}/(16)$ класів лишків цілих чисел за модулем 16 (див. п. 13.2) з елементами $\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{14}, \overline{15}$, то легко впевнитись, що він має чотири різні корені $\bar{0}, \bar{4}, \bar{8}, \overline{12}$.

23.3. Існування коренів многочлена. Поле розкладу. В попередньому пункті було з'ясовано, що число коренів многочлена не перевищує степеня цього многочлена. Однак питання про те, чи існує хоч один корінь будь-якого многочлена (ненульового степеня), залишилось відкритим. Тепер ми дамо відповідь на це питання.

Звичайно, легко навести приклади многочленів над полем P, які не мають жодного кореня у цьому полі. Так, многочлен $f(x) = 3x^2 - 1$ над полем $\mathbf{Q}$ не має раціональних коренів. Многочлен

[1] Ж. Л. Лагранж (1736—1813) — видатний французький математик і механік.

[2] Від латинського слова interpolatio — знаходження чи встановлення проміжних елементів між даними.

$g(x) = x^2 + 1$ над полем $\mathbf{R}$ не має дійсних коренів тощо. Але кожний з цих многочленів має корені у деякому р о з ш и р е н н і розглядуваного поля, а саме: $f(x)$ має корені $\pm \frac{1}{\sqrt{3}} \in \mathbf{R}$, $g(x)$ — корені $\pm i \in \mathbf{C}$.

Постає питання: чи відображають ці приклади загальну закономірність? Ствердну відповідь дає така важлива теорема.

Теорема 4 (Кронекера). *Якщо $f(x)$ — довільний многочлен над полем P, для якого $\deg f \geqslant 1$, то існує розширення K поля P, в якому є корінь $f(x)$.*

Д о в е д е н н я. Нехай $p(x)$ — один з незвідних множників многочлена $f(x)$ у полі P (якщо $f(x)$ — незвідний у полі P, то $p(x) = f(x)$). Тоді $f(x) = p(x)\,s(x)$ і досить довести твердження теореми для многочлена $p(x)$. Розглянемо головний ідеал (p) кільця $P[x]$, породжений елементом $p(x)$, і фактор-кільце $K = P[x]/(p)$. Оскільки $P[x]$ — кільце головних ідеалів, а (p) — ідеал, породжений простим елементом $p(x)$, то K є п о л е (теорема 6, п. 14.2). Це поле складається з класів лишків за модулем (p), представниками яких у даному разі є всі можливі остачі від ділення $h(x) \in P[x]$ на $p(x)$, тобто усі можливі многочлени не вище певного степеня з $P[x]$. Це поле містить усі константи $\bar{c}$, де $c \in P$, тобто є р о з ш и р е н н я м поля P. Саме у цьому розширенні існує корінь многочлена $p(x)$, таким коренем є елемент $\bar{x} \in K$. Справді, якщо $g(x)$ — будь-який многочлен з $K[x]$, то для знаходження $g(\bar{x})$ слід узяти остачу від ділення $g(x)$ на $p(x)$. Тому $p(\bar{x}) = \bar{0}$. Отже, в K справді існує корінь многочлена $p(x)$, а тому й $f(x)$. Теорему доведено.

Одним з важливих наслідків теореми Кронекера є таке твердження:

Теорема 5. *Для будь-якого многочлена $f(x) \in P[x]$ степеня $\deg f \geqslant 1$ існує таке розширення L поля P, в якому $f(x)$ розкладається на лінійні множники.*

Іншими словами, існує розширення L поля P, в якому степінь усіх незвідних множників многочлена $f(x)$ дорівнює 1.

Д о в е д е н н я. Нехай $\deg f = n$. За теоремою Кронекера, існує розширення K_1 поля P, в якому $f(x)$ має корінь α_1 і тому може бути поданий у вигляді $f(x) = (x - \alpha_1) f_1(x)$, де $f_1(x) \in K_1[x]$, $\deg f_1 = n - 1$. Застосовуючи тепер теорему Кронекера до поля K_1 і многочлена $f_1(x)$ (якщо $\deg f > 1$), дістанемо розширення K_2 поля K_1, в якому існує корінь α_2 многочлена $f_1(x)$. Зрозуміло, що α_2 є також коренем $f(x)$, а K_2 є розширенням поля P, в якому справедливий розклад

$$f(x) = (x - \alpha_1)(x - \alpha_2) f_2(x), \quad f_2(x) \in K_2[x], \quad \deg f = n - 2.$$

Цей процес можна продовжити, дістаючи розширення $K_3, \ldots, K_n$ поля P, корені $\alpha_3, \ldots, \alpha_n$ і дільники $f_3(x), \ldots, f_n(x)$ многочлена $f(x)$, поки не дістанемо $\deg f_{n+1} = 0$, тобто $f_{n+1} = c \in K_n$. Поле K_n і є шукане поле L, бо воно є розширенням поля P, в якому

$$f(x) = c(x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n). \tag{6}$$

Теорему доведено.

Означення 1. *Поле L, в якому многочлен $f(x)$ розкладається на лінійні множники, називають полем розкладу цього многочлена.*

Отже, теорема 5 означає, що для будь-якого многочлена $f(x) \in P[x]$ ненульового степеня існує поле розкладу L, яке є розширенням поля P. Звичайно, може бути, що $L = P$.

П р и к л а д 1. Многочлен $f(x) = x^4 - 2 \in \mathbf{Q}[x]$ не можна розкласти на множники у полі $\mathbf{Q}$ раціональних чисел. В кільці многочленів над полем чисел виду $a + b\sqrt{2}$ (a, b раціональні) маємо розклад:

$$x^4 - 2 = (x^2 - \sqrt{2})(x^2 + \sqrt{2}).$$

Дальше розкладання на множники у цьому кільці неможливе. Але в кільці $\mathbf{R}[x]$ многочленів над полем дійсних чисел дістанемо:

$$x^4 - 2 = (x - \sqrt[4]{2})(x + \sqrt[4]{2})(x^2 + \sqrt{2}).$$

Переходячи до ще більш широкого кільця многочленів $\mathbf{C}[x]$ над полем комплексних чисел, матимемо:

$$x^4 - 2 = (x - \sqrt[4]{2})(x + \sqrt[4]{2})(x - \sqrt[4]{2}i)(x + \sqrt[4]{2}i).$$

Оскільки всі множники тепер лінійні, робимо висновок, що $\mathbf{C}$ є поле розкладу многочлена $f(x)$.

Означення 2. *Поле P називається алгебраїчно замкнутим, якщо воно є полем розкладу для будь-якого многочлена $f(x) \in P[x]$ ненульового степеня.*

Іншими словами, *P є алгебраїчно замкнуте поле, якщо усі корені будь-якого многочлена $f(x) \in P[x]$ належать цьому самому полю.*

Важливим прикладом алгебраїчно замкнутих полів є поле $\mathbf{C}$ комплексних чисел. Твердження про алгебраїчну замкнутість цього поля за традицією часто називають *основною теоремою теорії многочленів або основною теоремою алгебри* (див. § 28).

Розклад (6) многочлена $f(x)$ на лінійні множники дає змогу дістати ряд важливих наслідків.

Наслідок 1. *Многочлен $f(x) \in P[x]$ n-го степеня має у полі розкладу n коренів.*

Оскільки $f(x)$ в жодному розширенні поля P не може мати більше за n коренів, то можна сказати, що *поле розкладу многочлена містить усі його корені.*

Наслідок 2. *У полі розкладу многочлен $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ має канонічний розклад виду*

$$f(x) = a_n(x - \alpha_1)^{k_1}(x - \alpha_2)^{k_2} \ldots (x - \alpha_m)^{k_m} \quad (k_1 + k_2 + \ldots + k_m = n), \tag{7}$$

де $\alpha_1, \alpha_2, \ldots, \alpha_m$ — різні корені многочлена $f(x)$.

Справді, оскільки у розкладі (6) можуть бути однакові множники, то його можна записати так:

$$f(x) = c(x - \alpha_1)^{k_1}(x - \alpha_2)^{k_2} \ldots (x - \alpha_m)^{k_m}; \quad (k_1 + k_2 + \ldots + k_m = n), \tag{8}$$

де α_j попарно різні. При цьому $x - \alpha_j$ є попарно різні незвідні множники, тобто (8) є канонічний розклад $f(x)$. Що ж до константи c, то,

прирівнюючи коефіцієнти при x^n у многочленів в обох частинах рівності (8), дістанемо $c = a_n$. Цим (7) доведено.

Прирівнюючи й інші коефіцієнти многочленів в обох частинах (7), дістанемо формули, які пов'язують між собою корені і коефіцієнти многочлена $f(x)$ у будь-якому полі розкладу цього многочлена.

Теорема 6 (Вієта) [1] *Якщо* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *— корені многочлена* $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in P[x]$, *то*

$$\begin{aligned}
&\alpha_1 + \alpha_2 + \cdots + \alpha_n = -\frac{a_{n-1}}{a_n},\\
&\alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_1\alpha_n + \alpha_2\alpha_3 +\\
&+ \cdots + \alpha_2\alpha_n + \cdots + \alpha_{n-1}\alpha_n = \frac{a_{n-2}}{a_n},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\sum_{C_n^k} \alpha_{j_1}\alpha_{j_2} \ldots \alpha_{j_k} = (-1)^k \frac{a_{n-k}}{a_n},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\alpha_1\alpha_2 \ldots \alpha_{n-1}\alpha_n = (-1)^n \frac{a_0}{a_n}.
\end{aligned} \tag{9}$$

Символ $\sum_{C_n^k}$ слід тут розуміти так, що сума береться по всіх C_n^k комбінаціях з n індексів 1, 2, 3, ..., n по k.

Для доведення формул (9) досить виконати множення у правій частині рівності

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = a_n(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

звести подібні члени і прирівняти коефіцієнти при однакових степенях x в обох частинах.

Співвідношення (9) називають формулами Вієта. Звичайно, читач помітив, що вони узагальнюють відомі з шкільного курсу формули Вієта для коренів квадратного тричлена.

Приклад 2. Для многочлена $f(x) = x^n - 1$ поле $\mathbf{C}$ є поле розкладу. Його коренями є $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}$ — корені n-го степеня з одиниці (1, § 16). За формулами Вієта маємо:

$$\begin{aligned}
&\varepsilon_0 + \varepsilon_1 + \cdots + \varepsilon_{n-1} = 0\\
&\varepsilon_0\varepsilon_1 + \varepsilon_0\varepsilon_2 + \cdots + \varepsilon_{n-2}\varepsilon_{n-1} = 0\\
&\cdots\cdots\cdots\cdots\cdots\\
&\varepsilon_0\varepsilon_1 \ldots \varepsilon_{n-2}\varepsilon_{n-1} = (-1)^{n+1}.
\end{aligned}$$

23.4. Похідна від многочлена. З курсу математичного аналізу відомо, що кожний многочлен n-го степеня з дійсними коефіцієнтами

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \tag{10}$$

[1] Ф. Вієт (1540—1603) — видатний французький математик.

розглядуваний як функція на множині всіх дійсних чисел, має в кожній точці x похідну $f'(x)$, яка також є многочленом, але вже $(n-1)$-го степеня:

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1. \tag{11}$$

У тих випадках, коли $f(x)$ є многочлен над іншим (взагалі кажучи, абстрактним) полем P, його можна розглядати як функцію, задану на деякому розширенні поля P (п. 21.5). Проте поняття г р а н и ц і, за допомогою якого вводять похідну в аналізі, застосовне не в усякому полі. Тому означимо поняття похідної від многочлена ф о р м а л ь н о, домовившись вираз (11) завжди називати похідною від многочлена (10) незалежно від того, до якого поля належать його коефіцієнти, і незалежно від множини його задання.

Означення. *Похідною від многочлена*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

називається многочлен

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1.$$

Похідну від многочлена нульового степеня, а також похідну від нуль-многочлена беруть рівною нулю.

Безпосередньо з означення випливає, що *похідна від многочлена над полем P є знову многочлен над полем P*. Справді, з формули (11) видно, що коефіцієнти $f'(x)$ дістаємо з коефіцієнтів $f(x)$ за допомогою множення на натуральні числа, тобто коефіцієнти $f'(x)$ є кратними елементів поля P і тому є також елементами поля P.

Досі ми не накладали на поле P ніяких обмежень. Проте у дальшому викладі завжди вважатимемо, що P *є поле характеристики* 0. При цій додатковій умові можна твердити, що $\deg f' = \deg f - 1$, якщо $\deg f \geqslant 1$, бо з $a_n \neq 0$ при будь-якому n випливає $na_n \neq 0$. У випадку ж поля скінченної характеристики $p \geqslant 1$ многочлен $g(x) = x^p$ степеня p має похідною нуль-многочлен, бо $n \cdot p = 0$.

Для многочленів над полем дійсних чисел (як і для всіх диференційованих функцій) справедливі, як відомо, такі правила диференціювання:

$$[f(x) + g(x)]' = f'(x) + g'(x), \tag{12}$$

$$[f(x)\,g(x)]' = f'(x)\,g(x) + f(x)\,g'(x) \tag{13}$$

і, зокрема,

$$\begin{aligned}
&[cf(x)]' = cf'(x) && (c \text{ — константа}),\\
&\{[f(x)]^k\}' = k[f(x)]^{k-1}f'(x) && (k \text{ — натуральне число}).
\end{aligned} \tag{14}$$

Рівності (12), (13), (14) залишаються в силі для похідних від многочленів над д о в і л ь н и м полем. Справді, (12) і (13) є рівностями, що пов'язують між собою деякі многочлени з кільця $P[x]$ многочленів над даним полем. Проте справедливість цих рівностей не залежить від того, над яким полем ці многочлени розглядаються, бо вони означають просто рівність відповідних коефіцієнтів многочленів в обох частинах рівностей.

Пояснимо цю думку на прикладі формули (12). Якщо дано два многочлени з кільця $\mathbf{R}[x]$:

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$
$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 \ (n \geqslant m), \text{ то}$$
$$f(x) + g(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + (a_m + b_m)x^m + \cdots + (a_1 + b_1)x + (a_0 + b_0).$$

Переходячи до похідних за правилом (11), помічаємо, що формула (12) означає просто тотожну рівність многочленів

$$na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + m(a_m + b_m)x^{m-1} + \cdots + (a_1 + b_1) = [na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + ma_m x^{m-1} + \cdots + a_1] + [mb_m x^{m-1} + \cdots + b_1]. \quad (15)$$

Ця тотожна рівність над полем $\mathbf{R}$ *рівносильна рівності відповідних коефіцієнтів* тих многочленів, які утворюються в обох частинах (15) після виконання певних раціональних дій над многочленами. Але тоді зрозуміло, що коли коефіцієнти $f(x)$ і $g(x)$ належать довільному іншому полю P характеристики 0, то формула (12) залишається справедливою.

У зв'язку з цим ми можемо вільно користуватись для довільних многочленів правилами диференціювання суми і добутку, які були виведені в курсі аналізу для функцій дійсного змінного за допомогою властивостей границь.

В алгебрі розглядають також другу, третю, ..., k-ту похідні від многочлена $f(x)$, відповідно позначаючи їх $f''(x)$, $f'''(x)$, ..., $f^{(k)}(x)$; при цьому $f^{(k)}(x)$ визначають як похідну від $f^{(k-1)}(x)$. Якщо $f(x)$ є многочлен n-го степеня, то, очевидно, $f^{(k)}(x) = 0$ при $k > n$, а $f^{(n)}(x) = n!a_n$ (де a_n — старший коефіцієнт многочлена $f(x)$).

23.5. Відокремлення кратних множників. У попередньому параграфі було доведено, що всякий многочлен над полем P можна єдиним способом подати у вигляді добутку многочленів нижчих степенів, незвідних у цьому полі:

$$f(x) = [p_1(x)]^{k_1}[p_2(x)]^{k_2} \cdots [p_l(x)]^{k_l}. \quad (16)$$

Отже, вивчати властивості многочленів і, зокрема, знаходити їх корені було б значно легше, якби для кожного многочлена був відомий канонічний розклад (16); тоді було б досить розглядати лише незвідні множники, степінь яких, як правило, значно нижчий за степінь даного многочлена. Але насправді ми встановили лише принципову можливість побудувати для многочлена розклад (16). Що ж до фактичного здійснення цього розкладу, тобто знаходження незвідних множників многочлена, то у нашому розпорядженні ще немає ніяких загальних методів розв'язання цієї задачі. Більше того, виявляється, що в загальному випадку ця задача рівнозначна знаходженню коренів многочлена, тобто тій проблемі, для розв'язання якої ми взялися за розкладання многочлена на множники.

Проте ми покажемо, що в деяких випадках можна розробити загальний метод розкладання многочлена на множники, хоч цей розклад не буде таким повним, як зображення (16).

Виберемо у розкладі (16) ті незвідні множники $p_i(x)$, кратність яких k_i дорівнює одиниці, і позначимо добуток цих множників через $\varphi_1(x)$:

$$\varphi_1(x) = p_{i_1}(x)\,p_{i_2}(x) \cdots p_{i_s}(x).$$

Тепер утворимо добуток тих множників $p_j(x)$, кратність яких k_j дорівнює 2, тобто тих, які входять у розклад (16) у другому степені:

$$\varphi_2(x) = p_{j_1}(x)\,p_{j_2}(x) \cdots p_{j_r}(x).$$

Зауважимо, що $\varphi_2(x)$ означає добуток самих незвідних множників кратності 2, а не їх квадратів, так що в розклад входить $[\varphi_2(x)]^2$. Аналогічно, через $\varphi_3(x)$ позначимо добуток незвідних множників з розкладу (16), що мають кратність 3, і т. д.

Тоді розклад (16) можна записати у такому вигляді:

$$f(x) = \varphi_1(x)[\varphi_2(x)]^2[\varphi_3(x)]^3 \cdots [\varphi_m(x)]^m,$$

або, випускаючи для спрощення запису позначення змінної x,

$$f = \varphi_1\varphi_2^2\varphi_3^3 \cdots \varphi_m^m. \quad (17)$$

Якщо множників кратності $k < m$ в розкладі (16) немає, природно вважати $\varphi_k = 1$.

Ми зараз викладемо загальний метод зображення многочленів у вигляді (17). Зрозуміло, що таке розкладання многочлена доцільне лише тоді, коли в зображенні (16) справді існують множники, кратність яких перевищує одиницю, тобто кратні множники. У противному разі $f(x) = \varphi_1(x)$, тобто ніякого розкладу на множники не буде.

Приклад 1. Нехай розклад многочлена $f(x)$ на незвідні множники в полі дійсних чисел має вигляд:

$$f(x) = x^{13} - 5x^{12} + 6x^{11} + 4x^{10} - 9x^9 + 5x^8 - 6x^7 - 4x^6 + 8x^5 = x^5(x-2)^3(x^2+1)(x+1)^2(x-1).$$

Тут

$$\varphi_1(x) = (x^2+1)(x-1) = x^3 - x^2 + x - 1,$$
$$\varphi_2(x) = x + 1, \quad \varphi_3(x) = x - 2,$$
$$\varphi_4(x) = 1, \quad \varphi_5(x) = x,$$

а $f(x) = \varphi_1(x)[\varphi_2(x)]^2[\varphi_3(x)]^3[\varphi_4(x)]^4[\varphi_5(x)]^5$.

Як бачимо, $f(x)$ — многочлен 13-го степеня, а степінь многочленів $\varphi_k(x)$ не перевищує 3. Тому в цьому прикладі розклад многочлена $f(x)$ на множники $\varphi_k(x)$ дає змогу повністю знайти всі його корені, бо ми вміємо розв'язувати рівняння 1—3 степенів.

Задача зображення многочлена у вигляді (17) називається *відокремленням кратних множників*. До розгляду цієї задачі ми й переходимо.

Є досить простий спосіб, який дає змогу, користуючись похідною (п. 23.4), визначати, чи має многочлен кратні множники.

Теорема 7. *Якщо незвідний у даному полі P характеристики 0 многочлен $q(x)$ є множником кратності $k \geqslant 2$ для многочлена $f(x)$, то він є множником кратності k — і для похідної $f'(x)$. Якщо $q(x)$ є множником першої кратності многочлена $f(x)$, то він не входить у розклад похідної $f'(x)$ на незвідні множники.*

Доведення. Те, що многочлен $q(x)$ є множником кратності k для многочлена $f(x)$, можна записати так: $f(x) = [q(x)]^k \varphi(x)$, де $\varphi(x)$ — многочлен, який не ділиться на незвідний множник $q(x)$ і тому взаємно простий з ним (п. 22.6).

За правилами (12), (13), (14) маємо:

$$f'(x) = k[q(x)]^{k-1} \cdot q'(x)\varphi(x) + [q(x)]^k \cdot \varphi'(x) =$$
$$= [q(x)]^{k-1}\{kq'(x)\varphi(x) + q(x)\varphi'(x)\}.$$

Звідси ми бачимо, що $f'(x)$ ділиться на $[q(x)]^{k-1}$. Для того щоб довести, що $q(x)$ є для $f'(x)$ множником саме $(k-1)$-ї кратності, треба ще показати, що многочлен

$$\psi(x) = kq'(x)\varphi(x) + q(x)\varphi'(x)$$

не ділиться на $q(x)$. Якби $\psi(x)$ ділився на $q(x)$, то й

$$q'(x)\varphi(x) = \frac{1}{k}[\psi(x) - q(x)\varphi'(x)]$$

ділився б на $q(x)$. Оскільки $\varphi(x)$ взаємно простий з $q(x)$, це означало б, що $q'(x)$ ділиться на $q(x)$, а це неможливо, бо $q'(x)$ має степінь, менший за степінь $q(x)$.

Якщо кратність k множника $q(x)$ дорівнює 1, то дістаємо:

$$f'(x) = q'(x)\varphi(x) + q(x)\varphi'(x).$$

З наведених вище міркувань видно, що $f'(x)$ на $q(x)$ не ділиться. Це й означає, що в цьому разі $q(x)$ не входить у розклад многочлена $f(x)$ на незвідні множники у даному полі. Теорему доведено.

Наслідок. *Для того щоб многочлен $f(x)$ не мав кратних множників, необхідно і достатньо, щоб $f(x)$ був взаємно простим із своєю похідною $f'(x)$.*

Справді, якщо всі незвідні множники многочлена $f(x)$ мають кратність 1, то в розкладі $f'(x)$ на незвідні множники не буде жодного множника, спільного з множниками многочлена $f(x)$. Тому, за теоремою 9 § 22, $(f, f') = 1$. Якщо ж $f(x)$ має хоч один кратний множник $q(x)$, то (f, f') ділиться на $q(x)$ і тому не може бути константою.

З доведеного наслідку, зокрема, випливає, що наявність чи відсутність кратних множників у даного многочлена залежить лише від його коефіцієнтів, а не від того поля, над яким його розглядають.

Приклад 2. У п. 22.5 було встановлено, що многочлени

$$f(x) = x^4 + x^3 + x^2 + x + 1 \text{ і } g(x) = 4x^3 + 3x^2 + 2x + 1$$

взаємно прості. Легко бачити, що $g(x) = f'(x)$. Отже, многочлен $f(x)$ кратних множників не має.

Як ми знаємо, кожний многочлен у даному полі можна подати у вигляді (17)

$$f = \varphi_1\varphi_2^2\varphi_3^3 \ldots \varphi_m^m.$$

Наше завдання полягає в тому, щоб, знаючи лише коефіцієнти многочлена $f(x)$, визначити многочлени $\varphi_1, \varphi_2, \ldots, \varphi_m$.

Оскільки $\varphi_1(x)$ є добутком незвідних множників многочлена $f(x)$, які мають кратність $k = 1$, то в $f'(x)$ жодний з цих множників (на підставі теореми 7) входити не буде. $\varphi_2(x)$ є добутком незвідних множників $f(x)$ кратності 2. У $f'(x)$ усі ці множники входять з кратністю одиниця, тобто $f'(x)$ має своїм множником добуток $\varphi_2(x)$ усіх цих незвідних множників, але вже у першому степені. Аналогічно, якщо $f(x)$ має множником $[\varphi_k(x)]^k$, то $f'(x)$ матиме множник $[\varphi_k(x)]^{k-1}$.

Таким чином, можемо записати:

$$f' = \varphi_2\varphi_3^2 \ldots \varphi_m^{m-1}\psi_1,$$

де ψ_1 не ділиться на $\varphi_1, \varphi_2, \ldots, \varphi_m$. Тоді, на підставі теореми 9 § 22, спільний найбільший дільник (f, f') є добутком усіх множників, які входять у розклади як $f(x)$, так і $f'(x)$:

$$d_1 = (f, f') = \varphi_2\varphi_3^2 \ldots \varphi_m^{m-1}.$$

Знайдемо тепер d_1' : $d_1' = \varphi_3\varphi_4^2 \ldots \varphi_m^{m-2}\psi_2$, де ψ_2 не ділиться на φ_i $(i = 2, \ldots, m)$. Далі,

$$d_2 = (d_1, d_1') = \varphi_3\varphi_4^2 \ldots \varphi_m^{m-2}.$$

Аналогічно можна обчислити:

$$\begin{aligned}
d_3 &= (d_2, d_2') = \varphi_4\varphi_5^2 \ldots \varphi_m^{m-3},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
d_{m-2} &= (d_{m-3}, d_{m-3}') = \varphi_{m-1}\varphi_m^2,\\
d_{m-1} &= (d_{m-2}, d_{m-2}') = \varphi_m,\\
d_m &= (d_{m-1}, d_{m-1}') = 1.
\end{aligned} \tag{18}$$

$d_1, d_2, \ldots, d_m$ вже не містять непотрібних нам множників ψ_j. Проте наша мета полягає в тому, щоб знайти кожний з множників φ_i окремо.

Для цього поділимо f на d_1. Дістанемо

$$q_1 = \frac{f}{d_1} = \varphi_1\varphi_2 \ldots \varphi_m. \tag{19}$$

Аналогічно:

$$q_2 = \frac{d_1}{d_2} = \varphi_2\varphi_3 \ldots \varphi_m,$$

$$q_3 = \frac{d_2}{d_3} = \varphi_3\varphi_4 \ldots \varphi_m,$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$q_{m-1} = \frac{d_{m-2}}{d_{m-1}} = \varphi_{m-1}\varphi_m, \tag{20}$$

$$q_m = \frac{d_{m-1}}{d_m} = \varphi_m.$$

З формул (19) і (20) шукані множники φ_k дістаємо вже безпосередньо:

$$\varphi_1 = \frac{q_1}{q_2}, \quad \varphi_2 = \frac{q_2}{q_3}, \quad \ldots, \quad \varphi_{m-1} = \frac{q_{m-1}}{q_m}, \quad \varphi_m = q_m. \tag{21}$$

Отже, ми приходимо до такого висновку:

У довільного многочлена над полем P можна відокремити кратні множники за допомогою скінченного числа раціональних дій над деякими многочленами.

Оскільки, як нам відомо, похідні та найбільші спільні дільники не залежать від того, над яким полем розглядаються задані многочлени, то й результати обчислень за формулами (18), (19), (20), (21) не залежать від цього поля, а лише від коефіцієнтів заданого многочлена.

Отже, розклад (17) многочлена над даним полем, на відміну від розкладу (16), не залежить від того, до якого основного поля P відносимо коефіцієнти цього многочлена.

Приклад 3. Відокремити кратні множники многочлена

$$f(x) = x^6 - 6x^4 - 4x^3 + 9x^2 + 12x + 4.$$

Знайдемо спочатку многочлени $d_l(x) \cdot d_1(x)$ — це НСД многочленів $f(x)$ і $f'(x) = 6x^5 - 24x^3 - 12x^2 + 18x + 12$. Застосувавши алгоритм Евкліда, який було розглянуто в п. 22.5, дістанемо $d_1(x) = x^4 + x^3 - 3x^2 - 5x - 2$. Далі маємо $d_1'(x) = 4x^3 + 3x^2 - 6x - 5$ і знаходимо $d_2(x) = (d_1, d_1')$. Дістанемо $d_2(x) = x^2 + 2x + 1$; $d_2'(x) = 2x + 2$. Знаходимо $d_3(x) = (d_2, d_2')$. Маємо $d_3(x) = x + 1$, $d_3'(x) = 1$, тому $d_4(x) = 1$.

Обчислимо тепер многочлени $q_l(x)$:

$$q_1(x) = \frac{f(x)}{d_1(x)} = x^2 - x - 2; \quad q_2(x) = \frac{d_1(x)}{d_2(x)} = x^2 - x - 2;$$

$$q_3(x) = \frac{d_2(x)}{d_3(x)} = x + 1; \quad q_4(x) = d_3(x) = x + 1.$$

Тепер уже можна знайти множники $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$, $\varphi_4(x)$:

$$\varphi_1(x) = \frac{q_1(x)}{q_2(x)} = 1; \quad \varphi_2(x) = \frac{q_2(x)}{q_3(x)} = x - 2;$$

$$\varphi_3(x) = \frac{q_3(x)}{q_4(x)} = 1; \quad \varphi_4(x) = q_4(x) = x + 1.$$

Маємо остаточно: $f(x) = \varphi_1(x)\,\varphi_2^2(x)\,\varphi_3^3(x)\,\varphi_4^4(x)$, тобто $f(x) = (x-2)^2\,(x+1)^4$.

Отже, ми многочлен 6-го степеня звели до досить простої форми; в даному разі розклад (17) на незвідні множники такий самий, як і розклад (16), бо кожний з множників виявився лінійним, а тому й незвідним у довільному полі.

Відокремлення кратних множників не тільки спрощує дослідження і знаходження коренів многочленів, але й іноді, як буде з'ясовано в розділі VII, є передумовою для застосування деяких методів дослідження і розв'язування алгебраїчних рівнянь.

У таких випадках немає потреби знаходити кожний з множників $\varphi_l(x)$. Завдання полягає в тому, щоб позбутись кратних множників, побудувавши за многочленом $f(x)$ такий многочлен $q(x)$, який має всі ті незвідні множники, які має й $f(x)$, але вже першої кратності. Як показує формула (19), для цього досить поділити $f(x)$ на НСД многочлена $f(x)$ і його похідної $f'(x)$.

На закінчення цього параграфа розглянемо ще способи встановлення кратності кореня многочлена.

Згідно з означенням 3 п. 23.1, для визначення кратності кореня α досить послідовним діленням $f(x)$ на $x - \alpha$ знайти таке k, щоб $f(x)$ ділився на $(x-\alpha)^k$, але не ділився на $(x-\alpha)^{k+1}$. Очевидно, k і буде кратністю кореня α.

Приклад 4. Многочлен $f(x) = x^6 - 6x^4 - 4x^3 + 9x^2 + 12x + 4$ має в полі $\mathbf{Q}$ корінь $\alpha = 2$, як це можна встановити безпосередньою перевіркою. Послідовно ділитимемо $f(x)$ на $x - 2$, використовуючи схему Горнера:

	1	0	−6	−4	9	12	4
2	1	2	−2	−8	−7	−2	0
2	1	4	6	4	1	0	
2	1	6	18	40	81		

Многочлен $f(x)$ ділиться на $(x-2)^2$, але не ділиться на $(x-2)^3$. Отже, кратність кореня 2 дорівнює двом.

У ряді випадків буває зручно користуватись такою ознакою кратності кореня.

Теорема 8. *Для того, щоб α був коренем кратності k многочлена $f(x)$, необхідно і достатньо, щоб*

$$f(\alpha) = f'(\alpha) = \cdots = f^{(k-1)}(\alpha) = 0; \quad f^{(k)}(\alpha) \neq 0. \tag{22}$$

Доведення. *Необхідність.* Нехай α є коренем $f(x)$ кратності k. Це означає, що $x - \alpha$ є незвідним множником k-ї кратності многочлена $f(x)$. Але тоді, за теоремою 7, $(x - \alpha)$ є незвідним множником похідної $f'(x)$ кратності $k - 1$, тобто α є коренем $(k-1)$-ї кратності многочлена $f'(x)$. Аналогічно, α є коренем $(k-2)$-ї кратності многочлена $f''(x)$, $(k-3)$-ї кратності многочлена $f'''(x)$ і т. д. Нарешті, $f^{(k-1)}(x)$ має $x - \alpha$ своїм множником кратності 1, а $f^{(k)}(x)$ цього множника не має зовсім, і тому, за наслідком з теореми Безу (п.22.3), $f^{(k)}(\alpha) \neq 0$. Отже, умова (22) справджується.

Достатність. Нехай справджується умова 22. Тоді α є коренем многочлена $f(x)$. Позначимо кратність цього кореня через l і покажемо, що $l = k$. З доведеного вище випливає, що

$$f(\alpha) = f'(\alpha) = \cdots = f^{(l-1)}(\alpha) = 0; \quad f^{(l)}(\alpha) \neq 0. \tag{23}$$

Якби було $l < k$, то (22) свідчило б про те, що $f^{(l)}(\alpha) = 0$, а це суперечить (23). Аналогічно відкидаємо припущення $l > k$. Отже, $l = k$. Теорему доведено.

Приклад 5. Застосовуємо цю ознаку до того самого многочлена $f(x) = x^6 - 6x^4 - 4x^3 + 9x^2 + 12x + 4$ (приклад 4). Маємо

$$f'(x) = 6(x^5 - 4x^3 - 2x^2 + 3x + 2); \quad f'(2) = 0;$$
$$f''(x) = 6(5x^4 - 12x^2 - 4x + 3); \quad f''(2) = 162 \neq 0.$$

Отже, кратність кореня 2 дорівнює двом.

Наведені міркування показують, що, знаючи корінь многочлена, легко визначити його кратність. Тому на практиці дослідження многочленів, які мають кратні корені, у більшості випадків зводять до дослідження многочленів нижчих степенів, що вже не мають кратних коренів. Це завжди можна зробити відокремленням кратних множників методами, описаними вище.

Приклад 6. Розглянемо знову многочлен $f(x) = x^6 - 6x^4 - 4x^3 + 9x^2 + 12x + 4$. Як відомо з прикладу 3 $(f, f') = x^4 + x^3 - 3x^2 - 5x - 2$. Це свідчить про наявність кратних множників. Щоб позбутися їх, поділимо $f(x)$ на (f, f'). Дістанемо $q(x) = x^2 - x - 2$.
Цей многочлен має ті самі незвідні множники, що й $f(x)$, але вже першої кратності. Оскільки $q(x)$ — квадратний тричлен, легко знайти його корені. Маємо $\alpha_1 = 2$, $\alpha_2 = -1$. Тепер можна з'ясувати їх кратність. Вище було показано, що 2 є двократним коренем. Просте врахування степенів показує, що $x + 1$ є чотирикратним множником многочлена $f(x)$, тобто -1 є коренем кратності 4.

§ 24. ПОЛЕ РАЦІОНАЛЬНИХ ДРОБІВ

24.1. Раціональні дроби. В п. 12.4 було показано, що будь-яку область цілісності можна «вкласти» у деяке поле. Більш конкретно було доведено, що для будь-якої області цілісності R існує єдине (з точністю до ізоморфізму) поле Q, яке містить R (тобто підмножину ізоморфну R) і має ту властивість, що кожний елемент з Q можна подати як частку двох елементів з R. При цьому Q називають полем часток (відношень) області цілісності R.

Стосовно області цілісності $P[x]$ це загальне твердження означає справедливість такої теореми:

Теорема 1. *Для будь-якого поля P існує єдине* (з точністю до ізоморфізму) *поле $P(x)$, яке містить кільце $P[x]$ многочленів над полем P і кожний елемент якого можна подати у вигляді частки*

$$\frac{f(x)}{g(x)}, \quad \text{де} \quad f(x), \quad g(x) \in P[x], \quad g(x) \neq 0 \tag{1}$$

$P(x)$ називається полем раціональних дробів над P, а кожний його елемент — раціональним дробом над P.

Важливо підкреслити, що згідно з загальною теорією елементом поля $P(x)$ є не кожна окрема частка (1) (яка визначається упорядкованою парою многочленів $f(x)$, $g(x)$), а клас рівних між собою часток,

причому цю рівність за означенням слід тлумачити так:

$$\forall_{\substack{f_1(x), f_2(x), g_1(x), g_2(x) \in P[x] \\ g_1(x) \neq 0, g_2(x) \neq 0}} \left[\frac{f_1(x)}{g_1(x)} = \frac{f_2(x)}{g_2(x)} \Leftrightarrow f_1(x) g_2(x) = f_2(x) g_1(x) \right]. \tag{2}$$

Тому кожний раціональний дріб можна подати у формі (1) багатьма способами, вибираючи по-різному многочлени $f(x)$ і $g(x)$ з додержанням умови рівності (2). Як правило, ми будем раціональний дріб представляти тією з часток $\frac{f(x)}{g(x)}$ відповідного класу, для якої $(f, g) = 1$; таку частку називають нескоротною. Ця домовленість спирається на той факт, що *в кожному класі рівних між собою часток нескоротна частка $\frac{f(x)}{g(x)}$ існує і є єдина* (з точністю до сталих множників при многочленах $f(x)$ і $g(x)$). Справді, якщо $\frac{\varphi}{\psi}$ — будь-яка частка з даного класу K (позначення змінної для скорочення не пишемо) і $d = (\varphi, \psi)$, то $\varphi = fd$, $\psi = gd$, де $(f, g) = 1$. При цьому $\frac{\varphi}{\psi} = \frac{f}{g}$, бо $\varphi g = fdg = f\psi$. Отже, $\frac{f}{g} \in K$ і є нескоротною часткою. Цим доведено існування нескоротної частки в будь-якому класі K. Єдиність (з точністю до сталих множників) випливає з таких міркувань. Якщо $\frac{f}{g}$, $\frac{f_1}{g_1}$ — нескоротні частки і $\frac{f}{g} = \frac{f_1}{g_1}$, то $fg_1 = f_1 g$; тому на підставі властивостей подільності многочленів (§ 22) можна твердити, що або $f = f_1 = 0$, або $f \vdots f_1$ [адже $fg_1 \vdots f_1 \wedge (f_1, g_1) = 1$] і $f_1 \vdots f$ [бо $f_1 g \vdots f \wedge (f, g) = 1$]. Отже, $f_1 = cf$, де c — константа. Аналогічно, $g_1 = c_1 g$ (c_1 — константа).

Якщо раціональний дріб подано у формі частки $\frac{f(x)}{g(x)}$, то $f(x)$ часто називають чисельником, $g(x)$ — знаменником цього дробу. Як видно з попереднього, чисельник і знаменник раціонального дробу визначаються неоднозначно. Многочлени $f(x) \in P[x]$, що утворюють підкільце поля $P(x)$, подаватимемо як раціональні дроби з знаменником 1: $f(x) = \frac{f(x)}{1}$. Ототожнення многочлена $f(x)$ з раціональним дробом, що визначається часткою $\frac{f(x)}{1}$, є по суті встановлення ізоморфізму між кільцем $P[x]$ і деяким підкільцем поля $P(x)$.

Дії над раціональними дробами (класами часток), як було показано в загальній теорії поля відношень області цілісності (§ 12), можна означити через дії над частками (представниками класів). Для випадку поля $P(x)$ правила дій набувають такого вигляду

$$\frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)} = \frac{f_1(x) g_2(x) + f_2(x) g_1(x)}{g_1(x) g_2(x)}, \quad (g_1(x) \neq 0, \ g_2(x) \neq 0); \tag{3}$$

$$\frac{f_1(x)}{g_1(x)} - \frac{f_2(x)}{g_2(x)} = \frac{f_1(x) g_2(x) - f_2(x) g_1(x)}{g_1(x) g_2(x)}, \quad (g_1(x) \neq 0, \ g_2(x) \neq 0); \tag{4}$$

$$\frac{f_1(x)}{g_1(x)} \cdot \frac{f_2(x)}{g_2(x)} = \frac{f_1(x) f_2(x)}{g_1(x) g_2(x)}, \quad (g_1(x) \neq 0,\ g_2(x) \neq 0); \tag{5}$$

$$\frac{f_1(x)}{g_1(x)} : \frac{f_2(x)}{g_2(x)} = \frac{f_1(x) g_2(x)}{f_2(x) g_1(x)}, \quad (g_1(x) \neq 0,\ g_2(x) \neq 0,\ f_2(x) \neq 0). \tag{6}$$

Звичайно, кожний читач має значний досвід практичного використання правил (3) — (6) при перетворенні алгебраїчних виразів. Ми згадуємо їх тут у зв'язку з тим, що з них (3) і (5) є означеннями дій, які перетворюють множину класів символів $\frac{f(x)}{g(x)}$ у поле $P(x)$, а (4) і (6) випливають з властивостей цього поля. Ми не будемо спинятися на питанні про властивості дій над раціональними дробами, бо вони є такими ж, як і властивості дій у будь-якому полі. Зауважимо тільки, що як специфічну дію у полі раціональних дробів можна розглядати так зване скорочення, що полягає у використанні рівності:

$$\underset{\substack{f(x),g(x),s(x) \in P[x] \\ g(x) \neq 0, s(x) \neq 0}}{\forall} \left[\frac{f(x)\, s(x)}{g(x)\, s(x)} = \frac{f(x)}{g(x)} \right]. \tag{7}$$

Рівність (7) випливає безпосередньо з означення (2).

24.2. Функціональне тлумачення раціональних дробів. Відомо (п.21.5), що многочлени над полем (або областю цілісності) P характеристики 0 можна розглядати і як функції, задані на множині елементів P. Таке функціональне тлумачення можна поширити і на раціональні дроби, але з деякими застереженнями.

Отже, нехай P — поле характеристики 0, $f(x)$ і $g(x)$ — взаємно прості многочлени з кільця $P[x]$, причому $g(x) \neq 0$ (тобто не є нуль-многочленом). Для довільного елемента $\alpha \in P$ відповідні значення многочленів $f(\alpha)$ і $g(\alpha)$ є якісь елементи поля P. Якщо $g(\alpha) \neq 0$ (тобто, якщо α не є коренем $g(x)$), то їх частка $\frac{f(\alpha)}{g(\alpha)}$ є цілком певним елементом поля P. Позначимо сукупність коренів многочлена $g(x)$, які належать до поля P, через K_g (звичайно, може бути $K_g = \varnothing$). Тоді, очевидно, можна задати функцію

$$\varphi : P \setminus K_g \to P$$

умовою

$$\underset{\alpha \in P \setminus K_g}{\forall} \left[\varphi(\alpha) = \frac{f(\alpha)}{g(\alpha)} \right].$$

Постає питання: *як пов'язана ця функція з раціональним дробом, вираженим нескоротною часткою* $\frac{f(x)}{g(x)}$?

Кожному раціональному дробу $r(x) \in P(x)$ поставимо у відповідність функцію φ_r в такий спосіб. Подамо $r(x)$ у формі нескоротного дробу $r(x) = \frac{f(x)}{g(x)}$, $(f, g) = 1$. Як було зазначено, це завжди можна зробити і притому єдиним способом. Щоб уникнути неоднозначності, пов'язаної з сталими множниками, будемо вважати, що старший коефіцієнт знаменника $g(x)$ завжди дорівнює 1.

Тепер означимо функцію φ_r умовою

$$\underset{\alpha \in P \setminus K_g}{\forall} \left[\varphi_r(\alpha) = \frac{f(\alpha)}{g(\alpha)} \right],$$

де K_g, як і вище, є сукупність елементів поля P, що є коренями многочлена $g(x)$.

Покажемо, що відповідність $r(x) \to \varphi_r$ є ізоморфізм.

Однозначність відповідності вже було підкреслено. Взаємна однозначність випливає з таких міркувань. Нехай $r_1(x) = \frac{f_1(x)}{g_1(x)}$, $(f_1, g_1) = 1$ і $r_2(x) = \frac{f_2(x)}{g_2(x)}$, $(f_2, g_2) = 1$. Слід показати, що при $\varphi_{r_1} = \varphi_{r_2}$ (в розумінні рівності двох функцій, тобто тотожної рівності усіх значень на спільній множині задання), матимемо $r_1(x) = r_2(x)$ (у розумінні рівності раціональних дробів). Множиною задання $\varphi_{r_1} \in P \setminus K_{g_1}$, множиною задання $\varphi_{r_2} \in P \setminus K_{g_2}$. Отже, спільна множина задання функцій φ_{r_1} і $\varphi_{r_2} \in P \setminus (K_{g_1} \cup K_{g_2})$ df P_1. Оскільки P — поле характеристики 0, то воно має безліч різних елементів, множини ж K_{g_1} і K_{g_2} є скінченні (адже число коренів многочлена не перевищує його степеня — див. п. 23.2); тому P_1 має безліч різних елементів. Покажемо, що

$$\underset{\alpha \in P_1}{\forall} \left[\frac{f_1(\alpha)}{g_1(\alpha)} = \frac{f_2(\alpha)}{g_2(\alpha)} \right] \Rightarrow \frac{f_1(x)}{g_1(x)} = \frac{f_2(x)}{g_2(x)}. \tag{9}$$

Внаслідок властивостей поля P

$$\frac{f_1(\alpha)}{g_1(\alpha)} = \frac{f_2(\alpha)}{g_2(\alpha)} \Leftrightarrow f_1(\alpha)\, g_2(\alpha) = g_1(\alpha)\, f_2(\alpha),$$

а за означеннями (2)

$$\frac{f_1(x)}{g_1(x)} = \frac{f_2(x)}{g_2(x)} \Leftrightarrow f_1(x)\, g_2(x) = g_1(x)\, f_2(x).$$

Отже, (9) рівносильне умові

$$\underset{\alpha \in P_1}{\forall} [f_1(\alpha)\, g_2(\alpha) = g_1(\alpha)\, f_2(\alpha)] \Rightarrow f_1(x)\, g_2(x) = g_1(x)\, f_2(x). \tag{10}$$

Але це справді так, бо, відповідно до результатів п. 21,5, два многочлени з $P[x]$ напевно рівні між собою, якщо вони приймають однакові значення у безлічі різних точок поля P (а саме про це свідчить (10), якщо згадати, що в P_1 є безліч різних елементів).

Отже, відповідність $r(x) \to \varphi_r$ взаємно однозначна. Щоб переконатись, що вона ізоморфна, досить перевірити, що

$$\underset{r_1(x), r_2(x) \in P(x)}{\forall} [\varphi_{r_1 + r_2} = \varphi_{r_1} + \varphi_{r_2}]; \tag{11}$$

$$\underset{r_1(x), r_2(x) \in P(x)}{\forall} [\varphi_{r_1 r_2} = \varphi_{r_1} \cdot \varphi_{r_2}]. \tag{12}$$

Встановимо справедливість одного з цих тверджень, скажімо, (12) (справедливість другого перевіряють аналогічно). Нехай

$$r_1(x) = \frac{f_1(x)}{g_1(x)}, \quad r_2(x) = \frac{f_2(x)}{g_2(x)}, \quad (f_1, g_1) = 1, \quad (f_2, g_2) = 1.$$

Тоді

$$r_1(x)\, r_2(x) = \frac{f_1(x)\, f_2(x)}{g_1(x)\, g_2(x)} = \frac{f(x)}{g(x)}, \quad \text{де} \quad (f, g) = 1. \tag{13}$$

Рівність (13) означає, зокрема, що

$$f_1(x)\, f_2(x)\, g(x) = g_1(x)\, g_2(x)\, f(x). \tag{14}$$

З (14) випливає, що для всіх елементів $\alpha \in P_1 = P \setminus (K_{g_1} \cup K_{g_2})$

$$f_1(\alpha)\, f_2(\alpha)\, g(\alpha) = g_1(\alpha)\, g_2(\alpha)\, f(\alpha). \tag{15}$$

Зауважимо, що $g(\alpha) \neq 0$ для $\alpha \in P_1$, бо, як видно з (14), $g(\alpha) = 0 \Rightarrow g_1(\alpha) = 0 \vee g_2(\alpha) = 0$ (адже $f(x)$ взаємно простий з $g(x)$ і тому $f(\alpha) \neq 0$).

За означенням функції $\varphi_{r_1 r_2}$ маємо:

$$\underset{\alpha \in P_1}{\forall} \left[\varphi_{r_1 r_2}(\alpha) = \frac{f(\alpha)}{g(\alpha)} \right]. \tag{16}$$

З другого боку, поділивши рівність (15) на $g(\alpha)\, g_1(\alpha)\, g_2(\alpha)$, дістанемо

$$\forall_{\alpha \in P_1} \left[\varphi_{r_1}(\alpha)\, \varphi_{r_2}(\alpha) = \frac{f_1(\alpha)}{g_1(\alpha)} \cdot \frac{f_2(\alpha)}{g_2(\alpha)} = \frac{f(\alpha)}{g(\alpha)} \right]. \qquad (17)$$

З (16) і (17) випливає, що $\forall_{\alpha \in P_1} [\varphi_{r_1}(\alpha)\, \varphi_{r_2}(\alpha) = \varphi_{r_1 r_2}(\alpha)]$, тобто функціональна рівність $\varphi_{r_1 r_2} = \varphi_{r_1} \cdot \varphi_{r_2}$.

Отже, ми показали, що *поле раціональних дробів $P(x)$ у випадку, коли P має характеристику 0, ізоморфне сукупності дробово-раціональних функцій φ_r.* Це й означає, що алгебраїчний та функціональний погляд на раціональні дроби над полями характеристики 0 цілком рівноправні. Зокрема, вони рівноправні у випадку, коли P — числове поле, що дає змогу застосувати в аналізі усі відомості щодо раціональних дробів, про які ми дізнаємося в курсі алгебри[1].

24.3. Розкладання раціональних дробів на елементарні. Ми розглянемо тут одне питання, яке має важливе значення для математичного аналізу (зокрема для інтегрального числення), але за своїм математичним змістом є суто алгебраїчним і безпосередньо пов'язаним з розкладанням многочлена на незвідні множники.

Означення 1. *Раціональний дріб $\frac{f(x)}{g(x)}$ називається правильним, якщо степінь $f(x)$ менший за степінь $g(x)$. У противному разі дріб називається неправильним.*

Як і раніше, вважатимемо, що раціональні дроби задаються нескоротними частками і що старший коефіцієнт знаменника дорівнює 1. Слід розуміти, що властивість дробу бути правильним не залежить від вибору представника класу рівних часток. Якщо $\frac{f}{g} = \frac{f_1}{g_1}$, то $fg_1 = gf_1$, тобто (при $f \neq 0$, $f_1 \neq 0$)

$$\deg f + \deg g_1 = \deg f_1 + \deg g.$$

Звідси зрозуміло, що

$$\deg f < \deg g \Rightarrow \deg f_1 < \deg g_1.$$

Наприклад, дроби $\frac{x-1}{x^2+1}$, $\frac{1}{x+5}$ правильні, а дроби $\frac{x-1}{x+1}$, $\frac{x^3}{x^2-2x+2}$ неправильні.

Це означення правильного раціонального дробу нагадує відповідне означення правильного дробу в арифметиці, причому, як і раніше, степінь многочлена виконує в ньому роль, яку для чисел виконує абсолютна величина. Однак, якщо в арифметиці сума правильних дробів могла бути неправильним дробом $\left(\text{наприклад, } \frac{1}{2} + \frac{3}{4} = \frac{5}{4}\right)$, то для раціональних дробів справедлива така лема:

[1] Іноді в аналізі функції $\varphi(x) = \frac{x^2-1}{x^3-1}$ та $\psi(x) = \frac{x+1}{x^2+x+1}$ не вважають рівними (оскільки $\varphi(x)$ «втрачає сенс» при $x = 1$). З викладеного тут погляду вони рівні, бо належать одному класу часток, причому друга частка нескоротна і є якраз представником класу. Проте по суті і в аналізі приходять до цього ж висновку, «доозначивши» $\varphi(x)$ в точці $x = 1$ за допомогою значення $\varphi(1) = \psi(1) = \frac{2}{3}$.

Лема 1. *Сума правильних раціональних дробів є правильний раціональний дріб.*

Доведення. Очевидно, що лему досить довести для двох доданків. Нехай

$$\frac{f(x)}{g(x)} = \frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)},$$

причому степінь $f_1(x)$ менший за степінь $g_1(x)$, а степінь $f_2(x)$ менший за степінь $g_2(x)$. Нам треба довести, що степінь $f(x)$ менший за степінь $g(x)$. Маємо:

$$\frac{f(x)}{g(x)} = \frac{f_1(x)\, g_2(x) + f_2(x)\, g_1(x)}{g_1(x)\, g_2(x)}.$$

Досить показати, що дріб, який стоїть у правій частині цієї рівності, правильний, бо якщо чисельник і знаменник цього дробу мають спільний дільник, то внаслідок скорочення їх степені зменшуються на те саме число. Степінь многочлена $f_1(x)\, g_2(x)$ менший за степінь многочлена $g_1(x)\, g_2(x)$, бо степінь $f_1(x)$ менший за степінь $g_1(x)$. Аналогічно, степінь $f_2(x)\, g_1(x)$ менший за степінь $g_1(x)\, g_2(x)$. Отже, і сума многочленів $f_1(x)\, g_2(x) + f_2(x)\, g_1(x)$ має степінь, менший за степінь $g_1(x)\, g_2(x)$, що й треба було довести.

Означення 2. *Елементарним дробом у полі P називається раціональний дріб виду $\frac{f(x)}{[g(x)]^k}$, де $g(x)$ — незвідний многочлен у полі P, $f(x)$ належить $P[x]$ і має степінь менший, ніж $\deg g(x)$, а k — будь-яке натуральне число.*

З цього означення відразу випливає, що всякий елементарний дріб правильний.

Приклади. **1.** Раціональний дріб $\frac{x-1}{x^2+1}$ є елементарним у полі дійсних чисел. Але цей самий дріб не є елементарним у полі комплексних чисел, бо $x^2 + 1$ у цьому полі є звідним многочленом: $x^2 + 1 = (x + i)(x - i)$.

2. Раціональний дріб $\frac{1}{x+5}$ є елементарним у полі раціональних, а також у полях дійсних чисел і комплексних чисел.

3. Раціональний дріб $\frac{x^2+1}{x^3-3}$ є елементарним у полі раціональних чисел, але не є елементарним у полі дійсних чисел.

4. Раціональний дріб $\frac{x^2+1}{(x-1)^3}$ не є елементарним, бо степінь $f(x) = x^2 + 1$ вищий, ніж степінь $g(x) = x - 1$.

5. Дріб $\frac{x^2-1}{x^2+1}$ не є елементарним, бо він неправильний.

Розглянемо питання про можливість зображення довільного раціонального дробу у вигляді суми елементарних дробів у даному полі. Зрозуміло, що неправильний дріб не допускає такого зображення, бо сума елементарних дробів за лемою 1 завжди є правильним дробом. Що ж до правильних дробів, то виявляється, що для них таке зображення завжди можливе.

Лема 2. *Якщо* $g_1(x)$ *і* $g_2(x)$ *— взаємно прості многочлени над полем* P *і* $\frac{f(x)}{g_1(x)\,g_2(x)}$ *— правильний раціональний дріб над цим полем, то в кільці* $P[x]$ *завжди можна знайти такі многочлени* $f_1(x)$ *і* $f_2(x)$, *що*

$$\frac{f(x)}{g_1(x)\,g_2(x)} = \frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)},$$

причому дроби у правій частині цієї рівності правильні.

Доведення. За наслідком з теореми 5, п. 22.5, можна знайти в $P[x]$ такі многочлени $u(x)$ і $v(x)$, що $g_1(x)\,u(x) + g_2(x)\,v(x) = 1$, тому

$$f(x)\,g_1(x)\,u(x) + f(x)\,g_2(x)\,v(x) = f(x). \tag{18}$$

Поділимо $f(x)\,u(x)$ на $g_2(x)$ з остачею: $f(x)\,u(x) = g_2(x)\,u_1(x) + f_2(x)$, причому степінь $f_2(x)$ нижчий за степінь $g_2(x)$. Підставляючи цей вираз у (18) і групуючи члени, маємо: $f_2(x)\,g_1(x) + g_2(x)\,[u_1(x)\,g_1(x) + f(x)\,v(x)] = f(x)$, або, вводячи означення $u_1(x)\,g_1(x) + f(x)\,v(x) = f_1(x)$, $f_2(x)\,g_1(x) + f_1(x)\,g_2(x) = f(x)$. Нарешті, поділивши цю рівність на $g_1(x)\,g_2(x)$, дістанемо:

$$\frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)} = \frac{f(x)}{g_1(x)\,g_2(x)}.$$

Дріб $\frac{f_2(x)}{g_2(x)}$ правильний, бо степінь $f_2(x)$, як уже було зазначено, менший, ніж степінь $g_2(x)$. Дріб $\frac{f_1(x)}{g_1(x)}$ також правильний як сума двох правильних дробів (лема 1):

$$\frac{f_1(x)}{g_1(x)} = \frac{f(x)}{g_1(x)\,g_2(x)} + \frac{-f_2(x)}{g_2(x)}.$$

Лему 2 доведено.

Наслідок. *Якщо* $g_1(x)$, $g_2(x)$, ..., $g_m(x)$ *— попарно взаємно прості многочлени над полем* P *і* $\frac{f(x)}{g_1(x)\,g_2(x)\,\dots\,g_m(x)}$ *— правильний раціональний дріб над цим полем, то в кільці* $P[x]$ *завжди можна знайти такі многочлени* $f_1(x)$, $f_2(x)$, ..., $f_m(x)$, *що*

$$\frac{f(x)}{g_1(x)\,g_2(x)\,\dots\,g_m(x)} = \frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)} + \dots + \frac{f_m(x)}{g_m(x)}, \tag{19}$$

причому всі дроби у правій частині рівності (19) *правильні.*

Доведення. Скористаємося методом індукції за m. При $m = 1$ рівність (19) очевидна. Припустимо тепер, що розклад (19) можливий для $m - 1$ попарно взаємно простих многочленів $g_1(x)$, $g_2(x)$, ..., $g_{m-1}(x)$. Якщо $g_m(x)$ — многочлен, взаємно простий з кожним з $g_1(x)$, $g_2(x)$, ..., $g_{m-1}(x)$, то він взаємно простий і з їх добутком (п. 22.5), тому за лемою 2 маємо:

$$\frac{f(x)}{g_1(x)\,g_2(x)\,\dots\,g_{m-1}(x)\,g_m(x)} = \frac{s(x)}{g_1(x)\,g_2(x)\,\dots\,g_{m-1}(x)} + \frac{f_m(x)}{g_m(x)}.$$

За припущенням індукції для першого з дробів, що стоять у правій частині цієї рівності, розклад (19) можливий. Тому дістаємо остаточно:

$$\frac{f(x)}{g_1(x)\,\dots\,g_{m-1}(x)\,g_m(x)} = \frac{f_1(x)}{g_1(x)} + \dots + \frac{f_{m-1}(x)}{g_{m-1}(x)} + \frac{f_m(x)}{g_m(x)}.$$

Наслідок доведено.

Лема 3. *Всякий правильний дріб над полем* P *виду* $\frac{f(x)}{[g(x)]^k}$, *де* $g(x)$ *— многочлен, незвідний у полі* P, *а* k *— довільне натуральне число, можна подати як суму двох правильних дробів над полем* P *виду*

$$\frac{f(x)}{[g(x)]^k} = \frac{f_1(x)}{[g(x)]^k} + \frac{s_1(x)}{[g(x)]^{k-1}}, \tag{20}$$

з яких перший є елементарним у полі P.

Доведення. Поділимо многочлен $f(x)$ на $g(x)$ з остачею:

$$f(x) = g(x)\,s_1(x) + f_1(x),$$

де степінь $f_1(x)$ менший за степінь $g(x)$. Отже,

$$\frac{f(x)}{[g(x)]^k} = \frac{g(x)\,s_1(x) + f_1(x)}{[g(x)]^k} = \frac{f_1(x)}{[g(x)]^k} + \frac{s_1(x)}{[g(x)]^{k-1}}.$$

Перший з цих дробів є елементарним у полі P, бо $g(x)$ незвідний у цьому полі многочлен, степінь якого більший за степінь чисельника $f_1(x)$. Другий дріб є правильним за лемою 1 як сума двох правильних дробів:

$$\frac{s_1(x)}{[g(x)]^{k-1}} = \frac{f(x)}{[g(x)]^k} + \frac{-f_1(x)}{[g(x)]^k}.$$

Лему доведено.

Наслідок. *Кожний правильний дріб над полем* P *виду* $\frac{f(x)}{[g(x)]^k}$, *де* $g(x)$ *— многочлен, незвідний у полі* P, *а* k *— довільне натуральне число, можна подати як суму елементарних дробів у цьому полі:*

$$\frac{f(x)}{[g(x)]^k} = \frac{f_1(x)}{g(x)} + \frac{f_2(x)}{[g(x)]^2} + \dots + \frac{f_k(x)}{[g(x)]^k}. \tag{21}$$

Для доведення досить застосувати метод індукції по k.

Тепер уже легко довести основну теорему.

Теорема 1. *Усякий правильний дріб над полем* P *можна подати як суму елементарних дробів у цьому полі.*

Доведення. Нехай дано правильний дріб $\frac{f(x)}{g(x)}$. Якщо многочлен $g(x)$ незвідний у полі P, то цей дріб уже є елементарним. Якщо ж $g(x)$ звідний, то розкладемо його на незвідні множники у полі P, що завжди можливо (наслідок з теореми 8 § 22):

$$g(x) = [g_1(x)]^{k_1} \dots [g_m(x)]^{k_m}. \tag{22}$$

Многочлени $[g_j(x)]^{k_j}$ ($j = 1, 2, \dots, m$) попарно взаємно прості, бо не мають однакових незвідних множників. Тому за наслідком з леми 2

матимемо такий розклад:

$$\frac{f(x)}{g(x)} = \frac{f(x)}{[g_1(x)]^{k_1} \dots [g_m(x)]^{k_m}} =$$

$$= \frac{f_1(x)}{[g_1(x)]^{k_1}} + \frac{f_2(x)}{[g_2(x)]^{k_2}} + \dots + \frac{f_m(x)}{[g_m(x)]^{k_m}}, \qquad (23)$$

причому всі дроби в правій частині (23) правильні.

За наслідком з леми 3 правильний дріб $\frac{f_1(x)}{[g_1(x)]^{k_1}}$ можна подати як суму елементарних дробів за формулою (21):

$$\frac{f_1(x)}{[g_1(x)]^{k_1}} = \frac{f_{11}(x)}{g_1(x)} + \frac{f_{12}(x)}{[g_1(x)]^2} + \dots + \frac{f_{1k_1}(x)}{[g_1(x)]^{k_1}}.$$

Аналогічно

$$\frac{f_2(x)}{[g_2(x)]^{k_2}} = \frac{f_{21}(x)}{g_2(x)} + \frac{f_{22}(x)}{[g_2(x)]^2} + \dots + \frac{f_{2k_2}(x)}{[g_2(x)]^{k_2}},$$

. . . . . .

$$\frac{f_m(x)}{[g_m(x)]^{k_m}} = \frac{f_{m1}(x)}{g_m(x)} + \frac{f_{m2}(x)}{[g_m(x)]^2} + \dots + \frac{f_{mk_m}(x)}{[g_m(x)]^{k_m}}. \qquad (24)$$

Підставляючи (24) у формулу (23), дістанемо шукане зображення дробу $\frac{f(x)}{g(x)}$ у вигляді суми елементарних дробів у полі P:

$$\frac{f(x)}{g(x)} = \frac{f(x)}{[g_1(x)]^{k_1} \dots [g_m(x)]^{k_m}} = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \frac{f_{ij}(x)}{[g_i(x)]^j}. \qquad (25)$$

Теорему доведено.

Для практичних потреб важливо знати, що, незалежно від способу, який застосовано для розкладання дробу на елементарні дроби, результат завжди буде той самий.

Теорема 2. *Розклад правильного раціонального дробу на елементарні дроби у даному полі єдиний.*

Доведення. Припустимо супротивне, тобто що можливі два різні розклади дробу $\frac{f(x)}{g(x)}$ на суму елементарних дробів. Це означає, що в цих двох розкладах не всі доданки будуть однакові (у зв'язку з громіздкістю розкладу дробу на елементарні ми цих двох розкладів не виписуватимемо). Віднімаючи ці розклади один від одного, в одній частині рівності дістанемо нуль, а в другій — алгебраїчну суму елементарних дробів. Якщо серед цих дробів є дроби з однаковими знаменниками, зведемо їх. Зрозуміло, що всі дроби при такому зведенні знищитись не можуть, бо тоді б усі доданки цих розкладів були б попарно рівні між собою, тобто розклади не були б різними. Зауважимо також, що сума двох елементарних дробів з однаковими знаменниками є знову елементарний дріб.

Нехай після такого зведення елементарні дроби, що залишаються,

мають знаменники: $g_1(x)$, $[g_1(x)]^2$, ..., $[g_1(x)]^{n_1}$; $g_2(x)$, $[g_2(x)]^2$, ..., $[g_2(x)]^{n_2}$; ..., $g_k(x)$, $[g_k(x)]^2$, ..., $[g_k(x)]^{n_k}$, де $g_i(x)$ — різні незвідні многочлени у даному полі.

Помножимо знайдену рівність на многочлен

$$[g_1(x)]^{n_1-1} \cdot [g_2(x)]^{n_2} \dots [g_k(x)]^{n_k}.$$

Тоді всі доданки суми, крім одного, що мав знаменник $[g_1(x)]^{n_1}$, стануть многочленами, тобто дістанемо

$$\frac{p(x)[g_1(x)]^{n_1-1} \cdot [g_2(x)]^{n_2} \dots [g_k(x)]^{n_k}}{[g_1(x)]^{n_1}} + q(x) = 0, \qquad (26)$$

де $q(x)$ — деякий многочлен, а $p(x)$ — чисельник елементарного дробу із знаменником $[g_1(x)]^{n_1}$; отже, $(p, g_1) = 1$. Після скорочення на $[g_1(x)]^{n_1-1}$ і очевидних перетворень рівність (26) можна записати в такому вигляді:

$$p(x)[g_2(x)]^{n_2} \dots [g_k(x)]^{n_k} = -q(x)g_1(x). \qquad (27)$$

Але рівність (27) показує, що ми прийшли до суперечності. Справді, права частина цієї рівності ділиться на незвідний многочлен $g_1(x)$, тоді як ліва частина взаємно проста з цим многочленом, бо він взаємно простий з многочленами $p(x)$, $g_2(x)$, ..., $g_k(x)$, а отже, і з їх добутком. Теорему доведено.

Теореми 1 і 2 є теоретичною основою методу невизначених коефіцієнтів, який на практиці застосовують для розкладання дробово-раціональної функції на елементарні дроби. Цей метод полягає в тому, що, знаючи розклад (22) знаменника $g(x)$ на незвідні множники, записують розклад (25) на можливі елементарні дроби. Знаменники цих дробів відомі; що ж до чисельників $f_{ij}(x)$, то їх записують як многочлени з невизначеними коефіцієнтами, степінь яких на одиницю менший за степінь відповідного многочлена $g_i(x)$ (інакше цей дріб не був би елементарним). Після цього обидві частини рівності множать на спільний знаменник і прирівнюють коефіцієнти при однакових степенях x. Відповідно до теореми 2 коефіцієнти при цьому завжди повинні визначатися однозначно.

Приклад 6. Розкладемо $\frac{x}{(x+1)^2(x^2+1)}$ на елементарні дроби у полі дійсних чисел. Тут $g_1(x) = x+1$, $k_1 = 2$; $g_2(x) = x^2+1$, $k_2 = 1$.

Відповідно до формули (25) маємо:

$$\frac{x}{(x+1)^2(x^2+1)} = \frac{A}{x+1} + \frac{B}{(x+1)^2} + \frac{Cx+D}{x^2+1}.$$

Звідси дістаємо:

$$x = (A+C)x^3 + (A+B+2C+D)x^2 + (A+C+2D)x + (A+B+D).$$

Очевидно, що $A + C = 0$; $A + B + 2C + D = 0$; $A + C + 2D = 1$, $A + B + D = 0$. Розв'язуючи цю систему, дістаємо: $A = 0$, $B = -\frac{1}{2}$, $C = 0$, $D = \frac{1}{2}$, тобто

$$\frac{x}{(x+1)^2(x^2+1)} = -\frac{1}{2(x+1)^2} + \frac{1}{2(x^2+1)}.$$

Формулу (25) фактично можна застосувати лише тоді, коли відомий розклад (22) знаменника $g(x)$ на незвідні множники. Проте в нашому розпорядженні поки що немає ніяких загальних методів для такого розкладання. Отже, по суті ми лише звели задачу розкладання правильного дробу на елементарні дроби до задачі розкладання многочлена на незвідні множники.

Розкласти на елементарні дроби можна лише правильний раціональний дріб. Що ж до неправильного дробу, то справедливе таке твердження.

Теорема 3. *Всякий неправильний дріб* $\frac{f(x)}{g(x)}$ *можна подати як суму многочлена і правильного дробу.*

Доведення. Поділимо $f(x)$ на $g(x)$ з остачею: $f(x) = g(x)\,s(x) + f_1(x)$, де степінь $f_1(x)$ нижчий, ніж степінь $g(x)$. Ділячи цю рівність на $g(x)$, дістанемо:

$$\frac{f(x)}{g(x)} = s(x) + \frac{f_1(x)}{g(x)}, \tag{28}$$

де $s(x)$ — многочлен, а $\frac{f_1(x)}{g(x)}$ правильний дріб. Теорему доведено.

Многочлен $s(x)$ називається цілою частиною дробу $\frac{f(x)}{g(x)}$, а зображення неправильного дробу у вигляді (28) — *відокремленням цілої частини.*

Наприклад, дріб $\frac{x^3}{x^2-2x+2}$ можна подати у вигляді:

$$\frac{x^3}{x^2-2x+2} = (x+2) + \frac{2x-4}{x^2-2x+2}.$$

Отже, довільний раціональний дріб над полем P можна подати як суму деякого многочлена (який може бути і 0) і елементарних дробів у цьому полі.

Розділ VI

МНОГОЧЛЕНИ ВІД КІЛЬКОХ ЗМІННИХ

§ 25. КІЛЬЦЕ МНОГОЧЛЕНІВ ВІД КІЛЬКОХ ЗМІННИХ

25.1. Побудова кільця многочленів. У попередньому розділі ми досить докладно розглянули поняття многочлена і властивості многочленів від одної змінної. Природним і важливим узагальненням цього поняття є многочлени від кількох змінних. Вони й будуть предметом вивчення в цьому розділі; при цьому значною мірою ми будемо спиратися на теорію многочленів від однієї змінної.

Щоб полегшити розуміння загального означення кільця многочленів від n змінних, розглянемо спочатку окремий випадок, а саме — сукупність многочленів від 2-х змінних.

Нехай R — якась область цілісності з одиницею, а $R[x]$ — сукупність усіх многочленів від однієї змінної x над R. Як відомо (теорема 1, § 21), $R[x]$ є також областю цілісності з одиницею. Отже, ми можемо побудувати кільце многочленів над $R[x]$ від однієї змінної (позначимо цю змінну через y), тобто сукупність усіх многочленів виду

$$a_n(x)\,y^n + a_{n-1}(x)\,y^{n-1} + \cdots + a_1(x)\,y + a_0(x), \tag{1}$$

коефіцієнти яких $a_n(x)$, $a_{n-1}(x)$, ..., $a_1(x)$, $a_0(x)$ є елементи області цілісності $R[x]$ (многочлени від змінної x над кільцем R). Природно позначити сукупність многочленів виду (1) через $R[x][y]$.

Згідно з теоремою 1 § 21, *кільце* $R[x][y]$ *є областю цілісності*. З'ясуємо, яку будову мають її елементи. Для цього запишемо коефіцієнти $a_j(x) \in R[x]$ $(j = 0, 1, \ldots, n)$ у канонічній формі:

$$\begin{aligned}
a_0(x) &= a_{m_0 0}x^{m_0} + a_{m_0-1,0}x^{m_0-1} + \cdots + a_{10}x + a_{00},\\
a_1(x) &= a_{m_1 1}x^{m_1} + a_{m_1-1,1}x^{m_1-1} + \cdots + a_{11}x + a_{01},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
a_{n-1}(x) &= a_{m_{n-1}n-1}x^{m_{n-1}} + a_{m_{n-1}-1,n-1}x^{m_{n-1}-1} + \cdots + a_{1,n-1}x + a_{0,n-1},\\
a_n(x) &= a_{m_n,n}x^{m_n} + a_{m_n-1,n}x^{m_n-1} + \cdots + a_{1,n}x + a_{0n}.
\end{aligned} \tag{2}$$

Підставляючи вирази (2) в (1) і користуючись властивостями дій в кільці $R[x][y]$ (з урахуванням того, що $R[x]$ — нідкільце $R[x][y]$), дістанемо вираз

$$\sum_{j=0}^{n}\sum_{i=0}^{m_j} a_{ij}x^i y^j, \quad a_{ij} \in R,$$

тобто якусь скінченну суму членів виду $a_{\mu\nu}x^\mu y^\nu$, де $a_{\mu\nu} \in R$, а μ і ν — якісь невід'ємні цілі числа. Навпаки, будь-яку скінченну суму виду

$$\sum_{\mu,\nu} a_{\mu\nu}x^\mu y^\nu, \quad a_{\mu\nu} \in R, \quad \mu, \nu \in \mathbf{Z}_+ \tag{3}$$

можна розглядати як елемент кільця $R[x][y]$. Адже кожний член виду $a_{\mu\nu}x^\mu y^\nu$ є многочлен від змінної y з коефіцієнтом $a_{\mu\nu}x^\mu \in R[x]$, тобто належить кільцю $R[x][y]$. Але тоді й сума (3), в якій додавання розуміємо як дію в кільці $R[x][y]$ є елементом кільця $R[x][y]$. Отже, $R[x][y]$ — сукупність усіх можливих сум виду (3).

Вважатимемо, що записи $a_{\mu\nu}x^\mu y^\nu$ і $a_{\mu\nu}y^\nu x^\mu$ означають той самий член. Таке ототожнення можна зробити тому, що воно не впливає на означення суми і добутку многочленів. Але тоді кільце $R[x][y]$ ототожнюється з кільцем $R[y][x]$. Надалі для цього кільця вживатимемо позначення $R[x, y]$ або $R[y, x]$. Елементи $R[x, y]$ називатимемо *многочленами від 2-х змінних* x *і* y *над* R — позначатимемо $f(x, y)$, $g(x, y)$ тощо.

Приклад 1. Подамо суму

$$f(x, y) = 3x^2y - 5xy^3 + x^7 - 1 + x^2y^3 + 2y - 7xy^4,$$

як многочлен від y над $Q[x]$. Маємо:

$$f(x, y) = -7xy^4 + (x^2 - 5x)\, y^3 + (3x^2 + 2)\, y + (x^7 - 1).$$

Перейдемо тепер до загальних означень.

Означення 1. *Кільцем многочленів $R[x_1, x_2, \ldots, x_{n-1}, x_n]$ від n змінних $x_1, x_2, \ldots, x_{n-1}, x_n$ над областю цілісності R називається кільце многочленів від змінної x_n над кільцем $R[x_1, x_2, \ldots, x_{n-1}]$, тобто*

$$R[x_1, x_2, \ldots, x_{n-1}, x_n] = R[x_1, x_2, \ldots, x_{n-1}][x_n]. \qquad (4)$$

Це означення має індуктивний характер. При $n = 1$ воно зводиться до означення кільця многочленів від однієї змінної x_1 над областю цілісності R (природно вважати, що при $n = 1$ $R[x_1, \ldots, x_{n-1}] = R$). Якщо ж уже означено кільце $R[x_1, \ldots, x_{n-1}]$ при $n \geqslant 1$, то за допомогою (4) дістаємо означення кільця $R[x_1, x_2, \ldots, x_{n-1}, x_n]$. Отже, для довільного натурального n означено кільце многочленів від n змінних $x_1, x_2, \ldots, x_n$.

Теорема 1. *Кільце многочленів $R[x_1, x_2, \ldots, x_{n-1}, x_n]$ над областю цілісності R є область цілісності.*

Доведення. Твердження правильне при $n = 1$ (теорема 1, § 21). Припустимо, що воно правильне при $n = m$ і розглянемо кільце $R[x_1, \ldots, x_m, x_{m+1}]$. Згідно з означенням 1, $R[x_1, \ldots, x_m, x_{m+1}]$ є кільце многочленів над $R_m \underset{df}{=} R[x_1, \ldots, x_m]$. За припущенням індукції, R_m є область цілісності. Отже, знову, застосовуючи теорему 1 § 21, дістаємо, що і $R_m[x_{m+1}] = R[x_1, \ldots, x_m, x_{m+1}]$ є область цілісності. За принципом індукції, $R[x_1, \ldots, x_n]$ є область цілісності при довільному натуральному n. Теорему доведено.

Зрозуміло, що коли R — область цілісності з одиницею, то й $R[x_1, \ldots, x_n]$ — область цілісності з одиницею.

Наступна теорема встановлює будову елементів області цілісності $R[x_1, x_2, \ldots, x_n]$.

Теорема 2. *Кожний елемент $f \in R[x_1, \ldots, x_n]$ можна подати у вигляді скінченної суми*

$$f = \sum_{i=1}^{N} A_i x_1^{k_{1i}} x_2^{k_{2i}} \ldots x_n^{k_{ni}}, \quad A_i \in R, \quad k_{ji} \in \mathbf{Z}_+. \qquad (5)$$

$(i = 1, 2, \ldots, N;\ j = 1, 2, \ldots, n)$.

Навпаки, будь-який вираз виду (5) *є елементом кільця $R[x_1, x_2, \ldots, x_n]$.*

Доведення проведемо індукцією по n. При $n = 1$ твердження правильне. Припустимо, що воно правильне при $n = m$ і перевіримо його правильність при $n = m + 1$. За означенням 1, кожний елемент $f \in R[x_1, \ldots, x_m, x_{m+1}]$ є многочлен від x_{m+1} над областю цілісності $R[x_1, \ldots, x_m]$, і тому його можна подати у вигляді суми

$$f = \sum_{j=1}^{l} a_j(x_1, \ldots, x_m)\, x_{m+1}^j, \quad a_j(x_1, \ldots, x_m) \in R[x_1, \ldots, x_m] \qquad (6)$$

$$(j = 0, 1, \ldots, l).$$

За припущенням індукції, кожний многочлен $a_j(x_1, \ldots, x_m)$ від m змінних можна подати у вигляді скінченної суми

$$a_j(x_1, \ldots, x_m) = \sum_{i=1}^{N_j} A_i^{(j)} x_1^{k_{1i}^{(j)}} x_2^{k_{2i}^{(j)}} \ldots x_m^{k_{mi}^{(j)}}, \quad A_i^{(j)} \in R, \quad k_{si}^{(j)} \in \mathbf{Z}_+ \qquad (7)$$

$(i = 1, 2, \ldots, N_j;\ s = 1, 2, \ldots, m;\ j = 0, 1, 2, \ldots, l)$.

Підставивши вираз (7) в (6) і виконавши відповідні дії (в розумінні дій у кільці $R[x_1, \ldots, x_m, x_{m+1}]$ з урахуванням того, що воно містить $R[x_1, \ldots, x_m]$ як підкільце), дістанемо скінченну суму виду

$$f = \sum_{r=1}^{N} B_r x_1^{k_{1r}} x_2^{k_{2r}} \ldots x_m^{k_{mr}} x_{m+1}^{k_{m+1,\,r}}, \qquad (8)$$

де $B_r \in R$ $(r = 1, \ldots, N)$, бо кожне B_r є якесь з $A_i^{(j)}$.

Навпаки, кожна сума виду (8) є елемент кільця $R[x_1, x_2, \ldots, x_{m+1}]$: адже будь-який її доданок $B_r x^{k_{1r}} \ldots x_m^{k_{mr}} x_{m+1}^{k_{m+1,\,r}}$ може розглядатись як многочлен від x_{m+1} з коефіцієнтом $B_r x_1^{k_{1r}} \ldots x_m^{k_{mr}} \in R[x_1, \ldots, x_m]$, а тому й уся сума належить кільцю $R[x_1, \ldots, x_m, x_{m+1}]$.

Отже, твердження теореми правильне і при $n = m + 1$, тобто за принципом індукції теорему доведено.

Означення 2. *Кожний елемент кільця $R[x_1, \ldots, x_n]$ називають многочленом від n змінних $x_1, x_2, \ldots, x_n$ над R і позначають $f(x_1, \ldots, x_n)$, $g(x_1, \ldots, x_n)$ і т. п.*

Згідно з теоремою 2, будь-який многочлен з $R[x_1, \ldots, x_n]$ можна подати у формі суми (5)

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{N} A_i x_1^{k_{1i}} \ldots x_n^{k_{ni}}, \quad A_i \in R, \quad k_{ji} \in \mathbf{Z}_+. \qquad (9)$$

Кожний доданок $A_i x_1^{k_{1i}} \ldots x_n^{k_{ni}}$ цієї суми називають членом многочлена $f(x_1, \ldots, x_n)$, відповідний елемент $A_i \in R$ — коефіцієнтом члена (і многочлена). Два члени, які відрізняються лише коефіцієнтами, називають подібними; іншими словами, члени подібні, якщо усі змінні входять множниками в ці члени у попарно рівних степенях, наприклад $Ax_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ та $Bx_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$. При цьому порядок, в якому записано множники $x_j^{k_j}$, неістотний, тобто члени $Ax_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$, $Ax_2^{k_2} x_1^{k_1} \ldots x_n^{k_n}$, $Ax_2^{k_2} x_3^{k_3} \ldots x_1^{k_1}$ тощо вважаємо однаковими, рівними між собою. Відповідно до цього, $R[x_1, x_2, \ldots, x_n]$, $R[x_2, x_1, \ldots, x_n]$, $R[x_2, \ldots, x_n, x_1]$ і т. п. є різними формами запису того самого кільця многочленів від змінних $x_1, x_2, \ldots, x_n$ над областю цілісності R.

Виконання над многочленами з кільця $R[x_1, x_2, \ldots, x_n]$ дій додавання і множення зводиться внаслідок дистрибутивного закону, що діє в кільці, до дій над членами цих многочленів. При додаванні двох (або більшого числа) подібних членів дістаємо один член, подібний до кожного з даних:

$$Ax_1^{k_1} x_2^{k_2} \ldots x_n^{k_n} + Bx_1^{k_1} x_2^{k_2} \ldots x_n^{k_n} = (A + B)\, x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}.$$

Зрозуміло, що $A \in R \wedge B \in R \Rightarrow (A + B) \in R$. Таку заміну кількох подібних членів одним називають *зведенням подібних членів*. Множення членів многочлена (на основі властивостей кільця $R[x_1, \ldots, x_n]$) здійснюють за правилом

$$(Ax_1^{k_1}x_2^{k_2} \ldots x_n^{k_n})(Bx_1^{l_1}x_2^{l_2} \ldots x_n^{l_n}) = ABx_1^{k_1+l_1}x_2^{k_2+l_2} \ldots x_n^{k_n+l_n},$$

причому $A \in R \wedge B \in R \Rightarrow AB \in R$. Як бачимо, результати дій додавання і множення не залежать від порядку запису змінних. Саме це і дає змогу ототожнювати кільця $R[x_{i_1}x_{i_2} \ldots x_{i_n}]$ при довільній перестановці $i_1, i_2, \ldots, i_n$ індексів $1, 2, \ldots, n$.

25.2. Різні форми зображення многочленів. При записі многочлена $f(x_1, x_2, \ldots, x_n)$ у формі (9) ми завжди вважатимемо, що серед членів многочлена немає подібних (тобто здійснено попереднє зведення подібних членів). Таку форму многочлена називають канонічною або нормальною.

Канонічна форма має ту позитивну особливість, що вона єдина (з точністю до порядку членів). Цю єдиність слід розуміти так, що коли два многочлени $f(x_1, x_2, \ldots, x_n)$ і $g(x_1, x_2, \ldots, x_n)$ подані у канонічній формі, рівні між собою, то кожний член многочлена $f(x_1, \ldots, x_n)$ є також членом многочлена $g(x_1, \ldots, x_n)$ і навпаки.

Теорема 3. *Будь-який многочлен* $f(x_1, \ldots, x_n) \in R[x_1, \ldots, x_n]$ *можна подати в канонічній формі лише одним способом* (з точністю до порядку членів).

Доведення. Розглянемо спочатку окремий випадок нуль-многочлена $\theta(x_1, \ldots, x_n)$ (тобто нуля кільця $R[x_1, \ldots, x_n]$) і покажемо, що його єдиною канонічною формою є $0 \in R$ (тобто, що всі його коефіцієнти дорівнюють нулю кільця R). Міркуватимемо за індукцією. При $n = 1$ ця властивість відома (§ 21). Припустимо, що твердження правильне при $n = m$. Подамо $\theta(x_1, \ldots, x_m, x_{m+1}) \in R[x_1, \ldots, x_m, x_{m+1}]$ у канонічній формі многочлена від однієї змінної x_{m+1} над кільцем $R[x_1, \ldots, x_m]$:

$$\theta(x_1, \ldots, x_m, x_{m+1}) = \sum_{i=1}^{N} a_i(x_1, \ldots, x_m)\, x_{m+1}^i.$$

За властивістю многочленів від однієї змінної звідси випливає, що всі коефіцієнти многочлена є нулями того кільця, над яким многочлен задано, тобто:

$$\forall_i [a_i(x_1, \ldots, x_m) = \theta(x_1, \ldots, x_m)].$$

За припущенням індукції, робимо висновок, що коефіцієнти усіх многочленів $a_i(x_1, \ldots, x_m)$ $(i = 0, 1, \ldots, N)$ є нулі кільця R, але тоді зрозуміло, що й усі коефіцієнти многочлена $\theta(x_1, \ldots, x_m, x_{m+1})$ дорівнюють 0 (нулю кільця R).

Перейдемо тепер до загального випадку. Якщо многочлени $f(x_1, \ldots, x_n)$ і $g(x_1, \ldots, x_n)$ подано в канонічній формі і $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_n)$, то $f(x_1, \ldots, x_n) - g(x_1, \ldots, x_n) = \theta(x_1, \ldots, x_n)$. Це означає, що всі коефіцієнти многочлена $f(x_1, \ldots, x_n) - g(x_1, \ldots, x_n)$ дорівнюють 0. Оскільки подібних членів у записі кожного з цих многочленів немає, а нуль може утворитись лише при відніманні подібних членів з однаковими коефіцієнтами, то робимо висновок, що кожний член многочлена $f(x_1, \ldots, x_n)$ є також членом многочлена $g(x_1, \ldots, x_n)$ і, навпаки, кожний член $g(x_1, \ldots, x_n)$ є членом $f(x_1, \ldots, x_n)$. Цим теорему доведено.

На підставі доведеного у дальшому викладі нуль-многочлен кільця $R[x_1, \ldots, x_n]$ позначатимемо просто 0.

Домовимося також про таку термінологію.

Означення. *Степенем члена* $Ax_1^{k_1}x_2^{k_2} \ldots x_n^{k_n}$ *многочлена називається сума* $k_1 + k_2 + \ldots + k_n$. *Число* k_i $(i = 1, 2, \ldots, n)$ *називають степенем даного члена відносно* x_i. *Найбільший із степенів членів називається степенем многочлена, а член з найбільшим степенем називається старшим членом многочлена.*

Цілком зрозуміло, що різних старших членів многочлена може бути не один. Наприклад, у многочлені

$$f(x_1, x_2, x_3) = 2x_1x_2^3x_3^2 + x_1^4x_2x_3 + 2x_1^2x_2^2x_3 - x_1^2x_2^3 - 2x_1^3 + 5x_1x_3^3 + 6 + x_1^2x_2x_3^2 \qquad (10)$$

два члени $2x_1x_2^3x_3^2$ і $x_1^4x_2x_3$ є старшими, бо мають однаковий найбільший степінь 6.

Очевидно також, що єдиним многочленом з $R[x_1, \ldots, x_n]$, до якого поняття степеня незастосовне, є нуль-многочлен $\theta(x_1, \ldots, x_n)$. Многочлени нульового степеня є відмінні від 0 елементи кільця R. Їх разом з елементами 0 називатимемо константами. Степінь многочлена $f(x_1, \ldots, x_n)$ позначатимемо $\deg f$ як і у випадку многочленів від однієї змінної.

Якщо всі члени многочлена мають той самий степінь l, то многочлен називається однорідним многочленом або формою степеня l. Очевидно, що *будь-який многочлен можна подати як суму скінченного числа однорідних многочленів різних степенів.*

У кільці $R[x_1, \ldots, x_n]$, як і в кільці $R[x]$, очевидно, *степінь суми двох многочленів не може перевищувати степінь кожного з них;* справедливе також таке твердження:

Теорема 4. *Якщо* $f(x_1, \ldots, x_n)$, $g(x_1, \ldots, x_n)$ *— відмінні від нуля многочлени з* $R[x_1, \ldots, x_n]$, *де* R *— область цілісності, то*

$$\deg(fg) = \deg f + \deg g. \qquad (11)$$

Доведення: а) Розглянемо спочатку випадок, коли дані многочлени однорідні. Нехай

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{N} A_i x_i^{k_{1i}} \ldots x_n^{k_{ni}} \qquad (12)$$

форма l-го степеня, тобто $\forall_i [k_{1i} + k_{2i} + \ldots + k_{ni} = l]$, а

$$g(x_1, \ldots, x_n) = \sum_{j=1}^{M} B_j x_1^{k_{1j}} \ldots x_n^{k_{nj}} \qquad (13)$$

— форма m-го степеня, тобто $\forall_j [k_{1j} + k_{2j} + \ldots + k_{nj} = m]$.

При цьому, як завжди, вважатимемо зображення (12) і (13) канонічними (тобто в кожному з них немає подібних членів).

Кожний член добутку $f(x_1, \ldots, x_n)\, g(x_1, \ldots, x_n)$ матиме вигляд:

$$C_{ij} x_1^{k_{1i}+k_{1j}} \ldots x_n^{k_{ni}+k_{nj}},$$

а його степінь дорівнює $l+m$ бо $\forall\limits_{i,j} [(k_{1i}+k_{1j}) + \ldots + (k_{ni}+k_{nj}) = l+m]$. При цьому цей добуток напевно відмінний від нуля, бо $R[x_1, \ldots, x_n]$ є областю цілісності (теорема 1) і $f \neq 0 \wedge g \neq 0 \Rightarrow fg \neq 0$. Отже, добуток $f(x_1, \ldots, x_n)\, g(x_1, \ldots, x_n)$ є також однорідний многочлен степеня $l+m = \deg f + \deg g$. Цим (11) доведено.

б) Якщо тепер $f(x_1, \ldots, x_n)$ і $g(x_1, \ldots, x_n)$ — довільні многочлени, відмінні від 0, то подамо їх як суму форм різних степенів:

$$f(x_1, \ldots, x_n) = \varphi_{l_1}(x_1, \ldots, x_n) + \varphi_{l_2}(x_1, \ldots, x_n) + \cdots + \varphi_{l_s}(x_1, \ldots, x_n),$$

$$g(x_1, \ldots, x_n) = \psi_{m_1}(x_1, \ldots, x_n) + \psi_{m_2}(x_1, \ldots, x_n) + \cdots + \psi_{m_r}(x_1, \ldots, x_n).$$

Тут φ_{l_i} та ψ_{m_j} позначають однорідні многочлени степеня l_i та m_j відповідно. Очевидно, $\max\{l_1, l_2, \ldots, l_s\} = \deg f$, $\max\{m_1, m_2, \ldots, m_r\} = \deg g$; нехай саме $l_1 = \deg f$, $m_1 = \deg g$, отже, $l_i < l_1$ $(i = 2, 3, \ldots, s)$, $m_j < m_1$ $(j = 2, 3, \ldots, r)$.

Добуток даних многочленів тоді можна подати так:

$$f(x_1, \ldots, x_n)\, g(x_1, \ldots, x_n) = \varphi_{l_1}(x_1, \ldots, x_n)\, \psi_{m_1}(x_1, \ldots, x_n) + \sum_{i,j\ (i\neq 1 \vee j\neq 1)} \varphi_{l_i}(x_1, \ldots, x_n)\, \psi_{m_j}(x_1, \ldots, x_n), \qquad (14)$$

де символ $\sum\limits_{i,j\ (i\neq 1 \vee j\neq 1)}$ означає суму по всіх парах значень i, j $(i = 1, 2, \ldots, s;\ j = 1, 2, \ldots, r)$, крім пари $i = 1$, $j = 1$ (оскільки відповідний член суми записано окремо).

Зрозуміло, що перший член суми (14) є форма степеня $l_1 + m_1 = \deg f + \deg g$. Решта ж членів цієї суми є формами степеня $l_i + m_j < \deg f + \deg g$ (оскільки хоч один з індексів i, j відмінний від 1 і тому $l_i < \deg f \vee m_j < \deg g$). Отже, $\deg(fg) = \deg(\varphi_{l_1}\psi_{m_1}) = \deg f + \deg g$, що й треба було довести.

Наслідок. *У кільці $R[x_1, \ldots, x_n]$ дільниками одиниці можуть бути лише відмінні від нуля константи.*

Доведення. Одиницею кільця $R[x_1, \ldots, x_n]$ є, очевидно, константа 1. Рівність $f(x_1, \ldots, x_n)\, g(x_1, \ldots, x_n) = 1$ означає, що $f(x_1, \ldots, x_n) \neq 0$, $g(x_1, \ldots, x_n \neq 0$, $\deg(fg) = 0$. Оскільки $\deg(fg) = \deg f + \deg g$, то $\deg(fg) = 0 \Rightarrow \deg f = 0 \wedge \deg g = 0$, тобто дільниками одиниці є лише відмінні від нуля константи.

З попереднього викладу випливає, що для многочленів від багатьох змінних поняття степеня члена вже недостатнє для встановлення єдиного порядку розміщення членів, як це було для многочленів від однієї змінної. Ураховуючи зручність однозначного упорядкування членів при вивченні многочленів, бажано ввести якусь іншу характеристику членів многочлена з тим, щоб для членів, які не є подібними, ця

характеристика, на відміну від степеня, завжди була неоднаковою. Найбільш поширений в алгебрі так званий лексикографічний принцип упорядкування членів многочлена. Термін «лексикографічний» походить від грецького слова «лексикон», що означає словник. Як відомо, у словниках слова упорядковують за алфавітом, тобто з двох слів, які починаються з різних букв, раніше йде те слово, в якому перша буква стоїть в алфавіті раніше, ніж перша буква другого слова. Якщо ж перші букви двох слів однакові, то алфавітний принцип застосовується до других букв і т. д. У випадку членів многочлена роль першої, другої і т. д. букв виконують відповідно $x_1, x_2, \ldots$, алфавітному ж принципу впорядкування i-ї букви відповідає упорядкування за степенями змінної x_i.

Розглянемо будь-які два члени многочлена

$$T_1 = A x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}, \qquad (15)$$

$$T_2 = B x_1^{l_1} x_2^{l_2} \ldots x_n^{l_n}. \qquad (16)$$

Якщо ці члени не подібні, то не всі відповідні степені k_i і l_i рівні між собою, тобто існує принаймні одне таке натуральне число p $(1 \leqslant p \leqslant n)$, що $k_i = l_i$ при $i = 1, 2, \ldots, p-1$, але $k_p \neq l_p$. Якщо $k_p > l_p$, то член (15) називається вищим, ніж член (16). Якщо ж $k_p < l_p$, то член (15) називається нижчим, ніж член (16); це рівнозначно тому, що член (16) вищий, ніж член (15).

Як бачимо, з будь-яких двох неподібних членів многочлена один вищий, ніж другий. Очевидно, далі, що коли член T_1 вищий за член T_2, а член T_2 вищий за член T_3, то T_1 вищий за T_3. Звідси зрозуміло, що завжди можна так розмістити члени в канонічній формі многочлена, щоб вищі члени передували нижчим. Таке розміщення і називається лексикографічним.

Наприклад, многочлен (10) при лексикографічному розміщенні можна записати так:

$$f(x_1, x_2, x_3) = x_1^4 x_2 x_3 - 2x_1^3 - x_1^2 x_2^3 + 2x_1^2 x_2^2 x_3 + x_1^2 x_2 x_3^2 + 2x_1 x_2^3 x_3^2 + 5x_1 x_3^3 + 6.$$

Називатимемо перший по порядку член многочлена при лексикографічному розміщенні вищим членом многочлена.

Відносно вищих членів двох многочленів доведемо таку лему, яка буде потрібна далі.

Лема. *Вищий член добутку двох многочленів дорівнює добутку вищих членів цих многочленів.*

Доведення. Візьмемо два довільні многочлени $f(x_1, \ldots, x_n)$ та $g(x_1, \ldots, x_n)$ і розмістимо їх члени лексикографічно:

$$f(x_1, \ldots, x_n) = A x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n} + B x_1^{l_1} x_2^{l_2} \ldots x_n^{l_n} + \cdots + D x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n},$$

$$g(x_1, \ldots, x_n) = L x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} + M x_1^{\beta_1} x_2^{\beta_2} \ldots x_n^{\beta_n} + \cdots + K x_1^{\gamma_1} x_2^{\gamma_2} \ldots x_n^{\gamma_n}.$$

Насамперед зауважимо, що при множенні многочлена $f(x_1, \ldots, x_n)$ на будь-який член виду $S = Cx_1^{\omega_1}x_2^{\omega_2} \ldots x_n^{\omega_n}$ лексикографічне розміщення членів многочлена, очевидно, не порушується.

Утворимо добуток многочлена $f(x_1, x_2, \ldots, x_n)$ на многочлен $g(x_1, x_2, \ldots, x_n)$, помноживши спочатку $f(x_1, x_2, \ldots, x_n)$ на перший член многочлена $g(x_1, x_2, \ldots, x_n)$, потім на другий член і т. д. Із сказаного вище зрозуміло, що кожна група членів, утворена внаслідок послідовного множення, буде розміщена лексикографічно. Отже, вищий член усього добутку слід шукати лише серед вищих членів цих окремих груп, тобто серед членів

$$LAx_1^{\alpha_1+k_1}x_2^{\alpha_2+k_2} \ldots x_n^{\alpha_n+k_n}; \quad MAx_1^{\beta_1+k_1}x_2^{\beta_2+k_2} \ldots x_n^{\beta_n+k_n}; \ldots;$$
$$KAx_1^{\gamma_1+k_1}x_2^{\gamma_2+k_2} \ldots x_n^{\gamma_n+k_n}.$$

Але ці члени записані тут у такому порядку, в якому їх дістаємо при множенні многочлена $g(x_1, x_2, \ldots, x_n)$, розміщеного лексикографічно, на член $Ax_1^{k_1}x_2^{k_2} \ldots x_n^{k_n}$. Тому вищий серед них буде той, який утворений при множенні вищого члена $g(x_1, x_2, \ldots, x_n)$, тобто член

$$LAx_1^{\alpha_1+k_1}x_2^{\alpha_2+k_2} \ldots x_n^{\alpha_n+k_n},$$

що й треба було довести.

Крім лексикографічного розміщення членів многочлена нам доведеться досить часто користуватися розміщенням членів за степенями однієї змінної, тобто зображенням типу (6). У цьому випадку многочлен $f(x_1, x_2, \ldots, x_n)$ над кільцем R можна записати у вигляді:

$$f(x_1, x_2, x_3, \ldots, x_n) = A_s(x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n)x_p^s + A_{s-1}(x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n)x_p^{s-1} + \ldots + A_0(x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n),$$

де коефіцієнти $A_i(x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n)$ $(i = 0, 1, \ldots, s)$ є многочленами від $n-1$ змінних $x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n$ над кільцем R.

25.3. Функціональне тлумачення многочленів. Як і у випадку многочленів та раціональних дробів від однієї змінної, природно поставити питання про можливість тлумачити многочлени $f(x_1, x_2, \ldots, x_n)$ як функції від n змінних і про зв'язок між алгебраїчним і функціональним підходами.

Кожному многочлену $f(x_1, x_2, \ldots, x_n) \in R[x_1, x_2, \ldots, x_n]$ поставимо у відповідність функцію

$$\varphi_f : R^n \to R, \tag{17}$$

значення якої визначаються умовою

$$\forall_{\alpha_1, \alpha_2, \ldots \alpha_n \in R} [\varphi_f(\alpha_1, \alpha_2, \ldots, \alpha_n) = f(\alpha_1, \alpha_2, \ldots, \alpha_n)], \tag{18}$$

де $f(\alpha_1, \alpha_2, \ldots, \alpha_n)$ — елемент кільця R, який дістанемо, підставивши у вираз для многочлена $f(x_1, x_2, \ldots, x_n)$ замість x_1 елемент $\alpha_1 \in R$, замість x_2 — елемент $\alpha_2 \in R$, ..., замість x_n — $\alpha_n \in R$ і виконавши відповідні дії множення і додавання (в розумінні операцій, заданих в R).

Теорема 5. *Якщо R — область цілісності характеристики 0, то кільце $R[x_1, x_2, \ldots, x_n]$ ізоморфне сукупності усіх функцій φ_f, визначених умовами* (17) — (18).

Доведення. Кожному многочлену $f(x_1, x_2, \ldots, x_n) \in R[x_1, \ldots, x_n]$ поставимо у відповідність функцію φ_f згідно з (17) — (18).

Легко перевірити, що ця відповідність $f \to \varphi_f$ зберігає операції, тобто $\forall_{f,g \in R[x_1,\ldots,x_n]} [f \to \varphi_f \wedge g \to \varphi_g \Rightarrow f + g \to \varphi_f + \varphi_g \wedge fg \to \varphi_f\varphi_g]$. Якщо $s(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) + g(x_1, \ldots, x_n)$, то очевидно,

$$\forall_{\alpha_1, \alpha_2, \ldots, \alpha_n \in R} [s(\alpha_1, \ldots, \alpha_n) = f(\alpha_1, \ldots, \alpha_n) + g(\alpha_1, \ldots, \alpha_n)].$$

Але тоді з (18) очевидно, що $\varphi_s = \varphi_f + \varphi_g$, або $f + g \to \varphi_f + \varphi_g$. Випадок добутку многочленів розглядають аналогічно. Отже, відповідність $f \to \varphi_f$ є гомоморфізм.

Щоб показати, що ця відповідність — ізоморфізм, досить показати її взаємну однозначність. Однозначність відповідності $f \to \varphi_f$ випливає безпосередньо з означення (18). Покажемо, що

$$[\varphi_f = \varphi_g] \Rightarrow [f(x_1, \ldots, x_n) = g(x_1, \ldots, x_n)]. \tag{19}$$

Рівність у лівій частині імплікації (19) слід розуміти функціонально (тобто $\forall_{\alpha_1,\ldots,\alpha_n \in R} [\varphi_f(\alpha_1, \ldots, \alpha_n) = \varphi_g(\alpha_1, \ldots, \alpha_n)]$), а рівність у правій частині — як рівність елементів кільця $R[x_1, x_2, \ldots, x_n]$.

Оскільки $[\varphi_f = \varphi_g] \equiv [\varphi_f - \varphi_g = 0] \equiv [\varphi_{f-g} = 0]$, а $[f = g] \equiv [f - g = 0]$, то імплікація (19) рівнозначна імплікації $\varphi_q = 0 \Rightarrow q = 0$, або в розгорнутому вигляді:

$$\forall_{\alpha_1, \alpha_2, \ldots, \alpha_n \in R} [q(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0] \Rightarrow q(x_1, x_2, \ldots, x_n) = 0. \tag{20}$$

Імплікацію (20) доведемо індукцією по n. При $n = 1$ правильність її було доведено в п. 21.5. Припустимо, що вона правильна для многочленів від $(n-1)$-ї змінної, і доведемо, що тоді вона правильна і для многочленів від n змінних. Цим імплікацію (20), а разом з нею і теорему, буде доведено. Упорядкуємо многочлен $q(x_1, x_2, \ldots, \ldots, x_n)$ за степенями будь-якої однієї змінної, наприклад x_n. Тоді $q(x_1, \ldots, x_n)$ можна записати у вигляді:

$$q(x_1, x_2, \ldots, x_n) = A_l(x_1, x_2, \ldots, x_{n-1})x_n^l + A_{l-1}(x_1, x_2, \ldots, x_{n-1})x_n^{l-1} + \ldots + A_1(x_1, x_2, \ldots, x_{n-1})x_n + A_0(x_1, x_2, \ldots, x_{n-1}), \tag{21}$$

де вирази $A_i(x_1, x_2, \ldots, x_{n-1})$ є многочленами від $(n-1)$-ї змінної, причому мають ті самі коефіцієнти, які мав многочлен $q(x_1, \ldots, x_n)$ до упорядкування за степенями змінної x_n.

Надамо змінним $x_1, x_2, \ldots, x_{n-1}$ будь-яких значень $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ з кільця R. Тоді (21) можна розглядати як многочлен від однієї змінної x_n. За умовою він тотожно на R дорівнює нулю, тому з теореми 4 § 21 дістаємо для його коефіцієнтів

$$A_i(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}) = 0 \qquad (i = 0, 1, 2, \ldots, l).$$

Оскільки $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ — довільні значення відповідних змінних і для многочленів від $n-1$ змінної за припущенням імплікація (20) справедлива, то всі коефіцієнти многочленів $A_i(x_1, x_2, \ldots, x_{n-1})$ дорівнюють нулю. Але ці коефіцієнти є одночасно і коефіцієнтами даного многочлена. Отже, усі коефіцієнти многочлена $q(x_1, x_2, \ldots, x_n)$ дорівнюють нулю. Теорему доведено.

З теореми 5 випливає, що для випадку областей цілісності R з одиницею, характеристика яких дорівнює 0 (зокрема, для всіх числових полів), алгебраїчне і функціональне тлумачення многочленів з $R[x_1, x_2, \ldots, x_n]$ цілком рівноправні. Це і є обгрунтування застосовності алгебраїчних властивостей многочленів від кількох змінних в аналізі і теорії функцій.

25.4. Подільність у кільці многочленів від кількох змінних. Досі ми розглядали переважно такі властивості многочленів від багатьох змінних, які повторюють (з природним узагальненням) відпо-

відні властивості многочленів від однієї змінної. Але тепер ми підійшли до такого кола питань, в якому проявляється специфіка многочленів від кількох змінних. Як ми побачимо, теорія подільності у кільці таких многочленів істотно відрізняється від теорії, розглянутої в § 22.

Звичайно, означення подільності, дільника, загальні властивості відношення подільності, поняття і властивості незвідних многочленів переноситься на кільце многочленів від n змінних без усяких змін, оскільки вони властиві будь-якій області цілісності, зокрема й $R[x_1, x_2, \ldots, x_n]$. Нагадаємо ці означення і факти стосовно многочленів від кількох змінних без коментарів і доведень. При цьому, як і у випадку многочленів від однієї змінної, вважатимемо, що основна область цілісності R є полем, яке позначатимемо P. Для спрощення записів многочлени записуватимемо без змінних $x_1, x_2, \ldots, x_n$ і кільце $P[x_1, x_2, \ldots, x_n]$ скорочено позначатимемо P_n.

Означення 1. *Вважатимемо, що многочлен $f \in P_n$ ділиться на многочлен $g \in P_n$, відмінний від нуля, і записуватимемо $f \vdots g$, якщо існує такий многочлен $s \in P_n$, що $f = gs$. При цьому g називають дільником многочлена f.*

Відношення подільності многочленів в P_n має такі властивості:

1. $\forall\limits_{f,g,h \in P_n} [f \vdots g \wedge g \vdots h \Rightarrow f \vdots h]$.
2. $\forall\limits_{f,g,h \in P_n} [f \vdots h \wedge g \vdots h \Rightarrow (f+g) \vdots h \wedge (f-g) \vdots h]$.
3. $\forall\limits_{f,h \in P_n} [f \vdots h \Rightarrow \forall\limits_{g \in P_n} [fg \vdots h]]$.
4. $\forall\limits_{h, f_1, \ldots, f_m \in P_n} [f_1 \vdots h \wedge f_2 \vdots h \wedge \cdots \wedge f_m \vdots h \Rightarrow \forall\limits_{g_1, g_2, \ldots, g_m \in P_n} [f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \vdots h]]$.
5. $\forall\limits_{f \in P_n} \forall\limits_{\substack{c \in P, \\ c \neq 0}} [f \vdots c]$.
6. $\forall\limits_{f,g \in P_n} \forall\limits_{\substack{c \in P, \\ c \neq 0}} [f \vdots g \Rightarrow f \vdots cg]$.

Відомо, що дільниками 1 в P_n можуть бути лише відмінні від нуля константи. З властивості 5 випливає, що кожна відмінна від нуля константа справді є дільником одиниці. Очевидно, що *асоційованими многочленами в кільці P_n є такі і тільки такі многочлени, які відрізняються множником, що є відмінною від нуля константою.*

Означення 2. *Многочлен $p \in P_n$ називається незвідним у полі P, якщо* $\deg p \geqslant 1$ *і* $\forall\limits_{u,v \in P_n} [p = uv \Rightarrow \deg u = 0 \vee \deg v = 0]$. *Многочлен $q \in P_n$ називається звідним у полі P, якщо* $\deg q \geqslant 1$ *і* $\exists\limits_{u,v \in P_n} [q = uv \wedge \deg u \geqslant 1 \wedge \deg v \geqslant 1]$.

Найпростішими властивостями незвідних многочленів є такі:

1. *Якщо p незвідний у P, то і будь-який асоційований з ним многочлен cp незвідний у P.*
2. *Якщо p, q незвідні у P многочлени і $p \vdots q$, то p і q — асоційовані.*

3. *Будь-який многочлен $p \in P_n$ першого степеня незвідний у P.*

На цьому по суті закінчується аналогія теорії подільності в кільцях $P[x]$ і $P[x_1, x_2, \ldots, x_n]$ при $n \geqslant 2$. Специфіка останнього полягає в тому, що $P[x_1, x_2, \ldots, x_n]$ при $n \geqslant 2$ *не є кільцем головних ідеалів* і (тим більше) *евклідовим кільцем*. Тому безпосереднє поперенесення на це кільце результатів пп. 14.2—14.3 або пп. 22.5—22.7 неможливе.

Щоб переконатись у тому, що не усі ідеали кільця $P[x_1, x_2, \ldots, x_n]$ головні, розглянемо такий приклад. Нехай $n = 2$, $R[x, y]$ — кільце усіх многочленів $f(x, y)$ над полем R дійсних чисел. Позначимо через I сукупність усіх тих многочленів $f(x, y) \in R[x, y]$, в яких вільний член (тобто член виду $Ax^0y^0 = A \in R$) дорівнює нулю. Очевидно, I є ідеал: адже для будь-яких $f(x, y), \in R[x, y]$, $g(x, y) \in I$ маємо $f(x, y)\, g(x, y) \in I$.

Проте ідеал I не є головним. Справді, кожний головний ідеал кільця $R[x, y]$, як і будь-якого кільця з одиницею (п. 13.1), породжується деяким елементом $s(x, y) \in R[x, y]$, тобто має будову $\{s(x, y) f(x, y)\}$, де $s(x, y)$ — фіксований, а $f(x, y)$ — довільний многочлен з $R[x, y]$. Але ідеал I не може бути породжений жодним многочленом $s(x, y) \in R[x, y]$: якби $\deg s = 0$, то було б $I = R[x, y]$, що неправильно; якщо ж $\deg s \geqslant 1$, то I не міг би містити обидва многочлени $u(x, y) = x + y$ і $v(x, y) = x - y$, бо ці многочлени не мають спільного множника $s(x, y)$ ненульового степеня. Оскільки, за означенням I, $u(x, y) \in I$ і $v(x, y) \in I$, то I не є головний ідеал. Проведене міркування без будь-яких змін переноситься на випадок довільного кільця $P[x_1, x_2, \ldots, x_n]$ при $n \geqslant 2$.

Отже, кільце $P[x_1, x_2, \ldots, x_n]$ многочленів від кількох змінних не є кільцем головних ідеалів. Тому воно не може бути й евклідовим кільцем; адже кожне евклідове кільце є кільцем головних ідеалів (п. 14.3); це означає, що алгоритм Евкліда і його численні наслідки, розглянуті в § 22 для кільця $P[x]$, не поширюються на випадок многочленів від кількох змінних.

Проте далі буде показано, що один з основних результатів теорії подільності в кільці $P[x]$, а саме — *можливість і єдиність розкладу многочлена у добуток незвідних множників — залишається в силі і в кільці $P[x_1, x_2, \ldots, x_n]$ при $n \geqslant 2$.*

Зрозуміло, що спосіб доведення цього факту повинен бути істотно іншим, ніж у загальній теорії кілець головних ідеалів (п. 14.2) і використовувати специфічні властивості кільця многочленів. Адже для будь-якої області цілісності з одиницею, що не є кільцем головних ідеалів, твердження про єдиність розкладу на прості множники, взагалі кажучи, неправильне (п. 22.7).

Можливість розкладу будь-якого многочлена $f(x_1, \ldots, x_n) \in P[x_1, \ldots, x_n]$ у добуток незвідних множників можна встановити досить просто. На відміну від цього, єдиність такого розкладу є глибокий факт, доведення якого досить складне. Тому ми викладемо відповідні твердження окремо.

Теорема 6. *Будь-який многочлен $f(x_1, x_2, \ldots, x_n)$ над полем P ненульового степеня можна подати як добуток многочленів, незвідних у полі P.*

Доведення проведемо індукцією за степенем многочлена $f(x_1, \ldots, x_n)$. Якщо $\deg f = 1$, то $f(x_1, \ldots, x_n)$ — незвідний многочлен кільця $P[x_1, x_2, \ldots, x_n]$ і тому його можна вважати «добутком» з одним множником. Припустимо, що твердження теореми правильне для усіх многочленів $f(x_1, x_2, \ldots, x_n) \in P[x_1, x_2, \ldots, x_n]$ таких, що $1 \leqslant \deg f < m$, і перевіримо його правильність для многочленів m-го степеня.

Нехай $f(x_1, x_2, \ldots, x_n)$ — будь-який многочлен m-го степеня над полем P. Якщо він незвідний, то для нього теорема правильна безпосередньо. Якщо ж $f(x_1, x_2, \ldots, x_n)$ — звідний, то існують многочлени $g(x_1, x_2, \ldots, x_n), s(x_1, x_2, \ldots, x_n) \in P[x_1, \ldots, x_n]$ такі, що $f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n)\, s(x_1, x_2, \ldots, x_n)$, $\deg g < \deg s = m$; $\deg s < \deg f = m$. За припущенням індукції кожний з многочленів $g(x_1, x_2, \ldots, x_n)$ і $s(x_1, \ldots, x_n)$ можна подати як добуток многочленів, незвідних у полі P. Отже, їх добуток $f(x_1, \ldots, x_n)$ також можна розкласти на незвідні множники. Теорему доведено.

Перейдемо тепер до питання про єдиність розкладу многочлена на незвідні множники у кільці $P[x_1, x_2, \ldots, x_n]$.

Ураховуючи індуктивне означення кільця $P[x_1, x_2, \ldots, x_n]$ многочленів від n змінних (див. п. 25.1), досить встановити правильність такого твердження.

Теорема 7. *Якщо S — область цілісності з одиницею, в якій кожний елемент, відмінний від нуля і від дільника одиниці, однозначно розкладається в добуток простих елементів* [1], *то кільце $S[x]$ многочленів над S має такі самі властивості.*

Твердження про єдиність розкладу многочленів на незвідні множники безпосередньо випливає з теореми 7.

Наслідок. *Будь-який многочлен $f(x_1, x_2, \ldots, x_n) \in P[x_1, x_2, \ldots, x_n]$ ненульового степеня розкладається на незвідні множники єдиним способом* (з точністю до сталих множників і порядку співмножників).

Справді, якщо теорема 7 справедлива, то сформульований наслідок можна легко довести, використовуючи індукцію по n. При $n = 1$ це твердження правильне. Якщо ж воно правильне при $n = k$, тобто у кільці $P[x_1, \ldots, x_k]$, що є областю цілісності з одиницею (теорема 1), то внаслідок теореми 7 воно правильне і у кільці $P[x_1, \ldots, x_k, x_{k+1}]$, що є кільцем многочленів від x_{k+1} над $P[x_1, \ldots, x_k]$. Тобто наслідок правильний при будь-якому натуральному n.

Отже, справа зводиться до доведення теореми 7. Основна ідея доведення полягає в тому, щоб «вкласти» область цілісності S у поле відношень T, що завжди можливо (див. п. 12.4), а далі скористатися тим відомим фактом, що кільце многочленів від однієї змінної над полем є кільцем головних ідеалів (п. 22.4) і тому в ньому розклад на незвідні множники єдиний (п. 22.7). Щоб встановити зв'язок між кільцями $S[x]$ і $T[x]$, доведемо спочатку кілька допоміжних тверджень.

Всюди далі в цьому пункті S позначає область цілісності з одиницею, в якій кожний елемент, відмінний від нуля і від дільника одиниці, однозначно розкладається на прості множники.

Лема 1. *Для будь-якої скінченної системи елементів $a_1, a_2, \ldots, a_k \in S$, відмінних від нуля, існує єдиний* (з точністю до дільників одиниці) *найбільший спільний дільник.*

Доведення цього твердження нічим не відрізняється від доведення теореми 9 п. 22.7.

Означення 3. *Многочлен $f(x) \in S[x]$, відмінний від нуля, називається примітивним* (відносно S), *якщо НСД його коефіцієнтів дорівнює* 1.

Лема 2. *Добуток двох примітивних многочленів з $S[x]$ є знову примітивний многочлен.*

Доведення. Нехай $f(x)$, $g(x) \in S[x]$ — примітивні многочлени:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$
$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0.$$

Припустимо, що многочлен $q(x) = f(x)\, g(x)$ не є примітивним відносно S. Це означає, що всі його коефіцієнти мають НСД d, який не є дільником 1 і тому однозначно розкладається на прості множники. Нехай p — один з таких множників. Всі коефіцієнти a_i ($i = 0, 1, \ldots, n$) не можуть ділитись на p, бо $f(x)$ — примітивний многочлен. Нехай k — найменший з індексів коефіцієнтів a_i, що не діляться на p. Аналогічно, нехай l — найменший з індексів коефіцієнтів b_j, які не діляться на p. Розглянемо коефіцієнт c_{k+l} при x^{k+l} многочлена $q(x)$

$$c_{k+l} = \sum_{i+j=k+l} a_i b_j.$$

В цій сумі лише один доданок $a_k b_l$ не ділиться на p; решта ж доданків обов'язково діляться на p, бо в них або індекс i менший за k, або індекс j менший за l (як завжди $a_i = 0$ при $i > n$, $b_j = 0$ при $j > m$). Але тоді c_{k+l} не може ділитись на p, а це суперечить припущенню. Лему доведено.

Нехай тепер T — поле відношень області цілісності S. Розглянемо кільце $T[x]$ многочленів над полем T, тобто многочленів, які можна подати у формі

$$\varphi(x) = \frac{a_n}{b_n} x^n + \frac{a_{n-1}}{b_{n-1}} x^{n-1} + \cdots + \frac{a_1}{b_1} x + \frac{a_0}{b_0}; \quad a_i, b_i \in S \quad (22)$$
$$(i = 0, 1, \ldots, n).$$

В кільці $T[x]$, згідно з теоремою 8 п. 22.7, розклад на незвідні множники однозначний. Щоб перейти від $T[x]$ до нашого кільця $S[x]$, поставимо кожному многочлену $\varphi(x) \in T[x]$ у відповідність деякий примітивний многочлен $f_\varphi(x) \in S[x]$ згідно з таким правилом.

Нехай $\varphi(x)$ — ненульовий многочлен з $T[x]$, поданий у формі (22). Помноживши $\varphi(x)$ на елемент $b = b_0 b_1 \ldots b_{n-1} b_n \in S$, дістанемо многочлен $b\varphi(x)$ з кільця $S[x]$. Позначимо через a НСД коефіцієнтів многочлена $b\varphi(x)$ у кільці S. Тоді многочлен

$$f_\varphi(x) = \frac{b\varphi(x)}{a} \quad (23)$$

[1] Домовимось розуміти однозначність розкладу як його єдиність з точністю до дільників одиниці і порядку співмножників. Зауважимо, що у дальшому викладі ми вживаємо як загальний термін п р о с т и й елемент (стосовно довільного кільця), так і його конкретизацію н е з в і д н и й (стосовно саме кільця многочленів).

є примітивним многочленом у кільці $S[x]$ і називається *відповідним многочленом* $\varphi(x) \in T[x]$.

Лема 3. *Відповідність* $\varphi(x) \to f_\varphi(x)$ *взаємно однозначна з точністю до дільників одиниці.* Точніше: *для кожного відмінного від нуля многочлена* $\varphi(x) \in T[x]$ *відповідний примітивний многочлен* $f_\varphi(x) \in S[x]$ *єдиний з точністю до дільників одиниці кільця* S; *два многочлени* $\varphi(x)$, $\psi(x) \in T[x]$, *яким відповідає той самий примітивний многочлен* $f_\varphi(x) = f_\psi(x) \in S[x]$, *збігаються з точністю до дільників одиниці кільця* $T[x]$.

Доведення. а) Як видно з формули (23), що визначає $f_\varphi(x)$ за даним $\varphi(x)$, неоднозначність $f_\varphi(x)$ може бути пов'язана лише з неоднозначністю a — НСД коефіцієнтів многочлена $b\varphi(x)$. Оскільки НСД елементів кільця S визначається з точністю до дільників одиниці кільця S, то те саме справедливо і для $f_\varphi(x)$.

б) Нехай $\varphi(x)$, $\psi(x)$ — ненульові многочлени з $T[x]$, $f_\varphi(x) = \frac{b}{a}\varphi(x)$, $f_\psi(x) = \frac{d}{c}\psi(x)$ — відповідні примітивні многочлени з $S[x]$. Якщо $f_\varphi(x) = f_\psi(x)$, то

$$\frac{b}{a}\varphi(x) = \frac{d}{c}\psi(x) \Leftrightarrow \varphi(x) = \frac{ad}{bc}\psi(x).$$

Очевидно, що a, b, c, d — відмінні від нуля елементи з S, а тому $\frac{ad}{bc}$ — якийсь відмінний від нуля елемент поля T або дільник одиниці кільця $T[x]$. Лему доведено.

Лема 4. *Многочлен* $\omega(x) \in T[x]$ *звідний в* T *тоді і тільки тоді, коли многочлен* $f_\omega(x) \in S[x]$ *звідний в* S, *причому* $\omega(x) = \varphi(x)\psi(x) \Leftrightarrow f_\omega(x) = f_\varphi(x) f_\psi(x)$ (з точністю до відповідних дільників одиниці). *Многочлен* $\omega(x) \in T[x]$ *незвідний в* T *тоді і тільки тоді, коли многочлен* $f_\omega(x)$ *незвідний в* S.

Доведення. а) Нехай $\omega(x)$ — звідний в T многочлен і $\omega(x) = \varphi(x)\psi(x)$. За правилом (23), $f_\varphi(x) = \frac{b}{a}\varphi(x)$, $f_\psi(x) = \left(\frac{d}{c}\right) \times \psi(x)$, звідки

$$\omega(x) = \varphi(x)\psi(x) = \frac{a}{b} f_\varphi(x) \frac{c}{d} f_\psi(x) = \frac{ac}{bd} f_\varphi(x) f_\psi(x).$$

З другого боку, знайдемо $f_\omega(x)$ за правилом (23)

$$f_\omega(x) = \frac{q}{p}\omega(x) = \frac{qac}{pbd} f_\varphi(x) f_\psi(x). \qquad (24)$$

За лемою 2 $f_\varphi(x) f_\psi(x)$ є примітивний многочлен в $S[x]$. Оскільки $f_\omega(x)$ є також примітивний многочлен в $S[x]$, то з рівності (24) випливає, що $\frac{qac}{pbd} = \xi$ є елемент кільця S і притому дільник одиниці цього кільця. Справді, ξ не може бути нескоротним дробом, бо якби $\xi = \frac{l}{m}$, $(l, m) = 1$, $m \neq 1$, то всі коефіцієнти многочлена $f_\varphi(x) f_\psi(x)$ ділилися б на m, що суперечить тому, що $f_\varphi(x) f_\psi(x)$ — примітивний. Отже, $\xi \in S$. Оскільки всі коефіцієнти примітивного многочлена $f_\omega(x)$ діляться на ξ, то ξ є дільник одиниці. Отже, $f_\omega(x)$ з точністю до дільників одиниці збігається з $f_\varphi(x) f_\psi(x)$.

Навпаки, якщо $f_\omega(x) = f_\varphi(x) f_\psi(x)$, то $\frac{q}{p}\omega(x) = \frac{bd}{ac}\varphi(x)\psi(x)$, або $\omega(x) = \frac{pbd}{acq}\varphi(x)\psi(x)$, тобто $\omega(x) = \varphi(x)\varphi(x)$ з точністю до дільників одиниці кільця $T[x]$.

б) Якщо $\omega(x) \in T[x]$ — незвідний в T многочлен, то $\deg \omega \geqslant 1$ і $\omega(x)$ не є звідним в T. Але тоді з (24) випливає, що $\deg f_\omega \geqslant 1$; а з доведеного в п. а.), що $f_\omega(x)$ — незвідний в S. Цілком аналогічно з незвідності $f_\omega(x)$ в S дістаємо незвідність $\omega(x)$ в T. Лему доведено.

Доведення теореми 7.

1. Доведемо спочатку, що будь-який відмінний від нуля і від дільника одиниці примітивний многочлен $f(x)$ однозначно розкладається на незвідні (прості) множники в кільці $S[x]$. Якщо $\deg f = 0$, то $f(x) = a \in S$. За умовою теореми будь-яка константа $a \in S$ (відмінна від 0 і дільника 1) однозначно розкладається на прості множники. Отже, у цьому разі теорема справедлива. Якщо ж $\deg f \geqslant 1$, то візьмемо в $T[x]$ многочлен $\varphi(x)$ такий, що $f(x) = f_\varphi(x)$; такий многочлен існує і він єдиний (з точністю до дільників одиниці). Для $\varphi(x) \in T[x]$ розклад на незвідні множники єдиний (з точністю до дільників одиниці і порядку множників). За лемою 4 цьому розкладу відповідає єдиний (у тому ж розумінні) розклад $f(x)$ на незвідні примітивні множники в $S[x]$. Цим твердження теореми доведене для випадку примітивних многочленів.

2. Якщо $f(x) \in S[x]$ не є примітивним многочленом і $f(x) \neq 0$, то його можна подати у вигляді $f(x) = dq(x)$, де d — НСД усіх коефіцієнтів многочлена $f(x)$, а $q(x)$ — примітивний відносно S многочлен. Для елемента $d \in S$ розклад на прості множники однозначний за умовою теореми, а для $q(x)$ — за першою частиною доведення. Отже, й для $f(x)$ розклад на прості множники однозначний. Теорему доведено.

Кільце многочленів від n змінних над полем P $(n \geqslant 2)$ є важливим прикладом кільця, в якому не всі ідеали головні, але справедливе твердження про однозначність розкладу на прості (незвідні) множники. Саме про цей приклад ми згадували при викладі загальної теорії подільності в кільці.

Поряд з цим деякі властивості подільності в кільці $P[x_1, \ldots, x_n]$ при $n \geqslant 2$ істотно відрізняються від властивостей в кільці многочленів від однієї змінної. Так, відомо (п. 23.3), що для будь-якого поля P і довільного многочлена $f(x) \in P[x]$ при $\deg f \geqslant 2$ існує розширення поля P, в якому $f(x)$ звідний (наприклад, поле розкладу $f(x)$). В кільці многочленів від кількох змінних це не так. Наприклад, многочлен $f(x, y) = x^m + y$, де m — довільне натуральне число, незвідний в будь-якому полі (доведіть це!).

25.5. Раціональні дроби від кількох змінних. Кільце $P[x_1, \ldots, x_n]$ многочленів від n змінних, так само як і кільце $P[x]$ многочленів від однієї змінної, можна «вкласти» у поле відношень $P(x_1, x_2, \ldots, x_n)$. Точніше, можна побудувати таке поле $P(x_1, x_2, \ldots, \ldots, x_n)$, яке містить $P[x_1, \ldots, x_n]$ як підкільце і кожний елемент якого можна подати як

частку $\frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)}$ двох многочленів $f(x_1, \ldots, x_n)$, $g(x_1, \ldots, x_n) \in P[x_1, \ldots, x_n]$. Таке поле єдине з точністю до ізоморфізму.

Існування і єдиність поля $P(x_1, \ldots, x_n)$, а також усі інші відомості про раціональні дроби, наведені в п. 24.1 для випадку однієї змінної, справедливі і для випадку n змінних ($n \geqslant 2$), оскільки вони є безпосередніми наслідками загальної теорії поля відношень довільної області цілісності (п. 12.4): адже за теоремою 1, $P[x_1, \ldots, x_n]$ є областю цілісності з одиницею для будь-якого поля P і будь-якого натурального n. Читач без труднощів зуміє переформулювати згадані властивості [1]. Зауважимо, що ми вже по суті у попередньому викладі користувались існуванням і властивостями поля раціональних дробів кількох змінних над P, коли застосовували теорему 7 п. 25.4 до доведення можливості однозначного розкладу многочленів від n змінних на незвідні множники. Адже для кільця цілісності $S = P[x_1, \ldots, x_n]$ відповідне поле відношень T і є поле раціональних дробів $P(x_1, x_2, \ldots, x_n)$.

Залишається в силі і результат п. 24.2 щодо рівноправності алгебраїчного і функціонального тлумачення раціональних дробів над полем характеристики 0. Єдина трудність, що виникає при перенесенні його на випадок n змінних, полягає в тому, що при $n \geqslant 2$ може існувати нескінченне число систем значень $\alpha_1, \alpha_2, \ldots, \alpha_n \in P$, при яких знаменник $g(x_1, x_2, \ldots, x_n) \neq 0$ нескоротної частки $\frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)}$ перетворюється в нуль, а відповідна дробово-раціональна функція втрачає сенс (у випадку $n = 1$ многочлен $g(x)$ завжди мав скінченне число коренів). Так, многочлен $f(x, y) = x - y$ приймає значення 0 для безлічі пар значень $x = \alpha$, $y = \alpha$ (α — довільний елемент поля p). Однак для міркувань, за допомогою яких обгрунтовується ізоморфізм між раціональними дробами і відповідними дробово-раціональними функціями, було істотним лише те, що *сукупність систем* $\alpha_1, \ldots, \alpha_n$ ($\alpha_i \in P$), *для яких* $g(x_1, \ldots, x_n)$ *відмінний від нуля, нескінченна.* А цей факт, справедливий для довільного n, неважко довести звичайним методом індукції за n. Що ж до результатів п. 24. 3 (розклад раціональних дробів на елементарні), то їх безпосереднє перенесення на випадок n змінних неможливе в зв'язку з тим, що для обгрунтування їх в п. 24.3 значною мірою використано евклідовість кільця $P[x]$ многочленів від однієї змінної.

§ 26. СИМЕТРИЧНІ МНОГОЧЛЕНИ

26.1. Означення і елементарні властивості. Важливим класом многочленів від кількох змінних є клас так званих симетричних многочленів.

Означення. *Многочлен* $f(x_1, x_2, \ldots, x_n)$ *називається симетричним відносно змінних* $x_1, x_2, \ldots, x_n$, *якщо внаслідок довільної перестановки змінних* $x_1, x_2, \ldots, x_n$ *утворюється многочлен, який дорівнює даному.*

Приклад. Многочлени

$$f_1(x_1, x_2) = x_1^2 + x_2^2,$$

$$f_2(x_1, x_2) = x_1x_2^2 + 3x_1x_2 - 2x_1 - 2x_2 + x_1^2x_2 + 5$$

симетричні відносно змінних x_1 і x_2, бо не змінюються від перестановки цих змінних.

Многочлен $f_3(x_1, x_2, x_3) = 2x_2^2 + x_1x_2 + x_2x_3$ симетричний відносно змінних x_1 і x_3, але не симетричний відносно змінних x_2, x_3 або x_1, x_2. Тому він не симетричний і відносно всіх змінних x_1, x_2, x_3.

[1] Зокрема залишиться в силі і твердження про можливість подати раціональний дріб у формі нескоротної частки, бо хоч алгоритм Евкліда в кільці $P[x_1, x_2, \ldots, x_n]$ незастосовний, НСД існує для будь-якої пари ненульових многочленів з цього кільця (лема 1 п. 25.4).

З важливими прикладами симетричних многочленів ми, власне кажучи, вже зустрічались у формулах Вієта (п. 23.3). Якщо позначити через $x_1, x_2, \ldots, x_n$ корені многочлена $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$, то за формулами Вієта матимемо:

$$\begin{aligned} x_1 + x_2 + \cdots + x_n &= -a_{n-1}, \\ x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n &= a_{n-2}, \\ &\cdots\cdots\cdots\cdots \\ x_1x_2 \ldots x_n &= (-1)^n a_0. \end{aligned}$$

Позначимо ліві частини цих формул відповідно через $\sigma_1, \sigma_2, \ldots, \sigma_n$, тобто

$$\begin{aligned} \sigma_1 &\underset{\text{df}}{=} x_1 + x_2 + \cdots + x_n, \\ \sigma_2 &\underset{\text{df}}{=} x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n, \\ &\cdots\cdots\cdots\cdots \\ \sigma_n &\underset{\text{df}}{=} x_1x_2 \ldots x_n. \end{aligned} \tag{1}$$

Якщо під $x_1, x_2, \ldots, x_n$ розуміти незалежні змінні, то $\sigma_1, \sigma_2, \ldots, \sigma_n$ є, очевидно, многочлени, симетричні відносно цих змінних. Многочлени (1) називаються *основними симетричними функціями*. Причини цієї назви буде з'ясовано дещо нижче.

Встановимо тепер деякі елементарні властивості довільних симетричних многочленів.

1. *Сума, різниця і добуток симетричних многочленів над деяким полем* P *є знову симетричними многочленами над цим полем.*

Це твердження очевидне.

Наслідок. *Множина всіх симетричних многочленів над полем* P *утворює область цілісності з одиницею відносно дій додавання і множення.*

Зрозуміло, що це кільце є підкільцем всіх многочленів над полем P.

2. *Якщо симетричний многочлен* $f(x_1, x_2, \ldots, x_n)$ *містить деякий член*

$$Mx_1^{l_1}x_2^{l_2} \ldots x_i^{l_i} \ldots x_j^{l_j} \ldots x_n^{l_n}, \tag{2}$$

то він містить і член, утворений з (2) *внаслідок будь-якої перестановки показників* $l_1, l_2, \ldots, l_n$.

Оскільки, як відомо, від довільної перестановки показників $l_1, l_2, \ldots, l_n$ до всякої іншої перестановки цих показників можна перейти за допомогою скінченного числа транспозицій (1, § 25), то досить показати, що при транспозиції довільних двох показників степенів у члені (2) ми дістаємо знову деякий член симетричного многочлена $f(x_1, x_2, \ldots, x_n)$. Виконуючи, наприклад, транспозицію показників l_i та l_j, матимемо член

$$Mx_1^{l_1}x_2^{l_2} \ldots x_i^{l_j} \ldots x_j^{l_i} \ldots x_n^{l_n}. \tag{3}$$

За означенням симетричного многочлена

$$f(x_1, x_2, \ldots, x_i, \ldots, x_j, \ldots, x_n) =$$
$$= f(x_1, x_2, \ldots, x_j, \ldots, x_i, \ldots, x_n).$$

Але другий з цих многочленів повинен містити член (3), бо його дістаємо з члена (2) заміною x_i на x_j і навпаки. Тому внаслідок єдиності канонічної форми і даний многочлен повинен містити член (3).

Наслідок. *Якщо*

$$Ax_1^{l_1}x_2^{l_2} \ldots x_i^{l_i}x_{i+1}^{l_{i+1}} \ldots x_n^{l_n} \tag{4}$$

в вищий член симетричного многочлена, то $l_1 \geqslant l_2 \geqslant l_3 \geqslant \ldots \geqslant l_n$.

Справді, припустимо супротивне, тобто що при якомусь i $l_i < l_{i+1}$. На підставі властивості 2 даний многочлен разом з членом (4) містить і член

$$Ax_1^{l_1}x_2^{l_2} \ldots x_i^{l_{i+1}}x_{i+1}^{l_i} \ldots x_n^{l_n}. \tag{5}$$

Але з умови $l_{i+1} > l_i$ випливає, що член (5) вищий за член (4), тобто член (4) не може бути вищим у многочлені. Ця суперечність доводить наше твердження.

26.2. Основна теорема. Важливу роль в алгебрі відіграє така теорема.

Теорема 1. (Основна теорема теорії симетричних многочленів). *Всякий симетричний многочлен $f(x_1, x_2, \ldots, x_n)$ від n змінних над полем P можна подати у вигляді многочлена від основних симетричних функцій $\sigma_1, \sigma_2, \ldots, \sigma_n$ цих змінних, коефіцієнти якого належать тому самому полю P.*

Доведення. Зробимо насамперед такі зауваження.

1) Усіх членів певного степеня l, утворених з даних змінних $x_1, x_2, \ldots, x_n$ (не враховуючи подібних), може бути лише скінченне число; це число, очевидно, дорівнює числу способів, якими l можна подати як суму n невід'ємних цілих упорядкованих доданків. Наприклад, при $l = 5$, $n = 2$ маємо такі шість можливостей: $5 = 0 + 5$; $5 = 1 + 4$; $5 = 2 + 3$; $5 = 3 + 2$; $5 = 4 + 1$; $5 = 5 + 0$.

2) Теорему досить довести для однорідних симетричних многочленів, бо всякий симетричний многочлен можна подати як суму однорідних симетричних многочленів. Справді, як ми вже зазначали, всякий многочлен є сумою однорідних многочленів. Якщо ж даний многочлен симетричний, то й кожний складовий однорідний многочлен повинен бути симетричний, бо при переставлянні змінних $x_1, x_2, \ldots, x_n$ кожний член може перейти лише в член того самого степеня, тобто в інший член того самого однорідного складового многочлена.

3) Вищий член $x_1^{l_1}x_2^{l_2} \ldots x_n^{l_n}$ будь-якого симетричного многочлена можна подати як *вищий член деякого добутку основних симетричних функцій* $\sigma_1, \sigma_2, \ldots, \sigma_n$.

Справді, розглянемо добуток

$$\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3} \ldots \sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n}. \tag{6}$$

За наслідком з властивості 2, всі степені $l_1 - l_2$, $l_2 - l_3$, ..., $l_{n-1} - l_n$ — невід'ємні числа, тому (6) є многочленом від $x_1, x_2, \ldots, x_n$. За лемою п. 25.2, вищий член цього многочлена дорівнює добутку вищих членів многочленів $\sigma_1, \sigma_2, \ldots, \sigma_n$ (причому піднесення до степеня слід розглядати як множення однакових многочленів). Оскільки вищі члени $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \sigma_n$ дорівнюють відповідно x_1; x_1x_2; ...; $x_1x_2 \ldots x_{n-1}$; $x_1x_2 \ldots x_{n-1}x_n$, то вищий член добутку (6) дорівнює

$$(x_1)^{l_1-l_2}(x_1x_2)^{l_2-l_3} \ldots (x_1x_2 \ldots x_{n-1})^{l_{n-1}-l_n}(x_1x_2 \ldots x_{n-1}x_n)^{l_n},$$

тобто (як це видно після елементарних перетворень) збігається з заданим членом $x_1^{l_1}x_2^{l_2} \ldots x_{n-1}^{l_{n-1}}x_n^{l_n}$.

Після цих зауважень легко довести теорему 1.

Нехай вищий член симетричного многочлена $f(x_1, x_2, \ldots, x_n)$ (який ми в результаті зауваження 2 можемо вважати однорідним многочленом степеня N) дорівнює

$$Ax_1^{l_1}x_2^{l_2} \ldots x_n^{l_n}. \tag{7}$$

Побудуємо симетричний многочлен

$$g(x_1, x_2, \ldots, x_n) \underset{\text{df}}{=} A\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3} \ldots \sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n}.$$

Згідно з зауваженням 3, вищий член цього многочлена дорівнює (7). Крім того, він однорідний, бо такими є всі многочлени $\sigma_1, \sigma_2, \ldots, \sigma_n$, а тому, очевидно, і їх добуток. Степінь многочлена $g(x_1, x_2, \ldots, x_n)$ дорівнює степеню многочлена $f(x_1, x_2, \ldots, x_n)$, бо в них однакові вищі члени.

Візьмемо

$$f_1(x_1, x_2, \ldots, x_n) \underset{\text{df}}{=} f(x_1, x_2, \ldots, x_n) - g(x_1, x_2, \ldots, x_n).$$

Зрозуміло, що $f_1(x_1, x_2, \ldots, x_n)$ — також однорідний симетричний многочлен степеня N. Але $f_1(x_1, x_2, \ldots, x_n)$ вже не містить усіх членів цього степеня. Справді, він не містить вищого члена (7), який у цій різниці знищується. Крім того, в цій різниці знищуються всі $n!$ членів, які дістаємо з вищого члена перестановкою показників $l_1, l_2, \ldots, l_n$, бо ці члени, за властивістю 2, входять в обидва симетричні многочлени.

Тепер зрозуміло, що $f_1(x_1, x_2, \ldots, x_n)$ може містити лише члени, нижчі за (7). Застосовуємо до цього многочлена той самий метод. Нехай вищий член многочлена $f_1(x_1, x_2, \ldots, x_n)$ має вигляд

$$Bx_1^{m_1}x_2^{m_2} \ldots x_n^{m_n}. \tag{8}$$

Вважаючи

$$g_1(x_1, x_2, \ldots, x_n) \underset{\text{df}}{=} B\sigma_1^{m_1-m_2}\sigma_2^{m_2-m_3} \ldots \sigma_{n-1}^{m_{n-1}-m_n}\sigma_n^{m_n}$$

і утворюючи різницю

$$f_2(x_1, x_2, \ldots, x_n) = f_1(x_1, x_2, \ldots, x_n) - g_1(x_1, x_2, \ldots, x_n),$$

бачимо, що $f_2(x_1, x_2, \ldots, x_n)$ є симетричний і однорідний многочлен степеня N, який не може містити ні члена (7), ні члена (8), а тільки члени, нижчі за них. Оскільки, взагалі, різних членів степеня N може

бути лише скінченне число (зауваження 1), то, продовжуючи цей процес, ми на якомусь кроці обов'язково дістанемо, що різниця

$$f_{k+1}(x_1, x_2, \ldots, x_n) = f_k(x_1, x_2, \ldots, x_n) - g_k(x_1, x_2, \ldots, x_n)$$

не може містити жодного члена степеня N, тобто дорівнює нулю. Тоді з рівностей

$$f_1 = f - g,$$
$$f_2 = f_1 - g_1,$$
$$\ldots\ldots\ldots$$
$$f_k = f_{k-1} - g_{k-1},\quad 0 = f_k - g_k$$

випливає, що

$$f = g + g_1 + \cdots + g_{k-1} + g_k.$$

А оскільки всі $g_1(x_1, x_2, \ldots, x_n)$ виражені через добутки $\sigma_1, \sigma_2, \ldots, \sigma_n$, то многочлен $f(x_1, x_2, \ldots, x_n)$ подано як многочлен від основних симетричних функцій

$$f(x_1, x_2, \ldots, x_n) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n), \qquad (9)$$

коефіцієнти якого (як видно з правила побудови g_i) знайдено з коефіцієнтів даного многочлена за допомогою операцій додавання і віднімання і тому належать полю P. Теорему доведено. Справедлива також теорема про єдиність многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$.

Теорема 2. *Зображення симетричного многочлена у вигляді многочлена від основних симетричних функцій єдине.*

Доведення. Нехай маємо

$$f(x_1, x_2, \ldots, x_n) = \varphi_1(\sigma_1, \sigma_2, \ldots, \sigma_n),$$
$$f(x_1, x_2, \ldots, x_n) = \varphi_2(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

Тоді різниця

$$\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = \varphi_1(\sigma_1, \sigma_2, \ldots, \sigma_n) - \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

повинна дорівнювати нулю при будь-яких значеннях $x_1, x_2, \ldots, x_n$. Зауважимо, що многочлен $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ можна розглядати двояко: як многочлен від $x_1, x_2, \ldots, x_n$ (бо від цих змінних залежать $\sigma_1, \sigma_2, \ldots, \sigma_n$) і як многочлен від $\sigma_1, \sigma_2, \ldots, \sigma_n$; нам треба розглянути останнє. Єдиність зображення (9) полягає саме в тому, що многочлени, $\varphi_1(\sigma_1, \sigma_2, \ldots, \sigma_n)$ і $\varphi_2(\sigma_1, \sigma_2, \ldots, \sigma_n)$ мають однакові відповідні коефіцієнти, тобто що многочлен $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ має коефіцієнти A_i, які дорівнюють нулю, в усіх членах $A_i\sigma_1^{k_{1i}}\sigma_2^{k_{2i}} \ldots \sigma_n^{k_{ni}}$. Якби $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ був многочленом від незалежних змінних, рівність нулю всіх його коефіцієнтів випливала б з теореми 3 п. 25.2. Але σ_i залежні між собою, бо виражаються через ті самі змінні $x_1, x_2, \ldots, x_n$. У зв'язку з цим поряд з многочленом $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ від залежних змінних розглянемо такий самий многочлен $\varphi(y_1, y_2, \ldots, y_n)$ від незалежних змінних $y_1, y_2, \ldots, y_n$. Тепер нам треба довести, що коли $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0$, то й $\varphi(y_1, y_2, \ldots, y_n) = 0$. Те саме можна сформулювати й інакше: нам треба довести, що коли $\varphi(y_1, y_2, \ldots, y_n) \neq 0$, то тоді й $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) \neq 0$.

Доведемо це методом математичної індукції по n. Нехай $n = 1$ і $\varphi(y_1) \neq 0$. Через те, що σ_1 в цьому разі дорівнює x_1, то $\varphi(\sigma_1) \neq 0$, бо $\varphi(\sigma_1) = \varphi(x_1)$, що те саме, що й $\varphi(y_1)$.

Нехай тепер $n > 1$, і наше твердження правильне для будь-якого числа змінних, меншого n. Чи може бути воно несправедливим для якогось многочлена від n змінних? Припустимо, що це так і існує многочлен $\varphi(y_1, y_2, \ldots, y_n)$ такий, що $\varphi(y_1, y_2, \ldots, y_n) \neq 0$, але $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0$. Подамо $\varphi(y_1, y_2, \ldots, y_n)$ за степенями y_n

$$\varphi(y_1, y_2, \ldots, y_{n-1}, y_n) = \varphi_k(y_1, y_2, \ldots, y_{n-1})\, y_n^k +$$
$$+ \cdots + \varphi_1(y_1, y_2, \ldots, y_{n-1})\, y_n + \varphi_0(y_1, y_2, \ldots, y_{n-1}) \neq 0, \qquad (10)$$

де $\varphi_i(y_1, y_2, \ldots, y_{n-1})$ — многочлени від $y_1, y_2, \ldots, y_{n-1}$. З другого боку, за нашим припущенням

$$\varphi(\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \sigma_n) = \varphi_k(\sigma_1, \sigma_2, \ldots, \sigma_{n-1})\, \sigma_n^k +$$
$$+ \cdots + \varphi_1(\sigma_1, \sigma_2, \ldots, \sigma_{n-1})\, \sigma_n + \varphi_0(\sigma_1, \sigma_2, \ldots, \sigma_{n-1}) = 0. \qquad (11)$$

Оскільки $\varphi(y_1, y_2, \ldots, y_{n-1}, y_n) \neq 0$, то хоч би один з його коефіцієнтів в (10) не дорівнює нулю. Завжди можна вважати, що $\varphi_0(y_1, y_2, \ldots, y_{n-1}) \neq 0$. Якщо $\varphi_l = 0$ $(i < l)$, а $\varphi_l \neq 0$, то дальші міркування проводять відносно многочлена $\hat{\varphi}(y_1, y_2, \ldots, y_n)$, який дістаємо з $\varphi(y_1, \ldots, y_n)$ після скорочення на y_n^l. Виходить, що при $y_n = 0$

$$\varphi(y_1, y_2, \ldots, y_{n-1}, 0) = \varphi_0(y_1, y_2, \ldots, y_{n-1}) \neq 0. \qquad (12)$$

З другого боку, візьмемо в (11) $x_n = 0$. Тоді $\sigma_n = 0$, а інші σ_i перетворюються в основні симетричні функції від $n - 1$ змінних. Позначимо їх через $\sigma_1', \sigma_2', \ldots, \sigma_{n-1}'$. Отже, при $x_n = 0$ з (11) дістаємо:

$$\varphi(\sigma_1', \sigma_2', \ldots, \sigma_{n-1}', 0) = \varphi_0(\sigma_1', \sigma_2', \ldots, \sigma_{n-1}') = 0. \qquad (13)$$

Порівнюючи (12) з (13) бачимо, що ми прийшли до суперечності з припущенням індукції, а тому висловлене твердження справедливе і для n.

Єдиність зображення (9) доведено.

З основної теореми теорії симетричних многочленів можна зробити важливий висновок.

Нехай дано якийсь многочлен n-го степеня від одного змінного (в зведеному вигляді) над полем P:

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0. \qquad (14)$$

Позначимо корені цього многочлена через $\alpha_1, \alpha_2, \ldots, \alpha_n$; вони можуть і не належати полю P. Візьмемо тепер довільний симетричний многочлен $g(x_1, x_2, \ldots, x_n)$ над P від n змінних. За основною теоремою теорії симетричних многочленів, многочлен $g(x_1, x_2, \ldots, x_n)$ можна подати у вигляді многочлена від основних симетричних функцій $\sigma_1, \sigma_2, \ldots, \sigma_n$ з коефіцієнтами з поля P, тобто $g(x_1, x_2, \ldots, x_n) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$.

Візьмемо тепер тут $x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_n = \alpha_n$. Тоді за формулами Вієта всі основні симетричні функції дорівнюватимуть відповід-

ним коефіцієнтам многочлена (14) з належним знаком:

$$\sigma_1(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1 + \alpha_2 + \cdots + \alpha_n = -a_{n-1},$$
$$\sigma_2(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{n-1}\alpha_n = a_{n-2},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\sigma_n(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1\alpha_2 \ldots \alpha_n = (-1)^n a_0.$$

У зв'язку з цим

$$g(\alpha_1, \alpha_2, \ldots, \alpha_n) = \varphi(-a_{n-1}, a_{n-2}, \ldots, (-1)^n a_0).$$

Але тоді $\varphi(-a_{n-1}, a_{n-2}, \ldots, (-1)^n a_0)$ елемент поля P як результат виконання операцій додавання і множення над елементами з поля P. Таким чином, $g(\alpha_1, \alpha_2, \ldots, \alpha_n) \in P$. Отже, ми довели таке твердження.

Теорема 3. *Якщо $f(x)$ — многочлен від однієї змінної над полем P з коренями $\alpha_1, \alpha_2, \ldots, \alpha_n$ (які можуть не належати P), то будь-який симетричний многочлен $g(x_1, x_2, \ldots, x_n)$ над полем P при $x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_n = \alpha_n$ набуває значення, яке є елементом поля P.*

У ряді питань доводиться зустрічатися з задачею побудови за даним многочленом $f(x) \in P[x]$ з коренями $x_1, x_2, \ldots, x_n$ такого многочлена $g(y)$, корені якого y_i виражаються через відповідні корені x_i за допомогою деякого многочлена $y = \varphi(x)$ над полем P; $y_i = \varphi(x_i)$. Найпростіші задачі такого типу зустрічаються в шкільному курсі алгебри для $P = Q$. Оскільки коефіцієнти A_k многочлена $g(y)$ відповідно до формул Вієта визначаються рівностями

$$-A_{n-1} = \varphi(x_1) + \varphi(x_2) + \cdots + \varphi(x_n),$$
$$A_{n-2} = \varphi(x_1)\varphi(x_2) + \varphi(x_1)\varphi(x_3) + \cdots + \varphi(x_{n-1})\varphi(x_n),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(-1)^n A_0 = \varphi(x_1)\varphi(x_2) \ldots \varphi(x_n),$$

то вони є значеннями деяких симетричних многочленів над P, аргументи яких є коренями даного многочлена $f(x)$. З теореми 1 випливає, що завжди можна знайти вираз коефіцієнтів A_k через коефіцієнти даного многочлена, а з теореми 3 зрозуміло, що знайдений многочлен належатиме тому самому кільцю $P[x]$, що й даний многочлен.

Зауважимо, що сказане залишається справедливим і для більш загального випадку, коли $y_i = \varphi_i(x_1, x_2, \ldots, x_n)$, де $\varphi_i(x_1, x_2, \ldots, x_n)$ — довільні симетричні многочлени над полем P.

Розглянутий вище метод доведення основної теореми можна використати для практичного зображення симетричних многочленів через основні симетричні функції.

Приклад. Нехай треба подати симетричний многочлен над полем Q

$$f(x_1, x_2, x_3) = x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 - 4(x_1^2 + x_2^2 + x_3^2) + 5$$

через основні симетричні функції. Як і при доведенні теореми, запишемо цей многочлен як суму однорідних многочленів. Дістанемо:

$$f(x_1, x_2, x_3) = \varphi_1(x_1, x_2, x_3) - 4\varphi_2(x_1, x_2, x_3) + 5,$$

де $\varphi_1(x_1, x_2, x_3) = x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2$,

$$\varphi_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2.$$

Спочатку $\varphi_1(x_1, x_2, x_3)$ подамо через основні симетричні многочлени. Вищий його член є $x_1^2x_2$. Згідно з методикою доведення теореми, від $\varphi_1(x_1, x_2, x_3)$ слід відняти многочлен

$$g(x_1, x_2, x_3) = \sigma_1^{2-1}\sigma_2^{1-0}\sigma_3^0 = \sigma_1\sigma_2,$$

бо система показників у вищому члені є 2, 1, 0. Але немає потреби фактично виконувати це віднімання. Спираючись на можливість і єдиність зображення даного многочлена у вигляді многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$, досить визначити можливий вигляд членів $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ і скористатися методом невизначених коефіцієнтів.

У різниці $\varphi_1(x_1, x_2, x_3) - g(x_1, x_2, x_3)$ знищаться всі члени виду $Ax_1^{l_1}x_2^{l_2}x_3^{l_3}$ з довільною перестановкою показників 2, 1, 0. Проте одночасно можуть з'явитися члени того самого степеня 3, але з іншою, нижчою системою показників, а саме: 1, 1, 1. Отже, потім треба буде відняти симетричний многочлен

$$g_1(x_1, x_2, x_3) = \sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^1 = \sigma_3.$$

Тому можна записати:

$$\varphi_1(x_1, x_2, x_3) = \sigma_1\sigma_2 + a\sigma_3,$$

де a — невизначений поки що коефіцієнт, тобто:

$$x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 =$$
$$= (x_1 + x_2 + x_3)(x_1x_2 + x_1x_3 + x_2x_3) + ax_1x_2x_3.$$

Щоб знайти a, досить надати деяких числових значень змінним x_1, x_2, x_3, наприклад $x_1 = x_2 = x_3 = 1$. Тоді дістанемо $6 = 9 + a$. Отже, $a = -3$. Таким чином,

$$\varphi_1(x_1, x_2, x_3) = \sigma_1\sigma_2 - 3\sigma_3.$$

Аналогічно міркуватимемо відносно многочлена

$$\varphi_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2.$$

Можливі системи показників тут будуть 2, 0, 0 і 1, 1, 0. Отже, відніматимемо такі многочлени:

$$g_2(x_1, x_2, x_3) = \sigma_1^{2-0}\sigma_2^{0-0}\sigma_3^0 = \sigma_1^2,$$
$$g_3(x_1, x_2, x_3) = \sigma_1^{1-1}\sigma_2^{1-0}\sigma_3^0 = \sigma_2$$

і далі, аналогічно до попереднього, $\varphi_2(x_1x_2, x_3) = \sigma_1^2 + b\sigma_2$. При $x_1 = x_2 = x_1 = 1$ маємо $3 = 3^2 + b \cdot 3$, тобто $b = -2$ і тому

$$\varphi_2(x_1, x_2, x_3) = \sigma_1^2 - 2\sigma_2. \qquad (15)$$

Отже, дістаємо остаточно

$$f(x_1, x_2, x_3) = \sigma_1\sigma_2 - 3\sigma_3 - 4(\sigma_1^2 - 2\sigma_2) + 5.$$

На практиці розрахунки зручно подавати у формі таблиці. Приклад такого запису буде наведено нижче.

Теорія симетричних многочленів має досить широке застосування в алгебрі. Використовуючи основну теорему теорії симетричних многочленів, можна довести основну теорему теорії многочленів. Ця сама теорема застосовується в загальній теорії розв'язування алгебраїчних рівнянь у радикалах. У наступному параграфі ми використаємо її при розв'язуванні систем алгебраїчних рівнянь.

27.1. Вступні зауваження. Вивчення властивостей многочленів $f(x)$ від однієї змінної безпосередньо пов'язане з розв'язуванням алгебраїчних рівнянь виду $f(x) = 0$, адже корені цього рівняння і корені многочлена $f(x)$ — це одне й те саме.

У випадку многочлена $f(x_1, x_2, \ldots, x_n)$ від кількох змінних можна також розглядати відповідне алгебраїчне рівняння

$$f(x_1, x_2, \ldots, x_n) = 0. \tag{1}$$

Замість термі́на к о р і́ н ь многочлена $f(x_1, x_2, \ldots, x_n)$ або рівняння (1) вживають термін р о з в' я з о к. А саме, якщо $f(x_1, x_2, \ldots, x_n) \in P[x_1, x_2, \ldots, x_n]$, а Δ — будь-яке розширення поля P, то *розв'язком рівняння* (1) *або многочлена* $f(x_1, x_2, \ldots, x_n)$ *називають упорядковану систему* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *елементів поля* Δ *таку, що при* $x_1 = \alpha_1$, $x_2 = \alpha_2, \ldots, x_n = \alpha_n$ *відповідне значення многочлена* $f(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *дорівнює нулю.*

Якщо P — поле характеристики 0, то легко встановити існування розв'язку будь-якого рівняння виду (1) в P або у деякому його розширенні (при $\deg f \geqslant 1$). Для цього досить подати $f(x_1, x_2, \ldots, x_n)$ за степенями однієї із змінних (відносно якої степінь многочлена виявиться ненульовим), скажімо, x_n,

$$f(x_1, x_2, \ldots, x_{n-1}, x_n) = a_m(x_1, \ldots, x_{n-1})\,x_n^m + \cdots + a_1(x_1, \ldots, x_{n-1})\,x_n + a_0(x_1, \ldots, x_{n-1}), \quad m \geqslant 1, \tag{2}$$

вибрати в P довільні значення $x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_{n-1} = \alpha_{n-1}$, при яких $a_m(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}) \neq 0$ (такі значення обов'язково знайдуться, бо $a_m(x_1, x_2, \ldots, x_{n-1}) \not\equiv 0$), підставити їх у коефіцієнти многочлена (2) і дістати многочлен

$$a_m(\alpha_1, \ldots, \alpha_{n-1})\,x_n^m + \cdots + a_1(\alpha_1, \ldots, \alpha_{n-1})\,x_n + a_0(\alpha_1, \ldots, \alpha_n)$$

від однієї змінної степеня $m \geqslant 1$ над полем P. За теоремою Кронекера (п. 23.3), цей многочлен має корінь α_n у полі P або у деякому його розширенні. Але тоді $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \alpha_n$ — розв'язок рівняння (1).

У цьому параграфі вважатимемо, що усі розглядувані многочлени задані над полем характеристики 0.

Як видно з наведеного вище міркування, будь-який многочлен ненульового степеня від n змінних ($n \geqslant 2$) має розв'язки. Але з того ж міркування випливає, що таких розв'язків є б е з л і́ ч: адже вибрати $\alpha_1, \alpha_2, \ldots, \alpha_{n-1} \in P$ так, щоб $a_m(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}) \neq 0$ можна нескінченним числом способів. Таким чином, задача розв'язування рівнянь типу (1) при $n \geqslant 2$ (на відміну від випадку $n = 1$) є по суті неозначеною.

Більш природною, поширеною і практично застосовною є задача

розв'язування с и с т е м алгебраїчних рівнянь виду

$$\begin{cases} f_1(x_1, \ldots, x_n) = 0, \\ f_2(x_1, \ldots, x_n) = 0, \\ \ldots\ldots\ldots \\ f_m(x_1, \ldots, x_n) = 0, \end{cases} \tag{3}$$

тобто знаходження спільних розв'язків усіх многочленів $f_k(x_1, \ldots, x_n)$ $(k = 1, 2, \ldots, m)$.

Ми докладно розглядали розв'язування систем л і н і й н и х рівнянь (1, розд. VI), тобто систем виду (3), для яких $\deg f_k = 1$ $(k = 1, 2, 3, \ldots, m)$. У випадку ж многочленів $f_k(x_1, \ldots, x_n)$ довільних степенів побудувати загальну теорію досить важко. В основі існуючих методів розв'язування системи (3) лежить ідея в и к л ю ч е н н я н е в і́ д о м и х, метою якої є зведення задачі до розв'язування о д н о г о алгебраїчного рівняння з одним невідомим. Саме тому цей розділ алгебри називають *теорією виключення.*

Пояснимо цю ідею на простому прикладі з курсу алгебри середньої школи.

П р и к л а д. Потрібно знайти розв'язки системи двох рівнянь з двома невідомими:

$$f(x, y) \underset{\text{df}}{=} x^2 + y^2 - a = 0, \quad g(x, y) \underset{\text{df}}{=} xy - b = 0. \tag{4}$$

Щоб знайти розв'язки системи (4), виключимо з двох даних рівнянь якимсь способом одне з невідомих і побудуємо вивідне рівняння. Наприклад, визначимо x з другого рівняння і підставимо знайдене значення в перше рівняння. Матимемо: $x^2 = \frac{b}{y}$; $\frac{b^2}{y^2} + y^2 - a = 0$, або

$$y^4 - ay^2 + b^2 = 0. \tag{5}$$

Отже, розв'язування системи двох рівнянь з двома невідомими ми звели до розв'язування одного рівняння (5) з одним невідомим.

У цьому випадку вивідне рівняння біквадратне і його можна розв'язати елементарно. Коренями рівняння (5) є числа:

$$\beta_{1,2} = \pm\sqrt{\tfrac{1}{2}(a + \sqrt{a^2 - 4b^2})}; \quad \beta_{3,4} = \pm\sqrt{\tfrac{1}{2}(a - \sqrt{a^2 - 4b^2})};$$

Підставляючи замість y кожне із значень β_i в одне з рівнянь системи (4), матимемо відповідне значення α_i невідомого x:

$$\alpha_{1,2} = \pm\frac{b}{\sqrt{\frac{1}{2}(a + \sqrt{a^2 - 4b^2})}}; \quad \alpha_{3,4} = \pm\frac{b}{\sqrt{\frac{1}{2}(a - \sqrt{a^2 - 4b^2})}}.$$

Хід розв'язування системи (4) можна подати так. Назвемо ліву частину рівняння (5), знайденого в результаті виключення одного невідомого з рівнянь системи (4), р е з у л ь т а н т о м і позначимо його символом $R(f, g)$. Отже,

$$R(f, g) \underset{\text{df}}{=} y^4 - ay^2 + b^2.$$

Тоді розв'язання системи (4), як ми бачили, зводиться до такого. Для заданих многочленів $f(x, y)$ і $g(x, y)$ будують результант $R(f, g)$, який є многочленом від однієї змінної. Далі знаходять корені результанта. Нарешті, знаючи ці корені, обчислюють розв'язки даної системи.

Якщо рівняння системи довільного степеня, то хід міркувань залишається той самий. Розв'язування знову зводиться до побудови результанта многочленів $f(x, y)$ і $g(x, y)$ і знаходжеиня його коренів. При цьому, очевидно, відразу ж виникають такі питання: чи існує результант для довільних двох многочленів; як побудувати результант практично; який зв'язок між розв'язками даної системи і коренями результанта; як розв'язуються системи не з двома, а з більшим числом невідомих?

Спробуємо дати відповідь на ці питання.

27.2. Результант. Нехай задано систему двох алгебраїчних рівнянь з двома невідомими. Розмістимо члени цих рівнянь за степенями одного з невідомих (їх ліві частини розглядаються над полем P):

$$\begin{cases} f(x, y) \underset{\text{df}}{=} a_n(y)\,x^n + a_{n-1}(y)\,x^{n-1} + \cdots + a_1(y)\,x + a_0(y) = 0, \\ g(x, y) \underset{\text{df}}{=} b_m(y)\,x^m + b_{m-1}(y)\,x^{m-1} + \cdots + b_1(y)\,x + b_0(y) = 0, \end{cases} \tag{6}$$

і припустимо, що пара елементів α, β є розв'язком системи (6), тобто $f(\alpha, \beta) = 0$, $g(\alpha, \beta) = 0$.

Зрозуміло, що з системи (6) можна утворити цілий ряд вивідних рівнянь, для яких (α, β) також буде розв'язком. Природно поставити питання, чи не можна побудувати таку систему вивідних рівнянь, яку було б порівняно легко розв'язати і тим самим знайти розв'язки і заданої системи. Звичайно, бажано, щоб вивідні рівняння, крім потрібних нам розв'язків, не мали б інших, зайвих розв'язків. Доцільно було б побудувати з системи (6) таке вивідне рівняння, яке містило б лише о д н е з невідомих (як кажуть, в и к л ю ч и т и друге невідоме), і корені якого визначали б розв'язки заданої системи.

Підставляючи в рівняння системи (6), що має розв'язок (α, β), значення $y = \beta$, дістанемо два рівняння з одним невідомим x:

$$\begin{cases} f(x, \beta) = a_n(\beta)\,x^n + a_{n-1}(\beta)\,x^{n-1} + \cdots + a_1(\beta)\,x + a_0(\beta) = 0, \\ g(x, \beta) = b_m(\beta)\,x^m + b_{m-1}(\beta)\,x^{m-1} + \cdots + b_1(\beta)\,x + b_0(\beta) = 0, \end{cases} \tag{7}$$

причому $x = \alpha$ є спільним коренем обох цих рівнянь. Оскільки два довільно взяті рівняння з одним невідомим, взагалі кажучи, спільних коренів не мають, то рівняння системи (7), які мають спільний корінь, не можуть бути незалежними. Між їх коефіцієнтами повинен бути деякий зв'язок. Якщо ми знайдемо цей зв'язок між коефіцієнтами $a_i(\beta)$, $b_j(\beta)$, тобто співвідношення

$$R[a_n(\beta),\ a_{n-1}(\beta),\ \ldots,\ a_1(\beta),\ a_0(\beta);$$
$$b_m(\beta),\ b_{m-1}(\beta),\ \ldots,\ b_1(\beta),\ b_0(\beta)] = 0,$$

то тим самим дістанемо деяке рівняння

$$R[a_n(y),\ \ldots,\ a_1(y),\ a_0(y),\ b_m(y),\ \ldots,\ b_1(y),\ b_0(y)] = 0,$$

яке повинно задовольнятись при $y = \beta$, щоб пара (α, β) могла бути розв'язком системи.

Отже, насамперед треба розв'язати таку задачу. Нехай дано систему двох рівнянь з одним невідомим:

$$\begin{cases} f(x) \underset{\text{df}}{=} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0, \\ g(x) \underset{\text{df}}{=} b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 = 0. \end{cases} \tag{8}$$

Знайти, *при яких умовах ці рівняння можуть мати спільний корінь.*

Зауважимо, що між системами (8) і (7) є певна відмінність. При розгляді системи (8) природно вважати, що $a_n \neq 0$ і $b_m \neq 0$, тоді як у системі (7), утвореній з (6) при $y = \beta$, деякі з коефіцієнтів і, зокрема старші коефіцієнти $a_n(\beta)$ і $b_m(\beta)$, можуть дорівнювати нулю, хоч відповідні многочлени $a_n(y)$ і $b_m(y)$ не були нулями. Це зауваження ми використаємо пізніше.

Повернемося до системи (8). Очевидно, що с п і л ь н і корені рівнянь системи треба шукати лише серед коренів многочлена $f(x)$. Позначимо ці корені через $\alpha_1, \alpha_2, \ldots, \alpha_n$. З другого боку, α_i лише тоді буде спільним коренем, якщо $g(\alpha_i) = 0$.

Означення. *Результантом многочленів*

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \\ g(x) &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0, \end{aligned} \qquad (a_n \neq 0,\ b_m \neq 0) \tag{9}$$

називається вираз

$$R(f, g) \underset{\text{df}}{=} a_n^m g(\alpha_1)\,g(\alpha_2)\,\ldots\,g(\alpha_n), \tag{10}$$

де $\alpha_1, \alpha_2, \ldots, \alpha_n$ *— корені многочлена* $f(x)$.

З а у в а ж е н н я 1. У цьому означенні многочлени $f(x)$ і $g(x)$ нерівноправні; для $R(g, f)$. Дістаємо:

$$R(g, f) = b_m^n f(\gamma_1)\,f(\gamma_2)\,\ldots\,f(\gamma_m), \tag{11}$$

де $\gamma_1, \gamma_2, \ldots, \gamma_m$ — корені многочлена $g(x)$. На перший погляд здається, що це є недоліком введеного означення (бо в поставленій задачі $f(x)$ і $g(x)$ рівноправні), проте нижче буде показано, що $R(f, g)$ і $R(g, f)$ можуть відрізнятися лише знаком.

З а у в а ж е н н я 2. В означенні $R(f, g)$ ми користуємося тим, що $a_n \neq 0$, бо вважаємо, що $f(x)$ є многочленом саме n-го степеня і має n коренів; при $a_n = 0$ означення в цій формі втрачає смисл, бо не може бути многочлена з n кореняими $\alpha_1, \alpha_2, \ldots, \alpha_n$ і старшим коефіцієнтом $a_n = 0$. Аналогічно, при означенні $R(g, f)$ ми спираємося на те, що $b_m \neq 0$. Отже, щоб означити для двох многочленів один з результантів $R(f, g)$ або $R(g, f)$, потрібно, щоб х о ч о д и н з старших членів цих многочленів був відмінний від нуля.

На перший погляд здається, що для побудови результанта двох многочленів треба знати корені одного з многочленів. Але в дійсності результант виражається (і притому раціонально) через коефіцієнти даних многочленів.

Справді, $R(f, g)$ є с и м е т р и ч н и м многочленом від $\alpha_1, \alpha_2, \ldots, \alpha_n$, коефіцієнти якого раціонально виражаються через $a_n, b_m, b_{m-1}, \ldots, b_0$. З теореми 1 п. 26.2 випливає, що $R(f, g)$ можна подати як

многочлен від основних симетричних функцій

$$\sigma_1 = \alpha_1 + \alpha_2 + \dots + \alpha_n,$$
$$\sigma_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \dots + \alpha_{n-1}\alpha_n,$$
$$\dots\dots\dots\dots\dots\dots$$
$$\sigma_n = \alpha_1\alpha_2 \dots \alpha_n.$$

Оскільки $\alpha_1, \alpha_2, \dots, \alpha_n$ — корені многочлена $f(x)$, то $\sigma_1, \sigma_2, \dots, \sigma_n$, за формулами Вієта, раціонально виражаються через його коефіцієнти $a_n, a_{n-1}, \dots, a_1, a_0$. Отже, остаточно $R(f, g)$ можна раціонально виразити через коефіцієнти a_i і b_j обох заданих многочленів. Звідси, зокрема, випливає, що *результант довільних двох многочленів над полем P є елементом цього самого поля*. Цей самий висновок можна дістати з теореми 3 п. 26.2.

Із сказаного зрозуміло, як можна практично обчислювати результант двох многочленів.

Приклад 3. Знайдемо результант двох многочленів 2-го степеня

$$f(x) = a_2x^2 + a_1x + a_0, \quad g(x) = b_2x^2 + b_1x + b_0.$$

Якщо позначити корені $f(x)$ через α_1 і α_2, то

$$R(f, g) = a_2^2 g(\alpha_1)\, g(\alpha_2) = a_2^2 [b_2\alpha_1^2 + b_1\alpha_1 + b_0][b_2\alpha_2^2 + b_1\alpha_2 + b_0].$$

Перемножаючи вирази в дужках, матимемо:

$$R(f, g) = a_2^2 [b_2^2\alpha_1^2\alpha_2^2 + b_1b_2(\alpha_1 + \alpha_2)\alpha_1\alpha_2 + b_1^2\alpha_1\alpha_2 + b_0b_2(\alpha_1^2 + \alpha_2^2) + b_0b_1(\alpha_1 + \alpha_2) + b_0^2].$$

А через те що за формулами Вієта $\alpha_1 + \alpha_2 = -\frac{a_1}{a_2}$, $\alpha_1\alpha_2 = \frac{a_0}{a_2}$, то

$$R(f, g) = a_0^2b_2^2 - a_0a_1b_1b_2 + a_0a_2b_1^2 + (a_1^2 - 2a_0a_2)b_0b_2 - a_1a_2b_0b_1 + a_2^2b_0^2.$$

Розглянемо тепер деякі властивості результанта, які випливають з означення.

1. $$R(f, g) = a_n^m b_m^n \prod_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}} (\alpha_i - \gamma_j). \tag{12}$$

Справді, оскільки $g(x) = b_m(x - \gamma_1)(x - \gamma_2) \dots (x - \gamma_m)$, то

$$g(\alpha_i) = b_m(\alpha_i - \gamma_1)(\alpha_i - \gamma_2) \dots (\alpha_i - \gamma_m) = b_m \prod_{(1 \leqslant j \leqslant m)} (\alpha_i - \gamma_j);$$

тому

$$R(f, g) = a_n^m g(\alpha_1)\, g(\alpha_2) \dots g(\alpha_n) = a_n^m [b_m \prod_{(1 \leqslant j \leqslant m)} (\alpha_1 - \gamma_j)] \times$$
$$\times [b_m \prod_{(1 \leqslant j \leqslant m)} (\alpha_2 - \gamma_j)] \dots [b_m \prod_{(1 \leqslant j \leqslant m)} (\alpha_n - \gamma_j)] =$$
$$= a_n^m b_m^n \prod_{(1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant m)} (d_i - \gamma_j),$$

що й треба було довести.

2. $$R(g, f) = (-1)^{mn} R(f, g). \tag{13}$$

Застосовуючи формулу (12) до результанта $R(g, f)$, тобто міняючи ролі $f(x)$ і $g(x)$, маємо:

$$R(g, f) = b_m^n a_n^m \prod_{(1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant m)} (\gamma_j - \alpha_i).$$

Змінимо знаки в дужках на протилежні, тобто винесемо в кожному множнику за дужки число — 1. Ураховуючи, що число множників дорівнює mn, матимемо:

$$R(g, f) = (-1)^{mn} b_m^n a_n^m \prod_{(1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant m)} (\alpha_i - \gamma_j) = (-1)^{mn} R(f, g).$$

Формулу (13) доведено.

Значення результанта для поставленої нами задачі про умови існування розв'язків системи (8) визначається такою теоремою.

Теорема 1. *Для того щоб многочлени $f(x)$ і $g(x)$ мали спільний корінь, необхідно і достатньо, щоб їх результант дорівнював нулю.*

Доведення. *Необхідність.* Якщо якийсь корінь α_i многочлена $f(x)$ є коренем і многочлена $g(x)$, то $g(\alpha_i) = 0$. Тоді в результанті

$$R(f, g) = a_n^m g(\alpha_1)\, g(\alpha_2) \dots g(\alpha_i) \dots g(\alpha_n)$$

один з множників дорівнює нулю і тому $R(f, g) = 0$.

Достатність. Якщо $R(f, g) = 0$, тобто $a_n^m g(\alpha_1)\, g(\alpha_2) \dots g(\alpha_i) \dots g(\alpha_n) = 0$, то хоча б один з множників у лівій частині цієї рівності дорівнює нулю. Через те що $a_n \neq 0$, то дорівнює нулю хоча б один з інших множників, наприклад $g(\alpha_i)$. Але якщо $g(\alpha_i) = 0$, то корінь α_i многочлена $f(x)$ є одночасно і коренем многочлена $g(x)$, тобто задані многочлени мають спільний корінь.

Звернемо увагу на те, що при доведенні теореми 1 ми істотно використали нерівність нулю коефіцієнта a_n. Якби $a_n = 0$, а $b_m \neq 0$, то теорема була б справедливою, але довелося б розглядати результант $R(g, f)$. Отже, теорема 1 застосовна лише при умові, коли хоча б один з старших коефіцієнтів даних многочленів відмінний від нуля.

27.3. Дискримінант. Поняття результанта і теорему 1 можна безпосередньо застосувати до розв'язання питання про наявність кратних коренів многочлена. З теореми 8 п. 23.5 випливає, що кратний корінь многочлена $f(x)$ повинен бути спільним коренем многочлена $f(x)$ і його похідної $f'(x)$ (а також і похідних вищих порядків, коли кратність кореня $k > 2$).

Нехай дано многочлен $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$. Позначимо його корені через $\alpha_1, \alpha_2, \dots, \alpha_n$. Очевидно, що

$$f(x) = a_n(x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n). \tag{14}$$

Результантом многочленів $f(x)$ і $f'(x)$, згідно з означеннями, є вираз

$$R(f, f') = a_n^{n-1} f'(\alpha_1) f'(\alpha_2) \dots f'(\alpha_n). \tag{15}$$

Щоб обчислити цей вираз, знайдемо похідну від многочлена $f(x)$, записаного у вигляді (14). Маємо:

$$f'(x) = a_n[(x - \alpha_2)(x - \alpha_3) \dots (x - \alpha_i) \dots (x - \alpha_n) +$$

$$+(x-\alpha_1)(x-\alpha_3)\quad\ldots(x-\alpha_i)\ldots(x-\alpha_n)+$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+(x-\alpha_1)\ldots(x-\alpha_{i-1})(x-\alpha_{i+1})\ldots(x-\alpha_n)+$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+(x-\alpha_1)(x-\alpha_2)\quad\ldots(x-\alpha_i)\quad\ldots(x-\alpha_{n-1})\}.$$

Звідси випливає, що

$$f'(\alpha_i)=a_n(\alpha_i-\alpha_1)(\alpha_i-\alpha_2)\ldots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\ldots(\alpha_i-\alpha_n).$$

Отже,

$$R(f,f')=a_n^{n-1}f'(\alpha_1)f'(\alpha_2)\ldots f'(\alpha_n)=$$
$$=a_n^{n-1}[a_n(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)\ldots(\alpha_1-\alpha_n)]\times$$
$$\times[a_n(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)\ldots(\alpha_2-\alpha_n)]\times\ldots$$
$$\ldots\times[a_n(\alpha_n-\alpha_1)(\alpha_n-\alpha_2)\ldots(\alpha_n-\alpha_{n-1})].$$

Легко зрозуміти, що в знайдений добуток кожний множник виду $\alpha_i-\alpha_j$ $(i\neq j)$ входить двічі, причому з протилежними знаками (наприклад, $\alpha_1-\alpha_2$ і $\alpha_2-\alpha_1$, або $\alpha_1-\alpha_n$ і $\alpha_n-\alpha_1$). Кожний з добутків $(\alpha_i-\alpha_j)(\alpha_j-\alpha_i)$ можна записати у вигляді $-(\alpha_i-\alpha_j)^2$, де $i>j$.

Підрахувавши число таких квадратів, легко встановити, що повинен з'явитися множник $(-1)^{\frac{n(n-1)}{2}}$. Отже,

$$R(f,f')=(-1)^{\frac{n(n-1)}{2}}a_n^{2n-1}\prod_{(1\leqslant j<i\leqslant n)}(\alpha_i-\alpha_j)^2. \tag{16}$$

Введемо тепер поняття дискримінанта многочлена.

Означення. *Дискримінантом $D(f)$ многочлена $f(x)$ називається вираз*

$$D(f)\underset{\text{df}}{=}(-1)^{\frac{n(n-1)}{2}}a_n^{-1}R(f,f'), \tag{17}$$

де $R(f,f')$ — результант многочлена $f(x)$ і його похідної $f'(x)$.

Згідно з цим означенням і на основі (16) дістаємо:

$$D(f)=a_n^{2n-2}\prod_{(1\leqslant j<i\leqslant n)}(\alpha_i-\alpha_j)^2. \tag{18}$$

Ми бачимо, що дискримінант лише сталим множником, підібраним так, щоб була справедлива зручна рівність (18), відрізняється від результанта многочлена і його похідної. Це дає змогу з теореми 1 відразу дістати таке твердження:

Теорема 2. *Многочлен $f(x)$ має кратний корінь тоді і тільки тоді, коли його дискримінант дорівнює нулю.*

Справді, якщо $f(x)$ має спільний корінь з $f'(x)$, то $R(f,f')=0$, а отже, і $D(f)$ дорівнює нулю. Якщо ж, навпаки, дискримінант $D(f)=0$, то це означає, що дорівнює нулю і $R(f,f')$, тобто $f(x)$ і $f'(x)$ мають спільний корінь.

Очевидно, що теорему 2 можна використовувати як ознаку наявності кратних коренів у многочлена $f(x)$.

Розглянемо кілька прикладів.

Приклади. **1.** Знайдемо дискримінант квадратного тричлена $f(x)=a_2x^2+a_1x+a_0$, корені якого позначимо, як і раніше, через α_1, α_2.

Через те, що $f'(x)=2a_2x+a_1$, то, згідно з (17) і (15), маємо:

$$D(f)=(-1)^{\frac{2\cdot1}{2}}a_2^{-1}R(f,f')=-a_2^{-1}a_2(2a_2\alpha_1+a_1)(2a\alpha_2+a_1)=$$
$$=-[4a_2^2\alpha_1\alpha_2+2a_1a_2(\alpha_1+\alpha_2)+a_1^2].$$

Оскільки $\alpha_1+\alpha_2=-\frac{a_1}{a_2}$, а $\alpha_1\alpha_2=\frac{a_0}{a_2}$, то

$$D(f)=-[4a_2a_0-2a_1^2+a_1^2]=a_1^2-4a_2a_0.$$

Отже, знайдений дискримінант квадратного тричлена повністю збігається з дискримінантом, введеним ще в середній школі. Шкільний курс алгебри містить і відповідне до теореми 2 твердження: *якщо дискримінант $a_1^2-4a_2a_0=0$, то квадратний тричлен має два однакові корені.*

2. Знайдемо дискримінант кубічного тричлена $f(x)=x^3+px+q$.

Позначимо його корені через α_1, α_2, α_3. Аналогічно попередньому маємо:

$$D(f)=(-1)^{\frac{3\cdot2}{2}}f'(\alpha_1)f'(\alpha_2)f'(\alpha_3)=-(3\alpha_1^2+p)(3\alpha_2^2+p)(3\alpha_3^2+p)=$$
$$=-[27\alpha_1^2\alpha_2^2\alpha_3^2+9p(\alpha_1^2\alpha_2^2+\alpha_1^2\alpha_3^2+\alpha_2^2\alpha_3^2)+3p^2(\alpha_1^2+\alpha_2^2+\alpha_3^2)+p^3].$$

Через те, що $\alpha_1^2\alpha_2^2+\alpha_1^2\alpha_3^2+\alpha_2^2\alpha_3^2=(\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)^2-2\alpha_1\alpha_2\alpha_3(\alpha_1+\alpha_2+\alpha_3)$, $\alpha_1^2+\alpha_2^2+\alpha_3^2=(\alpha_1+\alpha_2+\alpha_3)^2-2(\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)$, і, крім того, $\alpha_1+\alpha_2+\alpha_3=0$; $\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3=p$; $\alpha_1\alpha_2\alpha_3=-q$, то

$$D(f)=-[27q^2+4p^3].$$

Знайдений дискримінант відрізняється лише множником $(-27\cdot4)$ від того дискримінанта, який ми ввели для кубічного рівняння (1, § 17), причому і в елементарній теорії ознакою існування кратних коренів є те, що дискримінант дорівнює нулю.

З наведених прикладів випливає, що відомі означення дискримінантів для квадратного і кубічного рівнянь є окремими випадками загального поняття дискримінанта. Звернемо увагу на те, що в цих окремих випадках дискримінантом ми називали, власне кажучи, певний підкореневий вираз у формулах коренів відповідних рівнянь. У зв'язку з цим могло скластися враження, що для рівнянь вище четвертого степеня (які, як відомо, не розв'язуються в радикалах) поняття дискримінаита ввести не можна. Однак тепер зрозуміло, що дискримінант є поняттям, яке узагальнюється на многочлени довільного степеня.

27.4. Результат у формі Сільвестра. При практичному обчисленні результанта двох многочленів часто користуються іншими зображеннями результанта, які не потребують застосування теорії симетричних многочленів.

Розглянемо одне з таких зображень результанта у вигляді так званого детермінанта Сільвестра [1].

[1] Д. Сільвестр (1814—1897) — відомий англійський математик.

У дальшому викладі теорію лінійних рівнянь і детермінантів, яку в першій частині курсу було подано для випадку числових коефіцієнтів, буде застосовано до рівнянь, коефіцієнти яких належать довільному полю P характеристики 0. Адже згадані теорії при цьому повністю залишаються в силі.

Нехай знову дано многочлени

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$
$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0,$$

коренями яких відповідно є елементи $\alpha_1, \alpha_2, \ldots, \alpha_n$ і $\gamma_1, \gamma_2, \ldots, \gamma_m$. Для конкретності вважатимемо, що $n > m$. Побудуємо тепер за многочленами $f(x)$ і $g(x)$ систему $m + n$ таких многочленів:

$$x^{m-1}f(x);\ x^{m-2}f(x);\ \ldots;\ xf(x);\ f(x);$$
$$x^{n-1}g(x);\ x^{n-2}g(x);\ \ldots;\ xg(x);\ g(x).$$

Очевидно, матимемо такі $n + m$ рівностей:

$$\begin{aligned}
x^{m-1}f(x) &= a_n x^{n+m-1} + a_{n-1}x^{n+m-2} + \cdots + a_1 x^m + a_0 x^{m-1},\\
x^{m-2}f(x) &= a_n x^{n+m-2} + \cdots + a_1 x^{m-1} + a_0 x^{m-2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
xf(x) &= a_n x^{n+1} + a_{n-1}x^n + \cdots + a_1 x^2 + a_0 x,\\
f(x) &= a_n x^n + \cdots + a_1 x + a_0,\\
x^{n-1}g(x) &= b_m x^{n+m-1} + b_{m-1}x^{n+m-2} + \cdots + b_1 x^n + b_0 x^{n-1},\\
x^{n-2}g(x) &= b_m x^{n+m-2} + \cdots + b_1 x^{n-1} + b_0 x^{n-2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
xg(x) &= b_m x^{m+1} + b_{m-1}x^m + \cdots + b_1 x^2 + b_0 x,\\
g(x) &= b_m x^m + \cdots + b_1 x + b_0.
\end{aligned}$$

Перенесемо в цих рівностях усі члени в одну сторону і візьмемо $x = \alpha_i$, де α_i — один з коренів многочлена $f(x)$. Ураховуючи, що $f(\alpha_i) = 0$, матимемо:

$$\begin{aligned}
a_n\alpha_i^{n+m-1} + a_{n-1}\alpha_i^{n+m-2} + \cdots + a_1\alpha_i^m + a_0\alpha_i^{m-1} &= 0,\\
a_n\alpha_i^{n+m-2} + \cdots + a_1\alpha_i^{m-1} + a_0\alpha_i^{m-2} &= 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
a_n\alpha_i^{n+1} + a_{n-1}\alpha_i^n + \cdots + a_1\alpha_i^2 + a_0\alpha_i &= 0,\\
a_n\alpha_i^n + \cdots + a_1\alpha_i + a_0 &= 0,\\
b_m\alpha_i^{n+m-1} + b_{m-1}\alpha_i^{n+m-2} + \cdots + b_1\alpha_i^n + [b_0 - g(\alpha_i)]\,\alpha_i^{n-1} &= 0,\\
b_m\alpha_i^{n+m-2} + \cdots + b_1\alpha_i^{n-1} + [b_0 - g(\alpha_i)]\,\alpha_i^{n-2} &= 0, \qquad (19)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
b_m\alpha_i^{m+1} + b_{m-1}\alpha_i^m + \cdots + b_1\alpha_i^2 + [b_0 - g(\alpha_i)]\,\alpha_i &= 0,\\
b_m\alpha_i^m + \cdots + b_1\alpha_i + [b_0 - g(\alpha_i)] &= 0.
\end{aligned}$$

Рівності (19) можна розглядати як результат підстановки деякого розв'язку в систему $n + m$ однорідних лінійних рівнянь з $n + m$ невідомими. Справді,

$$\left\{\begin{aligned}
a_n x_1 + a_{n-1}x_2 + \cdots + a_1 x_n + a_0 x_{n+1} &= 0,\\
a_n x_2 + \cdots + a_1 x_{n+1} + a_0 x_{n+2} &= 0,\\
&\cdots\cdots\cdots\cdots\\
a_n x_{m-1} + a_{n-1}x_m + \cdots + a_0 x_{n+m-1} &= 0,\\
a_n x_m + \cdots + a_1 x_{n+m-1} + a_0 x_{n+m} &= 0,\\
b_m x_1 + b_{m-1}x_2 + \cdots + b_1 x_m + [b_0 - g(\alpha_i)]\,x_{m+1} &= 0,\\
b_m x_2 + \cdots + b_1 x_{m+1} + [b_0 - g(\alpha_i)]\,x_{m+2} &= 0,\\
&\cdots\cdots\cdots\cdots\\
b_m x_{n-1} + b_{m-1}x_n + \cdots + b_1 x_{n+m-2} + [b_0 - g(\alpha_i)]\,x_{n+m-1} &= 0,\\
b_m x_n + \ldots + b_1 x_{n+m-1} + [b_0 - g(\alpha_i)]\,x_{n+m} &= 0
\end{aligned}\right. \qquad (20)$$

є системою $n + m$ однорідних лінійних рівнянь з $n + m$ невідомими $x_1, x_2, \ldots, x_m, \ldots, x_n, \ldots, x_{n+m-1}, x_{n+m}$.

Згідно з (19), вона, очевидно, має розв'язок

$$\alpha_i^{n+m-1},\ \alpha_i^{n+m-2},\ \ldots,\ \alpha_i^n,\ \ldots,\ \alpha_i^m,\ \ldots,\ \alpha_i,\ 1,$$

який є нетривіальним, бо серед цих елементів є 1.

У зв'язку з цим детермінант системи (20) дорівнює нулю (1, § 29). Позначимо цей детермінант буквою Δ. Отже, маємо:

$$\Delta = \begin{vmatrix}
a_n & a_{n-1} & \cdots & a_1 & a_0 & & & \\
 & a_n & a_{n-1} & \cdots & a_1 & a_0 & & \\
 & & \cdots\cdots & & & & & \\
 & & a_n & a_{n-1} & \cdots & a_1 & a_0 & \\
 & & & a_n & a_{n-1} & \cdots & a_1 & a_0 \\
b_m & b_{m-1} & \cdots & b_1 & [b_0 - g(\alpha_i)] & & & \\
 & b_m & b_{m-1} & \cdots & b_1 & [b_0 - g(\alpha_i)] & & \\
 & & \cdots\cdots & & & & & \\
 & & & b_m & b_{m-1} & \cdots & b_1 & [b_0 - g(\alpha_i)] \\
 & & & & b_m & b_{m-1} & \cdots & b_1 [b_0 - g(\alpha_i)]
\end{vmatrix} = 0 \qquad (21)$$

(вважатимемо, що на незаповнених місцях стоять нулі).

Оскільки α_i — довільний корінь $f(x)$, то детермінант Δ дорівнює нулю для кожного α_i ($i = 1, 2, \ldots, n$). У зв'язку з цим розглянемо

многочлен

$$F(u)=\begin{vmatrix}
a_n & a_{n-1} & \cdots & & a_1 & a_0 & & & \\
 & a_n & a_{n-1} & \cdots & & a_1 & & a_0 & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & a_n & a_{n-1} & \cdots & a_1 & a_0 & \\
 & & & & a_n & a_{n-1} & \cdots & a_1 & a_0 \\
b_m & b_{m-1} & \cdots & & b_1 & (b_0-u) & & & \\
 & b_m & b_{m-1} & \cdots & & b_1 & (b_0-u) & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & b_m & b_{m-1} & \cdots & b_1 & (b_0-u) & \\
 & & & & b_m & b_{m-1} & \cdots & b_1 & (b_1-u)
\end{vmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\}m\text{ рядків}\\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\}n\text{ рядків}\end{matrix}\qquad(22)$$

Зрозуміло, що $F(u)$ є многочленом n-го степеня відносно u, бо члени з $b_0 - u$ зустрічаються по одному разу в n різних рядках і різних стовпцях детермінанта. Коефіцієнт при старшому члені u^n дорівнює $(-1)^n a_n^m$, що можна легко встановити, враховуючи особливості будови детермінанта. Справді, через те що на незаповнених місцях стоять нулі, то u^n дістаємо лише при множенні елементів першої діагоналі.

Маючи на увазі (21), слід зробити висновок, що числа $g(\alpha_1)$, $g(\alpha_2)$, ..., $g(\alpha_n)$ є коренями многочлена $F(u)$. У зв'язку з цим за формулами Вієта маємо [вільний член $F(u)$ дорівнює $F(0)$]:

$$g(\alpha_1)\,g(\alpha_2)\,\ldots\,g(\alpha_n)=(-1)^n\frac{F(0)}{(-1)^n a_n^m}\qquad(a_n\neq 0),$$

звідки

$$F(0)=a_n^m g(\alpha_1)\,g(\alpha_2),\ \ldots,\ g(\alpha_n),$$

тобто $F(0) = R(f, g)$.

Але $F(0)$ можна дістати з детермінанта (22), якщо вважати в ньому $u = 0$. Отже,

$$R(f,g)=\begin{vmatrix}
a_n & a_{n-1} & \cdots & & a_1 & a_0 & & & \\
 & a_n & a_{n-1} & \cdots & & a_1 & a_0 & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & a_n & a_{n-1} & \cdots & & a_1 & a_0 & \\
 & & & a_n & a_{n-1} & \cdots & & a_1 & a_0 \\
b_m & b_{m-1} & & \cdots & & b_1 & b_0 & & \\
 & b_m & b_{m-1} & \cdots & & & b_1 & b_0 & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & b_m & b_{m-1} & \cdots & & b_1 & b_0 \\
 & & & & b_m & b_{m-1} & \cdots & b_1 & b_0
\end{vmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\}m\text{ рядків}\\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\}n\text{ рядків}\end{matrix}\qquad(23)$$

Ця форма результанта і називається ф о р м о ю С і л ь в е с т р а. Форма Сільвестра зручна тим, що результант у цій формі можна записати дуже легко, бо безпосередньо виражається через коефіцієнти обох

многочленів $f(x)$ і $g(x)$. Досить лише записати підряд m раз коефіцієнти многочлена $f(x)$, а потім n раз коефіцієнти многочлена $g(x)$, відповідно розміщуючи їх у детермінанті, як це показано в (23).

П р и к л а д. Обчислимо дискримінант многочлена $f(x) = x^2 + px + q$ за допомогою детермінанта Сільвестра. Маємо: $f'(x) = 2x + p$. Відповідно до (23) дістаємо ($m = 1$, $n = 2$):

$$R(f, f')=\begin{vmatrix}1 & p & q\\ 2 & p & 0\\ 0 & 2 & p\end{vmatrix}=p^2+4q-2p^2=4q-p^2.$$

Тому за формулою (17) матимемо:

$$D(f)=p^2-4q.$$

Форма Сільвестра важлива ще й тим, що вона дає можливість узагальнити поняття результанта, введене в п. 27. 2. Ми бачили, що при $a_n \neq 0$ детермінант Сільвестра збігається з $R(f, g)$, означеним формулою (10). Аналогічно, при $b_m \neq 0$ $R(g, f)$ дорівнює відповідному детермінанту Сільвестра. Але якщо $a_n = b_m = 0$, то означення (10) — (11) результанта $R(f, g)$ або $R(g, f)$ втрачає смисл, тоді як детермінант Сільвестра зберігає смисл, хоч і набуває значення нуль.

У зв'язку з цим *надалі під результантом $R(f, g)$ двох многочленів ми розумітимемо детермінант Сільвестра для цих многочленів.*

Зауважимо, що при цьому новому означенні результанта властивість 1 може втратити смисл (при $a_n = b_m = 0$), але залишається в силі рівність $R(f, g) = (-1)^{mn} R(g, f)$ (властивість 2), що можна перевірити, переставляючи рядки детермінанта (23).

Ми вже зазначали, що при доведенні теореми 1 істотну роль відігравала умова, що a_n і b_m одночасно не дорівнюють нулю. Оскільки тепер результант зберігає смисл і при $a_n = b_m = 0$, то природно розглянути питання про те, як слід формулювати цю теорему для узагальненого означення результанта. Відповідь на це питання дають теореми 3 і 4.

Теорема 3. *Якщо результант $R(f, g)$ дорівнює нулю, то або* а) *многочлени $f(x)$ і $g(x)$ мають спільний корінь, або* б) *обидва їх старші коефіцієнти дорівнюють нулю.*

Д о в е д е н н я. Очевидно, досить показати, що коли при $R(f, g) = 0$ випадок б) не має місця, то неодмінно справджується висновок а). Припустимо, що $R(f, g) = 0$ і що a_n і b_m одночасно не перетворюються в нуль, тобто, наприклад, $a_n \neq 0$. Але тоді нове означення (23) результанта збігається з означенням (10) і до того ж справджуються умови теореми 1. Отже, $f(x)$ і $g(x)$ мають спільний корінь.

Теорема 4. *Якщо многочлени $f(x)$ і $g(x)$ мають спільний корінь, то $R(f, g)$ дорівнює нулю.*

Д о в е д е н н я. Якщо хоч один із старших коефіцієнтів a_n, b_m відмінний від нуля, то (аналогічно попередньому) можна застосувати теорему 1. Якщо ж $a_n = b_m = 0$, то рівність нулю результанта випливає безпосередньо з того, що в детермінанті Сільвестра всі елементи першого стовпця є нулі.

27.5. Розв'язування системи алгебраїчних рівнянь. Нехай дано систему двох алгебраїчних рівнянь з двома невідомими з коефіцієнта-

ми із P. Розмістивши члени цих рівнянь за степенями одного невідомого, матимемо:

$$\begin{aligned} f(x,\ y) &= a_n(y)\,x^n + a_{n-1}(y)\,x^{n-1} + \cdots + a_1(y)\,x + a_0(y) = 0,\\ g(x,\ y) &= b_m(y)\,x^m + b_{m-1}(y)\,x^{m-1} + \cdots + b_1(y)\,x + b_0(y) = 0. \end{aligned} \tag{24}$$

Розглядаючи y як параметр, побудуємо для многочленів $f(x, y)$ і $g(x, y)$ результант $R(f, g)$:

$$R(f,g) = \begin{vmatrix} a_n(y) & a_{n-1}(y) & \cdots & a_1(y) & a_0(y) & & & \\ & a_n(y) & a_{n-1}(y) & \cdots & a_1(y) & a_0(y) & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & a_n(y) & a_{n-1}(y) & \cdots & a_1(y) & a_0(y) & \\ & & & a_n(y) & a_{n-1}(y) & \cdots & a_1(y) & a_0(y) \\ b_m(y) & b_{m-1}(y) & \cdots & b_1(y) & b_0(y) & & & \\ & b_m(y) & b_{m-1}(y) & \cdots & b_1(y) & b_0(y) & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & b_m(y) & b_{m-1}(y) & \cdots & b_1(y) & b_0(y) & \\ & & & b_m(y) & b_{m-1}(y) & \cdots & b_1(y) & b_0(y) \end{vmatrix} \tag{25}$$

Позначимо цей результант через $R(y)$:

$$R(f,\ g) \underset{\text{df}}{=} R(y).$$

Як видно з (25), $R(y)$ дістаємо з коефіцієнтів $a_i(y)$ і $b_j(y)$ за допомогою дій додавання і множення. Тому *результант многочленів $f(x, y)$ і $g(x, y)$ є многочленом від y над тим самим полем P. Позначимо степінь многочлена $R(y)$ через l. Можна довести* [1], *що степінь $R(y)$ не перевищує добутку степенів многочленів $f(x, y)$ і $g(x, y)$* (відносно обох змінних x і y). Многочлен $R(y)$ має в полі розкладу l коренів $\beta_1, \beta_2, \ldots, \beta_l$, тобто $R(\beta_k) = 0$ $(k = 1, 2, \ldots, l)$. Але те, що результант дорівнює нулю, на основі теореми 3 означає, що многочлени

$$\begin{aligned} f(x,\ \beta_k) &= a_n(\beta_k)\,x^n + a_{n-1}(\beta_k)\,x^{n-1} + \cdots + a_1(\beta_k)\,x + a_0(\beta_k),\\ g(x,\ \beta_k) &= b_m(\beta_k)\,x^m + b_{m-1}(\beta_k)\,x^{m-1} + \cdots + b_1(\beta_k)\,x + b_0(\beta_k), \end{aligned}$$

або мають снільний корінь, або їх старші коефіцієнти $a_n(\beta_k)$ і $b_m(\beta_k)$ дорівнюють нулю. Розглянемо обидва випадки.

а) *Хоча б один з коефіцієнтів $a_n(\beta_k)$, $b_m(\beta_k)$ не дорівнює нулю.*

У цьому випадку многочлени $f(x, \beta_k)$ і $g(x, \beta_k)$ мають снільний корінь. Позначимо його через α_k. Пара (α_k, β_k), очевидно, є одним з розв'язків системи (24), бо $f(\alpha_k, \beta_k) = g(\alpha_k, \beta_k) = 0$. Зауважимо, що для даного значення β_k ми можемо мати не один, а кілька спільних коренів многочленів $f(x, \beta_k)$ і $g(x, \beta_k)$, наприклад α_k' і α_k''. Тоді, очевидно, обидві пари (α_k, β_k) і (α_k, β_k) є розв'язками системи (24).

б) *Обидва коефіцієнти $a_n(\beta_k)$ і $b_m(\beta_k)$ дорівнюють нулю.*

У цьому випадку, незважаючи на те, що результант дорівнює нулю,

[1] Див., наприклад: О к у н є в Л. Я. Вища алгебра. К., «Радянська школа», 1954, § 55. Це твердження називається *теоремою Безу*.

$f(x, \beta_k)$ і $g(x, \beta_k)$ можуть не мати спільного кореня. Якщо це так, то корінь результанта β_k слід відкинути. Проте може статися, що й у цьому випадку многочлени $f(x, \beta_k)$ і $g(x, \beta_k)$ мають спільний корінь α_k. Тоді знову (α_k, β_k) є розв'язком системи (24).

Щоб знайти всі розв'язки системи (24), треба аналогічно розглянути всі корені результанта $R(y)$.

Зауважимо, що крім знайдених таким способом розв'язків ніяких інших розв'язків система мати не може. Справді, якщо (α, β) — довільний розв'язок системи (24), то многочлени

$$\begin{aligned} f(x,\ \beta) &= a_n(\beta)\,x^n + a_{n-1}(\beta)\,x^{n-1} + \cdots + a_1(\beta)\,x + a_0(\beta),\\ g(x,\ \beta) &= b_m(\beta)\,x^m + b_{m-1}(\beta)\,x^{m-1} + \cdots + b_1(\beta)\,x + b_0(\beta) \end{aligned}$$

повинні мати спільний корінь α. Але тоді за теоремою 4 результант $R(y)$ цих многочленів при $y = \beta$ має дорівнювати нулю: $R(\beta) = 0$. Отже, β є одним з коренів результанта.

Як висновок з усього сказаного можна запропонувати таку схему дій при розв'язуванні системи двох рівнянь з двома невідомими.

1. Побудувати результант (25) і знайти всі його корені.

2. Знайдений корінь β_k $(k = 1, 2, \ldots, l)$ підставити в многочлени системи (24). Дістанемо многочлени $f(x, \beta_k)$ і $g(x, \beta_k)$.

3. Знайти спільний найбільший дільник $d_k(x)$ многочленів $f(x, \beta_k)$ і $g(x, \beta_k)$.

4. Розв'язати рівняння $d_k(x) = 0$. Корені цього рівняння $\alpha_{k_1}, \alpha_{k_2}, \ldots, \alpha_{k_p}$ є, очевидно, спільними коренями многочленів $f(x, \beta_k)$ і $g(x, \beta_k)$.

5. Скласти систему пар (α_{k_1}, β_k), (α_{k_2}, β_k), $\ldots$, (α_{k_p}, β_k). Ці пари є розв'язками системи (24), які відповідають кореню β_k.

Дії 2—5 слід виконати окремо для кожного з коренів β_k $(k = 1, 2, \ldots, l)$. При всьому цьому слід мати на увазі, що для деяких коренів результанта може не існувати спільних коренів многочленів $f(x, \beta_k)$, $g(x, \beta_k)$. Це може бути тоді, коли $a_n(\beta_k) = b_m(\beta_k) = 0$. Ознакою відсутності спільних коренів, очевидно, є те, що многочлени $f(x, \beta_k)$ і $g(x, \beta_k)$ взаємно прості, тобто що $d_k(x) = 1$.

У тих випадках, коли корені α_i одного з многочленів $f(x, \beta_k)$, $g(x, \beta_k)$ легко знайти, шукати найбільший спільний дільник $d_k(x)$ недоцільно; досить перевірити, які з коренів α_i є коренями і другого многочлена.

Ми розглянули як розв'язується система двох алгебраїчних рівнянь з двома невідомими. Цю теорію можна застосувати і до розв'язування системи s алгебраїчних рівнянь з s невідомими. Розглянемо в загальних рисах, як розв'язується система 3-х рівнянь з 3-ма невідомими. Нехай маємо систему:

$$\begin{cases} f_1(x,\ y,\ z) = 0,\\ f_2(x,\ y,\ z) = 0,\\ f_3(x,\ y,\ z) = 0. \end{cases} \tag{26}$$

Розмістивши многочлени f_1, f_2, f_3 за степенями змінної x, знайдемо результанти

$$R_1(f_1, f_2) \quad \text{і} \quad R_2(f_2, f_3).$$

Очевидно, це будуть многочлени від двох невідомих y, z:

$$R_1(f_1, f_2) = R_1(y, z) \quad \text{і} \quad R_2(f_2, f_3) = R_2(y, z).$$

Далі, розмістивши ці многочлени за степенями змінної y, знаходимо їх результант

$$R(R_1, R_2) = R(z).$$

Нехай γ — якийсь з коренів результанта $R(z)$. Підставивши $z = \gamma$ в результанти $R_1(y, z)$ і $R_2(y, z)$, знаходимо спільний корінь $y = \beta$ многочленів $R_1(y, \gamma)$ і $R_2(y, \gamma)$; далі, підставивши $y = \beta$, $z = \gamma$ у многочлени f_1, f_2, f_3, знаходимо спільний корінь α трьох многочленів $f_1(x, \beta, \gamma)$, $f_2(x, \beta, \gamma)$, $f_3(x, \beta, \gamma)$.

Трійка елементів α, β, γ і є одним з розв'язків системи (26).

Приклади. 1. Розв'язати систему

$$\begin{cases} f(x, y) \underset{\text{df}}{=} (y-1)x^2 - yx + 1 = 0, \\ g(x, y) \underset{\text{df}}{=} (y-1)x^3 + 3x + y - 6 = 0. \end{cases}$$

$$R(y) = \begin{vmatrix} y-1 & -y & 1 & 0 \\ 0 & y-1 & -y & 1 \\ y-1 & 3 & y-6 & 0 \\ 0 & y-1 & 3 & y-6 \end{vmatrix} = 2(y-1)(y^3 - 9y^2 + 24y - 20).$$

Аналізуючи многочлен $y^3 - 9y^2 + 24y - 20$, встановлюємо, що він має раціональний корінь $y = 2$, а потім розкладаємо його на множники. Матимемо: $R(y) = 2(y-1)(y-2)^2(y-5)$. Отже, коренями результанта є числа $\beta_1 = 1$, $\beta_{2,3} = 2$, $\beta_4 = 5$. а) $\beta_1 = 1$.

$$f(x, \beta_1) = -x + 1,$$
$$g(x, \beta_1) = 3x - 5.$$

Спільного кореня немає. Розв'язку системи, який відповідав би кореню результанта $\beta_1 = 1$, немає. Це стає зрозумілим, коли врахувати, що при $\beta_1 = 1$ старші коефіцієнти перетворилися в нуль.

б) $\beta_{2,3} = 2$.

$$f(x, \beta_2) = x^2 - 2x + 1,$$
$$g(x, \beta_2) = x^2 + 3x - 4.$$

Тут спільний корінь $\alpha_2 = 1$ і, отже, $\alpha_2 = 1$, $\beta_2 = 2$ є розв'язком системи.

в) $\beta_4 = 5$.

$$f(x, \beta_4) = 4x^2 - 5x + 1,$$
$$g(x, \beta_4) = 4x^2 + 3x - 1.$$

Спільний корінь $\alpha_4 = \frac{1}{4}$. Отже, $\alpha_4 = \frac{1}{4}$, $\beta_4 = 5$ — ще один розв'язок системи.

2. Застосуємо викладену теорію до прикладу, який ми розглядали у вступі до цього параграфа:

$$f(x, y) \underset{\text{df}}{=} x^2 + y^2 - a = 0,$$
$$g(x, y) \underset{\text{df}}{=} xy - b = 0.$$

Виключаючи невідоме x відомими з школи методами, ми дістали рівняння (п. 27.1):

$$y^4 - ay^2 + b^2 = 0. \tag{28}$$

Знайдемо тепер результант многочленів $f(x, y)$ і $g(x, y)$:

$$R(y) = \begin{vmatrix} 1 & 0 & y^2 - a \\ y & -b & 0 \\ 0 & y & -b \end{vmatrix} = y^4 - ay^2 + b^2.$$

Отже, ліва частина рівняння (28) є нічим іншим, як результантом.

Зауважимо, що взагалі виключення невідомих методами елементарної алгебри в більшості випадків і є знаходженням результанта. Потім за коренем β результанта знаходять спільний корінь многочленів $f(x, \beta)$ і $g(x, \beta)$. Спільність кореня перевіряють підстановкою, що обов'язково слід робити при розв'язуванні таких систем у школі.

Ми в цьому параграфі виходили з того, що вміємо розв'язувати задачу знаходження усіх коренів многочлена від однієї змінної. Насправді ми лише знаємо, що таку задачу в принципі можна розв'язати, оскільки для будь-якого многочлена однієї змінної степеня n існує n коренів у полі розкладу. Що ж до методів знаходження коренів, то вони нам поки що відомі тільки для многочленів нижчих степенів (1, § 17). Деякі методи знаходження коренів многочленів над числовими полями будуть розглянуті в розділі VII.

Розділ VII

МНОГОЧЛЕНИ НАД ЧИСЛОВИМИ ПОЛЯМИ

§ 28. ОСНОВНА ТЕОРЕМА ТЕОРІЇ МНОГОЧЛЕНІВ

28.1. Вступні зауваження. Перейдемо тепер до вивчення спеціальних властивостей многочленів з числовими коефіцієнтами, тобто многочленів над числовими полями. Важливість цієї теми обумовлена тим, що багато задач з різних галузей математики, природознавства, техніки, економіки зводяться до розв'язування і дослідження алгебраїчних рівнянь чи систем таких рівнянь з дійсними або комплексними коефіцієнтами. У зв'язку з цим саме теорія многочленів з числовими коефіцієнтами становила основний предмет алгебраїчної науки до середини XIX ст., на базі якого виникли і розвинулись сучасні уявлення про кільця, поля та інші алгебраїчні структури і про многочлени над абстрактними полями.

Особливе значення має вивчення властивостей многочленів над числовими полями для вчителів, оскільки в школі розглядаються лише такі многочлени і відповідні алгебраїчні рівняння.

Як було показано в п. 21.5, для многочленів над числовими полями *алгебраїчне та функціональне тлумачення цілком рівноправні*. Це означає, що такі многочлени можна розглядати як функції дійсної чи ком-

плексної змінної і застосовувати до них означення та твердження, встановлені для таких функцій зокрема, поняття і властивості неперервності. У цьому розділі ми будемо дотримуватись функціонального погляду на многочлени, оскільки він використовується при встановленні існування і дослідженні числа та розміщення коренів рівняння з числовими коефіцієнтами і відповідає як історичному розвитку алгебри, так і змісту сучасної шкільної програми.

У попередніх розділах ми вивчали в основному ті властивості многочленів, які не залежали від того, до якого основного поля належать їх коефіцієнти, є спільними для різних таких полів і допускають єдине доведення. Так, хоч розклад многочлена на незвідні множники не буде одним і тим самим у різних полях, що містять його коефіцієнти, проте спільна властивість (саме вона нас і цікавила в розділі V) полягає в тому, що в довільному такому полі цей розклад можливий і для даного поля єдиний. Спільними для всіх полів P є також властивості операцій у кільці многочленів $P[x]$, основні факти теорії подільності многочленів, властивості симетричних многочленів тощо. Звичайно, всі ці властивості мають місце і для многочленів над довільним числовим полем.

Важливими властивостями многочлена є наявність, число і розміщення його коренів. Приступаючи до вивчення цих питань, ми вже не можемо розраховувати на те, що відповідні властивості многочленів не залежатимуть від вибору основного поля P. Адже той самий многочлен може мати корені в одному і не мати їх у другому полі. Так, многочлен $f(x) = x^2 + 1$ не має коренів у полі дійсних чисел, але має два корені $\pm i$ в полі комплексних чисел.

Як було встановлено в § 23, для кожного многочлена $f(x)$ з кільця $P[x]$ існує своє поле розкладу, а саме таке розширення L поля P, в якому многочлен $f(x)$ розкладається в добуток лінійних множників. Серед числових полів найбільш важливу властивість має поле $\mathbf{C}$ усіх комплексних чисел. Виявляється, що полем розкладу для будь-якого многочлена $f(x)$ над полем $\mathbf{C}$ є саме поле $\mathbf{C}$, тобто в полі комплексних чисел будь-який многочлен розкладається на лінійні множники. Іншими словами, *поле* $\mathbf{C}$ *алгебраїчно замкнуте* і є єдиним числовим полем, яке має цю фундаментальну властивість.

У зв'язку з цим важливим є вивчення властивостей многочленів з комплексними коефіцієнтами або цілих раціональних функцій комплексної змінної. У цьому розділі комплексну змінну позначатимемо буквою z.

Головним результатом дослідження питання про існування коренів алгебраїчних рівнянь над полем $\mathbf{C}$ є так звана основна теорема теорії многочленів, згідно з якою довільний многочлен ненульового степеня з комплексними коефіцієнтами має хоча б один комплексний корінь. Цей важливий факт, рівнозначний згаданій вище властивості алгебраїчної замкнутості поля $\mathbf{C}$, лежить в основі класичної теорії многочленів, методів дослідження і розв'язування алгебраїчних рівнянь.

Саме тому ця теорема раніше називалася *основною теоремою алгебри*. Проте цю традиційну назву тепер вважають застарілою, оскільки предмет сучасної алгебри далеко не вичерпується теорією алгебраїчних

рівнянь. Існує багато доведень основної теореми теорії многочленів. Ми наведемо доведення Ейлера — Гаусса, яке вважається «найбільш алгебраїчним». Хоч і воно спирається на функціональні властивості многочленів (зокрема, на їх неперервність у полі дійсних чисел), проте, порівняно з іншими доведеннями, використання таких властивостей зведено тут до мінімуму.

28.2. Властивості модуля многочлена. Розглянемо спочатку властивості модуля многочлена.

Теорема 1. *Якщо $f(z)$ — многочлен ненульового степеня, то для довільного додатного числа M можна знайти таке число N, що при $|z| > N$ виконується нерівність $|f(z)| > M$.*

Це твердження означає, що $|f(z)|$ необмежено зростає, коли точка z необмежено віддаляється від початку координат, бо яким би великим не було число M, $|f(z)|$ перевищуватиме M, як тільки відстань точки z від початку координат буде більша за відповідне N.

Доведення. Користуючись властивостями модуля комплексного числа, маємо:

$$|f(z)| = |a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0| \geqslant$$
$$\geqslant |a_n| \cdot |z|^n - |a_{n-1}z^{n-1} + \cdots + a_1 z + a_0|. \quad (1)$$

Але

$$|a_{n-1}z^{n-1} + \cdots + a_1 z + a_0| \leqslant |a_{n-1}| \cdot |z|^{n-1} + \cdots +$$
$$+ |a_1| \cdot |z| + |a_0| \leqslant A \cdot [|z|^{n-1} + \cdots + |z| + 1] = A\frac{|z|^n - 1}{|z| - 1}, \quad (2)$$

де A — найбільший з модулів коефіцієнтів $|a_{n-1}|, \ldots |a_1|, |a_0|$. Якщо накласти на змінну z (яка до цього часу була довільним комплексним числом) додаткову умову

$$|z| > 1, \quad (3)$$

то

$$A\frac{|z|^n - 1}{|z| - 1} < A\frac{|z|^n}{|z| - 1}. \quad (4)$$

Підсилюючи за допомогою нерівностей (2) і (4) нерівність (1), маємо:

$$|f(z)| > |a_n| \cdot |z|^n - A\frac{|z|^n}{|z| - 1} = |z|^n \cdot \frac{|a_n| \cdot |z| - |a_n| - A}{|z| - 1}. \quad (5)$$

При необмеженому зростанні $|z|$ стане більшим за число

$$N_1 = \frac{2A}{|a_n|} + 1. \quad (6)$$

Для таких значень z справджується нерівність

$$\frac{A}{|z| - 1} < \frac{|a_n|}{2},$$

і тому

$$\frac{|a_n| \cdot |z| - |a_n| - A}{|z| - 1} = |a_n| - \frac{A}{|z| - 1} > |a_n| - \frac{|a_n|}{2} = \frac{|a_n|}{2}. \quad (7)$$

Коли $|z|$, задовольняючи нерівність (3), задовольняє і нерівність (6), тобто коли

$$|z| > \max\{1,\ N_1\} = N_1, \tag{8}$$

то на підставі (5) і (7) можна записати:

$$|f(z)| > |z|^n \cdot \frac{|a_n|}{2}. \tag{9}$$

Покажемо тепер, що при достатньо великих $|z|$ величина $|f(z)|$ буде більшою від наперед заданого числа M. Справді, при

$$|z| > \sqrt[n]{\frac{2M}{|a_n|}} \tag{10}$$

справедлива нерівність

$$|z|^n \cdot \frac{|a_n|}{2} > \frac{2M}{|a_n|} \cdot \frac{|a_n|}{2} = M. \tag{11}$$

Якщо при цьому також справджується нерівність (9), то з (9) і (11) випливає $|f(z)| > M$. Через те що нерівність (9) справедлива для тих z, що задовольняють умову (8), а нерівність (11) — для тих, що задовольняють умову (10), то потрібна нам нерівність $|f(z)| > M$ справджуватиметься для всіх z, які задовольняють обидві ці умови, тобто для яких $|z| > N$, де

$$N = \max\left\{N_1,\ \sqrt[n]{\frac{2M}{|a_n|}}\right\}.$$

Зрозуміло, що таке N можна знайти для довільного додатного M. Теорему доведено.

З нерівностей, встановлених при доведенні цієї теореми, можна безпосередньо дістати такі важливі наслідки:

Наслідок 1. *Многочлен* $f(z) = a_n z^n + \ldots + a_1 z + a_0$ *може мати тільки такі корені, модуль яких менший від числа*

$$N_0 = 1 + \frac{A}{|a_n|}, \tag{12}$$

де A *є найбільший з модулів коефіцієнтів* $|a_{n-1}|, \ldots, |a_1|, |a_0|$.

Справді, якщо z — довільне число, причому $|z| \geqslant N_0$, тобто $|z| \geqslant 1 + \frac{A}{|a_n|}$, або $|a_n| \cdot |z| - |a_n| - A \geqslant 0$, то нерівність (5) показує, що $|f(z)| > 0$, тобто z не є коренем $f(z)$. Зауважимо, що нерівність (5) справедлива лише для $|z| > 1$, але коли $z \geqslant N_0$, то зрозуміло, що й $|z| > 1$, бо $N_0 = 1 + \frac{A}{|a_n|} > 1$.

Наслідок 2. *При* $|z| > N_0 = 1 + \frac{A}{|a_n|}$ *модуль старшого члена многочлена* $f(z)$ *більший за модуль суми всіх інших членів цього многочлена.* Справді, якщо $|z| > N_0$, то $|a_n| \cdot |z| - |a_n| - A > 0$, тому

$$|a_n| \cdot |z|^n - \frac{A|z|^n}{|z|-1} = |z|^n \frac{|a_n| \cdot |z| - |a_n| - A}{|z|-1} > 0,$$

тобто $|a_n z^n| > A\frac{|z|^n}{|z|-1}$.

Але з нерівностей (2) і (4) дістанемо:

$$A\frac{|z|^n}{|z|-1} > |a_{n-1}z^{n-1} + \cdots + a_1 z + a_0|.$$

Отже, остаточно

$$|a_n z^n| > |a_{n-1}z^{n-1} + \cdots + a_1 z + a_0| \tag{13}$$

при $|z| > N_0 = 1 + \frac{A}{|a_n|}$.

Застосування доведеної теореми і її наслідків до окремого випадку — многочлена непарного степеня над полем **R** дійсних чисел — дає змогу встановити такий важливий факт.

Теорема 2. *Многочлен непарного степеня над полем* **R** *дійсних чисел має принаймні один дійсний корінь.*

Доведення. Нехай

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

многочлен з дійсними коефіцієнтами, і змінна x набуває лише значення дійсних чисел. В цьому випадку $f(x)$ можна розглядати як функцію дійсної змінної. Як відомо з курсу аналізу, ця функція неперервна на всій дійсній осі. Згідно з наслідком 2 попередньої теореми, при досить великих числових значеннях $|x|$ модуль старшого члена $|a_n x^n|$ більший за модуль суми всіх інших членів цього многочлена. Тому при таких значеннях x числове значення многочлена $f(x)$ має знак, який збігається з знаком старшого члена $a_n x^n$. А оскільки n — непарне число, то при $x \to +\infty$ і при $x \to -\infty$ старший член $a_n x^n$ набуває протилежних знаків. Тому при досить великих $|x|$ числові значення $f(x)$ будуть різні за знаком залежно від того, додатним, чи від'ємним буде значення змінної x.

Отже, існують значення $x = a$ і $x = b$ такі, що $f(a)$ і $f(b)$ будуть різні за знаком. Але тоді, згідно з теоремою Больцано — Коші, відомою з математичного аналізу, в (a, b) існує принаймні одна така точка ξ, в якій $f(x)$ перетворюється в нуль, тобто $f(\xi) = 0$. Теорему доведено.

28.3. Доведення основної теореми теорії многочленів. Доведемо спочатку таку теорему.

Теорема 3. *Кожний многочлен степеня* $n \geqslant 1$ *з дійсними коефіцієнтами має принаймні один комплексний корінь.*

Доведення. Насамперед зауважимо, що будь-яке натуральне число n можна записати так: $n = 2^k \cdot q$, де k — ціле невід'ємне число, а q — деяке непарне натуральне число.

Нехай $f(z) = a_n z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$ — будь-який многочлен з дійсними коефіцієнтами степеня $n = 2^k \cdot q$. Теорему доводитимемо методом математичної індукції по k. При $k = 0$ показник степеня $n = q$ — непарне число, тому, за теоремою 2 (п. 28.2), доводжуване твердження справедливе. Припустимо тепер, що теорема 3 справедлива для будь-якого многочлена з дійсними коефіцієнтами степеня $2^{k-1} \cdot q$, тобто для многочлена, степінь якого ділиться на 2^{k-1} і не ділиться на 2^k, і доведемо, що тоді вона справедлива і для будь-якого многочлена з дійсними коефіцієнтами степеня $2^k \cdot q$.

Для многочлена $f(z)$, що розглядається над полем $\mathbf{C}$ комплексних чисел, існує поле розкладу $\bar{\mathbf{C}}$. У полі $\bar{\mathbf{C}}$ $f(z)$ має n коренів (п. 23.3). Позначимо їх символами $\alpha_1, \alpha_2, \ldots, \alpha_n$. Виберемо тепер довільне дійсне число r і візьмемо всі можливі елементи поля $\bar{\mathbf{C}}$, що мають вигляд $\beta_{ij} = \alpha_i\alpha_j + r(\alpha_i + \alpha_j)$, $i < j$. Число таких елементів β_{ij} дорівнює, очевидно, числу комбінацій з n елементів по два, тобто

$$C_n^2 = \frac{n(n-1)}{2} = \frac{2^k \cdot q(2^k \cdot q - 1)}{2} = 2^{k-1}q(2^k q - 1) = 2^{k-1} \cdot q_1,$$

де $q_1 = q(2^k q - 1)$ — непарне число.

Розглянемо тепер многочлен

$$\varphi(z) = \prod_{i,j(i<j)} (z - \beta_{ij}),$$

коренями якого є числа β_{ij} і тільки вони. Степінь многочлена $\varphi(z)$, очевидно, дорівнює $2^{k-1} \cdot q_1$. Оскільки r — дійсне число, то коефіцієнти $\varphi(z)$ є многочленами від β_{ij} і, отже, від α_i, α_j з дійсними коефіцієнтами. Легко зрозуміти, що будь-яке переставлення елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$, очевидно, приведе тільки до переставлення лінійних множників многочлена

$$\varphi(z) = \prod_{i,j(i<j)} (z - \beta_{ij}) = \prod_{i,j(i<j)} \{z - [\alpha_i\alpha_j + r(\alpha_i + \alpha_j)]\},$$

а сам многочлен $\varphi(z)$ від цього не зміниться. Оскільки многочлен $\varphi(z)$ не змінюється при будь-якому переставленні елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$, то, отже, не змінюються і його коефіцієнти. Тому коефіцієнти многочлена $\varphi(z)$ — симетричні многочлени від $\alpha_1, \alpha_2, \ldots, \alpha_n$ над полем R дійсних чисел.

Але оскільки $\alpha_1, \alpha_2, \ldots, \alpha_n$ — корені многочлена $f(z)$ з дійсними коефіцієнтами, то, за наслідком з основної теореми про симетричні многочлени (п. 26.2, теорема 3), коефіцієнти многочлена $\varphi(z)$ — дійсні числа. Тому, за припущенням індукції, многочлен $\varphi(z)$ має принаймні один комплексний корінь. Але оскільки коренями многочлена $\varphi(z)$ є тільки елементи $\alpha_i\alpha_j + r(\alpha_i + \alpha_j)$ $(i < j)$, то принаймні один з цих елементів повинен бути комплексним числом.

Отже, яке б ми не взяли дійсне число r, можна вказати таку пару індексів i, j $(1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n)$, що елемент $\alpha_i\alpha_j + r(\alpha_i + \alpha_j)$ поля $\bar{\mathbf{C}}$ є комплексним числом. Різним дійсним числам r і r' відповідатимуть у цьому розумінні різні пари індексів. Але оскільки множина дійсних чисел нескінченна, а число всіх можливих пар індексів скінченне, то можна вибрати такі два різні дійсні числа r_1 і r_2, що їм відповідатиме та сама пара індексів i, j, для яких

$$\alpha_i\alpha_j + r_1(\alpha_i + \alpha_j) = \gamma_1, \tag{14}$$

$$\alpha_i\alpha_j + r_2(\alpha_i + \alpha_j) = \gamma_2 \tag{15}$$

є комплексні числа.

Віднявши почленно ці рівності, дістанемо $(r_1 - r_2)(\alpha_i + \alpha_j) = \gamma_1 - \gamma_2$, звідки

$$\alpha_i + \alpha_j = \frac{\gamma_1 - \gamma_2}{r_1 - r_2}. \tag{16}$$

Підставивши знайдене значення суми в (14), дістанемо:

$$\alpha_i\alpha_j + r_1\frac{\gamma_1 - \gamma_2}{r_1 - r_2} = \gamma_1,$$

звідки

$$\alpha_i\alpha_j = \gamma_1 - r_1\frac{\gamma_1 - \gamma_2}{r_1 - r_2}. \tag{17}$$

Як бачимо, сума $\alpha_i + \alpha_j$ і добуток $\alpha_i\alpha_j$ — комплексні числа. Але тоді з (16) і (17) можна знайти α_i і α_j, які, очевидно, теж є комплексними числами. Числа α_i і α_j можна знайти, наприклад, як корені такого квадратного рівняння:

$$z^2 - (\alpha_i + \alpha_j)z + \alpha_i\alpha_j = 0.$$

Таким чином показано, що серед коренів $\alpha_1, \alpha_2, \ldots, \alpha_n$ многочлена $f(z)$ є навіть два комплексні корені. Цим теорему доведено.

Теорема 3 формулюється для многочленів з дійсними коефіцієнтами. Доведемо тепер аналогічну теорему для більш широкого класу многочленів з комплексними коефіцієнтами — основну теорему теорії многочленів.

Теорема 4. *Довільний многочлен ненульового степеня з комплексними коефіцієнтами*

$$f(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

має хоча б один комплексний корінь.

Доведення. Нехай

$$f(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

є многочлен степеня $n \geqslant 1$ з довільними комплексними коефіцієнтами. Візьмемо многочлен

$$\bar{f}(z) = \bar{a}_n z^n + \bar{a}_{n-1}z^{n-1} + \cdots + \bar{a}_1 z + \bar{a}_0,$$

де $\bar{a}_i$ $(i = 0, 1, \ldots, n)$ є комплексне число, спряжене з числом a_i, і розглянемо добуток

$$g(z) = f(z) \cdot \bar{f}(z) = b_{2n}z^{2n} + b_{2n-1}z^{2n-1} + \cdots + b_1 z + b_0,$$

де $b_k = \sum_{i+j=k} a_i\bar{a}_j$, $k = 0, 1, 2, \ldots, 2n$.

Згідно з властивостями спряжених комплексних чисел

$$\bar{b}_k = \sum_{i+j=k} \bar{a}_i a_j = b_i,$$

тобто всі коефіцієнти многочлена $g(z)$ — дійсні числа. Тому за теоремою 3 многочлен $g(z)$ має принаймні один комплексний корінь α і, отже, $\varphi(\alpha) = f(\alpha) \cdot \bar{f}(\alpha) = 0$. Таким чином, або $f(\alpha) = 0$, або $\bar{f}(\alpha) = 0$. У першому випадку число α є коренем многочлена $f(z)$. Якщо маємо

другий випадок, тобто якщо $\bar{f}(\alpha) = \bar{a}_n\alpha^n + \bar{a}_{n-1}\alpha^{n-1} + \ldots + \bar{a}_1\alpha + \bar{a}_0 = 0$, то, замінивши в цій рівності всі комплексні числа спряженими з ними числами, дістанемо рівність

$$a_n\bar{\alpha}^n + a_{n-1}\bar{\alpha}^{n-1} + \cdots + a_1\bar{\alpha} + a_0 = f(\bar{\alpha}) = 0,$$

тобто число $\bar{\alpha}$ є коренем многочлена $f(z)$, і тому теорема знову правильна.

§ 29. НАСЛІДКИ З ОСНОВНОЇ ТЕОРЕМИ ТЕОРІЇ МНОГОЧЛЕНІВ

29.1. Розклад многочлена над полем комплексних чисел у добуток лінійних множників. З основної теореми теорії многочленів дістаємо ряд важливих наслідків.

Теорема 1. *Кожний многочлен, степінь якого вищий за одиницю, звідний у полі комплексних чисел.*

Доведення. Нехай $f(z)$ — многочлен степеня $n \geqslant 2$. За основною теоремою теорії многочленів, існує хоча б один корінь z_0 цього многочлена. За наслідком з теореми Безу (п. 23.1) $f(z)$ ділиться на $z - z_0$, тобто $f(z) = (z - z_0) \cdot f_1(z)$. Через те що степінь $f(z)$ більший за 1, то $f_1(z)$ є многочленом ненульового степеня. Цим і доведено звідність $f(z)$ у полі комплексних чисел.

Наслідок. *Для того щоб многочлен був незвідним у полі комплексних чисел, необхідно і достатньо, щоб його степінь дорівнював одиниці.*

Теорема 2. *Кожний многочлен n-го степеня над полем комплексних чисел єдиним способом* (з точністю до порядку множників) *розкладається на лінійні множники в цьому полі*

$$f(z) = a_n(z - z_1)(z - z_2) \ldots (z - z_n), \qquad (1)$$

де $z_1, z_2, \ldots, z_n$ — корені, а a_n — старший коефіцієнт многочлена $f(z)$.

Доведення. За теоремою 8 п. 22.7, кожний многочлен над полем **C** можна розкласти у добуток незвідних многочленів у цьому полі, причому ці многочлени визначаються однозначно з точністю до сталого множника: $f(z) = f_1(z) \cdot f_2(z) \ldots f_m(z)$. Але в полі комплексних чисел кожний незвідний многочлен має перший степінь. Отже, число множників m повинно дорівнювати степеню даного многочлена n і кожний з них є лінійним двочленом. Далі, оскільки $f_k(z)$ визначаються з точністю до сталого множника, вважатимемо, що в кожному з них старший коефіцієнт дорівнює одиниці, тобто $f_k(z) = z + \alpha_k$. Тоді $f(z)$ може відрізнятися від добутку всіх $f_k(z)$ лише сталим множником, тобто

$$f(z) = A(z + \alpha_1)(z + \alpha_2) \ldots (z + \alpha_n).$$

Але легко бачити, прирівнюючи старші коефіцієнти в обох частинах цієї рівності, що $A = a_n$. Далі, $-\alpha_1, -\alpha_2, \ldots, -\alpha_n$ є коренями многочлена $f(z)$, бо $f(-\alpha_k) = 0$. Тому ці числа позначимо через $z_1, z_2, \ldots, z_n$. Таким чином, замінюючи A через a_n і $-\alpha_k$ через z_k, дістаємо шуканий розклад (1). Оскільки сталі множники для незвідних многочленів $f_k(z)$ тут цілком визначені, то розклад (1) однозначний з точністю до порядку множників. Теорему доведено.

З розкладу (1) випливає, що жодне комплексне число, відмінне від чисел $z_1, z_2, \ldots, z_n$, не може бути коренем многочлена $f(z)$.

Оскільки під числом коренів многочлена в даному полі розуміють *число лінійних множників многочлена* в цьому полі (п. 23.1), то переконуємось у справедливості такого твердження:

Теорема 3. *Многочлен n-го степеня має в полі комплексних чисел точно n коренів.*

Ця теорема свідчить про те, що результати наших досліджень у ч. I курсу відносно числа коренів 3-го і 4-го степенів, а також двочленних рівнянь довільного степеня не були випадковими, а відбивали загальну закономірність, властиву всім алгебраїчним рівнянням.

Ми бачимо також, що всі корені многочлена $f(z)$ над полем **C** комплексних чисел належать цьому самому полю **C**, тобто *полем розкладу будь-якого многочлена $f(z)$ з комплексними коефіцієнтами є поле* **C** *комплексних чисел.*

Отже, *поле* **C** *комплексних чисел є алгебраїчно замкнутим.*

Ці результати показують, що вивчаючи многочлени з числовими коефіцієнтами, лише при переході до комплексної області можна створити загальну теорію алгебраїчних рівнянь. Справді, відомо, що поле раціональних і поле дійсних чисел не є алгебраїчно замкнутими: рівняння, коефіцієнти якого належать до поля **Q** або **R**, може зовсім не мати коренів у цих полях, або може мати лише деякі корені; решта ж коренів лежать в інших полях, зокрема в полі **C**.

Як випливає з теореми 6 п. 23.3, для коренів $z_1, z_2, \ldots, z_n$ алгебраїчного рівняння n-го степеня

$$a_nz^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0 = 0 \qquad (2)$$

справедливі співвідношення (формули Вієта):

$$\begin{aligned}
&z_1 + z_2 + \cdots + z_n = -\frac{a_{n-1}}{a_n},\\
&z_1z_2 + z_1z_3 + \cdots + z_{n-1}z_n = \frac{a_{n-2}}{a_n},\\
&z_1z_2z_3 + z_1z_2z_4 + \cdots + z_{n-2}z_{n-1}z_n = -\frac{a_{n-3}}{a_n}, \qquad (3)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&z_1z_2 \ldots z_{n-1} + z_1z_2 \ldots z_{n-2}z_n + \cdots + z_2z_3 \ldots z_n = (-1)^{n-1}\frac{a_1}{a_n},\\
&z_1z_2 \ldots z_n = (-1)^n\frac{a_0}{a_n}.
\end{aligned}$$

Зрозуміло, що в розкладі 1 многочлена $f(z)$, у полі C можуть бути кратні множники. У цьому випадку розклад (1) матиме вигляд:

$$f(z) = a_n(z - z_1)^{k_1}(z - z_2)^{k_2} \ldots (z - z_m)^{k_m}, \qquad (4)$$

де $z_1, z_2, \ldots, z_m$ — корені $f(z)$, серед яких немає рівних між собою ($m \leqslant n$).

Теорема 3 і формули Вієта спираються на розклад многочлена на незвідні множники не у формі (4), а у формі (1). Порівняння цих

розкладів показує, що згадані твердження справедливі тільки при умові, що *кожний корінь враховується стільки разів, яка його кратність.*

29.2. Розклад многочленів над полем дійсних чисел на добуток незвідних множників. Рівняння з дійсними коефіцієнтами є поширеним і важливим для практичних застосувань окремим випадком алгебраїчних рівнянь з комплексними коефіцієнтами. Оскільки дійсні числа утворюють підполе поля **C** комплексних чисел, всі результати цього параграфа, зокрема теореми про існування комплексних коренів та їх число, залишаються справедливими і для многочленів з дійсними коефіцієнтами, тобто *будь-який многочлен n-го степеня з дійсними коефіцієнтами має точно n комплексних коренів.*

Але в багатьох випадках особливий інтерес становлять саме дійсні корені рівнянь з дійсними коефіцієнтами. Ми знаємо, що рівняння з дійсними коефіцієнтами може взагалі не мати жодного дійсного кореня (наприклад, рівняння $x^2 + 1 = 0$). Проте виявляється, що основна теорема теорії многочленів дає змогу зробити ряд висновків і щодо коренів рівнянь з дійсними коефіцієнтами.

Теорема 4. *Якщо комплексне число z_0 є коренем многочлена з дійсними коефіцієнтами*

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \tag{5}$$

то спряжене комплексне число $\bar{z}_0$ також є коренем цього многочлена.

Доведення. Обчислимо значення $f(z_0)$. Відокремивши дійсну і уявну частини, матимемо:

$$f(z_0) = a_n z_0^n + a_{n-1} z_0^{n-1} + \cdots + a_1 z_0 + a_0 = A + Bi. \tag{6}$$

Але z_0 є коренем многочлена (5), тому $A + Bi = 0$, звідки $A = B = 0$. Обчислимо тепер вираз $f(\bar{z}_0)$. Через те, що всі коефіцієнти a_k — дійсні числа, то $\bar{a}_k = a_k$ і тому

$$f(\bar{z}_0) = a_n (\bar{z}_0)^n + a_{n-1} (\bar{z}_0)^{n-1} + \cdots + a_1 \bar{z}_0 + a_0 =$$
$$= \bar{a}_n (\bar{z}_0)^n + \bar{a}_{n-1} (\bar{z}_0)^{n-1} + \cdots + \bar{a}_1 \bar{z}_0 + \bar{a}_0. \tag{7}$$

Порівнюючи (6) і (7), бачимо, що $f(\bar{z}_0)$ можна дістати з $f(z_0)$ в результаті заміни всіх чисел спряженими. Оскільки над цими числами виконуються лише дії додавання і множення, то на підставі властивостей комплексних чисел (1, § 16) $f(z_0)$ і $f(\bar{z}_0)$ є спряжені комплексні числа, тобто $f(\bar{z}_0) = \overline{f(z_0)} = A - Bi$. Але ми вже показали, що $A = B = 0$. Отже, $f(\bar{z}_0) = 0$, тому $\bar{z}_0$ є коренем даного рівняння. Теорему доведено.

Цю теорему природно доповнити таким твердженням.

Теорема 5. *Якщо комплексне число z_0 є коренем k-ї кратності ($k > 1$) многочлена $f(z)$ з дійсними коефіцієнтами, то спряжене комплексне число $\bar{z}_0$ є коренем многочлена $f(z)$ тієї ж кратності k.*

Доведення. Оскільки z_0 є коренем $f(z)$ кратності k, то

$$f(z_0) = f'(z_0) = \cdots = f^{(k-1)}(z_0) = 0; \quad f^{(k)}(z_0) \neq 0 \tag{8}$$

(див. п. 23.5). Але всі похідні від $f(z)$ мають також дійсні коефіцієнти. Тому, застосовуючи теорему 4 до многочленів $f^{(j)}(z)$, можемо зробити

висновок, що $f^{(j)}(\bar{z}_0) = 0$ $(j = 1, 2, \ldots, k-1)$. З другого боку, $f^{(k)}(\bar{z}_0) \neq 0$, бо в противному разі, за теоремою 4, число z_0, спряжене з $\bar{z}_0$, також було б коренем $f^{(k)}(z)$, що суперечить (8). Отже,

$$f(\bar{z}_0) = f'(\bar{z}_0) = \cdots = f^{(k-1)}(\bar{z}_0) = 0; \quad f^{(k)}(\bar{z}_0) \neq 0.$$

Це й означає, що $\bar{z}_0$ є коренем многочлена $f(z)$ кратності k.

Теореми 4, 5 пояснюють, чому при розв'язуванні і дослідженні рівнянь 2-го, 3-го і 4-го степенів з дійсними коефіцієнтами ми завжди діставали попарно спряжені комплексні корені, а також чому кубічне рівняння з дійсними коефіцієнтами завжди мало один або всі три дійсні корені.

Теорема 6. *Кожний многочлен над полем* **R**, *степінь якого перевищує 2, є звідним у цьому полі.*

Доведення. Позначимо даний многочлен степеня $n > 2$ над полем **R** через $f(z)$. Нехай z_0 — якийсь корінь цього многочлена. Якщо z_0 — дійсне число, то за теоремою Безу в полі дійсних чисел можливий розклад $f(z) = (z - z_0) f_1(z)$, причому $f_1(z) \in \mathbf{R}[z]$ і є многочленом ненульового степеня, бо степінь $f(z)$ перевищує 2. Отже, в цьому випадку $f(z)$ звідний у полі дійсних чисел.

Якщо ж z_0 — комплексний корінь многочлена $f(z)$, то, за теоремою 4, спряжене число $\bar{z}_0$ також є коренем многочлена $f(z)$. Тому $f(z)$ (розглядуваний як многочлен над полем **C**) ділиться на $z - z_0$, так і на $z - \bar{z}_0$, а отже, ділиться і на добуток

$$\varphi(z) = (z - z_0)(z - \bar{z}_0) = z^2 - (z_0 + \bar{z}_0) z + z_0 \bar{z}_0.$$

Але $\varphi(z)$ має дійсні коефіцієнти, бо сума $z_0 + \bar{z}_0$ і добуток $z_0 \bar{z}_0$ двох спряжених чисел є дійсні числа.

Отже, многочлен $f(z)$ з кільця $\mathbf{R}[z]$ ділиться на многочлен $\varphi(z)$ з цього ж кільця. Тому їх частка $p_1(z)$ також є многочленом над полем **R**, і в цьому полі справедливий розклад $f(z) = \varphi(z) \cdot f_1(z)$. Через те що $\varphi(z)$ має степінь 2, а $f(z)$ — степінь, вищий за 2, то $f(z)$ є многочленом ненульового степеня, чим доведено звідність $f(z)$ у полі дійсних чисел.

Що ж до многочленів 2-го степеня, то вони можуть бути незвідні в полі **R** (якщо мають комплексні корені). Отже, якщо в полі комплексних чисел незвідними були тільки многочлени першого степеня, то в полі дійсних чисел незвідними є многочлени першого степеня і деякі многочлени другого степеня.

Теорема 7. *Кожний многочлен $f(z)$ над полем дійсних чисел допускає єдиний розклад на незвідні множники в цьому полі виду:*

$$f(z) = a_n (z - z_1)^{k_1} \cdot (z - z_2)^{k_2} \ldots (z - z_l)^{k_l} \times$$
$$\times (z^2 + p_{l+1} z + q_{l+1})^{k_{l+1}} \ldots (z^2 + p_m z + q_m)^{k_m}. \tag{9}$$

Доведення. Як відомо з теорії подільності многочленів, для $f(z)$ у полі дійсних чисел можливий розклад виду

$$f(z) = [f_1(z)]^{k_1} \cdot [f_2(z)]^{k_2} \ldots [f_m(z)]^{k_m}, \tag{10}$$

причому $f_1(z), \ldots, f_m(z)$ — незвідні у полі R многочлени, які визначаються з точністю до сталого множника. Якщо поставити вимогу, щоб старші коефіцієнти цих многочленів дорівнювали 1, то вони визначатимуться однозначно. З теореми 6 випливає, що $f_k(z)$ є многочленами, не вище 2-го степеня. Припустимо, що $f_1(z), \ldots, f_l(z)$ є множники 1-го степеня, а $f_{l+1}(z), \ldots, f_m(z)$ — незвідні множники 2-го степеня (може трапитись, що $l = 0$ або $l = m$). Тоді (10) матиме вигляд:

$$f(z) = A(z+\alpha_1)^{k_1}(z+\alpha_2)^{k_2} \ldots (z+\alpha_l)^{k_l} \times$$
$$\times (z^2 + p_{l+1}z + q_{l+1})^{k_{l+1}} \ldots (z^2 + p_m z + q_m)^{k_m}.$$

Як і при виведенні формули (1), легко показати, що A дорівнює старшому коефіцієнту многочлена a_n, а $-\alpha_1, -\alpha_2, \ldots, -\alpha_l$ — його дійсні корені $z_1, z_2, \ldots, z_l$; отже, цей розклад збігається з (9) і теорему доведено.

Приклад. Многочлен $f(z) = z^4 + 4$ не має жодного дійсного кореня. Його коренями є попарно спряжені комплексні числа $z_1 = 1+i$, $z_2 = 1-i$, $z_3 = -1+i$, $z_4 = -1-i$. Відповідно до теореми 8, многочлен $f(z)$ повинен у полі R розкладатися на незвідні множники, не вище 2-го степеня. Щоб дістати цей розклад, розкладемо спочатку $f(z)$ на незвідні множники в полі комплексних чисел:

$$z^4 + 4 = [z-(1+i)]\cdot[z-(1-i)]\cdot[z-(-1+i)]\cdot[z-(-1-i)].$$

Перемножаючи попарно множники, які відповідають спряженим кореням, матимемо:

$$z^4 + 4 = (z^2 - 2z + 2)(z^2 + 2z + 2).$$

Цей самий розклад можна було б дістати й елемеитарно, помітивши, що

$$z^4 + 4 = (z^2+2)^2 - 4z^2.$$

Теорема 7 має істотне значення для розкладання дробово-раціональних функцій на елементарні дроби в полі дійсних чисел, що використовується в курсі математичного аналізу. Для розкладання правильного дробу на елементарні дроби у довільному полі P ми мали формулу (п. 24.3).

$$\frac{f(x)}{g(x)} = \sum_{l=1}^{m}\sum_{j=1}^{k_l}\frac{f_{lj}(x)}{[g_l(x)]^j}, \qquad (11)$$

де $g_l(x)$ — незвідні множники знаменника, а $f_{lj}(x)$ — многочлени, степінь яких менший за степінь $g_l(x)$.

Але в полі дійсних чисел многочлени $g_l(x)$ мають перший або другий степінь, а відповідні чисельники $f_{lj}(x)$ мають нульовий або перший степінь. Отже, на підставі розкладу (9) для знаменника $g(x)$ (нагадаємо, що цей многочлен вважаємо зведеним) можна конкретизувати формулу (11) для поля дійсних чисел, а саме:

$$\frac{f(x)}{g(x)} = \frac{f(x)}{(x-x_1)^{k_1}\ldots(x-x_l)^{k_l}(x^2+p_{l+1}x+q_{l+1})^{k_{l+1}}\ldots(x^2+p_mx+q_m)^{k_m}} =$$

$$= \frac{A_{11}}{x-x_1} + \frac{A_{12}}{(x-x_1)^2} + \cdots + \frac{A_{1k_1}}{(x-x_1)^{k_1}} +$$

. . . . . .

$$+ \frac{A_{l1}}{x-x_l} + \frac{A_{l2}}{(x-x_l)^2} + \cdots + \frac{A_{lk_l}}{(x-x_l)^{k_l}} + \frac{B_{l+1,1}x + C_{l+1,1}}{x^2+p_{l+1}x+q_{l+1}} +$$

$$+ \frac{B_{l+1,2}x + C_{l+1,2}}{(x^2+p_{l+1}x+q_{l+1})^2} + \cdots + \frac{B_{l+1,k_{l+1}}x + C_{l+1,k_{l+1}}}{(x^2+p_{l+1}x+q_{l+1})^{k_{l+1}}} + \qquad (12)$$

. . . . . .

$$+ \frac{B_{m1}x + C_{m1}}{x^2+p_mx+q_m} + \frac{B_{m2}x + C_{m2}}{(x^2+p_mx+q_m)^2} + \cdots + \frac{B_{mk_m}x + C_{mk_m}}{(x^2+p_mx+q_m)^{k_m}}.$$

Теореми 2 і 7 показують, що розклад многочлена $f(z)$ у полі комплексних чисел або в полі дійсних чисел на незвідні множники дає змогу знайти всі його корені (бо степінь незвідних множників не перевищує 2). Проте в загальному випадку, щоб знайти цей розклад, треба в свою чергу знайти корені даного многочлена. Задача знаходження коренів многочлена по суті рівнозначна задачі розкладання многочлена на незвідні множники в полі C або R.

§ 30. РОЗМІЩЕННЯ ДІЙСНИХ КОРЕНІВ МНОГОЧЛЕНА

30.1. Межі дійсних коренів. Теореми попередніх параграфів розв'язують ряд принципових питань щодо існування і числа коренів алгебраїчних рівнянь. Але щоб знайти корені рівняння з достатнім степенем точності, треба знати, як ці корені розміщені на комплексній площині або на дійсній осі. Зауважимо, що іноді навіть немає потреби знаходити числові значення коренів, а досить лише з'ясувати їх розміщення на площині (число дійсних, зокрема, додатних і від'ємних коренів тощо). Наприклад, одна з важливих проблем механіки — теорія стійкості — потребує з'ясування умов, при яких усі корені даного алгебраїчного рівняння мають від'ємні дійсні частини (тобто лежать на комплексній площині зліва від уявної осі). Питання цього циклу досить складні і потребують застосування теорії функцій комплексного змінного. Тому тут ми обмежимось розглядом питань, пов'язаних з розміщенням на дійсній осі коренів рівнянь з дійсними коефіцієнтами, що мають особливо важливе значення для задач практичного характеру.

Зробимо лише два зауваження щодо комплексних коренів многочленів. Ці зауваження є безпосередніми наслідками раніше з'ясованих фактів.

1. *Усі корені многочлена* $f(z) = a_nz^n + a_{n-1}z^{n-1} + \ldots + a_1z + a_0$ *лежать усередині круга з центром у точці* 0 *і радіусом*

$$N_0 = 1 + \frac{A}{|a_n|};\ A = \max\{|a_{n-1}|, |a_{n-2}|, \ldots, |a_1|, |a_0|\}. \qquad (1)$$

Це випливає з наслідку 1 теореми 1, п. 28.2.

2. *Комплексні корені многочлена з дійсними коефіцієнтами розміщені симетрично відносно дійсної осі.*

Це випливає з теореми 4, п. 29, 2 і з того, що комплексно спряжені числа розміщені симетрично відносно дійсної осі.

Переходячи тепер до розгляду дійсних коренів многочленів з дійсними коефіцієнтами, будемо знову позначати змінне буквою x, а не z.

З наведеного зауваження 1 дістаємо таке твердження:

Теорема 1. *Усі дійсні корені рівняння*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

містяться в інтервалі $(-N_0, N_0)$, *де* $N_0 = 1 + \frac{A}{|a_n|}$

$$(A = \max\{|a_{n-1}|, |a_{n-2}|, \ldots, |a_1|, |a_0|\}).$$

Справді, всі комплексні корені лежать у крузі $|z| < N_0$, а тому, якщо серед них є дійсні, то вони повинні потрапити в зазначений інтервал.

Приклад 1. Для рівняння

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

$A = 3$, $a_5 = 2$ і тому $N_0 = 1 + \frac{3}{2} = 2{,}5$. Отже, дійсні корені цього рівняння повинні лежати в інтервалі $(-2{,}5;\ 2{,}5)$.

Теорему 1 часто називають теоремою про межі коренів рівняння. Є чимало способів, які дають змогу з більшою точністю встановлювати межі дійсних коренів алгебраїчних рівнянь. Ми розглянемо лише один з них, так званий спосіб Ньютона.

Зробимо деякі попередні зауваження.

Число N_0, визначене за теоремою 1, дає одночасно верхню межу додатних коренів многочлена і нижню межу його від'ємних коренів, бо вказує інтервал $(-N_0, N_0)$, в якому лежать усі дійсні корені, якщо вони існують. Один з шляхів уточнення, звуження меж, між якими слід шукати дійсні корені, полягає в тому, щоб окремо *знаходити нижню і верхню межі додатних коренів та нижню і верхню межі від'ємних коренів* даного многочлена, тобто такі чотири числа m_+, M_+, m_-, M_-, що всі додатні корені многочлена лежать в інтервалі (m_+, M_+), а всі від'ємні — в інтервалі (m_-, M_-). Якщо многочлен має корінь нуль, досить розглянути многочлен, утворений з даного діленням на x.

Завдання полегшується тим, що фактично досить знати спосіб знаходження лише одного з цих чотирьох чисел, наприклад M_+ — верхньої межі додатних коренів. Знаходження інших трьох меж дійсних коренів рівняння $f(x) = 0$ легко звести до знаходження верхньої межі додатних коренів деяких допоміжних рівнянь.

Так, зробивши в рівнянні $f(x) = 0$ заміну змінного $x = \frac{1}{t}$, дістанемо рівняння $g(t) = 0$, корені якого t_i зв'язані з відповідними коренями x_i заданого рівняння співвідношенням $t_i = \frac{1}{x_i}$. Якщо M'_+ — верхня межа додатних коренів рівняння $g(t) = 0$, тобто $0 < t_i < M'_+$, то $x_i > \frac{1}{M'_+} > 0$, звідки видно, що за нижню межу додатних коренів рівняння $f(x) = 0$ можна взяти число $\frac{1}{M'_+}$:

$$m_+ = \frac{1}{M'_+}.$$

Аналогічно, заміна $x = -y$ переводить рівняння $f(x) = 0$ в рівняння $\varphi(y) = 0$, корені якого y_i зв'язані з відповідними коренями x_i рівняння $f(x) = 0$ рівністю $y_i = -x_i$. Якщо y_i $(i = 1, 2, \ldots, q)$ — всі додатні корені рівняння $\varphi(y) = 0$, то x_i $(i = 1, 2, \ldots, q)$ — всі від'ємні корені рівняння $f(x) = 0$. З нерівності $m''_+ < y_i < M''_+$ видно, що $-M''_+ < x_i < -m''_+$, тобто верхня і нижня межі від'ємних коренів рівняння $f(x) = 0$ виражаються через межі додатних коренів рівняння $\varphi(y) = 0$:

$$m_- = -M''_+; \quad M_- = -m''_+.$$

Отже, досить мати правило для знаходження верхньої межі додатних коренів многочлена.

Теорема 2 (Ньютона). *Число M є верхньою межею додатних коренів многочлена $f(x)$, якщо при $x = M$ многочлен $f(x)$ має додатне значення, а всі його похідні — невід'ємні значення.*

Доведення. Ураховуючи, що $f(x)$ є функція дійсної змінної, для якої справедлива формула Тейлора (відома з курсу аналізу), можемо записати:

$$f(x) = f(M) + \frac{f'(M)}{1}(x - M) + \frac{f''(M)}{2!}(x - M)^2 + \cdots + \frac{f^{(n)}(M)}{n!}(x - M)^n,$$

звідки безпосередньо видно, що при $x \geqslant M$ $f(x) > 0$, тобто всі дійсні корені многочлена $f(x)$ менші за M.

Оскільки знак многочлена і його похідних у точці M збігається з знаком відповідних коефіцієнтів розкладу за степенями $x - M$, на практиці числа M зручно підбирати за допомогою схеми Горнера послідовного ділення $f(x)$ на $x - M$ (п. 22.3). При цьому в більшості випадків немає потреби обчислювати всі коефіцієнти: як тільки в процесі ділення на $x - M$ дістаємо рядок з невід'ємних чисел,— можна прийняти M за верхню межу додатних коренів, бо дальше застосування схеми Горнера ніколи не приведе до від'ємних коефіцієнтів. Зокрема, якщо заданий многочлен $f(x)$ має невід'ємні коефіцієнти, можна вважати $M = 0$, тобто многочлен не має додатних коренів.

Приклад 2. Розглянемо рівняння

$$2x^5 + x^3 + x^2 - 2x - 3 = 0.$$

Поділивши його ліву частину на $x - 1$, маємо:

	2	0	1	1	−2	−3
1	2	2	3	4	2	−1

Тут є від'ємний коефіцієнт, але вже при $M = 1{,}1$ дістаємо рядок додатних чисел:

	2	0	1	1	—2	—3
1,1	2	2,2	3,42	4,76	3,24	0,56

Отже, можна взяти 1,1 за верхню межу додатних коренів: $M_+ = 1{,}1$.

Для знаходження нижньої межі додатних коренів замінимо в рівнянні (2) $x = \frac{1}{t}$. Дістанемо рівняння

$$3t^5 + 2t^4 - t^3 - t^2 - 2 = 0. \qquad (3)$$

Застосуємо до лівої частини цього рівняння схему Горнера при $M = 1$:

	3	2	—1	—1	0	—2
1	3	5	4	3	3	1

Це означає, що за верхню межу додатних коренів рівняння (3) можна взяти 1, тобто для нижньої межі додатних коренів нашого рівняння (2) дістаємо: $m_+ = \frac{1}{1} = 1$. Замінивши в (2) $x = -y$, дістанемо

$$2y^5 + y^3 - y^2 - 2y + 3 = 0. \qquad (4)$$

Ділимо на $y - 1$:

	2	0	1	—1	—2	3
1	2	2	3	2	0	3

Отже, для верхньої межі додатних коренів рівняння (4) ми дістали число 1, звідки для нижньої межі від'ємних коренів даного рівняння (2) маємо: $m_- = -1$.

Нарешті, замінивши в (4) $y = \frac{1}{z}$, дістанемо рівняння

$$3z^5 - 2z^4 - z^3 + z^2 + 2 = 0. \qquad (5)$$

Поділивши його ліву частину на $z - 1$, маємо:

	3	—2	—1	1	0	2
1	3	1	0	1	1	3

Отже, і $M_- = -1$. Разом із знайденим раніше результатом $m_- = -1$ це означає, що задане рівняння (2) зовсім не має від'ємних коренів.

Як бачимо метод Ньютона дає змогу істотно уточнити попередні відомості про межі коренів рівняння (2). Якщо застосування теореми 1 привело до досить «грубих» меж (—2,5; 2,5), то тепер ми знаємо, що додатні корені рівняння (2) (якщо вони існують) розміщені в інтервалі (1; 1,1), а від'ємних коренів це рівняння не має зовсім.

Метод Ньютона настільки елементарний, що його можна використати в середній школі, якщо тільки учні обізнані із схемою Горнера.

30.2. Число дійсних коренів. Знання числа і розміщення дійсних коренів многочленів є важливою передумовою застосування багатьох методів чисельного розв'язування рівнянь. В окремих випадках деякі відомості про число дійсних коренів можна дістати за допомогою досить поверхового аналізу. Так, з теореми 5, п. 29.2 можна зробити висновок, що число дійсних коренів многочлена з дійсними коефіцієнтами дорівнює степеню многочлена або на парне число менше. Іноді при знаходженні меж коренів виявляється, що многочлен не має додатних або від'ємних коренів. Однак для повної відповіді на питання про число дійсних коренів многочлена з дійсними коефіцієнтами (або навіть про число таких коренів на довільному, наперед заданому інтервалі дійсної осі) потрібні більш глибокі дослідження.

У багатьох випадках число дійсних коренів рівняння з дійсними коефіцієнтами можна визначити за простим правилом, яке дав Декарт. Перш ніж формулювати це правило, зробимо деякі зауваження.

1) Ми розглядатимемо **кількість змін знаків** у даній упорядкованій скінченній послідовності дійсних чисел

$$c_1, c_2, \ldots, c_m, \qquad (6)$$

розуміючи під цим *кількість пар сусідніх чисел цієї послідовності, які мають протилежні знаки.*

Наприклад, у послідовності —1, —2, 6, 3, —1, 4 є 3 зміни знаків, а в послідовності —1, —2, —6, —3, —1, —4 є 0 змін знаків.

Якщо які-небудь з чисел $c_1, c_2, \ldots, c_m$ дорівнюють нулю, то при підрахунку числа змін знаків їх до уваги не беруть.

Зауважимо, що *коли перше й останнє числа c_1 і c_m даної послідовності мають однакові знаки, то кількість змін знаків у послідовності* (6) *парна; якщо ж c_1 і c_m мають протилежні знаки, то кількість змін знаків — непарна.*

Справді, члени послідовності, які безпосередньо йдуть за кожною зміною знаків, мають знак, протилежний знаку тих членів, які передували зміні знаків. Отже, якщо остання зміна знаків має непарний номер, то числа послідовності, що йдуть за нею (і зокрема, c_m) матимуть знак, протилежний до c_1.

2) Припускатимемо, що розглядуваний многочлен не має кратних коренів, оскільки завжди можна відокремити кратні множники.

Правило Декарта. *Число додатних коренів многочлена з дійсними коефіцієнтами*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \qquad (7)$$

дорівнює числу змін знаків у послідовності його коефіцієнтів або на парне число менше.

Ми не будемо доводити це твердження, а лише пояснимо його на прикладах. (Доведення цього твердження можна знайти, наприклад, у книзі Костарчука В. М., Хацета Б. І. Курс вищої алгебри. Вид. 3. К., «Вища школа», 1969, § 28).

Приклади. 1. Для многочлена $f(x) = x^3 - 6x + 1$ число змін знаків у послідовності коефіцієнтів 1, 0, —6, 1 дорівнює двом. Отже, за теоремою Декарта він має або два або 0 додатних коренів.

2. Для лівої частини рівняння $x^4 + x^3 + x^2 + 2 = 0$ число змін знаків у послідовності коефіцієнтів дорівнює 0, тому додатних коренів це рівняння не має. Рівняння $x^3 + 2x^2 + x^2 - 2x - 3 = 0$ напевне має один додатний корінь.

Зауваження 1. *Правило Декарта можна застосувати і для оцінки числа від'ємних коренів рівняння з дійсними коефіцієнтами.*

Для цього в рівнянні

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

треба зробити заміну змінного $x = -y$. Зрозуміло, що число від'ємних коренів даного рівняння дорівнює числу додатних коренів рівняння $f(-y) = 0$, яке можна оцінити за правилом Декарта.

Приклади. 3. У рівнянні $x^3 - 6x + 1 = 0$ заміна $x = -y$ дає рівняння $y^3 - 6y - 1 = 0$. Тут число змін знаків дорівнює 1, отже, є один додатний корінь. Це означає, що задане рівняння має один від'ємний корінь.

4. Рівняння $x^4 + x^3 + x^2 + 2 = 0$ заміною $x = -y$ перетворюємо в рівняння $y^4 - y^3 + y^2 + 2 = 0$. Тут маємо дві зміни знаків. Отже, задане рівняння має або два від'ємних корені, або жодного.

5. Двочленне рівняння $x^n - a = 0$ при $a > 0$ має, за правилом Декарта, один додатний корінь. Якщо n парне, то число від'ємних коренів також дорівнює 1, при непарному ж n від'ємних коренів немає.

При $a < 0$ додатних коренів немає, а число від'ємних коренів дорівнює 1 при непарному n і нулю — при парному n.

Як бачимо, в цьому випадку правило Декарта дає повну відповідь на питання про число дійсних коренів рівняння.

Якщо дане рівняння повне, тобто жодний його коефіцієнт не дорівнює нулю, то число від'ємних коренів можна визначити і не виконуючи заміни $x = -y$. Справді, в цьому випадку число σ змін знаків у ряді коефіцієнтів многочлена $f(-x)$ дорівнює числу t збережень знаків у ряді коефіцієнтів многочлена $f(x)$. Отже, *число q від'ємних коренів повного рівняння дорівнює числу t збережень знаків у ряді його коефіцієнтів або на парне число менше.* Зауважимо також, що для такого рівняння $s + t = n$ (де n — степінь рівняння), бо ряд коефіцієнтів складається з $n + 1$ числа.

Зауваження 2. Коли наперед відомо, що всі корені даного рівняння $f(x) = 0$ дійсні, то правило Декарта дає точну відповідь на питання про число дійсних коренів, а саме: число додатних коренів дорівнює числу змін знаків у ряді коефіцієнтів многочлена $f(x)$, а число від'ємних коренів — числу змін знаків у ряді коефіцієнтів многочлена $f(-x)$.

Справді, нехай, як і вище, p і q — число додатних і від'ємних коренів даного многочлена $f(x)$ n-го степеня; s і σ — число змін знаків у ряді коефіцієнтів многочлена $f(x)$ і многочлена $f(-x)$ відповідно.

З умови, що всі корені дійсні, випливає: $p + q = n$. Якби рівняння було повним, то ми мали б також $s + \sigma = n$. Якщо ж деякі з коефіцієнтів многочлена $f(x)$ [а тому й многочлена $f(-x)$] перетворюються в нуль, то числа s і σ можуть тільки зменшитись. Тому в загальному випадку $s + \sigma \leqslant n$, звідки $s + \sigma \leqslant p + q$, або $(s - p) + (\sigma - q) \leqslant 0$. Але з правила Декарта ми знаємо, що $s - p \geqslant 0$, $\sigma - q \geqslant 0$. Тому насправді $s = p$, $\sigma = q$, чим наше твердження доведене.

Приклад 6. Рівняння (8) напевне має всі дійсні корені, бо його дискримінант $D = \frac{(-6)^3}{27} + \frac{1}{4} = -\frac{31}{4}$ від'ємний. Тому воно має два додатні і один від'ємний корінь.

Як бачимо, необхідною умовою того, щоб всі корені рівняння були дійсними, є рівність $s + \sigma = n$. Однак ця умова не є достатньою; так, рівняння $x^3 + x + 1 = 0$ задовольняє цю умову, але не має жодного дійсного кореня.

На жаль, у більшості випадків наперед невідомо, чи всі корені рівняння дійсні. У зв'язку з цим правило Декарта, хоч і зручне з точки зору простоти застосування, не дає повної відповіді на питання про число дійсних коренів рівнянь з дійсними коефіцієнтами та їх розподіл між додатною і від'ємною півосями.

30.3. Відокремлення коренів методом Штурма. Поставимо тепер питання про те, *скільки дійсних коренів рівняння з дійсними коефіцієнтами лежить у довільному, наперед заданому інтервалі (a, b) дійсної осі.*

Повну відповідь на це питання дає теорема Штурма [1], яку ми розглянемо в цьому пункті.

Нехай дано рівняння $f(x) = 0$. Насамперед побудуємо деяку послідовність многочленів, пов'язаних з многочленом $f(x)$,— так званий ряд функцій Штурма, який відіграє основну роль у методі Штурма. Припустимо, що $f(x)$ уже не має кратних коренів.

Знайдемо похідну $f'(x)$ і побудуємо для $f(x)$ та $f'(x)$ алгоритм, подібний до алгоритму Евкліда; відмінність полягатиме в тому, що всі остачі $R_k(x)$ ми братимемо з протилежними знаками, тобто замінюватимемо многочленами $F_k(x) = -R_k(x)$. Матимемо:

$$\begin{aligned}
f(x) &= f'(x) \cdot \overline{\Phi}_1(x) - F_1(x),\\
f'(x) &= F_1(x) \cdot \overline{\Phi}_2(x) - F_2(x),\\
F_1(x) &= F_2(x) \cdot \overline{\Phi}_3(x) - F_3(x),\\
&\cdots\cdots\cdots\cdots\cdots\\
F_{k-1}(x) &= F_k(x) \cdot \overline{\Phi}_{k+1}(x) - F_{k+1}(x),\\
&\cdots\cdots\cdots\cdots\cdots\\
F_{m-2}(x) &= F_{m-1}(x) \cdot \overline{\Phi}_m(x) - F_m,\\
F_{m-1}(x) &= F_m \cdot \overline{\Phi}_{m+1}(x).
\end{aligned} \tag{9}$$

Ми тут пишемо F_m, не зазначаючи аргументу, бо $f(x)$ та $f'(x)$ взаємно прості (за припущенням, $f(x)$ не має кратних коренів) і тому

[1] Ж. Ш. Штурм (1803—1855) — французький математик.

$F_m = \text{const}$. Послідовність многочленів

$$f(x),\ f'(x),\ F_1(x),\ F_2(x),\ \ldots,\ F_{m-1}(x),\ F_m \qquad (10)$$

і називається рядом функцій Штурма, або просто рядом Штурма для многочлена $f(x)$. Іноді для зручності позначатимемо

$$f'(x) = F_0(x), \quad f(x) = F_{-1}(x).$$

У методі Штурма нас цікавитимуть не самі функції ряду Штурма або їх значення, а лише з н а к и числових значень цих функцій. У зв'язку з цим функції ряду (10) можна знаходити з точністю до сталого д о д а т н о г о множника, тобто, виконуючи ділення з остачею, домножати на сталі множники; ці множники обов'язково повинні бути додатні, щоб не змінювались знаки значень многочленів.

П р и к л а д 1. Знайдемо ряд Штурма для многочлена

$$f(x) = x^3 - 6x + 1.$$

Через те що $f'(x) = 3x^2 - 6 = 3(x^2 - 2)$, то за $F_0(x)$ можна взяти $x^2 - 2$. Маємо:

$$\begin{array}{r|l} x^3 - 6x + 1 & x^2 - 2 \\ \underline{x^3 - 2x} & \overline{x} \\ -4x + 1 & \end{array}$$

Остача $R_1(x) = -4x + 1$, тому за $F_1(x)$ слід узяти $-R_1(x) = 4x - 1$. Далі ділимо $F_0(x)$ на $F_1(x)$:

$$\begin{array}{r|l} x^2 - 2 & 4x - 1 \\ \text{(помножаємо на 4)}\quad \underline{4x^2 - 8} & \overline{x + 1} \\ 4x^2 - x & \\ \hline x - 8 & \end{array}$$

$$\begin{array}{r} 4x - 32 \\ \underline{4x - 1} \\ -31 \end{array} \qquad \text{(помножаємо на 4)}$$

Остача R_2 дорівнює -31, а $-R_2 = 31$, тобто можна узяти $F_2 = 1$. Отже, остаточно маємо:

$$F_{-1}(x) = f(x) = x^3 - 6x + 1, \quad F_0(x) = x^2 - 2,$$

$$F_1(x) = 4x - 1, \quad F_2(x) = 1.$$

Введемо поняття ч и с л а з м і н з н а к і в у р я д і Ш т у р м а. Візьмемо в ряді функцій (10) $x = a$, де a — якесь дійсне число. Тоді скінченна послідовність функцій (10) перетворюється в послідовність чисел

$$f(a),\ f'(a),\ F_1(a),\ F_2(a),\ \ldots,\ F_{m-1}(a),\ F_m.$$

Число змін знаків у цій послідовності позначатимемо через $s(a)$ і називатимемо його числом змін знаків у ряді Штурма в точці a.

П р и к л а д 2. Для розглянутого вище многочлена $f(x) = x^3 - 6x + 1$

а) при $x = 2$ маємо: $F_{-1}(2) = -3$, $F_0(2) = 2$, $F_1(2) = 7$, $F_2 = 1$. У послідовності чисел $-3, 2, 7, 1$, очевидно, тільки одна зміна знаків, тобто $s(2) = 1$;

б) при $x = -1$ маємо: $F_{-1}(-1) = 6$, $F_0(-1) = -1$, $F_1(-1) = -5$, $F_2 = 1$. Тобто для цього значення x дістаємо дві зміни знаків: $s(-1) = 2$.

Зауважимо, що при зростанні x від $x = -1$ до $x = 2$ число змін знаків у ряді Штурма змінилося: $s(2) - s(-1) = -1$. Як виявиться далі, це пов'язано з тим, що в інтервалі $(-1, 2)$ міститься корінь многочлена $f(x) = x^3 - 6x + 1$.

Розглянемо основні властивості ряду функцій Штурма.

Лема 1. *Ніякі дві сусідні функції ряду Штурма* (10) *не мають спільних коренів.*

Д о в е д е н н я. Припустимо супротивне: нехай α є спільним коренем $F_k(x)$ і $F_{k+1}(x)$, тобто $F_k(\alpha) = F_{k+1}(\alpha) = 0$. Тоді з (9) маємо, що й

$$F_{k-1}(\alpha) = F_k(\alpha)\,\overline{\Phi}_{k+1}(\alpha) - F_{k+1}(\alpha) = 0.$$

Так само переконуємось, що $F_{k-2}(\alpha) = F_{k-3}(\alpha) = \ldots = F_1(\alpha) = f'(\alpha) = f(\alpha) = 0$. Але $f'(\alpha) = f(\alpha) = 0$ означає, що многочлен $f(x)$ має кратний корінь α, а, за припущенням, $f(x)$ кратних коренів не має. Отже, ми прийшли до суперечності, яка й доводить лему 1.

Лема 2. *Якщо α є коренем однієї з проміжних функцій ряду Штурма, то значення сусідніх з нею функцій ряду Штурма мають у цій точці протилежні знаки.*

Д о в е д е н н я. Нехай $F_k(\alpha) = 0$. Тоді, за лемою 1, $F_{k-1}(\alpha) \neq 0$, $F_{k+1}(\alpha) \neq 0$. Далі, з (9) маємо:

$$F_{k-1}(\alpha) = F_k(\alpha)\,\overline{\Phi}_{k+1}(\alpha) - F_{k+1}(\alpha) = -F_{k+1}(\alpha)$$

і лему 2 доведено.

Оскільки кожна функція ряду Штурма є многочлен і тому неперервна на всій дійсній осі, то вона може змінити знак лише при проходженні аргументу x через її корінь. Отже, якщо x, зростаючи, не проходить через корінь жодної функції ряду Штурма, то знаки всіх функцій цього ряду, а тому й число змін знаків у ньому залишаються незмінними.

Розглянемо тепер, як впливатиме на число змін знаків у ряді Штурма проходження x через корінь якоїсь з функцій цього ряду.

Лема 3. *Якщо x, зростаючи, проходить через корінь якої-небудь проміжної функції ряду Штурма, але не проходить через корінь $f(x)$, то число змін знаків у ряді Штурма при цьому не змінюється.*

Д о в е д е н н я. Нехай α — корінь якоїсь функції $F_k(x)$, де k — одне з чисел $0, 1, 2, \ldots, (m-1)$. Зауважимо, що $k \neq m$ (бо F_m є сталe число, відмінне від нуля) і $k \neq -1$, бо $F_{-1}(\alpha) = f(\alpha) \neq 0$ за умовою.

У точці α значення $F_{k-1}(\alpha)$ і $F_{k+1}(\alpha)$ відмінні від нуля (лема 1) і мають протилежні знаки (лема 2). Внаслідок неперервності цих функцій можна знайти такий окіл $(\alpha - \delta, \alpha + \delta)$ точки α, в якому функції $F_{k-1}(x)$ і $F_{k+1}(x)$ зберігають свої знаки. Адже, за відомою теоремою аналізу, якщо $f(x)$ неперервна в точці α і $f(\alpha) \neq 0$, то існує такий окіл точки α, в якому функція $f(x)$ зберігає знак $f(\alpha)$; тут ми вибираємо спільний окіл для двох функцій $F_{k-1}(x)$ і $F_{k+1}(x)$.

Дослідимо число змін знаків у ряді трьох функцій $F_{k-1}(x)$, $F_k(x)$ і $F_{k+1}(x)$, якщо x, зростаючи, змінюється в межах цього околу. Можливі при цьому варіанти розподілу знаків можна зобразити схематично такими табличками:

x	Знак $F_{k-1}(x)$	Знак $F_k(x)$	Знак $F_{k+1}(x)$	$s(x)$
$\alpha-\delta<$				
$<x<\alpha$	$-$		$+$	1
$x=\alpha$	$-$	0	$+$	1
$\alpha<x<$				
$<\alpha+\delta$	$-$		$+$	1

x	Знак $F_{k-1}(x)$	Знак $F_k(x)$	Знак $F_{k+1}(x)$	$s(x)$
$\alpha-\delta<$				
$<x<\alpha$	$+$		$-$	1
$x=\alpha$	$+$	0	$-$	1
$\alpha<x<$				
$<\alpha+\delta$	$+$		$-$	1

У цих табличках не заповнені місця для знаків $F_k(x)$ в околі точки $x=\alpha$, бо вони не впливають на число змін знаків: які б знаки не стояли на незаповнених місцях, у кожному матимемо одну зміну знака. Отже, і при $x<\alpha$, і при $x>\alpha$ число змін знаків у ряді функцій $F_{k-1}(x)$, $F_k(x)$, $F_{k+1}(x)$ однакове.

Щодо інших функцій ряду Штурма, то немає потреби розглядати докладно, як змінюються їх знаки при зростанні x в межах інтервалу $(\alpha-\delta, \alpha+\delta)$. Справді, можливі лише два випадки: або всі інші функції ряду Штурма не перетворюються в нуль при $x=\alpha$ і тому зберігають свої знаки в достатньо малому околі точки $x=\alpha$, або якась з проміжних функцій [але не сусідня з $F_k(x)$] перетворюється в нуль при $x=\alpha$, але тоді для неї і для двох сусідніх функцій число змін знаків залишається одним і тим самим, як і для трьох розглянутих функцій.

Отже, в цілому число змін знаків залишається сталим для всього ряду Штурма. Лему 3 доведено.

Лема 4. *Якщо x, зростаючи, проходить через корінь многочлена $f(x)$, то число змін знаків у ряді Штурма зменшується на одиницю.*

Доведення. Нехай α — корінь многочлена $f(x)$. Тоді $f(\alpha)=0$, але $f'(\alpha)\neq 0$, бо многочлен за умовою не має кратних коренів. Виберемо δ так, щоб в околі $(\alpha-\delta, \alpha+\delta)$ функція $f'(x)$ не змінювала знака. Можливі два випадки:

1) $f'(x)>0$ в цьому околі. Тоді $f(x)$ — зростаюча функція, тому при $\alpha-\delta<x<\alpha$ маємо $f(x)<0$, а при $\alpha<x<\alpha+\delta$ маємо $f(x)>0$. Відповідна таблиця знаків буде така:

x	Знак $f(x)$	Знак $f'(x)$	$s(x)$
$\alpha-\delta<$			
$<x<\alpha$	$-$	$+$	1
$x=\alpha$	0	$+$	0
$\alpha<x<$			
$<\alpha+\delta$	$+$	$+$	0

2) $f'(x)<0$ в околі $(\alpha-\delta, \alpha+\delta)$. Тоді $f(x)$ — спадна функція, тому дістаємо таку таблицю знаків:

x	Знак $f(x)$	Знак $f'(x)$	$s(x)$
$\alpha-\delta<$			
$<x<\alpha$	$+$	$-$	1
$x=\alpha$	0	$-$	0
$\alpha<x<$			
$<\alpha+\delta$	$-$	$-$	0

В обох випадках число змін знаків у цій частині ряду Штурма зменшується на одиницю. Щодо інших функцій ряду Штурма, то на підставі леми 3 для них число змін знаків не змінюється, навіть якщо x проходить при цьому через корені деяких з них. Лему 4 доведено.

Леми 3 і 4 показують, що на число змін знаків у ряді Штурма впливає лише проходження x через корені многочлена $f(x)$. Отже, зміна цього числа на певному проміжку може характеризувати число дійсних коренів многочлена $f(x)$ на цьому проміжку.

Теорема 3 (Штурма). *Якщо a і b $(a<b)$ — довільні дійсні числа, які не є коренями многочлена $f(x)$, то число p дійсних коренів многочлена $f(x)$ в інтервалі (a, b) дорівнює $p=s(a)-s(b)$, де $s(a)$ і $s(b)$ є число змін знаків у ряді Штурма відповідно в точках a і b.*

Доведення. Якщо x, зростаючи від a до b, не пройде через жодний корінь $f(x)$, то за лемою 3 $s(a)=s(b)$. Якщо ж x, зростаючи, пройде через p коренів многочлена $f(x)$, то при проходженні через кожний корінь число змін знаків зменшуватиметься на одиницю (за лемою 4), так що $s(b)$ буде на p одиниць менше ніж $s(a)$, тобто $s(a)-s(b)=p$. Теорему доведено.

Зауваження 1. У теоремі зазначено, що a і b не є коренями многочлена $f(x)$. Ця умова практично не становить труднощів при застосуванні теореми Штурма. Справді, якщо a (або b) є коренем многочлена $f(x)$, то питання про розміщення цього дійсного кореня розв'язується само собою, а для визначення положення інших коренів елід

змінити межі вибраного інтервалу або розглядати многочлен, який дістанемо діленням $f(x)$ на лінійний двочлен $x - a$.

Теорема Штурма справедлива і для випадку, коли кінці інтервалу можуть бути коренями многочлена. Тільки тоді $s(a) - s(b)$ є число коренів не на інтервалі (a, b), а на півінтервалі $(a, b]$.

З а у в а ж е н н я 2. Якщо якась з проміжних функцій ряду Штурма $F_k(x)$ не має дійсних коренів, то можна наступних функцій Штурма не знаходити і користуватися в теоремі Штурма «укороченим» рядом

$$f(x),\ f'(x),\ F_1(x),\ \ldots,\ F_k(x).$$

Справді, число змін знаків у «залишковому» ряді Штурма

$$F_k(x),\ F_{k+1}(x),\ \ldots,\ F_{m-1}(x),\ F_m \qquad (11)$$

є сталим при будь-якому x. Адже x може пройти лише через корінь проміжної функції ряду (11), що, за лемою 3, не впливає на число змін знаків у цьому ряді. Отже, «залишковий» ряд (11) не впливає на різницю $s(a) - s(b)$.

З а у в а ж е н н я 3. Метод Штурма можна застосувати і без попереднього відокремлення кратних коренів. Якщо $f(x)$ має кратні корені, то остання функція ряду Штурма $F_m(x)$ вже не є сталою. Але тоді $F_m(x)$ є спільним дільником $f(x)$, $f'(x)$ та всіх проміжних функцій ряду Штурма і можна розглянути ряд многочленів

$$\frac{f(x)}{F_m(x)},\ \frac{f'(x)}{F_m(x)},\ \frac{F_1(x)}{F_m(x)},\ \ldots,\ \frac{F_{m-1}(x)}{F_m(x)},\ 1, \qquad (12)$$

який вже має всі властивості, зазначені в лемах 1—4. Через те що число змін знаків у ряді (12) збігається з числом змін знаків у звичайному ряді Штурма $f(x)$, $f'(x)$, $F_1(x)$, ..., $F_{m-1}(x)$, $F_m(x)$, то теорема Штурма залишається в силі. Слід лише ураховувати, що вона дає в цьому випадку число дійсних коренів не самого многочлена $f(x)$, а многочлена $\frac{f(x)}{F_m(x)}$ (в якому вже немає кратних коренів), тобто число р і з н и х коренів многочлена $f(x)$ в інтервалі (a, b) без урахування їх кратності.

П р и к л а д 3. Застосуємо теорему Штурма до многочлена $f(x) = x^3 - 6x + 1$, для якого ми в попередньому прикладі побудували ряд Штурма і для інтервалу $(-1, 2)$ дістали такі результати щодо числа змін знаків у цьому ряді:

x	$f(x)$	$F_0(x)$	$F_1(x)$	F_2	$s(x)$
-1	6	-1	-5	1	2
2	-3	2	7	1	1

Ми бачимо, що $s(-1) - s(2) = 1$, тобто інтервал $(-1, 2)$ містить один корінь даного многочлена.

Зауважимо, що практично немає потреби обчислювати значення функцій ряду Штурма в точках a і b, можна лише визначати їх знаки в цих точках (виконуючи обчислення наближено). Відповідно до цього ми в таблицях наводитимемо лише знаки функцій Штурма.

Теорема Штурма дає змогу розв'язувати найрізноманітніші задачі щодо розміщення коренів многочленів на дійсній осі. Розглянемо дві такі задачі.

1. За допомогою ряду Штурма для довільного многочлена $f(x)$ над полем дійсних чисел можна *точно визначити загальне число дійсних коренів*, а також *число його додатних і від'ємних коренів*. Для цього досить застосувати теорему Штурма до інтервалів $(-N_0, 0)$ і $(0, N_0)$, де N_0 — межа модуля коренів, бо поза інтервалом $(-N_0, N_0)$ многочлен $f(x)$ дійсних коренів не має.

На практиці, щоб не підставляти чисел $\pm N_0$ у функції ряду Штурма, замість інтервалів $(-N_0, 0)$ і $(0, N_0)$ розглядають інтервали $(-\infty, 0)$ і $(0, +\infty)$. При цьому користуються тим, що, як випливає з наслідку теореми 1 п. 28.2, при $x \geqslant N_0$ знак многочлена визначається знаком його старшого члена. Тому під знаком многочлена «при $x = \infty$» розуміють знак його старшого члена при додатному x, а під знаком многочлена «при $x = -\infty$» — знак його старшого члена при від'ємному x.

П р и к л а д 4. Знайдемо число дійсних (додатних і від'ємних) коренів многочлена $f(x) = x^3 - 6x + 1$. Нагадаємо, що його функціями Штурма є: $F_0(x) = x^2 - 2$, $F_1(x) = 4x - 1$, $F_2 = 1$.

Складемо таблицю:

x	$f(x)$	$F_0(x)$	$F_1(x)$	F_2	$s(x)$
$-\infty$	$-$	$+$	$-$	$+$	3
0	$+$	$-$	$-$	$+$	2
$+\infty$	$+$	$+$	$+$	$+$	0

Отже, многочлен $f(x)$ має три дійсних корені, з них два додатних і один від'ємний. Це збігається з результатом дослідження цього многочлена за допомогою правила Декарта.

Якщо число окремо додатних і окремо від'ємних коренів нас не цікавить, то теорему Штурма застосовуємо відразу до інтервалу $(-\infty, \infty)$.

2. За допомогою теореми Штурма можна здійснювати так зване в і д о к р е м л е н н я д і й с н и х к о р е н і в. Відокремлення коренів полягає в знаходженні таких інтервалів, у кожному з яких лежить точно один дійсний корінь многочлена. Ця задача дуже важлива, бо більшість методів наближеного обчислення коренів потребує попереднього відокремлення їх. Практично відокремлення коренів зводиться до підбору кінців потрібних інтервалів. Покажемо на прикладі, як роблять такий підбір.

П р и к л а д 5. Відокремимо дійсні корені многочлена $f(x) = x^3 - 6x + 1$. За теоремою 1, дійсні корені $f(x)$ лежать в інтервалі $(-N_0, N_0)$, де $N_0 = 1 + \frac{A}{|a_n|}$. У цьому разі $N_0 = 7$. Отже, за найлівішу точку дослідження можна взяти

—7. Для $f(x)$ маємо такі функції ряду Штурма: $F_0(x) = x^2 - 2$; $F_1(x) = 4x - 1$, $F_2 = 1$.

Складемо таблицю:

x	$f(x)$	$F_0(x)$	$F_1(x)$	F_2	$s(x)$
-7	$-$	$+$	$-$	$+$	3
-3	$-$	$+$	$-$	$+$	3
-2	$+$	$+$	$-$	$+$	2
0	$+$	$-$	$-$	$+$	2
1	$-$	$-$	$+$	$+$	1
2	$-$	$+$	$+$	$+$	1
3	$+$	$+$	$+$	$+$	0
7	$+$	$+$	$+$	$+$	0

З цієї таблиці видно, що один корінь лежить в інтервалі (—3, —2), другий — в інтервалі (0, 1), а третій — в інтервалі (2, 3). Корені відокремлено.

Зауважимо, що на практиці відокремлюють корені і визначають загальне число дійсних коренів, як правило, одночасно, у зв'язку з чим спочатку в таблиці проставляють знаки функцій ряду Штурма при $x = -\infty$, $x = 0$, $x = +\infty$, а потім уже заповнюють проміжні рядки.

При цьому часто доцільно, щоб зменшити число спроб, визначити знаки функцій Штурма в точках, які приблизно є серединами вже досліджених проміжків. Так, у розглянутому вище прикладі ми з'ясували, що в інтервалі (—7, 0) лежить один корінь многочлена. Визначивши s (—3) ми з'ясовуємо, що один корінь лежить у правій «половині» цього інтервалу, а саме між —3 і 0 і т. д.

3. За допомогою ряду Штурма можна знайти просту ознаку того, що всі n коренів многочлена $f(x)$ n-го степеня є дійсні різні числа. Для цього, очевидно, потрібно, щоб у ряді Штурма при зростанні x від $-\infty$ до $+\infty$ число змін знаків зменшилось на n. У свою чергу, для цього насамперед потрібно, щоб число функцій у ряді Штурма було не меншим за $n + 1$. Оскільки за самою побудовою цього ряду воно не може бути більшим за $n + 1$, то у випадку всіх дійсних коренів ряд Штурма складається точно з $n + 1$ функцій, причому кожна наступна функція цього ряду є многочленом на одиницю нижчого степеня, ніж попередня. Тепер видно, що всі корені будуть дійсними, якщо $s(-\infty) = n$, а $s(+\infty) = 0$. Зрозуміло, що це має місце тоді і тільки тоді, коли старші коефіцієнти всіх функцій Штурма одного знака. Отже, *для того щоб усі корені многочлена $f(x)$ степеня n були дійсні й різні, необхідно і достатньо, щоб відповідний ряд Штурма складався з $n + 1$ многочленів, старші коефіцієнти яких усі того самого знака.* Зауважимо, що розглянутий вище многочлен $f(x) = x^3 - 6x + 1$ задовольняє ці умови.

Отже, теорема Штурма дає змогу повністю розв'язати всі питання, пов'язані з розміщенням дійсних коренів алгебраїчних рівнянь. Недоліком методу Штурма є деяка громіздкість його. Є й інші способи побудови послідовності функцій, що має всі властивості ряду функцій

Штурма (10) і тому може замінити його у теоремі Штурма, проте ці побудови не простіші за побудову ряду (10). Іноді застосовують також деякі «спрощені ряди функції Штурма», але вони в загальному випадку не дають точної відповіді на питання про число коренів (подібно до правила Декарта).

Відокремлення дійсних коренів є істотною передумовою застосування багатьох чисельних методів розв'язування алгебраїчних рівнянь. Ці методи, які мають велике практичне значення і для багатьох класів рівнянь, що не розв'язуються в радикалах, є єдиними способами визначення коренів, вивчаються в курсах математичного аналізу та обчислювальної математики і програмування.

§ 31. МНОГОЧЛЕНИ НАД ПОЛЕМ РАЦІОНАЛЬНИХ ЧИСЕЛ

31.1. Звідність і незвідність многочленів у полі раціональних чисел. Наявність раціональних коренів у довільно взятого алгебраїчного рівняння — явище досить рідкісне. Тому знаходження таких коренів не має великого практичного значення. Але, якщо многочлен $f(x)$ над полем Q раціональних чисел, або, що те саме, рівняння

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

з раціональними коефіцієнтами має раціональні корені, то в багатьох випадках ці корені можна знайти за допомогою цілком елементарних способів. Знаючи навіть один корінь r, можна спростити дане рівняння, звівши його до рівняння $(n - 1)$-го степеня діленням на $x - r$.

У зв'язку з цим доцільно ознайомитися з елементарними способами знаходження раціональних коренів многочленів з кільця Q $[x]$. Знання цих прийомів особливо важливе для вчителя математики, оскільки в практиці шкільного викладання алгебраїчні рівняння, як правило, зустрічаються з раціональними коефіцієнтами і досить часто з раціональними коренями. Тому знання специфічних властивостей многочленів з раціональними коефіцієнтами потрібне вчителю для вибору доцільного методу розв'язування конкретних рівнянь у шкільному курсі.

Основна відмінність многочленів над полем Q раціональних чисел від многочленів над полем R всіх дійсних чисел або полем C всіх комплексних чисел полягає в тому, що *існують многочлени з раціональними коефіцієнтами як завгодно високого степеня, незвідні у полі раціональних чисел,* тоді як у кільці C $[x]$ звідним є довільний многочлен, степінь якого вищий від одиниці (п. 29.1), а в кільці R $[x]$ звідним є кожний многочлен, степінь якого перевищує 2, навіть якщо цей многочлен не має жодного дійсного кореня (п. 29.2). Перш ніж доводити існування незвідних многочленів будь-якого степеня в полі Q, розглянемо деякі властивості многочленів з раціональними коефіцієнтами.

Насамперед зауважимо, що *будь-яке алгебраїчне рівняння з раціональними коефіцієнтами множенням на спільний знаменник усіх коефіцієнтів можна звести до рівняння з цілими коефіцієнтами.*

Приклад 1. Рівняння $\frac{1}{3}x^3+\frac{1}{2}x^2+x-\frac{2}{3}=0$ множенням на 6 можна звести до вигляду

$$2x^3+3x^2+6x-4=0. \qquad (1)$$

Оскільки зручніше мати справу з цілими, а не з дробовими числами, ми далі намагатимемося зводити всі питання щодо многочленів над полем Q до відповідних питань відносно многочленів з цілими коефіцієнтами. Зокрема, так можна зробити з питаниям про звідність многочлена в полі Q. Для цього нагадаємо спочатку означення примітивного многочлена відносно кільця Z (п. 25.4).

Означення. *Многочлен $p(x)$ з цілими коефіцієнтами називається примітивним, якщо його коефіцієнти не мають спільних дільників, відмінних від ± 1.*

Приклад 2. Многочлен, що стоїть у лівій частині рівняння (1), примітивний, тоді як многочлен $2x^4-4x^3+6x^2-12$ не є примітивним.

Як відомо (п. 25.4, лема 2), справедливе таке твердження:

Лема. *Добуток двох примітивних многочленів є примітивним многочленом.*

Розглянемо тепер питання про звідність многочлена з цілими коефіцієнтами в полі раціональних чисел.

Теорема 1. *Для того щоб многочлен $f(x)$ з цілими коефіцієнтами був звідним у полі Q раціональних чисел, необхідно і достатньо, щоб він був звідним у кільці Z цілих чисел, тобто щоб існували многочлени $f_1(x)$ і $f_2(x)$ ненульового степеня з цілими коефіцієнтами такі, що $f(x)=f_1(x)\cdot f_2(x)$.*

Доведення. *Необхідність.* Нехай дано многочлен з цілими коефіцієнтами $f(x)$, звідний у полі раціональних чисел, тобто $f(x)=g_1(x)\cdot g_2(x)$, де $g_1(x)$, $g_2(x)$ — многочлени ненульового степеня з раціональними коефіцієнтами. Ми повинні довести, що існують многочлени ненульового степеня $f_1(x)$ і $f_2(x)$ з цілими коефіцієнтами, добуток яких дорівнює $f(x)$.

Нехай після зведення коефіцієнтів до спільного знаменника і винесення цього знаменника за дужки многочлен $g_1(x)$ має вигляд: $g_1(x)=\frac{\alpha}{\beta}s_1(x)$, де $s_1(x)$ — многочлен з цілими коефіцієнтами, β — спільний знаменник коефіцієнтів $g_1(x)$, а α — найбільший спільний дільник коефіцієнтів многочлена, що утворюється з $g_1(x)$ після зведення до спільного знаменника. Зрозуміло, що $s_1(x)$ є примітивний многочлен. Можна вважати, що $(\alpha,\beta)=1$, бо інакше можна було б виконати скорочення. Аналогічно, для $g_2(x)$ маємо: $g_2(x)=\frac{\gamma}{\delta}s_2(x)$, де $s_2(x)$ — примітивний многочлен, а $(\gamma,\delta)=1$. Отже,

$$f(x)=\frac{\alpha\gamma}{\beta\delta}\cdot s_1(x)\cdot s_2(x).$$

Доведемо, що $\frac{\alpha\gamma}{\beta\delta}$ дорівнює цілому числу. Справді, припустимо, що $\frac{\alpha\gamma}{\beta\delta}=\frac{p}{q}$, де p і q — взаємно прості числа. Добуток $s_1(x)\cdot s_2(x)=s(x)$ за лемою є примітивний многочлен. Нехай c_k — якийсь коефіцієнт $s(x)$. Добуток $\frac{p}{q}c_k$ має бути цілим числом при будь-якому k, бо $f(x)=\frac{p}{q}s(x)$ має цілі коефіцієнти. Але p взаємно просте з q; тому c_k має ділитися на q. Оскільки те саме повинно справджуватись для всіх коефіцієнтів c_k, дістаємо суперечність з тим, що $s(x)$ примітивний многочлен. Отже, $\frac{\alpha\gamma}{\beta\delta}=m$, де m — ціле число. Узявши $f_1(x)=ms_1(x)$, $f_2(x)=s_2(x)$, дістанемо $f(x)=f_1(x)\cdot f_2(x)$, де $f_1(x)$, $f_2(x)$ — многочлени ненульового степеня з цілими коефіцієнтами.

Достатність. Якщо $f(x)$ звідний у кільці цілих чисел, то він і поготів звідний у полі раціональних чисел, бо кожний многочлен з цілими коефіцієнтами є многочленом над полем раціональних чисел. Теорему доведено.

Теорема 1 повністю зводить питання про звідність многочленів у полі Q до звідності многочленів у кільці Z цілих чисел.

Теорема 2 (Ейзенштейна[1]). *Якщо в многочлені з цілими коефіцієнтами*

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

коефіцієнти $a_0, a_1, \ldots, a_{n-1}$ діляться на деяке просте число p, причому a_0 не ділиться на p^2, а старший коефіцієнт a_n не ділиться на p, то многочлен $f(x)$ незвідний у полі раціональних чисел.

Доведення. Згідно з теоремою 1, досить показати, що $f(x)$ при цих умовах не може бути добутком двох многочленів ненульового степеня з цілими коефіцієнтами. Припустимо супротивне, тобто що

$$f(x)=(b_rx^r+b_{r-1}x^{r-1}+\cdots+b_1x+b_0)(c_sx^s+c_{s-1}x^{s-1}+\cdots$$
$$\cdots+c_1x+c_0) \qquad (r+s=n).$$

Вважатимемо для конкретності, що $r\geqslant s$. Маємо:

$$\begin{aligned}
a_0&=b_0c_0,\\
a_1&=b_1c_0+b_0c_1,\\
a_2&=b_2c_0+b_1c_1+b_0c_2,\\
&\cdots\cdots\cdots\cdots\\
a_r&=b_rc_0+b_{r-1}c_1+\cdots+b_{r-s}c_s,\\
&\cdots\cdots\cdots\cdots\cdots\\
a_n&=b_rc_s.
\end{aligned} \qquad (2)$$

За умовою, a_0, тобто b_0c_0, повинно ділитися на p, але не може ділитися на p^2. Отже, на p ділиться лише одне з чисел: b_0 або c_0. Нехай, наприклад, b_0 ділиться на p, а c_0 не ділиться. Але тоді з другої рівності системи (2) дістаємо, що b_1 ділиться на p (бо a_1 ділиться на p за умовою, а c_0 не ділиться). Тепер уже b_0 і b_1 діляться на p, тому з третьої рівності видно, що й b_2 ділиться на p. Так можна показати, що всі коефіцієнти $b_0, b_1, b_2, \ldots, b_{r-1}$ і b_r ділиться на p. Але це неможливо, бо тоді й a_n ділилося б на p (це випливає з останньої рівності (2)), що суперечить умові теореми. Теорему доведено.

[1] Ф. Ейзенштейн (1823—1852) — німецький математик.

За допомогою теореми 2, яку часто називають критерієм Ейзенштейна, можна розв'язати питання про незвідність у полі раціональних чисел ряду многочленів з кільця $Q[x]$.

Приклад 3. Многочлен $f(x) = x^4 - 2x^3 - 4x^2 + 2x - 6$ напевне незвідний у полі раціональних чисел, бо його коефіцієнти задовольняють умови критерію Ейзенштейна при $p = 2$.

Проте основне значення теореми 2 полягає в тому, що з неї випливає *існування многочленів довільного степеня з цілими коефіцієнтами, незвідних у полі раціональних чисел*. Зокрема, при всякому натуральному n і простому p многочлен $f(x) = x^n + p$ напевне незвідний у полі Q. Зрозуміло, що такі многочлени можна побудувати багатьма способами.

Отже, на основі критерію Ейзенштейна ми довели справедливість такого твердження.

Теорема 3. *У кільці многочленів над полем раціональних чисел є многочлени довільного степеня, незвідні у полі* Q.

Теорема 2 дає достатню умову незвідності многочлена в полі Q. Легко дати і достатню умову звідності многочлена у полі Q, а саме:

Теорема 4. *Якщо многочлен $f(x)$ з раціональними коефіцієнтами, степінь якого більший за одиницю, має хоча б один раціональний корінь r, то $f(x)$ звідний у полі раціональних чисел.*

Доведення. Справді, за наслідком з теореми Безу, $f(x)$ ділиться на $x - r$, тобто $f(x) = (x - r) f_1(x)$, причому частка $f_1(x)$ є многочленом ненульового степеня над тим самим полем раціональних чисел. Теорему доведено.

Зауважимо, що *коли $f(x)$ має цілі коефіцієнти, а r — цілий корінь, то й $f_1(x)$ має цілі коефіцієнти.*

Це видно з того, що коефіцієнти частки за схемою Горнера обчислюють за допомогою дій додавання і множення.

Твердження, обернене до теореми 4, неправильне: многочлен $f(x)$ може не мати жодного раціонального кореня, але бути звідним у полі раціональних чисел. Наприклад, многочлен $x^4 - 4$ звідний у полі Q: $x^4 - 4 = (x^2 - 2)(x^2 + 2)$, але раціональних коренів не має.

Проте у випадку многочлена третього степеня обернена теорема справедлива. Доведемо її, оскільки вона буде потрібна нам далі.

Теорема 5. *Якщо многочлен третього степеня $f(x)$ з раціональними коефіцієнтами не має раціональних коренів, то він незвідний у полі раціональних чисел.*

Доведення. Припустимо супротивне. Нехай $f(x) = f_1(x) \times f_2(x)$, де $f_1(x)$ і $f_2(x)$ — многочлени ненульового степеня з кільця $Q[x]$. Оскільки сума степенів $f_1(x)$ і $f_2(x)$ дорівнює 3, то один з цих многочленів обов'язково має степінь 1, а другий — степінь 2. Нехай $f_1(x)$ є многочленом 1-го степеня з раціональними коефіцієнтами $f_1(x) = ax + b$. Але тоді число $x_0 = -\frac{b}{a}$ є раціональним коренем многочлена $f_1(x)$, а тому й многочлена $f(x)$. Виходить, що $f(x)$ має раціональні корені, що суперечить умові. Теорему доведено.

Приклад 4. Многочлен $x^3 - 4$ напевне незвідний у полі Q, бо не має раціональних коренів. За критерієм Ейзенштейна, питання про незвідність цього многочлена безпосередньо розв'язати не можна.

31.2. Раціональні корені многочленів з раціональними коефіцієнтами. Розглянемо елементарні способи знаходження раціональних коренів рівнянь з раціональними коефіцієнтами. Ми вже зазначали (п. 31.1), що рівняння над полем Q завжди можна вважати рівнянням з цілими коефіцієнтами.

Основне практичне значення для цього питання має така теорема.

Теорема 6. *Щоб число $\frac{p}{q}$, де p і q — взаємно прості числа, було коренем рівняння*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \qquad (3)$$

з цілими коефіцієнтами, необхідно, щоб p було дільником вільного члена a_0 а q — дільником старшого коефіцієнта a_n цього рівняння. —

Доведення. Нехай $\frac{p}{q}$ є коренем рівняння (3). Тоді

$$a_n \left(\frac{p}{q}\right)^n + a_{n-1} \left(\frac{p}{q}\right)^{n-1} + \cdots + a_1 \left(\frac{p}{q}\right) + a_0 = 0,$$

або

$$a_n p^n + a_{n-1} q p^{n-1} + \cdots + a_1 q^{n-1} p + a_0 q^n = 0.$$

Через те що всі доданки, крім останнього, діляться на p і сума (нуль) ділиться на p, то й $a_0 q^n$ ділиться на p. Але p і q взаємно прості, тому p і q^n також взаємно прості. Отже, a_0 ділиться на p. Аналогічно, всі доданки, крім першого, діляться на q, отже, і $a_n p^n$ ділиться на q, звідки a_n ділиться на q. Теорему доведено.

Наслідок. *Якщо старший коефіцієнт рівняння з цілими коефіцієнтами дорівнює* 1, *то всі раціональні корені цього рівняння є цілі числа і дільники вільного члена.*

Теорема 6 дає змогу за коефіцієнтами a_0 і a_n даного рівняння знайти всі раціональні числа, які можуть бути коренями цього рівняння. Далі, підставляючи ці числа в рівняння, можна знайти, які з них є його коренями.

Приклад 1. Рівняння

$$6x^4 + 19x^3 - 7x^2 - 26x + 12 = 0 \qquad (4)$$

може мати раціональними коренями лише такі числа:

$$\pm 1;\ \pm 2,\ \pm 3;\ \pm 4;\ \pm 6;\ \pm 12,$$

$$\pm \frac{1}{2};\ \pm \frac{3}{2};\ \pm \frac{1}{3};\ \pm \frac{2}{3};\ \pm \frac{4}{3};\ \pm \frac{1}{6},$$

тобто ті нескоротні дроби $\frac{p}{q}$, чисельники яких є дільниками числа 12, а знаменники — дільниками числа 6.

На практиці частіше користуються не самою теоремою 6, а її наслідком. Адже кожне рівняння з цілими коефіцієнтами можна звести до

рівняння з цілими коефіцієнтами, в якому старший коефіцієнт дорівнює 1. Для цього треба помножити рівняння (3) на a_n^{n-1} і зробити заміну $a_n x = y$.

Приклад 2. Для рівняння (1) (п. 31.1) після множення на 2^2 і заміни $y = 2x$ дістаємо:

$$y^3 + 3y^2 + 12y - 16 = 0. \qquad (5)$$

До такого рівняння вже можна застосувати наслідок з теореми 6. Його раціональними коренями можуть бути лише цілі числа — дільники вільного члена, тобто ± 1; ± 2; ± 4; ± 8; ± 16.

Проте таке зведення не завжди доцільно робити, як показує приклад рівняння (4), в якому після множення на 6^3 дістаємо вільний член $12 \cdot 6^3 = 2592$, який має дуже багато дільників. Зауважимо, що при цьому зведенні додатковий множник, а тому й число спроб, зростає при збільшенні степеня рівняння, тоді як при безпосередньому використанні теореми 6 число спроб залежить лише від числових значень коефіцієнтів.

Теорема 6 дає необхідну умову того, щоб раціональне число було коренем даного рівняння з цілими коефіцієнтами. Проте бажано мати кілька необхідних умов, щоб за допомогою їх зменшувати число проб — підставлянь у рівняння.

Теорема 7. *Для того щоб* $\frac{p}{q}$, де $(p, q) = 1$, *було раціональним коренем многочлена з цілими коефіцієнтами*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

необхідно, щоб при довільному цілому k *число* $f(k)$ *ділилося на* $p - qk$ (якщо тільки $p - qk \neq 0$).

Доведення. Поділимо $f(x)$ на $x - k$. За теоремою Безу маємо:

$$f(x) = (x - k)(b_{n-1}x^{n-1} + \cdots + b_1 x + b_0) + f(k). \qquad (6)$$

Усі коефіцієнти частки b_{n-1}, ..., b_1, b_0 є цілими числами. Підставляючи в (6) $x = \frac{p}{q}$ і враховуючи, що $f\left(\frac{p}{q}\right) = 0$, матимемо:

$$-f(k) = \left(\frac{p}{q} - k\right)\left[b_{n-1}\left(\frac{p}{q}\right)^{n-1} + \cdots + b_1\left(\frac{p}{q}\right) + b_0\right].$$

Помноживши обидві частини на q^n, дістанемо

$$-q^n f(k) = (p - qk)[b_{n-1}p^{n-1} + \cdots + b_1 pq^{n-2} + b_0 q^{n-1}],$$

звідки видно, що $q^n f(k)$ ділиться на $p - qk$, якщо $p - qk \neq 0$.

Покажемо тепер, що q і $p - qk$ взаємно прості. Справді, якби вони мали спільний дільник α, $|\alpha| \neq 1$, то і $p = qk + (p - qk)$ мало б цей дільник, що неможливо, бо $(p, q) = 1$. Отже, q і $p - qk$, а тому і q^n і $p - qk$ взаємно прості. Через те що добуток $q^n f(k)$ ділиться на $p - qk$, то це означає, що $f(k)$ ділиться на $p - qk$. Теорему доведено.

Наслідок. *Якщо старший коефіцієнт* a_n *даного многочлена* $f(x)$ *з цілими коефіцієнтами дорівнює одиниці, то його раціональними коренями можуть бути лише такі цілі числа* p, *для яких* $f(k)$ *ділиться на* $p - k$ *при всякому цілому* k, *при якому* $p - k \neq 0$.

такий, що кожне наступне поле є квадратичним розширенням попереднього і останнє Δ_{k_1} містить число q_1. Приєднавши до цього ланцюжка поле $\Delta_{k_1+1} = \Delta_{k_1}(\sqrt{q_1})$, дістанемо ланцюжок послідовних квадратичних розширень

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_{k_1} \subseteq \Delta_{k_1+1},$$

останнє з яких містить $\sqrt{q_1}$.

Таке саме міркування можна провести для числа q_2. Але тепер за вихідне поле ланцюжка квадратичних розширень візьмемо не Δ, а Δ_{k_1+1}. Це можна зробити тому, що всі числа поля Δ містяться і в Δ_{k_1+1}. Отже, існує ланцюжок квадратичних розширень

$$\Delta_{k_1+1} \subseteq \Delta_{k_1+2} \subseteq \cdots \subseteq \Delta_{k_2} \subseteq \Delta_{k_2+1},$$

причому $\sqrt{q_1} \in \Delta_{k_2+1}$ і $\sqrt{q_2} \in \Delta_{k_2+1}$. Продовжуючи цей процес, побудуємо ланцюжки квадратичних розширень

$$\Delta_{k_2+1} \subseteq \Delta_{k_2+2} \subseteq \cdots \subseteq \Delta_{k_3} \subseteq \Delta_{k_3+1} \qquad (\sqrt{q_1}, \sqrt{q_2}, \sqrt{q_3} \in \Delta_{k_3+1}),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\Delta_{k_{n-1}+1} \subseteq \Delta_{k_{n-1}+2} \subseteq \cdots \subseteq \Delta_{k_n} \subseteq \Delta_{k_n+1}$$

(всі $\sqrt{q_i} \in \Delta_{k_n+1}$ $(i = 1, 2, \ldots, n)$).

Зауважимо, що деякі з цих ланцюжків можуть складатися з самого лише поля Δ_{k_i+1} (це буде у випадках, коли $\sqrt{q_{i+1}} \in \Delta_{k_i+1}$).

Оскільки поле Δ_{k_n+1} містить усі числа $\sqrt{q_i}$ $(i = 1, 2, \ldots, n)$, то воно містить і число ξ. Отже, ланцюжок полів

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_{k_1} \subseteq \Delta_{k_1+1} \subseteq \Delta_{k_2+2} \subseteq \cdots \subseteq \Delta_{k_2} \subseteq$$

$$\subseteq \Delta_{k_2+1} \subseteq \ldots \Delta\, \Delta_{k_n} \subseteq \Delta_{k_n+1}$$

задовольняє умови 1—3. Цим доведено, що умови теореми справедливі для будь-якого числа ξ, для якого $p_\xi = m$. Згідно з принципом індукції, це твердження правильне для усіх натуральних p_ξ, тобто для всіх чисел, які виражаються у квадратних радикалах через числа поля Δ.

Теорему доведено повністю.

Приклади. 2. Побудуємо ланцюжок квадратичних розширень для числа α, вираженого в квадратних радикалах через раціональні числа у формі (7).

Очевидно, таким ланцюжком буде $\Delta \subseteq \Delta_1 \subseteq \Delta_2$, де $\Delta = \mathbf{Q}$, $\Delta_1 = \mathbf{Q}(\sqrt{5})$, $\Delta_2 = \Delta_1(\sqrt{2\sqrt{5} - 10})$. Зауважимо, що Δ_2 є квадратичне розширення поля Δ_1, бо його дістаємо з Δ_1 приєднанням елемента $\alpha_1 = \sqrt{2\sqrt{5} - 10}$, що є коренем многочлена 2-го степеня $f(x) = x^2 - (2\sqrt{5} - 10)$ над полем Δ_1.

3. Для числа $\xi = \sqrt{\sqrt{1 + \sqrt{2}} + \sqrt{5}}$ відповідний ланцюжок складається з 5 полів: $\Delta = \mathbf{Q}$, $\Delta_1 = \mathbf{Q}(\sqrt{5})$, $\Delta_2 = \Delta_1(\sqrt{2})$, $\Delta_3 = \Delta_2(\sqrt{1 + \sqrt{2}})$, $\Delta_4 = \Delta_3(\sqrt{\sqrt{1 + \sqrt{2}} + \sqrt{5}})$.

рівняння з цілими коефіцієнтами, в якому старший коефіцієнт дорівнює 1. Для цього треба помножити рівняння (3) на a_n^{n-1} і зробити заміну $a_n x = y$.

Приклад 2. Для рівняння (1) (п. 31.1) після множення на 2^2 і заміни $y = 2x$ дістаємо:

$$y^3 + 3y^2 + 12y - 16 = 0. \tag{5}$$

До такого рівняння вже можна застосувати наслідок з теореми 6. Його раціональними коренями можуть бути лише цілі числа — дільники вільного члена, тобто ± 1; ± 2; ± 4; ± 8; ± 16.

Проте таке зведення не завжди доцільно робити, як показує приклад рівняння (4), в якому після множення на 6^3 дістаємо вільний член $12 \cdot 6^3 = 2592$, який має дуже багато дільників. Зауважимо, що при цьому зведенні додатковий множник, а тому й число спроб, зростає при збільшенні степеня рівняння, тоді як при безпосередньому використанні теореми 6 число спроб залежить лише від числових значень коефіцієнтів.

Теорема 6 дає необхідну умову того, щоб раціональне число було коренем даного рівняння з цілими коефіцієнтами. Проте бажано мати кілька необхідних умов, щоб за допомогою їх зменшувати число проб — підставлянь у рівняння.

Теорема 7. *Для того щоб* $\frac{p}{q}$, де $(p, q) = 1$, *було раціональним коренем многочлена з цілими коефіцієнтами*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

необхідно, щоб при довільному цілому k *число* $f(k)$ *ділилося на* $p - qk$ (якщо тільки $p - qk \neq 0$).

Доведення. Поділимо $f(x)$ на $x - k$. За теоремою Безу маємо:

$$f(x) = (x - k)(b_{n-1}x^{n-1} + \cdots + b_1 x + b_0) + f(k). \tag{6}$$

Усі коефіцієнти частки $b_{n-1}, \ldots, b_1, b_0$ є цілими числами. Підставляючи в (6) $x = \frac{p}{q}$ і враховуючи, що $f\left(\frac{p}{q}\right) = 0$, матимемо:

$$-f(k) = \left(\frac{p}{q} - k\right)\left[b_{n-1}\left(\frac{p}{q}\right)^{n-1} + \cdots + b_1\left(\frac{p}{q}\right) + b_0\right].$$

Помноживши обидві частини на q^n, дістанемо

$$-q^n f(k) = (p - qk)[b_{n-1}p^{n-1} + \cdots + b_1 pq^{n-2} + b_0 q^{n-1}],$$

звідки видно, що $q^n f(k)$ ділиться на $p - qk$, якщо $p - qk \neq 0$.

Покажемо тепер, що q і $p - qk$ взаємно прості. Справді, якби вони мали спільний дільник α, $|\alpha| \neq 1$, то і $p = qk + (p - qk)$ мало б цей дільник, що неможливо, бо $(p, q) = 1$. Отже, q і $p - qk$, а тому і q^n і $p - qk$ взаємно прості. Через те що добуток $q^n f(k)$ ділиться на $p - qk$, то це означає, що $f(k)$ ділиться на $p - qk$. Теорему доведено.

Наслідок. *Якщо старший коефіцієнт* a_n *даного многочлена* $f(x)$ *з цілими коефіцієнтами дорівнює одиниці, то його раціональними коренями можуть бути лише такі цілі числа* p, *для яких* $f(k)$ *ділиться на* $p - k$ *при всякому цілому* k, *при якому* $p - k \neq 0$.

Справді, за наслідком з теореми 6, при $a_n = 1$ маємо $q = 1$, звідки й випливає наше твердження.

Теорема 7 дає змогу дістати довільну кількість необхідних умов того, щоб число $\frac{p}{q}$ було коренем даного рівняння, бо числу k можна надавати довільних цілих значень. Найбільш поширені на практиці умови, що відповідають $k = \pm 1$, бо вирази $f(1)$ і $f(-1)$ легко обчислити. Їх можна сформулювати так: *Щоб число* $\frac{p}{q}$ *було раціональним коренем многочлена з цілими коефіцієнтами, треба щоб* $\frac{f(1)}{p-q}$ *і* $\frac{f(-1)}{p+q}$ *були цілими числами.*

Приклад 3. Застосуємо наслідок теореми 7 до рівняння (4):

$$6x^4 + 19x^3 - 7x^2 - 26x + 12 = 0.$$

Воно може мати, як ми бачили, лише такі раціональні корені:

$$\pm 1;\ \pm 2;\ \pm 3;\ \pm 4;\ \pm 6;\ \pm 12;$$
$$\pm \frac{1}{2};\ \pm \frac{3}{2};\ \pm \frac{1}{3};\ \pm \frac{2}{3};\ \pm \frac{4}{3};\ \pm \frac{1}{6}. \tag{7}$$

Для даного рівняння $f(1) = 4$, $f(-1) = 18$. Подивимось, для яких з чисел (7), крім ± 1, відношення $\frac{4}{p-q}$ є цілим числом. Такими є числа

$$+2;\ \pm 3;\ +\frac{1}{2};\ +\frac{3}{2};\ \pm \frac{1}{3};\ +\frac{2}{3};\ +\frac{4}{3}. \tag{8}$$

Як бачимо, замість 24 чисел (7) маємо вже тільки 9 чисел (8), які можуть бути коренями рівняння (4). Тепер застосуємо до чисел (8) умову, щоб $\frac{18}{p+q}$ було цілим числом. Залишаються числа 2; -3; $\frac{1}{2}$; $-\frac{1}{3}$. Ці числа доцільно вже перевірити безпосередньо. Якби їх було ще багато, можна було б застосувати теорему 7 при $k = 2$.

Перевірка показує, що лише два з цих чисел, а саме $\frac{1}{2}$ і -3 є раціональними коренями даного рівняння. Ділячи ліву частину рівняння на $x + 3$ і $x - \frac{1}{2}$ за схемою Горнера, дістанемо квадратне рівняння, що має корені $x_{1,2} = \frac{-1 \pm \sqrt{13}}{3}$. Разом з числами $\frac{1}{2}$ і -3 маємо всі чотири корені рівняння (4).

На практиці доцільно поєднувати наведені тут прийоми знаходження раціональних коренів рівнянь із способами обчислення меж дійсних коренів (п. 29.1). Так, для рівняння (4) вже найпростіший метод обмеження коренів (теорема 1, п. 29.1) показує, що всі дійсні його корені лежать в інтервалі $\left(-5\frac{1}{3}, 5\frac{1}{3}\right)$, бо для цього рівняння $N_0 = 1 + \frac{26}{6} = 5\frac{1}{3}$; тому потреба у перевірці деяких з чисел (7) відпадає відразу. Тим більше слід ураховувати межі коренів, якщо користуватись наслідками з теорем 6 і 7.

Зауважимо, що замість того, щоб підставляти раціональне число r у многочлен $f(x)$, можна поділити $f(x)$ на $x - r$. Якщо остача дорівнюватиме нулю, то r буде коренем $f(x)$, у противному разі r не буде

коренем $f(x)$ (п. 23.1). Цей метод має ту перевагу перед звичайною підстановкою, що він дає змогу відразу дістати коефіцієнти частки від ділення на $x - r$, якщо r буде коренем, і перейти до розгляду рівняння нижчого степеня.

Розділ VIII

АЛГЕБРАЇЧНІ РОЗШИРЕННЯ ЧИСЛОВИХ ПОЛІВ

§ 32. АЛГЕБРАЇЧНІ ЧИСЛА І СКІНЧЕННІ РОЗШИРЕННЯ ЧИСЛОВИХ ПОЛІВ

Означення 1. *Число називається алгебраїчним, якщо воно є коренем деякого многочлена з раціональними коефіцієнтами.*

Очевидно, будь-яке раціональне число r алгебраїчне, бо його можна розглядати як корінь многочлена $f(x) = x - r$ з раціональними коефіцієнтами. Проте ірраціональні числа також можуть бути алгебраїчними. Наприклад, числа $\sqrt{2}$, $\sqrt[3]{5}$ алгебраїчні, бо вони є кореняими многочленів $x^2 - 2$, $x^3 - 5$ (відповідно) над полем $\mathbf{Q}$.

Проте не всі ірраціональні числа алгебраїчні. Існує безліч ірраціональних чисел, які не є коренями жодного многочлена над полем $\mathbf{Q}$. Такі числа називаються трансцендентними. Прикладами трансцендентних чисел можуть бути числа π, $\lg 2$, $2^{\sqrt{2}}$ та інші.

Поняття алгебраїчного числа можна узагальнити, увівши таке означення.

Означення 2. *Число α називається алгебраїчним відносно числового поля Δ, якщо воно є коренем деякого многочлена над полем Δ. Число, яке не є алгебраїчним відносно поля Δ, називається трансцендентним відносно Δ.*

Оскільки в цьому розділі ми розглядатимемо тільки числові поля, слово «числове» надалі, як правило, не писатимемо.

Очевидно, алгебраїчні в розумінні означення 1 числа — це числа, алгебраїчні відносно поля $\mathbf{Q}$ раціональних чисел. Оскільки $\mathbf{Q}$ є підполем будь-якого поля Δ, то ці числа алгебраїчні відносно довільного поля. Зрозуміло також, що кожне число поля Δ є алгебраїчним відносно цього самого поля Δ.

Нехай α — корінь многочлена степеня n над полем Δ:

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0. \qquad (1)$$

Вважатимемо, що многочлен (1) незвідний у полі Δ, бо в противному разі ми могли б розглядати той незвідний множник многочлена, який має α своїм коренем.

Якщо $g(x)$ — будь-який многочлен над полем Δ, коренем якого є α, то внаслідок незвідності $f(x)$ многочлен $g(x)$ ділиться на $f(x)$ (взаємно простими вони бути не можуть, бо мають спільний множник

$x - \alpha$) і тому має степінь, не нижчий за n. Зокрема, якщо $g(x)$ — також незвідний многочлен, то він збігається з $f(x)$ з точністю до сталого множника. Отже, зведений многочлен $f(x)$ — *єдиний незвідний многочлен над полем Δ, який має α своїм коренем, а його степінь n найнижчий серед степенів усіх многочленів з коренем α.*

Означення 3. *Зведений многочлен $f(x)$, незвідний у полі Δ, який має α своїм коренем, називається мінімальним многочленом числа α, а його степінь n — степенем алгебраїчного числа α відносно поля Δ.*

Якщо α — число першого степеня відносно Δ, то $\alpha \in \Delta$. При $n > 1$ з незвідності $f(x)$ випливає, що $\alpha \notin \Delta$. Справді, якби $\alpha \in \Delta$, то подільність многочлена $f(x)$ на лінійний двочлен $x - \alpha$ означала б, що $f(x)$ звідний у полі Δ.

32.2. Просте алгебраїчне розширення поля. Нехай тепер дано довільну числову множину M. Очевидно, завжди знайдуться числові поля, які містять всі числа множини M, наприклад поле всіх комплексних чисел.

Мінімальним полем $P\{M\}$, що містить дану числову множину M, називається поле, яке є перетином усіх числових полів, що містять множину M. Це означення спирається на відомий факт (1, § 13), що непорожній перетин довільної сукупності полів знову є полем.

Зрозуміло, що для будь-якої числової множини M мінімальне поле $P\{M\}$ завжди існує і є підполем довільного іншого поля, яке містить множину M.

Приклади. 1. Нехай множина M складається лише з одного числа 1. Тоді кожне числове поле містить цю множину. Мінімальним полем, яке містить це число, є, очевидно, поле $\mathbf{Q}$ раціональних чисел. Справді, поле $\mathbf{Q}$ належить усім числовим полям. З другого боку, ніяке ірраціональне число не може належати всім числовим полям, бо воно не належить хоча б числовому полю $\mathbf{Q}$. Зауважимо, що $\mathbf{Q}$ природно назвати просто *мінімальним числовим полем.*

2. Розглянемо тепер мінімальне поле, що містить число $\sqrt{2}$. Очевидно, що це є поле $\mathbf{Q}(\sqrt{2})$ чисел виду $a + b\sqrt{2}$, де a, b — довільні раціональні числа. Справді, ця числова множина утворює поле, яке, очевидно, містить $\sqrt{2}$. З другого боку, кожне інше поле P, яке містить $\sqrt{2}$, повинно містити і все поле $\mathbf{Q}(\sqrt{2})$: адже разом з раціональними числами і числом $\sqrt{2}$ в P повинні входити всі числа $a + b\sqrt{2}$, що є результатами додавання і множення згаданих чисел.

Нехай Δ — деяке числове поле і α — число, яке не належить цьому полю ($\alpha \notin \Delta$). Розглянемо мінімальне поле $P\{\Delta, \alpha\}$, яке містить і поле Δ, і число α. Очевидно, $P\{\Delta, \alpha\}$ є розширенням поля Δ, причому мінімальним розширенням, яке містить число α. Справді, всяке розширення поля Δ, яке містить α, міститиме і $P\{\Delta, \alpha\}$ за означенням мінімального поля.

Відомо (1, § 13), що мінімальне розширення поля Δ, яке містить число $\alpha \notin \Delta$, називають також *розширенням поля Δ, утвореним приєднанням числа α*, і позначають через $\Delta(\alpha)$. Аналогічно можна розглядати розширення $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$, утворене приєднанням кількох чисел $\alpha_1, \alpha_2, \ldots, \alpha_k$ до поля Δ, тобто мінімальне поле $P\{\Delta, \alpha_1, \alpha_2, \ldots, \alpha_k\}$, яке містить як Δ, так і числа $\alpha_1, \alpha_2, \ldots, \alpha_k$. Розширення, утворені приєднанням одного числа, часто називають простими.

Відповідно до цієї термінології розглянуте у прикладі 2 поле чисел виду $a + b\sqrt{2}$ (a, b — раціональні) можна назвати простим розширенням поля раціональних чисел, утвореним приєднанням числа $\sqrt{2}$.

Цей приклад є окремим випадком важливого класу простих розширень, який ми зараз розглянемо докладніше.

Означення 1. *Поле $\Delta(\alpha)$, утворене приєднанням до поля Δ числа α, алгебраїчного відносно поля Δ, називається простим алгебраїчним розширенням поля Δ.*

Будова простого алгебраїчного розширення характеризується такою теоремою.

Теорема 1. *Поле $\Delta(\alpha)$, утворене з поля Δ приєднанням кореня α, незвідного у полі Δ многочлена n-го степеня*

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

складається з усіх чисел виду

$$\zeta = c_0 + c_1\alpha + c_2\alpha^2 + \cdots + c_{n-1}\alpha^{n-1}, \qquad (2)$$

де $c_0, c_1, \ldots, c_{n-1}$ — довільні числа з поля Δ.

Доведення. Насамперед покажемо, що всі числа виду (2) утворюють поле. Те, що сума і різниця чисел виду (2) належить до тієї самої сукупності,— очевидно. Розглянемо добуток і частку таких чисел. Зрозуміло, що число виду (2) можна розглядати як результат підстановки α замість x у деякий многочлен $q(x)$ над полем Δ не вище $(n-1)$-го степеня: $\zeta = q(\alpha)$.

Нехай маємо два числа $\zeta_1 = q_1(\alpha)$, $\zeta_2 = q_2(\alpha)$. Тоді добуток $\zeta_1 \cdot \zeta_2 = q_1(\alpha) \cdot q_2(\alpha) = q(\alpha)$, де $q(x)$ — многочлен, степінь якого вже може перевищувати $n-1$. Поділимо $q(x)$ на $f(x)$ з остачею. Маємо:

$$q(x) = f(x)\varphi(x) + r(x), \qquad (3)$$

де степінь остачі $r(x)$ менший за степінь многочлена $f(x)$, тобто не перевищує $n-1$. Підставляючи α в тотожність (3), маємо $q(\alpha) = r(\alpha)$, тобто $\zeta_1 \cdot \zeta_2 = r(\alpha)$. Іншими словами, добуток чисел ζ_1 і ζ_2 є числом виду (2), бо $r(x)$ — многочлен, степінь якого не перевищує $n-1$.

Перейдемо до розгляду частки. Досить показати, що для будь-якого числа $\zeta = q(\alpha) \neq 0$ виду (2) $\frac{1}{r}$ також буде числом виду (2). Оскільки $f(x)$ — незвідний у полі Δ многочлен, то многочлен $q(x)$ або взаємно простий з $f(x)$, або ділиться на $f(x)$. Проте останній випадок неможливий, бо $q(x)$ має степінь, менший ніж степінь мінімального многочлена $f(x)$, і тому $(f, q) = 1$.

Отже (п. 22.5), існує єдина пара многочленів $\varphi_1(x)$ і $\varphi_2(x)$ таких, що справджується тотожність $f(x) \cdot \varphi_1(x) + q(x) \cdot \varphi_2(x) = 1$.

Узявши тут $x = \alpha$ і врахувавши, що $f(\alpha) = 0$, дістанемо $q(\alpha) \times \varphi_2(\alpha) = 1$, тобто $\zeta \cdot \varphi_2(\alpha) = 1$. Отже, $\frac{1}{\zeta} = \varphi_2(\alpha)$. Якщо многочлен $\varphi_2(x)$ має степінь, менший за n, то твердження доведено. Якщо ж $\varphi_2(x)$ має степінь, не менший за n, то ділимо $\varphi_2(x)$ на многочлен $f(x)$ з остачею, тобто $\varphi_2(x) = f(x) \cdot \varphi(x) + r(x)$, звідки $\varphi_2(\alpha) = r(\alpha) = \frac{1}{\zeta}$, і степінь $r(x)$ менший за степінь $f(x)$. Отже, $\frac{1}{\zeta}$ є числом виду (2).

Таким чином, числа виду (2) справді утворюють поле. Позначимо це поле через Δ_1.

Залишається довести, що $\Delta_1 = \Delta(\alpha)$. Оскільки поле Δ_1 містить як поле Δ, так і число α, то воно містить і $\Delta(\alpha)$, яке за означенням є мінімальним полем з такими властивостями, тобто $\Delta_1 \supseteq \Delta(\alpha)$. З другого боку, будь-яке поле, яке включає α і поле Δ, повинно включати і всі числа виду (2), які утворюються з чисел поля Δ і числа α за допомогою дій множення і додавання. Отже, $\Delta(\alpha) \supseteq \Delta_1$. З обох встановлених співвідношень випливає, що $\Delta_1 = \Delta(\alpha)$, чим і завершується доведення теореми.

Для дальшого викладу важливо виділити такий окремий випадок теореми 1.

Наслідок. *Якщо α — корінь многочлена другого степеня над полем Δ*

$$f(x) = x^2 + px + q,$$

причому $\alpha \notin \Delta$, то просте алгебраїчне розширення $\Delta(\alpha)$ поля Δ, утворене приєднанням числа α, складається з усіх чисел виду $a + b\alpha$, де a, b — довільні числа з поля Δ.

Приклади 3. Поле $\mathbf{Q}(\sqrt{2})$ утворюється приєднанням до поля $\mathbf{Q}$ кореня $\sqrt{2}$ незвідомого у полі Q многочлена другого степеня $f(x) = x^2 - 2$. Елементы поля Q $(\sqrt{2})$ мають вигляд $a + b\sqrt{2}$, де a, b — раціональні числа.

4. Розглянемо будову елементів поля $\mathbf{Q}(\sqrt[3]{2})$. Число $\alpha = \sqrt[3]{2}$ є коренем незвідного в полі $\mathbf{Q}$ многочлена третього степеня $f(x) = x^3 - 2$. Тому всі елементи ζ поля $\mathbf{Q}(\sqrt[3]{2})$, за теоремою 1, мають вигляд $\zeta = c_0 + c_1\sqrt[3]{2} + c_2\sqrt[3]{2^2}$, де c_0, c_1, c_2 — раціональні числа.

5. Поле $\mathbf{C}$ комплексних чисел утворюється, як відомо, з поля $\mathbf{R}$ дійсних чисел приєднанням до нього кореня i незвідного в $\mathbf{R}$ многочлена другого степеня $f(x) = x^2 + 1$. З доведеної теореми випливає, що всі елементи $\zeta \in \mathbf{C}$, тобто всі комплексні числа, мають вигляд $\zeta = a + bi$, де a, b — дійсні числа.

Введемо таке означення.

Означення 2. *Якщо корінь α квадратного тричлена над полем Δ не належить полю Δ, то просте алгебраїчне розширення $\Delta(\alpha)$, утворене з поля Δ приєднанням до нього числа α, називається квадратичним розширенням поля Δ.*

Розглянуте вище поле $\mathbf{Q}(\sqrt{2})$ є, очевидно, квадратичним розширенням поля Q раціональних чисел, утворене приєднанням кореня многочлена $f(x) = x^2 - 2$.

У дальшому викладі квадратичні розширення числових полів відіграватимуть важливу роль.

32.3. Знищення ірраціональності в знаменнику. Нехай дано дріб $\frac{p(\alpha)}{q(\alpha)}$, де $p(x)$ і $q(x)$ — многочлени над полем Q, а α — ірраціональний корінь незвідного многочлена $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ з раціональними коефіцієнтами (при цьому, звичайно, $q(\alpha) \neq 0$). Слід виконати такі тотожні перетворення даного дробу, щоб позбутись ірраціональності у знаменнику.

З самого доведення теореми 1 зрозуміло, що треба робити. Якщо $\deg q \geqslant n$, то, ділячи $q(x)$ на $f(x)$ з остачею, дістаємо рівність $q(x) = f(x)s(x) + r(x)$.

Підставляючи значення $x = \alpha$, дістаємо $q(\alpha) = r(\alpha)$, тому $\frac{p(\alpha)}{q(\alpha)} = \frac{p(\alpha)}{r(\alpha)}$, де $\deg r < \deg f$. Отже, завжди можна вважати степінь знаменника заданого дробу меншим за n. Але тоді зрозуміло, що $g(x)$ і $f(x)$ взаємно прості, адже $f(x)$ — незвідний многочлен.

Нехай тепер $\varphi_1(x)$ та $\varphi_2(x)$ — такі многочлени над $\mathbf{Q}$, що

$$\varphi_1(x) \cdot f(x) + \varphi_2(x)\, q(x) = 1. \tag{4}$$

Тоді $\frac{1}{q(\alpha)} = \varphi_2(\alpha)$ і

$$\frac{p(\alpha)}{q(\alpha)} = p(\alpha)\, \varphi_2(\alpha). \tag{5}$$

Таким чином, щоб знищити ірраціональність у знаменнику дробу $\frac{p(\alpha)}{q(\alpha)}$, де α — корінь незвідного многочлена $f(x)$, потрібно виконати такі дії:

1) Якщо $\deg q \geqslant n$, то замінити $q(x)$ числом $r(\alpha)$, де $r(x)$ — остача від ділення $q(x)$ на $f(x)$.

2) Знайти многочлени $\varphi_1(x)$ та $\varphi_2(x)$, які задовольняють рівність (4).

3) Обчислити $\varphi_2(\alpha)$ і подати дріб $\frac{p(\alpha)}{q(\alpha)}$ за формулою (5).

Приклад. Розглянемо дріб

$$\frac{\sqrt[3]{2}+4}{2-\sqrt[3]{2}}.$$

Тут $f(x) = x^3 - 2$, $p(x) = x + 4$, $q(x) = -x + 2$, $\deg q < \deg f$. Знаходимо $\varphi_1(x)$ та $\varphi_2(x)$ такі, що

$$\varphi_1(x)(x^3 - 2) + \varphi_2(x)(-x + 2) = 1.$$

Виконавши відповідні обчислення, дістаємо

$$\varphi_1(x) = \frac{1}{6}, \quad \varphi_2(x) = \frac{1}{6}x^2 + \frac{1}{3}x + \frac{2}{3}.$$

Тепер за формулою (5) маємо:

$$\frac{\sqrt[3]{2}+4}{2-\sqrt[3]{2}} = (\sqrt[3]{2}+4)\left(\frac{1}{6}\sqrt[3]{2^2} + \frac{1}{3}\sqrt[3]{2} + \frac{2}{3}\right) = \sqrt[3]{4} + 2\sqrt[3]{2} + 3.$$

32.4. Скінченні розширення полів. Теорема 1 і наведені приклади показують, що числа ζ поля $\Delta(\alpha)$ мають специфічну структуру. Вони являють собою суму виду

$$\zeta = c_0 + c_1\alpha + c_2\alpha^2 + \cdots + c_{n-1}\alpha^{n-1},$$

де кожний член є добутком елемента c_k поля Δ на елемент α^k ($k = 0, 1, \ldots, n-1$) поля $\Delta(\alpha)$. Отже, можна сказати, що довільний елемент ζ поля $\Delta(\alpha)$, де α є коренем незвідного в Δ многочлена степеня n, є *лінійною комбінацією елементів* $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ *з коефіцієнтами з поля* Δ. Оскільки сума елементів $\Delta(\alpha)$ і добуток їх на числа з поля Δ є знову елементи поля $\Delta(\alpha)$, то $\Delta(\alpha)$ *можна розглядати як лінійний простір над полем* Δ. Більше того, $\Delta(\alpha)$ є алгеброю над полем Δ; проте тут для нас істотні лише властивості $\Delta(\alpha)$ як лінійного простору.

Узагальнюючи це спостереження, розглянемо деяке поле Ω і його підполе Δ. Ω є лінійний простір над полем Δ. Елементами цього простору є числа з поля Ω, а операціями — додавання елементів поля Ω і множення їх на числа з поля Δ.

Розглянемо питання про базис і розмірність цього лінійного простору, нагадавши перед цим деякі означення і факти, відомі з теорії лінійних просторів (1, розд. VIII), стосовно до ситуації, яка нас тут цікавить.

Сукупність чисел $\alpha_1, \alpha_2, \ldots, \alpha_k$ з поля Ω називатимемо лінійно незалежною системою елементів відносно поля Δ, якщо рівність

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k = 0, \tag{6}$$

де $\lambda_1, \lambda_2, \ldots, \lambda_k$ належать полю Δ, можлива лише при всіх $\lambda_i = 0$. Якщо ж рівність (6) справджується і тоді, якщо хоч одне з λ_i не дорівнює нулю (нагадаємо, що $\lambda_i \in \Delta$), то система елементів $\alpha_1, \alpha_2, \ldots, \alpha_k$ називається *лінійно залежною відносно поля* Δ.

Проілюструємо ці означення прикладами.

Приклади. **1.** У полі $\mathbf{Q}(\sqrt{2})$ числа $\alpha_1 = 1$, $\alpha_2 = \sqrt{2}$ є лінійно незалежною системою елементів відносно поля $\mathbf{Q}$. Справді, нехай $\lambda_1, \lambda_2 \in \mathbf{Q}$. Запишемо рівність (6) для цього випадку: $\lambda_1 \cdot 1 + \lambda_2 \cdot \sqrt{2} = 0$. Покажемо, що тут $\lambda_1 = \lambda_2 = 0$. Припустимо супротивне, тобто що $\lambda_1 \neq 0$, $\lambda_2 \neq 0$ (випадок, коли одне з λ_1, λ_2 дорівнює, а друге не дорівнює нулю, неможливий, що видно безпосередньо). Тоді маємо $\sqrt{2} = -\frac{\lambda_1}{\lambda_2}$, тобто $\sqrt{2}$ є раціональне число, що неможливо.

2. У тому самому полі $\mathbf{Q}(\sqrt{2})$ два числа $\alpha_1 = a_1 + b_1\sqrt{2}$ і $\alpha_2 = a_2 + b_2\sqrt{2}$ можуть бути лінійно залежною системою елементів відносно поля $\mathbf{Q}$. Справді, з рівності $\lambda_1\alpha_1 + \lambda_2\alpha_2 = 0$, при умові, наприклад, $\lambda_2 \neq 0$, маємо $\alpha_2 = -\frac{\lambda_1}{\lambda_2}\alpha_1$, або $a_2 + b_2\sqrt{2} = -\frac{\lambda_1}{\lambda_2}(a_1 + b_1\sqrt{2})$. Отже, числа α_1 і α_2 утворюють лінійно залежну систему елементів відносно поля $\mathbf{Q}$, якщо вони задовольняють умови $a_2 = -\frac{\lambda_1}{\lambda_2}a_1$, $b_2 = -\frac{\lambda_1}{\lambda_2}b_1$, тобто коли справджується рівність $\frac{a_2}{a_1} = \frac{b_2}{b_1}$.

3. У полі $\mathbf{Q}(\sqrt[3]{2})$ система елементів $1, \sqrt[3]{2}, \sqrt[3]{2^2}$ є лінійно незалежною відносно поля $\mathbf{Q}$. Рекомендуємо читачеві довести це самостійно.

Нагадаємо тепер деякі прості властивості лінійно залежних і лінійно незалежних систем елементів.

1) *Кожна частина лінійно незалежної системи елементів відносно поля* Δ *є також лінійно незалежною системою елементів відносно поля* Δ.

2) *Будь-яка система елементів* $\alpha_1, \alpha_2, \ldots, \alpha_k, 0$ *поля* Ω, *яка включає число нуль, є лінійно залежною відносно поля* Δ.

3) *Якщо система елементів* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *поля* Ω *є лінійно залежною відносно поля* Δ, *то хоч один з елементів системи є лінійною комбінацією інших елементів з коефіцієнтами з поля* Δ.

4) *Якщо хоч один з елементів системи* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *поля* Ω *є лінійною комбінацією інших елементів цієї системи з коефіцієнтами з поля* Δ, *то система* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *є лінійно залежною відносно поля* Δ.

Повернемося тепер знову до результатів теореми 1. Як ми зазначали, ця теорема з'ясовує структуру розширення $\Delta(\alpha)$ поля Δ, де α —

корінь незвідного в полі Δ многочлена $f(x)$ n-го степеня. Поле $\Delta(\alpha)$ побудоване так. Існує n елементів цього поля $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ таких, що кожний елемент $\zeta \in \Delta(\alpha)$ є лінійною комбінацією цих елементів з коефіцієнтами з поля Δ.

Покажемо, що *сукупність чисел* $1, \alpha, \ldots, \alpha^{n-1}$ *є лінійно незалежною системою елементів відносно поля* Δ.

Справді, запишемо рівність виду (6):

$$\lambda_0 \cdot 1 + \lambda_1 \cdot \alpha + \cdots + \lambda_{n-1} \cdot \alpha^{n-1} = 0,$$

де $\lambda_i \in \Delta$. Якщо ця рівність справджується, коли не всі λ_i дорівнюють нулю, то це означає, що α є коренем деякого многочлена $\varphi(x) = \lambda_0 + \lambda_1 x + \ldots + \lambda_{n-1} x^{n-1}$ з коефіцієнтами з поля Δ, степінь якого не перевищує $n-1$. Але це неможливо, бо, як відомо, многочлен $f(x)$ степеня n є мінімальним многочленом числа α.

Отже, коли α — алгебраїчне число n-го степеня відносно поля Δ, то елементи розширення $\Delta(\alpha)$ є лінійними комбінаціями (з коефіцієнтами з поля Δ) елементів лінійно незалежної відносно Δ системи $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$.

У загальному випадку розглянемо деяке числове поле Δ і його розширення Ω ; припустимо, що в полі Ω існує лінійно незалежна відносно Δ система елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$ така, що кожний елемент $\zeta \in \Omega$ подається у вигляді лінійної комбінації чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$ з коефіцієнтами з поля Δ.

Систему $\alpha_1, \alpha_2, \ldots, \alpha_n$ можна назвати б а з и с о м розширення Ω відносно поля Δ, бо вона утворює базис лінійного простору Ω над полем Δ. Кількість елементів цього базису скінченна і дорівнює n.

Такі розширення поля мають назву с к і н ч е н н и х. Дамо повне означення згаданих понять.

Означення. *Розширення* Ω *поля* Δ *називається скінченним, якщо в полі* Ω *існує така лінійно незалежна відносно поля* Δ *система елементів* $\alpha_1, \alpha_2, \ldots, \alpha_n$, *що будь-який елемент* $\zeta \in \Omega$ *є лінійною комбінацією цих елементів з коефіцієнтами з поля* Δ:

$$\zeta = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n. \qquad (7)$$

Система елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$ *називається базисом поля* Ω *відносно поля* Δ.

Базис скінченного розширення Ω можна вибрати не одним способом. Проте всі базиси поля Ω мають те саме число елементів n. Більш того: *довільна лінійно незалежна система з* n *елементів є базисом* (1, § 32). Отже, число n є характеристикою скінченного розширення Ω поля Δ, незалежною від вибору базису. Число n називається *степенем розширення* Ω над полем Δ і позначається символом $(\Omega : \Delta)$. Зрозуміло, що $(\Omega : \Delta)$ є *розмірністю лінійного простору* Ω над полем Δ.

Розглянемо деякі властивості базисів.

Якщо n *— степінь розширення* Ω *над полем* Δ, *то будь-яка лінійно незалежна відносно* Δ *система* $\alpha_1, \alpha_2, \ldots, \alpha_m$ *елементів поля* Ω *не може містити більш як* n *чисел.*

Справді, якщо $m > n$, то виберемо з нашої лінійно незалежної системи частину елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$. Ця нова система також лінійно незалежна і містить n чисел. Тому її можна прийняти за базис поля Ω відносно поля Δ. Отже, кожний елемент поля Ω, зокрема α_m, буде лінійною комбінацією елементів $\alpha_1, \alpha_2, \ldots, \alpha_n$. Виходить, що система $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_m$ не є лінійно незалежною, бо один з її елементів є лінійною комбінацією інших. Тим більше вся система $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_{n+1}, \ldots, \alpha_m$ не може бути лінійно незалежною.

Звідси видно, що *степінь* n *скінченного розширення* Ω *над полем* Δ *дорівнює максимальному числу* l *елементів поля* Ω, *які можуть утворювати лінійно незалежну систему.* Справді, за доведеним, $n \geqslant l$; з другого боку, в Ω завжди є лінійно незалежна система з n елементів (довільний з базисів Ω). Отже, $n = l$.

Справедливе й обернене твердження, а саме:

Якщо l *— максимальне число елементів розширення* Ω *поля* Δ, *що утворюють лінійно незалежну систему відносно поля* Δ, *то* Ω *є скінченним розширенням над полем* Δ *степеня* l.

Досить показати, що в Ω існує базис відносно поля Δ, який містить l елементів. Нехай $\beta_1, \beta_2, \ldots, \beta_l$ — лінійно незалежна система елементів поля Ω (така система існує за самим означенням числа l). Покажемо, що ця система є базисом поля Ω відносно Δ. Якщо ζ — довільний елемент поля Ω, відмінний від $\beta_1, \beta_2, \ldots, \beta_l$, то система елементів $\zeta, \beta_1, \beta_2, \ldots, \beta_l$ — лінійно залежна, бо містить $l+1$ елемент. Це означає, що $\lambda_0\zeta + \lambda_1\beta_1 + \ldots + \lambda_l\beta_l = 0$ $(\lambda_i \in \Delta)$, коли не всі λ_i дорівнюють нулю. При цьому $\lambda_0 \neq 0$, бо при $\lambda_0 = 0$ з рівності $\lambda_1\beta_1 + \lambda_2\beta_2 + \ldots + \lambda_l\beta_l = 0$ і лінійної незалежності $\beta_1, \beta_2, \ldots, \beta_l$ випливало б, що й $\lambda_1 = \lambda_2 = \ldots = \lambda_l = 0$. Отже, можна подати ζ через елементи $\beta_1, \beta_2, \ldots, \beta_l$:

$$\zeta = \left(-\frac{\lambda_1}{\lambda_0}\right)\beta_1 + \cdots + \left(-\frac{\lambda_l}{\lambda_0}\right)\beta_l.$$

Звідси зрозуміло, що кожний елемент $\zeta \in \Omega$ є лінійною комбінацією елементів $\beta_1, \beta_2, \ldots, \beta_l$, які, отже, утворюють базис поля Ω.

П р и к л а д и. **3.** Поле $\mathbf{Q}(\sqrt{2})$ чисел виду $a + b\sqrt{2}$, де a, b — довільні раціональні числа, є розширенням поля $\mathbf{Q}$ степеня 2, бо існує базис поля $\mathbf{Q}(\sqrt{2})$ відносно поля $\mathbf{Q}$, який складається з двох елементів. За базис можна взяти числа 1 і $\sqrt{2}$. Сукупність цих чисел, як було показано, є лінійно незалежною системою відносно поля $\mathbf{Q}$. За базис можна взяти й інші числа, наприклад $1 - \sqrt{2}$ і $1 + \sqrt{2}$ або взагалі два числа $a_1 + b_1\sqrt{2}$ і $a_2 + b_2\sqrt{2}$, аби тільки система цих чисел була лінійно незалежною відносно $\mathbf{Q}$, тобто справджувалась умова $\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$.

4. Поле $\mathbf{Q}(\sqrt[3]{2})$ є скінченним розширенням степеня 3 поля $\mathbf{Q}$. Базисом цього поля відносно поля $\mathbf{Q}$ є, наприклад, числа $1, \sqrt[3]{2}, \sqrt[3]{2^2}$, або $1 + \sqrt[3]{2}, 2 - \sqrt[3]{2}, 1 + \sqrt[3]{2^2}$

5. Поле $\mathbf{C}$ комплексних чисел відносно поля $\mathbf{R}$ дійсних чисел є скінченним розширенням степеня 2. Базисом поля $\mathbf{C}$ відносно поля $\mathbf{R}$ є, наприклад, система чисел 1 і i.

Зауважимо, що не кожне розширення поля є скінченним. Так, поле $\mathbf{R}$ дійсних чисел є розширенням поля $\mathbf{Q}$ раціональних чисел. Проте це розширення не є скінченним, бо в ньому не існує скінченного базису, через який лінійно виражалося б будь-яке дійсне число

Розширення Ω першого степеня над полем Δ просто збігається з полем Δ. Справді, довільний елемент $\zeta \in \Omega$ в цьому випадку, згідно з (7), має вигляд $\zeta = \lambda\beta$, де β — деякий фіксований елемент поля Ω, а λ — будь-яке число з поля Δ. Якщо взяти, наприклад, $\zeta = 1$, то знайдеться $\lambda_1 \in \Delta$ таке, що $1 = \lambda_1\beta$, звідки $\beta = \frac{1}{\lambda_1}$. Як бачимо, β_1, а разом з ним і кожне ζ належать Δ, тобто $\Omega \subseteq \Delta$. Оскільки ж Ω — розширення Δ, то дістанемо $\Omega = \Delta$.

§ 33. АЛГЕБРАЇЧНІ РОЗШИРЕННЯ ЧИСЛОВИХ ПОЛІВ

33.1. Поняття алгебраїчного розширення. Для дальшого викладу нам потрібно ввести деякі нові типи розширень числових полів і з'ясувати зв'язок їх між собою та з скінченними розширеннями.

Розширення Ω поля Δ, утворене за допомогою кількох послідовно виконаних простих алгебраїчних розширень, називається *складним алгебраїчним розширенням поля* Δ.

Означення 1. *Ω є складним алгебраїчним розширенням поля Δ, якщо існує такий ланцюжок розширень*

$$\Delta \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \ \ldots \ \subseteq \Delta_k = \Omega,$$

що $\Delta_1 = \Delta(\alpha_1)$, $\Delta_2 = \Delta_1(\alpha_2)$, ..., $\Delta_k = \Delta_{k-1}(\alpha_k)$, причому кожне α_i є алгебраїчним числом над полем Δ_{i-1} (при $i = 1$ $\Delta_{i-1} = \Delta$).

Це означення не вимагає, щоб всі числа α_i були алгебраїчними відносно початкового поля Δ, бо алгебраїчність числа α_i відносно поля Δ_{i-1} безпосередньо ще не означає його алгебраїчності відносно його підполя Δ.

Складне алгебраїчне розширення поля Δ позначатимемо символом $\Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$. Це позначення не слід змішувати з символом $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$, який означає розширення поля Δ, утворене одночасним приєднанням до нього чисел $\alpha_1, \alpha_2, \ldots, \alpha_k$ (п. 32.2).

Природно розглядати складне алгебраїчне розширення як узагальнення простого, тобто вважати просте розширення $\Delta(\alpha_1)$ також складним з $k = 1$, або, що те саме, при $\alpha_2 \in \Delta(\alpha_1)$, $\alpha_3 \in \Delta(\alpha_1)$, ..., $\alpha_k \in \Delta(\alpha_1)$.

Крім поняття *простого* чи *складного алгебраїчного розширення* введемо ще поняття *алгебраїчного розширення числового поля*.

Означення 2. *Розширення Ω поля Δ називається алгебраїчним, якщо всі його елементи є алгебраїчними відносно поля Δ.*

Як буде показано далі, усі раніше введені типи розширень (просте і складне алгебраїчне, скінченне) належать до категорії алгебраїчних розширень. Звичайно, не кожне розширення поля є алгебраїчним: так, можна показати, що поле дійсних чисел $\mathbf{R}$ не є алгебраїчним розширенням поля $\mathbf{Q}$ раціональних чисел. Розширення, які не є алгебраїчними, називають трансцендентними. Тут ми обмежимося вивченням лише алгебраїчних розширень. Зауважимо лише, що, як можна показати, просте трансцендентне розширення поля P (тобто мінімальне поле, яке містить P і трансцендентний відносно

P елемент x) збігається з полем $P(x)$ відношень многочленів (раціональних дробів) над P (§ 24).

***33.2. Скінченність простих і складних алгебраїчних розширень.** Перейдемо до з'ясування співвідношень між різними типами розширень числових полів.

Насамперед, теорема 1, доведена в п. 32.2, означає справедливість такого твердження.

Теорема 1. *Просте алгебраїчне розширення $\Delta(\alpha)$, утворене з Δ приєднанням алгебраїчного відносно Δ числа α, є скінченним розширенням поля Δ. Степінь розширення $\Delta(\alpha)$ над полем Δ дорівнює степеню числа α відносно Δ.*

Справді, в теоремі 1 § 32 було доведено, що довільне число $\zeta \in \Delta(\alpha)$ можна подати у вигляді

$$\zeta = c_0 + c_1\alpha + \ \cdots \ + c_{n-1}\alpha^{n-1}.$$

Крім того, в п. 32.4 було доведено, що елементи $1, \alpha, \ldots, \alpha^{n-1}$ є базисом поля $\Delta(\alpha)$ відносно поля Δ. Отже, $\Delta(\alpha)$ є скінченним розширенням поля Δ степеня n. Теорему доведено.

Наслідок. *Степінь будь-якого квадратичного розширення числового поля дорівнює 2.*

Покажемо тепер, що *складне алгебраїчне розширення поля Δ є також скінченним розширенням цього поля.* Доведемо для цього спочатку такі твердження.

Лема 1. *Якщо Δ_2 — скінченне розширення поля Δ_1, а Δ_1 — скінченне розширення поля Δ, то Δ_2 — скінченне розширення поля Δ, причому $(\Delta_2 : \Delta) = (\Delta_2 : \Delta_1)(\Delta_1 : \Delta)$.*

Доведення. Нехай $(\Delta_2 : \Delta_1) = n$, $(\Delta_1 : \Delta) = m$. Кожний елемент $\omega \in \Delta_2$ можна подати у формі

$$\omega = \sum_{i=1}^{n} c_i\xi_i, \tag{1}$$

де сукупність елементів $\xi_1, \xi_2, \ldots \xi_n$ є базисом поля Δ_2 відносно поля Δ_1, а елементи $c_1, c_2, \ldots, c_n$ належать Δ_1.

З другого боку, через те що Δ_1 є розширенням поля Δ степеня m, маємо:

$$c_i = \sum_{k=1}^{m} a_{ik}\eta_k \qquad (i = 1, 2, \ldots, n), \tag{2}$$

де $\eta_1, \eta_2, \ldots \eta_m$ — деякий базис Δ_1 відносно Δ, а всі a_{ik} належать Δ. Підставляючи (2) в (1), дістанемо:

$$\omega = \sum_{i=1}^{n}\sum_{k=1}^{m} a_{ik}\eta_k\xi_i. \tag{3}$$

Рівність (3) показує, що кожне число $\omega \in \Delta_2$ є лінійною комбінацією елементів $\zeta_{ik} = \eta_k\xi_i$, які належать Δ_2, з коефіцієнтами a_{ik} з поля Δ.

З цього випливає, що Δ_2 є скінченне розширення поля Δ. Покажемо тепер, що степінь цього розширення дорівнює mn. Для цього досить

встановити, що mn елементів ζ_{ik} утворюють лінійно незалежну систему відносно поля Δ.

Нехай λ_{ik} — деякі числа з поля Δ такі, що справджується рівність:

$$\sum_{i=1}^{n}\sum_{k=1}^{m}\lambda_{ik}\zeta_{ik}=0. \tag{4}$$

Покажемо тепер, що ця рівність справджується лише при умові, що всі $\lambda_{ik}=0$. Виконуючи елементарні перетворення, маємо:

$$\sum_{i=1}^{n}\sum_{k=1}^{m}\lambda_{ik}\zeta_{ik}=\sum_{i=1}^{n}\sum_{k=1}^{m}\lambda_{ik}\eta_k\xi_i=\sum_{i=1}^{n}\left(\sum_{k=1}^{m}\lambda_{ik}\eta_k\right)\xi_i=\sum_{i=1}^{n}\mu_i\xi_i,$$

де

$$\mu_i=\sum_{k=1}^{m}\lambda_{ik}\eta_k,\quad \mu_i\in\Delta_1. \tag{5}$$

Рівність (4) можна переписати тепер так:

$$\sum_{i=1}^{n}\mu_i\xi_i=0.$$

Але $\xi_1, \xi_2, \ldots, \xi_n$ є базис поля Δ_2 відносно поля Δ_1, отже, ці елементи утворюють лінійно незалежну систему відносно Δ_1. Тому $\mu_1=\mu_2=\ldots=\mu_n=0$. Але тоді з рівностей (5) випливає:

$$\sum_{k=1}^{n}\lambda_{ik}\eta_k=0\qquad (i=1, 2, \ldots, n),$$

де $\eta_1, \eta_2, \ldots, \eta_m$ — базис поля Δ_1 відносно поля Δ і, отже, всі $\lambda_{ik}=0$.

Тим самим доведено, що система елементів $\zeta_{ik}=\eta_k\xi_i$ є базисом поля Δ_2 відносно поля Δ, тобто, що поле Δ_2 є розширенням степеня mn поля Δ. Іншими словами, $(\Delta_2:\Delta)=(\Delta_2:\Delta_1)(\Delta_1:\Delta)$.

Справедливе також твердження, в певному розумінні обернене до леми 1.

Лема 2. *Якщо Δ_2 — скінченне розширення поля Δ степеня m, а Δ_1 — довільне розширення поля Δ, що міститься в Δ_2:*

$$\Delta\subseteq\Delta_1\subseteq\Delta_2,$$

то Δ_1 — також скінченне розширення Δ, причому його степінь $(\Delta_1:\Delta)$ є дільником числа m.

Доведення. Кожний елемент β, який належить Δ_1, належить також і Δ_2. Степінь розширення Δ_2 відносно поля Δ дорівнює m. Тому довільна лінійно незалежна система елементів з Δ_2 відносно поля Δ містить не більш як m елементів (п. 32. 4). Це стосується і елементів поля Δ_1, яке за умовою є підполем поля Δ_2: ніяка лінійно незалежна відносно Δ система елементів з Δ_1 не може містити більше як m елементів. Звідси випливає, що існує базис Δ_1 відносно поля Δ, який складається з q елементів поля Δ_1, причому $q\leqslant m$. Отже, Δ_1 є скінченним розширенням Δ. Очевидно також, що поле Δ_2 є скінченним розширенням поля Δ_1. Справді, оскільки Δ_2 є скінченним розширенням поля Δ, то довільний елемент $\beta\in\Delta_2$ лінійно виражається через базис $\alpha_1, \alpha_2, \ldots, \alpha_m$ відносно поля Δ з коефіцієнтами з Δ.

Оскільки ж $\Delta\subseteq\Delta_1$, то виходить, що β лінійно виражається[1] через елементи $\alpha_1, \alpha_2, \ldots, \alpha_m$ з коефіцієнтами з поля Δ_1.

У цьому випадку справджуються всі умови леми 1: Δ_2 — скінченне розширення Δ_1, Δ_1 — скінченне розширення Δ. Отже, згідно з лемою 1, $(\Delta_2:\Delta)=(\Delta_2:\Delta_1)(\Delta_1:\Delta)$,
тобто $m=(\Delta_2:\Delta_1)(\Delta_1:\Delta)$ і $(\Delta_1:\Delta)$ є дільником числа m. Лему доведено.

З теореми 1 і леми 1 тепер безпосередньо випливає така теорема:

Теорема 2. *Складне алгебраїчне розширення $\Delta(\alpha_1)(\alpha_2)$ є скінченним розширенням поля Δ. Степінь цього розширення дорівнює добутку степеня розширення $\Delta(\alpha_1)$ відносно поля Δ на степінь розширення $\Delta(\alpha_1)(\alpha_2)$ відносно поля $\Delta(\alpha_1)$.*

Доведення. Позначимо $\Delta_1=\Delta(\alpha_1)$, $\Delta_2=\Delta_1(\alpha_2)$. Очевидно, маємо $\Delta\subseteq\Delta_1\subseteq\Delta_2$. Кожне з розширень Δ_1 і Δ_2 є простим алгебраїчним розширенням. За теоремою 1, ці розширення скінченні. Припустимо, що $(\Delta_2:\Delta_1)=n$, $(\Delta_1:\Delta)=m$. Використовуючи тепер лему 1, маємо, що Δ_2 є скінченним розширенням поля Δ, причому $(\Delta_2:\Delta)=mn$, що й треба було довести.

Наслідок 1. *Складне алгебраїчне розширення $\Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$ є скінченним розширенням поля Δ. Степінь цього розширення дорівнює добутку степенів усіх послідовних простих розширень.*

Для доведення використаємо метод математичної індукції. При $k=2$ наше твердження правильне, бо збігається з теоремою 2. Припустимо тепер, що твердження правильне при $k=i-1$, і доведемо його справедливість при $k=i$. Введемо позначення: $\Delta_{i-1}=\Delta(\alpha_1)(\alpha_2)\ldots(\alpha_{i-1})$; $\Delta_i=\Delta_{i-1}(\alpha_i)$. З припущення індукції випливає, що Δ_{i-1} — скінченне розширення поля Δ і його степінь дорівнює $\mu_{i-1}=(\Delta_1:\Delta)(\Delta_2:\Delta_1)\ldots(\Delta_{i-1}:\Delta_{i-2})$. Тоді, за теоремою 2, Δ_i є також скінченним розширенням поля Δ, причому його степінь дорівнює

$$\mu_i=\mu_{i-1}(\Delta_i:\Delta_{i-1})=(\Delta_1:\Delta)(\Delta_2:\Delta_1)\ldots(\Delta_{i-1}:\Delta_{i-2})(\Delta_i:\Delta_{i-1}),$$

що й треба було довести.

Оскільки у випадку ланцюжка квадратичних розширень степінь кожного послідовного простого розширення дорівнює 2, приходимо до такого наслідку:

Наслідок 2. *Якщо $\Delta\subseteq\Delta_1\subseteq\ldots\subseteq\Delta_k$ є ланцюжком квадратичних розширень, то Δ_k є скінченним розширенням поля Δ степеня 2^k.*

33. 3. Алгебраїчність і простота скінченних розширень. Ми показали, що як прості, так і складні алгебраїчні розширення поля є завжди скінченними розширеннями. Для доведення того, що всі згадані розширення *алгебраїчні*, досить встановити алгебраїчність довільного скінченного розширення поля.

Теорема 3. *Будь-яке скінченне розширення поля є його алгебраїчним розширенням.*

---

[1] Зауважимо, що система $\alpha_1, \alpha_2, \ldots, \alpha_m$ може не бути базисом відносно Δ_1, бо лінійна незалежність відносно поля Δ не означає лінійної незалежності відносно поля Δ_1. Але такий базис можна дістати, послідовно відкинувши з елементів $\alpha_1, \alpha_2, \ldots, \alpha_m$ ті, які є лінійними комбінаціями решти цих елементів з коефіцієнтами з поля Δ_1.

Доведення. Нехай Δ_1 — скінченне розширення n-го степеня поля Δ. Отже, будь-які $n+1$ елементів у полі Δ_1 є лінійно залежними відносно Δ (п. 32.4). Тому якщо взяти в Δ_1 довільний елемент α і за його допомогою скласти $n+1$ елементів: $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}, \alpha^n$, то ця система елементів є лінійно залежною відносно Δ, тобто існують такі $\lambda_i \in \Delta$, з яких не всі дорівнюють нулю, що справджується рівність

$$\lambda_0 + \lambda_1\alpha + \lambda_2\alpha^2 + \cdots + \lambda_{n-1}\alpha^{n-1} + \lambda_n\alpha^n = 0.$$

Отже, α є коренем деякого многочлена $f(x) = \lambda_n x^n + \ldots + \lambda_1 x + \lambda_0$ з коефіцієнтами з поля Δ, тобто є алгебраїчним числом відносно поля Δ. Теорему доведено.

З доведення безпосередньо видно, що *кожне число скінченного розширення n-го степеня поля Δ є алгебраїчним числом, степінь якого відносно Δ не вищий за n.*

Зокрема, оскільки будь-яке квадратичне розширення є скінченним розширенням другого степеня, то *кожний елемент γ з квадратичного розширення поля Δ є коренем деякого многочлена над полем Δ, степінь якого не перевищує* 2. Якщо $\gamma \notin \Delta$, то напевно γ є коренем незвідного квадратного тричлена над полем Δ.

Теореми 1—3 дають змогу прийти до такого висновку:

Кожне просте або складне алгебраїчне розширення числового поля Δ є алгебраїчним розширенням цього поля.

Справедливим є твердження, що *скінченне розширення поля Δ є не тільки алгебраїчним, але й простим алгебраїчним розширенням Δ.* Це твердження ми дістанемо як наслідок трьох наступних теорем.

Теорема 4. *Розширення $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$, утворене з Δ приєднанням алгебраїчних відносно Δ чисел $\alpha_1, \alpha_2, \ldots, \alpha_k$, збігається з складним алгебраїчним розширенням $\Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$.*

Доведення. Якщо побудувати послідовність розширень

$$\Delta \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \ldots \subseteq \Delta_{k-1} \subseteq \Delta_k,$$

де $\Delta_1 = \Delta(\alpha_1)$, $\Delta_2 = \Delta_1(\alpha_1)$, ..., $\Delta_i = \Delta_{i-1}(\alpha_i)$, ..., $\Delta_k = \Delta_{k-1}(\alpha_k)$, то всі ці розширення є простими алгебраїчними розширеннями. Справді, кожне α_i за умовою є алгебраїчним числом відносно поля Δ, а отже, є алгебраїчним і відносно розширення Δ_{i-1} цього поля. Тому $\Delta_i = \Delta_{i-1}(\alpha_i)$ є простим алгебраїчним розширенням. Зауважимо, що Δ_i може збігатися з Δ_{i-1}, якщо $\alpha_i \in \Delta_{i-1}$. У цьому випадку $\Delta(\alpha_1) \times (\alpha_2)\ldots(\alpha_k) = \Delta(\alpha_1)\ldots(\alpha_{i-1})(\alpha_{i+1})\ldots(\alpha_k)$.

Розглянемо поле $\Delta_k = \Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$. На основі теореми 1 п. 32.2, всі елементи цього поля є якісь многочлени над полем Δ від елементів $\alpha_1, \alpha_2, \ldots, \alpha_k$. Тим самим це поле має бути підполем поля $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$, тобто $\Delta_k \subseteq \Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$. З другого боку, поле $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$ мінімальне з тих, що містять $\Delta, \alpha_1, \ldots, \alpha_k$ і, отже, $\Delta_k \supseteq \Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$. Таким чином, розширення $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$ збігається з складним розширенням $\Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$.

Теорему доведено.

Теорема 4 показує, що розширення $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$ скінченне і алгебраїчне, бо таким є складне алгебраїчне розширення $\Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$. Розширення $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_k)$ іноді називають алгебраїчно породженим. Таким чином, клас алгебраїчно породжених розширень збігається з класом складних алгебраїчних розширень.

Теорема 5. *Будь-яке скінченне розширення поля Δ є складним алгебраїчним розширенням цього поля.*

Доведення. Нехай Δ_1 — скінченне розширення n-го степеня поля Δ. У теоремі 3 було показано, що Δ_1 є алгебраїчне розширення поля Δ. Отже, якщо $\beta_1, \beta_2, \ldots, \beta_n$ — який-небудь базис поля Δ_1, то всі числа $\beta_1, \beta_2, \ldots, \beta_n$ алгебраїчні відносно поля Δ. Покажемо тепер, що $\Delta_1 = \Delta(\beta_1, \beta_2, \ldots, \beta_n)$. Оскільки Δ_1 є поле і містить $\beta_1, \beta_2, \ldots, \beta_n$, а $\Delta(\beta_1, \ldots, \beta_n)$ — мінімальне поле з такими елементами, то

$$\Delta(\beta_1, \beta_2, \ldots, \beta_n) \subseteq \Delta_1. \tag{6}$$

За означенням поля, $\Delta(\beta_1, \beta_2, \ldots, \beta_n)$ містить усі лінійні комбінації елементів $\beta_1, \beta_2, \ldots, \beta_n$ з коефіцієнтами з поля Δ, тобто містить усе поле Δ_1:

$$\Delta(\beta_1, \beta_2, \ldots, \beta_n) \supseteq \Delta_1. \tag{7}$$

З (6) і (7) випливає, що $\Delta_1 = \Delta(\beta_1, \beta_2, \ldots, \beta_n) = \Delta(\beta_1)(\beta_2)\ldots(\beta_n)$. Теорему доведено.

Звичайно, складне алгебраїчне розширення може виявитись, зокрема, простим, якщо $\beta_2 \in \Delta(\beta_1)$, $\beta_3 \in \Delta(\beta_1)$, ..., $\beta_n \in \Delta(\beta_1)$.

Теорема 6. *Будь-яке складне алгебраїчне розширення $\Omega \overset{\text{df}}{=} \Delta(\alpha_1)(\alpha_2)\ldots(\alpha_k)$ поля Δ є простим розширенням цього поля, тобто існує таке число ω, алгебраїчне відносно Δ, що $\Omega = \Delta(\omega)$.*

Доведення. Застосуємо метод математичної індукції і розглянемо спочатку випадок $k = 2$. У цьому випадку $\Omega = \Delta(\alpha_1)(\alpha_2)$. Треба знайти таке число ω, алгебраїчне відносно Δ, що $\Omega = \Delta(\omega)$.

Оскільки α_1 і α_2 алгебраїчні відносно Δ, то можна розглянути їхні мінімальні многочлени $f(x)$ та $g(x)$ над полем Δ. Нехай $\xi_1, \xi_2, \ldots, \xi_n$ — корені $f(x)$, $\eta_1, \eta_2, \ldots, \eta_m$ — корені $g(x)$. Очевидно, всі числа ξ_i попарно різні (якби $f(x)$ мав кратні корені, він був би звідним і не міг би бути мінімальним), причому одне з ξ_i дорівнює α_i (наприклад, $\xi_1 = \alpha_1$); аналогічно, η_j попарно різні числа і $\eta_1 = \alpha_2$.

Утворимо тепер число $\omega = \alpha_1 + c\alpha_2$, де c — деяке число з поля Δ, яке ми виберемо так, щоб при всіх $i = 1, 2, \ldots, n$ і $j = 2, 3, \ldots, m$ було

$$\omega \neq \xi_i + c\eta_j. \tag{8}$$

Такий вибір завжди можна зробити, бо числове поле Δ має безліч елементів, а чисел c, для яких

$$\alpha_1 + c\alpha_2 = \xi_i + c\eta_j \Leftrightarrow c = \frac{\alpha_1 - \xi_i}{\eta_j - \alpha_2} \qquad (i = 1, 2, \ldots, n;\ j = 2, \ldots, m),$$

скінченна кількість.

Доведемо, що означене так число ω — шукане. Перш за все, $\omega \in \Omega$ і тому алгебраїчне відносно Δ. Крім того, зрозуміло, що $\Delta(\omega) \subseteq \Omega$. Залишається довести, що $\Delta(\omega)$ містить кожне число з Ω: $\Omega \subseteq \Delta(\omega)$.

Розглянемо многочлен $f_1(x) = f(\omega - cx)$. Його коефіцієнти належать полю $\Delta(\omega)$, бо коефіцієнти многочлена $f(x)$ і c належать Δ, а

$\omega \in \Delta(\omega)$. Серед коренів $f_1(x)$ є число α_2, бо $f_1(\alpha_2) = f(\omega - c\alpha_2) = f(\alpha_1) = 0$. Отже, $f_1(x)$ та $g(x)$ мають спільний корінь α_2. Але ці многочлени не мають жодного іншого спільного кореня, бо

$$[f_1(\eta_j) = 0 \ (j = 2, \ldots, m)] \Leftrightarrow [f(\omega - c\eta_j) = 0 \ (j = 2, \ldots, m)] \Leftrightarrow$$
$$\Leftrightarrow [\omega - c\eta_j = \xi_i \quad (i = 1, \ldots, n; \quad j = 2, \ldots, m)],$$

що суперечить (8).

Отже, $f_1(x)$ і $g(x)$ мають єдиний спільний корінь α_2 і тому їх НСД $(f_1, g) = x - \alpha_2$. Оскільки $f_1(x)$ та $g(x)$ — многочлени над полем $\Delta(\omega)$, яке є розширенням поля Δ, то і $x - \alpha_2$ — многочлен над $\Delta(\omega)$, тобто $\alpha_2 \in \Delta(\omega)$. Тепер зрозуміло, що й $\alpha_1 = \omega - c\alpha_2 \in \Delta(\omega)$. Але тоді мінімальне поле $\Delta(\alpha_1, \alpha_2)$, яке містить поле Δ і числа α_1, α_2, мусить бути підполем $\Delta(\omega)$. За теоремою 4 $\Delta(\alpha_1, \alpha_2) = \Delta(\alpha_1)(\alpha_2) = \Omega$; отже, ми показали, що $\Omega \subseteq \Delta(\omega)$, чим доведення теореми при $k = 2$ закінчується.

Припустимо тепер, що теорема правильна при $k = l \geqslant 2$, і доведемо правильність її при $k = l + 1$. Нехай $\Omega = \Delta(\alpha_1)(\alpha_2) \ldots (\alpha_l)(\alpha_{l+1})$, де α_i $(i = 1, 2, \ldots, l+1)$ — алгебраїчні відносно Δ числа. Розглянемо поле $\Omega_1 = \Delta(\alpha_1)(\alpha_2) \ldots (\alpha_l)$. За припущенням індукції $\Omega_1 = \Delta(\omega_1)$, де ω_1 — якесь алгебраїчне відносно Δ число. Але тоді $\Omega = \Delta(\omega_1)(\alpha_{l+1})$ і, за доведеним вище (при $k = 2$) твердженням теореми, Ω є просте алгебраїчне розширення поля Δ. Теорему доведено.

Наслідок. *Будь-яке розширення другого степеня поля Δ є квадратичним розширенням поля* Δ.

Отже, було розглянуто такі п'ять класів розширень числового поля Δ:

K_1 : прості алгебраїчні розширення;
K_2 : складні алгебраїчні розширення;
K_3 : скінченні розширення;
K_4 : алгебраїчно породжені розширення;
K_5 : алгебраїчні розширення.

Встановлені щодо цих класів факти можна символічно подати так: *Теорема* 1: $K_1 \subseteq K_3$; *теорема* 2 : $K_2 \subseteq K_3$; *теорема* 3 : $K_3 \subseteq K_5$; *теорема* 4 : $K_2 = K_4$; *теорема* 5 : $K_3 \subseteq K_2$; *теорема* 6 : $K_2 \subseteq K_1$.

Зрозуміло, що $K_2 = K_3$ (теореми 2 і 5), $K_1 = K_2$ (теорема 6 і очевидне співвідношення $K_1 \subseteq K_2$); $K_2 = K_4$ (теорема 4), тобто $K_1 = K_2 = K_3 = K_4 \underset{\text{df}}{=} K_0$. Іншими словами, поняття *простого і складного алгебраїчних розширень, алгебраїчно породженого розширення і скінченного розширення по суті збігаються*. Різні за способом побудови, всі вони являють собою ту саму алгебраїчну структуру.

Що ж до класу K_5 алгебраїчних розширень, то поки що нам лише відомо, що $K_0 \subseteq K_5$ (теорема 3). У наступному пункті буде показано, що $K_0 \subset K_5$, тобто існують алгебраїчні розширення, які не є скінченими (а тому й простими, або складними, або алгебраїчно породженими).

33.4. Поле алгебраїчних чисел. Розглянемо сукупність $A(\Delta)$ усіх алгебраїчних чисел відносно поля Δ, тобто множину коренів усіх многочленів з $\Delta[x]$.

Теорема 7. *Сукупність $A(\Delta)$ алгебраїчних чисел відносно поля Δ є поле.*

Доведення. Досить показати, що разом з будь-якими двома числами ξ, η сукупність $A(\Delta)$ містить $\xi + \eta$, $\xi \cdot \eta$, $-\xi$ та $\frac{1}{\xi}$ і що $0 \in A(\Delta)$, $1 \in A(\Delta)$. Останній факт випливає з того, що (п. 32.1).

$$\mathbf{Q} \subseteq \Delta \subseteq A(\Delta). \tag{9}$$

Нехай тепер $\xi \in A(\Delta)$, $\eta \in A(\Delta)$. Розглянемо алгебраїчно породжене розширення $\Delta(\xi, \eta)$. Це розширення алгебраїчне (див. п. 33.3), а тому кожне число з нього належить $A(\Delta)$. Зокрема, числа $\xi + \eta$, $\xi \cdot \eta$, $-\xi$, $\frac{1}{\xi}$, які належать $\Delta(\xi, \eta)$, належать і $A(\Delta)$. Теорему доведено.

Наслідок. *Множина A усіх алгебраїчних чисел* (тобто $A = A(\mathbf{Q})$) *є поле.*

З (9) зрозуміло, що $A(\Delta)$ є алгебраїчне розширення поля Δ. Покажемо тепер, що *в загальному випадку це розширення не є скінченним.* Для цього досить перевірити, що поле A не є скінченним розширенням поля $\mathbf{Q}$ раціональних чисел. Але це насправді так, бо у кільці $\mathbf{Q}[x]$ існують незвідні многочлени як завгодно великого степеня (п. 31.1). Отже, в A існують числа як завгодно великого степеня відносно поля $\mathbf{Q}$. Але тоді A не може бути скінченним розширенням поля $\mathbf{Q}$, бо, як випливає з доведення теореми 3, степінь скінченного розширення поля Δ не нижчий від степеня кожного елемента цього розширення відносно Δ.

З наведених міркувань випливає, що *існують алгебраїчні розширення полів, які не є скінченними,* тобто $K_5 \supset K_0$, але $K_5 \neq K_0$, як і було зазначено у попередньому пункті. Доведений факт не означає, що будь-яке розширення $A(\Delta)$ є нескінченним. Так, поле $\mathbf{C}$ комплексних чисел є, як ми знаємо, скінченним розширенням поля $\mathbf{R}$ дійсних чисел. В той же час $\mathbf{C} = A(R)$, бо корені усіх многочленів з дійсними коефіцієнтами належать $\mathbf{C}$, і будь-яке комплексне число $a + bi$ є коренем многочлена над $\mathbf{R}$ (наприклад, многочлена $f(x) = x^2 - 2ax + (a^2 + b^2)$).

Як відомо (§ 28), поле $\mathbf{C}$ алгебраїчно замкнуте, тобто всі алгебраїчні числа над цим полем належать самому полю $\mathbf{C}$: $A(\mathbf{C}) = \mathbf{C}$. Виявляється, що цю властивість має сукупність $A(\Delta)$ алгебраїчних чисел над довільним числовим полем Δ : $A[A(\Delta)] = A(\Delta)$.

Теорема 8. *Поле алгебраїчних чисел $A(\Delta)$ над довільним числовим полем Δ алгебраїчно замкнуте.*

Доведення. Нехай $A[A(\Delta)] = \Omega$. Нам потрібно показати, що $\Omega = A(\Delta)$. Оскільки включення $A(\Delta) \subseteq \Omega$ очевидне, досить показати, що будь-який елемент $\omega \in \Omega$ є елементом $A(\Delta)$, тобто алгебраїчний відносно Δ. Якщо

$$g(x) = x^n + \alpha_{n-1}x^{n-1} + \ldots + \alpha_1 x + \alpha_0 \ (\alpha_j \in A(\Delta),$$
$$j = 0, 1, \ldots, n-1)$$

є мінімальний многочлен над A (Δ) числа ω (такий многочлен існує, бо ω — алгебраїчне число відносно A (Δ)), то розглянемо алгебраїчно ироджене розширення $\Delta(\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$ поля Δ. За раніше доведеним, $\Delta(\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$ — алгебраїчне розширення Δ, тобто будь-який його елемент, зокрема ω, алгебраїчний над Δ. Звідси $\omega \in A$ (Δ), що й треба було довести.

Основна теорема теорії многочленів (§ 28) стверджує алгебраїчну замкненість поля $\mathbf{C}$ усіх комплексних чисел. Поле $A \underset{\text{df}}{=} A(\mathbf{Q})$ алгебраїчних чисел є підполем поля $\mathbf{C}$: $A \subset \mathbf{C}$ (причому правильним підполем, бо в $\mathbf{C}$ існують трансцендентні числа). З теореми 8 випливає, що і підполе A поля $\mathbf{C}$ має властивість алгебраїчної замкненості.

§ 34. РОЗВ'ЯЗНІСТЬ АЛГЕБРАЇЧНИХ РІВНЯНЬ У КВАДРАТНИХ РАДИКАЛАХ

34.1. Поняття розв'язності у квадратних радикалах. Як відомо (1, § 17), не кожне алгебраїчне рівняння можна розв'язати у радикалах, тобто виразити всі його корені через коефіцієнти за допомогою скінченного числа дій додавання, віднімання, множення, ділення і добування кореня з цілим показником степеня. Зокрема, славнозвісна теорема Руффіні — Абеля твердить, що рівняння n-го степеня з довільними буквеними коефіцієнтами при $n \geqslant 5$ не можна розв'язати в радикалах.

Разом з тим існують окремі класи рівнянь вищих степенів, які можна розв'язати у радикалах. Загальне дослідження проблеми розв'язності алгебраїчних рівнянь у радикалах є предметом важливої галузі загальної алгебри — так званої теорії Галуа[1]. Виклад цієї теорії виходить за межі цього курсу. Проте одне з питань теорії розв'язності рівнянь у радикалах ми розглянемо тут (в елементарній формі), ураховуючи його особливе значення для традиційного курсу шкільної математики. Йдеться про *розв'язність алгебраїчних рівнянь у квадратних радикалах*. Саме до цього питання зводиться дослідження можливості чи неможливості розв'язати певну геометричну *задачу на побудову за допомогою циркуля і лінійки*.

Означення 1. *Вважатимемо, що алгебраїчне рівняння*

$$f(x) \underset{\text{df}}{=} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \quad (a_n \neq 0) \tag{1}$$

можна розв'язати у квадратних радикалах, якщо кожний з його n коренів можна подати через коефіцієнти a_j ($j = 0, 1, \ldots, n$) за допомогою скінченного числа дій додавання, віднімання, множення, ділення та добування квадратного кореня.

Якщо дії додавання, віднімання, множення і ділення назвати (як це загальноприйнято) р а ц і о н а л ь н и м и операціями, то формулювання означення 1 можна дещо скоротити, а саме: *рівняння* (1) *розв'язне у квадратних радикалах, якщо кожний його корінь можна подати через коефіцієнти a_j ($j = 0, 1, \ldots, n$) за допомогою скінченного числа раціональних операцій та дій добування квадратного кореня.*

Очевидно, будь-яке лінійне рівняння

$$ax + b = 0 \tag{2}$$

та будь-яке квадратне рівняння

$$ax^2 + bx + c = 0 \tag{3}$$

розв'язуються у квадратних радикалах. Справді, корені цих рівнянь можна подати через коефіцієнти (при умові $a \neq 0$) формулами

$$\alpha = -\frac{b}{a}, \quad \alpha_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad \alpha_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

(де α — корінь рівняння (2); α_1, α_2 — корені рівняння (3)), тобто за допомогою скінченного числа дій додавання, віднімання, множення, ділення та добування квадратного кореня.

Існують рівняння, степінь яких перевищує 2 і які розв'язуються у квадратних радикалах.

П р и к л а д 1. Розглянемо рівняння

$$x^5 - 1 = 0. \tag{4}$$

Коренями його, як відомо (1, п. 17.8), є числа

$$\alpha_1 = 1, \quad \alpha_2 = \frac{-1-\sqrt{5}}{4} + \frac{\sqrt{2\sqrt{5}-10}}{4}, \quad \alpha_3 = \frac{-1-\sqrt{5}}{4} - \frac{\sqrt{2\sqrt{5}-10}}{4},$$

$$\alpha_4 = \frac{-1+\sqrt{5}}{4} + \frac{\sqrt{-10-2\sqrt{5}}}{4}, \quad \alpha_5 = \frac{-1+\sqrt{5}}{4} - \frac{\sqrt{-10-2\sqrt{5}}}{4}. \tag{5}$$

Якщо урахувати, що всі цілі числа, які фігурують у формулах (5), легко можна дістати з коефіцієнтів рівняння (4) (тобто чисел 1, —1,0) за допомогою скінченного числа раціональних дій, то стає зрозумілим, що рівняння (4) розв'язується у квадратних радикалах.

Очевидно, виконання раціональних операцій рівносильне розв'язуванню лінійного рівняння виду $ax + b = c$, а добування квадратного кореня — розв'язуванню двочленного квадратного рівняння $x^2 - a = 0$. Тому *можливість розв'язати деяке рівняння в квадратних радикалах означає, що його можна звести до скінченного ланцюжка двочленних рівнянь, степінь яких не вищий від 2, а коефіцієнти раціонально виражаються через коефіцієнти даного рівняння та корені проміжних рівнянь ланцюжка.*

Проілюструємо це на прикладах.

П р и к л а д и. 2. Розглянемо квадратне рівняння

$$ax^2 + bx + c = 0 \quad (a \neq 0).$$

Звичайне виведення формули для розв'язків цього рівняння рівносильне послідовному розв'язанню такого ланцюжка двочленних рівнянь:

$$\begin{cases} y^2 - \dfrac{b^2 - 4ac}{4a^2} = 0, \\ x + \dfrac{b}{2a} - y = 0. \end{cases}$$

[1] Е в а р і с т Г а л у а (1811—1832) — видатний французький математик. Про його незвичайне життя і коротку, але блискучу наукову діяльність радимо прочитати у книзі Д а л ь м а А. Э. Галуа, революционер и математик. М., Физматгиз, 1960.

3. Рівняння $x^5 - 1 = 0$ спочатку зведемо до ланцюжка рівнянь, степінь яких не перевищує 2:

$$\begin{cases} x - 1 = 0, \\ z^2 + z - 1 = 0, \\ x + \frac{1}{x} = z. \end{cases}$$

Відповідний ланцюжок двочленних рівнянь має вигляд:

$$\begin{cases} x - 1 = 0, \\ y^2 - \left(1 + \frac{1}{4}\right) = 0, \\ z + \frac{1}{2} - y = 0, \\ u^2 + 1 - \frac{z^2}{4} = 0, \\ x - \frac{z}{2} - u = 0. \end{cases}$$

Знайшовши всі можливі значення x, які задовольняють цю систему, дістанемо корені $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ рівняння $x^5 - 1 = 0$ (див. приклад 1).

34.2. Зв'язок з розширеннями числових полів. Подивимось тепер на розв'язність рівняння в квадратних радикалах з точки зору алгебраїчних розширень числових полів.

Нехай $f(x)$ — ліва частина даного рівняння (1) — є многочлен над числовим полем Δ. При цьому вважатимемо, що Δ є мінімальне числове поле, яке містить коефіцієнти a_0 a_1, ..., a_n многочлена $f(x)$, тобто

$$\Delta = \mathbf{Q}(a_0, a_1, \ldots, a_n),$$

оскільки будь-яке числове поле містить підполе $\mathbf{Q}$ усіх раціональних чисел.

Означення 1. *Основним полем* Δ (або областю раціональності) *рівняння* $a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0$ *називають алгебраїчне розширення* $\mathbf{Q}(a_0, a_1, \ldots, a_n)$ *поля* $\mathbf{Q}$ *раціональних чисел, утворене приєднанням коефіцієнтів даного рівняння.*

Доцільність введення до розгляду основного поля Δ видно з такої леми.

Лема. *Для того щоб рівняння* (1) *було розв'язним у квадратних радикалах, необхідно і достатньо, щоб кожний з його коренів можна було виразити через деякі числа поля* Δ *за допомогою скінченного числа раціональних операцій та дій добування квадратного кореня.*

При доведенні леми і далі нам зручно вживати вислів типу «число ξ виражається у квадратних радикалах через числа $b_1, b_2, \ldots, b_m$». Це означає, що число ξ можна подати через числа b_j $(j = 1, 2, \ldots, m)$ за допомогою скінченного числа раціональних операцій та дій добування квадратного кореня.

Доведення леми. Згідно з означеннями 1 п. 34.1, рівняння (1) розв'язне у квадратних радикалах, якщо кожний його корінь виражається у квадратних радикалах через його коефіцієнти. Отже, доведення леми зводиться до перевірки того, що певне число ξ виражається у квадратних радикалах через коефіцієнти рівняння (1) тоді і тільки тоді, коли воно виражається у квадратних радикалах через деякі числа основного поля Δ рівняння (1).

Якщо число ξ виражається у квадратних радикалах через коефіцієнти a_j $(j = 0, 1, \ldots, n)$, то тим самим воно виражається у квадратних радикалах через числа поля Δ, адже

$$a_i \in \Delta = \mathbf{Q}(a_0, a_1, \ldots, a_n).$$

Нехай тепер ξ виражається у квадратних радикалах через числа $b_1, b_2, \ldots, b_m$ з поля Δ. Кожне з чисел b_j, в свою чергу, виражається раціонально (тобто за допомогою раціональних дій) через певні раціональні числа і коефіцієнти a_i $(i = 0, 1, 2, \ldots, n)$ даного рівняння. Проте будь-яке раціональне число також раціонально виражається через коефіцієнти a_i $(i = 0, 1, \ldots, n)$, адже серед останніх є хоч одне число, відмінне від нуля, наприклад a_n. Тому числа 0, 1 і -1 раціонально виражаються через a_i: $0 = a_n - a_n$, $1 = \frac{a_n}{a_n}$. А довільне раціональне число раціонально виражається через 0, 1, -1:

$$\frac{m}{n} = \frac{\overbrace{1 + \cdots + 1}^{m}}{\underbrace{1 + \cdots + 1}_{n}} \text{ при } \frac{m}{n} > 0 \text{ і } \frac{m}{n} = \frac{\overbrace{(-1) + \cdots + (-1)}^{m}}{\underbrace{1 + \cdots + 1}_{n}} \text{ при } \frac{m}{n} < 0.$$

Отже, остаточно, кожне b_j $(i = 1, \ldots, m)$, а тому й ξ раціонально виражаються через коефіцієнти $a_0, a_1, \ldots, a_n$. Лему доведено.

Тепер зрозуміло, що *розв'язність рівняння у квадратних радикалах означає можливість виразити всі його корені у квадратних радикалах через числа основного поля* Δ.

З другого боку, очевидно, що *можливість виразити якесь число* ξ *у квадратних радикалах через числа поля* Δ *означає можливість виразити усі числа поля* $\Delta(\xi)$ *у квадратних радикалах через числа поля* Δ. Це безпосередньо випливає з того, що кожне число $\eta \in \Delta(\xi)$ раціонально виражається через ξ і деякі числа поля Δ, а ξ — у квадратних радикалах через числа поля Δ.

Означення 2. *Якщо* Δ *— основне поле рівняння*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0,$$

а $\alpha_1, \alpha_2, \ldots, \alpha_n$ *— корені цього рівняння, то поле* $\Omega = \Delta(\alpha_1, \alpha_2, \ldots, \alpha_n)$, *утворене приєднанням до* Δ *усіх коренів* α_j $(j = 1, \ldots, n)$ *називається нормальним полем* (нормою) *або полем розкладу даного рівняння.*

На підставі попереднього викладу легко переконатись у справедливості такого твердження:

Теорема 1. *Для того щоб рівняння* (1) *розв'язувалось у квадратних радикалах, необхідно і достатньо, щоб будь-яке число з його нормального поля* Ω *виражалось у квадратних радикалах через числа основного поля* Δ.

Доведення. Якщо будь-яке число з поля $\Omega = \Delta(\alpha_1, \alpha_2, \ldots, \alpha_n)$ виражається у квадратних радикалах через числа поля Δ, то,

зокрема, й корені рівняння $\alpha_1, \alpha_2, \ldots, \alpha_n$ виражаються у квадратних радикалах через числа поля Δ, тобто рівняння розв'язується у квадратних радикалах. Навпаки, якщо $\alpha_1, \alpha_2, \ldots, \alpha_n$ виражаються у квадратних радикалах через числа поля Δ, то й усі числа полів $\Delta(\alpha_1)$, $\Delta(\alpha_1)(\alpha_2), \ldots, \Delta(\alpha_1)(\alpha_2)\ldots(\alpha_n) = \Delta(\alpha_1, \alpha_2, \ldots, \alpha_k) = \Omega$ виражаються у квадратних радикалах через числа поля Δ, що й треба було довести.

Наслідок. *Якщо Δ_1 — квадратичне розширення поля Δ, то будь-яке число $\xi \in \Delta_1$ виражається у квадратних радикалах через число поля Δ.*

Справді, за означенням квадратичного розширення (п. 32.2), $\Delta_1 = \Delta(\alpha_1)$, де α_1 — корінь деякого квадратного рівняння $ax^2 + bx + c = 0$, коефіцієнти якого належать Δ, а корені α_1, α_2 — не належать. Очевидно, $\alpha_2 \in \Delta(\alpha_1)$ (бо $\alpha_2 = -\frac{b}{a} - \alpha_1$), тому $\Delta_1 = \Delta(\alpha_1) = \Delta(\alpha_1, \alpha_2)$ є нормою даного квадратного рівняння. Тепер зрозуміло, що розглядуване твердження справді є наслідком попередньої теореми.

Отже, питання про розв'язність алгебраїчного рівняння у квадратних радикалах звелось до питання про можливість виразити усі числа деякого поля Ω у квадратних радикалах через числа його підполя Δ.

34.3. Числа, що виражаються у квадратних радикалах. Число ξ, яке виражається у квадратних радикалах через числа поля Δ, як випливає з означення (п. 34.2), можна подати у формі

$$\xi = r(\sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_m}), \tag{6}$$

де $r(x_1, x_2, \ldots, x_n)$ — раціональна функція над полем Δ, а $q_1, \ldots, q_m$ — числа, які виражаються у квадратних радикалах через числа поля Δ.

Приклад 1. Число

$$\alpha = \frac{-1-\sqrt{5}}{4} + \frac{\sqrt{2\sqrt{5}-10}}{4}, \tag{7}$$

з яким ми зустрічались у попередньому викладі, виражається у квадратних радикалах через числа поля $\mathbf{Q}$. Його можна подати у вигляді

$$\alpha = r(\sqrt{q_1}, \sqrt{q_2}),$$

де $r = r(x_1, x_2) \underset{\text{df}}{=} \frac{-1-x_1+x_2}{4}$ — раціональна функція над полем $\mathbf{Q}$. У даному разі $q_1 = 5$, $q_2 = 2\sqrt{5} - 10$, тобто в свою чергу виражаються у квадратних радикалах через числа поля $\mathbf{Q}$: $q_1 \in \mathbf{Q}$, $q_2 = r'(\sqrt{q_1})$, де $r'(x) = 2x - 10$.

У загальному випадку кореневий вираз $\sqrt{q_i}$ побудовано так, що у ньому кілька квадратних коренів добуваються один з одного. Назвемо порядком даного кореневого виразу *число послідовних квадратних радикалів, що стоять один під одним*. Так, вираз $\sqrt{5}$ має порядок 1; вираз $\sqrt{2\sqrt{5}-10}$ — порядок 2; вираз виду $\sqrt{a+b\sqrt{c+d\sqrt{c}}}$ має порядок 3. Позначимо через p найбільший з порядків виразів $\sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_m}$ у представлені (6) числа ξ.

Для числа α, поданого виразом (7), $p_\alpha = 2$. Для числа

$$\beta = \sqrt{2\sqrt{3-\sqrt{2}}} + \frac{1}{\sqrt{2+\sqrt{2+\sqrt{3}}}}, \quad p_\beta = 3.$$

Зауважимо, що число p_ξ не визначається числом μ однозначно, а залежить від конкретного способу вираження ξ у квадратних радикалах через числа поля Δ. Так число $\xi = \sqrt{3+2\sqrt{2}+\sqrt{3}}$ можна подати і у вигляді $\xi = \sqrt{2}+\sqrt{3}+1$. Основним полем тут є поле $\mathbf{Q}$. У першому випадку $p_\xi = 2$, в другому — $p_\xi = 1$.

У дальшому викладі, вживаючи величину порядку p_ξ, матимемо на увазі, що вона відноситься до певного (взагалі кажучи, не єдино можливого) зображення цього числа у квадратних радикалах, є показником цього зображення. Якщо $\xi \in \Delta$, вважатимемо $p_\xi = 0$.

Теорема 2. *Для того щоб число ξ виражалось у квадратних радикалах через числа поля Δ, необхідно і достатньо, щоб існувала скінченна сукупність полів $\Delta_1, \Delta_2, \ldots, \Delta_k$ таких, що:*

1) Δ_1 є квадратичним розширенням поля Δ;

2) Δ_{j+1} є квадратичним розширенням поля Δ_j $(j = 1, 2, \ldots, k-1)$;

3) число ξ належить полю Δ_k.

Коротше кажучи, ξ виражається у квадратних радикалах через числа поля Δ тоді і тільки тоді, коли існує ланцюжок полів

$$\Delta \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \cdots \subseteq \Delta_{k-1} \subseteq \Delta_k,$$

з яких перше є заданим полем, кожне наступне — квадратичним розширенням попереднього, а останнє містить число ξ.

Зауважимо, що разом з ξ в Δ_k входить усе поле $\Delta(\xi)$, тобто мінімальне поле, що містить ξ і Δ (бо Δ_k також містить як ξ, так і Δ).

Доведення теореми.

Достатність. Нехай існує ланцюжок полів з властивостями 1)—3). Тоді на підставі наслідку з теореми 1 всі числа поля Δ_1 (квадратичного розширення поля Δ) виражаються у квадратних радикалах через числа поля Δ. Повторне застосування цього наслідку показує[1], що всі елементи полів $\Delta_2, \Delta_3, \ldots, \Delta_{k-1}$ і, нарешті, поля Δ_k, серед яких і число ξ, виражаються у квадратних радикалах через числа поля Δ.

Необхідність. Нехай нам дано поле Δ і відомо, що число ξ виражається у квадратних радикалах через числа цього поля. Припустимо, що якесь з відповідних зображень числа ξ через числа поля Δ характеризується порядком p_ξ.

Потрібно довести існування ланцюжка полів з властивостями 1)—3). Доведення проведемо індукцією по p_ξ. Очевидно, для будь-якого зображення числа ξ у квадратних радикалах p_ξ є нулем (у випадку $\xi \in \Delta$) або натуральним числом. Оскільки теорема стає тривіальною для випадку $\xi \in \Delta$ (тоді $\Delta_1 = \Delta_2 = \ldots = \Delta_k = \Delta$), вважатимемо $p_\xi \geqslant 1$.

[1] Тут використовується такий очевидний факт: якщо число ξ виражається у квадратних радикалах через числа поля Δ', а всі числа поля Δ' виражаються у квадратних радикалах через числа поля Δ'', то ξ виражається у квадратних радикалах через числа поля Δ''.

При $p_\xi = 1$ можна ξ подати у формі

$$\xi = r(\sqrt{q_1}, \ldots, \sqrt{q_k}),$$

де $r(x_1, \ldots, x_k)$ — раціональна функція над полем Δ, $q_i \in \Delta$, а $\sqrt{q_i} \notin \Delta$. Очевидно, в цьому випадку $\xi \in \Delta(\sqrt{q_1}, \ldots, \sqrt{q_k})$, тобто розширенню поля Δ, утвореному приєднанням радикалів $\sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_k}$. Але згідно з теоремою 4 п. 33.3

$$\Delta(\sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_k}) = \Delta(\sqrt{q_1})(\sqrt{q_2}) \ldots (\sqrt{q_k}),$$

тобто існує такий ланцюжок полів

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_k, \tag{8}$$

що $\Delta_1 = \Delta(\sqrt{q_1})$, $\Delta_2 = \Delta_1(\sqrt{q_2})$, $\ldots$, $\Delta_k = \Delta_{k-1}(\sqrt{q_k})$.

Зрозуміло, що Δ_1 є квадратичне розширення поля Δ, бо утворене з Δ приєднанням $\sqrt{q_1}$ — кореня квадратного двочлена $x^2 - q_1$ над полем Δ. Якщо $\sqrt{q_2} \in \Delta_1$, то $\Delta_2 = \Delta_1$; якщо ж $\sqrt{q_2} \notin \Delta_1$, то Δ_2 є квадратичне розширення поля Δ_1. Взагалі, якщо $\sqrt{q_l} \in \Delta_{l-1}$ ($2 \leqslant l \leqslant k$), то $\Delta_l = \Delta_{l-1}$; в противному разі Δ_l є квадратичне розширення поля Δ_{l-1}. Отже, відкидаючи у послідовності (8) такі поля, які збігаються з попередніми, дістанемо ланцюжок

$$\Delta \subseteq \Delta_1 \subseteq \qquad \Delta_{n_1} \subseteq \cdots \subseteq \Delta_{n\,s}, \tag{9}$$

де $n_1, n_2, \ldots, n_s$ — якісь з чисел 2, 3, $\ldots$, k (причому $n_1 < n_2 < \ldots < n_s$). Цей ланцюжок задовольняє умови теореми, бо Δ_1 — квадратичне розширення поля Δ, Δ_{n_l} — квадратичне розширення поля Δ_{n_l-1} (вважаючи $n_0 = 1$), а поле Δ_{n_s} містить число ξ.

Припустимо тепер, що наше твердження правильне для усіх чисел η, для яких порядок $p_\eta < m$. У зв'язку з неоднозначністю показника p_ξ для даного числа, це слід розуміти так, що *хоч для одного зображення* цього числа у квадратних радикалах через числа поля Δ має місце $p < m$. Доведемо справедливість твердження для будь-якого числа ξ, для якого $p_\xi = m$.

Оскільки для ξ існує зображення у квадратних радикалах з порядком p_ξ, то можна записати

$$\xi = r(\sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_n}), \tag{10}$$

де $r(x_1, x_2, \ldots, x_n)$ — раціональна функція над полем Δ, а $q_1, q_2, \ldots, q_n$ можна виразити у квадратних радикалах через числа поля Δ. При цьому

$$p_{q_i} \leqslant m - 1 \qquad (i = 1, 2, \ldots, n), \tag{11}$$

бо якби хоч для одного i (скажімо, i_0) було $p_{q_{i_0}} \geqslant m$, то $\sqrt{q_{i_0}}$ мав би порядок $p_{q_{i_0}} + 1 > m$, що суперечить припущенню $p_\xi = m$ для зображення (10).

За припущенням індукції для кожного з чисел q_i (на підставі умови (11)) справджуються умови теореми. Зокрема, існує ланцюжок полів

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_{k_s}$$

такий, що кожне наступне поле є квадратичним розширенням попереднього і останнє Δ_{k_s} містить число q_1. Приєднавши до цього ланцюжка поле $\Delta_{k_s+1} = \Delta_{k_s}(\sqrt{q_1})$, дістанемо ланцюжок послідовних квадратичних розширень

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_{k_s} \subseteq \Delta_{k_s+1},$$

останнє з яких містить $\sqrt{q_1}$.

Таке саме міркування можна провести для числа q_2. Але тепер за вихідне поле ланцюжка квадратичних розширень візьмемо не Δ, а Δ_{k_s+1}. Це можна зробити тому, що всі числа поля Δ містяться і в Δ_{k_s+1}. Отже, існує ланцюжок квадратичних розширень

$$\Delta_{k_s+1} \subseteq \Delta_{k_s+2} \subseteq \cdots \subseteq \Delta_{k_3} \subseteq \Delta_{k_3+1},$$

причому $\sqrt{q_1} \in \Delta_{k_2+1}$ і $\sqrt{q_2} \in \Delta_{k_2+1}$. Продовжуючи цей процес, побудуємо ланцюжки квадратичних розширень

$$\Delta_{k_3+1} \subseteq \Delta_{k_3+2} \subseteq \cdots \subseteq \Delta_{k_3} \subseteq \Delta_{k_3+1} \qquad (\sqrt{q_1}, \sqrt{q_2}, \sqrt{q_3} \in \Delta_{k_3+1}),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\Delta_{k_{n-1}+1} \subseteq \Delta_{k_{n-1}+2} \subseteq \cdots \subseteq \Delta_{k_n} \subseteq \Delta_{k_n+1}$$

(всі $\sqrt{q_i} \in \Delta_{k_n+1}$ ($i = 1, 2, \ldots, n$)).

Зауважимо, що деякі з цих ланцюжків можуть складатися з самого лише поля Δ_{k_i+1} (це буде у випадках, коли $\sqrt{q_{i+1}} \in \Delta_{k_i+1}$).

Оскільки поле Δ_{k_n+1} містить усі числа $\sqrt{q_i}$ ($i = 1, 2, \ldots, n$), то воно містить і число ξ. Отже, ланцюжок полів

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_{k_1} \subseteq \Delta_{k_1+1} \subseteq \Delta_{k_1+2} \subseteq \cdots \subseteq \Delta_{k_2} \subseteq$$

$$\subseteq \Delta_{k_2+1} \subseteq \ldots \Delta\, \Delta_{k_n} \subseteq \Delta_{k_n+1}$$

задовольняє умови 1—3. Цим доведено, що умови теореми справедливі для будь-якого числа ξ, для якого $p_\xi = m$. Згідно з принципом індукції, це твердження правильне для усіх натуральних p_ξ, тобто для всіх чисел, які виражаються у квадратних радикалах через числа поля Δ.

Теорему доведено повністю.

Приклади. 2. Побудуємо ланцюжок квадратичних розширень для числа α, вираженого в квадратних радикалах через раціональні числа у формі (7).

Очевидно, таким ланцюжком буде $\Delta \subseteq \Delta_1 \subseteq \Delta_2$, де $\Delta = \mathbf{Q}$, $\Delta_1 = \mathbf{Q}(\sqrt{5})$, $\Delta_2 = \Delta_1(\sqrt{2\sqrt{5} - 10})$. Зауважимо, що Δ_2 є квадратичне розширення поля Δ_1, бо його дістаємо з Δ_1 приєднанням елемента $\alpha_1 = \sqrt{2\sqrt{5} - 10}$, що є коренем многочлена 2-го степеня $f(x) = x^2 - (2\sqrt{5} - 10)$ над полем Δ_1.

3. Для числа $\xi = \sqrt{\sqrt{1 + \sqrt{2}} + \sqrt{5}}$ відповідний ланцюжок складається з 5 полів: $\Delta = \mathbf{Q}$, $\Delta_1 = \mathbf{Q}(\sqrt{5})$, $\Delta_2 = \Delta_1(\sqrt{2})$, $\Delta_3 = \Delta_2(\sqrt{1 + \sqrt{2}})$, $\Delta_4 = \Delta_3(\sqrt{\sqrt{1 + \sqrt{2}} + \sqrt{5}})$.

34.4. Ознаки того, що число виражається у квадратних радикалах. Теорема 2, доведена у попередньому пункті, є критерієм можливості виразити число у квадратних радикалах через числа даного поля Δ. Цей критерій має той недолік, що його важко застосовувати при дослідженні конкретної задачі. Адже побудова ланцюжка полів Δ_l здійснюється порівняно легко тільки тоді, коли ξ явно виражене через числа поля Δ за допомогою квадратних радикалів. Якщо ж (як це і буває у більшості задач на побудову) число ξ задано лише деякою умовою, що пов'язує його з елементами поля Δ, то безпосередньо з'ясувати питання про можливість чи неможливість побудови ланцюжка полів Δ_l практично неможливо.

Доведемо, спираючись на теорему 2, ряд тверджень, які у багатьох випадках дають змогу повністю розв'язати питання про можливість або неможливість виразити конкретне число у квадратних радикалах через числа основного поля.

Теорема 3. *Всі числа, які можна виразити у квадратних радикалах через числа поля Δ, алгебраїчні над цим полем.*

Доведення. Відповідно до теореми 2, будь-яке число ξ, яке можна виразити у квадратних радикалах через числа поля Δ належить деякому полю Δ_k, утвореному з Δ за допомогою ланцюжка квадратичних розширень: $\Delta \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \ldots \subseteq \Delta_k$.

За наслідком 2 з теореми 2, п. 33.2, поле Δ_k є скінченним розширенням поля Δ. Але тоді, за теоремою 3, п. 33.3, воно є й алгебраїчним розширенням цього поля, тобто усі числа поля Δ_k, зокрема ξ, алгебраїчні над полем Δ. Теорему доведено.

Нехай ξ — деяке число, і нам треба з'ясувати, чи можна його виразити у квадратних радикалах через числа поля Δ. Ураховуючи щойно доведену теорему, завжди можна вважати, що ξ є коренем деякого многочлена над полем Δ:

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0, \qquad (12)$$

бо для трансцендентного (неалгебраїчного) над полем Δ числа це питання вже з'ясоване — таке число не можна виразити у квадратних радикалах через числа поля Δ.

Отже, бажано було б мати ознаку, за допомогою якої з властивостей многочлена (12) можна було б зробити висновок про можливість чи неможливість виразити корінь ξ у квадратних радикалах через елементи основного поля Δ.

Встановимо спочатку необхідну умову такої можливості. Доведення достатньої умови потребуватиме розгляду деяких додаткових питань (п. 34.6).

Теорема 4. (Необхідна умова можливості виразити корінь многочлена у квадратних радикалах). *Якщо корінь ξ незвідного у полі Δ многочлена* (12) *виражається у квадратних радикалах через числа поля Δ, то степінь n многочлена $f(x)$ є числом виду 2^m* (m—ціле невід'ємне).

Доведення. Оскільки ξ виражається у квадратних радикалах через числа поля Δ, то за теоремою 2, п. 34.3 існує скінченна послідовність квадратичних розширень: $\Delta \subseteq \Delta_1 \subseteq \ldots \subseteq \Delta_k$ така, що поле Δ_k містить у собі ξ, а тому й мінімальне поле $\Delta(\xi)$. Згідно з теоремою 1, п. 33.2, $\Delta(\xi)$ є скінченне розширення поля Δ степеня n. Поле ж Δ_k є скінченним розширенням поля Δ степеня 2^k (як це випливає з наслідку 2 теореми 2, п. 33.2). Оскільки $\Delta(\xi)$ є нідполе поля Δ_k, то за лемою 2 п. 33.2 число n є дільником числа 2^k. Це й означає, що n є числом виду 2^m. Теорему доведено.

Відомо, що будь-яка необхідна умова для виконання якогось факту є одночасно достатньою умовою для невиконання цього факту. Це означає, що з доведеної теореми можна дістати такий наслідок:

Наслідок. *Корені многочлена $f(x)$, незвідного в полі Δ, степінь якого не є степенем числа 2, не виражаються в квадратних радикалах через числа цього поля.*

34.5. Розв'язність у квадратних радикалах рівнянь 3-го і 4-го степенів. Багато цікавих і важливих задач зводиться до з'ясування питания про можливість виразити у квадратних радикалах корені многочленів 3-го степеня. Для таких многочленів, використовуючи доведену теорему 4, легко знайти зручний і простий критерій розв'язності у квадратних радикалах.

Теорема 5. *Для того щоб усі корені кубічного многочлена над полем Δ виражались у квадратних радикалах через числа поля Δ, необхідно і достатньо, щоб цей многочлен був звідним у полі Δ.*

Доведення. Справді, якщо многочлен $f(x)$ незвідний у полі Δ, то у зв'язку з тим, що його степінь дорівнює 3, за теоремою 4 всі його корені не можна виразити у квадратних радикалах. Якщо ж многочлен $f(x)$ звідний у полі Δ, то може бути два випадки. У першому випадку многочлен $f(x)$ розкладається на добуток трьох лінійних двочленів з коефіцієнтами з поля Δ. Тоді всі корені належать заданому полю Δ і, отже, виражаються у квадратних радикалах. У другому випадку многочлен розкладається на добуток лінійного двочлена і незвідного квадратного тричлена. Приєднуючи до Δ корінь α цього квадратного тричлена, ми переходимо до квадратичного розширення $\Delta(\alpha)$, в якому знаходяться всі корені кубічного многочлена і, отже, всі корені виражаються у квадратних радикалах через числа поля Δ.

Теорему доведено.

Теорему 5 можна, очевидно, сформулювати ще й так:

Алгебраїчне рівняння 3-го *степеня*

$$f(x) \underset{\text{df}}{=} ax^3 + bx^2 + cx + d = 0,$$

де $f(x)$ — многочлен над полем Δ, розв'язується у квадратних радикалах тоді і тільки тоді, коли $f(x)$ звідний у полі Δ.

Зауважимо, що умова для всіх коренів кубічного многочлена рівносильна умові хоч для одного кореня цього многочлена. Справді, з можливості виразити якийсь корінь многочлена 3-го степеня у квадратних радикалах випливає (за теоремою 4) звідність цього многочлена, а із звідності (за теоремою 5) — така можливість для усіх трьох коренів.

Приклад. Нехай над полем **Q** задано многочлен

$$\varphi(t) = t^3 + 16t - 32. \tag{13}$$

Щоб з'ясувати питання про можливість виразити його корені у квадратних радикалах, слід встановити звідність чи незвідність $\varphi(t)$ в полі **Q**. Для цього, в свою чергу, досить з'ясувати, чи має $\varphi(t)$ раціональні корені.

Як відомо (§ 31), всі раціональні корені зведеного многочлена з цілими коефіцієнтами є цілі числа і дільники вільного члена. Отже, ці корені слід шукати серед чисел ± 1; ± 2; ± 4; $-\pm 8$; ± 16; ± 32. Легко впевнитись безпосередньо, що жодне з цих чисел не є коренем многочлена (13). Отже, $\varphi(t)$ — незвідний у полі раціональних чисел многочлен, тому жодний його корінь не виражається у квадратних радикалах через числа поля **Q**.

Відомо, що знаходження коренів многочлена 4-го степеня зводиться до розв'язання допоміжного кубічного рівняння — резольвенти (1, § 17). Можна довести, що *рівняння 4-го степеня розв'язується у квадратних радикалах тоді і тільки тоді, коли його резольвента розв'язується у квадратних радикалах* (див. Костарчук В. М., Хацет Б. І. Про можливе і неможливе в геометрії циркуля і лінійки. К., «Радянська школа»,1971).

34.6. Загальний критерій розв'язності у квадратних радикалах. Розглянемо тепер довільне рівняння

$$f(x) \underset{\text{df}}{=} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{14}$$

з основним полем Δ і поставимо питання про критерій розв'язності цього рівняння у квадратних радикалах.

Одну з форм такого критерію було вже розглянуто вище (п. 34.2). А саме, згідно з теоремою 1, рівняння (14) розв'язується у квадратних радикалах тоді і тільки тоді, коли всі числа з його нормального поля Ω виражаються у квадратних радикалах через числа основного поля Δ.

Тепер ми маємо змогу надати цьому критерію більш конкретної форми.

Теорема 6. (Критерій розв'язності рівняння у квадратних радикалах). *Для того щоб рівняння* (14) *з основним полем* Δ *і нормою* Ω *розв'язувалось у квадратних радикалах, необхідно і достатньо, щоб степінь норми* Ω *відносно поля* Δ *був цілим невід'ємним степенем числа* 2, *тобто* $(\Omega : \Delta) = 2^m$ ($m \geqslant 0$, ціле).

Доведення. *Необхідність.* Нехай многочлен $f(x)$ має корені $\alpha_1, \alpha_2, \ldots, \alpha_n$. Якщо корінь α_1 многочлена $f(x)$ виражається у квадратних радикалах, то, згідно з теоремою 2 (п. 34.3), існує така послідовність квадратичних розширень:

$$\Delta \subseteq \Delta_1 \subseteq \ldots \subseteq \Delta_k, \tag{15}$$

що розширення Δ_k містить $\Delta(\alpha_1)$. Якщо корінь α_2 при цьому не належить Δ, то, в зв'язку з можливістю виразити α_2 у квадратних радикалах через числа з поля Δ і, тим більше, з поля Δ_k, послідовність (15) можна продовжити, побудувавши дальші квадратичні розширення

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_k \subseteq \Delta_{k+1} \subseteq \cdots \subseteq \Delta_l$$

так, щоб розширення Δ_l містило і корінь α_1, і корінь α_2, тобто містило поле $\Delta(\alpha_1, \alpha_2)$. Продовжуючи такі міркування, ми впевнюємося, що існує така послідовність квадратичних розширень

$$\Delta \subseteq \Delta_1 \subseteq \cdots \subseteq \Delta_s,$$

що поле Δ_s включає норму $\Omega = \Delta(\alpha_1, \alpha_2, \ldots, \alpha_n)$ многочлена $f(x)$. Степінь $(\Delta_s : \Delta) = 2^s$. Через те що $\Delta_s \supseteq \Omega$, то за лемою 2 (п.33.2) степінь поля Ω відносно поля Δ є дільником числа 2^s. Отже, $(\Omega : \Delta)$ є числом виду 2^m, тобто $(\Omega : \Delta) = 2^m$, що й треба було довести.

Достатність. Ідея доведення достатності полягає ось у чому. Якщо $(\Omega : \Delta) = 2^m$, то встановлюється існування між нормою Ω і полем Δ такого квадратичного розширення Δ_1 поля Δ, що $\Delta \subseteq \Delta_1 \subseteq \Omega$. Іншими словами, можна побудувати такий квадратний тричлен, незвідний у полі Δ, що приєднання його коренів до поля Δ приводить до квадратичного розширення Δ_1, яке належить нормі Ω. Розглядаючи тепер поля Δ_1 і Ω і зазначаючи, що $(\Omega : \Delta_1) = 2^{m-1}$, знову встановлюють існування такого незвідного в Δ_1 квадратного тричлена над Δ_1, приєднання коренів якого до поля Δ_1 приводить до квадратичного розширення Δ_2 поля Δ_1 такого, що $\Delta_2 \subseteq \Omega$. Міркуючи так і далі, встановлюють існування такої послідовності квадратичних розширень $\Delta \subseteq \Delta_1 \subseteq \ldots \subseteq \Delta_k$, що $\Delta_k = \Omega$. Тим самим доведення теореми завершується.

Докладний виклад реалізації цієї ідеї ми тут наводити не будемо (див. Костарчук В. М., Хацет Б. І. Про можливе і неможливе в геометрії циркуля і лінійки. К., «Радянська школа», 1971).

Сформульований критерій, застосовний у принципі до будь-якого алгебраїчного рівняння, має (як і критерій, поданий у теоремі 2) той недолік, що стосовно до окремих конкретних задач він часто виявляється недостатньо ефективним. Це пов'язано з тим, що в більшості випадків для даного многочлена над полем Δ зовсім не просто побудувати норму Ω і визначити її степінь $(\Omega : \Delta)$.

Проте з цього загального критерію можна вивести ряд ефективних ознак розв'язності для тих чи інших окремих класів задач. Цікавий і важливий приклад побудови такої ознаки на основі наведеного загального критерію читач знайде в § 35.

§ 35. ПОБУДОВНІСТЬ ЧИСЕЛ ЦИРКУЛЕМ І ЛІНІЙКОЮ

Результати попереднього параграфа можуть бути застосовані до встановлення критеріїв розв'язності геометричних задач на побудову за допомогою циркуля і лінійки. Як відомо, таке обмеження засобів розв'язування конструктивних геометричних задач, що склалось історично, є основою теорії геометричних побудов, яка вивчається в середній і вищій школі. Нижче буде виведено необхідну і достатню умову побудовності числа циркулем і лінійкою, що дає змогу в принципі з'ясувати можливість чи неможливість розв'язати будь-яку конструктивну задачу за допомогою цих інструментів. Ми обмежимось розглядом лише планіметричних задач на побудову.

Виходячи з того, що читач обізнаний з основами теорії геометричних побудов на площині, яка вивчається в курсі геометрії, наведемо деякі означення та нагадаємо окремі твердження цієї теорії у формі, яка зручна для застосування до даної проблеми теорії числових полів.

35.1. Поняття побудовності чисел. У кожній конструктивній задачі можна вважати деяку систему точок заданою і деяку іншу — шуканою. Якщо ці точки розглядати відносно певної системи координат на площині, то можна вважати заданою деяку сукупність чисел (координати заданих точок), а шуканою — іншу сукупність чисел (координати шуканих точок). Якщо дану задачу на побудову можна розв'язати за допомогою циркуля і лінійки, вважатимемо, що *кожне з шуканих чисел може бути побудоване* (або просто *побудовне*) *циркулем і лінійкою, виходячи з даної сукупності чисел.*

Можна говорити про побудовність або непобудовність довільного дійсного числа ξ, виходячи з множини M заданих дійсних чисел, і незалежно від тієї чи іншої конкретної конструктивної задачі.

Множина M заданих чисел визначає систему $\mathfrak{M}$ точок площини, обидві координати яких належать множині M.

Означення 1. *Число ξ побудовне циркулем і лінійкою, виходячи з множини M, якщо побудовна циркулем і лінійкою, виходячи з системи $\mathfrak{M}$, хоч одна точка площини, для якої ξ є однією з координат.*

Проблема, яка нас цікавить, полягає в тому, щоб знайти *критерій можливості або неможливості побудувати певне число ξ за допомогою циркуля і лінійки, виходячи з заданої сукупності чисел M.*

Поки що, говорячи про побудовність чисел, ми під словом «число» розуміли д і й с н е число. Адже відповідно до принципів аналітичної геометрії задані точки на площині завжди мають дійсні координати, а нові точки утворюються за допомогою прямих та кіл лише тоді, коли відповідні рівняння для їх координат мають дійсні розв'язки.

Проте іноді в конструктивних задачах зручно інтерпретувати кожну точку площини не як пару (x, y) дійсних чисел, а як комплексне число $z = x + iy$ (п. 35.5). Тому доцільно узагальнити поняття побудовності на випадок к о м п л е к с н о г о числа. При цьому числа заданої множини M ми, як і раніше, вважатимемо дійсними; такий підхід достатній для аналізу будь-яких реальних конструктивних задач.

Означення 2. *Вважатимемо, що комплексне число $\zeta = \xi + i\eta$ побудовне циркулем і лінійкою, виходячи з заданої множини M дійсних чисел, якщо такими є дійсні числа ξ і η, які визначають дійсну і уявну частини числа ζ.*

35.2. Побудовність і числові поля. Позначимо, як і раніше, сукупність заданих чисел через M, а сукупність усіх чисел, які можуть бути побудовані циркулем і лінійкою, виходячи з множини M,— через N (поки що вважатимемо всі розглядувані числа дійсними). Зрозуміло, що $N \supseteq M$. Крім того, завжди припускатимемо, що числа 0 і 1 належать множині M (це не обмежує загальності розглядуваних конструктивних задач). Важливу характеристику множини N дає таке твердження, яке є безпосереднім наслідком того, що сума, різниця, добуток і частка двох побудовних чисел є побудовне число.

Теорема 1. *Сукупність N чисел, які можна побудувати циркулем і лінійкою, виходячи з множини M заданих чисел, утворює числове поле.*

Наслідок. *Всі раціональні числа побудовні циркулем і лінійкою.*

Справді, кожне числове поле включає поле раціональних чисел. Отже, незалежно від того, яка сукупність чисел M задана, всі раціональні числа входять в N.

Це твердження має велике значення. Оскільки за допомогою раціональних чисел можна з довільним степенем точності подати будь-яке дійсне число, то з нього випливає, що *кожна задача на побудову може бути розв'язана циркулем і лінійкою, наближено з довільним степенем точності.*

Теорема 2. *Нехай M — дана числова множина, а $P\{M\}$ — мінімальне поле, яке її містить. Тоді кожне число поля $P\{M\}$ побудовне циркулем і лінійкою виходячи з множини M.*

Справді, сукупність N всіх чисел, які можна побудувати, виходячи з множини M, є полем, яке містить множину M (теорема 1). Тому мінімальне поле $P\{M\}$ є підполем поля N. Отже, всі числа з $P\{M\}$ належать множині N, тобто побудовні, виходячи з множини M.

Теорема 2 показує, що в кожній задачі на побудову *можна вважати заданою множиною чисел деяке числове поле.* Справді, якщо задана числова множина M, то кожне число з мінімального поля $P\{M\}$ побудовне лінійкою і циркулем і тому $P\{M\}$ також може розглядатися як задане. Надалі ми задане числове поле завжди позначатимемо через Δ. Очевидно, що для задач, в яких задаються лише числа 0 і 1, заданим полем слід вважати поле $\mathbf{Q}$ раціональних чисел — мінімальне поле, що містить число 1.

Нехай тепер треба з'ясувати, чи може деяке число x бути побудоване циркулем і лінійкою, виходячи з поля Δ. Якщо $x \in \Delta$, то питання відпадає само собою, отже, завжди можна вважати $x \notin \Delta$. Розглянемо мінімальне поле $\Delta(x)$, яке містить Δ і число x (або просте розширення поля Δ, утворене приєднанням числа x). Якщо x побудовне, то такими є всі числа поля $\Delta(x)$. Справді, сукупність N усіх побудовних (виходячи з поля Δ) чисел містить як Δ, так і x, а тому містить і мінімальне поле $\Delta(x)$. Навпаки, якщо все поле $\Delta(x)$ побудовне, то таким є, зокрема, і число x. Отже, ми приходимо до такого висновку.

Питання про можливість побудови циркулем і лінійкою деякого числа x, виходячи з даної множини чисел, рівносильне питанню про можливість побудови всіх елементів деякого числового поля $\Delta(x)$, виходячи з певного підполя Δ цього поля.

Алгебраїчна суть використання саме циркуля і лінійки в геометричних побудовах характеризується такими твердженнями.

1) *Виходячи з поля Δ, за допомогою лише лінійки не можна побудувати жодної нової точки* (тобто точки, яка не належить Δ).

2) *Якщо число ξ побудоване внаслідок одноразового застосування циркуля, виходячи з поля Δ, то воно належить або полю Δ, або деякому квадратичному розширенню поля Δ.*

Ці твердження в тій чи іншій формі встановлюються в теорії геометричних побудов (див. К у р а н т Р., Р о б б и н с Г. Что такое математика. М., «Просвещение», 1967). Вони є наслідком того, що рівняння прямої на площині — лінійне, а рівняння кола — другого степеня. Тому координати точок перетину таких ліній виражаються через координати заданих точок, які визначають ці лінії, або раціонально, або за допомогою одного квадратного радикала.

35.3. Критерій побудовності числа циркулем і лінійкою. Нехай дано числове поле Δ, елементами якого є дійсні числа ($\Delta \subseteq \mathbf{R}$). З'ясуємо, при яких умовах певне число ζ (дійсне або комплексне) може бути побудоване циркулем і лінійкою, виходячи з поля Δ.

Нам потрібна буде така лема, яка стосується довільних числових полів (в тому числі й тих, що містять комплексні числа).

Лема 1. *Якщо Δ_1 — квадратичне розширення довільного числового поля Δ, то усі числа з Δ_1 побудовні циркулем і лінійкою, виходячи з поля Δ.*

Д о в е д е н н я. За означенням квадратичного розширення $\Delta_1 = \Delta(\alpha)$, де α — корінь квадратного тричлена $f(x)$ над полем Δ. Якщо $f(x) = x^2 + px + q^2$, то $\alpha = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$ і будь-яке чис-

ло з поля Δ_1 має вигляд (наслідок теореми 1 п. 32.2)

$$a + bd = a + b\left(-\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}\right) \qquad (a, b, p, q \in \Delta),$$

тобто виражається через числа поля Δ за допомогою раціональних дій і операції добування квадратного кореня. Отже, наше твердження буде доведене, якщо ми покажемо, що для довільних побудовних комплексних чисел $\mu = a_1 + ia_2$, $\nu = b_1 + ib_2$ побудовні також числа $\xi = \mu \pm \nu$, $\eta = \mu\nu$, $\zeta = \frac{\mu}{\nu}$ і $\omega = \sqrt{\mu}$. Побудовність чисел μ і ν означає побудовність їх компонент a_1, a_2, b_1, b_2. Компоненти чисел ξ, η, ζ, ω виражаються, як відомо, дійсними числами. Якщо $\xi = \xi_1 + i\xi_2$, $\eta = \eta_1 + i\eta_2$, $\zeta = \zeta_1 + i\zeta_2$, $\omega = \omega_1 + i\omega_2$, то $\xi_1 = a_1 \pm b_1$, $\eta_1 = a_1b_1 - a_2b_2$, $\zeta_1 = \frac{a_1b_1 + a_2b_2}{b_1^2 + b_2^2}$, $\omega_1 = \sqrt{\frac{\sqrt{a_1^2 + a_2^2} + a_1}{2}}$, $\xi_2 = a_2 \pm b_2$; $\eta_2 = a_1b_2 + a_2b_1$, $\zeta_2 = \frac{a_2b_1 - a_1b_2}{b_1^2 + b_2^2}$, $\omega_2 = \sqrt{\frac{\sqrt{a_1^2 + a_2^2} - a_1}{2}}$.

Очевидно, всі ці числа є побудовними, якщо побудовні a_1, a_2, b_1, b_2.

Тепер ми можемо довести основну теорему, яка встановлює рівносильність побудовності числа з можливістю виразити його у квадратних радикалах.

Теорема 3. *Для того щоб число* ζ (дійсне або комплексне) *було побудовним циркулем і лінійкою, виходячи з поля* $\Delta \subseteq R$, *необхідно і достатньо, щоб* ζ *виражалось у квадратних радикалах через числа поля* Δ.

Доведення. *Достатність.* Нехай ζ виражається у квадратних радикалах через числа поля Δ. Тоді, за теоремою 2 п. 34.3, існує ланцюжок числових полів $\Delta \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \ldots \subseteq \Delta_m$ такий, що $\zeta \in \Delta_m$ і кожне з полів ланцюжка є квадратичним розширенням попереднього. За щойно доведеною лемою усі числа з поля Δ_1 побудовні, виходячи з поля Δ. Усі числа з поля Δ_2 побудовні, виходячи з поля Δ_1 а тому і виходячи з поля Δ (адже при побудові можна користуватись будь-якою раніше побудованою точкою). Міркуючи далі в такий же спосіб, дістаємо, що всі числа з поля Δ_m, зокрема і ζ, побудовні, виходячи з поля Δ.

Необхідність. Нехай тепер відомо, що ζ — побудовне циркулем і лінійкою, виходячи з поля Δ. Оскільки ζ, взагалі кажучи,— комплексне число, вважатимемо $\zeta = \zeta_1 + i\zeta_2$; за означенням побудовності комплексного числа, ζ_1 і ζ_2 — побудовні числа.

Візьмемо число ζ_1 і виконуватимемо його побудову циркулем і лінійкою, виходячи з поля Δ.

Застосування лінійки не виведе нас за межі поля Δ. Застосовуючи один раз циркуль, ми побудуємо числа, які належать або полю Δ, або деякому квадратичному розширенню поля Δ (п. 35.2). Позначимо це розширення через Δ_1 і повторимо наше міркування, виходячи вже з поля Δ_1. Нове застосування лінійки не виведе за межі Δ_1, а застосування (одноразове) циркуля або не виведе з Δ_1, або приведе до чисел, що належать полю Δ_2, яке є квадратичним розширенням поля Δ_1. Міркуючи в такий спосіб, будуємо поля Δ_3, Δ_4, ..., Δ_j, ..., кожне з яких є квадратичним розширенням попереднього. Оскільки, за припущенням, ζ_1 побудовне, тобто для побудови його циркуль використовується скінченне число разів, то на певному кроці ми повинні прийти до поля Δ_k, що містить число ζ_1.

Отже, існує ланцюжок квадратичних розширень

$$\Delta \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \cdots \subseteq \Delta_k$$

такий, що $\zeta_1 \in \Delta_k$. Але тоді, за теоремою 2 п. 34.3, число ζ_1 *виражається у квадратних радикалах через числа поля* Δ.

Цілком аналогічні міркування показують, що і ζ_2 виражається у квадратних радикалах через числа поля Δ. Тепер зрозуміло, що таким є і число

$$\zeta = \zeta_1 + i\zeta_2 = \zeta_1 + \sqrt{-1}\,\zeta_2$$

(адже $-1 \in \Delta$). Теорему доведено.

При аналізі можливості розв'язувати ту чи іншу конструктивну задачу за допомогою циркуля і лінійки ми тепер можемо застосувати усі ознаки, встановлені у § 34 при з'ясуванні питання про можливість виразити число у квадратних радикалах через числа заданого поля.

35.4. Класичні задачі. Близько двох з половиною тисяч років тому геометри Греції виявили деякі задачі, які неможливо було розв'язати за допомогою циркуля і лінійки. Найбільш відомими серед них є проблеми *подвоєння куба, трисекції кута і квадратури круга.* Цими задачами грецькі геометри зацікавились і усвідомили їх трудність досить рано, як тільки було в достатній мірі з'ясоване поняття конструктивної задачі. У V ст. до н. е. ми зустрічаємо вже цілий ряд спроб розв'язати згадані задачі. Оскільки циркулем і лінійкою (тобто за допомогою проведення кіл і прямих ліній) виконати потрібні побудови не вдалося, грецькі математики використали деякі інші методи (наприклад, метод «вставки») та деякі інші криві (конічні перерізи, квадратрису, конхоїду тощо), частину з яких спеціально було введено для розв'язування даних задач.

Проте питання про розв'язання цих проблем лише за допомогою кіл і прямих, тобто, як тоді вважалося, єдиним справжньо геометричним методом, продовжувало в усі часи привертати увагу математиків. Елементарність постановки зазначених задач, поєднана з їх загадковою трудністю, спричинила надзвичайну популярність цих проблем і серед нематематиків: навіть у наші дні зустрічаються спроби розв'язати циркулем і лінійкою задачі квадратури круга або трисекції кута, хоч у XIX ст. було доведено неможливість такої побудови. Ця обставина ще раз підкреслює важливість для вчителя математики середньої школи бути обізнаним з науковою постановкою і розв'язанням зазначених проблем.

У цьому пункті, спираючись на встановлені критерії побудовності чисел циркулем і лінійкою, буде показано, що згадані класичні задачі не можна розв'язати за допомогою цих інструментів.

Подвоєння куба. Задача полягає *у побудові куба, об'єм якого вдвоє більший за об'єм даного куба.*

Якщо позначити довжину ребра даного куба через a, а довжину ребра куба вдвоє більшого об'єму через x, то матимемо співвідношення $x^3 = 2a^3$. Узявши $a = 1$, ми, очевидно, зведемо нашу задачу до побудови числа $\sqrt[3]{2}$, виходячи з поля $\mathbf{Q}$ раціональних чисел.

Але $\sqrt[3]{2}$ є коренем многочлена $f(x) = x^3 - 2$ над полем $\mathbf{Q}$. Многочлен $f(x)$, як легко перевірити, не має раціональних коренів і тому є незвідним у полі $\mathbf{Q}$.

Таким чином, мова йде про побудову кореня незвідного у полі $\mathbf{Q}$ многочлена третього степеня, виходячи з поля $\mathbf{Q}$. Відповідно до на-

слідку з теореми 4 п. 34.4 такий корінь не виражається у квадратних радикалах через числа поля **Q** і тому побудову його виконати циркулем і лінійкою неможливо.

Отже, *задача подвоєння куба не може бути розв'язана циркулем і лінійкою.*

Строге доведення[1] неможливості подвоєння куба циркулем і лінійкою було знайдене лише в 1837 р. після того, як близько 23-х століть тривали спроби розв'язати цю проблему.

Трисекція кута. Ця задача, яка полягає *у поділі довільного кута на три рівні частини,* є безпосереднім узагальненням елементарної задачі про поділ кута на дві рівні частини. Але в той час, як остання побудова легко виконується циркулем і лінійкою,— трисекцію довільного кута виконати цими інструментами вже неможливо. З'ясуємо причину цього факту.

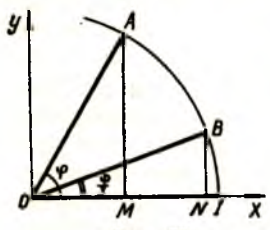


Рис. 72

Нехай φ — заданий кут (рис. 7). Проведемо коло одиничного радіуса з центром у вершині O цього кута і виберемо систему координат, початок якої збігається з точкою O, а вісь абсцис — з однією із сторін кута. Точку A, а отже — відрізок $OM = \cos\varphi$ можна вважати заданими разом з кутом φ. Задача буде розв'язана, якщо ми побудуємо точку B або відрізок $ON = \cos\frac{\varphi}{3}$. Таким чином, вихідним полем Δ у цій задачі є мінімальне поле, що містить 1 і $\cos\varphi = a$, тобто або поле **Q** (якщо a раціональне), або поле $\mathbf{Q}(a)$ (якщо a ірраціональне). Виходячи з цього поля, слід побудувати число $x_0 = \cos\frac{\varphi}{3}$.

Оскільки $\cos\varphi = 4\cos^3\frac{\varphi}{3} - 3\cos\frac{\varphi}{3}$, то $a = 4x_0^3 - 3x_0$, тобто x_0 є коренем многочлена

$$f(x) = 4x^3 - 3x - a \qquad (1)$$

над полем Δ.

На основі теореми 5 п. 34.5 ми можемо тепер твердити, що задачу трисекції кута можна розв'язати циркулем і лінійкою тоді і тільки тоді, коли многочлен (1) звідний у полі Δ.

Це показує, що здійсненність чи нездійсненність трисекції кута φ циркулем і лінійкою залежить від числового значення параметра $a = \cos\varphi$. Можна навести скільки завгодно прикладів таких значень φ, а отже, і значень $a = \cos\varphi$, при яких многочлен $f(x) = 4x^3 - 3x - a$ незвідний у полі Δ, а також приклади таких φ, для яких многочлен $f(x)$ звідний і трисекція можлива.

Приклади. **1.** При $\varphi = 60°$, $a = \cos 60° = \frac{1}{2}$, $\Delta = \mathbf{Q}$, $f(x) = 4x^3 - 3x - \frac{1}{2}$. Цей многочлен незвідний у полі раціональних чисел, бо, як легко перевірити, не має раціональних коренів. Отже, трисекція кута $\varphi = 60°$ циркулем і лінійкою неможлива.

[1] Це доведення дав П. Вантцель (1814—1848), який встановив нерозв'язність у квадратних радикалах незвідного кубічного рівняння.

2. При $\varphi = 90°$, $a = \cos 90° = 0$, $\Delta = \mathbf{Q}$, $f(x) = 4x^3 - 3x = x(4x^2 - 3)$, тобто многочлен $f(x)$ звідний у полі **Q**. Отже, трисекцію кута $\varphi = 90°$ можна виконати циркулем і лінійкою. Читачеві, звичайно, відомо, як цими інструментами побудувати кут в 30°.

3. При $\varphi = 45°$, $a = \frac{\sqrt{2}}{2}$, $\Delta = \mathbf{Q}(\sqrt{2})$, $f(x) = 4x^3 - 3x - \frac{\sqrt{2}}{2}$. Цей многочлен має корінь $-\frac{\sqrt{2}}{2}$ і тому звідний у полі $\mathbf{Q}(\sqrt{2})$. Отже, і кут 45° можна поділити циркулем і лінійкою на три рівні частини.

4. Як вправу рекомендуємо читачеві довести неможливість трисекції кута $\varphi = 30°$. У цьому випадку $a = \frac{\sqrt{3}}{2}$. Тому задача зводиться до встановлення незвідності многочлена $4x^3 - 3x - \frac{\sqrt{3}}{2}$ у полі чисел виду $a + b\sqrt{3}$ (a, b — раціональні).

Оскільки не для довільного φ можлива побудова кута $\frac{\varphi}{3}$, то *задача трисекції кута у наведеному вище формулюванні не може бути розв'язана циркулем і лінійкою.*

Квадратура круга. Задача полягає *у побудові квадрата, рівновеликого даному кругу.* Ця задача була найбільш відомою і разом з тим найбільш важкою з усіх класичних проблем.

Якщо взяти радіус даного круга за одиницю, то його площа, як відомо, дорівнює π. Отже, квадратура круга зводиться до побудови квадрата з стороною $\sqrt{\pi}$, якщо дано одиничний відрізок. Інакше кажучи, питання про можливість розв'язати проблему квадратури круга циркулем і лінійкою — це питання про побудовність числа $\sqrt{\pi}$, виходячи з поля **Q** раціональних чисел. Оскільки побудовність числа $\sqrt{\pi}$ рівнозначна побудовності числа π, то вся проблема зводиться до дослідження природи числа π.

Ми знаємо тепер, що π — *трансцендентне число,* тобто не може бути коренем жодного алгебраїчного рівняння з раціональними коефіцієнтами. Це й означає, на основі теореми 3.п.34.4, що воно не виражається у квадратних радикалах через раціональні числа і тому не може бути побудоване циркулем і лінійкою, виходячи з поля **Q**, тобто, що задача квадратури круга не може бути розв'язана циркулем і лінійкою.

Факт трансцендентності числа π встановив тільки в 1882 р. німецький математик Ліндеман. Цим завершився багатовіковий період спроб розв'язати проблему квадратури круга, а також численні дослідження арифметичної природи числа π.

35.5. Поділ кола (побудова правильних многокутників). Завдання поділу кола *полягає у побудові циркулем і лінійкою правильного n-кутника при довільному натуральному n.* Зрозуміло, що цей многокутник завжди можна уявляти собі вписаним у коло одиничного радіуса.

Уже в стародавній Греції було помічено, що в той час як деякі n-кутники (наприклад, при $n = 3, 4, 5, 6, 8, 10, 12$) можна побудувати циркулем і лінійкою досить легко, при інших значеннях n (7, 9, 11, 13 і ін.) це не вдається.

Нам тепер цілком зрозуміло, що не всякий правильний многокутник можна побудувати циркулем і лінійкою. Так, не можна побудувати правильний 9-кутник. Справа в тому, що нам задано в цій задачі лише радіус, тобто числа 0 і 1, що рівно-

значно заданню поля раціональних чисел. Але, як було показано вище, виходячи з поля $\mathbf{Q}$, не можна здійснити трисекцію кута в 60°, або, що те саме, побудувати циркулем і лінійкою кут у 20°. Побудовність же 9-кутника означала б і побудовність 18-кутника, тобто можливість побудувати циркулем і лінійкою кут в $\frac{360°}{18} = 20°$, що суперечить доведеному.

Ще в стародавні часи постало питання про те, які правильні многокутники можна побудувати циркулем і лінійкою, а які — ні. Повну відповідь на це питання дав видатний німецький математик К. Ф. Гаусс в кінці XVIII ст.

Переходячи до викладу результатів Гаусса, зробимо деякі попередні зауваження.

1) *Якщо якийсь правильний n-кутник побудовний циркулем і лінійкою, то таким буде і всякий многокутник з числом сторін $2^k \cdot n$.*

Це випливає з того, що, як відомо з шкільного курсу математики, довільну дану дугу завжди можна поділити на дві рівні частини за допомогою циркуля і лінійки. Наприклад, разом з правильним трикутником, побудовними є правильні многокутники з числом сторін 6, 12, 24, 48, ..., $2^k \cdot 3$.

2) *Якщо побудовний правильний многокутник з числом сторін $n = n_1 \cdot n_2$, то побудовними є і многокутники з числом сторін n_1 та n_2.*

Справді, можливість побудови n-кутника означає можливість побудови циркулем і лінійкою дуги, що є $\frac{1}{n}$ частиною одиничного кола. Але тоді побудовна і $\frac{1}{n_1}$ частина кола $\left(\text{бо } \frac{1}{n_1} = \frac{1}{n} \cdot n_2\right)$, і $\frac{1}{n_1}$ частина кола $\left(\text{бо } \frac{1}{n_2} = \frac{1}{n} \cdot n_1\right)$.

3) *Якщо побудовні многокутники з числом сторін n_1 та n_2, причому числа n_1, n_2 взаємно прості, то побудовний і многокутник з числом сторін $n = n_1 \cdot n_2$.*

Справді, як відомо (див. п. 14.2), можна знайти такі натуральні числа m_1 і m_2, що $n_1 m_1 - n_2 m_2 = 1$, бо n_1 і n_2 взаємно прості. Але тоді

$$\frac{1}{n_1 n_2} = m_1 \cdot \frac{1}{n_2} - m_2 \cdot \frac{1}{n_1}.$$

Ця рівність показує, що з побудовності $\frac{1}{n_1}$ частини кола і $\frac{1}{n_2}$ частини кола випливає побудовність $\frac{1}{n_1 n_2} = \frac{1}{n}$ частини кола, тобто правильного n-кутника.

Нагадаємо тепер, що довільне натуральне число n можна єдиним способом подати у вигляді

$$n = 2^k \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_r^{k_r},$$

де $p_1, \ldots, p_r$ — *різні* прості дільники числа n, причому $p_i \geqslant 3$.

На підставі зауважень 1) — 3) легко встановити справедливість такого твердження.

Теорема 4. *Для того щоб правильний многокутник з числом сторін*

$$n = 2^k \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_r^{k_r}$$

був побудовний циркулем і лінійкою, необхідно і достатньо, щоб були побудовані многокутники з числом сторін $p_1^{k_1}, p_2^{k_2}, \ldots, p_r^{k_r}$.

Справді, якщо n-кутник побудовний, то на підставі зауваження 2) такими є і правильні многокутники з числом сторін $p_1^{k_1}, p_2^{k_2}, \ldots, p_r^{k_r}$. Навпаки, якщо побудовні останні многокутники, то на підставі зауваження 3) побудовним є многокутник з числом сторін $n_1 = p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_r^{k_r}$ (бо всі числа $p_1^{k_1}, p_2^{k_2}, \ldots p_r^{k_r}$ взаємно прості), а на підставі зауваження 1) — многокутник з числом сторін $n = 2^k \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \times \times p_r^{k_r}$.

Тепер зрозуміло, що досить розглянути питання про побудовність многокутника, число сторін якого є простим або степенем простого числа.

Теорема 5. *Якщо $p \geqslant 3$ — просте число, то правильний p-кутник побудовний циркулем і лінійкою тоді і тільки тоді, коли p має вигляд $p = 2^m + 1$.*

Доведення. Для побудовності правильного p-кутника, вписаного в одиничне коло, необхідно і достатньо, щоб було побудовним (виходячи з поля $\mathbf{Q}$) комплексне число [1]

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

Справді, якщо це число побудовне, то відомі дві сусідні вершини правильного p-кутника (1 та ε_1), тобто $\frac{1}{p}$ частина кола. З другого боку, якщо p-кутник побудовний, то, беручи його центр за початок координат, а одну з вершин — за точку (1, 0), дістанемо, що сусідня (при русі проти годинникової стрілки) вершина зображатиме комплексне число ε_1.

Тому розглянемо питання про побудовність кореня ε_1 двочленного рівняння

$$z^p - 1 = 0. \tag{2}$$

До рівняння (2) незручно застосовувати результати § 34, бо многочлен $z^p - 1$, очевидно, звідний у полі раціональних чисел

$$z^p - 1 = (z-1)(z^{p-1} + z^{p-2} + \ldots + z + 1).$$

Зрозуміло, що ζ_1 є коренем многочлена

$$f(z) = z^{p-1} + z^{p-2} + \ldots + z + 1 \tag{3}$$

над полем $\mathbf{Q}$. Покажемо, що многочлен (3) вже незвідний у полі $\mathbf{Q}$.

Для цього перетворимо цей многочлен так, щоб до нього можна було застосувати критерій Ейзенштейна (п. 31.1). Беручи $z = u + 1$, дістанемо

$$f(z) = z^{p-1} + \ldots + z + 1 = \frac{z^p - 1}{z - 1} = \frac{(u+1)^p - 1}{u} =$$

$$= u^{p-1} + pu^{p-2} + \frac{p(p-1)}{2}u^{p-3} + \ldots + p.$$

Многочлен $f_1(u) = u^{p-1} + pu^{p-2} + \ldots + p$ має вільний член p. Кожний коефіцієнт його є цілим числом, причому всі коефіцієнти, крім старшого, діляться на p. Справді, коефіцієнт при u^{p-k-1} дорівнює числу C_p^k комбінацій з p елементів по k; це число можна подати у вигляді добутку $p \cdot \frac{(p-1)(p-2)\ldots(p-k+1)}{k!}$. Оскільки цей добуток є цілим числом, причому просте число p взаємно просте з $k!$ (бо $k < p$), то і

$$\frac{(p-1)(p-2)\ldots(p-k+1)}{k!} = q$$

є цілим числом. Отже, розглядуваний коефіцієнт має вигляд pq, тобто ділиться на p.

Як бачимо, многочлен $f_1(u)$ задовольняє всі умови критерію Ейзенштейна і тому є незвідним у полі $\mathbf{Q}$. Але тоді і даний многочлен $f(z)$ незвідний у цьому полі.

На підставі теореми 4 п. 34.4 і теореми 3 п. 35.3, для побудовності числа ε_1 н е о б х і д н о, щоб степінь незвідного многочлена (5) був числом виду 2^m, тобто щоб $p - 1 = 2^m$, $p = 2^m + 1$.

---

[1] Нагадаємо, що корені p-го степеня з 1, або, що те саме, розв'язки двочленного рівняння $z^p - 1 = 0$, є комплексні числа $\varepsilon_k = \cos\frac{2\pi k}{p} + i\sin\frac{2\pi k}{p}$ ($k = 0, 1, \ldots, p-1$), які геометрично зображуються точками комплексної площини, розташованими у вершинах правильного p-кутника, вписаного в коло одиничного радіуса (див. ч. 1, § 16).

Покажемо тепер і д о с т а т н і с т ь цієї умови для побудовності ε_1, спираючись на теорему 6 п. 34.6. Розглянемо для цього поле розкладу Ω многочлена $f(z)$. Легко бачити, що в даному окремому випадку поле Ω збігається з мінімальним полем $\mathbf{Q}(\varepsilon_1)$. Адже всі корені $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{p-1}$ многочлена (3), тобто всі корені многочлена $z^p - 1$ (крім 1) можна дістати раціонально з ε_1:

$$\varepsilon_2 = \varepsilon_1^2, \quad \varepsilon_3 = \varepsilon_1^3, \quad \ldots, \quad \varepsilon_{p-1} = \varepsilon_1^{p-1},$$

тобто вони належать полю $\mathbf{Q}(\varepsilon_1)$. Таким чином, поле розкладу Ω, утворене приєднанням до поля $\mathbf{Q}$ всіх коренів многочлена (3), збігається з полем, утвореним приєднанням лише одного кореня ε_1. Оскільки, за теоремою 1 п. 33.2, $(\mathbf{Q}(\varepsilon_1) : \mathbf{Q}) = p - 1$, то й $(\Omega : \mathbf{Q}) = p - 1$. Тому при $p = 2^m + 1$ матимемо $(\Omega : \mathbf{Q}) = 2^m$, і, за теоремою 6 п. 34.6, число ε_1 виражається у квадратних радикалах, тому побудовне. Теорему доведено.

На основі теореми 5 можна з'ясувати побудовність чи непобудовність циркулем і лінійкою довільного p-кутника, коли p — просте число. Наприклад, многокутники з числом сторін $3 = 2 + 1$, $5 = 2^2 + 1$, $17 = 2^4 + 1$, $257 = 2^8 + 1$ побудовні циркулем і лінійкою, тоді як при $p = 7, 11, 13, 19$ правильний p-кутник побудувати цими засобами неможливо. Отже, дана конструктивна проблема зведена до теоретико-числової задачі — знаходження всіх простих чисел виду $2^m + 1$ (див. п. 8.1).

Залишається розглянути випадок, коли n є степенем простого числа.

Теорема 6. *Якщо $p \geqslant 3$ — просте число, а $k \geqslant 2$ — довільне натуральне число, то правильний n-кутник з числом сторін $n = p^k$ не можна побудувати циркулем і лінійкою.*

Д о в е д е н н я. Досить показати, що при $k = 2$ побудова неможлива. Адже при $k > 2$ $n = p^2 \cdot p^{k-2}$ і побудовність p^k-кутника означала б побудовність p^2-кутника.

Як відомо, для побудовності p^2-кутника необхідно і достатньо, щоб був побудовним перший корінь η_1 многочлена

$$\varphi(z) = z^{p^2} - 1,$$

тобто число $\eta_1 = \cos\dfrac{2\pi}{p^2} + i \sin\dfrac{2\pi}{p^2}$.

Многочлен $\varphi(z)$ звідний у полі $\mathbf{Q}$: він ділиться, наприклад, на $\omega(z) = z^p - 1$. Оскільки число η_1 не є коренем многочлена $\omega(z)$ [бо $\omega(\eta_1) = \cos\dfrac{2\pi}{p} + i \sin\dfrac{2\pi}{p} - 1 \neq 0$], то воно є коренем частки

$$\psi(z) = \frac{\varphi(z)}{\omega(z)} = \frac{z^{p^2} - 1}{z^p - 1} = z^{p(p-1)} + z^{p(p-2)} + \cdots + z^p + 1. \qquad (4)$$

Покажемо, що $\psi(z)$ — незвідний многочлен у полі $\mathbf{Q}$ аналогічно тому, як це було зроблено для многочлена (3).

Зауважимо спочатку, що при діленні многочлена на інший многочлен із старшим коефіцієнтом 1 всі коефіцієнти частки визначаються через коефіцієнти діленого і дільника за допомогою дій віднімання і множення. Тому, якщо всі коефіцієнти діленого і дільника (крім старших) діляться на якесь просте число p, то й усі коефіцієнти частки (крім старшого) ділитимуться на p.

Зробимо тепер у многочлені (4) заміну $z = u + 1$. Дістанемо:

$$\psi(z) = \psi_1(u) = \frac{\varphi(u+1)}{\omega(u+2)} = \frac{(u+1)^{p^2} - 1}{(u+1)^p - 1}. \qquad (4)$$

Старший коефіцієнт многочлена $\omega(u+1) = (u+1)^p - 1$ дорівнює одиниці. Решта його коефіцієнтів, як було показано при доведенні теореми 7, діляться на p. Цілком аналогічно у многочлена $\varphi(u+1) = (u+1)^{p^2} - 1$ старший коефіцієнт дорівнює одиниці, а решта коефіцієнтів діляться[1] на p. На основі зробленого вище зауваження приходимо до висновку, що всі коефіцієнти частки $\psi_1(u)$, крім старшого, діляться на p.

Користуючись поданням (4), подамо $\psi_1(u)$ у формі

$$\psi_1(u) = (u+1)^{p(p-1)} + (u+1)^{p(p-2)} + \cdots + (u+1)^p + 1. \qquad (5)$$

Звідси бачимо, що старший коефіцієнт многочлена $\psi_1(u)$ дорівнює 1, а вільний член дорівнює p (числу членів суми (5)).

Отже, всі коефіцієнти многочлена $\psi_1(u)$, крім старшого, діляться на p, вільний член при цьому не ділиться на p^2, тому за критерієм Ейзенштейна $\psi_1(u)$, а тому й $\psi(z)$ незвідні у полі раціональних чисел.

Відповідно до теореми 4 п. 34.4 для побудовності η_1 необхідно, щоб степінь многочлена $\psi(z)$, тобто $p(p-1)$, задовольняв умові $p(p-1) = 2^m$.

Проте ця умова не може справджуватись для жодного простого числа $p \geqslant 3$. Цим наша теорема доведена повністю.

Об'єднуючи твердження, встановлені у теоремах 4—6, ми можемо тепер сформулювати остаточний результат.

Теорема 7. *Для того щоб правильний n-кутник міг бути побудований циркулем і лінійкою, необхідно і достатньо, щоб n було числом виду $n = 2^k \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_r$, де всі $p_1, p_2, \ldots, p_r$ — різні прості числа виду $p_j = 2^{m_j} + 1$.*

Це твердження дає змогу для довільного n відразу ж з'ясувати, чи побудовний правильний n-кутник. Єдина трудність, яка може виникнути при цьому,— це теоретико-числова трудність розкладу великих чисел на прості множники.

---

[1] Коефіцієнтом при u^{p^2-k} є ціле число $\dfrac{p^2(p^2-1)\ldots(p^2-k+1)}{k!}$. Тут p^2 може не бути взаємно простим з $k!$, бо k може бути більшим за p. Але p^2 і $k!$ можуть мати не більше одного спільного множника p (бо $k < p^2$), тому цей коефіцієнт, в усякому разі, ділиться на p.

ЛІТЕРАТУРА

Александров П. С. Вступ до теорії груп. К., «Радянська школа», 1955.
Артін Е. Теорія Галуа. К., «Радянська школа», 1963.
Бородін О. І. Теорія чисел, вид. 3. К., «Вища школа», 1970.
Бурбаки Н. Алгебраические структуры. Линейная и полилинейная алгебра. М., Физматгиз, 1962.
Бурбаки Н. Алгебра. Многочлены и поля. Упорядоченные группы. М., «Наука», 1965.
Бухштаб А. А. Теория чисел. М., «Просвещение», 1966.
Ван дер Варден Б. Л. Современная алгебра, ч. I, II. М., Гостехиздат, 1947.
Виноградов И. М. Основы теории чисел. М., «Наука», 1965.
Гейл Д. Теория линейных экономических моделей. М., Изд-во иностр. лит., 1963.
Дрозд Ю. А., Кириченко В. В. Высшая алгебра (методическая разработка), ч. II. К., Изд-во Киев. ун-та, 1973.
Дэвенпорт Г., Высшая арифметика. М., «Наука», 1965.
Зарисский О., Самюэль П. Коммутативная алгебра, т. I. М., Изд-во иностр. лит., 1963.
Зуховицкий С. И., Авдеева Л. И. Линейное и выпуклое программирование. М., «Наука», 1967.
Калужнин Л. А. Введение в общую алгебру. М., «Наука», 1973.
Костарчук В. М., Хацет Б. І. Курс вищої алгебри, вид. 3. К., «Вища школа», 1969.
Костарчук В. М., Хацет Б. І. Про можливе і неможливе в геометрії циркуля і лінійки. К., «Радянська школа», 1971.
Курант Р., Роббинс Г. Что такое математика. М., «Просвещение», 1967.
Курош А. Г. Курс высшей алгебры. М., «Наука», 1971.
Курош А. Г. Лекции по общей алгебре. М., «Наука», 1962.
Курош А. Г. Теория групп. М., «Наука», 1967.
Ленг С. Алгебра. М., «Мир», 1968.
Ляпин Е. С. Курс высшей алгебры. М., Учпедгиз, 1955.
Ляпин Е. С., Евсеев А. Е. Алгебра и теория чисел, ч. I. М., «Просвещение», 1974.
Мальцев А. И. Алгебраические системы. М., «Наука», 1970.
Маркушевич А. И. Деление с остатком в арифметике и алгебре. М.—Л., АПН РСФСР, 1949.
Окунев Л. Я. Высшая алгебра. М., «Просвещение», 1966.
Окунев Л. Я. Основы современной алгебры. М., Учпедгиз, 1941.
Окунев Л. Я. Сборник задач по высшей алгебре. М., «Просвещение», 1964.
Постников М. М. Теория Галуа. М., Физматгиз, 1963.
Солодовников А. С. Введение в линейную алгебру и линейное программирование. М., «Просвещение», 1966.
Фаддеєв Д. К., Сомінський І. С. Збірник задач з вищої алгебри. К., «Вища школа», 1971.
Хинчин А. Я. Цепные дроби, изд. 3. М., Физматгиз, 1961.
Черников С. Н. Линейные неравенства. М., «Наука», 1968.
Энциклопедия элементарной математики, кн. II. М.—Л., ГИТТЛ, 1951; кн. IV, Физматгиз, 1963.
Юшкевич А. П. История математики в России. М., «Наука», 1968.

ЗМІСТ

Сергей Трофимович Завало
Виктор Николаевич Костарчук
Борис Исаакович Хацет

Алгебра и теория чисел

Часть вторая

Допущено Министерством просвещения УССР в качестве учебника для студентов физико-математических факультетов педагогических институтов

(На украинском языке)

Издательское объединение «Вища школа»
Головное издательство
Киев — 1976

Редактор О. С. Дзюба
Літредактор О. П. Ковальчук
Обкладинка художника Г. М. Балюна
Художній редактор І. Р. Ойхман
Технічний редактор Л. Ф. Волкова
Коректор І. Б. Мілевська

Здано до набору 17.02. 1976 р. Підписано до друку 14.09. 1976 р. Формат паперу 60 x 90¹/₁₆. Папір друк. № 1. Друк. арк. 24. Обл.-видавн. арк. 25,0. Тираж 8000. Видавн. № 2705. БФ 15893. Ціна 1 крб. 02 коп. Зам. № 6-405.

Головне видавництво видавничого об'єднання «Вища школа», 252054, Київ, 54, Гоголівська, 7.

Надруковано з матриць Головного підприємства республіканського виробничого об'єднання «Поліграфкнига» Держкомвидаву УРСР, м. Київ, Довженка, 3 на Харківській книжковій фабриці «Комуніст» республіканського виробничого об'єднання «Поліграфкнига» Держкомвидаву УРСР, Харків, Енгельса, 11.

1 крб. 02 коп.