Н. ОБРЕШКОВ

ВИСША АЛГЕБРА

ДЪРЖАВНО ИЗДАТЕЛСТВО „НАУКА И ИЗКУСТВО"

Проф. д-р Н. ОБРЕШКОВ

ВИСША АЛГЕБРА

ПЕТО ИЗДАНИЕ

ДЪРЖАВНО ИЗДАТЕЛСТВО „НАУКА И ИЗКУСТВО“
СОФИЯ — 1962

Учебникът съдържа материала по Висша алгебра, който е застъпен в лекциите по тази математическа дисциплина във Физико-математическия факултет на Софийския университет.

В него освен класическата висша алгебра е дадено цялостно изложение на линейната алгебра, която обикновено у други автори е обект на отделна книга, както и важните отдели на съвременната алгебра — теория на групите, теория на пръстените и полетата, теория на матриците и пр.

Учебникът е предназначен за студентите по математика и физика от двата профила при Физико-математическия факултет и също така за всички, които изучават тази основна математическа дисциплина.

Редактор Зл. Петрова

Художник М. Кънева
Худ. редактор Ем. Рашков
Техн. редактор С. Бояджийски
Коректори: М. Илчева
Л. Михайлова

Дадена за набор на 9. I. 1962 г.
Печатни коли 35
Формат 16/65/95
Кн.тяло 1,75 лв.

Изд. №11244/I-4
Тем. № 571

Цена 1,98 лв.

Подписана за печат на 25. IX. 1962 г.
Издателски коли 35
Тираж 4071
Подвързия 0,23 лв.

Държавна печатница „Г. Димитров" — София. Пор. № 559

СЪДЪРЖАНИЕ

Част I

Основни свойства на полиномите

Глава I. Увод

Глава II. Комплексни числа

Глава III. Свойства на полиномите

Част II

Теория на детерминантите и приложения

Глава I. Основни свойства на детерминантите

Глава II. Системи линейни уравнения

Глава III. Други свойства

Част III

Линейна алгебра

Част VIII

Абелеви и биномни уравнения

Глава I. Абелеви уравнения.

Глава II. Алгебрично решение на биномните уравнения

Глава III. Алгебрична нерешимост на уравненията от степен, по-висока от четири

Глава IV. Неразложимият случай при кубичното уравнение

Част IX

Приложение на теорията на групите за алгебричното решаване на уравненията.

Глава I. Субституции

Глава II. Полиноми, принадлежащи на една група от субституции

Глава III. Теория на Галоа

Глава IV. Приложение на теорията на Галоа

ЧАСТ I

ОСНОВНИ СВОЙСТВА НА ПОЛИНОМИТЕ

Глава I

Увод

1. **Променливи и функции.** Отличаваме променливи и постоянни числа. Едно число се нарича постоянно, ако то има една определена стойност, а променливо, ако може да вземе различни стойности. Постоянните числа или количества бележим с началните букви на латинската азбука, a, b, c,..., а променливите — с последните, x, y, z,... Ако на всяка стойност, която може да взема x, отговаря по някакъв начин една определена стойност на y, то казваме, че y е функция на x. Функциите бележим така:

$$f(x),\ \varphi(x),\ F(x),\ldots$$

По аналогия можем да имаме функции на повече променливи:

$$f(x),\ \varphi(x, y),\ldots,\ f(x, y, z),\ldots$$

2. **Рационални функции.** Ако в една функция променливите са подложени на тъй наречените рационални действия: събиране, изваждане, умножение и деление, извършени в краен брой, то тя се нарича рационална. Ако в рационалната функция не е извършено деление върху променливите, то тя се казва цяла рационална или полином. Ако рационалната функция не е цяла, тя се нарича дробна рационална. Следователно един полином представлява сума от членове, в които фигурират произведения от променливи с постоянни числа, наречени коефициенти. Най-високата степен, която се среща в членовете, се нарича степен на полинома. Общата форма на полинома от степен n с едно променливо ще бъде

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n,$$

а общата форма на дробна рационална функция на едно променливо ще бъде

$$\frac{a_0x^k+a_1x^{k-1}+\ldots+a_k}{b_0x^m+b_1{}^{m-1}+\ldots+b_m}.$$

Полиномите от първа степен се наричат линейни функции, от втора степен — квадратни и т. н. Така

$$a+bx+cy$$

е линейна функция, а

$$a+bx+cy+dx^2+exy+fy^2$$

е квадратна. Общо един полином от променливите $x_1, x_2, \ldots, x_m$ ще има вида

$$f(x_1, x_2, \ldots, x_m)=\sum cx_1^{\alpha_1} x_2^{\alpha_2} \ldots x_m^{\alpha_m},$$

гдето $\alpha_1, \alpha_2, \ldots, \alpha_m$ са цели положителни числа, някои от които могат да бъдат нули. Степента на написания член е

$$\alpha_1+\alpha_2+ \ldots +\alpha_m,$$

а c е коефициентът му. Най-голямото от числата

$$(\alpha_1+\alpha_2+ \ldots +\alpha_m)$$

е, както поменахме, степента на полинома.

Ако всички членове на полинома имат една и съща степен, то той се нарича **хомогенен** или **форма**. Тогава очевидно, ако n е степента му, за всяко λ ще имаме

$$f(\lambda x_1, \lambda x_2, \ldots, \lambda x_m)=\lambda^n f(x_1, x_2, \ldots, x_m).$$

Ако формата е от първа степен, то тя се казва **линейна**, от втора — **квадратна**, от трета — **кубична** и т. н. Ако броят на променливите е две независимо от степента, то формата се казва **бинерна**, ако променливите са три — **тернерна** и т. н.

Ако в една функция върху променливите освен рационалните действия приложим и коренуване, то функцията се нарича **ирационална**. Така

$$a+b\sqrt{x}, \frac{x+\sqrt{y}}{1-x\sqrt{z}}$$

са ирационални функции.

3. **Уравнения.** Ако A и B са два израза, от които единият е само преработване на другия, то трябва $A=B$ за всички стойности на променливите, които влизат. Едно такова равенство се нарича **тъждество**, така например

$$(x-a)(x^2+ax+a^2)=x^3-a^3.$$

Обаче можем да търсим такива значения на x, за които един полином $f(x)$ да приема стойност нула. Тези стойности ще удовлетворяват на условието

$$f(x)=0,$$

което се нарича алгебрично уравнение. Степента на $f(x)$ се нарича **степен** на уравнението, а стойностите, които го удовлетворяват, се наричат **корени**. Общата форма на уравнението от n-та степен с едно неизвестно ще бъде

$$a_0x^n+a_1x^{n-1}+ \ldots +a_{n-1}x+a_n=0.$$

Числата $a_0, a_1, \ldots, a_n$ са коефициенти на уравнението. Ако в това уравнение коефициентите са полиноми на едно променливо t, то корените му ще бъдат функции на t, които се наричат в анализа алгебрични функции. Неалгебричните функции се наричат трансцендентни. В настоящия курс ще разбираме по-ограничено под алгебрична функция рационалните и ирационалните функции. Функциите e^x, $\log x$, $\sin x$ са трансцендентни.

Глава II

Комплексни числа

1. **Дефиниция.** Понеже каквото и да е реалното число x, квадратът му е положителен, то уравнението $x^2 = -1$ няма реално решение. Така се налага да се въведе една по-обща система от числа, които обаче притежават следните свойства: първо, тая система трябва да съдържа реалните числа, второ, законите, които са валидни за реалните числа, да са валидни и за числата от тази система и, трето, да има поне едно число x от системата, което удовлетворява на уравнението $x^2 = -1$.

Оказва се, че такава система съществува и се състои от тъй наречените комплексни числа, които имат голямо приложение. Под комплексно число разбираме един нареден чифт (a, b) реални числа, като някои операции, които можем да извършваме с тези числа, са дадени. Чифтът се нарича нареден, понеже числата (a, b) и (b, a) при $a \neq b$ са различни помежду си. Две комплексни числа (a, b), (a_1, b_1) само тогава считаме за равни, ако $a = a_1$, $b = b_1$. Числото $(0, 1)$ означаваме с i. За да изпълним първото условие, т. е. реалните числа да са частност от комплексните, ще считаме, че числото $(a, 0)$ е равно на реалното число a и (a, b) е само тогава нула, когато $a = b = 0$.

2. **Сума.** Под сума на комплексните числа

$$\alpha = (a, b), \quad \alpha_1 = (a_1, b_1)$$

разбираме числото $(a + a_1, b + b_1)$, което бележим с $\alpha + \alpha_1$. Ако $b = b_1 = 0$, т. е. α и α_1 са реални числа, то това води до сума на две реални числа, както трябва да бъде.

Оттук веднага се вижда валидността на следните основни закони, гдето α, β, γ са произволни комплексни числа:

$\alpha + \beta = \beta + \alpha$ — комутативен закон,

$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ — асоциативен закон,

$\alpha + 0 = \alpha$.

3. **Разлика.** Ако $\alpha = (a, b)$ и $\beta = (a_1, b_1)$ са две комплексни числа, то ще намерим комплексното число $x = (\zeta, \zeta_1)$, което удовлетворява на условието $\beta + x = \alpha$, т. е. като използуваме горната дефиниция на сума, ще имаме

$$a_1 + \zeta = a, \quad b_1 + \zeta_1 = b,$$

отгдето $\zeta=a-a_1$, $\zeta_1=b-b_1$. Обратно, веднага се вижда, че числото $x=(a-a_1, b-b_1)$ удовлетворява на условието $\beta+x=\alpha$. Това число се нарича разлика на числата α и β и се бележи с $\alpha-\beta$.

4. **Произведение.** Веднага можем да получим произведението на едно цяло положително число n с комплексното число $\alpha=(a, b)$. За това трябва да съберем n числа, равни на α, и според по-горното

$$n\,.\,(a, b)=(na, nb).$$

Ще постулираме това равенство за всяко реално число n. Тогава можем да пишем

$$\alpha=(a, b)=(a, 0)+(0, b)=a+b\,.\,(0, 1)=a+bi,$$

така че всяко комплексно число може да се представи така. Числото a се нарича реална част на α, а bi се нарича имагинерна част. Ако сега искаме да дефинираме произведение на две комплексни числа,

$$\alpha=(a, b), \quad \beta=(a_1, b_1),$$

то като предположим, че $i^2=(0, 1)^2=-1$ и пресметнем $(a+bi)(a_1+b_1 i)$, като вземем под внимание, че трябва основните закони да са в сила, получаваме

$$\begin{aligned}(a, b)(a_1, b_1)&=(a+bi)(a_1+b_1 i)=\\ &=aa_1+ab_1 i+a_1 bi+bb_1(-1)=\\ &=aa_1-bb_1+i(ab_1+ba_1)=(aa_1-bb_1, ab_1+ba_1).\end{aligned}$$

Следователно правилото за умножаване гласи:

$$(1)\qquad (a, b)(a_1, b_1)=(aa_1-bb_1, ab_1+ba_1).$$

Обратно, ако за произведение вземем тази дефиниция, лесно се убеждаваме, че всички закони за реалните числа остават в сила и за комплексните. Така веднага имаме

$$ab=(a, 0)(b, 0)=(ab, 0)=ab,$$

$$i^2=(0, 1)(0, 1)=(-1, 0)=-1.$$

$\alpha\beta=\beta\alpha$ — **комутативен закон,**

$\alpha(\beta\gamma)=(\alpha\beta)\gamma$ — асоциативен закон,

$(\alpha+\beta)\gamma=\alpha\gamma+\beta\gamma$ — дистрибутивен закон,

$1\,.\,\alpha=\alpha,$

$0\,.\,\alpha=0.$

Тези резултати читателят ще провери лесно.

5. **Частно.** Нека $\alpha=(a, b)$, $\beta=(a_1, b_1)$ са две произволни числа. Да търсим число $x=(\zeta, \zeta_1)$, което удовлетворява на условието $\beta x=\alpha$. Ако $\beta=0$, то това е възможно само ако $\alpha=0$ и тогава всяко число x

удовлетворява на това условие. Като изключим този случай, считаме, че $\beta \neq 0$. Но тогава за ζ и ζ_1 получаваме следните уравнения:

$$a_1\zeta - b_1\zeta_1 = a, \quad a_1\zeta_1 + b_1\zeta = b,$$

които, като решим, дават

$$\zeta = \frac{aa_1 + bb_1}{a_1^2 + b_1^2}, \quad \zeta_1 = \frac{a_1 b - b_1 a}{a_1^2 + b_1^2}.$$

Обратно, веднага се вижда, че числото

$$x = \left(\frac{aa_1 + bb_1}{a_1^2 + b_1^2}, \quad \frac{a_1 b - b_1 a}{a_1^2 + b_1^2} \right)$$

удовлетворява на условието $\beta x = \alpha$. Това число наричаме частно на α и β и бележим с $\frac{\alpha}{\beta}$. Горната формула лесно бихме получили директно със следното умножение:

$$\frac{\alpha}{\beta} = \frac{a+bi}{a_1+b_1 i} = \frac{(a+bi)(a_1-b_1 i)}{a_1^2+b_1^2} = \frac{aa_1+bb_1}{a_1^2+b_1^2} + \frac{ba_1-ab_1}{a_1^2+b_1^2} i.$$

6. **Конюговани комплексни числа.** Двете комплексни числа $a+bi$ и $a-bi$ се наричат конюговани. Те се отличават само по знака на имагинерната им част. Ако означим $\alpha = a+bi$, то конюгованото означаваме така: $\bar{\alpha} = a-bi$. Сумата и произведението на две конюговани числа са реални, а разликата им е чисто имагинерно число, т. е. комплексно число, на което реалната част е равна на нула. Действително имаме

$$\alpha + \bar{\alpha} = a + bi + a - bi = 2a,$$

$$\alpha - \bar{\alpha} = a + bi - (a - bi) = 2bi,$$

$$\alpha\bar{\alpha} = (a+bi)(a-bi) = a^2 + b^2.$$

Лесно се проверяват и равенствата:

$$\overline{\alpha+\beta} = \bar{\alpha} + \bar{\beta},$$

$$\overline{\alpha-\beta} = \bar{\alpha} - \bar{\beta},$$

$$\overline{\alpha\beta} = \bar{\alpha} \cdot \bar{\beta}.$$

7. **Геометрично представяне на комплексните числа.** В равнината вземаме една правоъгълна координатна система XOY. На числото $\alpha = a+bi$ отговаря тогава една точка с абсциса a и ордината b. Така на всяко комплексно число отговаря една точка и, обратно, на всяка точка отговаря едно комплексно число. Оста X наричаме реална ос, а Y — имагинерна ос. Точката, която отговаря на числото α, бележим обикновено пак с α.

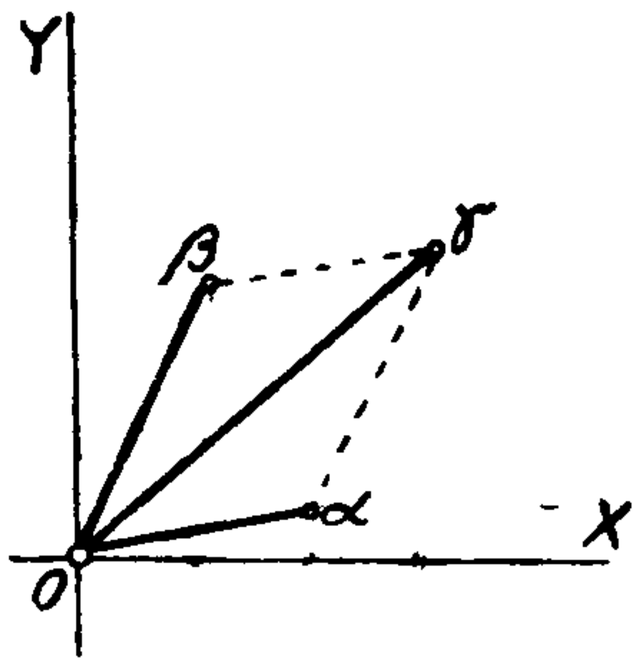
Ако дължината на отсечката $O\alpha$ означим с ρ, а ъгъла между $O\alpha$ и реалната ос с φ, то ще имаме

$$a=\rho\cos\varphi,\quad b=\rho\sin\varphi,\quad \rho^2=a^2+b^2.$$

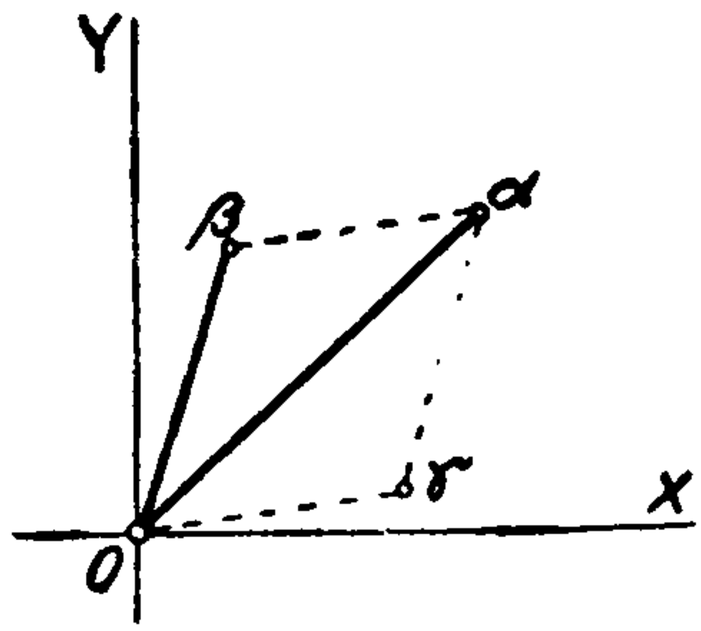
$$\alpha=\rho(\cos\varphi+i\sin\varphi),\quad \operatorname{tg}\varphi=\frac{b}{a}.$$

Последната форма е тригонометрична форма на комплексните числа, ρ се нарича модул (абсолютна стойност) и φ се нарича аргумент. Модулът се пише и така: $|\alpha|$.

Лесно се намира сумата и разликата на две комплексни числа по геометричен път. Именно сумата $\gamma=\alpha+\beta$ се получава веднага, като върху $O\alpha$ и $O\beta$ построим един паралелограм, то γ е другият връх на този паралелограм (черт. 1). Действително от чертежа се вижда, че



Черт. 1



Черт. 2

абсцисата на γ е сума от абсцисите на α и β. Същото е и за ординатата на γ. Тази постройка може да се извърши и така: от точката α теглим отсечка, равна, успоредна и еднакво насочена на $O\beta$. Тогава краят на тази отсечка е γ.

За да получим разликата $\alpha-\beta=\gamma$, достатъчно е да вземем под внимание, че α е сума от β и γ. Конструкцията е дадена в черт. 2. Същата конструкция може да се тълкува така: от α теглим отсечка $\alpha\gamma$, равна и успоредна на βO, със същото направление, т. е. вектора $\overrightarrow{\beta O}$ пренасяме в α.

За да получим произведението на две комплексни числа, ще използуваме тригонометричното им представяне. Ако

$$\alpha=r(\cos\varphi+i\sin\varphi),\ \beta=r_1(\cos\varphi_1+i\sin\varphi_1),$$

то имаме

$$\alpha\beta=rr_1[\cos\varphi\cos\varphi_1-\sin\varphi\sin\varphi_1+i(\sin\varphi\cos\varphi_1+\cos\varphi\sin\varphi_1)]$$

или

$$\alpha\beta=rr_1[\cos(\varphi+\varphi_1)+i\sin(\varphi+\varphi_1)],$$

т. е. при умножаване на две комплексни числа модулите се умножават, а аргументите се събират. Очевидно това

правило е валидно и за произведението на произволен брой комплексни числа.

За частното на две комплексни числа $\frac{\alpha}{\beta}$ ще имаме: ако

$$\frac{\alpha}{\beta}=\gamma=\rho(\cos\psi+i\sin\psi),$$

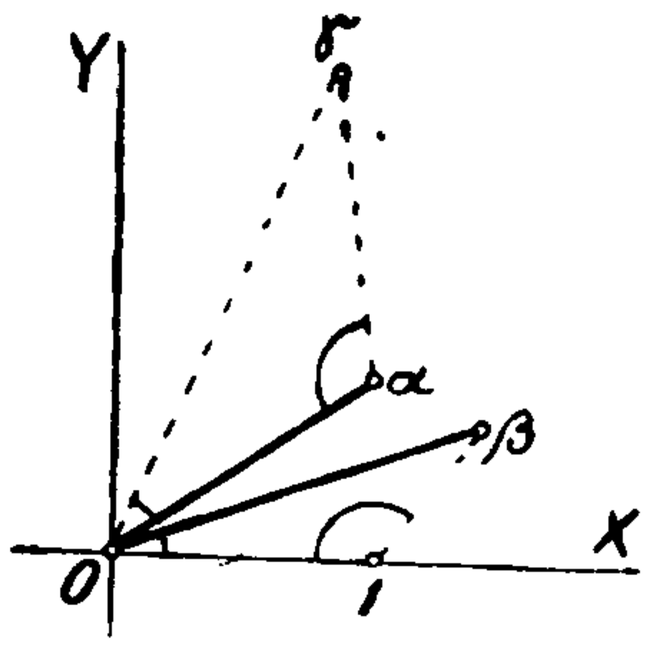
то

$$\alpha=\beta\gamma,\ r(\cos\varphi+i\sin\varphi)=\rho r_1[\cos(\psi+\varphi_1)+i\sin(\psi+\varphi_1)],$$

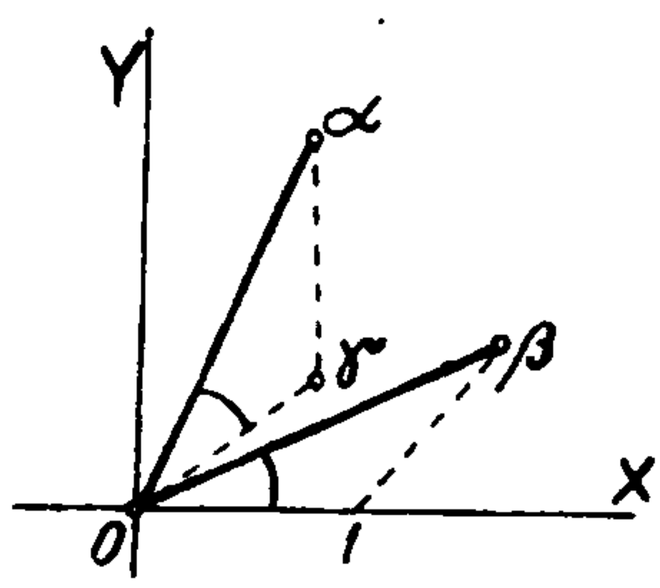
$$\rho=\frac{r}{r_1},\ \psi=\varphi-\varphi_1.$$

Следователно при деление модулите се делят, а аргументите се вадят.

За да построим точката, която отговаря на $\gamma=\alpha\beta$, постъпваме така: върху единия множител, например α, построяваме триъгълник, подобен и еднакво разположен с триъгълника $O1\beta$, като върховете O



Черт. 3



Черт. 4

α съответствуват на O, 1. Третият връх γ представлява произведението (черт. 3).

Действително, ако с R и ψ означим модула и аргумента на γ, то от построението следва, че

$$\frac{R}{r_1}=\frac{r}{1},\ R=rr_1,\ \psi=\varphi+\varphi_1,$$

което доказва, че $\gamma=\alpha\beta$.

За да построим частното $\frac{\alpha}{\beta}=\gamma$, трябва да вземем под внимание, че α е равно на произведението на β и γ. Тази конструкция е дадена на черт. 4, гдето α и β са дадени, като е построен триъгълник $O\alpha\gamma$, подобен на $O\beta1$, като на върховете O, γ съответствуват O, 1.

8. **Теореми за модулите.** Нека са дадени числата

$$\alpha=r(\cos\varphi+i\sin\varphi),\ \beta=r_1(\cos\varphi_1+i\sin\varphi_1),$$

тогава получаваме

$$|\alpha+\beta|^2=(r\cos\varphi+r_1\cos\varphi_1)^2+(r\sin\varphi+r_1\sin\varphi_1)^2=$$
$$=r^2+r_1^2+2r_1r\cos(\varphi-\varphi_1).$$

Понеже $\cos(\varphi-\varphi_1)$ варира между -1 и $+1$, то очевидно най-голяма стойност на $|\alpha+\beta|^2$ при постоянни r и r_1 ще имаме, когато $\cos(\varphi-\varphi_1)=1$, т. е. $\varphi=\varphi_1$, а най-малка — при $\cos(\varphi-\varphi_1)=-1$, т. е. $\varphi-\varphi_1=\pi$. Така че ще имаме

$$|\alpha+\beta|^2\leqq r^2+r_1^2+2rr_1=(r+r_1)^2,$$
$$|\alpha+\beta|^2\geqq r^2+r_1^2-2rr_1=(r-r_1)^2,$$

отгдето

$$|\alpha|-|\beta|\leqq|\alpha+\beta|\leqq|\alpha|+|\beta|.$$

Дясното неравенство се обобщава за повече от две събираеми.

Действително имаме

$$|\alpha+\beta+\gamma|\leqq|\alpha+\beta|+|\gamma|\leqq|\alpha|+|\beta|+|\gamma|.$$

Модулът на една сума от комплексни числа е най-много равен на сумата от модулите на отделните числа. Равенство има само тогава, когато аргументите на комплексните числа са еднакви. Модулът на сумата от две комплексни числа е най-малко равен на разликата от модулите им. Равенство може да има само тогава, когато аргументите им се различават с π.

Читателят лесно ще се убеди, че тези теореми са очевидни геометрически, като вземе под внимание, че едната страна на триъгълника е по-малка от сумата на другите две.

9. **Формула на Moivre.** Видяхме по-рано, че при умножаване на комплексни числа модулите им се умножават, а аргументите се събират. Ако предполо̀жим, че всички числа са равни на $\alpha=r(\cos\varphi+i\sin\varphi)$, то така получаваме формулата на Moivre:

$$[r(\cos\varphi+i\sin\varphi)]^n=r^n(\cos n\varphi+i\sin n\varphi),$$

за цял положителен показател n. Лесно се вижда, че същата формула е валидна при n цяло и отрицателно. Действително имаме при $n=-m$, $m>0$,

$$[r(\cos\varphi+i\sin\varphi)]^n=\frac{1}{[r(\cos\varphi+i\sin\varphi)]^m}=$$

$$=\frac{\cos 0+i\sin 0}{r^m(\cos m\varphi+i\sin m\varphi)}=r^{-m}[\cos(-m\varphi)+i\sin(-m\varphi)],$$

което трябваше да се докаже.

Да намерим формулата за дробен показател. Ако поставим

$$\sqrt[n]{r(\cos\varphi+i\sin\varphi)}=\rho(\cos\psi+i\sin\psi),$$

то по условието

$$r(\cos\varphi+i\sin\varphi)=\rho^n(\cos n\psi+i\sin n\psi),$$

отгдето получаваме

$$r=\rho^n,\quad \varphi+2k\pi=n\psi,$$

гдето k е произволно цяло число. Следователно имаме

$$(1)\qquad \sqrt[n]{r(\cos\varphi+i\sin\varphi)}=\sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n}+i\sin\frac{\varphi+2k\pi}{n}\right)=X_k.$$

Привидно така получаваме безбройно много стойности на n-ия корен, обаче лесно е да се види, че само n от числата X_k са различни помежду си. Именно, ако на k дадем стойности e и m, то ако $e-m$ е кратно на n, $X_e=X_m$, понеже от $e-m=np$ следва

$$\frac{\varphi+2\pi e}{n}-\frac{\varphi+2\pi m}{n}=2\pi p,$$

така че синусите и косинусите ще имат една и съща стойност. Обратно, от $X_e=X_m$ веднага следва, че $e-m$ е кратно на n. Така че формулата (1) дава всички стойности на n-те корена, като положим

$$k=0,\ 1,\ 2,\ldots,n-1.$$

Оттук се получава формулата на Moivre за дробен показател. Именно имаме

$$[r(\cos\varphi+i\sin\varphi)]^{\frac{p}{q}}=r^{\frac{p}{q}}\left[\cos\frac{p(\varphi+2k\pi)}{q}+i\sin\frac{p(\varphi+2k\pi)}{q}\right].$$

10. **Някои геометрични приложения.** Числата

$$\alpha=a+bi,\quad \alpha=a-bi$$

са конюговани. Абсцисите на точките, които отговарят, са равни, а ординатите им са равни по абсолютна стойност, но обратни по знак. Така че тези две точки са симетрични спрямо реалната ос. Да разгледаме делението, и то отначало най-простото: $z_1=\frac{1}{z}$. Ако поставим

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

$$z_1=r_1(\cos\varphi_1+i\sin\varphi_1),$$

то от горната релация веднага се получава

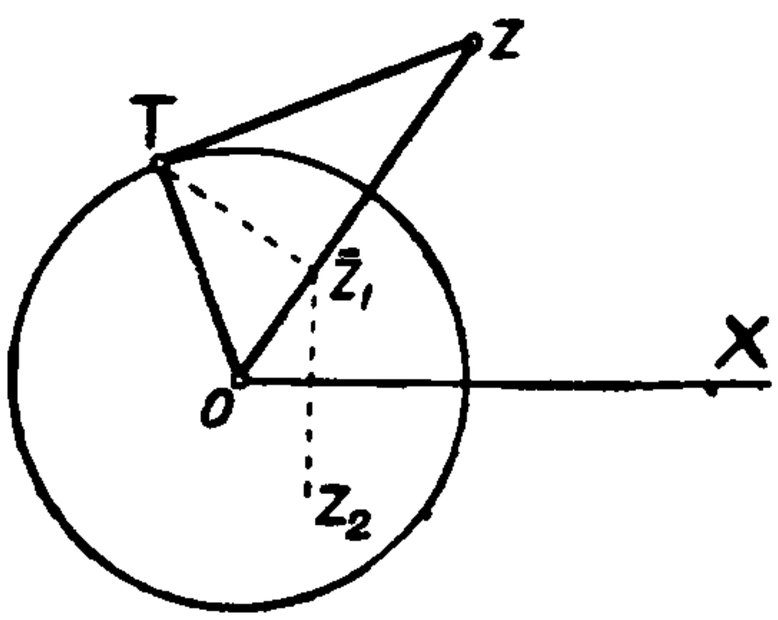
$$r_1=\frac{1}{r},\ \varphi_1=-\varphi.$$

Но понеже $z_1 = r_1(\cos\varphi - i\sin\varphi)$, то тази точка може да се построи така (черт. 5): построяваме с трансформиране с инверзия на точката z инверзния ѝ образ $\bar{z}_1$ спрямо окръжността с център $z=0$ и радиус 1. В чертежа триъгълниците OTz и $O\bar{z}_1T$ са правоъгълни, така че ще имаме

$$|\bar{z}_1|\,|z| = 1, \quad |\bar{z}_1| = \frac{1}{r},$$

и аргументът е същ, което доказва конструкцията.

Ако сега вземем огледалния образ на $\bar{z}_1$ спрямо реалната ос, то ще получим точката $z_1 = \frac{1}{z}$. Ако точката z описва права, то точката $\bar{z}_1$ (и оттам точката z_1) описва, както е известно от елементарната геометрия, окръжност, минаваща през O. Ако z описва окръжност, то и $\bar{z}_1$, а оттам и z_1 описва окръжност. Лесно е да установим това директно, като поставим $z = x + iy$, $z_1 = X + iY$ и намерим връзката между X и Y.



Черт. 5

От горното се вижда, че същото свойство притежава и $u = a + \frac{1}{z}$, където a е произволно фиксирано комплексно число. Така, ако z описва права L, минаваща още през една точка z_0, то понеже при $z = z_0$, $u = a_0 + \frac{1}{z_0} = u_0$, а при $z = \infty$, $u = a$, точката u ще описва окръжност, минаваща през a и u_0.

11. **История.** За въвеждането на комплексните числа е спомогнал стремежът да се решат квадратни уравнения, които нямат за корени реални числа. Така уравнението $x^2 = \gamma$ при $\gamma > 0$ има корени $\pm\sqrt{\gamma}$, но ако $\gamma < 0$, то старите математици са пишели същия израз за x напълно формално, без да влагат за това определено ясно разбиране. Смятали са даже, че някой резултат, доказан посредством имагинерните числа (някои са ги наричали невъзможни), трябва да се провери и директно неговата вярност. Именно това е било така, защото комплексните числа не са били напълно дефинирани ясно, както и действията с тях. Знакът $i = \sqrt{-1}$ е въведен от Ойлер, който вече в много въпроси е употребявал комплексните числа. От Гаус (1799) е въведен терминът комплексно число. Същият в 1811 г. дава геометрическото представяне на тези числа; ето защо равнината, в която се нанасят, често се нарича Гаусова равнина на числата.

Геометричната теория на комплексните числа е била също по-рано разгледана от други математици, на които работите тогава останали в голяма или по-малка неизвестност. Именно това са трудовете на Gaspar Wessel (1799), J. R. Argand (1806) и C. V. Mourey (1820). Теорията с чифтове, която изложихме, е дадена от W. R. Hamilton (1837).

Глава III

Свойства на полиномите

1. Теорема за големи стойности на $|z|$. В полинома

$$f(z)=a_0 z^n+a_1 z^{n-1}+\ldots+a_{n-1} z+a_n,\ a_0\neq 0,$$

на който коефициентите са реални или комплексни, абсолютната стойност на z може да се избере толкова голяма, че модулът на първия член да бъде по-голям от модула на сумата на останалите членове. Или по-другояче казано: съществува число R такова, че при $|z|\geqq R$ да имаме

$$a_0 z^n|>|a_1 z^{n-1}+\ldots+a_n|. \tag{1}$$

Действително нека $\alpha_0,\ \alpha_1,\ldots,\alpha_n$ са модулите на числата $a_0,\ a_1,\ a_2,\ldots,a_n$ и нека α е най-голямото от числата $\alpha_1,\ \alpha_2,\ldots,\alpha_n$. Ако поставим $|z|=r$, то (1) ще бъде сигурно удовлетворено, ако е изпълнено неравенството

$$|a_0 z^n|>|a_1 z^{n-1}|+\ldots+|a_n|,$$

т. е.

$$a_0 r^n>\alpha_1 r^{n-1}+\ldots+\alpha_n,$$

което ще се усили, ако заменим $\alpha_1,\ \alpha_2,\ldots,\alpha_n$ с α, т. е.

$$\alpha_0 r^n>\alpha(r^{n-1}+\ldots+1)=\alpha\frac{r^n-1}{r-1}.$$

Да допуснем, че $r>1$ и изпуснем -1 в числителя, с което още повече усилваме неравенството

$$\alpha_0\geqq\alpha\frac{1}{r-1},\quad r\geqq 1+\frac{\alpha}{\alpha_0}=R.$$

По обратен път заключаваме, че щом $|z|\geqq R$, то следва (1), с което теоремата е доказана. От този резултат следва, че модулът на първия член за достатъчно големи $|z|$ става по-голям от k пъти модула на сумата от останалите членове, гдето $k>0$ е произволно число. Затова достатъчно е $|z|\geqq 1+\frac{k\alpha}{\alpha_0}$. Следствие: Съществува едно определено число $R>0$ така, че $f(z)$ не се анулира вън от кръга с радиус R. Действително нека допуснем, че за $z=z_0$, $|z_0|\geqq$

$\geqq R=1+\frac{\alpha}{\alpha_0}$, $f(z_0)=0$. Тогава от $a_0 z_0^n=-a_1 z_0^{n-1}-\ldots-a_n$ следва

$$|a_0 z_0^n|=|a_1 z_0^{n-1}+\ldots+a_n|,$$

което според теоремата е невъзможно.

Може $R=R(\varepsilon)$ да се избере така, че модулът на $f(z)$ да се намира между границите

$$\alpha_0 r^n(1-\varepsilon) \quad \text{и} \quad \alpha_0 r^n(1+\varepsilon),$$

където $r=|z|$, $\alpha_0=|a_0|$, $\varepsilon>0$ е произволно малко число.

2. **Теорема за малки стойности на** $|z|$. Нека цялата рационална функция $f(z)$ е наредена по растящи степени на z, следователно

$$f(z)=a_0 z^m+a_1 z^{m+1}+\ldots+a_p z^{m+p},\ a_0\neq 0.$$

Винаги можем да изберем абсолютната стойност на z достатъчно малка така, че абсолютната стойност на първия член да бъде по-голяма от абсолютната стойност на сумата от останалите членове. Или с други думи съществува число $\rho>0$ такова, че при $|z|\leqq\rho(z\neq 0,\ m>0)$

$$|a_0 z^m|>|a_1 z^{m+1}+\ldots+a_p z^{m+p}|. \tag{2}$$

Доказателството е напълно аналогично на предшествуващата теорема. Нека $|z|=r$, $|a_i|=\alpha_i$, $i=0, 1, 2,\ldots,p$, α е най-голямото от числата $\alpha_1, \alpha_2,\ldots,\alpha_p$. Очевидно неравенството (2) ще бъде изпълнено, ако е изпълнено неравенството

$$\alpha_0 r^m>\alpha_1 r^{m+1}+\ldots+\alpha_p r^{m+p}$$

или усиленото неравенство

$$\alpha_0 r^m>\alpha(r^{m+1}+\ldots+r^{m+p})=\alpha r^{m+1}\frac{1-r^p}{1-r}.$$

Като положим $r<1$, изпуснем r^p, с което го усилваме, получаваме неравенството

$$\alpha_0 r^m\geqq\alpha\frac{r^{m+1}}{1-r},\quad a_0\geqq\frac{\alpha r}{1-r},$$

от което

$$r\leqq\frac{\alpha_0}{\alpha_0+\alpha}=\rho.$$

Оттук се вижда, че ако е удовлетворено последното неравенство, то ще бъде изпълнено и неравенството (2).

Следствие: Ако $\varepsilon>0$ е произволно малко число, то $\rho(\varepsilon)$ може така да се избере, че за $f(z)$,

$$f(z)=a_0 z^m+\ldots+a_p z^{m+p},$$

при $|z|\leqq\rho(\varepsilon)$ да имаме

$$\alpha_0(1-\varepsilon)r^m\leqq|f(z)|\leqq\alpha_0(1+\varepsilon)r^m,\ \alpha_0=|a_0|,\ r=|z|.$$

3. **Формула на Тейлор и производни.** Нека $f(x)$ е полином от n-та степен:

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_{n-1} x+a_n.$$

Тогава, ако вместо x поставим $x+h$, ще имаме

$$f(x+h)=a_0(x+h)^n+a_1(x+h)^{n-1}+\ldots+a_{n-1}(x+h)+a_n.$$

Ако развием степените на $x+h$ по формулата на Нютон и наредим по степените на h, ще получим

(3)
$$f(x+h)=$$
$$=f(x)+\frac{h}{1!}f'(x)+\frac{h^2}{2!}f''(x)+\frac{h^3}{3!}f'''(x)+\ldots+\frac{h^n}{n!}f^{(n)}(x),$$

гдето

$$f'(x)=na_0x^{n-1}+(n-1)\,a_1x^{n-2}+\ldots+a_{n-1},$$
$$f''(x)=n(n-1)\,a_0\,x^{n-2}+(n-1)(n-2)\,a_1x^{n-3}+\ldots+1\,.\,2a_{n-2},$$
$$f'''(x)=$$
$$=n(n-1)(n-2)\,a_0x^{n-3}+(n-1)(n-2)(n-3)\,a_1x^{n-4}+\ldots+3!\,a_{n-3},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f^{(n-1)}(x)=n(n-1)\ldots2\,.\,a_0x+(n-1)\ldots1\,.\,a_1,$$
$$f^{(n)}(x)=n(n-1)\ldots2\,.\,1\,.\,a_0.$$

Функциите $f'(x)$, $f''(x)$,... са цели рационални съответно от степени $n-1$, $n-2$,... Те се наричат **първа, втора и т. н. производна** на $f(x)$. Първата производна на $f(x)$ се получава, като всеки член се умножава със степента му и тя се намалява с единица. От това е очевидно, че втората производна $f''(x)$ е първа производна на $f'(x)$ и т. н. Формулата (3) се **нарича формула на Тейлор.**

От (3) следва

(4)
$$f(x+h)-f(x)=\frac{h}{1!}f'(x)+\frac{h^2}{2!}f''(x)+\ldots+\frac{h^n}{n!}f^{(n)}(x).$$

Коефициентите на h, h^2,... са полиноми на x и следователно при крайна стойност на x имат крайни стойности. Тогава по основните свойства на полиномите, понеже дясната част на (4) е полином на h, за всяко произволно малко $\varepsilon>0$ може да се намери съответно $r>0$ така, че щом $|h|\leqq r$,

$$|f(x+h)-f(x)|<\varepsilon$$

или другояче, както и h да клони към нула, разликата $f(x+h)-f(x)$ ще клони към нула за всяко крайно x. Следователно всеки полином $f(x)$ е непрекъсната функция за всяко крайно x.

От (4) се получава

$$\frac{f(x+h)-f(x)}{h}=f'(x)+\frac{h}{2!}f''(x)+\ldots+\frac{h^{n-1}}{n!}f^{(n)}(x).$$

Ако оставим x постоянно, а h да клони по произволен начин към нула, то дясната част очевидно ще клони към $f'(x)$, т. е. ще имаме

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x).$$

Оттук се вижда, че производната на полинома $f(x)$ може да се дефинира като граница на отношението на разликата от стойностите на полинома $f(x+h)-f(x)$ към разликата h на стойностите на променливото, когато тази разлика клони към нула. Тази дефиниция се прилага в анализа за произволни функции. Оттук става ясно, че правилата за диференциране при реални променливи остават в сила.

4. **Биномна и полиномна формула.** От елементарната алгебра е известна формулата на Нютон:

$$(x+y)^n=x^n+\binom{n}{1}x^{n-1}y+\binom{n}{2}x^{n-2}y^2+\ldots+y^n,$$

гдето n е цяло положително число. Биномните коефициенти $\binom{n}{k}$, като поставим $\binom{n}{0}=1$, притежават свойството

$$\binom{n}{k}=\frac{n!}{k!\,(n-k)!}\quad\text{и}\quad\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1},$$

което веднага следва от умножението на $(x+y)^{n-1}$ с $(x+y)$.

Като обобщение на биномната теорема е полиномната, която се отнася до развитието на

$$(x_1+x_2+\ldots+x_p)^n.$$

Понеже тази функция е хомогенна спрямо $x_1, x_2, \ldots, x_p$, то ще имаме

$$(x_1+x_2+\ldots+x_p)^n=\sum C_{\nu_1,\nu_2,\ldots,\nu_p}x_1^{\nu_1}x_2^{\nu_2}\ldots x_p^{\nu_p}$$
$$(\nu_1+\nu_2+\ldots+\nu_p=n),$$

гдето коефициентите C подлежат на определяне. По биномната формула имаме

$$(x_1+x_2+\ldots+x_p)^n=\sum_{\nu_p=0}^{n}\binom{n}{\nu_p}(x_1+x_2+\ldots+x_{p-1})^{n-\nu_p}x_p^{\nu_p}.$$

$$(x_1+x_2+\ldots+x_{p-1})^{n-\nu_p}=$$
$$=\sum_{\nu_{p-1}=0}^{n-\nu_p}\binom{n-\nu_p}{\nu_{p-1}}(x_1+x_2+\ldots+x_{p-2})^{n-\nu_p-\nu_{p-1}}x_{p-1}^{\nu_{p-1}},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(x_1+x_2)^{n-\nu_p-\ldots-\nu_2}=$$
$$=\sum_{\nu_2=0}^{n-\nu_p-\ldots-\nu_3}\binom{n-\nu_p-\ldots-\nu_3}{\nu_2}x_1^{n-\nu_p-\ldots-\nu_3-\nu_2}x_2^{\nu_2}.$$

Ако постепенно заместим тези резултати, започвайки от крайната формула, то ще получим като коефициент на

$$x_1^{n-\nu-\ldots-\nu_2} x_2^{\nu_2} \ldots x_p^{\nu} p = x_1^{\nu_1} x_2^{\nu_2} \ldots x_p^{\nu} p$$

в развитието на $(x_1+x_2+\ldots+x_p)^n$:

$$C_{\nu_1, \nu_2, \ldots, \nu_p} = \binom{n}{\nu_p}\binom{n-\nu_p}{\nu_{p-1}}\ldots\binom{n-\nu_p-\ldots-\nu_3}{\nu_2} =$$

$$= \frac{n!}{\nu_p!\,(n-\nu_p)!}\,\frac{(n-\nu_p)!}{\nu_{p-1}!\,(n-\nu_p-\nu_{p-1})!}\ldots\frac{(n-\nu_p-\ldots-\nu_3)!}{\nu_2!\,(n-\nu_p-\ldots-\nu_2)!} =$$

$$= \frac{n!}{\nu_1!\,\nu_2!\ldots\nu_p!}.$$

Така получаваме формулата

$$(5)\qquad (x_1+x_2+\ldots+x_p)^n = \sum \frac{n!}{\nu_1!\,\nu_2!\ldots\nu_p!}\, x_1^{\nu_1} x_2^{\nu_2}\ldots x_p^{\nu} p$$

$$(\nu_1+\nu_2+\ldots+\nu_p=n),$$

гдето сумирането е разпространено върху всички цели положителни или нули индекси ν.

Можем последната формула да изведем директно. Именно, като извършим умножението на n-те линейни форми

$$(x_1+x_2+\ldots+x_p)\,(x_1+x_2+\ldots+x_p)\ldots(x_1+x_2+\ldots+x_p),$$

ще получим една сума от p^n члена:

$$(6)\qquad (x_1+x_2+\ldots+x_p)^n = \sum_{\lambda_1=1}^{p}\sum_{\lambda_2=}^{p}\cdots\sum_{\lambda_n=1}^{p} x_{\lambda_1} x_{\lambda_2}\ldots x_{\lambda_n}.$$

$$(\nu_1+\nu_2+\ldots+\nu_p=n).$$

От тези членове някои се повтарят. Да видим колко члена са равни на

$$(7)\qquad x_1^{\nu_1} x_2^{\nu_2}\ldots x_p^{\nu} p \quad (\nu_1+\nu_2+\ldots+\nu_p=n).$$

Ясно е, че такива членове влизат в дясната част на (6), когато ν_1 индекси λ са равни на 1, ν_2 на 2 и т. н. Следователно броят на тези членове ще бъде равен на броя на пермутациите на n елемента, от които ν_1 се повтарят, ν_2 други се повтарят и т. н. до последна група от ν_p пак повтарящи се елементи, т. е. е равен на

$$\frac{n!}{\nu_1!\,\nu_2!\ldots\nu_p!}.$$

Това число следователно ще бъде коефициентът на (7) в развитието. Така получаваме отново формулата (5).

ЧАСТ II

ТЕОРИЯ НА ДЕТЕРМИНАНТИТЕ И ПРИЛОЖЕНИЯ

Глава I

Основни свойства на детерминантите

1. **История.** Едва в средата на миналия век била доста пълно разработена теорията на детерминантите. Както много теории в математиката, така и тази теория е възникнала с решаването на конкретни задачи. Именно пръв Leibniz в 1693 г. забелязва едно общо правило при решаването на системи линейни уравнения, което води до дефинирането на детерминантите. По-късно Cramer (Introduction à l'analyse des lignées courbées algébriques, Genève, 1750) дава строга дефиниция, като намира и правилото за решаване на линейните уравнения. След това са работили Laplace, Lagrange, Vandermonde, но теорията им е била главно завършена в стройна система с работите на Cauchy и Jacobi в средата на XIX век.

2. **Дефиниция на детерминанти.** Нека са дадени 2 уравнения с две неизвестни:

$$a_{11}x_1+a_{12}x_2=a_1,$$

$$a_{21}x_1+a_{22}x_2=a_2.$$

Умножаваме първото уравнение с a_{22}, второто с $-a_{12}$ и ги събираме. Също умножаваме първото с $-a_{21}$, второто с a_{11} и ги събираме. Така получаваме следните резултати:

$$(a_{11}a_{22}-a_{12}a_{21})x_1=a_1a_{22}-a_2a_{12},$$

$$(a_{11}a_{22}-a_{12}a_{21})x_2=a_2a_{11}-a_1a_{21}.$$

Ако изразът $a_{11}a_{22}-a_{12}a_{21}$ е отличен от нула, то можем веднага да намерим x_1 и x_2. Изразът

$$a_{11}a_{22}-a_{12}a_{21}$$

наричаме детерминанта от втори ред и го бележим така:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Числата a_{ik}, $i, k=1, 2$, се наричат елементи на детерминантата. С това означение веднага се вижда, че решението на дадената система е следното:

$$x_1=\frac{\begin{vmatrix} a_1 & a_{12} \\ a_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2=\frac{\begin{vmatrix} a_{11} & a_1 \\ a_{21} & a_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Ако решим 3 уравнения с три неизвестни, то ще получим аналогични правила, които по-сетне ще дадем като приложение на детерминантите.

От n елемента, както е известно от елементарната алгебра, можем да образуваме всичко $n!$ пермутации. Една от тези пермутации нека считаме за главна. Ако в една дадена пермутация два елемента са в обратен ред на реда им в главната пермутация, то казваме, че те образуват една инверзия. По-нататък ще считаме, че елементите са числата $1, 2, 3, \ldots, n$, а главната пермутация е $123 \ldots n$. Тогава очевидно два елемента в една пермутация образуват инверзия, ако по-големият е пред по-малкия. Ако броят на инверзиите е четен, то казваме, че пермутацията е четна, или от I клас, а ако е нечетен, пермутацията е нечетна, или от II клас. Броят на инверзиите в пермутацията $\alpha\beta\gamma \ldots \lambda$ ще означаваме така:

$$[\alpha\beta\gamma \ldots \lambda].$$

Така ще имаме

$$[1\ 2\ 3\ 4]=0,\ [1\ 4\ 3\ 2]=3,\ [2\ 3\ 1\ 4]=2,\ [4\ 3\ 2\ 1]=6,\ [4\ 2\ 3\ 1]=5.$$

Очевидно е, че най-много инверзии има пермутацията

$$n,\ n-1,\ n-2, \ldots, 2,\ 1,$$

броят на които е

$$n-1+n-2+\ldots+2+1=\frac{n(n-1)}{2}.$$

Лема. Ако в една пермутация заместим два елемента един с друг, то тя променя класа си.

Действително нека дадената пермутация е

$$Aa\alpha_1\alpha_2 \ldots \alpha_p bB, \tag{1}$$

където A са елементите пред a и B тези след b. Ако сменим a и b помежду им, т. е. извършим така наречената транспозиция (ab), то получаваме пермутацията

$$Ab\alpha_1\alpha_2 \ldots \alpha_p aB. \tag{2}$$

При това сменяване броят на инверзиите на a и b с елементите в A и B не се променя. Така че изменение може да стане само в броя

на инверзиите с α и помежду им. За да получим (2) от (1), постъпваме така: извършваме транспозицията $(a\ \alpha_1)$ и получаваме

$$A\alpha_1 a\alpha_2 \ldots \alpha_p bB, \tag{3}$$

в която пермутация броят на инверзиите е също както в (1) само с тази разлика, че ако между a и α_1 е имало инверзия, сега няма и обратно. Така класът се променя. Ако сега сменим a с α_2, получаваме

$$A\alpha_1 \alpha_2 a\alpha_3 \ldots \alpha_p bB, \tag{4}$$

като класът се изменя още един път. Продължавайки така, получаваме пермутацията

$$A\alpha_1 \alpha_2 \ldots \alpha_p b\, a\, B,$$

като класът се изменя $(p+1)$ пъти. Ако сега сменяваме така последователно b с $\alpha_p, \alpha_{p-1}, \ldots, \alpha_1$, то класът ще се измени още p пъти и ще получим пермутацията (2). Следователно (2) се получава от (1), като класът се изменя $2p+1$ пъти, т. е. нечетно число пъти, което показва, че тези пермутации са от различни класове.

Като непосредствено приложение на тази лема ще намерим броя на пермутациите от двата класа. От всички пермутации от I клас с една транспозиция получаваме пермутациите от II клас и, обратно, т. е. ако r и s са броят на пермутациите от I и II клас съответно, то ще имаме

$$r \leqq s, \quad s \leqq r,$$

отгдето $r = s = \frac{n!}{2}$.

Ако са дадени nm елемента a_{ik}, гдето $i = 1, 2, 3, \ldots, n$, $k = 1, 2, 3, \ldots, m$ то правоъгълната схема

$$\left\| \begin{array}{cccc} a_{11} & a_{12} & a_{13} \ldots & a_{1m} \\ a_{21} & a_{22} & a_{23} \ldots & a_{2m} \\ a_{31} & a_{32} & a_{33} \ldots & a_{3m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} \ldots & a_{nm} \end{array} \right\|$$

се нарича матрица от типа (n, m). Ако $m = n$, то тя се казва още и квадратна.

Под детерминанта Δ от n^2 елемента a_{ik}, $i, k = 1, 2, \ldots, n$, разбираме сумата от $n!$ члена:

$$\sum (-1)^{[i_1, i_2, i_3, \ldots, i_n]} a_{1\, i_1} a_{2\, i_2} \ldots a_{n i_n}, \tag{5}$$

гдето $i_1\ i_2 \ldots i_n$ са пермутациите на елементите $1, 2, 3, \ldots, n$, и я бележим

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \ldots a_{1n} \\ a_{21} & a_{22} & a_{23} \ldots a_{2n} \\ a_{31} & a_{32} & a_{33} \ldots a_{3n} \\ \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} \ldots a_{nn} \end{vmatrix}.$$

Често Δ се бележи съкратено и така: $\Delta = |a_{ik}|$, $i, k = 1, 2, \ldots, n$.

Ясно е, че детерминантата се получава от члена

$$a_{11}\ a_{22}\ a_{33} \ldots a_{nn},$$

наречен главен, като пермутираме вторите индекси по всевъзможни начини и получените членове вземем със знака $+$, ако пермутацията от вторите индекси е четна, и със знака $-$, ако е нечетна. Δ се нарича **детерминанта от n-ти ред**. Елементите $a_{i1}\ a_{i2} \ldots a_{in}$ образуват i-ия хоризонтален ред, а $a_{1i}\ a_{2i} \ldots a_{ni}$ образуват i-ия вертикален ред или стълб. Общо хоризонталните и вертикалните редове (или стълбове) се наричат **редове на детерминантата**.

Да разгледаме за члена

$$a_{\alpha_1\ \beta_1}\ a_{\alpha_2\ \beta_2} \ldots a_{\alpha_n\ \beta_n}, \tag{6}$$

в който $\alpha_1\ \alpha_2 \ldots \alpha_n$, $\beta_1\ \beta_2 \ldots \beta_n$ са две произволни пермутации от елементите $1, 2, 3, \ldots, n$. Сумата от инверзиите е

$$s = [\alpha_1\ \alpha_2 \ldots \alpha_n] + [\beta_1\ \beta_2 \ldots \beta_n].$$

Ако в (6) разменим два елемента $a_{\alpha i\ \beta i}$ с $a_{\alpha k\ \beta k}$ помежду им, то по лемата двете пермутации от първите и вторите индекси ще си променят класа, отгдето е очевидно, че s ще се мени с четно число. Така че от члена

$$(-1)^s\, a_{\alpha_1\beta_1} a_{\alpha_2\beta_2} \ldots a_{\alpha_n\ \beta_n}$$

получаваме члена

$$(-1)^{s_1} a_{1\gamma_1} a_{2\gamma_2} \ldots a_{n\gamma_n},$$

гдето $s_1 = [\gamma_1\ \gamma_2 \ldots \gamma_n]$, и така сме разменили елементите, че първите индекси да станат в натурален ред. От тази забележка е явно, че за Δ ще имаме

$$\Delta = \sum (-1)^{[i_1\ i_2 \ldots\ i_n] + [j_1\ j_2 \ldots\ j_n]}\ a_{i_1 j_1}\ a_{i_2 j_2} \ldots a_{i_n j_n},$$

гдето сумирането е разпространено върху всички пермутации на елементите $1, 2, \ldots, n$, като повтарящите се членове са написани еднократно.

Да разгледаме детерминантата от втори ред като пример:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Δ ще се състои от два члена $a_{11}a_{22}$ и $a_{12}a_{21}$, от които първия трябва да вземем със знак $+$, а втория с $-$, понеже $[2, 1]=1$, така че

$$\Delta = a_{11}\,a_{22} - a_{12}\,a_{21}.$$

Като друг пример да разгледаме детерминантата от третия ред:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Лесно се получава

$$\Delta = a_{11}\,a_{22}\,a_{33} + a_{12}\,a_{23}\,a_{31} + a_{13}\,a_{21}\,a_{32} - \tag{7}$$
$$- a_{12}\,a_{21}\,a_{33} - a_{13}\,a_{22}\,a_{31} - a_{11}\,a_{23}\,a_{32}.$$

В детерминантата елементите $a_{11}\ a_{22}\ldots a_{nn}$ образуват главния диагонал, а елементите $a_{n1}\,a_{n-1,2}\ldots a_{1n}$ — втория диагонал.

За получаване стойността на детерминантите от трети ред може да се ползуваме от следното правило на Сарус: Преписваме първите два стълба надясно от Δ и след това умножаваме елементите на главния диагонал и паралелните му два диагонала, като получените произведения вземаме със знак $+$, също умножаваме елементите във втория диагонал и паралелните на него други два, но получените произведения вземаме със знак $-$. Едно просто пресмятане ни показва, че действително така получаваме израза (7) за Δ. Така за

$$\Delta = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

пишем

$$\begin{array}{ccccc} a_1 & a_2 & a_3 & a_1 & a_2 \\ b_1 & b_2 & b_3 & b_1 & b_2 \\ c_1 & c_2 & c_3 & c_1 & c_2 \end{array}$$

и получаваме

$$\Delta = a_1\,b_2\,c_3 + a_2\,b_3\,c_1 + a_3\,b_1\,c_2 - a_3\,b_2\,c_1 - a_1\,b_3\,c_2 - a_2\,b_1\,c_3.$$

3. **Елементарни свойства на детерминантите.** От дефиницията на детерминантата следват непосредствено следните свойства:

1. **Ако заместим редовете със стълбовете и обратно,** детерминантата не се изменя, т. е.

$$\begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} \dots a_{n1} \\ a_{12} & a_{22} \dots a_{n2} \\ \dots & \dots \\ a_{1n} & a_{2n} \dots a_{nn} \end{vmatrix}.$$

Действително на члена в първата детерминанта

$$(-1)^s a_{\mu_1\lambda_1} a_{\mu_2\lambda_2} \dots a_{\mu_n\lambda_n}, \quad s=[\mu_1\ \mu_2 \dots \mu_n]+[\lambda_1\ \lambda_2 \dots \lambda_n]$$

отговаря членът

$$(-1)^{s_1} a_{\lambda_1\ \mu_1} a_{\lambda_2\ \mu_2} \dots a_{\lambda_n\mu_n}, \quad s_1=s,$$

от което предложението е очевидно.

2. Ако разменим два паралелни реда помежду им, в детерминанта, то новата детерминанта е равна на $-\Delta$.

Нека разместим r-ия с s-ия ред и обратно, т. е.

$$\Delta=\begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ \dots & \dots \\ a_{r-1,1} & a_{r-1,2} \dots a_{r-1,n} \\ a_{r1} & a_{r2} \dots a_{rn} \\ \dots & \dots \\ a_{s1} & a_{s2} \dots a_{sn} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}, \quad \Delta'=\begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ \dots & \dots \\ a_{s1} & a_{s2} \dots a_{sn} \\ \dots & \dots \\ a_{r1} & a_{r2} \dots a_{rn} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}.$$

Ако елементите на Δ' означим с b_{ij}, то те са равни съответно на a_{ij} с изключение на $b_{rk}=a_{sk}$, $b_{sk}=a_{rk}$, $k=1, 2, \dots, n$, така че за Δ' ще имаме

$$\Delta'=\sum(-1)^{[\alpha\beta\dots\mu\dots\nu\dots\delta]} a_{1\alpha} a_{2\beta} \dots a_{s\mu} \dots a_{r\nu} \dots a_{n\delta},$$

гдето $\alpha\beta\dots\mu\dots\nu\dots\delta$ е пермутация на $1, 2, \dots, n$. За Δ ще имаме

$$\Delta=\sum(-1)^{[\alpha\beta\dots\nu\dots\mu\dots\delta]} a_{1\alpha} a_{2\beta} \dots a_{r\nu} \dots a_{s\mu} \dots a_{n\delta}$$

и понеже на основание на доказаната лема

$$(-1)^{[\alpha\beta\dots\nu\dots\mu\dots\delta]} = -(-1)^{[\alpha\beta\dots\mu\dots\nu\dots\delta]},$$

то $\Delta'=-\Delta$, с което предложението се доказва. Ако следователно направим p такива разместванията по на два паралелни реда, детерминантата се умножава с $(-1)^p$. Така за .

$$\Delta_1 = \begin{vmatrix} a_{i_1k_1} & a_{i_1k_2} & \dots & a_{i_1k_n} \\ a_{i_2k_1} & a_{i_2k_2} & \dots & a_{i_2k_n} \\ \dots & \dots & \dots & \dots \\ a_{i_nk_1} & a_{i_nk_2} & \dots & a_{i_nk_n} \end{vmatrix}$$

ще имаме $\Delta_1 = \pm\Delta$. Членът

$$a_{i_1k_1} a_{i_2k_2} \dots a_{i_nk_n}$$

в Δ_1 е със знак $+$, а в Δ с

$$(-1)^{[i_1i_2\dots i_n]+[k_1k_2\dots k_n]},$$

така че ще имаме

$$\Delta_1 = \Delta(-1)^{[i_1i_2\dots i_n]+[k_1k_2\dots k_n]}.$$

3. Детерминантата е равна на нула, ако два паралелни реда са равни.

Действително, ако в Δ сменим тези два паралелни равни редове помежду им, то по свойство 2 трябва новата детерминанта да е равна на $-\Delta$, т. е. $\Delta = -\Delta$, отгдето $\Delta = 0$.

4. **Адюнгирани количества и детерминантата като функция на елементите от един ред.** От дефиницията на детерминантата е очевидно, че всеки член в развитието ѝ съдържа само елементи от различни хоризонтални редове и стълбове. В развитието на

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}$$

$\Delta = \sum(-1)^{[\alpha\beta\dots\delta]} a_{1\alpha} a_{2\beta} \dots a_{n\delta}$ нека от всички членове, които съдържат a_{ik}, го извадим пред скоби. Изразът в скобите, които бележим с A_{ik}, се нарича адюнгирано количество на елемента a_{ik}. Ясно е, че A_{ik} не съдържа елементи от i-ия ред и k-ия стълб. Следователно ще имаме

$$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik}, \tag{8}$$

което е търсеното представяне по i-ия ред. По s-ия стълб ще имаме аналогично

$$\Delta = a_{1s}A_{1s} + a_{2s}A_{2s} + \dots + a_{ns}A_{ns} = \sum_{k=1}^{n} a_{ks}A_{ks}. \tag{9}$$

Въз основа на това развитие лесно ще получим редица свойства.

4. Ако елементите на един ред са нули, то детерминантата е равна на нула.

Това се вижда от (8) и (9) веднага.

5. Ако всички елементи на един ред умножим на едно число g, то детерминантата се умножава на това число.

Действително, ако умножим i-ия ред с g, то новата детерминанта Δ', понеже адюнгираните количества A_{ik} не зависят от елементите на i-ия ред, ще бъде

$$\Delta' = ga_{i1}A_{i1} + ga_{i2}A_{i2} + \ldots + ga_{in}A_{in} = g\Delta.$$

6. Ако в една детерминанта елементите на един ред, например i-ия, са суми от две събираеми, то тя е равна на сума от две детерминанти, в които i-ия ред са първите събираеми в едната и вторите събираеми в другата, а останалите елементи са еднакви.

Действително, ако имаме $a_{ik} = \alpha_{ik} + \beta_{ik}$, $k = 1, 2, \ldots, n$, то

$$\Delta = (a_{i1} + \beta_{i1})A_{i1} + (\alpha_{i2} + \beta_{i2})\, A_{i2} + \ldots + (\alpha_{in} + \beta_{in})\, A_{in} =$$

$$= \alpha_{i1}A_{i1} + \alpha_{i2}A_{i2} + \ldots + \alpha_{in}A_{in} + \beta_{i1}A_{i1} + \beta_{i2}A_{i2} + \ldots +$$

$$+ \beta_{in}A_{in} = \Delta' + \Delta''.$$

7. Детерминантата не си изменя стойността, ако към елементите на един ред прибавим съответните елементи на друг паралелен нему ред, умножени с едно число g.

Действително по свойствата 3, 5, 6 имаме ($i < s$)

$$\begin{vmatrix} a_{11} \ldots a_{1i} + ga_{1s} \ldots a_{1s} \ldots a_{1n} \\ a_{21} \ldots a_{2i} + ga_{2s} \ldots a_{2s} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} \ldots a_{ni} + ga_{ns} \ldots a_{ns} \ldots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} \ldots a_{1i} \ldots a_{1s} \ldots a_{1n} \\ a_{21} \ldots a_{2i} \ldots a_{2s} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} \ldots a_{ni} \ldots a_{ns} \ldots a_{nn} \end{vmatrix} +$$

$$+ g \begin{vmatrix} a_{11} \ldots a_{1s} \ldots a_{1s} \ldots a_{1n} \\ a_{21} \ldots a_{2s} \ldots a_{2s} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} \ldots a_{ns} \ldots a_{ns} \ldots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} \ldots a_{1i} \ldots a_{1s} \ldots a_{1n} \\ a_{21} \ldots a_{2i} \ldots a_{2s} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} \ldots a_{ni} \ldots a_{ns} \ldots a_{nn} \end{vmatrix}.$$

Така например

$$\begin{vmatrix} a & b + \beta & c + \gamma \\ a_1 & b_1 + \beta_1 & c_1 + \gamma_1 \\ a_2 & b_2 + \beta_2 & c_2 + \gamma_2 \end{vmatrix} = \begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} + \begin{vmatrix} a & b & \gamma \\ a_1 & b_1 & \gamma_1 \\ a_2 & b_2 & \gamma_2 \end{vmatrix} +$$

$$+\begin{vmatrix} a & \beta & c \\ a_1 & \beta_1 & c_1 \\ a_2 & \beta_2 & c_2 \end{vmatrix} + \begin{vmatrix} a & \beta & \gamma \\ a_1 & \beta_1 & \gamma_1 \\ a_2 & \beta_2 & \gamma_2 \end{vmatrix}.$$

5. **Поддетерминанти и развитие по тях.** Нека в Δ зачеркнем елементите от i-ия ред и k-ия стълб и от останалите елементи, без да изменяме взаимното им положение, да образуваме една детерминанта Δ_{ik} от ред $n-1$. Така получената детерминанта се нарича поддетерминанта на елемента a_{ik}. Тя ще има следната форма:

$$\Delta_{ik} = \begin{vmatrix} a_{11} \dots a_{1,k-1} & a_{1,k+1} \dots a_{1n} \\ \dots\dots\dots\dots\dots\dots \\ a_{i-1,1} \dots a_{i-1,k-1} & a_{i-1,k+1} \dots a_{i-1,n} \\ a_{i+1,1} \dots a_{i+1,k-1} & a_{i+1,k+1} \dots a_{i+1,n} \\ \dots\dots\dots\dots\dots\dots \\ a_{n1} \dots a_{n,k-1} & a_{n,k+1} \dots a_{n,n} \end{vmatrix}.$$

Ще докажем, че между адюнгираното количество и поддетерминантата на един елемент има простата зависимост:

$$A_{ik} = (-1)^{i+k}\Delta_{ik}. \tag{10}$$

Отначало ще установим, че $A_{11} = \Delta_{11}$, отгдето ще получим и общия случай. От развитието

$$\Delta = \sum (-1)^{[\alpha\beta\dots\gamma]} a_{1\alpha}\, a_{2\beta} \dots a_{n\gamma}$$

е очевидно, че за да получим A_{11}, т. е. коефициента на a_{11}, трябва да поставим $\alpha = 1$ и да сумираме върху всички пермутации от елементите $2, 3, \dots, n$. Следователно

$$A_{11} = \sum (-1)^{[\beta\delta\dots\gamma]} a_{2\beta}\, a_{3\delta} \dots a_{n\gamma} = \begin{vmatrix} a_{22} & a_{23} \dots a_{2n} \\ a_{32} & a_{33} \dots a_{3n} \\ \dots\dots\dots\dots \\ a_{n2} & a_{n3} \dots a_{nn} \end{vmatrix} = \Delta_{11},$$

понеже $(-1)^{[\beta\delta\dots\gamma]} = (-1)^{[\beta\delta\dots\gamma]}$, тъй като всички $\beta, \delta, \dots, \gamma$ са по-големи от 1.

За да установим формулата (10), преработваме Δ, като сменяваме i-ия ред последователно с $i-1, i-2, \dots,$ докато дойде на първо място. Също правим с k-ия стълб. Така детерминантата се умножава с $(-1)^{i+k}$, т. е. ще имаме

$$(-1)^{i+k}\Delta=\begin{vmatrix} a_{ik} & a_{i1} & \ldots a_{i,k-1} & a_{i,k+1} & \ldots a_{in} \\ a_{1k} & a_{11} & \ldots a_{1,k-1} & a_{1,k+1} & \ldots a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{i-1,k} & a_{i-1,1} & \ldots a_{i-1,k-1} & a_{i-1,k+1} & \ldots a_{i-1,n} \\ a_{i+1,k} & a_{i+1,1} & \ldots a_{i+1,k-1} & a_{i+1,k+1} & \ldots a_{i+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{nk} & a_{n1} & \ldots a_{n,k-1} & a_{nk+1} & \ldots a_{nn} \end{vmatrix}.$$

Очевидно е, че поддетерминантата на a_{ik} в горната детерминанта е точно Δ_{ik}. Тогава от

$$(-1)^{i+k}(A_{i1}a_{i1}+\ldots+A_{ik}a_{ik}+\ldots A_{in}a_{in})=a_{ik}\Delta_{ik}+\ldots$$

с приравняване на коефициента пред a_{ik} получаваме

$$\Delta_{ik}=(-1)^{i+k}A_{ik},$$

отгдето следва и (10).

От получените резултати следват развитията по елементите на i-ия ред и k-ия стълб:

$$\Delta=(-1)^{i+1}a_{i1}\Delta_{i1}+(-1)^{i+2}a_{i2}\Delta_{i2}+\ldots+(-1)^{i+n}a_{in}\Delta_{in},$$

$$\Delta=(-1)^{k+1}a_{1k}\Delta_{1k}+(-1)^{k+2}a_{2k}\Delta_{2k}+\ldots+(-1)^{k+n}a_{nk}\Delta_{nk}.$$

Например имаме

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-b_1\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}+c_1\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}=$$

$$=-a_2\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}+b_2\begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix}-c_2\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}=$$

$$=a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}-b_3\begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}+c_3\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Ако елементите на един ред са нули с изключение на един, то детерминантата е равна на произведението на този елемент с адюнгираното му количество. Оттук веднага се вижда, че ако елементите от едната страна на главния диагонал са нули, то детерминантата е равна на главния си член. Така

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots \\ 0 & a_{22} & a_{23} & \dots \\ 0 & 0 & a_{33} & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots \\ 0 & a_{33} & \dots \\ \dots & \dots & \dots \\ \dots & \dots & a_{nn} \end{vmatrix} =$$

$$= \dots = a_{11} a_{22} \dots a_{nn}.$$

Видяхме, че

$$\Delta = a_{i1} A_{i1} + a_{i2} A_{i2} + \dots + a_{in} A_{in}.$$

Изразът

$$u = a_{i1} A_{k1} + a_{i2} A_{k2} + \dots + a_{in} A_{kn},$$

гдето $k \neq i$ представя една детерминанта, в която i-ият и k-ият редове са еднакви, следователно $u = 0$.

Глава II

Системи линейни уравнения

1. Обща система. Да разгледаме n линейни уравнения с n известни

$$(1) \qquad \begin{cases} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = a_1, \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n = a_2, \\ \dots \dots \dots \dots \dots \dots \dots \\ a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n = a_n. \end{cases}$$

Детерминантата A, образувана от коефициентите пред неизвестните, се нарича детерминанта на системата. Нека A_{ik} да означава адюнгираното количество на a_{ik}. Да умножим първото уравнение с A_{1k}, второто с A_{2k} и т. н., а последното с A_{nk} и получените резултати да съберем. Коефициентът пред x_k ще бъде равен на

$$a_{1k} A_{1k} + a_{2k} A_{2k} + \dots + a_{nk} A_{nk} = A,$$

а коефициентът пред x_s, $s \neq k$:

$$a_{1s} A_{1k} + a_{2s} A_{2k} + \dots + a_{ns} A_{nk} = 0.$$

Изразът вдясно на резултата ще бъде

$$A_k = a_1 A_{1k} + a_2 A_{2k} + \dots + a_n A_{nk},$$

който представя детерминантата, получена от A със сменяване на елементите от k-ия стълб със свободните членове. Така ще имаме

$$Ax_1 = A_1, \; Ax_2 = A_2, \dots, Ax_n = A_n,$$

отгдето, ако $A \neq 0$, получаваме формулите на Крамер:

$$x_1=\frac{A_1}{A},\quad x_2=\frac{A_2}{A},\ldots,\quad x_n=\frac{A_n}{A}.$$

Случаят, когато $A=0$, ще разгледаме при теоремата на Rouché.

2. **Система хомогенни уравнения.** Нека са дадени n линейни хомогенни уравнения с n неизвестни:

$$(2)\qquad \begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,\\ \ldots\ldots\ldots\ldots\ldots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0. \end{cases}$$

Очевидно е, че $x_1=x_2=\ldots=x_n=0$ представлява решение на системата. С изключение на това решение на всяка хомогенна система ние ще казваме, че уравненията (2) имат решение, ако поне едно от числата $x_1, x_2,\ldots, x_n$, които го удовлетворяват, е отлично от нула. Да допуснем, че $x_n\neq 0$. Тогава, следвайки начина в по-раншния параграф, получаваме лесно

$$Ax_n=0,$$

отгдето $A=0$. Следователно, за да има системата (2) решение, трябва детерминантата A на системата да е равна на нула. Да разделим уравненията (2) с x_n и изхвърлим последното уравнение.

Така получаваме $n-1$ уравнения с $n-1$ неизвестни:

$$\frac{x_1}{x_n},\quad \frac{x_2}{x_n},\ldots,\frac{x_{n-1}}{x_n}.$$

$$(3)\qquad \begin{cases} a_{11}\dfrac{x_1}{x_n}+a_{12}\dfrac{x_2}{x_n}+\ldots+a_{1,n-1}\dfrac{x_{n-1}}{x_n}=-a_{1n},\\ a_{21}\dfrac{x_1}{x_n}+a_{22}\dfrac{x_2}{x_n}+\ldots+a_{2,n-1}\dfrac{x_{n-1}}{x_n}=-a_{2n},\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{n-1,1}\dfrac{x_1}{x_n}+a_{n-1,2}\dfrac{x_2}{x_n}+\ldots+a_{n-1,n-1}\dfrac{x_{n-1}}{x_n}=-a_{n-1,n}. \end{cases}$$

Ако $A_n\neq 0$, решена, системата дава

$$A_{nn}\frac{x_k}{x_n}=-\begin{vmatrix} a_{11}\ldots a_{1,k-1} & a_{1n} & a_{1,k+1}\ldots a_{1,n-1}\\ a_{21}\ldots a_{2,k-1} & a_{2n} & a_{2,k+1}\ldots a_{2,n-1}\\ \ldots\ldots\ldots\ldots & \ldots & \ldots\ldots\ldots\\ a_{n-1,1}\ldots a_{n-1,k-1} & a_{n-1,n} & a_{n-1,k+1}\ldots a_{n-1,n-1} \end{vmatrix}$$

и като преместим k-ия стълб на последно място, получаваме

$$A_{nn}\frac{x_k}{x_n}=(-1)^{n-k}\begin{vmatrix} a_{11} & a_{12}\ldots & a_{1,k-1} & a_{1,k+1}\ldots a_{1n} \\ a_{21} & a_{22}\ldots & a_{2,k-1} & a_{2,k+1}\ldots a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2}\ldots & a_{n-1,k-1} & a_{n-1,k+1}\ldots a_{n-1,n} \end{vmatrix}.$$

Следва $\frac{x_k}{x_n}=\frac{A_{nk}}{A_{nn}}$, или

$$\frac{x_1}{A_{n1}}=\frac{x_2}{A_{n2}}=\frac{x_3}{A_{n3}}=\ldots=\frac{x_n}{A_{nn}},$$

т. е. неизвестните са пропорционални на адюнгираните количества на един ред.

От горното следва, че когато една детерминанта е равна на нула, то адюнгираните количества на два паралелни реда са пропорционални, което по-строго ще установим по-нататък.

Да разгледаме сега n уравнения с $n-1$ неизвестни:

$$(4)\quad \begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1,n-1}x_{n-1}+a_{1n}=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2,n-1}x_{n-1}+a_{2n}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{n,n-1}x_{n-1}+a_{nn}=0. \end{cases}$$

Тази система се получава от (3), като поставим $x_n=1$. Оттук или аналогично с умножаване следва веднага, че за да има системата (4) решение, трябва детерминантата, образувана от коефициентите и свободните членове, да бъде равна на нула.

3. **Обобщение. Теорема на Руше.** Сега ще си поставим за задача да изследваме най-общо решението на m линейни уравнения с n неизвестни:

$$(5)\quad \begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=a_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=a_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=a_m. \end{cases}$$

Преди обаче да формулираме предложението на Руше, ще се запознаем с някои дефиниции. Видяхме в началото какво наричаме матрица. Нека A е една матрица:

$$(A)=\begin{Vmatrix} a_{11} & a_{12}\ldots & a_{1n} \\ a_{21} & a_{22}\ldots & a_{2n} \\ a_{31} & a_{32}\ldots & a_{3n} \\ \ldots & \ldots & \ldots \\ a_{m1} & a_{m2}\ldots & a_{mn} \end{Vmatrix}.$$

Ако μ е число най-много равно на по-малкото от двете числа n и m, то в (A) да вземем μ произволни реда и стълба. От елементите на (A), в които се пресичат тези редове и стълбове, да образуваме една детерминанта от ред μ. Казваме, че тази детерминанта принадлежи на (A). Ако има поне една детерминанта от ред r, принадлежаща на (A), отлична от нула, а всички други детерминанти от по-висок от r-ти ред са равни на нула, то казваме, че матрицата (A) е от ранг r. Очевидно можем да предположим само, че детерминантите от ред $(r+1)$ са нули, защото всяка детерминанта от ред $(r+2)$ се развива по детерминанти от $(r+1)$-ви ред и е нула и т. н. Така например матрицата

$$\left\| \begin{matrix} 1 & 1 & 2 & 2 \\ 2 & 3 & 4 & 5 \\ 3 & 5 & 6 & 8 \end{matrix} \right\|$$

е от ранг две, понеже $\begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 1 \neq 0$, а всички детерминанти от трети ред са нули.

Очевидно е, че ако матрицата (B) е получена от (A) с разместване на паралелни редове, то те са от един и същ ранг. Също, ако към елементите на един ред прибавим съответните им елементи на друг, паралелен нему ред, умножени с едно произволно число, то рангът не се изменя. Така рангът на двете матрици

$$\left\| \begin{matrix} a_{11}+qa_{1k} & a_{12} \ldots a_{1n} \\ a_{21}+qa_{2k} & a_{22} \ldots a_{2n} \\ \cdot & \cdot \quad \ldots \\ \cdot & \cdot \quad \ldots \\ a_{m1}+qa_{mk} & a_{m2} \ldots a_{mn} \end{matrix} \right\}, \qquad \left\| \begin{matrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \cdot & \cdot \quad \ldots \\ \cdot & \cdot \quad \ldots \\ a_{m1} & a_{m2} \ldots a_{mn} \end{matrix} \right\|$$

е един и същ, в което лесно се убеждаваме на основание на елементарните свойства на детерминантите. Да означим сега с (A_0) матрицата

$$(A_0) = \left\| \begin{matrix} a_1 & a_{11} & a_{12} \ldots a_{1n} \\ a_2 & a_{21} & a_{22} \ldots a_{2n} \\ \cdot & \cdot & \cdot \quad \ldots \\ a_m & a_{m1} & a_{m2} \ldots a_{mn} \end{matrix} \right\|,$$

която получаваме, като към (A) прибавим свободните членове в уравненията (5). Имаме следната

Теорема на Руше. Необходимо и достатъчно условие системата (5) да има решение е матриците (A) и (A_0) да имат един и същ ранг r. Решенията ще зави-

сят от $n-r$ произволни параметъра, т. е. броят им ще бъде ∞^{n-r}.

Доказателство. Нека рангът на (A) е r. Тогава ще има поне една детерминанта от ред r, принадлежаща на (A), която е отлична от нула, а всички от ред $(r+1)$ са нули. С преместване на уравненията и означенията на неизвестните можем да допуснем, че

$$\theta=\begin{vmatrix} a_{11} & a_{12} \ldots a_{1r} \\ a_{21} & a_{22} \ldots a_{2r} \\ \ldots\ldots\ldots \\ a_{r1} & a_{r2} \ldots a_{rr} \end{vmatrix}$$

е отлична от нула. Да означим с

$$f_i=a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n,\quad i=1,\ 2,\ 3,\ \ldots,m \tag{6}$$

и да образуваме детерминантите от $(r+1)$-ви ред:

$$T_\alpha=\begin{vmatrix} f_\alpha & a_{\alpha 1} & a_{\alpha 2} \ldots a_{\alpha r} \\ f_1 & a_{11} & a_{12} \ldots a_{1r} \\ f_2 & a_{21} & a_{22} \ldots a_{2r} \\ \ldots\ldots\ldots\ldots \\ f_r & a_{r1} & a_{r2} \ldots a_r \end{vmatrix},\quad \alpha=r+1,\ r+2,\ldots,m$$

Ще докажем, че $T_\alpha=0$, каквито и да бъдат $x_1,\ x_2,\ldots,x_n$. Действително, като заменим $f_\alpha,\ f_1,\ldots,f_r$ с равните им по (6), получаваме

$$T_\alpha=x_1\begin{vmatrix} a_{\alpha 1} & a_{\alpha 1} & a_{\alpha 2},\ldots,a_{\alpha r} \\ a_{11} & a_{11} & a_{12} \ldots a_{1r} \\ a_{21} & a_{21} & a_{22} \ldots a_{2r} \\ \ldots\ldots\ldots\ldots \\ a_{r1} & a_{r1} & a_{r2} \ldots a_{rr} \end{vmatrix}+\ldots+x_r\begin{vmatrix} a_{\alpha r} & a_{\alpha 1} & a_{\alpha 2} \ldots a_{\alpha r} \\ a_{1r} & a_{11} & a_{12} \ldots a_{1r} \\ a_{2r} & a_{21} & a_{22} \ldots a_{2r} \\ \ldots\ldots\ldots\ldots \\ a_{rr} & a_{r1} & a_{r2} \ldots a_{rr} \end{vmatrix}+$$

$$+x_{r+1}\begin{vmatrix} a_{a,r+1} & a_{\alpha 1} & a_{\alpha 2} \ldots a_{\alpha r} \\ a_{1,r+1} & a_{11} & a_{12} \ldots a_{1r} \\ a_{2,r+1} & a_{21} & a_{22} \ldots a_{2r} \\ \ldots\ldots\ldots\ldots \\ a_{r,r+1} & a_{r1} & a_{r2} \ldots a_{rr} \end{vmatrix}+\ldots+x_n\begin{vmatrix} a_{\alpha n} & a_{\alpha 1} & a_{\alpha 2} \ldots a_{\alpha r} \\ a_{1n} & a_{11} & a_{12} \ldots a_{1r} \\ a_{2n} & a_{21} & a_{22} \ldots a_{2r} \\ \ldots\ldots\ldots\ldots \\ a_{rn} & a_{r1} & a_{r2} \ldots a_{rr} \end{vmatrix},$$

в което развитие коефициентите на $x_1,\ x_2,\ldots,x_r$ са нули, понеже са детерминанти с два равни паралелни стълба, а коефициентите на $x_{r+1},\ldots,x_n$

са детерминанти от ред $(r+1)$, принадлежащи на (A), следователно са равни на нула. Ако сега $\theta_{\alpha 1}, \theta_{\alpha 2}, \ldots, \theta_{\alpha r}$ са адюнгираните количества на $f_1, f_2, \ldots, f_r$ в T_α, то ще имаме

$$f_\alpha\theta+f_1\theta_{\alpha 1}+f_2\theta_{\alpha 2}+\ldots+f_r\theta_{\alpha r}=0,$$

отгдето, понеже $\theta\neq 0$, получаваме

$$f_\alpha=-\frac{\theta_{\alpha 1}}{\theta}f_1-\frac{\theta_{\alpha 2}}{\theta}f_2-\ldots-\frac{\theta_{\alpha r}}{\theta}f_r. \tag{7}$$

От израза (7) се вижда, че функциите f_i, $i=r+1, \ldots, m$ са линейни хомогенни функции на $f_1, f_2, \ldots, f_r$.

Да допуснем, че системата (5) има решение. Тогава ще докажем, че рангът на (A_0) е r. Действително рангът на (A_0) не може да бъде по-малък от r, понеже матрицата (A) е част от (A_0). За да докажем нашето твърдение, трябва да установим, че всички детерминанти от ред $(r+1)$, принадлежащи на (A_0), са равни на нула. Ако такава детерминанта няма елементи от първия стълб на (A_0), то тя принадлежи на (A) и следователно е равна на нула. Нека тя съдържа елементи от първия стълб, т. е. да има форма

$$\alpha=\begin{vmatrix} a_{i_1} & a_{i_1 s_1} \ldots a_{i_1 s_r} \\ a_{i_2} & a_{i_2 s_1} \ldots a_{i_2 s_r} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{i_r+1} & a_{i_r+1\, s_1} \ldots a_{i_r+1 s_r} \end{vmatrix},$$

гдето $i_1, i_2, \ldots, i_{r+1}$ са $(r+1)$ числа от 1, 2, 3, ..., m, а $s_1, s_2, \ldots, s_r$ са r числа от 1, 2, 3, ..., n. Аналогично на $T_i=0$ се доказва, че

$$\alpha_f=\begin{vmatrix} f_{i_1} & a_{i_1 s_1} \ldots a_{i_1 s_r} \\ f_{i_2} & a_{i_2 s_1} \ldots a_{i_2 sr} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \\ f_{i_r+1} & a_{i_r+1 s_1} \ldots a_{i_r+1 s_r} \end{vmatrix}=0,$$

за всички $x_1, x_2, \ldots, x_n$. Ако системата (5), която може да се пише

$$f_i=a_i,\ i=1,\ 2, \ldots,\ m,$$

има решение, то като поставим в α_f значенията на $x_1, x_2, \ldots, x_n$, понеже f_i се обръща в a_i, α_f се обръща в α, с което е доказано, че $\alpha=0$, т. е. рангът на (A_0) е r. Обратно, да допуснем, че (A) и (A_0) имат един и същ ранг r и него $\theta\neq 0$ (което, както видяхме, можем винаги да допуснем). Тогава освен $T_\alpha=0$ ще имаме

$$S_\alpha=\begin{vmatrix} a_\alpha & a_{\alpha 1} \ldots a_{\alpha r} \\ a_1 & a_{11} \ldots a_{1r} \\ \cdot & \cdot\ \cdot\ \cdot \\ a_r & a_{r1} \ldots a_{rr} \end{vmatrix}=0,$$

като детерминанта от ред $(r+1)$ и принадлежаща на (A_0). Оттук

$$T_\alpha - S_\alpha = \begin{vmatrix} f_\alpha - a_\alpha & a_{\alpha 1} \ldots a_{\alpha r} \\ f_1 - a_1 & a_{11} \ldots a_{1r} \\ f_2 - a_2 & a_{21} \ldots a_{2r} \\ \cdots\cdots\cdots\cdots \\ f_r - a_r & a_{r1} \ldots a_{rr} \end{vmatrix} = 0.$$

Като развием по първия стълб, получаваме

$$(8) \qquad f_\alpha - a_\alpha =$$

$$= -\frac{\theta_{\alpha 1}}{\theta}(f_1 - a_1) - \frac{\theta_{\alpha 2}}{\theta}(f_2 - a_2) - \ldots - \frac{\theta_{\alpha r}}{\theta}(f_r - a_r),$$

гдето $\alpha = r+1, r+2, \ldots, m$. Ако някои стойности на $x_1, x_2, \ldots, x_n$ удовлетворяват уравненията $f_1 = a_1, f_2 = a_2, \ldots, f_r = a_r$, то на основание на (8) те ще удовлетворяват и

$$f_{r+1} = a_{r+1}, f_{r+2} = a_{r+2}, \ldots, f_m = a_m,$$

т. е. система (5). Но уравненията

$$f_i = a_i, \quad i = 1, 2, 3, \ldots, r,$$

могат да се напишат в следната форма:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = a_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = a_2 - a_{2,r+1}x_{r+1} - \ldots - a_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = a_r - a_{r,r+1}x_{r+1} - \ldots - a_{rn}x_n. \end{cases}$$

Понеже детерминантата $\theta \neq 0$, то ако A_{ik} означава адюнгираното количество на a_{ik} в θ, ще получим

$$\theta x_\alpha = (a_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n)A_{1\alpha} + \ldots +$$

$$+ (a_r - a_{r,r+1}x_{r+1} - \ldots - a_{rn}x_n)A_{r\alpha},$$

гдето $\alpha = 1, 2, 3, \ldots, r$. Оттук е очевидно, че системата (5) има решения, зависещи от $(n-r)$ произволни параметъра $x_{r+1}, x_{r+2}, \ldots, x_n$, т. е. има ∞^{n-r} решения.

На решението на системата (5) може да се даде една по-симетрична форма. Да изразим $x_{r+1}, \ldots, x_n$ посредством нови параметри $q_1, q_2, \ldots, q_{n-r}$,

$$(9) \qquad \begin{cases} x_{r+1} = b_{r+1,0} + b_{r+1,1}q_1 + \ldots + b_{r+1,n-r}q_{n-r}, \\ x_{r+2} = b_{r+2,0} + b_{r+2,1}q_1 + \ldots + b_{r+2,n-r}q_{n-r}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = b_{n0} + b_{n1}q_1 + \ldots + b_{n,n-r}q_{n-r}. \end{cases}$$

Тези формули трябва да са такива, че уравненията да могат да се решат спрямо $q_1, q_2, \ldots, q_{n-r}$, понеже $x_{r+1}, \ldots, x_n$ могат да вземат произволни стойности. За тази цел трябва детерминантата

$$\begin{vmatrix} b_{r+1,1} & \ldots & b_{r+1,n-r} \\ b_{r+2,1} & \ldots & b_{r+2,n-r} \\ \ldots & \ldots & \ldots \\ b_{n,1} & \ldots & b_{n,n-r} \end{vmatrix}$$

да бъде отлична от нула.

Тогава решенията приемат следната симетрична форма:

$$x_\alpha = b_{\alpha 0} + b_{\alpha 1} q_1 + b_{\alpha 2} q_2 + \ldots + b_{\alpha,n-r} q_{n-r},$$
$$\alpha = 1, 2, 3, \ldots, n.$$

Примери. Нека е дадена системата

$$\begin{aligned} f_1 &\equiv 3x + 4y + 5z + t - 2u = 3, \\ f_2 &\equiv x - y + 3z + 2t + 3u = 5, \\ f_3 &\equiv 2x + 3y + z - t + u = 2, \\ f_4 &\equiv 6x + 6y + 9z + 2t + 2u = 10, \\ f_5 &\equiv 7x + 12y + 8z - t - 6u = 3. \end{aligned} \tag{10}$$

Матриците (A) и (A_0) са следните:

$$(A) = \begin{Vmatrix} 3 & 4 & 5 & 1 & -2 \\ 1 & -1 & 3 & 2 & 3 \\ 2 & 3 & 1 & -1 & 1 \\ 6 & 6 & 9 & 2 & 2 \\ 7 & 12 & 8 & -1 & -6 \end{Vmatrix}, \quad (A_0) = \begin{Vmatrix} 3 & 3 & 4 & 5 & 1 & -2 \\ 5 & 1 & -1 & 3 & 2 & 3 \\ 2 & 2 & 3 & 1 & -1 & 1 \\ 10 & 6 & 6 & 9 & 2 & 2 \\ 3 & 7 & 12 & 8 & -1 & -6 \end{Vmatrix},$$

за които лесно се убеждаваме, че са от трети ранг. Следователно системата има решения, които ще зависят от два произволни параметъра. Понеже детерминантата

$$\theta = \begin{vmatrix} 3 & 4 & 5 \\ 1 & -1 & 3 \\ 2 & 3 & 1 \end{vmatrix} = 16 \neq 0,$$

то f_4, f_5 се изразяват посредством f_1, f_2, f_3. Формулите, които получихме по-горе, ни дават

$$f_4=f_1+f_2+f_3,\quad f_5=2f_1-f_2+f_3$$

и

$$f_4-10=(f_1-3)+(f_2-5)+(f_3-2),$$
$$f_5-3=2(f_1-3)-(f_2-5)+(f_3-2).$$

Така че решението на системата (10) се свежда към това на системата

$$\begin{aligned} 3x+4y+5z&=3-t+2u,\\ x-y+3z&=5-2t+3u,\\ 2x+3y+z&=2+t-u \end{aligned}$$

спрямо x, y, z. Решенията ще зависят от два произволни параметъра t и u. За x, y, z получаваме

$$\begin{aligned} 15x&=59+5t-70u,\\ 15y&=-28+5t+35u,\\ 15z&=-4-10t+20u. \end{aligned}$$

Ако искаме да приложим формулите (9), нека поставим

$$\begin{aligned} t&=1+3p+q,\\ u&=1+3p-q, \end{aligned}\qquad \begin{vmatrix} 3 & 1\\ 3 & -1 \end{vmatrix}=-6\neq 0,$$

то ще получим

$$\begin{cases} x=-\frac{2}{5}-13p+5q,\\ y=\frac{4}{5}+8p-2q,\\ z=\frac{2}{5}+2p-2q,\\ t=1+3p+q,\\ u=1+3p-q. \end{cases}$$

Друг пример. Нека разгледаме три равнини

$$\begin{cases} a_1x+b_1y+c_1z=d_1,\\ a_2x+b_2y+c_2z=d_2,\\ a_3x+b_3y+c_3z=d_3 \end{cases}$$

и да изучим пресечните им точки.

За тази система имаме

$$(A)=\left\|\begin{array}{ccc} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array}\right\|, \qquad (A_0)=\left|\begin{array}{cccc} d_1 & a_1 & b_1 & c_1 \\ d_2 & a_2 & b_2 & c_2 \\ d_3 & a_3 & b_3 & c_3 \end{array}\right|.$$

Да означим с δ, α, β, γ детерминантите от трети ред, получени от матрицата (A_0) с отстраняване съответно на първия, втория, третия и четвъртия стълб: Ако $\delta \neq 0$, (A) и (A_0) са от трети ранг, равнините имат една обща пресечна точка. Ако $\delta=0$, но поне една от детерминантите α, β, γ е отлична от нула, то уравненията (11) нямат решение, т. е. трите равнини нямат обща точка на крайно разстояние. Ако $\alpha=\beta=\gamma=\delta=0$, но има поне една детерминанта от втори ред в (A), отлична от нула, то x, y, z ще бъдат линейни функции на един параметър, т. е. трите равнини се пресичат в една права. Ако обаче няма такава детерминанта в (A), а рангът на (A_0) е две, то общата права отива в ∞, т. е. трите равнини са успоредни. Най-сетне, ако (A) и (A_0) са от първи ранг, то решенията зависят от два параметъра, т. е. трите равнини се сливат.

Да разгледаме частния случай на хомогенни уравнения:

$$\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=0. \end{array} \tag{11}$$

Очевидно в този случай матрицата (A_0) ще има същия ранг на матрицата (A). Следователно системата (11) има винаги решение. Ако рангът на матрицата (A) е равен на r, то решенията ще зависят от $n-r$ произволни параметъра. При $r=n$ единственото решение ще бъде нулевото $x_1=x_2=\ldots=x_n=0$.

Нека рангът на (A) е r. С разместване на уравненията и неизвестните можем да приемем, че детерминантата

$$\theta=\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \ldots & \ldots & \ldots & \ldots \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix}$$

е отлична от нула. Да означим с f_i линейните функции

$$f_i=a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n, \quad 1 \leqq i \leqq m$$

и нека предположим, че между функциите $f_1, f_2, \ldots, f_r$ имаме линейната връзка (за всички стойности на променливите)

$$\lambda_1 f_1+\lambda_2 f_2+\ldots+\lambda_r f_r=0, \tag{12}$$

където $\lambda_1, \lambda_2, \ldots, \lambda_r$ са числа. Като приравним към нула коефициентите на $x_1, x_2, x_3, \ldots$ в лявата част на (12), получаваме

$$\begin{aligned} a_{11}\lambda_1 + a_{21}\lambda_2 + \ldots + a_{r1}\lambda_r &= 0, \\ a_{12}\lambda_1 + a_{22}\lambda_2 + \ldots + a_{r2}\lambda_r &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{1r}\lambda_1 + a_{2r}\lambda_2 + \ldots + a_{rr}\lambda_r &= 0. \end{aligned}$$

Понеже детерминантата от коефициентите на неизвестните $\lambda_1, \lambda_2, \ldots, \lambda_r$ е отлична от нула, то следва, че $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$. Следователно между линейните форми $f_1, f_2, \ldots, f_r$ не може да има връзка от вида (12), гдето поне едно от числата $\lambda_1, \lambda_2, \ldots, \lambda_r$ е отлично от нула. Казваме в случая, че тези линейни форми са линейно независими. От формулата (7) виждаме, че всичките форми f_i, $1 \leqq i \leqq m$ са линейни комбинации на формите $f_1, f_2, \ldots, f_r$. Така установихме, че ако матрицата (A) е от ранг r, то има r линейно независими форми $f_1, f_2, \ldots, f_r$ и всяка линейна форма f_i е линейна комбинация на тези r форми.

Глава III

Други свойства

1. **Миньори.** Детерминантите, принадлежащи на матрицата на една детерминанта Δ:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots\ldots\ldots & \\ a_{n1} & a_{n2} \ldots a_{nn} \end{vmatrix},$$

се наричат миньори или поддетерминанти. Общата форма на един миньор от ред p ще бъде

$$\Delta_1 = \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} \ldots a_{i_1 j_p} \\ a_{i_2 j_1} & a_{i_2 j_2} \ldots a_{i_2 j_p} \\ \ldots\ldots\ldots\ldots & \\ a_{i_p j_1} & a_{i_p j_2} \ldots a_{i_p j_p} \end{vmatrix},$$

гдето $i_1 < i_2 < \ldots < i_p$; $j_1 < j_2 < \ldots < j_p$ са числа от 1, 2, 3, ..., n. Ако $i_1 = j_1, i_2 = j_2, \ldots, i_p = j_p$, миньорът се нарича главен. Два миньора се наричат конюговани, ако хоризонталните редове на единия имат същите индекси, както вертикалните редове на другия и обратно. Всеки главен миньор е конюгован на себе си.

Можем да получим миньора Δ_1, като в Δ зачеркнем $n-p$ реда и толкова стълба. Ако сега в Δ зачеркнем хоризонталните редове и

стълбове, които влизат в Δ_1, то така ще получим нов миньор Δ_2 от ред $n-p$, който наричаме адюнгиран на първия. Ако с

$$s_1<s_2<\ldots<s_{n-p}$$

означим числата, които остават в редицата $1, 2, 3, \ldots, n$, като сме премахнали $i_1, i_2, \ldots, i_p$, също с $u_1<u_2<\ldots<u_{n-p}$ числата от $1, 2, \ldots, n$, които остават след премахването на $j_1, j_2, \ldots, j_p$, то Δ_2 е миньорът, който има за индекси на хоризонталните си редове числата s, а за индекси на стълбовете си — числата u. Ако $\sigma=i_1+i_2+\ldots+i_p+j_1+j_2+\ldots+j_p$, то $(-1)^{\sigma}\Delta_2$ се нарича адюнгирано количество на Δ_1. Ако σ' е подобно число за Δ_2, то $(-1)^{\sigma'}\Delta_1$ е адюнгираното количество на Δ_2. Но $\sigma+\sigma'$ е сумата на индексите на всички хоризонтални и вертикални редове на Δ, така че

$$\sigma+\sigma'=2(1+2+\ldots+n).$$

Следователно σ и σ' са от еднаква четност, т. е. $(-1)^{\sigma}=(-1)^{\sigma'}$, отгдето имаме: адюнгираното количество на Δ_2 е Δ_1 или — Δ_1 според това, дали адюнгираното количество на Δ_1 е Δ_2, или — Δ_2.

2. **Правило на Лаплас.** С помощта на миньорите ще получим едно правило, дадено от Лаплас, за развитие на детерминантите по няколко реда.

Лема. Произведението от един миньор с неговото адюнгирано количество е една част от развитието на детерминантата Δ.

Произведението на всеки член от Δ_1 с всеки член от Δ_2, като не обръщаме внимание на знака, принадлежи на Δ, понеже съдържа елементи от различни стълбове и редове. Първите индекси в това произведение образуват пермутация на $i_1, i_2, \ldots, i_p, s_1, s_2, \ldots, s_{n-p}$, а вторите на $j_1, j_2, \ldots, j_p, u_1, u_2, \ldots, u_{n-p}$. Ако сега първите две пермутации от i и j имат α и β инверзии, а вторите γ и δ, то знакът на този член в произведението $(-1)^{\sigma}\Delta_1.\Delta_2$ се дават от

$$(-1)^{\alpha+\beta+\gamma+\delta+\sigma}.$$

Ако сега с ε означим броя на инверзиите между $i_1, i_2, \ldots, i_p$ и $s_1, s_2, \ldots, s_{n-p}$, който брой не се изменя очевидно при разместването на $i_1 \ldots i_p$ помежду им, а с η съответния брой на инверзиите между $j_1, j_2, \ldots, j_p$ и $u_1, u_2, \ldots, u_{n-p}$, то знакът на този член в Δ е

$$(-1)^{\alpha+\beta+\gamma+\delta+\varepsilon+\eta}.$$

Сега ще определим ε и η. Понеже $i_1, i_2, \ldots, i_p$ са наредени по растяща големина, то i_r може да образува инверзии само с числата $1, 2, 3, \ldots, i_{r-1}$. От тези обаче елементи трябва да махнем числата $i_1, i_2, \ldots, i_{r-1}$, които принадлежат на първата пермутация. Следователно i_r образува с $s_1, s_2, \ldots, s_{n-p}$

$$i_r-1-(r-1)=i_r-r$$

инверзии. Оттук получаваме

$$\varepsilon=(i_1-1)+(i_2-2)+\ldots+(i_p-p)=$$

$$=i_1+i_2+\ldots+i_p-\frac{1}{2}p(p+1),$$

$$\eta=(j_1-1)+(j_2-2)+\ldots+(j_p-p)=$$

$$=j_1+j_2+\ldots+j_p-\frac{1}{2}p(p+1),$$

които дават

$$\varepsilon+\eta=\sigma-p(p+1).$$

Понеже числото $p(p+1)$ е четно, то

$$(-1)^{\alpha+\beta+\gamma+\delta+\varepsilon+\eta}=(-1)^{\alpha+\beta+\gamma+\delta+\sigma},$$

с което лемата е доказана.

Теорема на Лаплас. Всяка детерминанта е равна на сумата на всички миньори от ν-ти ред, съдържащи се в ν-паралелни реда, умножени със съответните им адюнгирани количества. От лемата се вижда, че всички така получени членове принадлежат на детерминантата. Освен това очевидно е, че те са все различни. Следователно, ако установим, че броят им е $n!$, с това се доказва теоремата на Лаплас. Броят на миньорите от ν-ти ред, които се съдържат в ν паралелни реда, е точно равен на броя на комбинациите на n по ν, т. е. на $\binom{n}{\nu}$. Всеки миньор има $\nu!$ члена, а съответно адюнгирано количество (което е, като не обръщаме внимание на знака, миньор от $n-\nu$-ти ред) има $(n-\nu)!$ члена, така че се получават

$$\nu!\,(n-\nu)!\binom{n}{\nu}=n!$$

члена.

Така да разгледаме детерминантата

$$\begin{vmatrix} a & a_1 & a_2 & a_3 \\ b & b_1 & b_2 & b_3 \\ c & c_1 & c_2 & c_3 \\ d & d_1 & d_2 & d_3 \end{vmatrix}$$

и да положим $p=2$, $i_1=1$, $i_2=2$; имаме .

$$\Delta=\begin{vmatrix} a & a_1 \\ b & b_1 \end{vmatrix}\cdot\begin{vmatrix} c_2 & c_3 \\ d_2 & d_3 \end{vmatrix}-\begin{vmatrix} a & a_2 \\ b & b_2 \end{vmatrix}\cdot\begin{vmatrix} c_1 & c_3 \\ d_1 & d_3 \end{vmatrix}+\begin{vmatrix} a & a_3 \\ b & b_3 \end{vmatrix}\cdot\begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix}+$$

$$+\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\cdot\begin{vmatrix} c & c_3 \\ d & d_3 \end{vmatrix}-\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\cdot\begin{vmatrix} c & c_2 \\ d & d_2 \end{vmatrix}+\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\cdot\begin{vmatrix} c & c_1 \\ d & d_1 \end{vmatrix}.$$

За $i_1=2,\ i_2=4$ имаме

$$\Delta=-\begin{vmatrix} b & b_1 \\ d & d_1 \end{vmatrix}\cdot\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}+\begin{vmatrix} b & b_2 \\ d & d_2 \end{vmatrix}\cdot\begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix}-\begin{vmatrix} b & b_3 \\ d & d_3 \end{vmatrix}\cdot\begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}-$$

$$-\begin{vmatrix} b_1 & b_2 \\ d_1 & d_2 \end{vmatrix}\cdot\begin{vmatrix} a & a_3 \\ c & c_3 \end{vmatrix}+\begin{vmatrix} b_1 & b_3 \\ d_1 & d_3 \end{vmatrix}\cdot\begin{vmatrix} a & a_2 \\ c & c_2 \end{vmatrix}-\begin{vmatrix} b_2 & b_3 \\ d_2 & d_3 \end{vmatrix}\cdot\begin{vmatrix} a & a_1 \\ c & c_1 \end{vmatrix}.$$

Детерминантата D наричаме **изменена по контура**, ако тя се получава от една детерминанта A от по-нисък ред с прибавяне на едно определено число хоризонтални редове и толкова стълбове. Ние ще разгледаме само тези, които се получават с прибавяне на един стълб и един хоризонтален ред, т. е. които имат формата

$$D=\begin{vmatrix} a_{11} & a_{12}\ldots a_{1n} & y_1 \\ a_{21} & a_{22}\ldots a_{2n} & y_2 \\ \ldots\ldots\ldots & & \\ a_{n1} & a_{n2}\ldots a_{nn} & y_n \\ x_1 & x_2\ldots x_n & z \end{vmatrix},$$

$$A=\begin{vmatrix} a_{11} & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}\ldots a_{2n} \\ \ldots\ldots\ldots & \\ a_{n1} & a_{n2}\ldots a_{nn} \end{vmatrix}.$$

Ние ще развием D по произведенията $x_i\,y_j$. Ясно е, че в развитието на D членовете, които съдържат z, ще бъдат Az, а за да видим какъв ще бъде коефициентът на $x_i\,y_j$, да разгледаме миньора

$$\begin{vmatrix} a_{ij} & y_i \\ x_j & z \end{vmatrix}.$$

Ако използуваме по-раншния резултат, то адюнгираният миньор на този в D се получава, като в матрицата на тази детерминанта зачеркнем i-ия и $(n+1)$-ия ред и j-ия и $(n+1)$-ия стълб, а това е точно поддетерминантата Δ_{ij} на a_{ij} в A. От получената по-горе лема се вижда, че знакът на произведението

$$\text{(1)}\qquad \begin{vmatrix} a_{ij} & y_i \\ x_j & z \end{vmatrix}\Delta_{ij}$$

в развитието на детерминантата D ще бъде

$$(-1)^{i+n+1+j+n+1}=(-1)^{i+j}.$$

Очевидно е обаче, че произведението x_jx_i ще се среща само в произведението (1), т. е. ще има коефициент

$$-(-1)^{i+j}\Delta_{ij}.$$

Като вземем под внимание, че в развитието на D всеки член ще съдържа или множител z, или едно произведение $x_j\, y_j$, то получаваме следното развитие (Коши):

$$D=Az-\sum(-1)^{i+j}\, x_j\, y_i\, \Delta_{ij},$$

гдето $i,\ j=1,\ 2,\ldots,n$.

По същия начин можем да получим развитие на такива детерминанти, получени с прибавяне на повече от един хоризонтален ред и стълб.

3. **Умножение на детерминанти.** Ще покажем, че произведението на две детерминанти може да се представи пак като детерминанта.

Преди това ще установим една помощна теорема. Нека матрицата на една детерминанта Δ се състои от четири матрици, от които две са квадратни, а едната от другите две се състои само от нули. Ще докажем, че Δ е равна на произведението на две детерминанти.

По-точно, ако Δ има форма

$$\Delta=\begin{vmatrix} a_{11}\ a_{12}\ldots a_{1n} & 0\ \ 0\ldots 0 \\ a_{21}\ a_{22}\ldots a_{2n} & 0\ \ 0\ldots 0 \\ \ldots\ldots\ldots\ldots & \ldots\ldots \\ a_{n1}\ a_{n2}\ldots a_{nn} & 0\ \ 0\ldots 0 \\ \ldots\ldots & b_{11}\ b_{12}\ldots b_{1m} \\ \ldots\ldots & b_{21}\ b_{22}\ldots b_{2m} \\ \ldots\ldots\ldots\ldots & \ldots\ldots \\ \ldots\ldots & b_{m1}\ b_{m2}\ldots b_{mm} \end{vmatrix},$$

то имаме

$$\Delta=\begin{vmatrix} a_{11}\ a_{12}\ldots a_{1n} \\ a_{21}\ a_{22}\ldots a_{2n} \\ \ldots\ldots\ldots \\ a_{n1}\ a_{n2}\ldots a_{nn} \end{vmatrix}\cdot\begin{vmatrix} b_{11}\ b_{12}\ldots b_{1m} \\ b_{21}\ b_{22}\ldots b_{2m} \\ \ldots\ldots\ldots \\ b_{m1}\ b_{m2}\ldots b_{mm} \end{vmatrix}.$$

Доказателство. В Δ означаваме всички елементи с a_{lk}, гдето $i,\ k=1,\ 2,\ldots,n,\ n+1,\ldots,n+m$. Тогава ще имаме

$$a_{ik}=0,\quad i\leqq n,\quad k>n,$$

$$a_{ik}=b_{i-n,k-n},\quad i>n,\quad k>n.$$

А за Δ имаме

$$\Delta=\sum(-1)^{[\alpha,\ \beta,\ \ldots,\ \delta,\ \mu,\ \nu,\ \ldots,\ \tau]}\, a_{1\alpha}\, a_{2\beta}\ldots a_{n\delta}\, a_{n+1,\mu}\ldots a_{n+m,\tau}.$$

Всички членове, в които поне едно от числата $\alpha,\ \beta,\ldots,\ \delta$ е по-голямо от n, ще са равни на нула. Следователно можем да предполагаме, че $\alpha,\ \beta,\ldots,\ \delta$ са числата $1,\ 2,\ldots,\ n$. Тогава $\mu,\ \nu,\ldots,\ \tau$ ще са числата $n+1,\ldots,n+m$ в някой ред. Но тогава, ако означим с

$$\mu-n=g,\quad \nu-n=h,\ldots,\ \tau-n=e,$$

понеже

$$[\alpha, \beta, \ldots, \delta, \mu, \nu, \ldots, \tau]=[\alpha, \beta, \ldots, \delta]+[\mu, \nu, \ldots, \tau],$$

$$a_{n+1,\mu}=b_{1g},\ a_{n+2,\nu}=b_{2h}, \ldots,\ a_{n+m,\tau}=b_{me},$$

то

$$\Delta=\sum(-1)^{[\alpha, \beta, \ldots, \delta]}a_{1\alpha}\,a_{2\beta}\ldots a_{n\delta}\sum(-1)^{[g, h, \ldots, e]}\,b_{1g}\,b_{2h}\ldots b_{me},$$

гдето първата сума е разпростряна върху всички пермутации на 1, 2, . . ., n, а втората — върху всички пермутации на числата 1, 2,...,m, с което теоремата е доказана.

На основание на тази теорема ще изведем правилото за умножение на две детерминанти, което се състои в следното:

$$(2)\qquad \begin{vmatrix} a_{11}a_{12}\ldots a_{1n}\\ a_{21}a_{22}\ldots a_{2n}\\ \ldots\ldots\\ a_{n1}a_{n2}\ldots a_{nn}\end{vmatrix}\cdot\begin{vmatrix} b_{11}\,b_{12}\ldots b_{1n}\\ b_{21}\,b_{22}\ldots b_{2n}\\ \ldots\ldots\\ b_{n1}\,b_{n2}\ldots b_{nn}\end{vmatrix}=\begin{vmatrix} c_{11}\,c_{12}\ldots c_{1n}\\ c_{21}\,c_{22}\ldots c_{2n}\\ \ldots\ldots\\ c_{n1}\,c_{n2}\ldots c_{nn}\end{vmatrix},$$

гдето

$$(3)\qquad c_{pq}=a_{p1}\,b_{q1}+a_{p2}\,b_{q2}+\ \ldots\ +a_{pn}\,b_{qn},\ p,\ q=1,\ 2,\ \ldots,n.$$

За простота ще се ограничим на $n=3$, което благодарение на симетричността не представлява ограничение. Да разгледаме детерминантата

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0\\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0\\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0\\ -1 & 0 & 0 & b_{11} & b_{21} & b_{31}\\ 0 & -1 & 0 & b_{12} & b_{22} & b_{32}\\ 0 & 0 & -1 & b_{13} & b_{23} & b_{33}\end{vmatrix},$$

която е равна на произведението

$$\begin{vmatrix} a_{11}\ a_{12}\ a_{13}\\ a_{21}\ a_{22}\ a_{23}\\ a_{31}\ a_{32}\ a_{33}\end{vmatrix}\cdot\begin{vmatrix} b_{11}\ b_{12}\ b_{13}\\ b_{21}\ b_{22}\ b_{23}\\ b_{31}\ b_{32}\ b_{33}\end{vmatrix}.$$

В Δ умножаваме четвъртия ред с a_{11}, петия с a_{12}, шестия с a_{13} и ги прибавяме към първия ред. Също умножаваме четвъртия ред с a_{21}, петия с a_{22}, шестия с a_{23} и ги прибавяме към втория ред. Най-после умножаваме четвъртия ред с a_{31}, петия с a_{32}, шестия с a_{33} и ги прибавяме към третия. Така получаваме

$$\Delta=\begin{vmatrix} 0 & 0 & 0 & c_{11} & c_{12} & c_{13} \\ 0 & 0 & 0 & c_{21} & c_{22} & c_{23} \\ 0 & 0 & 0 & c_{31} & c_{32} & c_{33} \\ -1 & 0 & 0 & b_{11} & b_{21} & b_{31} \\ 0 & -1 & 0 & b_{12} & b_{22} & b_{32} \\ 0 & 0 & -1 & b_{13} & b_{23} & b_{33} \end{vmatrix},$$

гдето

$$c_{11}=a_{11}\,b_{11}+a_{12}\,b_{12}+a_{13}\,b_{13}$$

$$c_{12}=a_{11}\,b_{21}+a_{12}\,b_{22}+a_{13}\,b_{23}$$

$$c_{13}=a_{11}\,b_{31}+a_{12}\,b_{32}+a_{13}\,b_{33}$$

$$c_{21}=a_{21}\,b_{11}+a_{22}\,b_{12}+a_{23}\,b_{13}$$

$$c_{22}=a_{21}\,b_{21}+a_{22}\,b_{22}+a_{23}\,b_{23}$$

$$c_{23}=a_{21}\,b_{31}+a_{22}\,b_{32}+a_{23}\,b_{33}$$

$$c_{31}=a_{31}\,b_{11}+a_{32}\,b_{12}+a_{33}\,b_{13}$$

$$c_{32}=a_{31}\,b_{21}+a_{32}\,b_{22}+a_{33}\,b_{23}$$

$$c_{33}=a_{31}\,b_{31}+a_{32}\,b_{32}+a_{33}\,b_{33}$$

Преместваме четвъртия стълб на първо място, петия на второ, шестия на трето; знакът ще се промени съответно на $(-1)^{3^2}$. Тогава Δ е равна на произведението на детерминантата

$$\begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix}=(-1)^3$$

с детерминантата

$$\begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix}.$$

Понеже $(-1)^{3^2+3}=+1$, то формулата (3) се доказва при $n=3$. В общия случай доказателството става по същия начин, като се вземе пред вид, че $(-1)^{n^2+n}=+1$.

Умножението (3) доказахме, като вземем редове от първата и редове от втората детерминанта. Понеже една детерминанта не се изменя, като разменим стълбовете с хоризонталните редове и обратно, то получаваме, че произведението може да се представи по четири начина. Именно: първо — хоризонтални редове с такива от втората,

второ — хоризонтални редове със стълбове, трето — стълбове с хоризонтални редове и четвърто — стълбове със стълбове. Ще имаме формулата

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix} = \begin{vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{vmatrix},$$

гдето

I) $$c_{pq} = a_{p1}\,b_{q1} + a_{p2}\,b_{q2} + \dots + a_{pn}\,b_{qn} = \sum_{k=1}^{n} a_{pk}\,b_{qk};$$

II) $$c_{pq} = a_{p1}\,b_{1q} + a_{p2}\,b_{2q} + \dots + a_{pn}\,b_{nq} = \sum_{k=1}^{n} a_{pk}\,b_{kq};$$

III) $$c_{pq} = a_{1p}\,b_{q1} + a_{2p}\,b_{q2} + \dots + a_{np}\,b_{qn} = \sum_{k=1}^{n} a_{kp}\,b_{qk};$$

IV) $$c_{pq} = a_{1p}\,b_{1q} + a_{2p}\,b_{2q} + \dots + a_{np}\,b_{nq} = \sum_{k=1}^{n} a_{kp}\,b_{kq}.$$

4. Примери:

$$\begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix} \quad \begin{vmatrix} \alpha & \beta \\ \alpha_1 & \beta_1 \end{vmatrix} = \begin{vmatrix} a\alpha + b\beta & a\alpha_1 + b\beta_1 \\ a_1\alpha + b_1\beta & a_1\alpha_1 + b_1\beta_1 \end{vmatrix} = \begin{vmatrix} a\alpha + b\alpha_1 & a\beta + b\beta_1 \\ a_1\alpha + b_1\alpha_1 & a_1\beta + b_1\beta_1 \end{vmatrix} =$$

$$= \begin{vmatrix} a\alpha + a_1\beta & a\alpha_1 + a_1\beta_1 \\ b\alpha + b_1\beta & b\alpha_1 + b_1\beta_1 \end{vmatrix} = \begin{vmatrix} a\alpha + a_1\alpha_1 & a\beta + b_1\beta_1 \\ b\alpha + b_1\alpha_1 & a_1\beta + b_1\beta_1 \end{vmatrix}$$

$$\begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix}^2 = \begin{vmatrix} a^2 + b^2 & aa_1 + bb_1 \\ aa_1 + bb_1 & a_1^2 + b_1^2 \end{vmatrix} = \begin{vmatrix} a^2 + a_1 b & ab + bb_1 \\ aa_1 + a_1 b_1 & a_1 b + b_1^2 \end{vmatrix} =$$

$$= \begin{vmatrix} a^2 + a_1^2 & ab + a_1 b_1 \\ ab + a_1 b_1 & b^2 + b_1^2 \end{vmatrix}.$$

Друг пример:

$$\begin{vmatrix} a & b \\ -b' & a' \end{vmatrix} \begin{vmatrix} \alpha & \beta \\ -\beta' & \alpha' \end{vmatrix} = \begin{vmatrix} a\alpha + b\beta & -a\beta' + b\alpha' \\ a'\beta - b'\alpha & a'\alpha' + b'\beta' \end{vmatrix} =$$

$$= \begin{vmatrix} a\alpha - b\beta' & a\beta + b\alpha' \\ -a'\beta' - b'\alpha & a'\alpha' - b'\beta \end{vmatrix} = \begin{vmatrix} a\alpha - b'\beta & -a\beta' - b'\alpha' \\ a'\beta + b\alpha & a'\alpha' - b\beta' \end{vmatrix} =$$

$$= \begin{vmatrix} a\alpha + b'\beta' & a\beta - b'\alpha' \\ -a'\beta' + b\alpha & a'\alpha' + b\beta \end{vmatrix}.$$

Ако сега поставим тук

$$a=a_1+ia_2,\quad b=b_1+ib_2,\quad \alpha=\alpha_1+i\alpha_2,\quad \beta=\beta_1+i\beta_2,$$
$$a'=a_1-ia_2,\quad b'=b_1-ib_2,\quad \alpha'=\alpha_1-i\alpha_2,\quad \beta'=\beta_1-i\beta_2$$

и означим с $N(a)=|a|^2=aa'$ и т. н., то получаваме следните тъждества:

$$[N(a)+N(b)][N(\alpha)+N(\beta)]=N(a\alpha+b\beta)+N(a'\beta-b'\alpha)=$$
$$=N(a\alpha-b\beta')+N(a\beta+b\alpha')=N(a\alpha+b'\beta)+N(a'\beta+b\alpha)=$$
$$=N(a\alpha+b'\beta')+N(a'\beta'-b\alpha).$$

Но като вземем под внимание, че

$$N(a\alpha+b\beta)=(a\alpha+b\beta)(a'\alpha'+b'\beta')=$$
$$=(a_1\alpha_1-a_2\alpha_2+b_1\beta_1-b_2\beta_2)^2+(a_1\alpha_2+a_2\alpha_1+b_1\beta_2+b_2\beta_1)^2,$$
$$N(a'\beta-b'\alpha)=(a'\beta-b'\alpha)(a\beta'-b\alpha')=$$
$$=(b_1\alpha_1+b_2\alpha_2-a_1\beta_1-a_2\beta_2)^2+(b_2\alpha_1-b_1\alpha_2+a_1\beta_2-a_2\beta_1)^2,$$

получаваме

$$(a_1^2+a_2^2+b_1^2+b_2^2)(\alpha_1^2+\alpha_2^2+\beta_1^2+\beta_2^2)=$$
$$=(a_1\alpha_1-a_2\alpha_2+b_1\beta_1-b_2\beta_2)^2+(a_1\alpha_2+a_2\alpha_1+b_1\beta_2+b_2\beta_1)^2+$$
$$+(b_1\alpha_1+b_2\alpha_2-a_1\beta_1-a_2\beta_2)^2+(b_2\alpha_1-b_1\alpha_2+a_1\beta_2-a_2\beta_1)^2$$

и други три подобни тъждества. Имаме следователно, че произведението на две суми от четири квадрата може да се представи по четири начина като сума от четири квадрата — резултат, даден от Ойлер и Лагранж.

Можем по изложеното правило да умножаваме детерминанти от различен ред, стига детерминанта, която е от по-нисък ред, да я представим като детерминанта от по-висок ред с въвеждане на елементи, равни на нула. Така например имаме

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \cdot \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \cdot \begin{vmatrix} a & \beta & 0 & 0 \\ \gamma & \delta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} =$$

$$= \begin{vmatrix} a_1\alpha+b_1\beta & a_1\gamma+b_1\delta & c_1 & d_1 \\ a_2\alpha+b_2\beta & a_2\gamma+b_2\delta & c_2 & d_2 \\ a_3\alpha+b_3\beta & a_3\gamma+b_3\delta & c_3 & d_3 \\ a_4\alpha+b_4\beta & a_4\gamma+b_4\delta & c_4 & d_4 \end{vmatrix}.$$

5. Умножение на матрици. Ще се занимаем с обобщението на правилото за умножение. Нека са дадени две матрици:

$$\left\| \begin{matrix} a_{11}\, a_{12} \ldots a_{1p} \\ a_{21}\, a_{22} \ldots a_{2p} \\ \ldots\ldots\ldots \\ a_{n1}\, a_{n2} \ldots a_{np} \end{matrix} \right\|, \quad \left\| \begin{matrix} b_{11}\, b_{12} \ldots b_{1p} \\ b_{21}\, b_{22} \ldots b_{2p} \\ \ldots\ldots\ldots \\ b_{n1}\, b_{n2} \ldots b_{np} \end{matrix} \right\|.$$

Ако умножим хоризонталните редове с хоризонтални редове, както при умножението на детерминанти, то получаваме n^2 елементи c_{ik} на една детерминанта

$$C = \begin{vmatrix} c_{11}\, c_{12} \ldots c_{1n} \\ c_{21}\, c_{22} \ldots c_{2n} \\ \ldots\ldots\ldots \\ c_{n1}\, c_{n2} \ldots c_{nn} \end{vmatrix},$$

$$c_{ik} = a_{i1}\, b_{k1} + a_{i2}\, b_{k2} + \ldots + a_{ip} b_{kp} = \sum_{r=1}^{p} a_{ir} b_{kr}.$$

Първият член на тази детерминанта е

$$c_{11} c_{22} \ldots c_{nn} = \sum_{r} a_{1r}\, b_{1r} \sum_{s} a_{2s} b_{2s} \ldots \sum_{v} a_{nv} b_{nv} =$$

$$= \sum_{r,s,t,\ldots,v} a_{1r} a_{2s} a_{3t} \ldots a_{nv} b_{1r} b_{2s} b_{3t} \ldots b_{nv},$$

гдето сумата е разпростряна за всички стойности на $r, s, t, \ldots, v$ от 1 до p. От този член се получават всички членове на C с пермутиране на вторите индекси на c. При това се пермутират само първите индекси на b; всички други остават непроменени. Следователно

$$C = \sum \pm c_{11} c_{22} \ldots c_{nn} =$$

$$= \sum (a_{1r} a_{2s} a_{3t} \ldots a_{nv} \cdot \sum \pm b_{1r} b_{2s} b_{3t} \ldots b_{nv}) =$$

$$= \sum a_{1r} a_{2s} a_{3t} \ldots a_{nv} . B_{r,s,t,\ldots,v},$$

гдето $B_{r,s,t,\ldots,v}$ е детерминанта от n-ти ред от матрицата на b, която съдържа n вертикални реда с индекси $r, s, t, \ldots, v$. В сумата всеки член е нула, ако някои от числата $r, s, t, \ldots, v$ са равни помежду си. Следователно, ако $p < n$, то $C = 0$; понеже между индексите $r, s, t, \ldots, v$, които са взети от числата $1, 2, 3, \ldots, p$, трябва да има равни, т. е. $B_{r,s,t,\ldots,v} = 0$, като детерминанта с равни стълбове.

Ако $p=n$, то индексите могат само по един начин да се комбинират и да бъдат различни и сумата Σ е разпростряна върху всички пермутации на числата 1, 2, 3,..., n. Ако с B означим детерминантата

$$B=\begin{vmatrix} b_{11}\, b_{12} \ldots b_{1n} \\ b_{21}\, b_{22} \ldots b_{2n} \\ \ldots\ldots\ldots \\ b_{n1}\, b_{n2} \ldots b_{nn} \end{vmatrix},$$

то тогава, както получихме в началото,

$$B_{r,s,t,\ldots,v}=(-1)^{[r,s,t,\ldots,v]}B,$$

така че

$$C=B\sum(-1)^{[r,s,t,\ldots,v]}a_{1r}a_{2s}a_{3t},\ldots a_{nv}=AB,$$

ако A е съответно детерминанта от a. Така получихме едно ново доказателство на правилото за умножение на детерминанти.

Ако обаче $p>n$, то може отначало в израза за C сумирането да разпростремем само върху пермутациите на една комбинация на n-те индекса r, s, $t,\ldots,v$. Така получаваме, както по-горе, произведението

$$A_{r,s,t,\ldots,v}\,B_{r,s,t,\ldots,v},$$

гдето $A_{r,s,t,\ldots,v}$ е детерминанта на n-ти ред, образувана от матрицата на a, в която стълбовете имат индекси r, s, $t,\ldots,v$. Следователно имаме

$$C=\sum A_{r,s,t,\ldots,v}\,B_{r,s,t,\ldots,v},$$

гдето сумирането е разпростряно върху всички комбинации $r, s, t,\ldots,v$ $(r<s<t<\ldots<v)$ на числата 1, 2, 3,..., p, взети по n. Така получаваме следното правило на Бине — Коши. Произведението на две матрици от по n хоризонтални реда и p стълба е равно на нула, ако

$$p<n,$$

и равно на произведението на двете детерминанти от матриците, ако $p=n$. В случай, че $p>n$, то това произведение е равно на сумата от всички детерминанти от n-ти ред, които може да образуваме от едната матрица, умножени със съответните и детерминанти от другата матрица.

За пример да вземем произведение на матриците

$$\left\| \begin{matrix} a & b & c \\ a_1 & b_1 & c_1 \end{matrix} \right\|. \qquad \left\| \begin{matrix} \alpha & \beta & \gamma \\ \alpha_1 & \beta_1 & \gamma_1 \end{matrix} \right\|.$$

Така получаваме

$$\begin{vmatrix} a\alpha+b\beta+c\gamma & a\alpha_1+b\beta_1+c\gamma_1 \\ a_1\alpha+b_1\beta+c_1\gamma & a_1\alpha_1+b_1\beta_1+c_1\gamma_1 \end{vmatrix} = (ab_1-a_1b)(\alpha\beta_1-\alpha_1\beta)+$$
$$+(bc_1-cb_1)(\beta\gamma_1-\gamma\beta_1)+$$
$$+(ac_1-ca_1)(\alpha\gamma_1-\alpha_1\gamma).$$

6. Адюнгирана детерминанта. От адюнгираните количества A_{ik} в детерминантата

$$A=\begin{vmatrix} a_{11} & a_{12}\dots a_{1n} \\ a_{21} & a_{22}\dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2}\dots a_{nn} \end{vmatrix}$$

да образуваме детерминантата

4)
$$A'=\begin{vmatrix} A_{11} & A_{12}\dots A_{1n} \\ A_{21} & A_{22}\dots A_{2n} \\ \dots & \dots \\ A_{n1} & A_{n2}\dots A_{nn} \end{vmatrix},$$

която се нарича адюнгирана на първата. Ако умножим двете детерминанти A и A' по хоризонтални редове, то понеже

$$a_{i1}A_{k1}+a_{i2}A_{k2}+\dots+a_{in}A_{kn}=\begin{cases} 0, \text{ ако } i\neq k, \\ A \quad \text{„} \quad i=k, \end{cases}$$

получаваме

$$AA'=\begin{vmatrix} A & 0 & 0\dots 0 \\ 0 & A & 0\dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & 0\dots A \end{vmatrix}=A^n,$$

отгдето, ако $A\neq 0$,

(5)
$$A'=A^{n-1}.$$

Следователно адюнгираната детерминанта е $n-1$-ва степен на дадената.

Това изведохме при $A\neq 0$. Обаче лесно е да се види, че ако $A=0$, то и $A'=0$. Това е очевидно, ако рангът на A е по-малък от $n-1$, понеже тогава всички A_{ik} като детерминанти от $n-1$ ред от матрицата на A са нули, така че $A'=0$. Ако рангът е $n-1$ и например, ако не всички детерминанти, образувани от първите $n-1$ стълба на A, са равни на нула, то с малко подходящо изменение на елементи от последния стълб може да се получи детерминанта A, отлична от

нула. Но понеже за тази е валидна формулата (5), то с граничен преход се доказва горното. Впрочем в това се убеждаваме и директно, като вземем под внимание, че A е функция на n^2 независими променливи a_{ik}, нетъждествено равна на нула, така че от $AA' = A^n$ следва (5).

Аналогични резултати ще получим за миньорите на A'. Така нека на миньора от първите m хоризонтални реда и стълба дадем формата

$$\begin{vmatrix} A_{11} & \dots & A_{1m} & A_{1,m+1} & \dots & A_{1n} \\ A_{21} & \dots & A_{2m} & A_{2,m+1} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ A_{m1} & \dots & A_{mm} & A_{m,m+1} & \dots & A_{m,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{vmatrix}.$$

Като умножим тази детерминанта с A по редове, получаваме

$$\begin{vmatrix} A & 0 & 0 & \dots & 0 & a_{1,m+1} & \dots & a_{1n} \\ 0 & A & 0 & \dots & 0 & a_{2,m+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A & a_{m,m+1} & \dots & a_{mn} \\ 0 & 0 & 0 & \dots & 0 & a_{m+1,m+1} & \dots & a_{m+1,n} \\ 0 & 0 & 0 & \dots & 0 & a_{m+2,m+1} & \dots & a_{m+2,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & a_{n,m+1} & \dots & a_{n,n} \end{vmatrix} = A^m \begin{vmatrix} a_{m+1,m+1} & \dots & a_{m+1,n} \\ a_{m+2,m+1} & \dots & a_{m+2,n} \\ \dots & \dots & \dots \\ a_{n,m+1} & \dots & a_{n,n} \end{vmatrix},$$

отгдето

$$\begin{vmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{vmatrix} = A^{m-1} \begin{vmatrix} a_{m+1,m+1} & \dots & a_{m+1,n} \\ a_{m+2,m+1} & \dots & a_{m+2,n} \\ \dots & \dots & \dots \\ a_{n,m+1} & \dots & a_{n,n} \end{vmatrix}.$$

Този резултат веднага се обобщава за един произволен миньор. Ако $r_1, r_2, \dots, r_m$; $s_1, s_2, \dots, s_m$ са индексите на хоризонталните ре-

дове и стълбовете и $g, h, i, \ldots$; $u, v, w, \ldots$ съответните за индексите на адюнгирания миньор в A, то

$$\begin{vmatrix} A_{r_1 s_1} & \ldots & A_{r_1 s_m} \\ A_{r_2 s_1} & \ldots & A_{r_2 s_m} \\ \ldots & \ldots & \ldots \\ a_{r_m s_1} & \ldots & A_{r_m s_m} \end{vmatrix}_m = (-1)^{\mu} A^{m-1} \begin{vmatrix} a_{gu} & a_{gv} & a_{gw} \ldots \\ a_{hu} & a_{hv} & a_{hw} \ldots \\ a_{iu} & a_{iv} & a_{iw} \ldots \\ \ldots & \ldots & \ldots \end{vmatrix},$$

m

где $\mu = \sum_{k=1}^{m} r_k + \sum_{k=1}^{m} s_k$. Ако $A=0$, убеждаваме се както по-горе, че релацията остава в сила.

Ако $A=0$, то всички миньори от ред, по-голям от единица, на A' са равни на нула. Така

$$\begin{vmatrix} A_{r1} & A_{rk} \\ A_{s1} & A_{sk} \end{vmatrix} = 0,$$

откъдето $\frac{A_{r1}}{A_{s1}} = \frac{A_{rk}}{A_{sk}}$, т. е. $\frac{A_{r1}}{A_{s1}} = \frac{A_{r2}}{A_{s2}} = \ldots = \frac{A_{rn}}{A_{sn}}$; и тъй, ако една детерминанта е нула, то адюнгираните количества на два паралелни реда (хоризонтални редове или стълбове) са пропорционални помежду си.

7. Теорема на Силвестер. Нека е дадена детерминантата $A=|a_{ik}|$, $i, k=1,2,\ldots, n$, и нека $\nu<n$, $i>\nu$, $k>\nu$. От матрицата на главния миньор

$$M_\nu = \begin{vmatrix} a_{11} a_{12} \ldots a_{1\nu} \\ a_{21} a_{22} \ldots a_{2\nu} \\ \ldots \\ a_{\nu 1} a_{\nu 2} \ldots a_{\nu\nu} \end{vmatrix} \tag{6}$$

с прибавяне на i-ия ред и k-ия стълб ще получим детерминанта b_{ik} от ред $(\nu+1)$:

$$b_{ik} = \begin{vmatrix} a_{11} \ldots a_{1\nu} a_{1k} \\ \ldots \\ a_{\nu 1} \ldots a_{\nu\nu} a_{\nu k} \\ a_{i1} \ldots a_{i\nu} a_{ik} \end{vmatrix}. \tag{7}$$

Имаме следната теорема на Силвестер:

Детерминантата от $(n-\nu)$-ти ред

$$B = \begin{vmatrix} b_{\nu+1,\nu+1} \ldots b_{\nu+1,n} \\ \ldots \\ b_{n,\nu+1} \ldots b_{n,n} \end{vmatrix} \tag{8}$$

е равна на произведението от A и $(n-\nu-1)$-вата степен на M_ν, т. е.

$$B = AM_\nu^{n-\nu-1}.$$

Видяхме в предния параграф, че главният миньор на A'

(9)
$$M = \begin{vmatrix} A_{\nu+1,\nu+1} \dots A_{\nu+1,n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ A_{n,\nu+1} \dots A_{n,n} \end{vmatrix}$$

има стойност

(10)
$$M = A^{n-\nu-1} M_\nu.$$

Ако в (9) зачеркнем реда с индекс i и стълба с индекс k, то от останалите елементи се получава един миньор от A' от ред $n-\nu-1$, който има стойност (по § 6), равна на

$$(-1)^{i+k} A^{n-\nu-2} b_{ik},$$

гдето b_{ik} е дадено със (7). От друга страна, този миньор, умножен с $(-1)^{i+k-\nu-\nu} = (-1)^{i+k}$, дава адюнгираното количество на A_{ik} в детерминантата (9). Следователно това адюнгирано количество е равно на

$$A^{n-\nu-2} b_{ik},$$

така че адюнгираната детерминанта на (9) има стойност

(11)
$$A^{(n-\nu-2)(n-\nu)} B.$$

От друга страна, по § 6 тази адюнгирана детерминанта има стойност $M^{n-\nu-1}$ и е равна следователно според (10) на

(12)
$$A^{(n-\nu-1)^2} M_\nu^{n-\nu-1},$$

така че, като приравним (11) и (12)

$$A^{(n-\nu-2)(n-\nu)} B = A^{(n-\nu-2)(n-\nu)} AM_\nu^{n-\nu-1},$$

получаваме търсената релация, като съкратим на общата степен на A, която като функцията на n независими променливи a_k не е идентично равна на нула.

Глава IV

Специални детерминанти

1. Детерминанта на Вандермианд. Така се нарича детерминантата

$$\nu_n = \begin{vmatrix} 1 & 1 & \dots 1 \\ x_1 & x_2 & \dots x_n \\ x_1^2 & x_2^2 & \dots x_n^2 \\ \cdot \quad \cdot & \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ x_1^{n-1} & x_2^{n-1} & \dots x_n^{n-1} \end{vmatrix}.$$

Да пресметнем ν_4:

$$\nu_4 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{vmatrix}.$$

Да извадим от четвъртия хоризонтален ред третия, умножен с x_1, от третия—втория, умножен с x_1, и от втория — първия, умножен пак с x_1; така получаваме

$$\nu_4 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & x_2-x_1 & x_3-x_1 & x_4-x_1 \\ 0 & x_2^2-x_1x_2 & x_3^2-x_1x_3 & x_4^2-x_1x_4 \\ 0 & x_2^3-x_1x_2^2 & x_3^3-x_1x_3^2 & x_4^3-x_1x_4^2 \end{vmatrix} =$$

$$= (x_2-x_1)\ (x_3-x_1)\ (x_4-x_1) \begin{vmatrix} 1 & 1 & 1 \\ x_2 & x_3 & x_4 \\ x_2^2 & x_3^2 & x_4^2 \end{vmatrix}.$$

По същия начин

$$\nu_3 = \begin{vmatrix} 1 & 1 & 1 \\ x_2 & x_3 & x_4 \\ x_2^2 & x_3^2 & x_4^2 \end{vmatrix} = (x_3-x_2)\ (x_4-x_2) \begin{vmatrix} 1 & 1 \\ x_3 & x_4 \end{vmatrix} =$$

$$= (x_3-x_2)(x_4-x_2)(x_4-x_3).$$

Както се вижда, този път е същият за ν_n и дава

$$\nu_n = (x_2-x_1)(x_3-x_1)\ldots(x_n-x_1)(x_3-x_2)\ldots(x_n-x_2)\ldots$$

$$\ldots(x_n-x_{n-1}) = \prod_{\alpha>\beta} (x_\alpha-x_\beta), \text{ където } \alpha, \beta = 1,2,\ldots, n.$$

Този резултат можем да получим направо. Детерминантата ν_n е цяла рационална функция на $x_1, x_2, \ldots, x_n$ от степен

$$0+1+2+\ldots+(n-1) = \frac{n(n-1)}{2}.$$

Ако извадим от α-ия стълб β-ия, то елементите на α-ия стълб ще станат

$$0,\ x_\alpha-x_\beta,\ x_\alpha^2-x_\beta^2, \ldots,\ x_\alpha^{n-1}-x_\beta^{n-1},$$

които се делят на $x_\alpha - x_\beta$. Следователно ν_n се дели на всички тези разлики, т. е. ще имаме

$$\nu_n = \theta \prod_{\alpha>\beta} (x_\alpha - x_\beta) \quad \alpha, \beta = 1, 2, \ldots, n.$$

Понеже степента на произведението $\prod$ е $\frac{n(n-1)}{2}$, то θ е константа. За да намерим тази константа, нека видим кои са коефициентите на $x_2 x_3^2 \ldots x_n^{n-1}$. В ν_n очевидно коефициентът е 1 и вдясно θ, отгдето следва $\theta = 1$, т. е. за ν_n имаме формулата (1). От тази формула става ясно, че само тогава $\nu_n = 0$, когато поне две от числата $x_1, x_2, \ldots, x_n$ са равни помежду си.

2. Циркуланти. Детерминантите, в които редовете се получават чрез циклично пермутиране на първия ред, се наричат циркуланти или кръгови детерминанти. Нека $\alpha_1, \alpha_2, \ldots, \alpha_n$ са корените на $x^n = 1$. За да получим стойността на циркулантата

$$D = \begin{vmatrix} a_1 & a_2 & a_3 \ldots a_n \\ a_n & a_1 & a_2 \ldots a_{n-1} \\ a_{n-1} & a_n & a_1 \ldots a_{n-2} \\ \cdot & \cdot & \cdot \ldots \cdot \\ a_2 & a_3 & a_4 \ldots a_1 \end{vmatrix},$$

умножаваме я с

$$\Delta = \begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 \ldots \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 \ldots \alpha_2^{n-1} \\ \cdot & \cdot & \cdot \ldots \cdot \\ 1 & \alpha_n & \alpha_n^2 \ldots \alpha_n^{n-1} \end{vmatrix}.$$

Ако означим с

$$f(x) = a_1 + a_2 x + a_3 x^2 + \ldots + a_n x^{n-1},$$

то произведението на всеки хоризонтален ред $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ от Δ с първия хоризонтален ред от D е равно на $f(\alpha)$ и произведението на същия хоризонтален ред от Δ с i-ия в D е

$$a_{n-i+2} + a_{n-i+3}\alpha + \ldots + a_n \alpha^{-2} + a_1 \alpha^{i-1} + a_2 \alpha^i + \ldots + a_{n-i+1} \alpha^{n-1}$$

или като вземем под внимание, че $\alpha^n = 1$,

$$a_1 \alpha^{i-1} + a_2 \alpha^i + \ldots + a_{n-i+2}\, \alpha^n + a_{n-i+3}\, \alpha^{n+1} + \ldots + a_n \alpha^{n+i-2} = \alpha^{i-1} f(\alpha).$$

Следователно

$$D\Delta = \begin{vmatrix} f(\alpha_1) & \alpha_1 f(\alpha_1) & \alpha_1^2 f(\alpha_1) \ldots \alpha_1^{n-1} f(\alpha_1) \\ f(\alpha_2) & \alpha_2 f(\alpha_2) & \alpha_2^2 f(\alpha_2) \ldots \alpha_2^{n-1} f(\alpha_2) \\ \cdot & \cdot & \cdot \ldots \cdot \\ f(\alpha_n) & \alpha_n f(\alpha_n) & \alpha_n^2 f(\alpha_n) \ldots \alpha_n^{n-1} f(\alpha_n) \end{vmatrix}.$$

Като извадим $f(\alpha_1)\ f(\alpha_2)\ldots f(\alpha_n)$ пред детерминантата, остава Δ, така че

$$D=f(\alpha_1),f(\alpha_2)\ldots f(\alpha_n).$$

Така например при $n=4$, $x^4=1$ има корени $1,-1,\ i,-i$, отгдето имаме, като означим с $a_1=a,\ a_2=b,\ a_3=c,\ a_4=d$:

$$f(1)=a+b+c+d,\ f(i)=a-c+(b-d)\ i,$$

$$f(-1)=a-b+c-d,\ f(-i)=a-c-(b-d)\,i,$$

$$\begin{vmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{vmatrix} =(a+b+c+d)\ (a-b+c-d)\ [(a-c)^2+(b-d)^2].$$

3. Континюанти. Континюанти се наричат такива детерминанти, на които елементите са нули с изключение на тези по главния диагонал, които са произволни, и тези, които са по съседните две линии, успоредни на главния диагонал, и имат стойност $+1$ и -1. Означаваме континюантите така:

$$(a_1\,a_2\ldots a_n)=\begin{vmatrix} a_1 & 1 & 0\ldots 0 & 0 \\ -1 & a_2 & 1\ldots 0 & 0 \\ 0 & -1 & a_3\ldots 0 & 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 0 & 0 & 0\ldots a_{n-1} & 1 \\ 0 & 0 & 0\ldots -1 & a_n \end{vmatrix}$$

Ако развием тази детерминанта по първите k стълба, получаваме

$$(6)\qquad (a_1a_2\ldots a_n)=(a_1a_2\ldots a_k)\ (a_{k+1}\ldots a_n)+(a_1a_2\ldots a_{k-1})\ (a_{k+2}\ldots a_n).$$

В частност имаме

$$(7)\qquad (a_1a_2\ldots a_n)=a_1\ (a_2\ldots a_n)+(a_3a_4\ldots a_n),$$

$$(8)\qquad (a_1a_2\ldots a_n)=a_n(a_1a_2\ldots a_{n-1})+(a_1a_2\ldots a_{n-2}).$$

От (7) получаваме

$$a_1=(a_1),\ a_1+\frac{1}{a_2}=\frac{(a_1a_2)}{(a_1)},\ a_1+\cfrac{1}{a_2+\cfrac{1}{a_3}}=\frac{(a_1a_2a_3)}{(a_2a_3)}.$$

По-общо

$$\frac{(a_1a_2\ldots a_n)}{(a_2a_3\ldots a_n)} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdot\ \cdot\ \cdot + \cfrac{1}{a_n}}}.$$

Да разгледаме сега континюантата

$$u_n = \begin{vmatrix} 1 & 1 & 0 \ldots 0 \\ -1 & 1 & 1 \ldots 0 \\ 0 & -1 & 1 \ldots 0 \\ \cdot & \cdot & \cdot \ldots \cdot \\ 0 & 0 & 0 \ldots 1 \end{vmatrix}.$$

Релацията (8) дава

$$u_n = u_{n-1} + u_{n-2},\ u_0 = 1,\ u_{-1} = 0, \tag{9}$$

което показва, че в редицата $u_1, u_2, u_3, \ldots$ всеки член е сума от два предшествуващи го члена. Стойностите им са

$$1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots$$

Този ред се нарича ред на Фибоначи. Числата от реда притежават интересни свойства. Така от (6) имаме

$$u_n = u_k u_{n-k} + u_{k-1} u_{n-k-1}$$

и в частност $u_{2k} = u_k^2 + u_{k-1}^2$. За да пресметнем u_n, нека α и β са корените на $x^2 - x - 1 = 0$:

$$\alpha = \frac{1}{2}\ (1 + \sqrt{5}),\ \beta = \frac{1}{2}\ (1 - \sqrt{5}).$$

Лесно се вижда, че (9) се удовлетворява, ако поставим

$$u_n = a\alpha^n + b\beta^n,$$

гдето a и b, за да се изпълняват условията $u_0 = 1$, $u_{-1} = 0$, трябва да удовлетворяват

$$\frac{a}{\alpha} + \frac{b}{\beta} = 0,\ a + b = 1,$$

отгдето получаваме

$$a = \frac{\alpha}{\alpha - \beta},\ b = -\frac{\beta}{\alpha - \beta},\ \text{т. е. } u_n = \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}.$$

Числото u_n ни дава броя на членовете в развитието на континюантата.

4. Детерминанта на Смит. Една друга детерминанта, която е във връзка с теорията на числата, е тази на Смит:

$$D=\begin{vmatrix}(1,\ 1)\ (1,\ 2)\ (1,\ 3)\ldots(1,\ n)\\(2,\ 1)\ (2,\ 2)\ (2,\ 3)\ldots(2,\ n)\\(3,\ 1)\ (3,\ 2)\ (3,\ 3)\ldots(3,\ n)\\\ldots\ldots\ldots\ldots\ldots\\(n,\ 1)\ (n,\ 2)\ (n,\ 3)\ldots(n,\ n)\end{vmatrix},$$

гдето $(i,\ j)$ означава най-големия общ делител на i и j. Нека с $\varphi\ (n)$ да означим броя на числата, по-малки от n и взаимно прости с него. Известно е, че сумата от стойностите на $\varphi\ (d)$, когато d взема стойности, равни на всички делители на n, е равна на n. От това следва, че $(i,\ j)$ е сума от всички стойности на $\varphi(n)$, които отговарят на общите делители на i и j. Ако a_{ij} означава 1, ако i се дели на j или нула, ако i не се дели на j така, че винаги $a_{ij}=0$ при $i<j$, то ще имаме

$$(10)\qquad (i,j)=a_{i1}\,a_{j1}\,\varphi(1)+a_{i2}a_{j2}\varphi(2)+\ldots+a_{in}a_{jn}\varphi(n).$$

Действително $\varphi(\nu)$ само тогава фигурира в тази сума, когато i и j се делят на ν.

Обаче релацията (10) показва, че D е произведение на детерминантата

$$A=\begin{vmatrix}a_{11} & a_{12}\ldots a_{1n}\\a_{21} & a_{22}\ldots a_{2n}\\\ldots & \ldots\\a_{n1} & a_{n2}\ldots a_{nn}\end{vmatrix}$$

с детерминантата

$$B=\begin{vmatrix}a_{11}\ \varphi(1)\ a_{12}\ \varphi(2)\ldots a_{1n}\ \varphi(n)\\a_{21}\ \varphi(1)\ a_{22}\ \varphi(2)\ldots a_{2n}\ \varphi(n)\\\ldots\ldots\ldots\ldots\ldots\\a_{n1}\ \varphi(1)\ a_{n2}\ \varphi\ (2)\ldots\ a_{nn}\ \varphi(n)\end{vmatrix}=A\,\varphi(1)\varphi(2)\ldots\varphi(n).$$

Понеже $a_{ik}=0$ при $k>i$ и $a_{ii}=1$, то в A от едната страна на главния диагонал елементите са равни на нула, отгдето следва, че $A=$ $=a_{11}\,a_{22}\ldots a_{nn}=1$, т. е.

$$D=\varphi(1)\,\varphi(2)\,\varphi(3)\ldots\varphi(n).$$

5. Симетрични детерминанти. Една детерминанта

$$A=|\,a_{ik}\,|,\ i,\ k=1,2,\ldots,\ n$$

се нарича симетрична, ако $a_{ik}=a_{ki}$. Така

$$\begin{vmatrix} a & b & d \\ b & c & f \\ d & f & e \end{vmatrix}$$

е симетрична. Лесно се вижда, че квадратът на всяка детерминанта може да се представи като симетрична детерминанта. Действително, ако $A=|a_{is}|$, $i, s=1, 2, \ldots, n$, то

$$A^2=|c_{kh}|, \text{ гдето } c_{kh}=\sum_{p=1}^{n} a_{kp}a_{hp} \text{ и } c_{hk}=\sum_{p=1}^{n} a_{hp}a_{kp}=c_{kh}.$$

Ако една детерминанта A е симетрична, то детерминантата от адюнгираните количества е пак симетрична. Ако означим с A_{ij} адюнгираното количество на a_{ij}, ще докажем, че $A_{ij}=A_{ji}$. Действително, ако представим A_{ij} като детерминанта от n-ти ред с елементи нули на i-ия хоризонтален ред и j-ия стълб с изключение на елемента от i-ия хоризонтален ред и k-ия стълб, който вземаме за 1, то A_{ij} не се изменя при променяне на стълбове с редове. Така например нека $n=5$, $i=2$, $j=3$, то

$$A_{23}=\begin{vmatrix} a_{11} & a_{12} & 0 & a_{14} & a_{15} \\ 0 & 0 & 1 & 0 & 0 \\ a_{31} & a_{32} & 0 & a_{34} & a_{35} \\ a_{41} & a_{42} & 0 & a_{44} & a_{45} \\ a_{51} & a_{52} & 0 & a_{54} & a_{55} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & a_{31} & a_{41} & a_{51} \\ a_{12} & 0 & a_{32} & a_{42} & a_{52} \\ 0 & 1 & 0 & 0 & 0 \\ a_{14} & 0 & a_{34} & a_{44} & a_{54} \\ a_{15} & 0 & a_{35} & a_{45} & a_{55} \end{vmatrix}$$

и понеже $a_{ik}=a_{ki}$, то тази детерминанта е равна на

$$\begin{vmatrix} a_{11} & 0 & a_{13} & a_{14} & a_{15} \\ a_{21} & 0 & a_{23} & a_{24} & a_{25} \\ 0 & 1 & 0 & 0 & 0 \\ a_{41} & 0 & a_{43} & a_{44} & a_{45} \\ a_{51} & 0 & a_{53} & a_{54} & a_{55} \end{vmatrix}=A_{32}.$$

Да разгледаме една друга детерминанта, която се среща в небесната механика, именно при изчисление на вековите изменения на планетните орбити. Именно дадено е уравнението

$$\begin{vmatrix} a_{11}-x & a_{12} & a_{13} & \ldots a_{1n} \\ a_{21} & a_{22}-x & a_{23} & \ldots a_{2n} \\ a_{31} & a_{32} & a_{33}-x & \ldots a_{3n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots a_{nn}-x \end{vmatrix}=0. \tag{11}$$

Ще докажем, че ако $a_{ik}=a_{ki}$, $i, k=1, 2,\ldots, n$, то всички корени на това уравнение са реални.

Първо ще изложим доказателството, дадено от Силвестер. Ако умножим $D_n(\ )$ с $D_n(-x)$, то ще получим една детерминанта с елементи c'_{ik}, дадени с

$$c'_{ik}=a_{i1}a_{k1}+\ldots+(a_{ii}-x)\,a_{ki}+\ldots+a_{ik}(a_{kk}+x)+$$

$$+\ldots+a_{in}a_{kn}=c_{ik},$$

ако $i\neq k$, гдето $c_{ik}=\sum_{s=1}^{n} a_{is}a_{ks}$, и

$$c'_{ii}=a_{i1}^2+a_{i2}^2+\ldots+(a_{ii}-x)\ (a_{ii}+x)+\ldots+a_{in}^2=c_{ii}-x^2.$$

Следователно уравнението $D_n(x)D_n(-x)=0$ може да се пише

$$\begin{vmatrix} c_{11}-x^2 & c_{12} & \ldots c_{1n} \\ c_{21} & c_{22}-x^2 & \ldots c_{2n} \\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \\ c_{n1} & c_{n2} & \ldots c_{nn}-x^2 \end{vmatrix}=0.$$

Ако сега представим всеки елемент $c_{ik}\,(i\neq k)$ така: $c_{ik}=c_{ik}+0$, то всички елементи на тази детерминанта са биноми и можем да я разложим на други детерминанти, в които ще фигурира само по едно събираемо. Ако означим тогава с $C=|c_{ik}|$, $i, k=1, 2,\ldots, n$ лесно се вижда, че горното уравнение може да се пише така:

$$(12)\qquad C+\delta_{n-1}(-x^2)+\delta_{n-2}(-x^2)^2+\ldots+\delta_1(-x^2)^{n-1}+(-x^2)^n=0,$$

гдето δ_ν е сумата на всички главни миньори от ред ν на детерминантата C. Но всеки главен миньор в C е квадратът на една матрица в $A=|a_{in}|$, $i, k=1, 2,\ldots, n$ и по теоремата на Бине – Коши той е сума от квадрати на детерминанти, следователно ще има положителна стойност. Така например имаме

$$\begin{vmatrix} c_{11} & c_{12}\ldots c_{1p} \\ c_{21} & c_{22}\ldots c_{2p} \\ & \cdot\ \cdot\ \cdot \\ c_{p1} & c_{p2}\ldots c_{pp} \end{vmatrix}=\left\| \begin{matrix} a_{11} & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}\ldots a_{2n} \\ & \cdot\ \cdot\ \cdot\ \cdot \\ a_{p1} & a_{p2}\ldots a_{pn} \end{matrix} \right\|^2.$$

По горното се убеждаваме, че числата $\delta_1, \delta_2,\ldots,\delta_{n-1}, C$ са положителни. Но не е трудно да се види, че уравнението (12) не може да има чисто имагинерен корен $x=qi$. Действително тогава $-x^2=q^2$, а лявата част на (12) е положителна. Да допуснем, че уравнението (11)

има комплексен корен $x=p+iq$. Тогава, ако поставим $a'_{ii}=a_{ii}-p$, $i=1, 2, \dots, n$, би следвало, че уравнението

$$\begin{vmatrix} a'_{11}-x & a_{12} & \dots a_{1n} \\ a_{21} & a'_{22}-x & \dots a_{2n} \\ \cdot & \cdot \; \cdot \; \cdot & \cdot \\ a_{n1} & a_{n2} & \dots a'_{nn}-x \end{vmatrix} = 0$$

има корен qi, което видяхме е невъзможно.

Ще дадем друго доказателство на същата теорема, основано на по-прости свойства на детерминантите. Нека $x=p+iq$ е един корен на (11). Тогава от условието $D_n(p+iq)=0$ следва, че има числа $x_1, x_2, \dots, x_n$ не всички равни на нула, които удовлетворяват системата

$$\begin{aligned} &(a_{11}-p-iq)\; x_1+a_{12}x_2+\dots+a_{1n}x_n=0, \\ &a_{21}\, x_1+(a_{22}-p-iq)\, x_2+\dots+a_{2n}x_n=0, \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ &a_{n1}\, x_1+a_{n2}\, x_2+\dots+(a_{nn}-p-iq)\, x_n=0. \end{aligned}$$

Нека $x_k=y_k+iz_k$, то получаваме

$$\sum_{k=1}^{n} a_{ik}\, x_k=(p+iq)x_i=(p+iq)\;(y_i+iz_i)=py_i+qz_i+i\;(pz_i+qy_i),$$

или

$$\sum_{k=1}^{n} a_{gk}\, y_k=py_g-qz_g, \quad \sum_{k=1}^{n} a_{gk}\, z_k=qy_g+pz_g.$$

Като умножим първото с $-z_g$, второто с y_g и съберем,

$$\sum_{k=1}^{n} a_{gk}\,(y_g z_k-y_k z_g)=q\,(y_g^2+z_g^2).$$

Сумираме за $g=1, 2, \dots, n$ и като вземем под внимание, че $a_{ik}=a_{ki}$, получаваме

$$\sum_{g=1}^{n}\sum_{k=1}^{n} a_{gk}(y_g z_k-y_k z_g)=q\sum_{g=1}^{n}(y_g^2+z_g^2),$$

т. е.

$$q\sum_{g=1}^{n}(y_g^2+z_g^2)=0,$$

което показва, че $q=0$, т. е. всички корени са реални.

Читателят лесно ще се убеди, че ако допуснем, че a_{ik} са имагинерни, но $a_{ik}=\bar{a}_{ki}$, $i,\ k=1, 2,\ldots, n$, то $D_n(x)=0$ има пак само реални корени.

6. Полусиметрични детерминанти. Една детерминанта $A=|a_{ik}|$, $i, k=1, 2,\ldots, n$ се нарича така, ако $a_{ik}=-a_{ki}$, следователно $a_{ii}=0$.

Полусиметричната детерминанта от нечетен ред е равна на нула. Действително да умножим хоризонталните редове с -1, при което детерминантата се умножава с $(-1)^n$, и в получената детерминанта да сменим хоризонталните редове със стълбовете и обратно. Като вземем под внимание, че $a_{ik}+a_{kl}=0$, вижда се, че така получаваме пак A, т. е.

$$(-1)^n A=A,$$

т. е. понеже n е нечетно, $A=-A$, $A=0$. Така например имаме

$$A=\begin{vmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{vmatrix},\quad -A=\begin{vmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{vmatrix}=\begin{vmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{vmatrix}=A.$$

Нека A_{ik} е адюнгираното количество на a_{ik} в A. Ако представим A_{ik} като детерминанта от n-ти ред, както това направихме при симетричните детерминанти, умножим с -1 всички хоризонтални редове с изключение на i-ия и променим тези редове със стълбовете, то лесно се вижда, че ще получим A_{ki}, т. е. $(-1)^{n-1}A_{ik}=A_{ki}$. Значи адюнгираната детерминанта на A е симетрична, ако n е нечетно, и полусиметрична, ако n е четно.

Ако n е нечетно, то понеже $A=0$, всеки миньор от ред, по-голям от единица, в адюнгираната детерминанта е равен на нула и ще имаме

$$A_{ii}A_{kk}-A_{ik}A_{ki}=0 \quad \text{или} \quad A_{ii}A_{kk}=A_{ik}^2.$$

Всяка полусиметрична детерминанта от четен ред е точен квадрат от една цяла рационална функция на нейните елементи.

Нека $B=|a_{ik}|$, $i, k=1, 2,\ldots, n$; $n=2m$, е дадената полусиметрична детерминанта. Нека $A=|a_{ik}|$, $i, k=1, 2,\ldots, n-1$ и A_{ik} да е адюнгираното количество на a_{ik} в A. Тогава, както видяхме (§ 2, глава III), B може да се представи така:

$$B=a_{nn}A-\sum_{\alpha,\beta} a_{\alpha n}a_{n\beta}A_{\alpha\beta}=-\sum a_{\alpha n}a_{n\beta}A_{\alpha\beta},$$

понеже $A=0$ като полусиметрична от нечетен ред. Ще преработим израза B. Ако $\alpha=\beta$, понеже $a_{n\beta}=-a_{\beta n}$, имаме $-a_{\alpha n}a_{n\alpha}A_{\alpha\alpha}=a^2_{\alpha n}A_{\alpha\alpha}$. Ако $\alpha\neq\beta$, групираме двата члена

$$-a_{\alpha n}a_{n\beta}A_{\alpha\beta}-a_{\beta\alpha}a_{n\alpha}A_{\beta\alpha}=a_{\alpha n}a_{\beta n}(A_{\alpha\beta}+A_{\beta\alpha})=2a_{\alpha n}a_{\beta n}A_{\alpha\beta}.$$

Следователно имаме

$$B=\sum(a^2{}_{\alpha\pi}A_{\alpha\alpha}+2a_{\alpha\pi}a_{\beta\pi}A_{\alpha\beta}).$$

Но видяхме, че $A_{\alpha\beta}=\sqrt{A_{\alpha\alpha}A_{\beta\beta}}$, понеже A е полусиметрична детерминанта от нечетен ред. Оттук

$$(13)\qquad B=\sum(a^2{}_{\alpha\pi}A_{\alpha\alpha}+2a_{\alpha\pi}a_{\beta\pi}\sqrt{A_{\alpha\alpha}A_{\beta\beta}})=(\sum a_{\alpha\pi}\sqrt{A_{\alpha\alpha}})^2,$$

гдето знака на единия радикал можем да вземем произволен, например знака на $\sqrt{A_1}$. Но тогава знакът на останалите радикали е напълно определен, понеже

$$\sqrt{A_{\alpha\alpha}}=\frac{A_{\alpha 1}}{\sqrt{A_{11}}}.$$

Да допуснем, че сме доказали теоремата за детерминанта от $(n-2)$-ти ред. Тогава $A_{\alpha\alpha}$ като полусиметрични детерминанти от $(n-2)$-ри ред ще са точни квадрати и тогава следва от (13), че и B ще е точен квадрат. Понеже при $n=2$ теоремата е валидна, тъй като

$$\begin{vmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{vmatrix}=a_{12}{}^2,$$

то следва валидността ѝ при всяко n. При $n=4$ имаме

$$\begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{vmatrix}=(a_{12}a_{34}-a_{13}a_{24}+a_{14}a_{23})^2.$$

7. **Пфафиан.** Изразът

$$(14)\qquad \sqrt{B}=\sum a_{\alpha\pi}\sqrt{A_{\alpha\alpha}}$$

се нарича Пфафиан от n-ти ред или полудетерминанта и се бележи с

$$P_n=(1,\ 2,\ 3,\ldots,\ n).$$

В дясната част на (14), понеже $a_{\pi\pi}=0$, имаме $n-1$ събираеми. Също всяко $\sqrt{A_{\alpha\alpha}}$ съдържа $n-3$ събираеми и т. н. до $n=2$, която има само един член. Следователно P_n се състои от $1.3.5\ldots(n-1)$ члена. $A_{\alpha\alpha}$ се получава от B, като зачеркнем α-ия и n-ия хоризонтален ред и стълбовете със същите индекси. Следователно елементите от $A_{\alpha\alpha}$ съдържат само индекси от реда

$$1,\ 2,\ 3,\ldots,\ \alpha-1,\ \alpha+1,\ldots,\ n-1.$$

Следвайки така, получаваме: ако

$$(15)\qquad a_{\alpha\beta}a_{\gamma\delta}\ldots a_{\sigma\pi}$$

е един член на P_n, то между индексите α, β, γ, δ, ..., σ, τ всяка стойност от 1, 2, 3, ..., $(n-1)$, n влиза само един път. P_n съдържа само такива членове. Действително един произволен индекс α може да се комбинира с един от $(n-1)$-те индекси β: след това един от останалите индекси γ да се комбинира с един от другите $(n-3)$ и т. н. Следователно в (15) има $(n-1)(n-3)\ldots 3.1$ такива членове.

8. Ортогонални детерминанти. Една детерминанта $A=|a_{ik}|$, $i, k=1, 2, \ldots, n$ се нарича ортогонална, ако между елементите ѝ има следните връзки:

$$(19)\qquad a_{i1}a_{k1}+a_{i2}a_{k2}+\ldots+a_{in}a_{kn}=\begin{cases}0, \text{ ако } i\neq k,\\ 1 \text{ ако } i=k.\end{cases}$$

Оттук се вижда, че между елементите ѝ има на брой

$$n+\frac{n(n-1)}{2}=\frac{n(n+1)}{2}$$

връзки. Ако умножаваме по редове въз основа на (19), виждаме, че A^2 представя детерминанта с елементи от главния диагонал, равни на 1, а всички други са равни на нула, т. е. $A^2=1$, $A=\pm 1$.

Друго важно свойство е следното: адюнгираното количество на един елемент е равно на произведението на A с този елемент. Действително от уравненията

$$\begin{aligned}
&a_{11}a_{i1}+a_{12}a_{i2}+\ldots+a_{1n}a_{in}=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{i1}a_{i1}+a_{i2}a_{i2}+\ldots+a_{in}a_{in}=1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{n1}a_{i1}+a_{n2}a_{i2}+\ldots+a_{nn}a_{in}=0,
\end{aligned}$$

ако умножим първото с A_{1j}, второто с A_{2j} и т. н., последното с A_{nj} и съберем, ще получим

$$(20)\qquad Aa_{ij}=A_{ij}.$$

Ако това равенство умножим с a_{is} и съберем за $i=1, 2, \ldots, n$, получаваме

$$A(a_{1j}a_{1s}+a_{2j}a_{2s}+\ldots+a_{nj}a_{ns})=A_{1j}a_{1s}+A_{2j}a_{2s}+\ldots+A_{nj}a_{ns},$$

отгдето веднага следва

$$(21)\qquad a_{1j}a_{1s}+a_{2j}a_{2s}+\ldots+a_{nj}a_{ns}=\begin{cases}0, \text{ ако } j\neq s,\\ 1, \text{ ако } j=s.\end{cases}$$

Обратно, по същия начин от (21) следва (19), така че тези уравнения са напълно еквивалентни.

Понеже между n^2 елемента на една ортогонална детерминанта съществуват $\frac{n(n+1)}{2}$ зависимости, то от тях само

$$n^2-\frac{n(n+1)}{2}=\frac{n(n-1)}{2}$$

са независими. Интересно е да се намерят прости изрази за всички елементи посредством $\frac{n(n-1)}{2}$ независими параметри. Такива ни дават формулите на Кейли. Да вземем една детерминанта, която има форма на полусиметрична, само че елементите, които са по главния диагонал, не са нули:

$$B=\begin{vmatrix} b & b_{12} & b_{13} \dots & b_{1n} \\ b_{21} & b & b_{23} \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & b_{n3} \dots & b \end{vmatrix},$$

гдето $b_{ik}=-b_{ki}$ при $i\neq k$. Нека B_{ik} да е адюнгираното количество на b_{ik} в B.

Тогава, ако поставим

$$a_{ik}=\frac{2b\,B_{ik}}{B}, \quad a_{ii}=\frac{2b\,B_{ii}}{B}-1, \tag{22}$$

ще докажем, че детерминантата $A=|a_{ik}|$ е ортогонална със стойност 1.

За тази цел нека разгледаме следните $3n$ величини, дадени с уравненията:

$$x_i=b_{1i}z_1+b_{2i}z_2+\dots+b_{ni}z_n,$$

$$y_i=b_{i1}z_1+b_{i2}z_2+\dots+b_{in}z_n.$$

Ако ги съберем, понеже $b_{ik}=-b_{ki}$, $b\neq i$, то получаваме

$$x_i+y_i=2z_i\,b.$$

Ако сега умножим $x_1, x_2, \dots, x_n$ съответно с $B_{k1}, B_{k2}, \dots, B_{kn}$ и съберем, ще получим

$$Bz_k=B_{k1}x_1+B_{k2}x_2+\dots+B_{kn}x_n,$$

аналогично

$$Bz_k=B_{1k}y_1+B_{2k}y_2+\dots+B_{nk}y_n.$$

Ако в тези уравнения вместо z_k поставим $\frac{x_k+y_k}{2b}$, имаме

$$By_k=2b\,B_{k1}x_1+\dots+(2b\,B_{kk}-B)x_k+\dots+2b\,B_{kn}x_n,$$
$$Bx_k=2b\,B_{1k}y_1+\dots+(2b\,B_{kk}-B)y_k+\dots+2b\,B_{nk}y_n$$

или като въведем числата a,

$$y_k = a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n,$$
$$x_k = a_{1k}y_1 + a_{2k}y_2 + \ldots + a_{nk}y_n.$$

Ако y, определени от първото, заместим във второто с приравнение коефициентите пред x, получаваме

$$a_{1i}a_{1k} + a_{2i}a_{2k} + \ldots + a_{ni}a_{nk} = \begin{cases} 0, & i \neq k, \\ 1, & i = k, \end{cases}$$

което показва, че A е ортогонална детерминанта.

Детерминантата A е равна на

$$A = \frac{1}{B^n}\begin{vmatrix} 2bB_{11}-B & 2bB_{12} & \ldots & 2bB_{1n} \\ 2bB_{21} & 2bB_{22}-B & \ldots & 2bB_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 2bB_{n1} & 2bB_{n2} & \ldots & 2bB_{nn}-B \end{vmatrix}.$$

Умножаваме детерминантата вдясно с детерминанта B по редове. Произведението на елементите от първата от i-ия ред с втората от k-ия ред дава

$$2bB_{i1}b_{k1} + \ldots + (2bB_{ii}-B)b_{ki} + \ldots + 2bB_{in}B_{kn} = \begin{cases} Bb_{ik}, & i \neq k, \\ Bb, & i = k. \end{cases}$$

Оттук следва, че това произведение е равно на

$$\begin{vmatrix} Bb & Bb_{12} & \ldots & Bb_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ Bb_{n1} & Bb_{n2} & \ldots & Bb \end{vmatrix} = B^{n+1},$$

отгдето следва, че $A = 1$.

Примери. При $n = 2$ нека изберем

$$b = 1, \quad b_{12} = -b_{21} = p,$$

то ортогоналната детерминанта по формулите на Кейли ще бъде

$$\begin{vmatrix} \frac{1-p^2}{1+p^2} & \frac{2p}{1+p^2} \\ -\frac{2p}{1+p^2} & \frac{1-p^2}{1+p^2} \end{vmatrix}.$$

При $n = 3$ детерминантата на B е

$$\begin{vmatrix} 1 & r & -q \\ -r & 1 & p \\ q & -p & 1 \end{vmatrix} = 1 + p^2 + q^2 + r^2.$$

Тогава по формулите на Кейли ортогоналната детерминанта ще бъде

$$\begin{vmatrix} \frac{1+p^2-q^2-r^2}{B} & \frac{2(r+pq)}{B} & \frac{2(-q+pr)}{B} \\ \frac{2(-r+qp)}{B} & \frac{1-p^2+q^2-r^2}{B} & \frac{2(p+qr)}{B} \\ \frac{2(q+rp)}{B} & \frac{2(-p+rq)}{B} & \frac{1-p^2-q^2+r^2}{B} \end{vmatrix}.$$

Ако с α, β, γ означим три числа, подчинени на условието $\alpha^2+\beta^2+\gamma^2=1$, то винаги можем да поставим

$$p=\alpha \operatorname{tg}\frac{\theta}{2}, \quad q=\beta \operatorname{tg}\frac{\theta}{2}, \quad r=\gamma \operatorname{tg}\frac{\theta}{2}.$$

Горната детерминанта се обръща в

$$\begin{vmatrix} \cos\theta+\alpha^2(1-\cos\theta) & \gamma\sin\theta+\alpha\beta(1-\cos\theta) & -\beta\sin\theta+\alpha\gamma(1-\cos\theta) \\ -\gamma\sin\theta+\beta\alpha(1-\cos\theta) & \cos\theta+\beta^2(1-\cos\theta) & \alpha\sin\theta+\beta\gamma(1-\cos\theta) \\ \beta\sin\theta+\gamma\alpha(1-\cos\theta) & -\alpha\sin\theta+\gamma\beta(1-\cos\theta) & \cos\theta+\gamma^2(1-\cos\theta) \end{vmatrix}.$$

10. **Безкрайни детерминанти. Теорема на Поанкаре.** В някои въпроси от математиката играят важна роля тъй наречените безкрайни детерминанти, т. е. детерминанти, които имат безкрайно много редове и стълбове:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} & \cdot & \cdot & \cdot \\ a_{31} & a_{32} & a_{33} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}.$$

Нека от първите n реда и стълба образуваме една детерминанта от n-ти ред D_n. Тогава, ако редицата $D_n (n=1, 2, 3, \ldots)$ клони към определена граница D, то казваме, че D е сходяща детерминанта от безкраен ред. Пръв Поанкаре даде достатъчното условие за сходимост на D.

Теорема на Поанкаре. Ако главните елементи a_{ii} са равни на 1 и редът $\sum_{i \neq k} |a_{ik}|$ е сходящ, то D е сходяща.

Действително детерминантата D_n се получава от развитието на произведението

$$\Pi_n = (1+|a_{12}|+|a_{13}|+\ldots+|a_{1,n}|).$$
$$(1+|a_{21}|+|a_{23}|+\ldots+|a_{2,n}|).$$
$$\ldots\ldots\ldots\ldots$$
$$(1+|a_{n1}|+\ldots+|a_{n,n-1}|),$$

като някои членове се вземат със знаците $+$, $-$ или се премахват. Така че очевидно $|D_n|<\Pi_n$.

Ако поставим в D_{n+p} или също в Π_{n+p} някои елементи, равни на нула, то получаваме D_n и Π_n. Така анулираните членове именно представят съответно разликите $D_{n+p}-D_n$ и $\Pi_{n+p}-\Pi_n$. От друга страна, членовете в Π_{n+p}, които привеждаме към нула, са положителни и някои от тях, като ги вземем с подходящ знак, съвпадат с членовете, които трябва да се анулират в D_{n+p}, за да се получи D_n. От това веднага заключаваме неравенството

$$|D_{n+p}-D_n|<\Pi_{n+p}-\Pi_n. \tag{23}$$

От друга страна, известно е, че безкрайното произведение

$$P=(1+u_1)(1+u_2)(1+u_3)\ldots$$

е сходящо, ако редът $\sum_{n=1}^{\infty}|u_n|$ е сходящ. Понеже редът

$$\sum_{i\neq j}|a_{ij}|$$

е сходящ, то сходящо е и произведението Π_n. По теоремата на Коши от (23) следва сходимостта на D_n. Ако диагоналните елементи a_{ii} не са равни на 1, то, както лесно се вижда, сходимостта се запазва, стига да допуснем, че още редът $\sum_{i=1}^{\infty}|1-a_{ii}|$ е сходящ. Действително в този случай разглеждаме произведението

$$\Pi_n=(1+|c_{11}|+|a_{12}|+\ldots). \qquad a_{ii}=1+c_{ii},$$
$$(1+|c_{22}|+|a_{21}|+\ldots).$$
$$\ldots\ldots\ldots\ldots$$
$$(1+|c_{nn}|+|a_{n1}|+\ldots+|a_{n,n-1}|)$$

и повтаряме същите разсъждения. Така получаваме теоремата на Хелге фон Кох. Детерминантата е сходяща, ако редовете

$$\sum_{i=1}^{\infty}|1-a_{ii}|,\ \sum_{i\neq j}|a_{ij}|$$

са сходящи.

Тези безкрайни детерминанти се наричат нормални.

Лесно се вижда от доказателството, че горните теореми остават в сила и когато елементите a_{ik} са комплексни числа.

11. **Теорема на Хадамар за максималната стойност на една детерминанта.** Нека

$$A=\begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \cdot & \cdot \cdot \cdot \cdot \cdot \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}$$

е една детерминанта с произволни реални или имагинерни елементи. Имаме следната теорема на Хадамар: Квадратът на модула на A е най-много равен на произведението от сумите от квадратите на модулите на елементите от n-те реда, т. е.

$$|A|^2 \leqq (|a_{11}|^2+|a_{12}|^2+\dots+|a_{1n}|^2)\dots(|a_{n1}|^2+|a_{n2}|^2+\dots+|a_{nn}|^2).$$

Нека означим, както по-рано приехме, с $\bar{\alpha}$ конюгованото число на α. Тогава очевидно ще имаме

$$\bar{A}=\begin{vmatrix} \bar{a}_{11} & \bar{a}_{12} \dots \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22} \dots \bar{a}_{2n} \\ \cdot & \cdot \cdot \cdot \cdot \cdot \\ \bar{a}_{n1} & \bar{a}_{n2} \dots \bar{a}_{nn} \end{vmatrix}.$$

Ще преработим детерминантата A, като я сведем на една ортогонална детерминанта. Нека B е ортогонална детерминанта с реални или имагинерни елементи:

$$B=\begin{vmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2n} \\ \cdot & \cdot \cdot \cdot \cdot \cdot \\ b_{n1} & b_{n2} \dots b_{nn} \end{vmatrix}.$$

Тогава между елементите ѝ има връзките

$$b_{i1}\bar{b}_{k1}+b_{i2}\bar{b}_{k2}+\dots+b_{in}\bar{b}_{kn}=0, \qquad i\neq k. \tag{24}$$

Нека B се получава от A, като положим $b_{1i}=a_{1i}$, $i=1, 2, \dots, n$, $b_{2i}=a_{2i}-m_{21}b_{1i}$, и въобще

$$b_{ri}=a_{ri}-\sum_{k=1}^{r-1} m_{rk}b_{ki}. \tag{25}$$

Очевидно $B=A$, понеже редовете на B се получават, като от съответните редове на A се извадят линейни комбинации на предшеству-

ващите редове. Ако умножим (25) с $\bar{b}_{si}$, $s<r$ и сумираме по i, то на основание на (24) ще имаме

$$m_{rs}\sum_{i=1}^{n} b_{si}\bar{b}_{si}=\sum_{i-1}^{n} a_{ri}\bar{b}_{si}, \tag{26}$$

която релация ни дава m_{rs}. От формулите (25) и (26), като вземем конюгованите стойности, ще имаме

$$\bar{b}_{ri}=\bar{a}_{ri}-\sum_{k=1}^{r-1}\bar{m}_{rk}\bar{b}_{ki}, \tag{27}$$

$$\bar{m}_{rs}\sum_{i=1}^{n} b_{si}\bar{b}_{si}=\sum_{i=1}^{n}\bar{a}_{ri}b_{si}. \tag{28}$$

Като умножим сега B с $\bar{B}$ по редове, ще получим

$$A\bar{A}=B\bar{B}=|c_{pq}|,$$

гдето $c_{pp}=\sum_{k=1}^{n} b_{pk}\bar{b}_{pk}, c_{pq}=\sum_{k=1}^{n} b_{pk}\bar{b}_{qk}=0$ при $p\neq q$ на основание на (24).

В детерминантата $|c_{pq}|$ освен елементите на главния диагонал всички други са равни на нула. Следователно ще имаме

$$|A|^2=\prod_{p=1}^{n}\left(\sum_{k=1}^{n}|b_{pk}|^2\right). \tag{29}$$

От друга страна, като умножим b_{ri} с $\bar{b}_{ri}$, използувайки изразите им (25) и (27), получаваме

$$b_{ri}\bar{b}_{ri}=a_{ri}\bar{a}_{ri}-\sum_{s=1}^{r-1}\bar{m}_{rs}a_{rs}\bar{b}_{si}-\sum_{s=1}^{r-1}m_{rs}\bar{a}_{ri}b_{si}+\sum_{s=1}^{r-1}\sum_{t=1}^{r-1}m_{rs}\bar{m}_{rt}b_{si}\bar{b}_{ti},$$

отгдето със сумиране спрямо i, използувайки (24) и (28), ще имаме

$$\sum_{i=1}^{n}|b_{ri}|^2=\sum_{i=1}^{n}|a_{ri}|^2-\sum_{s=1}^{r-1}|m_{rs}|^2\sum_{i=1}^{n}|b_{si}|^2.$$

Оттук е очевидно, че

$$\sum_{i=1}^{n}|b_{ri}|^2\leqq\sum_{i=1}^{n}|a_{ri}|^2, \tag{30}$$

гдето равенство може да имаме само тогава, ако $m_{ri}=0$, т. е. когато A е ортогонална детерминанта. От (29) и (20) следва веднага теоремата на Хадамар.

Ако елементите a_{ik} са реални, при $n=2,3$ теоремата има прост геометричен смисъл. Действително, ако 0 е началото на координатната система в равнината ($n=2$), A_1 е точка с координати (a_{11}, a_{12}) и $A_2(a_{21}, a_{22})$, теоремата се свежда на това, че лицето на паралелограма, построен върху OA_1 и OA_2 при постоянни дължини на страните е само тогава максимално, когато той е правоъгълник, което се установява и директно на основание на прости геометрични свойства. Ако $A_k(a_{k1}, a_{k2}, a_{k3})$, $k=1, 2, 3$ са три точки в пространството ($n=3$), теоремата се свежда на факта, че обемът на паралелепипеда, построен върху OA_1, OA_2, OA_3 е само тогава максимален, когато той е правоъгълен. При произволно n теоремата представя обобщение на споменатото свойство в пространство с много измерения.

Глава V

Матрици и действия с тях

1. Събиране на матрици и умножение с число. Както в началото видяхме, под матрица от nm числа a_{ij}, $i=1, 2, \ldots, n$; $j=1, 2, \ldots, m$ разбираме системата от тези числа, наредени в правоъгълна таблица

$$\mathbf{A}=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \ldots a_{1m} \\ a_{21} & a_{22} & a_{23} \ldots a_{2m} \\ a_{31} & a_{32} & a_{33} \ldots a_{3m} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} \ldots a_{nm} \end{Vmatrix}. \tag{1}$$

Матрицата **A** се нарича от типа (n, m). Числата a_{ij} се наричат елементи на матрицата. Матрицата се нарича нулева, ако всичките ѝ елементи са равни на нула. Нулевите матрици бележим с **O**. Под сума на матриците **A** и **B**, които са от един и същ тип, разбираме матрица, на която всеки елемент е сума от съответните елементи на двете матрици, т. е. ако **A** е матрицата (1) и **B** е матрицата

$$\mathbf{B}=\begin{Vmatrix} b_{11} & b_{12} & b_{13} \ldots b_{1m} \\ b_{21} & b_{22} & b_{23} \ldots b_{2m} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & b_{n3} \ldots b_{nm} \end{Vmatrix},$$

то сумата на двете матрици **A** и **B**, която бележим с **A**+**B**, е матрицата

$$\left\| \begin{array}{cccc} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \ldots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \ldots & a_{2m}+b_{2m} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & & \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & & \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & a_{n3}+b_{n3} \ldots & a_{nm}+b_{nm} \end{array} \right\|.$$

Очевидно е, че комутативният закон за събирането, както и асоциативният закон са в сила, т. е. имаме

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}, \quad \mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}.$$

Ако g е произволно число, под $g\mathbf{A}$ или $\mathbf{A}g$ разбираме матрица, получена от матрицата **A** с умножение на всичките ѝ елементи с g. Лесно се вижда, че

$$g(\mathbf{A}+\mathbf{B})=g\mathbf{A}+g\mathbf{B}, \quad \mathbf{A}+\mathbf{O}=\mathbf{A}, \quad \mathbf{O}\,.\,\mathbf{A}=\mathbf{O}, \quad (g+k)\,\mathbf{A}=g\mathbf{A}+k\mathbf{A}.$$

Две матрици **A** и **B** считаме само тогава равни, когато имат равни съответни елементи. Равенството на матриците означаваме с $\mathbf{A}=\mathbf{B}$. Матрицата $(-1)\mathbf{B}$ означаваме накъсо с $-\mathbf{B}$ и матрицата $\mathbf{A}+(-1)\mathbf{B}$ означаваме с $\mathbf{A}-\mathbf{B}$—разлика на матриците **A** и **B**. Естествено изваждането на матриците можем да дефинираме като обратно действие на събирането.

Ако матрицата има еднакъв брой редове и стълбове, то тя се нарича квадратна. Ако в една квадратна матрица всичките елементи, които не лежат на главния диагонал, са равни на нула, тя се нарича диагонална. В частност, ако в диагоналната матрица елементите на главния диагонал са равни помежду си, то тя се нарича скалар (скаларна матрица). Ако в последната матрица елементите на главния диагонал са равни на 1, матрицата се нарича единица и се бележи с **E**.

2. Умножение на матрици. Нека са дадени n линейни функции на m променливи

$$\begin{aligned} u_1 &= a_{11}x_1+a_{12}x_2+\ldots+a_{1m}x_m, \\ u_2 &= a_{21}x_1+a_{22}x_2+\ldots+a_{2m}x_m, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ u_n &= a_{n1}x_1+a_{n2}x_2+\ldots+a_{nm}x_m. \end{aligned} \tag{2}$$

Върху променливите $x_1, x_2, \ldots, x_m$ да извършим линейната субституция

$$\begin{aligned} x_1 &= b_{11}y_1+b_{12}y_2+\ldots+b_{1r}y_r, \\ x_2 &= b_{21}y_1+b_{22}y_2+\ldots+b_{2r}y_r, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_m &= b_{m1}y_1+b_{m2}y_2+\ldots+b_{mr}y_r. \end{aligned} \tag{3}$$

Линейните функции (2) стават

$$\begin{aligned} u_1 &= c_{11}y_1 + c_{12}y_2 + \ldots + c_{1r}y_r, \\ u_2 &= c_{21}y_1 + c_{22}y_2 + \ldots + c_{2r}y_r, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u_n &= c_{n1}y_1 + c_{n2}y_2 + \ldots + c_{nr}y_r, \end{aligned} \tag{4}$$

където имаме

$$\begin{aligned} &c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{im}b_{mk}, \\ &i = 1, 2, \ldots, n; \quad k = 1, 2, \ldots, r. \end{aligned} \tag{5}$$

Да означим с **A**, **B** и **C** съответно матриците, образувани от коефициентите на променливите и линейните форми (2), (3) и (4). Матрицата **C** се нарича произведение на матриците **A** и **B** и се бележи с **AB**. Елементите ѝ се получават по формулата (5), т. е. като умножаваме редовете на **A** със стълбовете на **B**. Очевидно между типа на матриците **A** и **B** има известна връзка. Именно броят на редовете на **B** е равен на броя на стълбовете на **A**. Ако матриците **A** и **B** са квадратни, то детерминантата **C** от елементите на матрицата **C** е равна на произведението на детерминантите A и B от матриците **A** и **B**. Детерминантата B се нарича модул на линейната субституция (3) и детерминантата A—инвариант на линейните форми (2).

Умножението е въобще некомутативно. Така например за матриците

$$\mathbf{A} = \begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}, \quad \mathbf{B} = \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

имаме

$$\mathbf{AB} = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}, \quad \mathbf{BA} = \begin{Vmatrix} 0 & -1 \\ 1 & 0 \end{Vmatrix}$$

и следователно $\mathbf{AB} \neq \mathbf{BA}$. Произведението на двете матрици може да бъде равно на нула, без никоя от тях да е нула. Например

$$\begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 0 \\ 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}.$$

Ще установим сега, че асоциативният закон за умножението е в сила, т. е. за кои да е три матрици имаме

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} = \mathbf{ABC}.$$

Действително нека $\mathbf{A} = (a_{ik})$ е от типа (n, m), $\mathbf{B} = (b_{ik})$ — от типа (m, r) и $\mathbf{C} = (c_{ik})$ — от типа (r, s). За елементите на матриците

$$\mathbf{AB} = (\alpha_{ik}), \quad \mathbf{BC} = (\beta_{ik}),$$

които са съответно от типа (n, r) и (m, s), ще имаме

$$\alpha_{ik}=\sum_{\tau=1}^{m} a_{i\tau}\, b_{\tau k},\ \beta_{ik}=\sum_{\tau=1}^{r} b_{i\tau}\, c_{\tau k}.$$

Ако $\mathbf{A}(\mathbf{BC})=(\gamma_{ik})$ и $(\mathbf{AB})\,\mathbf{C}=(\delta_{ik})$, то за елементите на тези матрици получаваме

$$\gamma_{ik}=\sum_{h=1}^{m} a_{ih}\ \beta_{hk}=\sum_{h=1}^{m}\sum_{\tau=1}^{r} a_{ik}\, b_{h\tau}\, c_{\tau k},$$

$$\delta_{ik}=\sum_{\tau=1}^{r} \alpha_{i\tau}\ c_{\tau k}=\sum_{\tau=1}^{r}\sum_{h=1}^{m} c_{\tau k}\, a_{ih}\, b_{h\tau}.$$

От предните равенства очевидно следва, че $\gamma_{ik}=\delta_{ik}$, $i=1, 2,\ldots, n$; $k=1, 2,\ldots, s$.

Произведението е дистрибутивно. Така за три матрици **A**, **B** и **C** от един и същ тип имаме

$$(\mathbf{A}+\mathbf{B})\,\mathbf{C}=\mathbf{AC}+\mathbf{BC},\quad \mathbf{C}\,(\mathbf{A}+\mathbf{B})=\mathbf{CA}+\mathbf{CB}.$$

Доказателството е лесно. Например първото равенство следва от тъждествата

$$\sum (a_{i\tau}+b_{i\tau})\, c_{\tau k}=\sum a_{i\tau}\, c_{\tau k}+\sum b_{i\tau}\, c_{\tau k}.$$

Очевидно за коя да е квадратна матрица A имаме

$$\mathbf{AE}=\mathbf{EA}=\mathbf{A}.$$

Две матрици **A** и **B** се наричат комутативни, ако произведението им остава също при разместване на множителите, т. е. когато $\mathbf{AB}=\mathbf{BA}$. Очевидно всяка квадратна матрица е комутативна с единичната матрица от същия ред. По-общо всяка квадратна матрица е комутативна с всяка скаларна матрица от същия ред. Обратно, ще установим, че ако една квадратна матрица **A** е комутативна с всяка квадратна матрица от същия ред, то матрицата **A** е скаларна.

Действително нека **A** и **B** са матриците

$$\mathbf{A}=\begin{Vmatrix} a_{11} & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}\ldots a_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2}\ldots a_{nn} \end{Vmatrix},\qquad \mathbf{B}=\begin{Vmatrix} b_{11} & b_{12} & b_{13}\ldots b_{1n} \\ b_{21} & b_{22} & \ldots\ b_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ b_{n1} & b_{n2} & \ldots\ b_{nn} \end{Vmatrix},$$

като предполагаме, че равенството $\mathbf{AB}=\mathbf{BA}$ е изпълнено за всяка матрица **B**. Като вземем пред вид правилото за умножение на матрици, получаваме равенствата

$$a_{i1}\, b_{1k}+a_{i2}\, b_{2k}+\ldots+a_{ii}\, b_{ik}+\ldots+a_{in}\, b_{kn}=b_{i1}\, a_{1k}+b_{i2}\, a_{2k}+\ldots+$$
$$+b_{ik}\, a_{kk}+\ldots+b_{in}\, a_{nk},$$

$$\vdots$$

които трябва да бъдат изпълнени, каквито и да са числата b_{ij}, $i, j = 1, 2, \ldots, n$. Оттук очевидно следва, че $a_{ij}=0$ при $i \neq j$, $b_{ik} a_{kk} = b_{ik} a_{ij}$, откъдето получаваме

$$a_{kk} = a_{ii},$$

които равенства показват, че $a_{11} = a_{22} = \ldots = a_{nn}$.

3. Степен на матрица. В разглежданията ще предполагаме, че матриците са квадратни и r ще означава реда им. Под n-та степен на матрицата **A** (n естествено число) разбираме произведението на n матрици, равни на **A**. Тази степен бележим с $\mathbf{A}^n$. Очевидно за произволни естествени числа n и m ще имаме

$$\mathbf{A}^n \mathbf{A}^m = \mathbf{A}^{n+m}.$$

Една матрица наричаме особена, ако детерминантата, образувана от елементите ѝ, е равна на нула. Ако въпросната детерминанта е отлична от нула, матрицата се нарича неособена. Нека матрицата

$$\mathbf{A} = \begin{Vmatrix} a_{11} & a_{12} \ldots a_{1r} \\ a_{21} & a_{22} \ldots a_{2r} \\ \cdot & \cdot \quad \cdot \quad \cdot \\ a_{r1} & a_{r2} \ldots a_{rr} \end{Vmatrix}$$

е неособена. Ако с $A_{\mu\nu}$ означим адюнгираното количество на елемента $a_{\mu\nu}$ в детерминантата A от матрицата **A**, да означим с $\mathbf{A}_1$ матрицата

$$\mathbf{A}_1 = \begin{Vmatrix} \frac{A_{11}}{A} & \frac{A_{21}}{A} \ldots \frac{A_{r1}}{A} \\ \frac{A_{12}}{A} & \frac{A_{22}}{A} \ldots \frac{A_{r2}}{A} \\ \cdot & \cdot \quad \cdot \quad \cdot \\ \frac{A_{1r}}{A} & \frac{A_{2r}}{A} \ldots \frac{A_{rr}}{A} \end{Vmatrix}.$$

Лесно се вижда, че

$$\mathbf{A}\mathbf{A}_1 = \mathbf{A}_1\mathbf{A} = \begin{Vmatrix} 1 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 \ldots 0 \\ 0 & 0 & 1 \ldots 0 \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ 0 & 0 & 0 \ldots 1 \end{Vmatrix} = \mathbf{E}.$$

Матрицата $\mathbf{A}_1$ се нарича обратна на **A** и се бележи с $\mathbf{A}^{-1}$. Детерминантата от матрицата $\mathbf{A}^{-1}$ е равна на A^{-1}, като с A сме означили детерминантата от матрицата **A**. Непосредствено се вижда, че **A** е също обратна на $\mathbf{A}^{-1}$, т. е. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$. Ще докажем и следното свойство:

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\ \mathbf{B}^{-1}\ \mathbf{A}^{-1}.$$

За тази цел да означим с **M** матрицата $\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$. Имаме

$$\mathbf{MA}=\mathbf{C}^{-1}\mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})=\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{E}=\mathbf{C}^{-1}\mathbf{B}^{-1},\quad \mathbf{MAB}=\mathbf{C}^{-1},\quad \mathbf{M}.\mathbf{ABC}=\mathbf{E},$$

с което свойството е установено.

Нека **A** е неособена матрица. Да намерим матрица **X**, която удовлетворява уравнението

$$\mathbf{AX}=\mathbf{E}.$$

С умножение вляво с $\mathbf{A}^{-1}$ получаваме

$$\mathbf{A}^{-1}\mathbf{AX}=\mathbf{A}^{-1}\mathbf{E},\ \mathbf{X}=\mathbf{A}^{-1}\mathbf{E},$$

откъдето имаме $\mathbf{X}=\mathbf{A}^{-1}$. Подобно за решението на уравнението $\mathbf{YA}=\mathbf{E}$ получаваме $\mathbf{Y}=\mathbf{A}^{-1}$.

Ако m е естествено число и **A** е неособена матрица, под степен $\mathbf{A}^{-m}$ разбираме m-тата степен на обратната матрица $\mathbf{A}^{-1}$, т. е. $\mathbf{A}^{-m}=(\mathbf{A}^{-1})^m$. Ако сега под $\mathbf{A}^0$ разбираме матрицата единица, то правилото за умножение на степени

$$\mathbf{A}^p\mathbf{A}^q=\mathbf{A}^{p+q},\ (\mathbf{A}^p)^q=\mathbf{A}^{pq}$$

остава в сила за каквито и да е цели числа p и q.

Видяхме, че произведението на две матрици може да е равно на нула без матриците да са нулеви. Ако обаче едната матрица е неособена, другата трябва непременно да е равна на нула. Действително нека **A** е неособена матрица и за някоя матрица **B** имаме

$$\mathbf{AB}=\mathbf{O}.$$

Като умножим това равенство с $\mathbf{A}^{-1}$, получаваме

$$\mathbf{A}^{-1}.\mathbf{AB}=\mathbf{A}^{-1}.\mathbf{O}=\mathbf{O},$$

т. е. $\mathbf{B}=\mathbf{O}$.

На основание на събиране и умножение на матрици можем да дефинираме и полином от една матрица **A**. Именно, ако

$$F(x)=a_0+a_1x+\ldots+a_nx^n$$

е произволен полином, то под полином $F(\mathbf{A})$ на **A** разбираме матрицата

$$F(\mathbf{A})=a_0\mathbf{E}+a_1\mathbf{A}+a_2\mathbf{A}^2+\ldots+a_n\mathbf{A}^n.$$

Казваме също, че $F(\mathbf{A})$ е значението на полинома $F(x)$ за $x=\mathbf{A}$. Върху полиномите от една матрица можем да извършваме събиране и умножение. Така нека са дадени полиномите

$$F(x)=b_0+b_1x+\ldots+b_nx^n$$
$$\varphi(x)=c_0+c_1x+\ldots+c_nx^n,$$

като за удобство сме приели степените им за равни, което не е ограничение, понеже полиномът от по-ниска степен можем да допълним с членове, равни на нула. За сумата и произведението им получаваме

$$g(x)=f(x)+\varphi(x)=b_0+c_0+(b_1+c_1)x+\ldots+(b_n+c_n)x^n$$
$$h(x)=f(x)\varphi(x)=b_0c_0+(b_0c_1+b_1c_0)x+\ldots+(b_0c_n+b_1c_{n-1}+\ldots+b_nc_0)x^n.$$

На основание на правилата за събиране и умножение на матрици получаваме тогава

$$f(\mathbf{A})+\varphi(\mathbf{A})=(b_0+c_0)\mathbf{E}+(b_1+c_1)\mathbf{A}+\ldots+(b_n+c_n)\mathbf{A}^n=g(\mathbf{A}),$$

$$f(\mathbf{A})\varphi(\mathbf{A})=b_0c_0\mathbf{E}+(b_0c_1+b_1c_0)\mathbf{A}+\ldots+(b_0c_n+b_1c_{n-1}+\ldots+b_nc_0)\mathbf{A}^n=h(\mathbf{A}).$$

Следователно значението на сумата от двата полинома $f(x)$ и $\varphi(x)$ за $x=\mathbf{A}$ е равно на сумата от значенията на тези полиноми за $x=\mathbf{A}$ и значението на произведението на двата полинома за $x=\mathbf{A}$ е равно на произведението от значенията на същите полиноми.

Така например от равенството

$$x^2-1=(x-1)(x+1)$$

следва, че

$$\mathbf{A}^2-\mathbf{E}=(\mathbf{A}-\mathbf{E})(\mathbf{A}+\mathbf{E})$$

и от

$$(x+1)^3=x^3+3x^2+3x+1$$

следва, че

$$(\mathbf{A}+\mathbf{E})^3=\mathbf{A}^3+3\mathbf{A}^2+3\mathbf{A}+\mathbf{E}.$$

Сумата и произведението на полиноми от дадена матрица очевидно не се променят при промяна на реда на събирането и на умножението. Също така и дистрибутивният закон за умножението е в сила.

4. Транспонирана и други видове матрици. Ако в една матрица

$$\mathbf{A}=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \ldots a_{1m} \\ a_{21} & a_{22} & a_{23} \ldots a_{2m} \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ a_{n1} & a_{n2} & a_{n3} \ldots a_{nm} \end{Vmatrix}$$

сменим редовете със стълбовете и, обратно, получаваме матрицата

$$\begin{Vmatrix} a_{11} & a_{21} \ldots a_{n1} \\ a_{12} & a_{22} \ldots a_{n2} \\ a_{13} & a_{23} \ldots a_{n3} \\ \cdot & \cdot \quad \cdot \quad \cdot \\ a_{1m} & a_{2m} \ldots a_{nm} \end{Vmatrix},$$

която се нарича транспонирана на дадената матрица. Нея бележим с $\mathbf{A}'$. От определението следва, че транспонираната матрица на $\mathbf{A}'$ е матрицата $\mathbf{A}$, т. е. $(\mathbf{A}')'=\mathbf{A}$. Ако α е произволно число и $\mathbf{B}$ е матрица от типа на $\mathbf{A}$, то лесно се вижда, че ще имаме

$$(\alpha\mathbf{A})'=\alpha\mathbf{A}',\ (\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}',$$

откъдето за произволни числа α и β получаваме

$$(\alpha\mathbf{A}+\beta\mathbf{B})'=\alpha\mathbf{A}'+\beta\mathbf{B}'.$$

Нека матрицата **A** е от типа (n, m) и матрицата **B** е от типа (m, p). Ще докажем равенството

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'. \tag{6}$$

Действително елементът c'_{ij} на произведението $B'A'$, който лежи на i-тия ред и j-тия стълб, е равен на

$$c'_{ij} = b_{1i}a_{j1} + b_{2i}a_{j2} + \ldots + b_{mi}a_{jm}.$$

Но тогава очевидно $c'_{ij} = c_{ji}$ е елементът на произведението **AB**, който лежи на j-тия ред и i-тия стълб в тази матрица.

Нека **A** е една неособена квадратна матрица. От равенството

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{E}$$

на основание на (6) получаваме

$$(\mathbf{A}^{-1})'\ \mathbf{A}' = \mathbf{E}.$$

Следователно имаме

$$(\mathbf{A}^{-1})' = \mathbf{A}'^{-1},$$

т. е. транспонираната матрица на обратната матрица на **A** *е обратната матрица на транспонираната на* **A** *матрица.*

Една матрица **A** се нарича симетрична, ако $\mathbf{A}' = \mathbf{A}$, и полусиметрична (или кососиметрична), ако $\mathbf{A}' = -\mathbf{A}$. Естествено матриците се предполагат квадратни. Ако матрицата **A** е полусиметрична за елементите по главния ѝ диагонал, ще имаме $a_{ii} = -a_{ii}$, т. е. те са равни на нула. Нека **A** и **B** са две симетрични или полусиметрични матрици едновременно от един и същ ред. Ако α и β са произволни числа, то матрицата $\alpha\mathbf{A} + \beta\mathbf{B}$ е също симетрична (респ. полусиметрична). Произведението на две симетрични матрици не винаги е симетрична матрица. *Ако обаче матриците са комутативни, то произведението им е симетрична матрица.* Това се вижда от равенството

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}' = \mathbf{BA} = \mathbf{AB}.$$

Една квадратна матрица **A** се нарича ортогонална, ако

$$\mathbf{A}\mathbf{A}' = \mathbf{E}. \tag{7}$$

От предното равенство следва, че детерминантата **A** от матрицата **A** е ортогонална детерминанта, която има стойност, равна на 1 или — 1. Също от (7) имаме $\mathbf{A}' = \mathbf{A}^{-1}$, т. е. транспонираната матрица на **A** е обратната матрица на **A**. *Транспонираната матрица на* **A** *е също ортогонална.* Това следва от равенствата

$$(\mathbf{A}^{-1})' = (\mathbf{A}')' = \mathbf{A} = (\mathbf{A}^{-1})^{-1}.$$

Произведението на две ортогонални матрици е също ортогонална матрица. Действително имаме

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}' = \mathbf{B}^{-1}\mathbf{A}^{-1} = (\mathbf{AB})^{-1}.$$

Ако **A** е матрица с елементи, които са комплексни числа, то под $\bar{\mathbf{A}}$ разбираме матрица, чиито елементи са конюгованите стойности на елементите на **A**.

Една матрица **A** с комплексни елементи се нарича Хермитова, ако имаме

$$\mathbf{A}=\bar{\mathbf{A}}'.$$

За елементите на главния ѝ диагонал a_{pp} следва, че ще имаме $a_{pp}=\bar{a}_{pp}$, т. е. тези елементи са реални числа. Ако α е произволно реално число, матрицата $\alpha\mathbf{A}$ е също така Хермитова. *Ако две Хермитови матрици* **A** *и* **B** *са комутативни, то произведението им* **AB** *е също Хермитова матрица.* Това свойство следва от равенствата

$$(\mathbf{AB})'=\mathbf{B}'\mathbf{A}'=\bar{\mathbf{B}}\bar{\mathbf{A}}=\bar{\mathbf{A}}\bar{\mathbf{B}}=\overline{\mathbf{AB}}.$$

Една матрица **A** с комплексни елементи се нарича унитарна, ако

$$\mathbf{A}\bar{\mathbf{A}}'=\mathbf{E}.$$

От предното равенство за детерминантата получаваме

$$|\mathbf{A}|\,|\bar{\mathbf{A}}'|=1$$

и понеже $|\bar{\mathbf{A}}'|=|\mathbf{A}|$, то следва, че $|\mathbf{A}|^2=1$, т. е. $|\mathbf{A}|=\pm 1$.

Матрицата **B** се нарича подобна на матрицата **A**, ако съществува неособена матрица **X**, така че имаме

$$\mathbf{B}=\mathbf{X}^{-1}\mathbf{A}\mathbf{X}. \tag{8}$$

Матрицата **B** се нарича и трансформирана на **A** посредством **X**. Ако умножим равенството (8) отдясно с $\mathbf{X}^{-1}$, а отляво с **X**, получаваме

$$\mathbf{X}\mathbf{B}\mathbf{X}^{-1}=\mathbf{A}$$

или

$$\mathbf{A}=(\mathbf{X}^{-1})^{-1}\mathbf{B}\mathbf{X}^{-1}.$$

Следователно матрицата **A** е подобна на матрицата **B**. Нека **B** и **C** са подобни на матрицата **A**, т. е. имаме

$$\mathbf{B}=\mathbf{X}^{-1}\mathbf{A}\mathbf{X},$$

$$\mathbf{C}=\mathbf{Y}^{-1}\mathbf{A}\mathbf{Y}.$$

От първото равенство получаваме, както преди $\mathbf{A}=\mathbf{X}\mathbf{B}\mathbf{X}^{-1}$. Като заместим във второто равенство **A** с предната стойност, получаваме

$$\mathbf{C}=\mathbf{Y}^{-1}\mathbf{X}\mathbf{B}\mathbf{X}^{-1}\mathbf{Y}=(\mathbf{X}^{-1}\mathbf{Y})^{-1}\mathbf{B}(\mathbf{X}^{-1}\mathbf{Y}),$$

т. е. матрицата **C** е подобна на матрицата **B**. *Следователно, ако две матрици са подобни на трета, то те са подобни помежду си.*

От лесно доказуемите равенства

$$\mathbf{X}^{-1}(\mathbf{A}_1+\mathbf{A}_2+\ldots+\mathbf{A}_m)\mathbf{X}=\mathbf{X}^{-1}\mathbf{A}_1\mathbf{X}+\mathbf{X}^{-1}\mathbf{A}_2\mathbf{X}+\ldots+\mathbf{X}^{-1}\mathbf{A}_m\mathbf{X},$$

$$\mathbf{X}^{-1}\mathbf{A}_1\mathbf{X}\,.\,\mathbf{X}^{-1}\mathbf{A}_2\mathbf{X}\ldots\mathbf{X}^{-1}\mathbf{A}_m\mathbf{X}=\mathbf{X}^{-1}(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_m)\mathbf{X}$$

следват свойствата:

Трансформираната матрица на сума от матрици е сума от трансформираните матрици на събираемите и трансформираната матрица на произведение от матрици е произведение от трансформираните матрици на множителите.

Матриците, състоящи се само от един стълб (с m елемента в него), се наричат m-членни вектори.

Ако с $\mathbf{r}$, $\mathbf{u}$ и $\mathbf{h}$ означим векторите

$$\mathbf{r}=\begin{Vmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ x_m \end{Vmatrix}, \qquad \mathbf{u}=\begin{Vmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \\ u_n \end{Vmatrix}, \qquad \mathbf{h}=\begin{Vmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \\ y_r \end{Vmatrix},$$

то от (2) (3) и (4) имаме

$$\mathbf{u}=\mathbf{Ar},\ \mathbf{r}=\mathbf{Bh},\ \mathbf{u}=\mathbf{ABh}.$$

За транспонираната матрица на r очевидно имаме

$$\mathbf{r}'=\| x_1 x_2 \ldots x_m \|.$$

Произведението $\mathbf{r}'\mathbf{h}$ (при $r=m$) се нарича скаларно (вътрешно) произведение на векторите $\mathbf{r}$ и $\mathbf{h}$.

5. Характеристично уравнение и теорема на Хамилтон-Кейли. Нека $\mathbf{A}$ е матрица от n-ти ред с елементи $a_{\mu\nu}$. Уравнението

$$(9)\qquad \varphi(x)=x\mathbf{E}-\mathbf{A}|=\begin{vmatrix} x-a_{11} & -a_{12}\ldots -a_{1n} \\ -a_{21} & x-a_{22}\ldots -a_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ -a_{n1} & -a_{n2}\ldots x-a_{nn} \end{vmatrix}=0$$

се нарича характеристично уравнение на матрицата $\mathbf{A}$. Очевидно произведението $(x-a_{11})\ (x-a_{22})\ldots(x-a_{nn})$ на елементите на главния диагонал съдържа x в най-високата степен. Понеже в другите членове в развитието на детерминантата (9) влизат най-много $n-2$ линейни на x множители, то полиномът $\varphi(x)$ има вида

$$\varphi(x)=x^n-(a_{11}+a_{22}+\ldots+a_{nn})x^{n-1}+\ldots$$

Следователно уравнението (9) е от n-та степен с коефициент пред x^n, равен на 1. Коефициентът пред x^{n-1} е равен на сумата от елементите на главния диагонал, взета с обратен знак. Въпросната сума се нарича следа на матрицата A. Корените на уравнението (9) се наричат характеристични числа на матрицата или собствени стойности. Лявата част на характеристическото уравнение се нарича характеристически полином.

Нека $A_{\mu\nu}$ означава адюнгираното количество на елемента от μ-тия ред и ν-тия стълб на детерминантата (9). Очевидно $A_{\mu\nu}$ ще представлява полином на x от степен $\leqq n-1$,

$$A_{\mu\nu}=\alpha_{\mu\nu}^{(0)}+\alpha_{\mu\nu}^{(1)}x+\ldots+\alpha_{\mu\nu}^{(n-1)}x^{n-1}. \tag{10}$$

Да умножим сега матрицата $x\mathbf{E}-\mathbf{A}$ с матрицата

$$\mathbf{K}=\begin{Vmatrix} A_{11} & A_{21}\ldots A_{n1} \\ A_{12} & A_{22}\ldots A_{n2} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ A_{1n} & A_{2n}\ldots A_{nn} \end{Vmatrix}.$$

Получаваме лесно

$$(x\mathbf{E}-\mathbf{A})\,\mathbf{K}=\varphi(x)\,\mathbf{E}. \tag{11}$$

На основание на (10) се вижда, че $\mathbf{K}$ има формата

$$\mathbf{K}=\mathbf{L}_0+\mathbf{L}_1x+\ldots+\mathbf{L}_{n-1}x^{n-1},$$

където $\mathbf{L}_p$ са матрици от елементите $\alpha_{\mu\nu}^{(s)}$. Но тогава равенството (11) става

$$(x\mathbf{E}-\mathbf{A})(\mathbf{L}_0+\mathbf{L}_1x+\ldots+\mathbf{L}_{n-1}x^{n-1})=\mathbf{E}\varphi(x)=\mathbf{E}(\alpha_0+\alpha_1x+\ldots+\alpha_nx^n),\ \alpha_n=1.$$

С приравняване на коефициентите пред степените на x получаваме

$$-\mathbf{AL}_0=\alpha_0\mathbf{E}$$

$$-\mathbf{AL}_1+\mathbf{L}_0=\alpha_1\mathbf{E}$$

$$-\mathbf{AL}_2+\mathbf{L}_1=\alpha_2\mathbf{E}$$

$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

$$-\mathbf{AL}_{n-1}+\mathbf{L}_{n-2}=\alpha_{n-1}\mathbf{E}$$

$$\mathbf{L}_{n-1}=\mathbf{E}.$$

Като умножим последователно тези равенства отляво с $\mathbf{E}, \mathbf{A}, \mathbf{A}^2,\ldots,\mathbf{A}^n$ и ги съберем, получаваме

$$\alpha_0\mathbf{E}+\alpha_1\mathbf{A}+\alpha_2\mathbf{A}^2+\ldots+\alpha_n\mathbf{A}^n=\mathbf{0}.$$

Така установяваме следната теорема на Хамилтон-Кейли:

Ако $\varphi(x)$ *е характеристическият полином за матрицата* $\mathbf{A}$, *то* $\varphi(\mathbf{A})=\mathbf{0}$.

Глава VI

Квадратични и билинеарни форми

1. Билинеарни форми. Ако една цяла рационална функция е линейна и хомогенна спрямо две редици от променливи $x_1, x_2, \ldots x_n$ и $y_1, y_2, \ldots, y_n$, то тя се нарича б и л и н е а р н а.

Следователно тя има вида

$$f = \sum_{i,k}^{1\ldots n} a_{ik} x_i y_k.$$

Детерминантата

$$D = \begin{vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \cdot & \cdot \cdot \cdot \cdot \cdot \\ a_{n1} & a_{n2} \ldots a_{nn} \end{vmatrix}$$

се нарича д е т е р м и н а н т а или д и с к р и м и н а н т а на билинеарната форма.

Нека $\mathbf{r}' = \| x_1 \; x_2 \ldots x_n \|$ и $\mathbf{h}' = \| y_1 \, y_2 \ldots y_n \|$. Като вземем под внимание, че f може да се пише така:

$$f = u_1 x_1 + u_2 x_2 + \ldots + u_n x_n,$$

гдето

$$u_i = a_{i1} y_1 + a_{i2} y_2 + \ldots + a_{in} y_n, \; i = 1, \; 2, \; 3, \ldots, \; n,$$

то f може да се напише така:

$$f \equiv \mathbf{r}' \mathbf{A} \mathbf{h};$$

$$\mathbf{A} = \left\| \begin{vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ \cdot & \cdot \cdot \cdot \cdot \cdot \\ a_{n1} & a_{n2} \ldots a_{nn} \end{vmatrix} \right\|.$$

Лесно е да се убедим, че ако

$$\mathbf{AB} = \mathbf{C},$$

то

$$\mathbf{B}'\mathbf{A}' = \mathbf{C}'.$$

Действително в първото произведение умножаваме редове от **A** със стълбове от **B**, а във второто умножаваме редове от **B**′, т. е. стълбове от **B**, със стълбове от **A**′, т. е. с редове от **A**, и индексите са само така разменени.

Ако трансформираме f с линейните трансформации

$$\mathbf{r}=\mathbf{Su},$$

$$\mathbf{h}=\mathbf{Tb},$$

тогава получаваме

$$f=\mathbf{r'Ah}=\mathbf{u'S'ATb},$$

така че **S'AT** е матрицата на трансформираната форма.

Ако A, T, S са детерминантите от матриците **A**, **T**, **S** то $A.\ T.\ S$ ще бъде детерминантата на трансформираната форма.

2. **Квадратични форми.** Ако в билинеарната форма двете редици променливи са еднакви и ако поставим $a_{ik}=a_{ki}$, $i,\ k=1,\ 2,\ldots,\ n$, то формата се обръща в тъй наречената квадратична форма:

$$u=\sum_{i,\ k}^{1\ldots n} a_{ik},\ x_i x_k.$$

Ако поставим

$$u=\sum_{m=1}^{n} a_{i}^{m} x_m, \tag{4}$$

то

$$u=u_1x_1+u_2x_2+\ldots+u_n\,x_n,$$

т. е. с матрици

$$u=\mathbf{r'Ar}.$$

Детерминантата

$$A=\begin{vmatrix} a_{11} & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}\ldots a_{2n} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2}\ldots a_{nn} \end{vmatrix}$$

се нарича детерминанта или дискриминанта на квадратичната форма. Тази дискриминанта, понеже $a_{ik}=a_{ki}$, е симетрична, така че очевидно матриците **A** и **A'** са еднакви.

Ако извършим линейната трансформация

$$\mathbf{r}=\mathbf{Bh},$$

то ще получим

$$\mathbf{u}=\mathbf{r'\ Ar}=\mathbf{h'B'ABh}.$$

Следователно матрицата на трансформираната форма ще бъде

$$\mathbf{C}=\mathbf{B'AB},$$

която е също симетрична, понеже

$$\mathbf{C'}=(\mathbf{B'AB})'=\mathbf{B'A'B}=\mathbf{B'AB}.$$

От горното следва, че дискриминантата на трансформираната форма ще бъде равна на

$$BA^2.$$

Дискриминантата на трансформираната форма е равна на дискриминантата на дадената квадратична форма, умножена с квадрата на модула на субституцията. Дискриминантата е един инвариант.

3. **Реципрочна форма.** Като решим уравненията (4) спрямо неизвестните $x_1, x_2, \ldots, x_n$, получаваме, ако $A \neq 0$,

$$\mathbf{r} = \mathbf{A}^{-1}\mathbf{u},$$

гдето

$$\mathbf{A}^{-1} = \left\| \begin{matrix} A^{11} & A^{12} \ldots A^{1n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ A^{n1} & A^{n2} \ldots A^{nn} \end{matrix} \right\|,$$

$$A^{ik} = \frac{A_{ki}}{A} = \frac{A_{ik}}{A}.$$

Матрицата $\mathbf{A}^{-1}$ наричаме инверсна или реципрочна на матрицата **A**. Лесно се вижда, че $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{E}$. Следователно ще имаме

$$\mathbf{u} = \mathbf{r}'\mathbf{A}\mathbf{r} = \mathbf{u}' \ (\mathbf{A}^{-1})'\mathbf{A}\mathbf{A}^{-1}\mathbf{u} = \mathbf{u}'\mathbf{A}^{-1}\mathbf{u}.$$

Наричаме $\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}$ реципрочна форма на $\mathbf{r}'\mathbf{A}\mathbf{r}$.

4. **Ранг на квадратичната форма.** Ще докажем, че рангът на матриците на дадената квадратична форма и трансформираната посредством линейна субституция, на която модулът е отличен от нула, е един и същ. Действително нека рангът на матрицата **A** е равен на r, а на трансформираната матрица $\mathbf{C} = \mathbf{B}'\mathbf{A}\mathbf{B}$ е r'. Понеже B^2 може да се представи пак като детерминанта от ред n, то като приложим правилото за умножение на матрици (стр. 65), получаваме, че всички детерминанти от един произволен ред на C са линейни хомогенни функции от детерминантите от същия ред на A. Понеже A е от ранг r, то всички миньори от ред $r+1$ са равни на нула, отгдето следва, че и всички миньори от ред $r+1$ на C са също равни на нула. Така че $r' \leqq r$. Но с реципрочната субституция от трансформираната квадратична форма получаваме дадената форма. От това следва на същото основание, че $r \leqq r'$, отгдето $r' = r$.

5. **Трансформиране на квадратичната форма на сума от квадрати.** С помощта на една линейна субституция може всяка квадратична форма да се представи във вида

$$u = \sum \lambda_i z_i^2, \tag{5}$$

т. е. да представлява сума от квадрати с постоянни коефициенти. Ще докажем отначало, че такова трансформиране е действително възможно.

Нека отначало допуснем, че има поне един коефициент $a_{ii}(i=1, 2, \ldots, n)$, например a_{11}, който е отличен от нула. Тогава формата

$$u-a_{11}\left(x_1+\sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k\right)^2=u_1$$

не съдържа x_1, т. е. представлява квадратична форма на $n-1$ променливи. Ако направим трансформацията

$$z_1=x_1+\sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k,$$

$$z_\lambda=x_\lambda,\ \lambda=2, 3, \ldots, n,$$

на която детерминантата е равна на 1, то получаваме

$$u=a_{11}z_1^2+u_1.$$

Ако всички a_{ii} са равни на нула, ще има поне едно $a_{ik}(i\neq k)$, отлично от нула. Тогава да направим субституцията

$$x_i=y_i+y_k,$$

$$x_k=y_i-y_k,$$

$$x_\lambda=y_\lambda,\ \lambda\neq i,\ \lambda\neq k,$$

на която детерминантата, както веднага се проверява, е отлична от нула. Тогава се преминава в една квадратична форма на $y_1, y_2, \ldots, y_n$, като коефициентът на y_i^2 е равен на $2a_{ik}\neq 0$. Върху тази форма можем да приложим вече горното, т. е. да я сведем към сума от един квадрат и една квадратична форма с $n-1$ променливи. Продължавайки така, се вижда, че u ще се обърне на сума от квадрати. Ако с $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_p$ означим матриците на последователните линейни субституции, то за u ще имаме

$$u=\mathbf{z}'\mathbf{S}_p'\mathbf{S}'_{p-1}\ldots\mathbf{S}_1'\mathbf{A}\mathbf{S}_1\mathbf{S}_2\ldots\mathbf{S}_p\mathbf{z},$$

така че с линейната трансформация

$$\mathbf{r}=\mathbf{S}_1\mathbf{S}_2\ldots\mathbf{S}_p\mathbf{z}=\mathbf{S}\mathbf{z}$$

можем направо да получим формата (5), именно

$$u=\sum\lambda_i z_i^2.$$

Тъй като рангът на формата не се изменя при последователните линейни трансформации и понеже на (5) матрицата е

$$\left\|\begin{array}{cccc}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & \lambda_n\end{array}\right\|,$$

на която рангът е равен на броя на отличните от нула λ, то следва, че броят на квадратите, на които се трансформира дадена квадратична форма, е равен на ранга ѝ.

По същия начин се установява, че билинеарната форма

$$\mathbf{r'Ah}$$

може да се приведе във вида

$$\sum \lambda_i u_i v_i.$$

Действително чрез субституциите

$$\mathbf{r = Su,\ h = Sv}$$

формата става

$$\mathbf{r'Ah = u'S'ASv}$$

и матрицата **S'AS** на трансформираната форма съдържа само диагонални елементи, а другите са нули.

6. **Ортогонална трансформация.** В по-раншния параграф ние видяхме един ничин за трансформиране на една квадратична форма в сума от квадрати. Нека с помощта на субституцията

$$(6)\qquad \begin{cases} x_1 = b_{11} y_1 + \ldots + b_{1n} y_n, \\ \quad \ldots\ldots\ldots\ldots \\ x_n = b_{n1} y_1 + \ldots + b_{nn} y_n, \end{cases}$$

т. е. на $\mathbf{r = Bh}$, сме дали на u формата

$$(7)\qquad u = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y^2 n.$$

Тази редукция е възможна по безбройно много начини, понеже неизвестните b_{ik} са подчинени само на $\frac{1}{2}\, n(n-1)$ условия, защото в трансформираната форма трябва да получим за коефициентите ѝ $c_{ik} = 0$, при $i \neq k$, което именно дава $\frac{n(n+1)}{2}$ условия. Остават свободни за разполагане само

$$n^2 - \frac{n(n-1)}{2} = \frac{n(n-1)}{2}$$

параметъра b_{ik}. Но понеже в една ортогонална трансформация имаме толкова условия между елементите ѝ, то ясно е, че ще съществува една напълно определена такава трансформация, посредством която можем да дадем на u формата (7). Ако решим уравненията (6) спрямо неизвестните $x_1, x_2, \ldots, x_n$, то благодарение на ортогоналността, която предполагаме, ще получим

$$\begin{aligned} y_1 &= b_{11} x_1 + b_{21} x_2 + \ldots + b_{n1} x_n, \\ y_2 &= b_{12} x_1 + b_{22} x_2 + \ldots + b_{n2} x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n &= b_{1n} x_1 + b_{2n} x_2 + \ldots + b_{nn} x_n. \end{aligned}$$

Формата (7) става

$$u=\sum_{i=1}^{n}\lambda_i(b_{1i}x_1+b_{2i}x_2+\ldots+b_{ni}x_n)^2$$

и като идентифицираме с $u=\sum a_{ij}\,x_ix_j$, получаваме

$$a_{ij}=\sum_{r=1}^{n}\lambda_r b_{ir}\,b_{jr},$$

отгдето за $j=1,\ 2,\ 3,\ldots,\ n$ имаме

$$\begin{cases} a_{i1}=\lambda_1 b_{i1}b_{11}+\lambda_2 b_{i2}b_{12}+\ldots+\lambda_n b_{in}b_{1n},\\ a_{i2}=\lambda_1 b_{i1}b_{21}+\lambda_2 b_{i2}b_{22}+\ldots+\lambda_n b_{in}b_{2n},\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{in}=\lambda_1 b_{i1}b_{n1}+\lambda_2 b_{i2}b_{n2}+\ldots+\lambda_n b_{in}b_{nn}. \end{cases}$$

Като ги решим спрямо $\lambda_1 b_{i1},\ \lambda_2 b_{i2},\ldots,\ \lambda_n b_{in}$, получаваме

$$B\lambda_j b_{ij}=a_{i1}B_{1j}+a_{i2}B_{2j}+\ldots+a_{in}B_{nj}$$

или по свойство на ортогоналните детерминанти

$$B_{ij}=Bb_{ij}$$

получаваме

$$\lambda_j b_j=a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{in}b_{nj}.$$

Ако поставим $i=1,\ 2,\ 3,\ldots,\ n$, то получаваме

$$(8)\qquad \begin{cases} (a_{11}-\lambda_j)b_{1j}+ a_{12}b_{2j}+\ldots+ a_{1n}b_{nj}=0,\\ a_{21}b_1+(a_{22}-\lambda_j)b_{2j}+\ldots+ a_{2n}b_{nj}=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{n1}b_{1j}+ a_{n2}b_{2j}+\ldots+(a_{nn}-\lambda_j)b_{nj}=0. \end{cases}$$

За да има стойности на b, които да ги удовлетворяват, трябва детерминантата да бъде равна на нула. Следователно $\lambda_1,\ \lambda_2,\ldots,\lambda_n$ са корени на уравнението

$$\begin{vmatrix} a_{11}-x & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22}-x & \ldots & a_{2n}\\ \ldots & \ldots & ; & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn}-x \end{vmatrix}=0.$$

Впрочем до този резултат можем да достигнем директно, като използуваме предишния параграф. Именно по условие формата

$$u-\lambda(x_1^2+x_2^2+\ldots+x_n^2)$$

с ортогонална субституция трябва да премине във формата $(\lambda_1-\lambda)y_1^2+$ $+(\lambda^2-\lambda)y_2^2+\ldots+(\lambda_n-\lambda)y_n^2$, и то така, че при $\lambda=\lambda_1, \lambda_2,\ldots, \lambda_n$ трябва да се редуцира на по-малко от n квадрата, т. е. трябва дискриминантата ѝ да е равна на нула за тези стойности на λ.

Лесно е да намерим елементите на трансформацията, като решим системата (8), считайки λ за известни. Тогава известно ни е, че неизвестните са пропорционални на адюнгираните количества на елементите на един ред в детерминантата от коефициентите:

$$\begin{vmatrix} a_{11}-\lambda_j & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}-\lambda_j\ldots a_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2}\ldots a_{nn}-\lambda_j \end{vmatrix}.$$

Като вземем под внимание, че тази детерминанта е симетрична, от което споменатите адюнгирани количества са пропорционални на квадратните корени на адюнгираните количества на главните елементи, то като означим с A_{ij} адюнгираното количество на $a_{ii}-\lambda_j$ в горната детерминанта, получаваме за всяко j

$$\frac{b_{1j}}{\sqrt{A_{1j}}}=\frac{b_{2j}}{\sqrt{A_{2j}}}=\ldots=\frac{b_{nj}}{\sqrt{A_{nj}}}=\frac{1}{\sqrt{A_{1j}+A_{2j}+\ldots+A_{nj}}},$$

понеже $b_{1j}^2+b_{2j}^2+\ldots+b_{nj}^2=1$. Оттук имаме

$$b_{ij}=\sqrt{\frac{A_{ij}}{A_{1j}+A_{2j}+\ldots+A_{nj}}},\ i,\ j=1,\ 2,\ldots,\ n.$$

7. Закон за инерцията на квадратичните форми. Както видяхме, една квадратична форма с подходяща трансформация с отлична от нула детерминанта може да се представи като сума от квадрати, броят на които е точно равен на ранга на квадратичната форма. Възможните такива представяния притежават едно важно свойство, което има голямо приложение. Именно във всичките такива трансформации на една квадратична форма в сума от квадрати броят на квадратите с положителни коефициенти е един и същ. Същото очевидно важи тогава и за броя на отрицателните квадрати.

Това е така нареченият закон за инерчността на квадратичните форми, открит от Силвестер и Хермит.

Така нека формата $u=\sum_{i,k}^{1\ldots n} a_{ik}x_ix_k$ с помощта на линейните субституции

$$\text{(9)}\qquad \begin{aligned} x_i&=b_{i1}y_1+\ldots+b_{in}y_n \\ x_i&=c_{i1}z_1+\ldots+c_{in}z_n \end{aligned} \qquad (i=1,\ 2,\ldots,\ n)$$

с отлични от нула детерминанти се привежда във формите

$$u=\lambda_1 y_1^2+\lambda_2 y_2^2+\ldots+\lambda_n y_n^2,$$

(10)

$$u=\mu_1 z_1^2+\mu_2 z_2^2+\ldots+\mu_n z_n^2.$$

Ясно е, че от (9) можем да изразим $y_1, y_2, \ldots, y_n$ посредством $z_1, z_2, \ldots, z_n$ и обратно. От (10) имаме

(11) $$\lambda_1 y_1^2+\ldots+\lambda_n y_n^2=\mu_1 z_1^2+\ldots+\mu_n z_n^2.$$

Да допуснем, че в дясната част има h отрицателни члена, а в лявата $h_1>h$ такива. Тогава можем в дясната част тези h члена да анулираме, като дадем стойности нула на съответните z. Също можем вляво да анулираме положителните членове на броя $n-h_1$. Остават още $n-h+h_1>n$ неизвестни y и z, които можем да подберем така, че да удовлетворяват уравненията (9). Но тогава вдясно в (11) остават само положителни членове, а вляво — само отрицателни, т. е. това уравнение не е възможно. Следователно допускането, че в двете части на (11) няма еднакъв брой положителни членове, е невярно.

Разликата от броя на положителните и отрицателните квадрати по абсолютна стойност се нарича с и г н а т у р а на квадратичната форма.

Ако всички квадрати са с еднакъв знак и броят им е равен на броя n на променливите, формата се нарича д е ф и н и т н а. Ако всички квадрати са положителни и следователно формата е положителна и взема стойност нула само тогава, когато всички променливи се анулират, то тя се нарича п о л о ж и т е л н о д е ф и н и т н а. Също ако всичките квадрати са отрицателни и формата се анулира само при анулирането на променливите, тя се нарича о т р и ц а т е л н о д е ф и н и т н а.

8. Дефинитни форми. Ще използуваме трансформацията в параграф 6. Нека предположим, че във всяка отделна линейна трансформация всички коефициенти a_{ii} не са нули, така че не ще става нужда за втората трансформация. При всяка трансформация тогава ще се явяват квадрати на променливите, които нека така бъдат подбрани, че индексите им да бъдат в натурален ред. Отделните трансформации ще имат вида

$$\begin{aligned}
x'_1&=x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n,\\
x'_2&=\quad x_2\\
x'_3&=\quad\quad x_3\\
&\cdots\cdots\cdots\\
x'_n&=\quad\quad\quad\quad x_n,\\
x''_1&=x'_1\\
x''_2&=x'_2+\alpha_{23}x'_3+\ldots+\alpha_{2n}x'_n\\
x''_3&=\quad x'_3\\
&\cdots\cdots\cdots\\
x''_n&=\quad\quad\quad\quad x'_n\\
&\cdots\cdots\cdots
\end{aligned}$$

Тогава, ако означим последователно трансформациите с

$$\mathbf{r}' = \mathbf{S}_1\mathbf{r},\ \mathbf{r}'' = \mathbf{S}_2\mathbf{r}', \ldots,\ \mathbf{u} = \mathbf{S}_n\mathbf{r}^{(n-1)},$$

то ще имаме

$$\mathbf{u} = \mathbf{S}\mathbf{r} = \mathbf{S}_n\mathbf{S}_{n-1}\ldots\mathbf{S}_1\mathbf{r},$$

гдето **S** ще бъде, както лесно се убеждаваме, матрица с елементи на главния диагонал 1 и с елементи от едната страна на главния диагонал, равни на нула. Матрицата на трансформираната форма ще бъде $\mathbf{S}'\mathbf{A}\mathbf{S} = \mathbf{B}$.

От това се вижда, че със субституцията **S**

$$\begin{aligned} y_1 &= x_1 + h_{12}x_2 + \ldots + h_{1n}x_n, \\ y_2 &= x_2 + h_{23}x_3 + \ldots + h_{2n}x_n, \\ &\cdots\cdots\cdots\cdots \\ y_n &= x_n \end{aligned}$$

формата $u = \sum_{i,k=1}^{n} a_{ik}x_ix_k$ се обръща във вида $u = \sum_{i=1}^{n} \lambda_i y_i^2$.

Ако поставим

$$A_m = \begin{vmatrix} a_{11} & a_{12} \ldots a_{1m} \\ a_{21} & a_{22} \ldots a_{2m} \\ \cdots & \cdots\cdots \\ a_{m1} & a_{m2} \ldots a_{mm} \end{vmatrix},\ m = 1,\ 2, \ldots,\ n,$$

то ясно е, че при $x_{m+1} = x_{m+2} = \ldots = x_n = 0$ формата $\sum_{i,k=1}^{m} a_{ik}x_ix_k$ с горната субституция се обръща в $\sum_{i=1}^{m} \lambda_i y_i^2$ така, че за дискриминантата на новата форма имаме

$$\lambda_1\lambda_2 \ldots \lambda_m = A_m.$$

Следователно ще имаме

$$\lambda_1 = A_1,\ \lambda_2 = \frac{A_2}{A_1}, \ldots,\ \lambda_n = \frac{A_n}{A_{n-1}}$$

при предположение, че A са отлични от нула. Че детерминантите A_m не се променят, се вижда лесно и на основание на специалната форма на матрицата **S**. От допускането ни следва действително, че тези числа са отлични от нула. Така първо $A_1 = a_{11} \neq 0$. Като премахнем при първата трансформация x_1^2, детерминантата A_2 остава непроменена и коефициентът пред новия квадратен член е равен на $\frac{A_2}{A_1}$, който по пред-

положение е отличен от нула, и т. н. Очевидно е тогава, че ако $A_k > 0$, $k=1, 2, \ldots, n$, то формата ще е положително дефинитна. Ако обаче $A_1 < 0$, $A_2 > 0$, $A_3 < 0, \ldots, (-1)^n A_n > 0$, то понеже числата λ са отрицателни, формата ще бъде отрицателно дефинитна.

Тези достатъчни условия са и необходими. Действително всичките a_{ii} трябва да бъдат отлични от нула. Понеже, ако например $a_{kk}=0$, то като поставим $x_k=1$, а останалите променливи $x_s=0$, формата взема стойност нула и не би била положително или отрицателно дефинитна.

С трансформацията

$$y_1 = x_1 + \frac{a_{12}}{a_{11}} x_2 + \ldots + \frac{a_{1n}}{a_{11}} x_n,$$
$$y_2 = \qquad x_2,$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$y_n = \qquad\qquad x_n$$

се явява квадратът $\lambda_1 y_1^2$. Със същата трансформация всички главни миньори на детерминантата A, в които влизат елементи от първия ред и стълб, т. е. всички миньори, на които матрицата е симетрична спрямо главния диагонал

$$a_{11}$$
$$a_{22}$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$a_{nn},$$

остават неизменени.

Понеже в първия ред и първия стълб остава само $a_{11} \neq 0$, то онези от главните миньори, които са от втори ред, разделени с a_{11}, дават коефициентите на квадратните членове, т. е. онези, които съдържат квадратите на променливите, което впрочем се и директно показва лесно с пресмятане. Ако формата трябва да бъде положително или отрицателно дефинитна, то никой от тях не може да бъде равен на нула. Защото тогава, като поставим $x_1=0$ и върху останалата форма приложим горните разсъждения, ще можем да дадем стойности на променливите, от които една поне не е нула, но формата да има стойност нула. Така последователно получаваме теоремата: **Необходимо и достатъчно условие, щото квадратичната форма $u=\sum_{i,k=1}^{n} a_{ik}x_i x_k$ да бъде положително дефинитна, е главните миньори $A_1, A_2, \ldots, A_n$ да бъдат положителни.**

За да бъде формата отрицателно дефинитна, необходимо и достатъчно е редицата

$$1, A_1, A_2, \ldots, A_n$$

да има n вариации на знаците, т. е. $(-1)^i A_i > 0$. Като следваме горния път, получаваме:

Необходимо и достатъчно условие една форма от ранг μ да не може да вземе отрицателни стойности се състои в това, че променливите да могат така да се номерират, че главните миньори $A_1, A_2, \ldots, A_\mu$ да бъдат положителни.

Необходимо и достатъчно условие една форма от ранг μ да не може да вземе положителни стойности се състои в това, че променливите да могат така да се номерират, че редицата $1, A_1, \ldots, A_\mu$ да има променливи знаци.

Че условията са достатъчни, се вижда, както това направихме по-рано. Необходимостта им се доказва така: ако $\mu=0$, то се анулират всички a_{ik} и тогава твърдението на теоремите е очевидно. Ако $\mu>0$, има поне едно число a_{ik}, отлично от нула. Ако всички a_{ii} са нули, но например $a_{12}\neq 0$, то поставяме

$$x_1=y_1+y_2,\ x_2=y_1-y_2,\ x_i=y_i\ \ (i>2).$$

Тогава в трансформираната форма се явяват двата члена $2a_{12}y_1^2$ и $-2a_{12}y_2^2$. Ако поставим $y_i=0$ за $i>2$, то се вижда, че формата мени знака си. Следователно поне едно $a_{ii}\neq 0$. Номерираме така променливите, че $a_{11}\neq 0$, и с първата трансформация се получава квадратът на новите променливи — именно членът $a_{11}y_1^2$.

Ако $\mu=1$, то новата форма се анулира идентично. Ако $\mu>1$, то трябва да има поне един главен миньор от втори ред, отличен от нула, понеже такива миньори са коефициентите на квадратните членове в новата форма и ако бяха равни всички на нула, последната, както вече видяхме, би променяла знака си. Продължавайки така, достигаме до доказване на теоремата.

9. **Хермитови форми.** Билинеарната форма

$$A(x, y)=\sum_{i,\,k}^{1\ldots n} a_k x_i y_k=\mathbf{r'Ah}$$

и квадратичната форма

$$A(x, \bar{x})=\sum_{i,\,k}^{1\ldots n} a_{ik} x_i \bar{x}_k=\mathbf{r'A\bar{r}}$$

се наричат хермитови, ако $a_{ik}=\bar{a}_{ki}$. Тогава за матрицата ѝ

$$A=\left\|\begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{array}\right\|$$

ще имаме $\mathbf{A'}=\bar{\mathbf{A}}$, гдето дясната част е матрицата, която се получава от $\mathbf{A}$, като вземем конюгованите стойности на елементите ѝ. Числата a_{ii} са реални, понеже $a_{ii}=\bar{a}_{ii}$. Ако приложим линейна трансформация

$$\mathbf{r}=\mathbf{Su},\quad \mathbf{h}=\bar{\mathbf{S}}\mathbf{v},$$

то формите се обръщат в

$$A'(u,\ v)=\mathbf{u'S'A\bar{S}v},$$

$$A'(u,\bar{u})=\mathbf{u'S'\,A\bar{S}u},$$

които са пак хермитови, понеже

$$\mathbf{(S'A\bar{S})'=\bar{S}'A'S=\bar{S}'\bar{A}S=\overline{(S'A\bar{S})}.}$$

Теорията на квадратичните форми може с малко изменение да се приложи напълно за хермитовите форми.

Така матрицата **S** може да се избере по такъв начин, че трансформираните форми да имат вида

$$\sum b_k u_k v_k,\qquad \sum b_k u_k \bar{u}_k.$$

Също рангът при линейната трансформация не се изменя и броят на различните от нула b_k е точно равен на ранга на формата.

Законът за инерчността на квадратичните форми остава в сила и за хермитовите форми. За пример ще посочим доказателството на това приложение, което гласи:

Ако една хермитова форма на n променливи

$$\sum_{p=1}^{n}\sum_{q=1}^{n} a_{pq}x_p\bar{x}_q,\ a_{pq}=\bar{a}_{qp}$$

с трансформацията

$$x_i=\sum_{\varrho=1}^{n}\gamma_{i\varrho}\,x'_\varrho \tag{12}$$

приема вида $\sum_{\varrho=1}^{n} b_\varrho x'_\varrho\bar{x}'_\varrho$ и ако с една друга такава трансформация приема вида $\sum_{\varrho=1}^{n} c_\varrho x''_\varrho\bar{x}''_\varrho$, то между числата b има толкова положителни, отрицателни или равни на нула, колкото има съответно между числата c.

Броят на числата b и c, които не са нули, е равен на ранга на формата. Нека втората трансформация е

$$\sum_{\varrho=1}^{n}\delta_{i\varrho}\,x''_\varrho\,. \tag{13}$$

Като решим уравненията (12) и (13) спрямо x' и x'', да имаме

$$x'_\varrho=\sum_{\lambda=1}^{n}\gamma'_{\varrho\lambda}\,x_\lambda,\ x''_\varrho=\sum_{\lambda=1}^{n}\delta'_{\varrho\lambda}\,x_\lambda.$$

Следователно имаме тъждеството

$$\sum_{\lambda=1}^{n} b_\varrho \,|\gamma'_{\varrho 1}\, x_1 + \ldots + \gamma'_{\varrho n}\, x_n|^2 =$$

(14)

$$= \sum_{\varrho=1}^{n} c_\varrho \,|\delta'_{\varrho 1}\, x_1 + \ldots + \delta'_{\varrho n}\, x_n|^2.$$

Нека от числата b_ϱ, m_1 са положителни и l_1 отрицателни, а между числата c_ϱ, m_2 са положителни и l_2 отрицателни. Тогава сумата $m_1+l_1=m_2+l_2$ е равна на ранга. Ако имаме $m_1>m_2$, т. е. $l_1<l_2$, то $m_2+l_1<m_1+l_1\leqq n$. Тогава можем да дадем на x_λ стойности, които не са всички равни на нула, такива, че квадратите на брой l_1+m_2, които принадлежат на отрицателни b и положителни c, да се анулират. Но в такъв случай вдясно ще имаме само отрицателни квадрати, а вляво положителни освен ако всички квадрати са нули. Но тогава трябва и x_λ да са равни на нула.

Дефинициите и свойствата на дефинитните форми се обобщават непосредствено и за хермитовите форми.

ЧАСТ III

ЛИНЕЙНА АЛГЕБРА

Глава I

Пространство от n-мерни вектори

1. Определение на вектор и линейна зависимост. Под n-мерен вектор разбираме една наредена система от n числа $(a_1, a_2, \ldots, a_n)$. Векторите ще означаваме с буквите **a**, **b**, **c**, ... Числата $a_1, a_2, \ldots, a_n$ се наричат координати на вектора $(a_1, a_2, \ldots, a_n)$. Два вектора $\mathbf{a}=(a_1, a_2, \ldots, a_n)$ и $\mathbf{b}=(b_1, b_2, \ldots, b_n)$ се считат равни само тогава, когато координатите им са еднакви, т. е. когато

$$a_1=b_1,\ a_2=b_2, \ldots, a_n=b_n.$$

Равенството им означаваме с $\mathbf{a}=\mathbf{b}$. Под нулев вектор **0** разбираме вектора $(0, 0, \ldots, 0)$.

Например двумерният вектор (a, b) определя положението на една точка в равнината, на която декартовите координати са a и b. На матрицата от m реда и n стълба

$$\left\|\begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{array}\right\|$$

отговарят m-те n-мерни вектора

$$(a_{i1}, a_{i2}, \ldots, a_{in}),\ i=1, 2, 3, \ldots, m,$$

които се определят напълно, или n-те m-мерни вектора

$$(a_{1k}, a_{2k}, \ldots, a_{mk}),\ k=1, 2, 3, \ldots, n.$$

Под сумата $\mathbf{a}+\mathbf{b}=\mathbf{c}$ на два произволни вектора

$$\mathbf{a}=(a_1, a_2, \ldots, a_n),\ \mathbf{b}=(b_1, b_2, \ldots, b_n)$$

разбираме вектора

$$\mathbf{c}=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n).$$

Очевидно ще имаме

$$\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}$$

за кой да е вектор **a**. Лесно се убеждаваме, че събирането на вектори притежава основните свойства на числата.

1. Разместително (комутативно) свойство:

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}.$$

2. Съчетателно (асоциативно) свойство:

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

Под произведение $\mathbf{a}\lambda=\lambda\mathbf{a}$ на вектора $\mathbf{a}=(a_1, a_2, \ldots, a_n)$ с произволно число λ разбираме вектора

$$\lambda\mathbf{a}=(\lambda a_1, \lambda a_2, \ldots, \lambda a_n).$$

Умножението на число притежава разместителното и разпределителното свойство. Имаме

$$c\mathbf{a}=\mathbf{a}c,$$

$$c(\mathbf{a}+\mathbf{b})=c\mathbf{a}+c\mathbf{b},$$

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

Да докажем например второто равенство. Нека

$$\mathbf{a}=(a_1, a_2, \ldots, a_n)\quad \mathbf{b}=(b_1, b_2, \ldots, b_n).$$

По определенията ще имаме

$$c(\mathbf{a}+\mathbf{b})=(ca_1+cb_1, ca_2+cb_2, \ldots, ca_n+cb_n),$$

$$c\mathbf{a}+c\mathbf{b}=(a_1c+b_1c, a_2c+b_2c, \ldots, a_nc+b_nc),$$

от които равенства следва непосредствено твърдението. Третото равенство следва непосредствено от второто.

Ще дефинираме разлика $\mathbf{a}-\mathbf{b}$ на векторите **a** и **b**. Под разлика разбираме такъв вектор **x**, сумата на който с **b** е равна на **a**,

$$\mathbf{x}+\mathbf{b}=\mathbf{a}.$$

Ако означим с

$$\mathbf{x}=(x_1, x_2, \ldots, x_n),$$

то горното равенство става

$$(x_1+b_1, x_2+b_2, \ldots, x_n+b_n)=(a_1, a_2, \ldots, a_n),$$

откъдето получаваме

$$x_1=a_1-b_1, \quad x_2=a_2-b_2, \ldots, \quad x_n=a_n-b_n.$$

Следователно разликата се определя с

$$\mathbf{a}-\mathbf{b}=(a_1, a_2, \ldots, a_n)-(b_1, b_2, \ldots, b_n)=$$
$$=(a_1-b_1, a_2-b_2, \ldots, a_n-b_n).$$

Разликата $0-\mathbf{a}$ се означава с $-\mathbf{a}$ и се нарича вектор, противоположен на $\mathbf{a}$. Следователно

$$\mathbf{a}+(-\mathbf{a})=0.$$

От определенията следват свойствата:

$$c\,(d\mathbf{a})=(cd)\,\mathbf{a},$$

$$c\,(\mathbf{a}-\mathbf{b})=c\mathbf{a}-c\mathbf{b},\ (c-d)\,\mathbf{a}=c\mathbf{a}-d\mathbf{a},$$

$$1\,.\,\mathbf{a}=\mathbf{a},$$

$$(-1)\,\mathbf{a}=-\mathbf{a},$$

$$0\,.\,\mathbf{a}=\mathbf{0},\ \alpha\,.\,0=\mathbf{0}.$$

Нека $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ е една система от n-мерни вектори. Казваме, че тази система от вектори са линейно зависими, ако съществуват m числа $p_1, p_2, \ldots, p_m$, поне едното от които е отлично от нула, така че да имаме

$$p_1\,\mathbf{a}_1+p_2\,\mathbf{a}_2+\ldots+p_m\,\mathbf{a}_m=\mathbf{0}.$$

Ако такова равенство е невъзможно за такива числа $p_1, p_2, \ldots p_m$, т. е. горното равенство е само тогава възможно, ако всичките числа $p_1, p_2, \ldots, p_m$ са равни на нула, то казваме, че системата вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ са линейно независими.

Например векторите

$$\mathbf{a}=(2,\ 1,\ 3),\quad \mathbf{b}=(1,\ 0,\ -5),\quad \mathbf{c}=(-4,\ -1,\ 7)$$

са линейно зависими, понеже между тях има линейната връзка

$$\mathbf{a}+2\mathbf{b}+\mathbf{c}=\mathbf{0}.$$

Лесно се проверява, че горните вектори са два по два линейно независими. Например да предположим, че $\mathbf{a}$ и $\mathbf{c}$ са линейно зависими, т. е.

$$p\mathbf{a}+q\mathbf{c}=\mathbf{0}.$$

Това равенство е възможно, ако

$$2p-4q=0,$$

$$p-q=0,$$

$$3p+7q=0.$$

От тези уравнения получаваме обаче $p=0$, $q=0$.

Системата от нулевия вектор $\mathbf{0}$ е линейно зависима, понеже за всяко число c имаме $c\,.\,\mathbf{0}=\mathbf{0}$. Ако $\mathbf{a}$ е вектор, отличен от $\mathbf{0}$, то равенството $c\,.\,\mathbf{a}=\mathbf{0}$ е само тогава възможно, ако $c=0$, т. е. системата от всеки отличен от $\mathbf{0}$ вектор е линейно независима. Системата от кои да е вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ и нулевия вектор $\mathbf{0}$ е линейно зависима, понеже имаме

$$1\,.\,\mathbf{0}+0\,.\,\alpha_1+0\,.\,\mathbf{a}_2+\ldots+0\,.\,\mathbf{a}_n=0.$$

Ако системата вектори $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_m$ е линейно независима, то всяка част от нея е също линейно независима система вектори.

Да предположим например, че системата вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$, $p<m$, е линейно зависима. Тогава съществуват p числа $c_1, c_2, \ldots, c_p$, едно поне от които е отлично от нула, за които имаме

$$c_1\mathbf{a}_1+c_2\mathbf{a}_2+\ldots+c_p\mathbf{a}_p=0.$$

Но тогава ще имаме

$$c_1\mathbf{a}_1+c_2\mathbf{a}_2+\ldots+c_p\mathbf{a}_p+0\cdot\mathbf{a}_{p+1}+0\cdot\mathbf{a}_{p+2}+\ldots+0\cdot\mathbf{a}_m=0,$$

от което равенство следва предложението.

Ако един n-мерен вектор $\mathbf{b}$ се изразява посредством n-мерните вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ с равенството

$$\mathbf{b}=p_1\mathbf{a}_1+p_2\mathbf{a}_2+\ldots+p_m\mathbf{a}_m,$$

където $p_1, p_2, \ldots, p_m$ са числа, то казваме, че векторът $\mathbf{b}$ е линейна комбинация на векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ или че $\mathbf{b}$ се изразява линейно посредством векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. Лесно се вижда, че линейна зависимост и линейно изразяване са две равносилни понятия. Действително, ако векторът $\mathbf{a}$ е линейна комбинация на векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$, то

$$\mathbf{a}=c_1\mathbf{a}_1+c_2\mathbf{a}_2+\ldots+c_m\mathbf{a}_m,$$

където $c_1, c_2, \ldots, c_m$ са числа. Но последното равенство може да се пише така:

$$c_1\mathbf{a}_1+c_2\mathbf{a}_2+\ldots+c_m\mathbf{a}_m+(-1)\mathbf{a}=0,$$

което показва линейната зависимост на векторите $\mathbf{a}, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$.

Обратно, ако векторите $\mathbf{a}, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ са линейно зависими, то ще имаме

$$c\mathbf{a}+c_1\mathbf{a}_1+c_2\mathbf{a}_2+\ldots+c_m\mathbf{a}_m=0,$$

където поне едно от числата $c, c_1, c_2, \ldots, c_m$ е отлично от нула. Нека $c\neq 0$. Тогава от горното равенство имаме

$$\mathbf{a}=-\frac{c_1}{c}\mathbf{a}_1-\frac{c_2}{c}\mathbf{a}_2-\ldots-\frac{c_m}{c}\mathbf{a}_m,$$

т. е. $\mathbf{a}$ е линейна комбинация на $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$.

2. Ранг на система от вектори. Под ранг на една система от n-мерни вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ разбираме най-голямото число линейно независими вектори от тях.

Например рангът на системата

$$\mathbf{a}=(1, 2, -1),\ \mathbf{b}=(2, -1, 3),\ \mathbf{c}=(3, 1, 2),\ \mathbf{d}=(0, 5, -5)$$

е равен на 2, понеже векторите $\mathbf{a}$ и $\mathbf{b}$ са линейно независими и

$$\mathbf{c}=\mathbf{a}+\mathbf{b},\quad \mathbf{d}=2\mathbf{a}-\mathbf{b}.$$

От дефиницията на ранг непосредствено следва предложението: Ако рангът на системата от n-мерни вектори

(1) $$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$$

е равен на p и $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_p}$ са p линейно независими вектори от нея, то всеки от векторите (1) се изразява линейно посредством векторите $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_p}$.

Действително нека $\mathbf{a}$ е произволен вектор, отличен от векторите $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_p}$. Понеже рангът е равен на p, то ще имаме равенството

(2) $$\lambda\mathbf{a}+\lambda_1\mathbf{a}_{i_1}+\lambda_2\mathbf{a}_{i_2}+\ldots+\lambda_p\mathbf{a}_{i_p}=0,$$

където едно поне от числата $\lambda, \lambda_1, \lambda_2, \ldots, \lambda_p$ е отлично на нула. Но числото λ е сигурно отлично от нула, понеже инак бихме имали

$$\lambda_1\mathbf{a}_{i_1}+\lambda_2\mathbf{a}_{i_2}+\ldots+\lambda_p\mathbf{a}_{i_p}=0,$$

което противоречи на линейната независимост на системата $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots \mathbf{a}_{i_p}$. Но тогава от (2) получаваме

$$\mathbf{a}=-\frac{\lambda_1}{\lambda}\mathbf{a}_{i_1}-\frac{\lambda_2}{\lambda}\mathbf{a}_{i_2}-\ldots-\frac{\lambda_p}{\lambda}\mathbf{a}_{i_p},$$

с което предложението е установено.

1. Ако в една система от n-мерни вектори

(3) $$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$$

отстраним вектор, който се изразява линейно посредством останалите вектори, то рангът на системата не се изменя. Също рангът на системата не се изменя при прибавяне към нея на вектор, който се изразява линейно посредством векторите на системата.

Нека векторът $\mathbf{a}_m$ се изразява линейно посредством останалите вектори на системата (3)

(4) $$\mathbf{a}_m=\mu_1\mathbf{a}_1+\mu_2\mathbf{a}_2+\ldots+\mu_{m-1}\mathbf{a}_{m-1}$$

и нека r е рангът ѝ. Ако $r=0$, то векторите (3) са нулеви и при отстраняване на нулевия вектор $\mathbf{a}_m$ очевидно рангът е пак равен на нула. Нека сега $r>0$. Да предположим, че при отстраняване на $\mathbf{a}_m$ рангът на системата (3) намалява. Тогава рангът би трябвало да стане равен на $r-1$, т. е. системата

(5) $$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_{m-1}$$

трябва да има ранг $r-1$. Ако $r=1$, то от (4) следва, че и $\mathbf{a}_m=\mathbf{0}$, т. е. рангът на системата (3) е равен на 0, което е противоречие.

Нека сега $r>1$ и да означим с

(6) $$\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_r}$$

r линейно независими вектори от системата (3). Не е трудно да се види, че векторът $\mathbf{a}_m$ трябва да се съдържа в редицата (6). Действително, ако $\mathbf{a}_m$ не е между векторите (6), то последните ще принадлежат на системата (5) и рангът ѝ би бил равен на $r-1$, което е противоречие. Нека тогава $\mathbf{a}_m=\mathbf{a}_{k_r}$. Векторите $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_{r-1}}$ са линейно независими като част от системата (6). Но понеже рангът на системата (5) е равен на r, то по предното предложение всеки вектор от системата (5) се изразява линейно посредством векторите $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \mathbf{a}_{k_3}, \ldots, \mathbf{a}_{k_{r-1}}$. От равенството (4) се вижда тогава, че векторът $\mathbf{a}_m=\mathbf{a}_{k_r}$ се изразява линейно посредством векторите $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_{r-1}}$, което противоречи на линейната независимост на системата вектори (6).

Втората част на теоремата следва непосредствено от първата. Нека предположим, че при прибавяне на вектор $\mathbf{a}_{m+1}$ към системата (3) рангът се повишава. Но тогава при премахване на този елемент рангът би трябвало да се понижи, което противоречи на първата част.

2. Ако рангът на една система от n-мерни вектори

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \tag{7}$$

е равен на r, то има в системата (7) r линейно независими вектори, чрез които се изразяват линейно всичките вектори на системата (7). Обратно, ако в системата (7) съществуват r линейно независими вектора, посредством които векторите на системата се изразяват линейно, то рангът на системата е равен на r.

В случай, че рангът е равен на нула, то векторите на системата (7) са всички нули и теоремата е очевидна. Ако рангът $r \geqq 1$, то в системата трябва да има r линейно независими вектори $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{kr}$ и по предложението всеки вектор $\mathbf{a}$ от системата ще се изразява линейно посредством векторите $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_r}$.

Обратно, нека векторите $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_r}$ са линейно независими и всеки вектор от системата (7) се изразява линейно посредством $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_r}$. Тогава, като премахнем векторите $\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{jm-r}$, които се изразяват линейно посредством $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_r}$, получаваме линейно независимите вектори

$$\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_r},$$

като рангът на тази система е очевидно равен на r. По теорема I рангът на първоначалната система (7) е също равен на r.

На основание на теорема 2 можем да дадем друго определение на ранг. Ранг на една система от вектори се нарича броят на такива линейно независими вектори от системата, посредством които всеки вектор от системата може да се изрази линейно.

Нека е дадена една система от m вектора

$$\mathbf{a}_p=(a_{p1}, a_{p2}, \ldots, a_{pn}),\ 1 \leqq p \leqq m. \tag{8}$$

Да означим с A матрицата

$$\left\| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right\|,$$

на която редовете представляват координатите на векторите $\mathbf{a}_p$. Ще установим следната основна теорема:

Рангът на системата вектори (8) е равен на ранга на матрицата **A**. Да означим с **A′** транспонираната матрица на **A**;

$$\left\| \begin{array}{cccc} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{array} \right\|.$$

Рангът на **A′** е равен очевидно на ранга на **A**. Нека r е рангът на **A**. Ще има следователно поне една детерминанта от ред r, която е отлична от нула, и всичките детерминанти от ред $r+1$, образувани от елементите на матрицата **A′**, ще са равни на нула. С подходящо разместване на редовете и стълбовете можем да предположим, че детерминантата

$$\theta = \left| \begin{array}{cccc} a_{11} & a_{21} & \dots & a_{r1} \\ a_{12} & a_{22} & \dots & a_{r2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{1r} & a_{2r} & \dots & a_{rr} \end{array} \right|$$

е отлична от нула. Но тогава векторите $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ ще бъдат линейно независими. Действително, ако имаме равенството

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \dots + \lambda_r\mathbf{a}_r = 0,$$

то би следвало, че

$$\begin{array}{l} \lambda_1 a_{11} + \lambda_2 a_{21} + \dots + \lambda_r a_{r1} = 0, \\ \lambda_1 a_{12} + \lambda_2 a_{22} + \dots + \lambda_r a_{r2} = 0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \lambda_1 a_{1n} + \lambda_2 a_{2n} + \dots + \lambda_r a_{rn} = 0. \end{array} \tag{9}$$

Понеже рангът на матрицата, образувана от коефициентите на числата $\lambda_1, \lambda_2, \dots, \lambda_r$, е равен на r, то по теоремата на Руше единственото решение на системата линейни хомогенни уравнения (9) ще бъде нулевото, т. е. $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$, с което е установена линейната независимост на векторите $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$.

Да разгледаме системата линейни уравнения: $r+1 \leqq p \leqq n$,

$$
\begin{aligned}
&a_{11}x_1+a_{21}x_2+\ldots+a_{r1}x_r=a_{p1},\\
&a_{12}x_1+a_{22}x_2+\ldots+a_{r2}x_r=a_{p2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{1n}x_1+a_{2n}x_2+\ldots+a_{rn}x_r=a_{pn}.
\end{aligned}
\tag{10}
$$

Понеже матриците

$$
\begin{Vmatrix}
a_{11} & a_{21}\ldots a_{r1}\\
a_{12} & a_{22}\ldots a_{r2}\\
\cdots & \cdots\cdots\\
a_{1n} & a_{2n}\ldots a_{rn}
\end{Vmatrix},
\qquad
\begin{Vmatrix}
a_{p1} & a_{11} & a_{21}\ldots a_{r1}\\
a_{p2} & a_{12} & a_{22}\ldots a_{r2}\\
\cdots & \cdots & \cdots\cdots\\
a_{pn} & a_{1n} & a_{2n}\ldots a_{rn}
\end{Vmatrix}
$$

са от един и същ ранг, равен на r, то системата (10) има определено решение $x_1=\nu_1,\ x_2=\nu_2,\ldots,\ x_r=\nu_r$. Но тогава лесно намираме, че

$$\mathbf{a}_p=\nu_1\,\mathbf{a}_1+\nu_2\,\mathbf{a}_2+\ldots+\nu_r\,\mathbf{a}_r,$$

т. е. всеки вектор от системата (8) е линейна комбинация на векторите $\mathbf{a}_1,\ \mathbf{a}_2,\ldots,\ \mathbf{a}_r$. Следователно рангът на системата (8) е равен на r.

Нека сега рангът на системата (8) е равен на r. Тогава всеки вектор от нея ще бъде линейна комбинация на r независими вектори, които да означим с

$$\mathbf{a}_1,\ \mathbf{a}_2,\ldots,\ \mathbf{a}_r.$$

Можем следователно да пишем

$$a_q=\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_r\mathbf{a}_r.$$

Като приведем това равенство към координатите на векторите, които влизат в него, ще имаме

$$
\begin{aligned}
&a_{q1}=\lambda_1a_{11}+\lambda_2a_{21}+\ldots+\lambda_ra_{r1},\\
&a_{q2}=\lambda_1a_{12}+\lambda_2a_{22}+\ldots+\lambda_ra_{r2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{qn}=\lambda_1a_{1n}+\lambda_2a_{2n}+\ldots+\lambda_ra_{rn}.
\end{aligned}
$$

От елементарните свойства на детерминантите заключаваме тогава, че всички детерминанти от ред $r+1$, образувани от елементите на матрицата $\mathbf{A}'$, ще са равни на нула. Понеже рангът на системата вектори (8) е равен на r, то има r независими линейно вектори, които означихме с $\mathbf{a}_1,\ \mathbf{a}_2,\ldots,\ \mathbf{a}_r$. Но тогава детерминантата (θ) ще бъде отлична от нула, с което доказваме, че рангът на матрицата $\mathbf{A}'$ е равен на r.

Едновременно установихме, че рангът на системата вектори

$$\mathbf{b}_i=(a_{1i},\ a_{2i},\ldots,a_{mi}),\ i=1,\ 2,\ 3,\ldots,\ n,$$

на които координатите са стълбовете на матрицата $\mathbf{A}$, е равен на ранга на матрицата $\mathbf{A}$.

3. **Векторно пространство**. Множеството от всичките n-мерни вектори

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

разглеждано заедно с дефинираните действия събиране на вектори и умножение на вектори с реални числа, се нарича n-мерно векторно пространство, което ще отбелязваме с V_n. Да разгледаме векторите

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0),$$

$$\mathbf{e}_2 = (0, 1, 0, \ldots, 0),$$

$$\ldots\ldots\ldots$$

$$\mathbf{e}_n = (0, 0, 0, \ldots, 1).$$

От прости съображения или от предната теорема се вижда, че тези вектори са линейно независими. За всеки вектор $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ от V_n ще имаме

$$\mathbf{a} = (a_1, 0, 0, \ldots, 0) + (0, a_2, 0, \ldots, 0) + \ldots + (0, 0, 0, \ldots, a_n) =$$

$$= a_1(1, 0, 0, \ldots, 0) + a_2(0, 1, 0, \ldots, 0) + \ldots + a_n(0, 0, 0, \ldots, 1) =$$

$$= a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \ldots + a_n\mathbf{e}_n,$$

т. е. всеки вектор от пространството V_n е линейна комбинация на векторите $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Наричаме векторите $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ базис на пространството V_n. Изобщо всяка система от n-линейно независими вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ от пространството V_n наричаме базис на това пространство. Ако означим с $a_{p1}, a_{p2}, \ldots, a_{pn}$ координатите на векторите $\mathbf{a}_p$, $1 \leqq p \leqq n$, то от основната теорема следва, че необходимото и достатъчно условие, щото векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ да образуват базис на пространството V_n, се състои в това, че детерминантата

$$(11) \qquad \begin{vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots & \ldots\ldots \\ a_{n1} & a_{n2} \ldots a_{nn} \end{vmatrix}$$

да бъде отлична от нула. Нека

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$

е произволен вектор от пространството V_n. Системата линейни уравнения

$$a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n = a_1,$$

$$a_{12}x_1 + a_{22}x_2 + \ldots + a_{n2}x_n = a_2,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{nn}x_n = a_n$$

ще има едно напълно определено решение $x_1=\lambda_1, x_2=\lambda_2, \ldots, x_n=\lambda_n$. Лесно се вижда тогава, че векторът **a** ще е равен на

$$\mathbf{a}=\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_n\mathbf{a}_n. \tag{12}$$

Така виждаме, че всеки вектор от V_n е линейна комбинация на векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. Представянето (12) е еднозначно. Това видяхме, че следва от самото получаване на това представяне. Впрочем еднозначността на представянето (12) може да се установи и така: нека за вектора **a** имаме и друго представяне:

$$\mathbf{a}=\mu_1\mathbf{a}_1+\mu_2\mathbf{a}_2+\ldots+\mu_n\mathbf{a}_n. \tag{13}$$

Но тогава от (12) и (13) ще имаме

$$(\lambda_1-\mu_1)\mathbf{a}_1+(\lambda_2-\mu_2)\mathbf{a}_2+\ldots+(\lambda_n-\mu_n)\mathbf{a}_n=0,$$

което равенство противоречи на факта, че векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ са линейно независими.

Нека $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ са k вектори от пространството V_n. Множеството от всичките вектори

$$\tau_1\mathbf{b}_1+\tau_2\mathbf{b}_2+\ldots+\tau_k\mathbf{b}_k \tag{14}$$

за произволни реални числа $\tau_1, \tau_2, \ldots, \tau_k$ е също векторно пространство, понеже сумата на вектори от вида (14) е от същата форма, както и произведението им с произволно реално число. Полученото така пространство, на което очевидно векторите принадлежат и на V_n, се нарича п о д п р о с т р а н с т в о на V_n. Нулевият вектор и самото пространство V_n са очевидно подпространства на V_n. Те се наричат н е с о б с т в е н и подпространства на V_n. Останалите подпространства се наричат собствени подпространства на V_n.

4. Линейно преобразуване. Ще разгледаме сега въпроса за тъй нареченото линейно преобразуване на едно векторно пространство, което има приложение в разни въпроси от различни области на математиката. Нека V_n е едно n-мерно векторно пространство и да направим да отговаря на всеки вектор **a** от него също вектор **a′** от същото пространство. Векторът **a′** се нарича образ на вектора **a** и това съответствие се означава с $\mathbf{a}\to\mathbf{a}'$. При това съответствие не се предполага непременно взаимно еднозначно, т. е. на вектора **a** отговаря напълно определен вектор **a′**, но образът **a′** може да отговаря на различни вектори от V_n. Съответствието $\mathbf{a}\to\mathbf{a}'$ се нарича линейно преобразуване на пространството V_n, ако удовлетворява условията:

1. Ако $\mathbf{a}\to\mathbf{a}'$, то за всяко реално число λ, $\lambda\mathbf{a}\to\lambda\mathbf{a}'$.

2. Ако $\mathbf{a}\to\mathbf{a}'$ и $\mathbf{b}\to\mathbf{b}'$, то $\mathbf{a}+\mathbf{b}\to\mathbf{a}'+\mathbf{b}'$.

Например, ако на вектора (a, b) отговаря векторът $(a, 0)$, то получаваме така едно линейно преобразуване на пространството V_n, съставено от всичките двумерни вектори. Проверката на условията 1 и 2 е непосредствена. Така от $(a, b)\to(a, 0)$ следва

$$c(a, b)=(ca, cb)\to(ca, 0)=c(a, 0)$$

и от $(a,\ b) \to (a,\ 0)$ и $(a_1,\ b_1) \to (a_1,\ 0)$ следва

$$(a,\ b)+(a_1,\ b_1)=(a+a_1,\ b+b_1) \to (a+a_1,\ 0)=(a,\ 0)+(a_1,\ 0).$$

Очевидно при това преобразуване на всеки вектор $(a,\ b)$ отговаря само един вектор $(a,\ b)$, но векторът $(a, 0)$ отговаря на безбройно много вектори.

Ако на всеки вектор $\mathbf{a}$ от n-мерното векторно пространство V_n поставим в съответствие същия вектор $\mathbf{a}$, то получаваме очевидно линейно преобразувание, което се нарича тъждествено и се означава с 1^*. В случай, че на всеки вектор $\mathbf{a}$ отговаря нулевият вектор $\mathbf{0}$, полученото линейно преобразувание се нарича нулево и се означава с $\mathbf{0}^*$.

От свойството 1 следва следното: ако $\mathbf{a} \to \mathbf{a}'$, то $-\mathbf{a} \to -\mathbf{a}'$. Понеже $\mathbf{a}+(-\mathbf{a})=0$, то следва, че на нулата при всяко линейно преобразувание трябва да отговаря също нулата.

Множеството M от образите на векторите от пространството V_n при дадено линейно преобразувание е някое подпространство на V_n.

Действително, ако векторът $\mathbf{a}'$ от V_n е образ на вектора $\mathbf{a}$, то векторът $\lambda\mathbf{a}'$, λ произволно число, е образ на вектора $\lambda\mathbf{a}$ и следователно принадлежи на M. Ако $\mathbf{b}'$ е произволен вектор от M, то векторът $\mathbf{a}'+\mathbf{b}'$ като образ на вектора $\mathbf{a}+\mathbf{b}$ принадлежи също на M. Съгласно с теоремата от предния параграф M е подпространство на V_n.

Да означим образа $\mathbf{a}'$ на вектора $\mathbf{a}$ с $A(\mathbf{a})$. Тогава условията 1 и 2 за линейно преобразувание добиват прегледната форма:

1. $A(c\mathbf{a})=cA(\mathbf{a})$ за произволен вектор $\mathbf{a}$ от V_n и произволно число c.

2. $A(\mathbf{a}+\mathbf{b})=A(\mathbf{a})+A(\mathbf{b})$ за произволни вектори $\mathbf{a}$ и $\mathbf{b}$ от V_n.

Символът $A(\mathbf{a})$, който удовлетворява предните две условия, се нарича и линеен оператор. Той е еквивалентен на линейно преобразувание в смисъл, че на вектора $\mathbf{a}$ отговаря векторът $\mathbf{a}'=A(\mathbf{a})$.

Нека $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ е някой базис на пространството V_n. Ако $\lambda_1, \lambda_2, \ldots, \lambda_n$ са произволни числа, то на основание 1 и 2 получаваме лесно равенството

$$A(\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_n\mathbf{a}_n)=\lambda_1A(\mathbf{a}_1)+\lambda_2A(\mathbf{a}_2)+\ldots+\lambda_nA(\mathbf{a}_n).$$

Векторите $A(\mathbf{a}_i)$ принадлежат на пространството V_n и ще бъдат линейни комбинации на векторите $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$,

$$A(\mathbf{a}_i)=a_{1i}\mathbf{a}_1+a_{2i}\mathbf{a}_2+\ldots+a_{ni}\mathbf{a}_n,\ i=1,\ 2,\ 3,\ldots,\ n.$$

Матрицата $\mathbf{A}$

$$\mathbf{A}=\begin{Vmatrix} a_{11} & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}\ldots a_{2n} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2}\ldots a_{nn} \end{Vmatrix},$$

на която стълбовете са равни на координатите на векторите $A(\mathbf{a}_i)$, се нарича матрица на линейното преобразувание A при базис $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

Така на всяко линейно преобразуване отговаря една напълно определена матрица. Обратно, на всяка матрица **A** отговаря едно напълно определено линейно преобразуване. Последното свойство следва от следната теорема:

Нека $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ е един базис на пространството V_n и $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ е една система от n произволни вектори от V_n. Тогава съществува едно и само едно линейно преобразуване, което привежда векторите от базиса $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ във векторите $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

За да установим теоремата, нека на всеки вектор

$$\mathbf{a}=\lambda_1\ \mathbf{u}_1+\lambda_2\ \mathbf{u}_2+\ldots+\lambda_n\,\mathbf{u}_n$$

от V_n да направим да съответствува векторът

$$\mathbf{a}'=\lambda_1\,\mathbf{v}_1+\lambda_2\,\mathbf{v}_2+\ldots+\lambda_n\,\mathbf{v}_n=A(\mathbf{a}).$$

За да установим, че така получаваме линейно преобразуване, трябва да проверим дали условията 1 и 2 са изпълнени. Понеже за всяко число λ имаме

$$\lambda\,\mathbf{a}=\lambda\lambda_1\,\mathbf{u}_1+\lambda\lambda_2\,\mathbf{u}_2+\ldots+\lambda\lambda_n\,\mathbf{u}_n$$

и на вектора $\lambda\,\mathbf{a}$ трябва да съответствува векторът

$$\lambda\lambda_1\,\mathbf{v}_1+\lambda\lambda_2\,\mathbf{v}_2+\ldots+\lambda\lambda_n\,\mathbf{v}_n=\lambda\,\mathbf{a}',$$

то виждаме, че условието 1 за линейното преобразуване е изпълнено.

Нека

$$\mathbf{b}=\mu_1\mathbf{u}_1+\mu_2\mathbf{u}_2+\ldots+\mu_n\mathbf{u}_n$$

е също вектор от пространството V_n. На него ще отговаря векторът

$$\mathbf{b}'=\mu_1\mathbf{v}_1+\mu_2\mathbf{v}_2+\ldots+\mu_n\mathbf{v}_n.$$

На сумата от двата вектора

$$\mathbf{a}+\mathbf{b}=(\lambda_1+\mu_1)\,\mathbf{u}_1+(\lambda_2+\mu_2)\,\mathbf{u}_2+\ldots+(\lambda_n+\mu_n)\,\mathbf{u}_n$$

ще отговаря векторът

$$(\lambda_1+\mu_1)\,\mathbf{v}_1+(\lambda_2+\mu_2)\mathbf{v}_2+\ldots+(\lambda_n+\mu_n)\mathbf{v}_n=\mathbf{a}'+\mathbf{b}',$$

т. е. сумата на векторите $\mathbf{a}'$ и $\mathbf{b}'$. Така установихме, че и условието 2 за линейност на преобразуването е изпълнено.

Ако означим с A полученото линейно преобразуване на пространството V_n, ще имаме

$$A(\mathbf{a})=A(\lambda_1\,\mathbf{u}_1+\lambda_2\,\mathbf{u}_2+\ldots+\lambda_n\,\mathbf{u}_n)=\lambda_1\mathbf{v}_1+\lambda_2\,\mathbf{v}_2+\ldots+\lambda_n\,\mathbf{v}_n.$$

Ако поставим тук $\lambda_n=1$ и $\lambda_s=0$ за $s=1, 2, \ldots, k-1, k+1, \ldots, n$, получаваме $A(\mathbf{u}_k)=\mathbf{v}_k$, $1\leqq k\leqq n$, т. е. всеки базисен вектор $\mathbf{u}_k$ при преобразуването преминава в съответния вектор $\mathbf{v}_k$. Лесно се вижда, че няма друго линейно преобразуване със същото свойство. Действително да допуснем, че B е линейно преобразуване, което трансфор-

мира векторите $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ съответно във векторите $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Тогава за кой да е вектор a ще имаме

$$B\,\mathbf{a}=\lambda_1 B\,\mathbf{u}_1+\lambda_2 B\,\mathbf{u}_2+\ldots+\lambda_n B\,\mathbf{u}_n=\lambda_1\mathbf{v}_1+\lambda_2\mathbf{v}_2+\ldots+\lambda_n\mathbf{v}_n=\mathbf{a}',$$

понеже $B\,\mathbf{u}_k=\mathbf{v}_k$, $1\leqq k\leqq n$. Горното равенство показва, че $B(\mathbf{a})=A(\mathbf{a})$ за всеки вектор $\mathbf{a}$ от пространството V_n, т. е. линейното преобразуване B съвпада с преобразуването A.

Видяхме, че на всяко линейно преобразуване съответствува една матрица. Ще покажем сега, че съответствието с матрицата е и взаимно еднозначно при предположение, че базисът е фиксиран. Именно нека

$$\begin{vmatrix} a_{11} & a_{12}\ldots a_{1n} \\ a_{21} & a_{22}\ldots a_{2n} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2}\ldots a_{nn} \end{vmatrix} \tag{15}$$

е някоя матрица от ред n. Да образуваме следната система от вектори:

$$\mathbf{v}_p=a_{1p}\,\mathbf{u}_1+a_{2p}\,\mathbf{u}_2+\ldots+a_{np}\,\mathbf{u}_n,\ p=1,\ 2,\ldots,\ n.$$

От доказаната теорема следва, че може да намерим едно линейно преобразуване A в пространството V_n, което привежда векторите $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ във векторите $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, т. е. $A\,\mathbf{u}_1=\mathbf{v}_1$, $A\,\mathbf{u}_2=\mathbf{v}_2, \ldots, A\,\mathbf{u}_n=\mathbf{v}_n$, и това преобразуване е единствено.

Да разгледаме сега въпроса за преобразуването на координатите на векторите при прилагане на едно линейно преобразуване A. Нека при базис $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ координатите на вектора $\mathbf{x}$ са $x_1, x_2, \ldots, x_n$, а тези на $A\,\mathbf{x}$ са $y_1, y_2, \ldots, y_n$, т. е.

$$\mathbf{x}=x_1\,\mathbf{u}_1+x_2\,\mathbf{u}_2+\ldots+x_n\,\mathbf{u}_n,$$
$$A\,\mathbf{x}=y_1\,\mathbf{u}_1+y_2\,\mathbf{u}_2+\ldots+y_n\,\mathbf{u}_n.$$

За $A\,\mathbf{x}$ получаваме

$$A\,\mathbf{x}=A(x_1\,\mathbf{u}_1+x_2\,\mathbf{u}_2+\ldots+x_n\,\mathbf{u}_n)=x_1A(\mathbf{u}_1)+x_2A(\mathbf{u}_2)+\ldots+x_nA(\mathbf{u}_n).$$

Като поставим тук стойностите на $A\mathbf{u}_k$ от $A\mathbf{u}_k=\sum_{k=1}^{n} a_{ik}\,\mathbf{u}_i$, получаваме

$$A\,\mathbf{x}=(a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n)\,\mathbf{u}_1+\ldots+(a_{n1}x_1+$$
$$+a_{n2}x_2+\ldots+a_{nn}x_n)\,\mathbf{u}_n.$$

Следователно ще имаме

$$y_p=a_{p1}x_1+a_{p2}x_2+\ldots+a_{pn}x_n,\ p=1,\ 2,\ 3,\ldots,\ n.$$

Матрицата на тези n линейни форми съвпада с матрицата (15).

По-нататъшни свойства на линейните преобразувания ще бъдат разгледани в отдела за общи линейни пространства.

5. Евклидово пространство. В разгледаните досега линейни пространства не са отразени метричните свойства на фигурите в геомет-

рията, като дължина, ъгъл и т. н. В смисъл на обобщение на тези свойства ще разгледаме сега понятието за тъй наречените евклидови пространства. Едно векторно пространство V_n ще наричаме евклидово, ако на всеки два вектора **a** и **b** от него направим да съответствува едно реално число, означено с (**a**, **b**), което дефинираме по следния начин: ако

$$\mathbf{a}=(a_1,\ a_2,\ \ldots,\ a_n),$$

$$\mathbf{b}=(b_1,\ b_2,\ \ldots,\ b_n),$$

то полагаме

$$(\mathbf{a},\ \mathbf{b})=a_1\,b_1+a_2\,b_2+\ \ldots\ +a_n\,b_n.$$

Числото (**a**, **b**) се нарича скаларно произведение на векторите **a** и **b**. Под дължина на вектора **a** разбираме числото $\sqrt{(\mathbf{a},\mathbf{a})}$. Означаваме дължината с $|\mathbf{a}|$. Следователно имаме за вектора **a**

$$|\mathbf{a}|=+\sqrt{(\mathbf{a},\ \mathbf{a})}=\sqrt{a_1^2+a_2^2+\ \ldots\ +a_n^2}.$$

Очевидно ще имаме $|\mathbf{a}|=0$ само тогава, когато векторът **a** е равен на нулевия вектор.

За скаларното произведение имаме следното неравенство на Коши—Буняковски:

$$(\mathbf{a},\ \mathbf{b})^2\leqq|\mathbf{a}|\,.\,|\mathbf{b}|\,. \tag{16}$$

Нека **a** и **b** са векторите

$$\mathbf{a}=(a_1,\ a_2,\ \ldots,\ a_n),$$

$$\mathbf{b}=(b_1,\ b_2,\ \ldots,\ b_n).$$

Тогава неравенството (16) следва непосредствено от тъждеството

$$|\mathbf{a}|^2\,|\mathbf{b}|^2-(\mathbf{a},\ \mathbf{b})^2=(a_1\,b_2-a_2\,b_1)^2+(a_1\,b_3-a_2\,b_3)^2+\ \ldots$$

$$+\ \ldots\ +(a_2b_3-c_3b_2)^2+\ \ldots\ +(a_{n-1}\,b_n-a_n\,b_{n-1})^2.$$

При това от предното тъждество се вижда, че равенство в (16) можем да имаме само тогава, когато $\mathbf{a}=\lambda\,\mathbf{b}$, където λ е произволно число.

Под ъгъл φ между два вектора **a** и **b** разбираме ъгъла φ, определен с равенството

$$\cos\varphi=\frac{(\mathbf{a},\ \mathbf{b})}{|\mathbf{a}|\,|\mathbf{b}|}.$$

Дясната част на това равенство има смисъл при $|\mathbf{a}|\,|\mathbf{b}|>0$, т. е. векторите **a** и **b** не са нулеви. Тогава от неравенството на Коши — Буняковски следва, че тази част се намира в границите от -1 до 1 и ъгълът φ ще бъде еднозначно определен в интервала $0\leqq\varphi<\pi$.

Ако скаларното произведение на два вектора е равно на нула, то те се наричат ортогонални. Понеже за всеки вектор **a** имаме $(\mathbf{a},\ \mathbf{0})=0$, то нулевият вектор е ортогонален на всеки вектор.

Скаларното произведение притежава следните свойства:

1. $(\mathbf{a}, \mathbf{b})=(\mathbf{b}, \mathbf{a})$.
2. $(c\mathbf{a}, \mathbf{b})=c(\mathbf{a}, \mathbf{b})$ за всяко число c.
3. $(\mathbf{a}+\mathbf{b}, \mathbf{c})=(\mathbf{a}, \mathbf{c})+(\mathbf{b}, \mathbf{c})$ за всеки два вектора **a**, **b** и вектор **c**.
4. $(\mathbf{a}, \mathbf{a}) \geqq 0$ за всеки вектор **a** и равенство имаме само при $\mathbf{a}=\mathbf{0}$.

Проверката на тези свойства е лесна. Примерно ще установим третото свойство. Нека

$$\mathbf{a}=(a_1, a_2, \ldots, a_n),$$

$$\mathbf{b}=(b_1, b_2, \ldots, b_n),$$

$$\mathbf{c}=(c_1, c_2, \ldots, c_n).$$

Имаме тогава

$$(\mathbf{a}+\mathbf{b}, \mathbf{c})=(a_1+b_1)c_1+(a_2+b_2)c_2+\ldots+(a_n+b_n)c_n=$$

$$=(a_1c_1+a_2c_2+\ldots+a_nc_n)+(b_1c_1+b_2c_2+\ldots+b_nc_n)=(\mathbf{a}, \mathbf{c})+(\mathbf{b}, \mathbf{c}).$$

От горните свойства следва, че за кои да са вектори $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$, **b** и кои да са числа $\lambda_1, \lambda_2, \ldots, \lambda_m$ ще имаме

$$(\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m, \mathbf{b})=\lambda_1(\mathbf{a}_1, \mathbf{b})+\lambda_2(\mathbf{a}_2, \mathbf{b})+\ldots+\lambda_m(\mathbf{a}_m, \mathbf{b}).$$

Имаме следната теорема:

Ако няколко вектора, никой от които не е равен на нулевия вектор, са два по два ортогонални, то те са линейно независими.

Да допуснем, че между векторите на една такава система $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ има линейната зависимост

$$\mu_1\mathbf{a}_1+\mu_2\mathbf{a}_2+\ldots+\mu_p\mathbf{a}_p=0.$$

Като умножим скаларно това равенство с $\mathbf{a}_1$, получаваме

$$\mu_1(\mathbf{a}_1, \mathbf{a}_1)+\mu_2(\mathbf{a}_2, \mathbf{a}_1)+\ldots+\mu_p(\mathbf{a}_p, \mathbf{a}_1)=0,$$

$$\mu_1(\mathbf{a}_1, \mathbf{a}_1)=0.$$

Следователно $\mu_1=0$. С подобни умножения получаваме, че и $\mu_2=0, \ldots, \mu_p=0$, с което теоремата е установена.

От предната теорема следва, че всяка система от n ортогонални два по два вектора, от които никой не е нулев, образуват базис от n-мерното пространство V_n, на което те принадлежат. Всеки базис от този вид се нарича ортогонален базис. Всеки ненулев вектор с умножение на подходящо число може да се сведе към вектор с дължина 1. Именно вместо вектора **a** вземаме вектора $\frac{\mathbf{a}}{|\mathbf{a}|}$. Векторите с дължина единица се наричат единични вектори (нормирани вектори). Като заместим всеки вектор от ортогоналния базис със съответния му нормиран вектор, при което очевидно ортогоналността не се изменя, получаваме базис, в който всеки вектор е нормиран.

Такъв базис се нарича ортогонален нормиран базис на пространството V_n.

Например векторите

$$\mathbf{e}_1=(1,\ 0,\ 0,\ldots,0),$$
$$\mathbf{e}_2=(0,\ 1,\ 0,\ldots,\ 0),$$
$$\ldots\ldots\ldots\ldots$$
$$\mathbf{e}_n=(0,\ 0,\ 0,\ldots,1)$$

са нормирани (единични) и ортогонални два по два. Те образуват следователно ортогонален нормиран базис на пространството V_n.

Да означим с $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ един ортогонален нормиран базис на пространството V_n. Всеки вектор $\mathbf{a}$ от пространството V_n се представя, както знаем, във формата

$$\mathbf{a}=x_1\mathbf{u}_1+x_2\mathbf{u}_2+\ldots+x_n\mathbf{u}_n, \tag{17}$$

където $x_1, x_2, \ldots, x_n$ са координатите на вектора $\mathbf{a}$ спрямо разглеждания базис. Да умножим предното равенство с базисния вектор $\mathbf{u}_p$:

$$(\mathbf{a},\ \mathbf{u}_p)=x_1(\mathbf{u}_1,\ \mathbf{u}_p)+x_2(\mathbf{u}_2,\ \mathbf{u}_p)+\ldots+x_p(\mathbf{u}_p,\ \mathbf{u}_p)+\ldots.+x_n(\mathbf{u}_n,\ \mathbf{u}_p)=x_p.$$

Така получаваме простата формула $x_p=(\mathbf{a},\ \mathbf{u}_p)$. Да умножим сега скаларно равенството (17) с произволен вектор $\mathbf{b}$ от пространството V_n. Получаваме

$$(\mathbf{a},\ \mathbf{b})=x_1(\mathbf{u}_1,\ \mathbf{b})+x_2(\mathbf{u}_2,\ \mathbf{b})+\ldots+x_n(\mathbf{u}_n,\ \mathbf{b}). \tag{18}$$

Ако координатите на вектора $\mathbf{b}$ са $y_1, y_2, \ldots, y_n$, т. е.

$$\mathbf{b}=y_1\mathbf{u}_1+y_2\mathbf{u}_2+\ldots+y_n\mathbf{u}_n,$$

видяхме, че $y_p=(\mathbf{b},\ \mathbf{u}_p)=(\mathbf{u}_p,\ \mathbf{b})$. Като заместим тогава произведенията

$$(\mathbf{u}_1,\ \mathbf{b}),\ (\mathbf{u}_2,\ \mathbf{b}),\ldots,\ (\mathbf{u}_n,\ \mathbf{b})$$

в равенството (18) с равните им изрази $y_1, y_2, \ldots, y_n$, получаваме

$$(\mathbf{a},\ \mathbf{b})=x_1y_1+x_2y_2+\ldots+x_ny_n.$$

В частност (при $\mathbf{a}=\mathbf{b}$) тази формула се свежда на следната:

$$(\mathbf{a},\ \mathbf{a})=|\mathbf{a}|^2=x_1^2+x_2^2+\ldots+x_n^2.$$

Ще разгледаме още едно свойство, което отразява едно свойство от елементарната геометрия. Нека $\mathbf{x}$ и $\mathbf{y}$ са два произволни вектора от пространството V_n, които предполагаме ортогонални помежду си. На основание на свойствата 1, 2 и 3 на скаларното произведение получаваме

$$(\mathbf{x}+\mathbf{y},\ \mathbf{x}+\mathbf{y})=(\mathbf{x}+\mathbf{y},\ \mathbf{x})+(\mathbf{x}+\mathbf{y},\ \mathbf{y})=(\mathbf{x},\ \mathbf{x})+2(\mathbf{x},\ \mathbf{y})+(\mathbf{y},\ \mathbf{y})=$$
$$=(\mathbf{x},\ \mathbf{x})+(\mathbf{y},\ \mathbf{y}),$$

т. е.

$$|\mathbf{x}+\mathbf{y}|^2=|\mathbf{x}|^2+|\mathbf{y}|^2.$$

Това равенство се нарича формула на Питагор по аналогия на познатата теорема за правоъгълния триъгълник.

По-нататък ще се запознаем с по-общи понятия за линейни пространства и евклидови пространства, към които спадат разгледаните пространства. Освен това ще видим, че по същество изучените досега пространства не се отличават от въпросните по-общи пространства.

6. Безкрайно пространство. Да разгледаме сега подобни пространства от безкрайно измерение. Под вектора **a** от такова пространство разбираме една редица от безбройно много реални числа

$$\mathbf{a}=(a_1,\ a_2,\ a_3,\ldots),$$

които са подчинени на условието, че редът

$$\sum_{n=1}^{\infty} a_n^2 \tag{19}$$

е сходящ. Множеството от тези вектори се нарича пространство на Хилберт. Обикновено се бележи с H. Ще покажем, че това пространство притежава основните свойства на разгледаните пространства от крайно измерение. Така за вектора

$$\lambda\,\mathbf{a}=(\lambda a_1,\ \lambda a_2,\ \lambda a_3,\ldots),$$

където λ е произволно число, съответният ред (19) е сходящ и следователно векторът $\lambda\,\mathbf{a}$ принадлежи на пространството H. Ако

$$\mathbf{b}=(b_1,\ b_2,\ b_3,\ldots)$$

е също вектор от H, то не е трудно да се покаже, че

$$(a_1+b_1,\ a_2+b_2,\ a_3+b_3,\ldots)$$

е вектор от H, който наричаме сума на векторите **a** и **b** и го отбелязваме с $\mathbf{a}+\mathbf{b}$. За да установим това, трябва да покажем, че редът

$$\sum_{n=1}^{\infty}(a_n+b_n)^2 \tag{20}$$

е сходящ при предположение, че редовете $\sum_1^\infty a_n^2$ и $\sum_1^\infty b_n^2$ са сходящи.

Но за всеки две реални числа a и b имаме неравенството

$$(a+b)^2\leqq 2(a^2+b^2). \tag{21}$$

Действително за разликата от дясната и лявата част на това неравенство получаваме

$$2\,a^2+2\,b^2-(a+b)^2=a^2+b^2-2\,ab=(a-b)^2.$$

На основание на (21) имаме тогава

$$\sum_{n=1}^{\infty}(a_n+b_n)^2 \leqq 2\sum_{n=1}^{\infty} a_n^2+2\sum_{n=1}^{\infty} b_n^2,$$

откъдето следва сходимостта на реда (20).

Аналогично на по-рано ще въведем понятието за скаларно произведение на два вектора:

$$\mathbf{a}=(a_1,\ a_2,\ a_3,\ldots),$$

$$\mathbf{b}=(b_1,\ b_2,\ b_3,\ldots).$$

По неравенството на Коши—Буняковски за всяко n имаме

$$\left(\sum_{i=1}^{n}|a_ib_i|\right)^2 \leqq \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2$$

или

$$\left(\sum_{i=1}^{n}|a_ib_i|\right)^2 \leqq \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2. \tag{22}$$

Оттук следва, че парциалните суми на реда

$$\sum_{i=1}^{\infty}|a_ib_i|$$

са ограничени и следователно този ред е сходящ. Но и тогава редът

$$\sum_{i=1}^{\infty} a_ib_i$$

ще бъде също сходящ. Сумата на този ред наричаме скаларно произведение на векторите **a** и **b** и го бележим както преди с (**a**, **b**). Два вектора наричаме ортогонални, ако скаларното им произведение е равно на нула. Нулевият вектор

$$0=(0,\ 0,\ 0,\ldots)$$

е ортогонален на всеки вектор от пространството H. Векторите

$$(a_1,\ 0,\ a_3,\ 0,\ a_5,\ 0,\ldots),$$

$$(0,\ b_2,\ 0,\ b_4,\ 0,\ b_6,\ldots)$$

$\left(\text{редовете } \sum_{0}^{\infty} a_{2n+1}^2 \text{ и } \sum_{1}^{\infty} b_{2n}^2, \text{ предположени сходящи}\right)$ са ортогонални.

Единичните вектори

$$\mathbf{e}_1=(1,\ 0,\ \ 0,\ldots),$$
$$\mathbf{e}_2=(0,\ 1,\ \ 0,\ldots),$$
$$\mathbf{e}_3=(0,\ 0,\ 1,\ 0,\ldots),$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

са очевидно ортогонални два по два.

От определението на скаларно произведение лесно се вижда, че то притежава свойствата 1, 2, 3 и 4 от предния параграф. Също така неравенството на Коши—Буняковски остава в сила, както това се вижда непосредствено от неравенството (22), като оставим n да расте неограничено.

Нека $\mathbf{a}_1,\ \mathbf{a}_2,\ldots,\ \mathbf{a}_n$ са вектори, два по два ортогонални. Подобно на по-рано получаваме

$$(\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_n,\ \mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_n)=(\mathbf{a}_1,\ \mathbf{a}_1)+(\mathbf{a}_2,\ \mathbf{a}_2)+\ldots+(\mathbf{a}_n,\ \mathbf{a}_n)$$

или

$$|\mathbf{a}_1+\mathbf{a}_2+\ldots+\mathbf{a}_n|^2=|\mathbf{a}_1|^2+|\mathbf{a}_2|^2+\ldots+|\mathbf{a}_n|^2.$$

Значи квадратът на дължината на сума от ортогонални вектори е сума от квадратите на дължините им. Това свойство е обобщение на теоремата на Питагор. Нека векторите $\mathbf{a}_1,\ \mathbf{a}_2,\ldots,\ \mathbf{a}_n$ са и нормирани, т. е. дължините им са равни на 1. Под величина на проекцията на произволен вектор $\mathbf{a}$ върху оста $\mathbf{a}_i$ разбираме скаларното произведение $(\mathbf{a},\ \mathbf{a}_i)=\lambda$. Сумата $\mathbf{b}=\sum_{i=1}^{n}\lambda_i\mathbf{a}_i$ е вектор на H и да означим с $\mathbf{u}$ вектора, който е равен на разликата между вектора $\mathbf{a}$ и вектора $\mathbf{b}$, т. е.

$$\mathbf{a}=\sum_{i=1}^{n}\lambda_i\mathbf{a}_i+\mathbf{u}.$$

Като умножим това равенство скаларно с вектора $\mathbf{a}_k$, $k=1,\ 2,\ 3,\ldots,\ n$,

$$(\mathbf{a},\ \mathbf{a}_k)=\sum_{i=1}^{n}\lambda_i(\mathbf{a}_i,\ \mathbf{a}_k)+(\mathbf{u},\ \mathbf{a}_k).$$

Понеже векторът $\mathbf{a}_k$ е ортогонален на останалите вектори $\mathbf{a}_n$, $i=1,\ 2,\ldots,\ k-1,\ k+1,\ldots,\ n$ и дължината му е единица, то следва от предното равенство, че $(\mathbf{u},\ \mathbf{a}_k)=0,\ 1\leqq k\leqq n$, т. е. векторът $\mathbf{u}$ е ортогонален на всичките вектори $\mathbf{a}_i,\ 1\leqq i\leqq n$. По теоремата на Питагор ще имаме равенството

$$|\mathbf{a}|^2=\sum_{i=1}^{n}\lambda_i^2+|u|^2.$$

Оттук следва непосредствено неравенството

$$\sum_{i=1}^{n} \lambda_i^2 \leqq |\mathbf{a}|^2, \tag{23}$$

което се нарича неравенство на Бесел. Очевидно в (23) ще имаме само тогава равенство, когато векторът $\mathbf{u}$ е нулев.

Да разгледаме сега една редица от безбройно много ортогонални и нормирани вектори:

$$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots$$

Ако означим с $\lambda_i = (\mathbf{a}, \mathbf{a}_i)$, $i = 1, 2, 3, \ldots$, големините на проекциите на произволен вектор $\mathbf{a}$ по осите a_i, $i = 1, 2, 3, \ldots$, то по горното неравенство за всяко n ще имаме

$$\sum_{i=1}^{n} \lambda_i^2 \leqq |\mathbf{a}|^2.$$

Като поставим тук n да расте неограничено, получаваме

$$\sum_{i=1}^{\infty} \lambda_i^2 \leqq |\mathbf{a}|^2.$$

Аналогично на случая на пространства от крайно измерение от значение е да се реши въпросът, кога в (24) имаме знака равенство и за кои случаи всеки вектор от пространството H може да се представи в ред по ортогоналните вектори $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots$ Тези и подобни въпроси се разглеждат в анализа и не ще се спираме по-нататък на тях

Глава II

Линейно (афинно) пространство

1. **Определение.** В математиката често се срещат обекти от разнообразно естество, върху които се извършват действията събиране и умножение с реални числа. Такива са например самите реални числа, комплексните числа. Определените по-рано n-мерни вектори

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$

са също подобни обекти. Именно под сума на два вектора $\mathbf{a}$ и $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ разбираме вектора

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

и под произведение на вектора $\mathbf{a}$ с реалното число c разбираме вектора $c\mathbf{a} = (ca_1, ca_2, \ldots, ca_n)$.

Ще дадем сега едно общо понятие. Множеството R от елементи $x, y, z, \ldots$, които ще наричаме и вектори, се нарича линейно (афинно) пространство, ако:

1) на всеки два елемента x и y се съпоставя в съответствие елемент z от R, наречен сума на елементите x и y и който се означава със $z=x+y$;

2) за всеки елемент от $x \in R$ и всяко реално число λ отговаря елемент u от R, наречен произведение на елемента x с число λ и който се означава с $u=\lambda x$;

3) определенията за сума и произведение удовлетворяват следните условия:

I. а) $x+y=y+x$ за произволни x и y от R;

б) $(x+y)+z=x+(y+z)$ за произволни x, y, z от R;

в) съществува елемент 0 (нулев елемент — нулев вектор) такъв, че $x+0=x$ за произволен елемент $x \in R$;

г) за всеки елемент $x \in R$ съществува противоположен елемен, който означаваме с $y=-x$, такъв, че $x+y=0$.

II. д) 1. $x=x$ за всяко $x \in R$;

е) $\alpha(\beta x)=(\alpha\beta)x$ за произволно $x \in R$ и произволни реални числа α и β;

III. ж) $(\alpha+\beta)x=\alpha x+\beta x$ за произволно $x \in R$ и произволни реални числа α и β;

з) $\alpha(x+y)=\alpha x+\alpha y$ за произволни x и y от R и произволно реално число α.

В случая пространството се нарича реално линейно пространство. Ако в горните условия числата λ, α, β, ... са комплексни числа, то се нарича комплексно линейно пространство.

На основание на свойствата ние можем да разместваме събираемите в сумата на повече от два елемента и да извършваме умноженията с произволни реални числа.

Лесно се вижда, че постулираната нула е единствена. Именно нека допуснем, че има два нулеви елемента 0 и 0_1, т. е. за всеки елемент x от R имаме

$$x+0=x, \quad x+0_1=x.$$

Като поставим в първото равенство $x=0_1$ и във второто $x=0$, получаваме

$$0_1+0=0_1, \quad 0+0_1=0.$$

Понеже по а) $0_1+0=0+0_1$, то от горните равенства получаваме $0_1=0$. На всеки елемент съответствува само един единствен противоположен елемент. Действително нека на елемента x отговарят два противоположни елемента y_1 и y_2, т. е. имаме $x+y_1=0$, $x+y_2=0$. Но тогава по аксиомите б) и в) ще имаме

$$y_2+(x+y_1)=(y_2+x)+y_1=0+y_1=y_1,$$

$$y_2+(x+y_1)=y_2+0=y_2,$$

откъдето следва, че $y_1=y_2$. От аксиомите ж) и д) следва, че за всеки елемент x от R имаме

$$0 \cdot x=0,$$

където вдясно 0 означава нулевия вектор. Действително за всеки вектор x имаме $0.x+1.x=(0+1)\,x=1.x=x$. Ако y е противоположният елемент на x, то $x+y=0$. Като прибавим y към предното равенство, получаваме $0.x+1.x+y=x+y=0$. Понеже $0.x+x+y=0.x+0$, то от предните две равенства следва, че $0.x=0$.

От горното получаваме: противоположният елемент на елемента x е равен на $(-1)\,x$. Наистина имаме

$$x+(-1)\,x=(1-1)\,x=0.x=0.$$

Елементът $(-1)\,x$ означаваме с $-x$.

Под разлика на векторите x и y ще разбираме сумата на векторите x и $-y$ и я означаваме с $x-y$.

Ще разгледаме някои примери на линейни пространства.

Пространство V_3. Съвкупността на всички свободни вектори от пространството образува линейно пространство. Също така съвкупността на векторите в равнината образува линейно пространство, което да означим с V_2. Векторите по една права образуват линейно пространство V_1.

Пространство T_n. Съвкупността от векторите от n измерения

$$a=(a_1,\ a_2,\ldots,\ a_n)$$

образува линейно пространство.

Пространство $C(a,\ b)$. Елементите на това пространство са всички непрекъснати функции $f(x)$ в затворения интервал $(a,\ b)$. Събирането и умножението с едно число се дефинират както в анализа. Елементът нула е тъждествено равната на нула функция.

2. **Линейна зависимост и измеримост.** Казваме, че между n вектора $x_1,\ x_2,\ldots,\ x_n$ от линейното пространство R съществува линейна зависимост, ако имаме равенството

$$\lambda_1 x_1+\lambda_2 x_2+\ldots+\lambda_n x_n=0, \tag{1}$$

където от числата $\lambda_1,\ \lambda_2,\ldots,\ \lambda_n$ поне едно е отлично от нула. В противен случай, т. е. когато равенството (1) е само тогава възможно, ако числата $\lambda_1,\ \lambda_2,\ldots,\ \lambda_n$ са равни на нула, векторите $x_1,\ x_2,\ldots,\ x_n$ се наричат линейно независими.

Ако един вектор x се представя посредством векторите $x_1,\ x_2,\ldots,\ x_m$ като линейна функция на тях, т. е. ако

$$x=\mu_1 x_1+\mu_2 x_2+\ldots+\mu_m x_m,$$

където $\mu_1,\ \mu_2,\ldots,\ \mu_m$ са числа, то казваме, че векторът x е линейна комбинация на векторите $x_1,\ x_2,\ldots,\ x_m$. Ако един вектор x е линейна комбинация на векторите $x_1,\ x_2,\ldots,\ x_m$, то очевидно между векторите $x,\ x_1,\ x_2,\ldots,\ x_m$ съществува линейна зависимост и, обратно, ако между векторите $x,\ x_1,\ x_2,\ldots,\ x_m$ има линейна зависимост, то поне единият от тях е линейна комбинация на останалите. Понеже от

$$\nu_0 x+\nu_1 x_1+\nu_2 x_2+\ldots+\nu_m x_m=0,$$

където поне едно от числата $\nu_k \neq 0$, $k=0, 1, \ldots, m$, следва

$$x_k = -\frac{\nu}{\nu_k} x - \frac{\nu_1}{\nu_k} x_1 - \ldots - \frac{\nu_{k-1}}{\nu_k} x_{k-1} - \frac{\nu_{k+1}}{\nu_k} x_{k+1} - \ldots - \frac{\nu_n}{\nu_k} x_n.$$

Линейното пространство R се нарича, че е от n измерения или по-накъсо n-мерно, ако в него съществуват n линейно независими вектора и няма в него повече на брой линейно независими вектори.

Ако в пространството не съществуват само системи от краен брой линейно независими вектори, т. е. колкото и да е голямо числото n, винаги има в него n линейно независими вектори, то пространството се нарича от безкрайно измерение или безкрайномерно.

3. **Базис и координати.** Всяка система от n линейно независими вектора $e_1, e_2, \ldots, e_n$ от n-мерното пространство R се нарича базис (база) на R.

Всеки вектор x от n-мерното пространство R може да се представи като линейна комбинация на векторите от базиса, и то по единствен начин.

Действително, понеже пространството R е n-мерно, то между векторите $x, e_1, e_2, \ldots, e_n$ трябва да съществува линейна зависимост, т. е.

$$\lambda x + \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0.$$

Тук числото λ трябва да е отлично от нула, понеже в противен случай, т. е. ако $\lambda=0$, би следвало, че между векторите $e_1, e_2, \ldots, e_n$ има линейна зависимост. Но тогава от горното уравнение получаваме

$$(1) \qquad x = -\frac{\lambda_1}{\lambda} e_1 - \frac{\lambda_2}{\lambda} e_2 - \ldots - \frac{\lambda_n}{\lambda} e_n,$$

с което първата част на теоремата е установена.

Да предположим, че за вектора x имаме двете представяния:

$$x = \mu_1 e_1 + \mu_2 e_2 + \ldots + \mu_n e_n,$$

$$x = \nu_1 e_1 + \nu_2 e_2 + \ldots + \nu_n e_n.$$

От тях получаваме

$$(\mu_1 - \nu_1) e_1 + (\mu_2 - \nu_2) e_2 + \ldots + (\mu_n - \nu_n) e_n = 0.$$

Понеже векторите $e_1, e_2, \ldots, e_n$ са линейно независими, то следва, че

$$\mu_1 - \nu_1 = 0, \ \mu_2 - \nu_2 = 0, \ldots, \ \mu_n - \nu_n = 0,$$

т. е.

$$\mu_1 = \nu_1, \ \mu_2 = \nu_2, \ldots, \ \mu_n = \nu_n,$$

с което единствеността на представяне на x е доказана.

Така за всеки вектор x от R получаваме представянето

$$x = \mu_1 e_1 + \mu_2 e_2 + \ldots + \mu_n e_n.$$

Числата $\mu_1, \mu_2, \ldots, \mu_n$ се наричат координати на вектора x в базиса $e_1, e_2, \ldots, e_n$.

Оттук за всяко число λ имаме

$$\lambda x=\lambda\mu_1 e_1+\lambda\mu_2 e_2+\ldots+\lambda\mu_n e_n.$$

Ако y е вектор също от R, то имаме

$$y=\mu_1' e_1+\mu_2' e_2+\ldots+\mu_n' e_n.$$

За сумата $x+y$ получаваме тогава

$$x+y=(\mu_1+\mu_1') e_1+(\mu_2+\mu_2') e_2+\ldots+(\mu_n+\mu_n',\mathrm{I})\, e_n.$$

т. е. векторът $x+y$ има за координати числата $\mu_1+\mu_1', \mu_2+\mu_2', \ldots, \mu_n+\mu_n'$.

Така получаваме предложението: При събиране на векторите координатите им се събират и при умножение с произволно число координатите им се умножават със същото число.

Да разгледаме някои примери от линейни пространства. 1. Множеството T_n от n-мерните вектори

$$x=(\xi_1, \xi_2, \ldots, \xi_n)$$

е n-мерно пространство. Векторите от него

$$e_1=(1, 0, 0, \ldots, 0),$$
$$e_2=(0, 1, 0, \ldots, 0),$$
$$\ldots\ldots\ldots\ldots$$
$$e_n=(0, 0, 0, \ldots, 1)$$

са линейно независими и образуват един базис. Всеки вектор x от T_n представлява линейна комбинация на предните вектори. Действително, ако

$$x=(\xi_1, \xi_2, \ldots, \xi_n),$$

то имаме

$$x=(\xi_1, 0, 0, \ldots, 0)+(0, \xi_2, 0, \ldots, 0)+\ldots+(0, 0, 0, \ldots, 0, \xi_n)=$$
$$=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n.$$

Можем да намерим и други базиси на пространството T_n. Например векторите

$$e_1'=(1, 1, 1, \ldots, 1),$$
$$e_2'=(0, 1, 1, \ldots, 1),$$
$$e_3'=(0, 0, 1, \ldots, 1),$$
$$\ldots\ldots\ldots\ldots$$
$$e_n'=(0, 0, 0, \ldots, 0, 1)$$

са линейно независими (проверката на това твърдение е лесна). Да намерим координатите на вектора $x=(\xi_1, \xi_2, \ldots, \xi_n)$ в този нов базис. Ако означим въпросните координати с $\eta_1, \eta_2, \ldots, \eta_n$, то ще имаме

$$(\xi_1, \xi_2, \ldots, \xi_n)=\eta_1 (1, 1, 1, \ldots, 1)+\eta_2(0, 1, 1, \ldots, 1)+\eta_3(0, 0, 1, 1, \ldots, 1)+$$
$$+\ldots+\eta_n(0, 0, \ldots, 0, 1)=(\eta_1, \eta_1+\eta_2, \eta_1+\eta_2+\eta_3, \ldots, \eta_1+\eta_2+\ldots+\eta_n).$$

Следователно числата $\eta_1, \eta_2, \ldots, \eta_n$ ще се определят от уравненията

$$\eta_1 = \xi_1,$$

$$\eta_1 + \eta_2 = \xi_2,$$

$$\eta_1 + \eta_2 + \ldots + \eta_n = \xi_n,$$

от които получаваме

$$\eta_1 = \xi_1,\ \eta_2 = \xi_2 - \xi_1,\ \eta_3 = \xi_3 - \xi_2, \ldots,\ \eta_n = \xi_n - \xi_{n-1}.$$

2. Нека R е множеството, съставено от всички полиноми на t от степен, ненадвишаваща $n-1$, $n \geqq 1$. Очевидно това множество представлява векторно пространство. Всеки вектор x от него като полином от степен $\leqq n-1$ ще има формата

$$P(t) = a_0 t^{n-1} + a_1 t^{n-2} + \ldots + a_{n-1}.$$

Векторите $1, t, t^2, \ldots, t^{n-1}$ образуват един базис на това пространство и коефициентите $a_0, a_1, \ldots, a_n$ са координатите на вектора $P(x)$ в този базис.

3. Да означим с $C(a, b)$ множеството от всички непрекъснати функции в интервала $a \leqq t \leqq b$. Понеже сумата на две непрекъснати функции в този интервал е непрекъсната функция в него и умножението с произволно число на една непрекъсната функция не променя непрекъснатостта на функцията, то множеството $C(a, b)$ е линейно векторно пространство. Прости съображения ни показват, че това пространство е от безкрайно измерение.

4. **Изоморфизъм на векторни линейни пространства.** От предните разглеждания забелязваме, че представянето на всеки вектор x от едно n-мерно пространство е напълно аналогично с представянето на векторите от друго n-мерно пространство и специално от пространството T_n. Това ни навежда на мисълта, че по същество такива пространства не са различни (са изоморфни). На основание на това ще въведем следната дефиниция:

Две линейни пространства R и R' се наричат изоморфни, ако между елементите им можем да установим взаимно еднозначно съответствие, т. е. на всеки вектор x от R да отговаря един и само един вектор x' от R' и всеки елемент от R' да отговаря на един и само един вектор от R, като това съответствие се подчинява и на условията:

1. Ако на векторите x и y от R съответствуват векторите x' и y' от R', то на вектора $x+y$ да съответствува векторът $x'+y'$.

2. На вектора λx да съответствува векторът $\lambda x'$.

Взаимно еднозначното съответствие на векторите x и x' означаваме накъсо с $x \longleftrightarrow x'$. Имаме следната основна теорема:

Ако две линейни пространства са изоморфни, то те имат еднакво измерение и, обратно, ако две линейни

пространства са от еднакво измерение, то те са изоморфни.

⇒ Нека пространствата R и R' са изоморфни. На нулевия вектор 0 от R ще отговаря нулевият вектор $0'$ от R', понеже от съответствието $x \longleftrightarrow x'$ получаваме $0=x+(-1)x \longleftrightarrow x'+(-1)x'=0'$. Ако n е измерението на пространството R, то в него ще съществуват $n+1$ линейно зависими вектора $x_1, x_2, x_3, \ldots, x_n, x_{n+1}$, т. е. ще имаме равенството

$$\lambda_1 x_1+\lambda_2 x_2+\ldots+\lambda_{n+1} x_{n+1}=0.$$

Ако $x_1', x_2', \ldots, x_{n+1}'$ са съответствуващите вектори от пространството R', то ще имаме равенството

$$\lambda_1 x_1'+\lambda_2 x_2'+\ldots+\lambda_{n+1} x_{n+1}'=0, \tag{1}$$

понеже векторът $\sum_{i=1}^{n+1}\lambda_i x_i'$ съответствува на вектора $\sum_{i=1}^{n+1}\lambda_i x_i$. Равенството (1) показва, че измерението на пространството R' не надминава n. По същия начин, изхождайки от пространството R', намираме, че измерението на пространството R не надминава това на R' и следователно пространствата R и R' имат еднаква измеримост.

⇐ Нека сега пространствата R и R' са от еднаква измеримост n. Да означим с $e_1, e_2, \ldots, e_n$ и $e_1', e_2', \ldots, e_n'$ произволни базиси съответно на пространството R и на R'. На всеки вектор x от R

$$x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n$$

да съпоставим вектора x' от R'

$$x'=\xi_1 e_1'+\xi_2 e_2'+\ldots+\xi_n e_n',$$

на който координатите са равни на координатите на вектора x. Очевидно това съответствие е взаимно еднозначно. Но лесно се убеждаваме, че пространствата R и R' са и изоморфни. Действително, ако y е също вектор от R

$$y=\eta_1 e_1+\eta_2 e_2+\ldots+\eta_n e_n$$

и y' е съответствуващият вектор от R', то имаме

$$x+y=(\xi_1+\eta_1)e_1+(\xi_2+\eta_2)e_2+\ldots+(\xi_n+\eta_n)e_n,$$

$$x'+y'=(\xi_1+\eta_1)e_1'+(\xi_2+\eta_2)e_2'+\ldots+(\xi_n+\eta_n)e_n'.$$

Тези равенства изразяват свойствата: от $x \longleftrightarrow x'$ и $y \longleftrightarrow y'$ следва $x+y \longleftrightarrow x'+y'$ и $\lambda x \longleftrightarrow \lambda x'$. Теоремата е установена.

Така всяко n-мерно пространство е изоморфно на пространството T_n от n-мерните вектори $(\xi_1, \xi_2, \ldots, \xi_n)$.

5. Подпространства. Всяко множество R_1 от вектори, принадлежащи на линейното пространство R, което от своя страна е линейно пространство, се нарича подпространство на R.

Понеже векторите от R_1 принадлежат на пространството R, то лесно се вижда, че условията 1—4 за линейност на пространството R_1

се свеждат на следните: от $x \in R_1$ и $y \in R_1$ следва $x+y \in R_1$ и $\lambda x \in R_1$ за всяко число λ.

Например нулевото пространство е подпространство на всяко пространство R. Също самото пространство R е подпространство на себе си. Тези две подпространства ги наричаме несобствени подпространства. Другите подпространства се наричат собствени подпространства.

Нека V_3 е триизмеримото пространство. Множеството от векторите, лежащи в една равнина, образува подпространство на V_3, което е собствено.

Нека R_1 и R_2 са две подпространства на пространството R. С R' да означим множеството на всичките вектори $x+y$, където $x \in R_1$ и $y \in R_2$. Това множество R' е очевидно подпространство на R и се нарича сума на подпространствата R_1 и R_2. Множеството R'' от всичките общи вектори на подпространствата R_1 и R_2 е също подпространство на R, което се нарича сечение на R_1 и R_2.

В пространството T_n да разгледаме множеството K на всичките вектори

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

координатите на които да удовлетворяват хомогенните линейни уравнения

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=0. \end{aligned}$$

По-рано видяхме, че ако $x_1', x_2', \ldots, x_n'$ и $x_1'', x_2'', \ldots, x_n''$ са решения на тази система, то и $x_1'+x_1'', x_2'+x_2'', \ldots, x_n'+x_n''$ и $\lambda x_1', \lambda x_2', \ldots, \lambda x_n'$ са също решения на системата. Следователно множеството K е линейно пространство, което е подпространство на T_n.

Понеже всяко подпространство е и линейно пространство, то за него остават в сила всичките доказани свойства за линейните пространства. Размерността на всяко подпространство на едно пространство R не може да надмине размерността на R, понеже не може да има в подпространството повече линейно независими вектори, отколкото в R.

Един естествен начин за образуване на подпространства на дадено линейно пространство R се състои в следното: нека да вземем няколко вектора $x_1, x_2, \ldots, x_p$ от R и да образуваме всички възможни линейни комбинации на тези вектори. Полученото така множество от вектори R' образува подпространство на R. Действително сборът на два вектора от R' е също линейна комбинация на векторите $x_1, x_2, \ldots, x_p$ и следователно принадлежи на R'. Същото е вярно и за произведението на вектор от R' с произволно число. Подпространството R' се нарича линейна обвивка на векторите $x_1, x_2, \ldots, x_p$. Очевидно линейната обвивка е най-малкото подпространство, съдържащо векторите $x_1, x_2, \ldots, x_p$. Линейната обвивка се означава с $L(x_1, x_2, \ldots, x_p)$.

Ако векторите y_1, y_2,... принадлежат на линейната обвивка $L(x_1, x_2,...)$ на векторите x_1, x_2,..., x_p, то линейната обвивка $L(y_1, y_2,...)$ се съдържа в $L(x_1, x_2,...)$. Действително всяка линейна комбинация на векторите y_1, y_2,... ще принадлежи на $L(x_1, x_2,...)$, понеже представлява линейна комбинация от векторите x_1, x_2,... Оттук непосредствено следва следното свойство: ако някой от векторите x_1, x_2,..., x_p на $L(x_1, x_2,..., x_p)$ е линейна комбинация от останалите вектори, то той може да се изостави, без да измени линейната обвивка.

Нека от векторите x_1, x_2,..., x_p на $L(x_1, x_2,..., x_p)$ да имаме k линейно независими, които да означим с x_1, x_2,..., x_k. Понеже всеки вектор от $L(x_1, x_2,..., x_p)$ е линейна комбинация на векторите x_1, x_2,..., x_p, то следва, че той ще бъде и такава комбинация от векторите x_1, x_2,..., x_k. Нека y_1, y_2,..., y_m са произволни вектори от $L(x_1, x_2,..., x_p)$, броят на които m надминава k. Ще имаме за тях

$$y_1=\lambda_{11}\ x_1+\lambda_{12}\ x_2+\ldots+\lambda_{1k}\ x_k,$$
$$y_2=\lambda_{21}\ x_1+\lambda_{22}\ x_2+\ldots+\lambda_{2k}\ x_k,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$y_m=\lambda_{m1}\ x_1+\lambda_{m2}\ x_2+\ldots+\lambda_{mk}\ x_k.$$

Но в матрицата

$$\ldots\left\|\begin{matrix}\lambda_{11} & \lambda_{12}\ldots\lambda_{1k}\\ \lambda_{21} & \lambda_{22}\ldots\lambda_{2k}\\ \ldots & \ldots\ldots\\ \lambda_{m1} & \lambda_{m2}\ldots\lambda_{mk}\end{matrix}\right\|\ldots$$

броят на редовете m надминава броя k на стълбовете. Следователно уравненията

$$\lambda_{11}\ \mu_1+\lambda_{21}\ \mu_2+\ldots+\lambda_{m1}\ \mu_m=0,$$
$$\lambda_{12}\ \mu_1+\lambda_{22}\ \mu_2+\ldots+\lambda_{m2}\ \mu_m=0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\lambda_{1k}\ \mu_1+\lambda_{2k}\ \mu_2+\ldots+\lambda_{mk}\ \mu_m=0$$

ще имат решения. Но тогава получаваме лесно равенството

$$\mu_1 y_1+\mu_2 y_2+\ldots+\mu_m y_m=0,$$

отгдето следва, че векторите y_1, y_2,..., y_m са линейно зависими. С това се установява, че подпространството $L(x_1, x_2,..., x_p)$ е от k измерения.

Най-простото подпространство освен нулевото, което не представлява интерес, е едномерното подпространство R_1. За базис на такова подпространство служи един кой да е ненулев вектор e_1. Векторите от това подпространство ще имат формата λe_1, където λ е произволно реално число. Съвкупността от векторите $x=x_0+\lambda e_1$, гдето векторът x_0 принадлежи на R, по аналогия с геометрията се нарича права в линейното пространство R. Ако e_1 и e_2 са два линейно независими вектора от R, то множеството вектори $\lambda_1 e_1+\lambda_2 e_2$, гдето λ_1 и λ_2 са произволни числа, образува двумерно подпространство на R.

Ако x_0 е някой вектор от R, то множеството вектори $x=x_0+\lambda_1 e_1+ +\lambda_2 e_2$ се нарича равнина в R. Изобщо, ако R' е някое подпространство на R и e е вектор от него, то множеството от векторите x_0+е, гдето $x_0 \in R$, се нарича хиперравнина.

6. Трансформация на координатите при промяна на базиса. Нека $e_1, e_2, \ldots e_n$ и $e_1', e_2', \ldots, e_n'$ са два базиса на n-мерното линейно пространство R. Понеже векторите $e_1', e_2', \ldots, e_n'$ принадлежат на R, то те са линейни комбинации на векторите $e_1, e_2, \ldots, e_n$, т. е.

$$
\begin{aligned}
e_1' &= a_{11} e_1 + a_{21} e_2 + \ldots + a_{n1} e_n, \\
e_2' &= a_{12} e_1 + a_{22} e_2 + \ldots + a_{n2} e_n, \\
&\cdots\cdots\cdots\cdots\cdots \\
e_n' &= a_{1n} e_1 + a_{2n} e_2 + \ldots + a_{nn} e_n.
\end{aligned} \tag{1}
$$

Лесно се вижда, че детерминантата от числата a_{ik}, $i, k=1, 2, \ldots, n$, трябва да бъде отлична от нула. Действително, ако тази детерминанта е равна на нула, то системата уравнения

$$
\begin{aligned}
a_{11}\ x_1 + a_{12}\ x_2 + \ldots + a_{1n}\ x_n &= 0, \\
a_{21}\ x_1 + a_{22}\ x_2 + \ldots + a_{2n}\ x_n &= 0, \\
&\cdots\cdots\cdots\cdots \\
a_{n1}\ x_1 + a_{n2}\ x_2 + \ldots + a_{nn}\ x_n &= 0
\end{aligned}
$$

ще има решение $x_1=\lambda_1, x_2=\lambda_2, \ldots, x_n=\lambda_n$. Но тогава ще имаме

$$\lambda_1 e_1' + \lambda_2 e_2' + \ldots + \lambda_n e_n' = 0,$$

което противоречи на условието, че векторите $e_1', e_2', \ldots, e_n'$ са линейно независими. Нека отбележим, че имаме и обратното твърдение: ако детерминантата $A=|a_{ik}|$ е отлична от нула, то векторите $e_1', e_2', \ldots, e_n'$ образуват базис на R. Действително, като решим системата (1) спрямо неизвестните $e_1, e_2, \ldots, e_n$, получаваме, че те са линейни хомогенни функции на векторите $e_1', e_2', \ldots, e_n'$. Но тогава всеки вектор α от R като линейна хомогенна функция на векторите $e_1, e_2, \ldots, e_n$ ще бъде такава функция и на векторите $e_1', e_2', \ldots, e_n'$.

Нека сега представянето на произволен вектор x от R спрямо двата базиса да бъде

$$
\begin{aligned}
x &= \zeta_1 e_1 + \zeta_2 e_2 + \ldots + \zeta_n e_n, \\
x &= \eta_1 e_1' + \eta_2 e_2' + \ldots + \eta_n e_n'.
\end{aligned}
$$

Следователно ще имаме равенството

$$\zeta_1 e_1 + \zeta_2 e_2 + \ldots + \zeta_n e_n = \eta_1 e_1' + \eta_2 e_2' + \ldots + \eta_n e_n'.$$

Като заместим тук векторите e_i' с равните му изрази от (1), получаваме

$$
\begin{aligned}
\zeta_1 e_1 + \zeta_2 e_2 + \ldots + \zeta_n e_n = {} & \eta_1 (a_{11} e_1 + a_{21} e_2 + \ldots + a_{n1} e_n) + \\
& + \eta_2 (a_{12} e_1 + a_{22} e_2 + \ldots + a_{n2} e_n) + \\
& \cdots\cdots\cdots\cdots\cdots\cdots \\
& + \eta_n (a_{1n} e_1 + a_{2n} e_2 + \ldots + a_{nn} e_n).
\end{aligned}
$$

Понеже векторите $e_1, e_2, \ldots, e_n$ са линейно независими, то от предното равенство следва, че коефициентите при тях в двете части трябва да бъдат равни. Така получаваме

$$
\begin{aligned}
\zeta_1 &= a_{11}\eta_1 + a_{12}\eta_2 + \ldots + a_{1n}\eta_n,\\
\zeta_2 &= a_{21}\eta_1 + a_{22}\eta_2 + \ldots + a_{2n}\eta_n,\\
&\cdots\cdots\cdots\cdots\cdots\\
\zeta_n &= a_{n1}\eta_1 + a_{n2}\eta_2 + \ldots + a_{nn}\eta_n.
\end{aligned}
\tag{2}
$$

Виждаме, че матрицата на линейната трансформация (2) е транспонираната матрица $\mathbf{A}'$ на матрицата $\mathbf{A}$ на трансформацията (1). Ако решим системата (2) спрямо $\eta_1, \eta_2, \ldots, \eta_n$, получаваме

$$
\begin{aligned}
\eta_1 &= b_{11}\zeta_1 + b_{12}\zeta_2 + \ldots + b_{1n}\zeta_n,\\
\eta_2 &= b_{21}\zeta_1 = b_{22}\zeta_2 + \ldots + b_{2n}\zeta_n,\\
&\cdots\cdots\cdots\cdots\cdots\\
\eta_n &= b_{n1}\zeta_1 + b_{n2}\zeta_2 + \ldots + b_{nn}\zeta_n,
\end{aligned}
$$

където числата b_{ij} са елементи на обратната матрица на матрицата $\boldsymbol{A}'$.

Глава III

Линейни функции в афинното пространство

1. **Линейни форми.** В теорията на функциите се разглеждат функции на едно или няколко променливи. Една функция $f(x_1, x_2, \ldots, x_n)$ на променливите $x_1, x_2, \ldots, x_n$ може да се разглежда като функция на вектора $(x_1, x_2, \ldots, x_n)$. Можем подобно да разглеждаме функции от векторите на едно произволно линейно пространство. Ние ще се ограничим да разгледаме линейни числови функции на векторни аргументи и линейни векторни функции на векторни аргументи, значенията на които са вектори от пространството, на което принадлежат аргументите. Векторните линейни функции се наричат обикновено **линейни оператори.**

Ако числовата функция $f(x)$, определена в линейното пространство R, удовлетворява условията

$$f(x+y) = f(x) + f(y)$$

за произволни x и y, принадлежащи на R, и

$$f(\lambda x) = \lambda f(x)$$

при произволен вектор x и произволно число λ, то тя се нарича линейна форма.

От тези условия за произволни вектори $x_1, x_2, \ldots, x_k$ от R и произволни числа $\lambda_1, \lambda_2, \ldots, \lambda_k$ получаваме

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k) = \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_k f(x_k).$$

Лесно е тогава да намерим общата форма на линейните функции в едно произволно n-мерно пространство R. Нека $e_1, e_2, \ldots, e_n$ е една база на R. Тогава всеки вектор x от R се изразява с

$$x=\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_n e_n.$$

На основание на горните две свойства ще имаме

$$f(x)=\alpha_1 f(e_1)+\alpha_2 f(e_2)+\ldots+\alpha_n f(e_n)=\sum_{i=1}^{n}\alpha_i c_i,\ c_i=f(e_i).$$

Следователно всяка линейна форма се изразява чрез координатите на вектора x линейно с определени коефициенти $c_1, c_2, \ldots, c_n$.

2. **Линейни оператори.** Ако на всеки вектор x от пространството R направим да съответствува вектор y от същото пространство, то казваме, че в това пространство е зададен операторът $y=A(x)$. Операторът $A(x)$ се нарича линеен, ако удовлетворява на условията:

1°. $$A(x+y)=A(x)+A(y)$$

при произволни вектори x и y от R.

2°. $$A(\lambda x)=\lambda A(x)$$

за произволен вектор x от R и произволно число λ.

1. Например операторът, който прави да отговаря на всеки вектор от пространството R векторът нула, е очевидно линеен. Този оператор се нарича нулев оператор.

2. Оператор на подобие се нарича този, който поставя в съответствие на всеки вектор x вектор kx, където k е фиксирано число. Очевидно този оператор е линеен.

3. Нека L е права, минаваща през нулевата точка на пространството с две измерения (равнината). Да съпоставим на всеки вектор x от нея проекцията му $x'=A(x)$ върху правата L. Лесно се проверява, че условията 1 и 2 са изпълнени и следователно този оператор е линеен.

4. Нека $e_1, e_2, \ldots, e_m, \ldots, e_n$ е базис на едно n-мерно пространство R. На всеки вектор $x=\sum_{1}^{n}\lambda_i e_i$ от R да съпоставим вектора $A(x)=\sum_{1}^{m}\lambda_i e_i$, където $m<n$. Операторът $A(x)$ е линеен, както лесно се вижда от проверката на условията 1 и 2 за него. Той се нарича оператор на проектиране върху подпространството, образувано с векторите $e_1, e_2, \ldots, e_m$.

5. Нека R е n-мерното пространство, образувано от всички полиноми от степен $\leqq n-1$. На всеки полином $P(t)$ да съпоставим полинома

$$AP(t)=P'(t)$$

Понеже

$$A[P_1(t)+P_2(t)]=[P_1(t)+P_2(t)]'=P_1'(t)+P_2'(t)=AP_1(t)+AP_2(t),$$

$$A[\lambda P(t)]=\lambda AP(t),$$

то операторът $A(P)$ (на диференциране) е линеен.

6. Нека K е пространството от безкрайно измерение, съставено от всички диференцируеми функции $x(t)$ в интервала $a \leqq t \leqq b$. Тогава операторът $Dx(t)=x'(t)$ е линеен.

7. Нека на всяка функция $x(t)$ от $C(a, b)$, т. е. всички непрекъснати в (a, b) функции, съпоставим функцията

$$Ax(t)=\int_a^t x(\tau)\, d\tau.$$

Проверяваме лесно, че операторът Ax е линеен.

8. Нека на всяка функция $x(t)$ от интегруемите функции в интервала (a, b) съпоставим функцията

$$y(t)=Ax(t)=\int_a^b K(t, s)\, x(s)\, ds,$$

където $K(t, s)$ е фиксирана функция, отговаряща на условия за интегруемост (например непрекъсната относно s). Операторът е линеен и се нарича оператор на Фредхолм. Функцията $K(t, s)$ се нарича ядро на оператора.

3. **Общ вид на линейния оператор.** Нека $e_1, e_2, \ldots, e_n$ е един произволен базис на едно n-мерно пространство R. Ще намерим общата форма на кой да е линеен оператор $A(x)$. За всеки вектор x от R имаме представянето

$$x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n.$$

Следователно за Ax на основание на свойствата 1 и 2 ще имаме

$$Ax=A(\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n)=\xi_1 A(e_1)+\xi_2 A(e_2)+\ldots+\xi_n A(e_n)=$$

$$=\xi_1 g_1+\xi_2 g_2+\ldots+\xi_n g_n,$$

където $g_i=Ae_i$, $1 \leqq i \leqq n$ са векторите, които отговарят съответно на векторите e_i, $1 \leqq i \leqq n$, от базиса на R.

Обратно, нека $g_1, g_2, \ldots, g_n$ са произволни n вектора от R. Лесно можем да докажем, че съществува линейно преобразуване Ax, и то единствено, което удовлетворява на условията $Ae_i=g_i$, $1 \leqq i \leqq n$. Действително нека на всеки вектор $x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n$ от R да съпоставим вектора $x'=\xi_1 g_1+\xi_2 g_2+\ldots+\xi_n g_n$. Понеже векторът x се представя еднозначно чрез векторите e_i, то на x се прави да отговаря напълно определен вектор $x'=Ax$. Непосредствено се проверява лесно, че така дефинираният оператор е линеен.

Ако $a_{1p}, a_{2p}, \ldots, a_{np}$ са координатите на векторите g_p, $1 \leqq p \leqq n$, то имаме

$$
\begin{aligned}
g_1 &= a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n,\\
g_2 &= a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
g_n &= a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n.
\end{aligned} \tag{1}
$$

Матрицата

$$
A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \ldots a_{1n} \\ a_{21} & a_{22} & a_{23} \ldots a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} \ldots a_{nn} \end{Vmatrix},
$$

в която стълбовете са координатите на векторите $g_1, g_2, \ldots, g_p$, се нарича матрица на линейното преобразуване Ax при базис $e_1, e_2, \ldots, e_n$.

Така на всяко такова преобразуване съответствува една матрица. Обратно, на всяка матрица съответствува едно напълно определено линейно преобразуване, определено с равенствата (1).

Да намерим сега изрази за координатите $\eta_1, \eta_2, \ldots, \eta_n$ на вектора Ax чрез координатите $\xi_1, \xi_2, \ldots, \xi_n$ на вектора x. Понеже

$$
\begin{aligned}
A(x) = A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) &= \xi_1(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n) + \\
+ \xi_2(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n) + \ldots + \xi_n(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n) &= \\
= \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n,&
\end{aligned}
$$

то със сравнение получаваме формулите:

$$
\begin{aligned}
\eta_1 &= a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n,\\
\eta_2 &= a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\eta_n &= a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n.
\end{aligned}
$$

Така достигаме до следния общ резултат:

Ако линейният оператор Ax има в даден базис $e_1, e_2, \ldots, e_n$ съответна матрица **A**, то базисните вектори се преобразуват с помощта на стълбовете на тази матрица, а координатите на произволни вектори — с помощта на хоризонталните й редове.

Да разгледаме някои примери. 1. Нека R е триизмеримо пространство и операторът Ax да представлява проектиране на всеки вектор x върху равнината Oxy. Да вземем единични вектори e_1, e_2, e_3 в направление по осите на координатите. Тогава очевидно ще имаме

$$
Ae_1 = e_1, \quad Ae_2 = e_2, \quad Ae_3 = e_3
$$

и съответствуващата матрица **A** ще има формата

$$
\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}.
$$

2. Нека в n-мерното пространство при базисни вектори $e_1, e_2, \ldots, e_n$ положим

$$A(e_1)=\lambda_1 e_1, \quad A(e_2)=\lambda_2 e_2, \ldots, A(e_n)=\lambda_n e_n.$$

Тогава матрицата **A** има формата

$$\begin{Vmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ 0 & \lambda_2 & 0 \ldots 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & \lambda_n \end{Vmatrix}$$

и се нарича диагонална. За произволен вектор $x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n$ ще имаме

$$A(x)=\lambda_1 \xi_1 e_1+\lambda_2 \xi_2 e_2+\ldots+\lambda_n \xi_n e_n.$$

Операторът Ax в случая се нарича диагонален.

Два оператора Ax и Bx считаме за равни, ако за всеки вектор x имаме $A(x)=B(x)$. Заключаваме лесно, че необходимо и достатъчно условие за равенството на операторите се дава с равенствата $A(e_i)=B(e_i)$, $1 \leqq i \leqq n$.

4. **Действия с линейни оператори.** Ще дефинираме сега действията събиране на оператори, умножение с число и умножение на оператори в дадено n-мерно пространство.

Сумата $C=A+B$ на два оператора A и B се определя с равенството

$$C(x)=A(x)+B(x).$$

Под произведение на оператора A с числото λ разбираме оператора B, $B=\lambda A$, дефиниран с формулата

$$B(x)=\lambda A(x).$$

Лесно се вижда, че с горните дефиниции получаваме линейни оператори. Например за сумата $C=A+B$ имаме за кои да са вектори x и y и числа λ, μ;

$$C(\lambda x+\mu y)=A(\lambda x+\mu y)+B(\lambda x+\mu y)=$$
$$=\lambda A(x)+\mu A(y)+\lambda B(x)+\mu B(y)=$$
$$=\lambda(Ax+Bx)+\mu(Ay+By)=\lambda Cx+\mu Cy.$$

Ако върху вектора x приложим оператора B и след това върху вектора B приложим оператора A, получаваме вектора $A(Bx)$, който означаваме с Cx и пишем накъсо $C=AB$. Лесно се вижда, че C е линеен оператор. Именно имаме

$$C(\lambda x+\mu y)=AB(\lambda x+\mu y)=A(\lambda Bx+\mu By)=$$
$$=\lambda ABx+\mu ABy=\lambda Cx+\mu Cy.$$

Степента на оператора A се определя по-нататък, както е възприето за числата. Именно под A^2 разбираме оператора $A.A$, под A^3 разбираме $A^2.A$ и т. н. Приема се означението $A^0=E$, където E е единичният оператор, т. е. $Ex=x$ за всеки вектор x. Очевидно е, че за кои да е цели неотрицателни числа m и n ще имаме

$$A^n.A^m=A^{n+m}.$$

Например в пространството R от полиномите от степен $\leqq n-1$ операторът на диференциране се определя с

$$DP(t)=P'(t).$$

Тогава имаме $D^2P(t)=D(DP)=D(P')=P''$,... При това за всеки вектор (полином) от това пространство ще имаме

$$D^nP=0.$$

Както видяхме, на всяко линейно преобразуване съответствува една определена матрица. Ще покажем, че на действията с линейните преобразувания съответствуват подобни действия с принадлежащите им матрици. Нека съответните матрици на линейните оператори A и B са $\mathbf{A}=\|a_{ik}\|$, $\mathbf{B}=\|b_{ik}\|$ при едни и същи базични вектори $e_1, e_2, \ldots, e_n$. Имаме

$$Ae_k=\sum_i a_{ik}e_i,\ Be_k=\sum_i b_{ik}e_i.$$

За оператора $C=A+B$ ще имаме

$$Ce_k=Ae_k+Be_k=\sum (a_{ik}+b_{ik})e_i.$$

Следователно на оператора C ще отговаря матрицата $\mathbf{C}$, която е сума от матриците $\mathbf{A}$ и $\mathbf{B}$. Да разгледаме сега оператора $\Gamma=AB$ Имаме

$$ABe_k=A\left(\sum_{s=1}^{n} b_{sk}e_s\right)=\sum_s b_{sk}Ae_s=\sum_{s,i} b_{sk}a_{is}e_i.$$

Следователно можем да пишем

$$\Gamma e_k=\sum_c c_{ik}e_i,\ c_{ik}=\sum_{s=1}^{n} a_{is}b_{sk}.$$

Така елементите на матрицата Γ, която отговаря на оператора Γ, се получават от умножение на елементите от редовете на матрицата $\mathbf{A}$ с елементите от стълбовете на матрицата $\mathbf{B}$, т. е. и имаме $\Gamma=\mathbf{AB}$. С това установяваме следното правило:

Матрицата на сумата от линейни оператори е равна на сумата от матриците на операторите. Матрицата на произведението на два оператора е равна на произведението на матриците на операторите, взето в същия ред.

Свойствата на матриците тогава непосредствено се пренасят върху операторите и за кои да е оператори A, B и C ще имаме равенствата

$$A+B=B+A,$$

$$(A+B)+C=A+(B+C),$$

$$A(BC)=(AB)C,$$

$$(A+B)C=AC+BC,$$

$$C(A+B)=CA+CB.$$

За произволни числа λ и μ ще имаме

$$\lambda(\mu A)=(\lambda\mu)A,$$

$$1.\ A=A,$$

$$(\lambda+\mu)A=\lambda A+\mu A,$$

$$\lambda(A+B)=\lambda A+\lambda B.$$

Ако $P(t)=a_0t^m+a_1t^{m-1}+\ldots+a_m$ е произволен полином, то под оператора $P(A)$ разбираме израза

$$P(A)=a_oA^m+a_1A^{m-1}+\ldots+a_mE,$$

който очевидно представлява линеен оператор.

Ако $\mathbf{A}$ е матрица, то в съответствие с $P(\mathbf{A})$ се означава матрицата

$$P(\mathbf{A})=a_0\mathbf{A}^m+a_1\mathbf{A}^{m-1}+\ldots+a_m\mathbf{E}.$$

Глава IV

Евклидово пространство

1. **Определение.** Едно линейно (афинно) пространство е характерно с това, че в него можем да събираме векторите и да ги умножаваме с произволни числа. При изучаване на евклидовата геометрия ние се срещаме с други факти и свойства, които не могат да се определят само с горните действия върху векторите. Така ние не можем да определим дължина на вектор, ъгъл между два вектора и т. н.

Ще казваме, че в линейното пространство R е определено скаларно произведение, ако на всеки чифт от два вектора x и y е поставено в съответствие реално число, което означаваме с (x, y). При това въведеното съответствие удовлетворява аксиомите:

1. $(x, y)=(y, x)$, т. е. скаларното произведение е симетрично.

2. $(\lambda x, y)=\lambda(x, y)$ за всяко реално число λ.

3. $(x_1+x_2, y)=(x_1, y)+(x_2, y)$ (разпределителност, дистрибутивност) на скаларното произведение.

4. За всеки вектор x, $(x, x)\geqq 0$ и е само тогава нула, когато векторът x е равен на нула.

Линейното пространство R се нарича тогава *евклидово.*

Под дължина на вектора x разбираме числото $\sqrt{(x, x)}$, което се означава с $|x|$. На основание на 4 дължината на един вектор е само тогава нула, когато векторът е нулев.

Да разгледаме някои примери.

В пространството R' от n-мерните вектори

$$x=(a_1, a_2, \ldots, a_n),\ y=(b_1, b_2, \ldots, b_n), \ldots$$

скаларното произведение се дефинира с

$$(x, y)=a_1 b_1+a_2 b_2+ \ldots +a_n b_n.$$

Непосредствено се вижда, че то удовлетворява на аксиомите 1 до 4.

Например имаме

$$(\lambda x, y)=\lambda a_1 . b_1+\lambda a_2 . b_2+ \ldots +\lambda a_n . b_n=\lambda(a_1 b_1+a_2 b_2+ \ldots +a_n b_n)=\lambda(x, y).$$

Дължината на вектора x ще бъде равна на

$$|x|=\sqrt{a_1^2+a_2^2+ \ldots +a_n^2}.$$

Да разгледаме сега пространството $C(a, b)$ от всички непрекъснати реални функции в интервала (a, b). Скаларно произведение на две функции $x=x(i)$ и $y=y(t)$ ще дефинираме с числото

$$(x, y)=\int_a^b x(t)\, y(t)\, dt.$$

Непосредствено се вижда, че аксиомите 1...4 са изпълнени.

Дължината на вектора x ще се дава с

$$x|=\sqrt{\int_a^b x^2(\lambda)\, d\lambda}.$$

Обикновено в този случай изразът $x|$ се означава с $\| x \|$ и се нарича норма на функцията $x(t)$.

От равенството

$$|\lambda x|=\sqrt{(\lambda x, \lambda x)}=\sqrt{\lambda^2 (\lambda, x)}=|\lambda| \sqrt{(x, x)}=|\lambda|\,|x|$$

получаваме, че дължината на един вектор x се умножава с $|\lambda|$, ако умножим вектора с числото λ.

Всеки вектор x с дължина 1 се нарича нормиран. Ако y е произволен вектор, то с умножение на $\frac{1}{|y|}$ той става нормиран.

Множеството F от R се нарича ограничено, ако за всеки вектор $x \in F$ имаме $|x|<M$, като M е фиксирано число. Множеството на всички вектори от R, на които дължината не надминава 1, се нарича *единична сфера.*

2. Неравенство на Коши—Буняковски. Нека x и y са два произволни вектора и t е произволно реално число. На основание на аксиома 4 ще имаме

$$(tx-y,\ tx-y) \geqq 0.$$

Като преработим лявата част на това неравенство съгласно с аксиоми 1, 2, 3, получаваме

$$t^2(x,\ x)-2t(x,\ y)+(y,\ y) \geqq 0, \tag{1}$$

което трябва да е изпълнено за всяко t. Следователно уравнението.

$$t^2(x,\ x)-2t(x,\ y)+(y,\ y)=0 \tag{2}$$

не трябва да има два реални различни корена, т. е. дискриминантата му $(x,\ y)^2-(x,\ x).(y,\ y)$ трябва да бъде неположителна:

$$(x,\ y)^2-(x,\ x)\ (y,\ y) \leqq 0.$$

Оттут получаваме неравенството на Коши—Буняковски:

$$|(x,\ y)| \leqq |x|\ |y|. \tag{3}$$

Ако тук имаме равенство, то уравнението (2) ще има един (двоен) корен реално число t_0, т. е. ще имаме

$$(t_0x-y,\ t_0x-y)=0.$$

На основание на аксиома 4 следва оттук, че $t_0x-y=0$, т. е. $y=t_0x$. В такъв случай векторите x, y наричаме *колинеарни*.

Например в пространството R' за векторите

$$x=(a_1,\ a_2,\ldots,\ a_n),\ y=(b_1,\ b_2,\ldots,\ b_n)$$

скаларното произведение се дефинира с

$$(x,\ y)=a_1b_1+a_2b_2+\ldots+a_nb_n.$$

Следователно неравенството (3) приема вида

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqq \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2.$$

Именно това неравенство е било установено от Коши в 1821 г.

В пространството $C(a,\ b)$ за векторите $x(t)$, $y(t)$ имаме

$$(x,\ y)=\int_a^b x(t)\ y(t)\,dt$$

и неравенството (3) става

$$\int_a^b x(t)\,y(t)\,dt \leqq \sqrt{\int_a^b x^2(t)\,dt \int_a^b y^2(t)\,dt}.$$

Това неравенство е било за пръв път дадено от руския математик В. Я. Буняковски в 1859 г.

Под ъгъл между два вектора x и y разбираме този ъгъл, съдържащ се между 0 и 180^0, на който косинусът е равен на

$$\frac{(x, y)}{|x|\,|y|}.$$

На основание на неравенството на Коши—Буняковски това отношение не надминава по абсолютна стойност 1 и следователно съществува напълно определен ъгъл φ от отношенията

$$\cos\varphi=\frac{(x, y)}{|x|\,|y|},\quad 0\leqq\varphi<180^0. \tag{4}$$

Предполагаме засега, че векторите x и y са ненулеви.

Векторите x и y се наричат ортогонални, ако скаларното им произведение (x, y) е равно на нула. От $(x, y)=0$ се вижда, че ъгълът между тях е равен на 90^0. Понеже $(0, y)=0$, то векторът нула е ортогонален на всеки вектор (при този случай изразът в (4) приема неопределената форма $\frac{0}{0}$).

Нека x и y са два ортогонални вектора. По анология с елементарната геометрия можем да считаме вектора $x+y$ за хипотенуза на правоъгълен триъгълник с катети x и y. По определение квадратът на дължината на вектора $x+y$ се дава с

$$(x+y,\ x+y).$$

На основание на дистрибутивното свойство получаваме

$$(x+y,\ x+y)=(x,\ x+y)+(y,\ x+y)=(x,\ x)+2\,(x,\ y)+(y,\ y). \tag{5}$$

Понеже $(x, y)=0$, то от горното равенство имаме

$$|x+y|^2=(x,\ x)+(y,\ y)=|x|^2+|y|^2. \tag{6}$$

Така получихме, че сборът от квадратите на дължината на две неуспоредни страни на правоъгълника е равен на квадрата на дължината на диагонала. Ако $x_1, x_2, \ldots, x_k$ са два по два ортогонални вектора, то по индуктивен път от (6) получаваме общото равенство

$$|x_1+x_2+\ldots+x_k|^2=|x_1|^2+|x_2|^2+\ldots+|x_k|^2.$$

Понеже по неравенството на Коши—Буняковски имаме $|(x, y)|\leqq|x|\,|y|$, то от (5) получаваме неравенството

$$|x+y|^2=(x+y, x+y)\leqq|x|^2+2|x|\,|y|+|y|^2=(|x|+|y|)^2,$$

т. е.

$$|x+y|\leqq|x|+|y|.$$

За произволни вектори x и y от евклидовото пространство имаме неравенството (неравенство за триъгълника)

$$|x+y| \leqq |x| + |y|.$$

По аналогия с геометрията под разстояние между векторите x и y се разбира дължината d на вектора $x-y$, т. е. $d=|x-y|$. При ортогонални вектори получаваме лесно, както по-горе, формулата

$$d^2 = |x|^2 + |y|^2.$$

3. **Ортогонална база.** Казваме, че векторите $e_1, e_2, \ldots, e_n$ от n-мерното евклидово пространство R, никой от които не е нулев, образуват ортогонална база, ако два по два са ортогонални помежду си.

За да оправдаем в дефиницията въведената дума база, трябва да установим, че векторите са линейно независими. Нека допуснем, че съществува равенството

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0.$$

Но тогава ще имаме например

$$(\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n, e_1) = \lambda_1 (e_1, e_1) + \lambda_2 (e_2 e_1) + \ldots + \lambda_n (e_n, e_1) = 0.$$

Понеже $(e_1, e_1) \neq 0$ и $(e_k, e_1) = 0$ за $2 \leqq k \leqq n$, то получаваме, че $\lambda_1 = 0$. Аналогично с умножение на $e_2, e_3, \ldots, e_n$ последователно получаваме, че и $\lambda_2 = 0$, $\lambda_3 = 0, \ldots, \lambda_n = 0$, с което се доказва линейната независимост на векторите $e_1, e_2, \ldots, e_n$.

Ще установим сега, че във всяко n-мерно пространство съществува ортогонална база. За тази цел ще използуваме начина на ортогонализация на Грам и Шмид. Нека $g_1, g_2, \ldots, g_n$ е една база от вектори на n-мерното пространство R. Ортогоналните базични вектори ще получим именно като линейни комбинации на предните вектори. Да положим $e_1 = g_1$ и да търсим e_2 във формата $e_2 = g_2 + \mu_1 e_1$. Трябва да подберем числото μ_1 по такъв начин, че този вектор да е ортогонален на e_1, т. е. да имаме

$$(g_2 + \mu_1 e_1, e_1) = 0 \quad (g_2, e_1) + \mu_1 (e_1, e_1) = 0.$$

Оттук получаваме за μ_1 стойността

$$\mu_1 = -\frac{(g_2, e_1)}{(e_1, e_1)}.$$

Продължаваме да търсим така вектор e_3, който е линейна комбинация на векторите e_1 и g_2 и е ортогонален на e_1 и e_2 и т. н. Нека за общност приемем, че сме намерили $(p-1)$-те вектори $e_1, e_2, \ldots, e_{p-1}$, които са ортогонални два по два и отлични от нула. Ще търсим $p-1$ числа, такива, че векторът

(7) $$e_p = g_p + \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_{p-1} e_{p-1}$$

да бъде ортогонален на векторите e_1, $e_2, \ldots$, e_{p-1}, които сме вече построили. Ще имаме равенствата:

$$(g_p+\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_{p-1} e_{p-1}, e_1)=0,$$
$$(g_p+\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_{p-1}, e_{p-1}, e_2)=0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(g_p+\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_{p-1} e_{p-1}, e_{p-1})=0.$$

Понеже векторите e_1, $e_2, \ldots$, e_{p-1} са два по два ортогонални, то предните равенства стават:

$$(g_p, e_1)+\alpha_1 (e_1, e_1)=0,$$
$$(g_p, e_2)+\alpha_2 (e_2, e_2)=0,$$
$$\ldots\ldots\ldots\ldots$$
$$(g_p, e_{p-1})+\alpha_{p-1} (e_{p-1}, e_{p-1})=0.$$

Оттук намираме

$$\alpha_1=-\frac{(g_p, e_1)}{(e_1, e_1)}, \quad \alpha_2=-\frac{(g_p, e_2)}{(e_2, e_2)}, \ldots, \quad \alpha_{p-1}=-\frac{(g_p, e_{p-1})}{(e_{p-1}, e_{p-1})}.$$

Ако в (7) заместим векторите e_1, $e_2, \ldots$, e_{p-1} с равните им изрази като линейни комбинации на векторите g_1, $g_2, \ldots$, g_{p-1}, то получаваме

$$e_p=g_p+l_1 g_1+l_2 g_2+\ldots+l_{p-1} g_{p-1},$$

гдето l_1, $l_2, \ldots$, l_{p-1} са числа. Оттук непосредствено следва, че e_p не е нулев вектор, понеже в противен случай би съществувала линейна зависимост между векторите g_1, $g_2, \ldots, g_{p-1}$, g_p, което противоречи на условието. Така получаваме n вектора e_1, $e_2, \ldots$, e_{n-1}, e_n, които образуват ортогонална база на n-мерното пространство и всеки вектор от него е линейна комбинация от тях. Ако заместим векторите e_k с векторите

$$e'_k=\frac{e_k}{|e_k|}, \quad 1\leqq k\leqq n,$$

получаваме нормирани базични вектори. Една такава база ще наричаме нормирана.

Нека e_1, $e_2, \ldots, e_n$ е една ортогонална нормирана база на n-мерното евклидово пространство R. За два кои да са вектора x, y,

$$x=\lambda_1 e_1+\lambda_2 e_2+\ldots+\lambda_n e_n, \quad y=\mu_1 e_1+\mu_2 e_2+\ldots+\mu_n e_n,$$

от R ще имаме

$$(x, y)=(\lambda_1 e_1+\lambda_2 e_2+\ldots+\lambda_n e_n, \mu_1 e_1+\mu_2 e_2+\ldots+\mu_n e_n)=$$
$$=\sum_{i=1}^{n} \lambda_i \mu_i (e_i, e_i)+\sum_{i\neq\tau} \lambda_j \mu_\tau (e_j, e_\tau)=\sum_{i=1}^{n} \lambda_i \mu_i.$$

Така получаваме, че скаларното произведение на два вектора е равно на сумата от произведенията на съответствуващите им координати.

Оттук за дължината $|x|$ на вектора x получаваме

$$|x|^2=(x,\ x)=\lambda_1{}^2+\lambda_2{}^2+\ldots+\lambda_n{}^2,$$

т. е.

$$|x|=\sqrt{\lambda_1{}^2+\lambda_2{}^2+\ldots+\lambda_n{}^2}\ .$$

Ако умножим равенството

$$x=\lambda_1 e_1+\lambda_2 e_2+\ldots+\lambda_s e_s+\ldots+\lambda_n e_n$$

с вектора e_s и вземем скаларното произведение, получаваме

$$(x,\ e_s)=\lambda_1(e_1,\ e_s)+\lambda_2(e_2,\ e_s)+\ldots+\lambda_s(e_s,\ e_s)+\ldots+\lambda_n(e_n,\ e_s)=\lambda_s.$$

Така виждаме, че координатите на вектора x в ортогонален нормиран базис са скаларните произведения на вектора x със съответни базисни вектори.

Нека от пространството C (—1, 1) разгледаме едно подпространство от $n+1$ измерения, съставено от всички полиноми, на които степента не надминава n. Скаларното произведение на два вектора $P(t)$, $Q(t)$ от това пространство е дефинирано с

$$(P, Q)=\int_{-1}^{1} P(t)\ Q(t)\,dt.$$

Очевидно векторите $1,\ t,\ t^2,\ldots,\ t^n$ образуват един базис на това пространство и при ортогонализация ще получим редицата от полиноми:

$$1, t, t^2-\frac{1}{3},\ t^3-\frac{3}{5}\ t,\ldots$$

(при нормировка на коефициента пред най-високата степен, взет за 1). Получените полиноми, които се определят при ортогонализацията до постоянен множител, се наричат полиноми на Лежандр. По една формула на Родриг полиномите на Лежандр могат да бъдат определени с

$$P_n(t)=\frac{1}{2^n n!}\ [(t^2-1)^n]^{(n)}. \tag{8}$$

Оттук се вижда, че $P_n(t)$ е полином от n-та степен, като $P_n(t)=$ $=A_n t^n+\ldots$ и $A_n=\ \frac{(2n)!}{2^n(n!)^2}$. Полиномите $P_v(t)$ са ортогонални помежду си, т. е. имаме

$$\int_{-1}^{1} P_v(t)\,P_\mu(t)\,dt=0,\ \mu\neq v.$$

Проверката на това равенство става лесно, като се докаже, че

$$\int_{-1}^{1} t^k P_v(t)\, dt = 0$$

за всяко цяло число $k, 0 \leqq k \leqq v-1$. Но последното равенство се получава лесно с интегриране по части, като в него се замести полиномът $P_v(t)$ с равната му стойност от (8). Така получаваме, че полиномите $P_0, P_1, P_2, \ldots, P_n$ образуват ортогонален базис на пространството C_{n+1}. За да нормираме векторите на този базис, трябва да пресметнем интегралите

$$\delta_n = \int_{-1}^{1} P_n^2(t)\, dt = A_n \int_{-1}^{1} t^n P_n(t)\, dt = \frac{A_n}{2^n n!} \int_{-1}^{1} t^n [(t^2-1)]^{(n)}\, dt.$$

С интегриране по части намираме $\delta_n = \frac{2}{2n+1}$ и следователно нормираните полиноми $Q_n(t)$ ще бъдат

$$Q_n(t) = \sqrt{n+\frac{1}{2}}\, P_n(t).$$

Тогава за всеки полином $f(t)$ от степен $\leqq n$ ще имаме

$$f(t) = c_0 Q_0(t) + c_1 Q_1(t) + \ldots + c_n Q_n(t),$$

като

$$c_k = \int_{-1}^{1} f(t)\, Q_k(t)\, dt,\ k = 0, 1, 2, \ldots, n.$$

Да разгледаме сега множеството на всички тъй наречени тригонометрични полиноми от n-ти ред:

$$P(t) = \frac{a_0}{2} + a_1 \cos t + b_1 \sin t + a_2 \cos 2t + b_2 \sin 2t + \ldots + a_n \cos nt + b_n \sin nt,$$

гдето $a_0, a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ са произволни реални числа. Очевидно това множество образува $2n+1$-мерно пространство R. За скаларно произведение за кои да е негови вектори $P_1(t)$ и $P_2(t)$ приемаме числото

$$(P_1, P_2) = \int_0^{2\pi} P_1(t)\, P_2(t)\, dt.$$

Лесно се вижда, че специалните полиноми

$$1, \cos t, \sin t, \cos 2t, \sin 2t, \ldots, \cos nt, \sin nt$$

са ортогонални два по два в интервала $(0, 2\pi)$, т. е. $(p \neq q)$,

$$\int_0^{2\pi} \cos pt \cos qt\, dt = 0, \quad \int_0^{2\pi} \cos pt \sin qt\, dt = 0, \quad \int_0^{2\pi} \sin pt \sin qt\, dt = 0.$$

Понеже

$$\int_0^{2\pi} \cos^2 pt\, dt = \pi,\quad \int_0^{2\pi} \sin^2 pt\, dt = \pi\,(p>0),\quad \int_0^{2\pi} dt = 2\pi,$$

функциите

$$e_0 = \frac{1}{\sqrt{2\pi}},\quad e_1 = \frac{\cos t}{\sqrt{\pi}},\quad e_2 = \frac{\sin t}{\sqrt{\pi}},\ \ldots,\ e_{2n-1} = \frac{\cos nt}{\sqrt{\pi}},\quad e_{2n} = \frac{\sin nt}{\sqrt{\pi}}$$

образуват една ортогонална нормирана база за R.

4. **Детерминанта на Грам.** Нека $\alpha_1, \alpha_2, \ldots, \alpha_m$ са m произволни вектора от едно евклидово пространство. Детерминантата

$$\Delta = \begin{vmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) \ldots (\alpha_1, \alpha_m) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) \ldots (\alpha_2, \alpha_m) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ (\alpha_m, \alpha_1) & (\alpha_m, \alpha_2) \ldots (\alpha_m, \alpha_m) \end{vmatrix}$$

се нарича детерминанта на Грам. Да предположим, че векторите са линейно зависими, т. е. имаме равенството

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \ldots + \lambda_m \alpha_m = 0,$$

гдето поне едно от числата $\lambda_1, \lambda_2, \ldots, \lambda_m$ е отлично от нула. Като умножим това равенство скаларно с $\alpha_1, \alpha_2, \ldots, \alpha_m$, получаваме равенствата:

$$\begin{aligned} &\lambda_1 (\alpha_1, \alpha_1) + \lambda_2 (\alpha_1, \alpha_2) + \ldots + \lambda_m (\alpha_1, \alpha_m) = 0, \\ &\lambda_1 (\alpha_2, \alpha_1) + \lambda_2 (\alpha_2, \alpha_2) + \ldots + \lambda_m (\alpha_2, \alpha_m) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\lambda_1 (\alpha_m, \alpha_1) + \lambda_2 (\alpha_m, \alpha_2) + \ldots + \lambda_m (\alpha_m, \alpha_m) = 0. \end{aligned}$$

Детерминантата от коефициентите на $\lambda_1, \lambda_2, \ldots, \lambda_m$ е точно равна на детерминантата Δ на Грам. Понеже системата хомогенни уравнения има нетривиалното решение $\lambda_1, \lambda_2, \ldots, \lambda_m$, то детерминантата Δ трябва да е равна на нула.

Нека сега предположим, че векторите $\alpha_1, \alpha_2, \ldots, \alpha_m$ са линейно независими. Чрез ортогонализиране ще получим от тях m ортогонални и нормирани вектора $e_1, e_2, \ldots, e_m$, които са също линейно независими. Дадените вектори $\alpha_1, \alpha_2, \ldots, \alpha_m$ са линейни комбинации на последните вектори, т. е. ще имаме

$$\alpha_i = a_{i1}\ e_1 + a_{i2}\, e_2 + \ldots + a_{im}\, e_m,\ i = 1,\ 2,\ 3, \ldots, m.$$

Понеже за скаларните произведения (α_i, α_j) имаме

$$(\alpha_i, \alpha_j) = a_{i1}\ a_{j1} + a_{i2}\, a_{j2} + \ldots + a_{im}\, a_{jm},$$

детерминантата на Грам Δ ще бъде равна на квадрата на детерминантата

$$D=\begin{vmatrix} a_{11} & a_{12} \dots a_{1m} \\ a_{21} & a_{22} \dots a_{2m} \\ \cdot & \cdot \cdot \cdot \cdot \cdot \\ a_{m1} & a_{m2} \dots a_{mm} \end{vmatrix}.$$

Но от линейната независимост на векторите $\alpha_1, \alpha_2, \dots, \alpha_m$ следва, че детерминантата D е отлична от нула, т. е. детерминантата Δ е положителна. Така установихме предложението:

Детерминантата на Грам на една система от вектори е неотрицателна. Тя е само тогава равна на нула, когато векторите от системата са линейно зависими.

В частност неравенството

$$\begin{vmatrix} (\alpha_1, \alpha_1)(\alpha_1, \alpha_2) \\ (\alpha_2, \alpha_1)(\alpha_2, \alpha_2) \end{vmatrix} \geqq 0$$

е идентично с неравенството на Коши—Буняковски:

$$(\alpha_1, \alpha_2)^2 \leqq (\alpha_1, \alpha_1)\ (\alpha_2, \alpha_2).$$

В пространството $C(a, b)$ от непрекъснати функции в (a, b) детерминантата на Грам е

$$\Delta' = \begin{vmatrix} \int_a^b f_1^2(x)\,dx & \int_a^b f_1(x) f_2(x)\,dx \dots & \int_a^b f_1(x) f_n(x)\,dx \\ \int_a^b f_1(x) f_2(x)\,dx & \int_a^b f_2^2(x)\,dx \dots & \int_a^b f_2(x)\ f_n(x)\,dx \\ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot & \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot & \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ \int_a^b f_1(x) f_n(x)\,dx & \int_a^b f_2(x) f_n(x)\,dx \dots & \int_a^b f_n^2(x)\,dx \end{vmatrix}.$$

Следователно необходимо и достатъчно условие функциите $f_1(x)$, $f_2(x), \dots, f_n(x)$ да бъдат линейно зависими се състои в това детерминантата Δ' да е равна на нула.

5. **Перпендикуляр към подпространство.** Аналогично на елементарната геометрия да решим следната задача: даден вектор f от едно евклидово пространство R да разложим на два вектора g и h, от които g да принадлежи на дадено подпространство R' на R и h да е ортогонален на това подпространство R'. Нака R' е p-мерно и $e_1, e_2, \dots, e_p$ е една ортогонална и нормирана база на R'. Понеже $g \in R'$, то ще имаме

$$g = \lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_p e_p.$$

Векторът $h = f - g$ трябва да е ортогонален на R' и следователно ще имаме

$$(h, e_k) = (f - g, e_k) = 0, \quad k = 1, 2, 3, \dots, p$$

или

$$(f-g,\ e_k)=(f-\lambda_1 e_1-\lambda_2 e_2-\ \ldots\ -\lambda_p e_p,\ e_k)=(f, e_k)-\lambda_k=0.$$

Оттук получаваме за числата λ_k изразите

$$\lambda_k=(f,\ e_k),\ k=1,\ 2,\ 3,\ldots,\ p.$$

Векторът g се нарича проекция на вектора f върху подпространството R', а h—перпендикуляр, спуснат от края на вектора f върху R'. По теоремата на Питагор имаме

$$|f|^2=|g|^2+|h|^2,$$

отгдето следва, че $|h|\leqq|f|$, т. е. дължината на перпендикуляра не надминава дължината на вектора (наклонената).

Ако $|h|=0$, то $f=g$, което означава, че f принадлежи на R'. Условието $|h|=|f|$ показва, че $g=0$, т. е. $f=0+h$, и значи в този случай f е ортогонален на подпространството R'. При всички останали положения дължината на вектора h е положителна величина, по-малка от дължината на f.

Ще решим същата задача, като изхождаме от една произволна база $\omega_1,\ \omega_2,\ldots,\ \omega_p$ на подпространството R'. За вектора

$$g=\mu_1\omega_1+\mu_2\ \omega_2+\ \ldots\ +\mu_p\omega_p$$

уравненията (1) стават

$$(h, \omega_1)=(f-g, \omega_1)=(f, \omega_1)-\mu_1(\omega_1, \omega_1)-\mu_2(\omega_2, \omega_1)-\ \ldots-\mu_p(\omega_p, \omega_1)=0,$$

$$(h, \omega_2)=(f-g, \omega_2)=(f, \omega_2)-\mu_1(\omega_1, \omega_2)-\mu_2(\omega_2, \omega_2)-\ \ldots\ -\mu_p(\omega_p, \omega_2)=0,$$

. . . . . .

$$(h,\ \omega_p)=(f-g, \omega_p)=(f, \omega_p)-\mu_1(\omega_1, \omega_p)-\mu_2(\omega_2, \omega_p)-\ \ldots\ -\mu_p(\omega_p, \omega_p)=0.$$

Тук неизвестни са числата $\mu_1,\ \mu_2,\ldots,\mu_p$. Детерминантата пред неизвестните е равна на

$$\Delta=\begin{vmatrix}(\omega_1,\ \omega_1)\ (\omega_2,\ \omega_1)\ldots(\omega_p,\ \omega_1)\\(\omega_2,\ \omega_1)\ (\omega_2,\ \omega_2)\ldots(\omega_p,\ \omega_2)\\ \ldots\ldots\ldots\ldots\\(\omega_1,\ \omega_p)\ (\omega_2,\ \omega_p)\ldots(\omega_p,\ \omega_p)\end{vmatrix}.$$

Понеже векторите $\omega_1,\ \omega_2,\ldots,\omega_p$ са линейно независими, то детерминантата на Грам за тях е отлична от нула и системата има едно напълно определено решение, което можем да получим по известен начин.

6. **Изоморфизъм на евклидови пространства.** Две евклидови пространства R и R' се наричат изоморфни, ако между елементите им $x, y,\ldots$ и $x', y',\ldots$ може да се установи такова еднозначно съответствие, означено с $x\longleftrightarrow x'$, че да бъдат изпълнени следните условия:

1. Ако $x\longleftrightarrow x'$ и $y\longleftrightarrow y'$, то $x+y\longleftrightarrow x'+y'$, т. е. ако на векторите x, y от R съответстват векторите x', y' от R', то на вектора $x+y$ съответства векторът $x'+y'$.

2. Ако $x \longleftrightarrow x'$, то $\lambda x \longleftrightarrow \lambda x'$.

3. Ако $x \longleftrightarrow x'$ и $y \longleftrightarrow y'$, то $(x, y)=(x', y')$, т. е. скаларните произведения на съответствуващи чифтове вектори са равни помежду си.

Ще установим сега основната теорема :

Всички евклидови пространства от дадено измерение n са изоморфни помежду си.

При доказателството ще изберем пространството от векторите

$$\alpha=(a_1, a_2, \ldots, a_n)$$

за основно. Скаларното произведение на векторите $\alpha=(a_1, a_2, \ldots, a_n)$ и $\beta=(b_1, b_2, \ldots, b_n)$ се дава, както вече знаем, с

$$(\alpha, \beta)=a_1 b_1+a_2 b_2+ \ldots +a_n b_n.$$

Нека R е произволно n-мерно евклидово пространство, изоморфно на T, и $e_1, e_2, \ldots, e_n$ е една негова ортогонална и нормирана база. На всеки вектор

$$x=\xi_1 e_1+\xi_2 e_2+ \ldots + \xi_n e_n$$

от R да съпоставим вектора

$$x'=(\xi_1, \xi_2, \ldots, \xi_n)$$

от пространството Т. Така очевидно получаваме съответствието, което е еднозначно. Не е трудно да се види, че условията 1, 2, 3 са изпълнени, т. е. пространствата R и T са изоморфни помежду си. Така свойствата 1 и 2 следват от

$$\lambda x'=(\lambda\xi_1, \lambda\xi_2, \ldots, \lambda\xi_n) \longleftrightarrow \lambda\xi_1 e_1+\lambda\xi_2 e_2+ \ldots +\lambda\xi_n e_n=\lambda x,$$

$$x'+y'=(\xi_1+\eta_1, \xi_2+\eta_2, \ldots, \xi_n+\eta_n) \longleftrightarrow (\xi_1+\eta_1) e_1+(\xi_2+\eta_2) e_2+ \ldots +$$

$$+(\xi_n+\eta_n) e_n=x+y.$$

По формулата за скаларно произведение при нормирана ортогонална база имаме $(x, y)=\xi_1 \eta_1+\xi_2 \eta_2+ \ldots +\xi_n \eta_n$ и понеже $(x', y')=$ $=\xi_1 \eta_1+\xi_2 \eta_2 + \ldots + \xi_n \eta_n$, то $(x, y)=x', y')$, с което е установено свойството 3.

Доказаната теорема показва, че всяко свойство, отнасящо се за събиране, умножение с числа и скаларни произведения на вектори от едно евклидово пространство, остава в сила и за всяко негово изоморфно евклидово пространство. Така произволно геометрично твърдение за два или три вектора можем да проверим, като го разгледаме за съответните вектори от известното триизмеримо пространство от елементарната геометрия. Например неравенството за два вектора x и y

$$|x+y| \leqq |x|+|y|$$

следва непосредствено от свойството в геометрията, че диагоналът на един кой да е паралелограм по дължина не надминава сумата от дължините на две неуспоредни негови страни.

Глава V

Билинейни и квадратични форми

1. Билинейни форми. Една функция $A(x; y)$ от векторите x и y на едно афинно пространство R се нарича билинейна форма (билинейна функция) на векторите x и y, ако удовлетворява следните условия:

1. При фиксиран вектор y, $A(x; y)$ е линейна функция на x.

2. При фиксиран вектор x, $A(x; y)$ е линейна функция на y.

На основание на дефиницията на линейна функция предните условия са еквивалентни на следните:

1. $A(x_1+x_2; y)=A(x_1; y)+A(x_2; y)$ и $A(\lambda x; y)=\lambda A(x; y)$ за всяко реално число λ.

2. $A(x; y_1+y_2)=A(x; y_1)+A(x; y_2)$ и $A(x, \mu y)=\mu A(x; y)$ за всяко реално число μ.

Една билинейна форма $A(x; y)$ се нарича симетрична, ако за всеки два вектора x и y имаме

$$A(x; y)=A(y; x). \tag{1}$$

Лесно се вижда, че скаларното произведение (x, y) на два вектора x и y от едно евклидово пространство е симетрична билинейна форма.

Ще намерим общата форма на $A(x, y)$. Нека $e_1, e_2, \ldots, e_n$ е произволен базис на n-мерното линейно пространство R. За векторите x и y имаме

$$x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n,$$

$$y=\eta_1 e_1+\eta_2 e_2+\ldots+\eta_n e_n.$$

На основание на свойствата 1 и 2 получаваме

$$\begin{aligned}A(x; y)&=A(\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n;\ \eta_1 e_1+\eta_2 e_2+\ldots+\eta_n e_n)=\\&=\xi_1\eta_1 A(e_1; e_1)+\xi_1\eta_2 A(e_1; e_2)+\ldots+\xi_1\eta_n A(e_2; e_n)+\\&+\xi_2\eta_1 A(e_2; e_1)+\xi_2\eta_2 A(e_2; e_2)+\ldots+\xi_2\eta_n A(e_2; e_n)+\\&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\&+\xi_n\eta_1 A(e_n; e_1)+\xi_n\eta_2 A(e_n; e_2)+\ldots+\xi_n\eta_n A(e_n; e_n)\end{aligned}$$

или

$$A(x; y)=\sum_{i,k=1}^{n} a_{ik}\xi_i\eta_k, \tag{2}$$

където a_{ik} означават числата $A(e_i; e_k)$. Така получаваме предложението:

Всяка билинейна форма $A(x; y)$ при базис $e_1, e_2, \ldots, e_n$ в n-мерно пространство може да се представи във вида

$$A(x; y)=\sum_{i,k=1}^{n} a_{ik}\xi_i\eta_k, \tag{3}$$

където $\xi_1, \xi_2, \ldots, \xi_n$ са координатите на вектора x, $\eta_1, \eta_2, \ldots, \eta_n$ са координатите на вектора y и числата a_{ik}, зависещи само от базиса, се дават с $a_{ik} = A(e_i; e_k)$.

Матрицата $\mathbf{A} = \| a_{ik} \|$ от числата a_{ik} се нарича матрица на билинейната форма при базис $e_1, e_2, \ldots, e_n$.

Ако билинейната форма $A(x, y)$ е симетрична, то ще имаме

$$a_{ik} = A(e_i; e_k) = A(e_k; e_i) = a_{ki}$$

и следователно матрицата $\mathbf{A}$ на формата $A(x; y)$ ще съвпада с транспонираната матрица $\mathbf{A}'$. Непосредствено се вижда, че ако матриците $\mathbf{A}$ и $\mathbf{A}'$ са равни, то билинейната форма $A(x; y)$ е симетрична.

2. Трансформацията на матрицата на билинейната форма при изменение на базиса. Нека в n-мерното пространство вземем друг базис $e_1', e_2', \ldots, e_n'$. Векторите $e_1', e_2', \ldots, e_n'$ се изразяват линейно чрез векторите $e_1, e_2, \ldots, e_n$:

$$\begin{aligned}
e_1' &= c_{11} e_1 + c_{21} e_2 + \ldots + c_{n1} e_n, \\
e_2' &= c_{12} e_1 + c_{22} e_2 + \ldots + c_{n2} e_n, \\
&\ldots\ldots\ldots\ldots \\
e_n' &= c_{1n} e_1 + c_{2n} e_2 + \ldots + c_{nn} e_n.
\end{aligned}$$

Тук числата $c_{1p}, c_{2p}, \ldots, c_{np}$ са координатите на вектара e_p' при базис $e_1, e_2, \ldots, e_n$. Матрицата

$$\mathbf{C} = \begin{Vmatrix} c_{11} & c_{12} \ldots c_{1n} \\ c_{21} & c_{22} \ldots c_{2n} \\ \ldots\ldots\ldots \\ c_{n1} & c_{n2} \ldots c_{nn} \end{Vmatrix},$$

в която стълбовете представляват координатите на векторите $e_1', e_2', \ldots, e_n'$, се нарича матрица на прехода от базиса $e_1, e_2, \ldots, e_n$ към базиса $e_1', e_2', \ldots, e_n'$. Нека

$$\mathbf{A} = \begin{Vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots\ldots\ldots \\ a_{n1} & a_{n2} \ldots a_{nn} \end{Vmatrix}$$

е матрицата на формата $A(x; y)$ при базис $e_1, e_2, \ldots, e_n$ и

$$\mathbf{B} = \begin{Vmatrix} b_{11} & b_{12} \ldots b_{1n} \\ b_{21} & b_{22} \ldots b_{2n} \\ \ldots\ldots\ldots \\ b_{n1} & b_{n2} \ldots b_{nn} \end{Vmatrix}.$$

е матрицата на същата форма при базис e_1', e_2'.... e_n'. Числата

$$b_{pq} = A(e_p'; e_q')$$

очевидно представляват значенията на билинейната форма $A(x; y)$ при $x = e_p'$ и $y = e_q'$. Следователно по формулата (3) те се намират, като вместо променливите $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$ поставим координатите $c_{1p}, c_{2p}, \ldots, c_{np}$; $c_{1q}, c_{2q}, \ldots, c_{nq}$ на векторите e_p' и e_q' в базиса $e_1, e_2, \ldots, e_n$. Значи ще имаме

$$b_{pq} = A(e_p'; e_q') = \sum_{i,k=1}^{n} a_{ik} c_{ip} c_{kq}.$$

Ако с $\mathbf{C}' = \| c'_{pi} \|$ означим транспонираната матрица на матрицата $\mathbf{C}$, то предното равенство става

$$b_{pq} = \sum_{i,k=1}^{n} c'_{pi} a_{ik} c_{kq}.$$

Това ни показва, че матрицата B се дава с равенството

$$\mathbf{B} = \mathbf{C}'\mathbf{A}\mathbf{C}.$$

Следователно, ако $\mathbf{A}$ и $\mathbf{B}$ са матриците на билинейната форма $A(x; y)$ съответно при базиси $e_1, e_2, \ldots, e_n$ и $e_1', e_2', \ldots e_n'$, то $\mathbf{B} = \mathbf{C}'\mathbf{A}\mathbf{C}$, където $\mathbf{C}$ е матрицата на прехода от базиса $e_1, e_2, \ldots, e_n$ към базиса $e_1', \ldots, e_2', e_n'$ и $\mathbf{C}'$ е транспонираната матрица на $\mathbf{C}$.

3. **Квадратични форми.** Ако $A(x; y)$ е произволна симетрична билинейна форма, то функцията $A(x; x)$ се нарича квадратична форма.

Следователно на основание на формула (3) при $y = x$ квадратичната форма $A(x; x)$ ще се представя така:

$$A(x; x) = \sum_{i,k=1}^{n} a_{ik} \xi_i \xi_k,$$

където $a_{ik} = a_{ki}$.

Обратно, на квадратичната форма $A(x; x)$ отговаря билинейната форма $A(x; y)$, и то еднозначно.

Действително на основание на определението на билинейна форма имаме

$$A(x+y; x+y) = A(x; x) + A(x; y) + A(y; x) + A(y; y).$$

Следствие на симетрията $A(x; y) = A(y; x)$ получаваме

$$A(x; y) = \frac{1}{2} [A(x+y; x+y) - A(x; x) - A(y; y)],$$

като вдясно членовете са стойности на квадратичната форма $A(x; x)$

Понеже числата a_{ik} са равни на $A(e_i; e_k)$, то квадратичната форма зависи от избрания базис. При избор на нов базис на пространството координатите на вектора x се преобразуват линейно; Следователно можем да приложим резултатите от глава V на част I. Ще имаме следните предложения:

Нека $A(x, y)$ е произволна квадратична форма в n-мерното пространство R. Тогава в R съществува базис $e_1, e_2, \ldots, e_n$, в който квадратичната форма $A(x; x)$ има вида

$$A(x; x)=\lambda_1 \zeta_1^2+\lambda_2 \zeta_2^2+\ldots+\lambda_n \zeta_n^2,$$

гдето $\zeta_1, \zeta_2, \ldots, \zeta_n$ са координатите на вектора x в новия базис.

Ако квадратичната форма е приведена по два различни начина (т. е. при различни базиси) в сума от квадрати, то числото на положителните квадрати, както и числото на отрицателните квадрати, и в двата случая е едно и също. Също така и борят на членовете с коефициенти, равни на нула, е еднакъв и при двете представяния.

Матрицата на квадратичната форма в различни базиси има един и същ ранг, равен на броя на членовете в представянето на формата в сума от квадрати, на които коефициентите са отлични от нула.

Квадратичната форма $A(x; x)$ е положително дефинитна, ако $A(x; x) > 0$ за всеки вектор $x \neq 0$. Ако $A(x; y)$ е съответната ѝ билинейна форма, то ще имаме

$$A(x;y)=A(y;x),$$

$$A(x_1+x_2; y)=A(x_1; y)+A(x_2; y),$$

$$A(\lambda x; y)=\lambda A(x; y),$$

$$A(x, x) \geqq 0 \text{ и } A(x; x) > 0 \text{ при } x \neq 0.$$

От тези равенства виждаме, че билинейната форма $A(x; y)$ има същите свойства, както скаларното произведение (x, y) на два вектора x и y. Така получаваме предложението:

Скаларното произведение е билинейна форма, съответствуваща на положително дефинитна квадратична форма, и всяка такава форма може да бъде взета за скаларно произведение.

Глава VI

Комплексно n-мерно пространство

1. Комплексно линейно пространство. В началото дефинирахме реално линейно пространство като множество от елементи, наречени вектори, които се подчиняват на условията 1, 2, 3, I, II, III, (гл. II) при реални числа λ. Ще разгледаме сега подобни множества, като числата λ могат да бъдат комплексни. Следователно едно множество R от еле-

менти-вектори ще наричаме комплексно линейно пространство, ако предположенията 1—3, I—III (гл. II) са изпълнени при произволни комплексни числа. Всички свойства на реалните пространства остават в сила и за компелксните пространства, като реалните числа се заместват с комплексни числа. Така всеки вектор x от n-мерното пространство ще има формата

$$x=\zeta_1 e_1+\zeta_2 e_2+\ldots+\zeta_n e_n,$$

гдето числата $\zeta_1, \zeta_2, \ldots, \zeta_n$, наречени координати, са комплексни числа и векторите $e_1, e_2, \ldots, e_n$ са линейно независими, т. е. образуват базис на пространството R.

Една функция $f(x)$ на произволен вектор x от R ще наричаме линейна от първи ред, ако тя удовлетворява условията

$$f(x+y)=f(x)+f(y),$$

$$f(\lambda x)=\lambda f(x)$$

за произволно комплексно число λ. Функцията $f(x)$ на вектора x се нарича линейна от втори род, ако удовлетворява условията

$$f(x+y)=f(x)+f(y),$$

$$f(\lambda x)=\bar{\lambda} f(x)$$

за произволно комплексно число λ.

На основание на предните дефиниции получаваме лесно, че функциите от първи род ще имат вида

$$f(x)=a_1\zeta_1+a_2\zeta_2+\ldots+a_n\zeta_n,$$

гдето $\zeta_1, \zeta_2, \ldots, \zeta_n$ са координатите на вектора x при базис $e_1, e_2, \ldots, e_n$ и $a_k=f(e_k)$, $1 \leqq k \leqq n$, а функциите от втори род ще имат вида

$$f(x)=b_1\bar{\zeta}_1+b_2\bar{\zeta}_2+\ldots+b_n\bar{\zeta}_n.$$

2. **Комплексно евклидово пространство.** Едно линейно комплексно пространство се нарича комплексно евклидово пространство, ако в него е въведено скаларно произведение, т. е. на всеки два вектора x, y, които могат да бъдат и равни, съответствува комплексно число (x, y), което удовлетворява следните аксиоми:

1. $(y, x)=\overline{(x, y)}$, като с $\bar{\alpha}$ се означава, както е прието, спрегнатото число на комплексното число α.

2. $(\lambda x, y)=\lambda(y, x)$ за всяко комплексно число λ.

3. $(x_1+x_2, y)=(x_1, y)+(x_2, y)$.

4. Съгласно с 1 имаме $(x, x)=\overline{(x, x)}$, т. е. числото (x, x) е реално. Предполагаме повече, че $(x, x) \geqq 0$ и само тогава имаме равенство, когато векторът x е нулев.

От 1, 2 и 3 следват равенствата

$$(x,\ \lambda y)=\bar{\lambda}(x,\ y),$$

$$(x,\ y_1+y_2)=(x,\ y_1)+(x,\ y_2).$$

Действително имаме

$$(x,\ \lambda y)=\overline{(\lambda y,\ x)}=\bar{\lambda}\,\overline{(y,\ x)}=\bar{\lambda}(x,\ y),$$

$$(x,\ y_1+y_2)=\overline{(y_1+y_2,\ x)}=\overline{(y_1,\ x)}+\overline{(y_2,\ x)}=(x,\ y_1)+(x,\ y_2).$$

Под дължина на вектора x, която означаваме с $|x|$, разбираме числото $\sqrt{(x,\ x)}$. Два вектора x и y се наричат ортогонални, ако скаларното им произведение $(x,\ y)$ е равно на нула.

Ще разгледаме някои примери на комплексно евклидово пространство. Под вектор x нека разбираме една система от n комплексни числа

$$x=(a_1,\ a_2,\ \ldots,\ a_n).$$

Очевидно множеството на всички такива вектори образува комплексно линейно пространство от n измерение. Ако за два кои да са вектора

$$x=(a_1,\ a_2,\ \ldots,\ a_n),\ y=(b_1,\ b_2,\ \ldots,\ b_n)$$

дефинираме скаларно произведение с формулата

$$(x,\ y)=a_1\bar{b}_1+a_2\bar{b}_2+\ \ldots\ +a_n\bar{b}_n,$$

то лесно се проверява валидността на аксиомите 1 до 4 за това произведение. Така получаваме комплексно евклидово пространство, като дължината на вектора x ще се дава с

$$|x|=\sqrt{|a_1|^2+|a_2|^2+\ \ldots\ +|a_n|^2}\ .$$

Множеството от всички непрекъснати функции $x(t)$ на променливото t в сегмента $[\alpha,\ \beta]$, вземащи комплексни стойности по този сегмент, при скаларно произведение, дефинирано с

$$(x(t),\ y(t))=\int_{\alpha}^{\beta}x(t)\bar{y}(t)\,dt,$$

е комплексно евклидово пространство.

Ако $e_1,\ e_2,\ \ldots,\ e_n$ са n линейно независими вектори в едно n-мерно комплексно евклидово пространство, то, както преди, можем от тях да получим с ортогонализация n ортогонални вектора, които са също линейно независими и следователно образуват базис на това пространство. Същите вектори могат да се нормират по употребен вече ничин. Нека $f_1,\ f_2,\ \ldots,\ f_n$ са n вектора, образуващи ортогонален и нормиран базис на пространството. Тогава скаларното произведение на два кои да са вектора $x,\ y$

$$x=\xi_1 f_1+\xi_2 f_2+\ \ldots\ +\xi_n f_n,\quad y=\eta_1 f_1+\eta_2 f_2+\ \ldots\ +\eta_n f_n$$

от R ще се дава с

$$(x,\ y)=\xi_1\bar{\eta}_1+\xi_2\bar{\eta}_2+\ldots+\xi_n\bar{\eta}_n.$$

Дължината $|x|$ на вектора x ще се дава с

$$|x|=\sqrt{|\xi_1|^2+|\xi_2|^2+\ldots+|\xi_n|^2}$$

и за координатите $\xi_1, \xi_2, \ldots, \xi_n$ на вектора x получаваме

$$\zeta_k=(x,\ f_k),\ k=1,\ 2,\ 3, \ldots, n.$$

Изоморфизмът на две комплексни линейни пространства се дефинира напълно подобно на случая на реални пространства и по същия начин, както се установява, че две кои да е комплексни афинни пространства от същото измерение са изоморфни и две кои да е комплексни евклидови пространства с еднакво измерение са изоморфни.

3. **Билинейни и квадратични форми.** Една функция $A(x; y)$ на векторите x и y се нарича билинейна форма, ако удовлетворява на условията :

1°
$$A(x_1+x_2;\ y)=A(x_1;\ y)+A(x_2;\ y),$$
$$A(\lambda x;\ y)=\lambda A(x;\ y).$$

2°
$$A(x;\ y_1+y_2)=A(x;\ y_1)+(x;\ y_2),$$
$$A(x;\ \lambda y)=\bar{\lambda}A(x;\ y),$$

т. е. тя е линейна функция от първи ред спрямо x (при фиксирано y) и такава от втори ред спрямо y (при фиксирано x).

Например скаларното произведение $(x,\ y)$ е билинейна форма. Нека

$$e_1,\ e_2, \ldots, e_n$$

е произволен базис на n-мерното комплексно пространство R. За векторите x и y ще имаме

$$x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n,\ y=\eta_1 e_1+\eta_2 e_2+\ldots+\eta_n e_n.$$

Тогава на основание на свойствата 1 и 2 получаваме

$$A(x;\ y)=A(\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n;\ \eta_1 e_1+\eta_2 e_2+\ldots+\eta_n e_n)=$$
$$=\sum_{i,k=1}^{n} A(e_i;\ e_k)\,\xi_i\bar{\eta}_k=\sum_{i,k=1}^{n} a_{ik}\,\xi_i\bar{\eta}_k,$$

като a_{ik} означават числата $A(e_i;\ e_k)$.

Ако в билинейната форма $A(x;\ y)$ поставим $x=y$, то получената функция се нарича квадратична форма. Не е трудно да се види, че всяка билинейна форма се определя еднозначно от квадратичната си форма. Действително от равенствата

$$A(x+y;\ x+y)=A(x;\ x)+A(y;\ x)+A(x;\ y)+A(y;\ y),$$
$$A(x+iy;\ x+iy)=A(x;\ x)+iA(y;x)-iA(x;\ y)+A(y;\ y),$$

(1) $$A(x-y;\ x-y)=A(x;\ x)-A(y;\ x)-A(x;\ y)+A(y;\ y),$$
$$A(x-iy;\ x-iy)=A(x;\ x)-iA(y;\ x)+iA(x;\ y)+A(y;\ y)$$

с умножение последователно на 1, $+i$, -1, $-i$ и събиране получаваме

(2) $$A(x;\ y)=\frac{1}{4}\left\{A(x+y;\ x+y)+iA(x+iy;\ x+iy)-\right.$$
$$\left.-A(x-y;\ x-y)-iA(x-iy;\ x-iy)\right\}.$$

Билинейната форма $A(x;\ y)$ се нарича хермитова, ако удовлетворява условието

(3) $$A(x;\ y)=\overline{A(y;\ x)}.$$

Необходимо и достатъчно условие формата $A(x:\ y)$ да бъде хермитова се състои в изпълнение на равенствата

$$a_{ik}=\bar{a}_{ki}.$$

Действително от условието (3) следва

$$a_{ik}=A(e_i;\ e_k)=\overline{A(e_k;\ e_i)}=\bar{a}_{ki}.$$

Обратно, от $a_{ik}=\bar{a}_{ki}$ получаваме

$$A(x;\ y)=\sum a_{ik}\xi_i\bar{\eta}_k=\sum\bar{a}_{ki}\xi_i\bar{\eta}_k=\overline{\sum a_{ki}\bar{\xi}_i\eta_k}=\overline{A(y;\ x)}.$$

Ако при един базис имаме $a_{ik}=\bar{a}_{ki}$, то и при кой да е друг базис това условие ще бъде изпълнено, т. е. ако формата $A(x;\ y)$ е хермитова при един базис, то тя е такава и при кой да е друг базис.

Съответната квадратична форма на една билинейна хермитова форма се нарича също хермитова форма. Необходимо и достатъчно условие една квадратична форма да бъде хермитова се състои в това, че тя да има реална стойност за кой да е вектор x. Действително, ако $A(x;\ y)=\overline{A(y;\ x)}$, то имаме

$$A(x;\ x)=\overline{A(x;\ x)}.$$

Обратно, нека $A(x;\ x)$ е реално число за всеки вектор x. Но тогава от формулата (2) и от следната формула:

$$A(v;\ x)=\frac{1}{4}\left\{A(x+y;\ x+y)-iA(x+iy;\ x+iy)-\right.$$
$$\left.-A(x-y;\ x-y)+iA(x-iy;\ x-iy)\right\},$$

която получаваме от (2) със смяна на векторите x и y помежду им, се вижда, че $A(x;\ y)$ и $A(y;\ x)$ са спрегнати комплексни числа.

Теорията на тези форми разгледахме по-рано.

Нека $x_1, x_2, \ldots, x_m$ са произволни m вектора от пространството R. Детерминантата на Грам се дава с

$$\Delta = \begin{vmatrix} (x_1, x_1) & (x_1, x_2) \ldots (x_1, x_m) \\ (x_2, x_1) & (x_2, x_2) \ldots (x_2, x_m) \\ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ (x_m, x_1) & (x_m, x_2) \ldots (x_m, x_m) \end{vmatrix}.$$

По аналогичен път, както по-горе, ще установим, че тази детерминанта е винаги едно реално, неотрицателно число и само тогава е равна на нула, когато векторите $x_1, x_2, \ldots, x_m$ са линейно зависими.

Ако векторите $x_1, x_2, \ldots, x_m$ са линейно зависими, виждаме напълно аналогично на по-рано, че Δ е равна на нула. Нека сега същите вектори са линейно независими и да означим с $e_1, e_2, \ldots, e_m$ векторите, които се получават с ортогонализирането им. Тогава за векторите $x_1, x_2, \ldots, x_m$ ще имаме

$$x_p = a_{p1} e_1 + a_{p2} e_2 + \ldots + a_{pm} e_m, \; p = 1, 2, \ldots, m,$$

като детерминантата

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

ще бъде отлична от нула. Понеже за скаларните произведения (x_p, x_q) имаме

$$(x_p, x_q) = a_{p1} \bar{a}_{q1} + a_{p2} \bar{a}_{q2} + \ldots + a_{pm} \bar{a}_{qm},$$

то непосредствено се вижда, че детерминантата на Грам Δ ще е равна на произведението на детерминатата D с конюгованата ѝ детерминанта, т. е. с тази, която се получава от D със смяна на всичките ѝ елементи с конюгованите им комплексни числа. Но тогава от равенството

$$\Delta = D \bar{D}$$

следва, че $\Delta > 0$. При $m = 2$ получаваме неравенството на Коши—Буняковски:

$$|(x, y)|^2 \leqq (x, x)\ (y, y).$$

Ще разгледаме пространството K от всички полиноми

$$a_0 + a_1 z + \ldots + a_n z^n,$$

в които коефициентите $a_0, a_1, \ldots, a_n$ са произволни комплексни числа и $z = x + iy$ е произволно комплексно променливо. Нека D е крайна проста област на равнината (x, y). За скаларно произведение на два вектора $P(z)$ и $Q(z)$ от K ще приемем интеграла

$$(P,Q) = \int\int_D P(z)\, \overline{Q(z)}\, dx\, dy.$$

Очевидно полученото така скаларно произведение удовлетворява условията ѝ за него K ще бъде едно евклидово комплексно пространство. Да разгледаме системата вектори

$$1,\ z,\ z^2,\ldots,\ z^n,$$

които очевидно са линейно независими. Чрез ортогонализация ще получим редицата от полиноми:

$$P_0(z),\ P_1(z),\ P_2(z),\ldots,\ P_n(z),$$

като степента на $P_\nu(z)$ е точно равна на ν, които два по два ще бъдат ортогонални, т. е. ще имаме

$$\iint_D P_\mu(z)\overline{P_\nu(z)}\,dx\,dy=0,\ \mu\neq\nu,\ \mu,\nu=0,\ 1,2,\ldots,n.$$

От самия начин на ортогонализация се вижда, че с увеличение на n получените по-рано полиноми не ще се изменят. По този начин получаваме една редица от ортогонални полиноми:

$$P_0(z),\ P_1(z),\ P_2(z),\ P_3(z),\ldots,$$

всеки от които има степен, равна на индекса му. Тези полиноми до постоянен множител за всеки са еднозначно определени. С подходящо деление на положителни числа ние можем да ги нормираме, т. е. да предполагаме, че

$$\iint_D |P_n(z)|^2\,dx\,dy=1,\ n=0,\ 1,\ 2\ \ 3,\ldots$$

За всеки полином $f(z)$ от степен m ще имаме представянето

$$f(z)=c_0P_0(z)+c_1P_1(z)+\ldots+c_mP_m(z),$$

гдето

$$c_\nu=\iint_D f(z)\overline{P_\nu(z)}\,dx\,dy,\ \nu=0,\ 1,\ 2,\ldots,m.$$

Глава VII

Спрегнати линейни оператори

1. **Връзка между билинейни форми и линейни оператори.** В n-мерното евклидово комплексно пространство R всеки вектор x, както видяхме, може да се представи във формата

$$x=\xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n, \tag{1}$$

гдето $\xi_1,\ \xi_2,\ldots,\ \xi_n$ са координатите на вектора x и $e_1, e_2,\ldots, e_n$ е една нормирана ортогонална база. Ако A е един линеен оператор, то координатите на вектора Ax ще са равни на

$$\zeta_1=a_{11}\xi_1+a_{21}\xi_2+\ldots+a_{n1}\xi_n,$$

$$\zeta_2 = a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{n2}\xi_n,$$

(2) $$\ldots\ldots\ldots\ldots\ldots$$

$$\zeta_n = a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{nn}\xi_n,$$

гдето $\|a_{lk}\|$ е матрицата на оператора A. Ако y е произволен вектор от пространството R с координати $\eta_1, \eta_2, \ldots, \eta_n$, т. е.

$$y = \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n,$$

то скаларното произведение (z, y) на векторите $z = Ax$ и y се дава с

$$(z, y) = \zeta_1 \overline{\eta}_1 + \zeta_2 \overline{\eta}_2 + \ldots + \zeta_n \overline{\eta}_n.$$

Като заместим тук стойностите на координатите $\zeta_1, \zeta_2, \ldots, \zeta_n$ от (2), виждаме, че това произведение има формата

$$\begin{aligned}(z, y) = {} & a_{11}\xi_1\overline{\eta}_1 + a_{12}\xi_1\overline{\eta}_2 + \ldots + a_{1n}\xi_1\overline{\eta}_n + \\ & + a_{21}\xi_2\overline{\eta}_1 + a_{22}\xi_2\overline{\eta}_2 + \ldots + a_{2n}\xi_2\overline{\eta}_n + \\ & + a_{n1}\xi_n\overline{\eta}_1 + a_{n2}\xi_n\overline{\eta}_2 + \ldots + a_{nn}\xi_n\overline{\eta}_n,\end{aligned}$$

т. е. представлява една билинейна форма $A(x; y)$ на векторите x и y и ще имаме

$$(z, y) = A(x; y) = (Ax, y).$$

Така получаваме, че на всеки линеен оператор A отговаря билинейна форма $A(x; y)$ със свойството

$$A(x; y) = (Ax, y).$$

Обратно, нека $A(x; y)$ е една билинейна форма на векторите x, y:

$$A(x; y) = \sum_{i,k=1}^{n} a_{ik}\xi_i\overline{\eta}_k.$$

Тази форма можем да пишем така:

$$\begin{aligned}A(x; y) = {} & (a_{11}\xi_1 + a_{21}\xi_2 + \ldots + a_{n1}\xi_n)\overline{\eta}_1 + \\ & + (a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{n2}\xi_n)\overline{\eta}_2 + \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ & + (a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{nn}\xi_n)\overline{\eta}_n.\end{aligned}$$

Като въведем вектора z с координати $\zeta_1, \zeta_2, \ldots, \zeta_n$, дефинирани с равенствата (2), получаваме, че

$$A(x; y) = \zeta_1\overline{\eta}_1 + \zeta_2\overline{\eta}_2 + \ldots + \zeta_n\overline{\eta}_n = (z, y),$$

гдето $z=A(x)$ е линеен оператор. Виждаме така, че на всяка билинейна форма $A(x;\ y)$ отговаря такъв линеен оператор A, за който да имаме

$$(Ax,\ y)=A(x;\ y).$$

Лесно е да видим, че операторът A се определя еднозначно от формата $A(x;\ y)$. Нека допуснем противното, т. е. че имаме

$$A(x;\ y)=(Ax,\ y),$$

$$A(x;\ y)=(Bx,\ y).$$

Но тогава от равенството $(Ax,\ y)=(Bx,\ y)$ получаваме

$$(Ax-Bx,\ y)=0,$$

което равенство трябва да бъде вярно за всеки вектор x и y. Следователно $Ax-Bx=0$ за всяко x, т. е. $A=B$. Така установихме теоремата:

1. Равенството

$$A(x;\ y)=(Ax,\ y) \tag{3}$$

установява взаимно еднозначно съответствие между билинейната форма $A(x;\ y)$ и линейния оператор A.

Нека отбележим, че на основание на формулата (3) еднозначното съответствие, както установихме, не зависи от избора на базиса.

Да направим сега друга преработка на билинейната форма $A(x;\ y)$, като групираме членовете ѝ спрямо координатите $\xi_1,\ \xi_2,\ \ldots,\ \xi_n$.

Както преди, получаваме

$$\begin{aligned}A(x;\ y)=&\xi_1(a_{11}\overline{\eta}_1+a_{12}\overline{\eta}_2+\ \ldots\ +a_{1n}\overline{\eta}_n)+\\&+\xi_2(a_{21}\overline{\eta}_1+a_{22}\overline{\eta}_2+\ \ldots\ +a_{2n}\overline{\eta}_n)+\\&\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\\&+\xi_n(a_{n1}\overline{\eta}_1+a_{n2}\overline{\eta}_2+\ \ldots\ +a_{nn}\overline{\eta}_n)=\\=&\xi_1\overline{\tau}_1+\xi_2\overline{\tau}_2+\ \ldots\ +\xi_n\overline{\tau}_n=(x,\tau)=(x,\ A^*y),\end{aligned}$$

гдето

$$A^*y=\tau_1 e_1+\tau_2 e_2+\ \ldots\ +\tau_n e_n.$$

Матрицата на оператора A^* се получава от матрицата на оператора A с транспониране и заместване на елементите ѝ с конюгованите им числа. Така виждаме, че на всеки оператор A от евклидовото пространство отговаря оператор A^*, и то еднозначно притежаващ свойството

$$(Ax,\ y)=(x,\ A^*y).$$

Операторът A^* се нарича спрегнат на оператора A. Матрицата на оператора A^* се получава от тази на оператора A с преминаване към транспонираната матрица и заместване на елементите ѝ със спрегнатите комплексни числа.

Еднозначността следва непосредствено от еднозначното съответствие на билинейната функция и оператора.

Преминаването от преобразуването A към A^* се подчинява при умножение и събиране на операторите на следните правила:

$$(AB)^* = B^* A^*.$$

$$(A^*)^* = A.$$

$$(A+B)^* = A^* + B^*.$$

$$(\lambda A)^* = \bar{\lambda} A^*.$$

$$E^* = E.$$

Примерно ще установим първото от тези равенства. Имаме

$$(ABx, y) = (Bx,\ A^* y) = (x, B^* A^* y).$$

От определението на спрегнат оператор имаме

$$(ABx, y) = (x, (AB)^* y).$$

Понеже операторът се определя еднозначно от съответната билинейна форма, то от горните две равенства получаваме

$$(AB)^* = B^* A^*,$$

което трябваше да се докаже.

2. **Спрегнат на себе си оператор.** Линейният оператор A се нарича спрегнат на себе си, ако $A^* = A$. Лесно се вижда, че необходимо и достатъчно условие, щото линейният оператор A да бъде спрегнат на себе си, се състои в това, че билинейната форма $A(x; y)$ да бъде хермитова. Действително условието операторът да бъде спрегнат на себе си се изразява с равенството

$$(Ax,\ y) = (x,\ Ay). \tag{4}$$

Формата $A(x; y)$ е хермитова, ако

$$(Ax; y) = \overline{(Ay, x)}. \tag{5}$$

Очевидно е, че равенствата (4) и (5) са еквивалентни.

Ще установим сега някои теореми за спрегнатите на себе си преобразувания, които имат приложения в разни въпроси от математиката.

2. Собствените значения на спрегнатите на себе си преобразувания са реални.

Нека x е собственият вектор на спрегнатото на себе си преобразувание, A и λ — собственото значение. Ще имаме значи

$$Ax = \lambda x,\ x \neq 0.$$

Понеже $A^* = A$, то

$$(Ax, x) = (x, A^* x) = (x, Ax).$$

Като вземем пред вид предното равенство, получаваме оттук

$$(\lambda x, x)=(x, \lambda x)$$

или $\lambda (x, x)=\bar{\lambda} (x, x)$. Понеже $(x, x)\neq 0$, то $\lambda=\bar{\lambda}$, т. е. числото λ е реално.

3. Собствените вектори, отговарящи на различни собствени значения, са ортогонални помежду си.

Нека e_1, e_2 са собствени функции на оператора A и λ_1, λ_2 са съответните собствени значения, като $\lambda_1\neq\lambda_2$. Като вземем пред вид равенствата

$$Ae_1=\lambda_1 e_1, \quad Ae_2=\lambda_2 e_2,$$

получаваме

$$(Ae_1, e_2)=(e_1, A^* e_2)=(e_1, Ae_2)=(e_1, \lambda_2 e_2)=\lambda_2 (e_1, e_2),$$

т. е.

$$\lambda_1 (e_1, e_2)=\lambda_2 (e_1, e_2)$$

или

$$(\lambda_1-\lambda_2) (e_1, e_2)=0.$$

Понеже $\lambda_1\neq\lambda_2$, то следва, че $(e_1, e_2)=0$, с което теоремата е установена.

Ще докажем сега следното помощно предложение:

4. Нека A е спрегнато на себе си линейно преобразуване в n-мерно евклидово пространство, R и e е неговият собствен вектор. Множеството от векторите x, ортогонални на e, е $(n-1)$-мерно подпространство, инвариантно относно преобразуването A.

Знаем, че множеството R_1 от векторите x, които са ортогонални на e, образува $(n-1)$-мерно подпространство. Не е трудно да се види, че R_1 е инвариантно спрямо e. Действително нека $x\in R_1$, т. е. $(x, e)=0$. Понеже преобразуването A е спрегнато на себе си, ще имаме

$$(Ax, e)=(x, A^* e)=(x, Ae)=(x, \lambda e)=\lambda (x, e)=0.$$

Равенството $(Ax, e)=0$ показва, че векторът Ax принадлежи също на R_1.

На основание на предните предложения ще установим следната теорема:

5. Нека A е спрегнато на себе си линейно преобразуване в едно n-мерно евклидово пространство R. Тогава съществуват n взаимно ортогонални собствени вектори на преобразуването A, на които съответствуващите собствени значения са реални.

Известно е, че в R съществува поне един собствен вектор α_1 на преобразуването A. По предното предложение множеството вектори x, ортогонални на α_1, образуват подпространство R_1 на R, инвариантно спрямо него. На същото основание в R_1 съществува собствен вектор α_2 на преобразуването A. Векторите от R_1, които са ортогонални на

α_2, образуват инвариантно подпространство R_2 на R_1 и A ще има поне един собствен вектор α_3 и т. н. Така получаваме n взаимно ортогонални собствени вектори $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Нека A е спрегнато на себе си линейно преобразуване в едно n-мерно евклидово пространство R. Тогава съществува ортогонален базис, в който матрицата на A е диагонална и с реални елементи. В сила е и обратното твърдение.

За базис на пространството R да изберем получените в предната теорема вектори $\alpha_1, \alpha_2, \ldots, \alpha_n$, които можем да считаме и за нормирани, т. е. с дължини, равни на единица. Ако $\lambda_1, \lambda_2, \ldots, \lambda_n$ са съответните собствени значения, то имаме

$$A\alpha_1 = \lambda_1 \alpha_1,$$
$$A\alpha_2 = \lambda_2 \alpha_2,$$
$$\ldots\ldots$$
$$A\alpha_n = \lambda_n \alpha_n$$

Следователно матрицата на преобразуването A в този базис ще има диагоналната форма

$$\begin{Vmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ 0 & \lambda_2 & 0 \ldots 0 \\ \ldots & \ldots & \ldots \\ 0 & 0 & 0 \ldots \lambda_n \end{Vmatrix}.$$

Обратно, нека матрицата на преобразуването A има предната форма с реални числа λ_i. Видяхме, че при ортогонален нормиран базис матрицата на спрегнатата трансформация A^* се получава с транспониране и заместване на елементите с конюгованите им числа. Но очевидно матрицата остава същата, т. е. преобразуванията A и A^* имат една и съща матрица. Това означава, че линейните преобразувания A и A^* са идентични, с което изказаната теорема е установена напълно.

От теорема 5 следва следният резултат:
Корените на уравнението

$$\begin{vmatrix} a_{11}-x & a_{12} & a_{13} \ldots a_{1n} \\ a_{21} & a_{22}-x & a_{23} \ldots a_{2n} \\ \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} \ldots a_{nn}-x \end{vmatrix} = 0, \tag{6}$$

където a_{ik} са произволни комплексни числа, като $a_{ik} = \bar{a}_{ki}$, са всичките реални.

Предните теореми са в сила и за реалните евклидови пространства. Нека A е едно линейно преобразуване в n-мерното реално евклидово пространство R. Тогава на него отговаря еднозначно една билинейна форма $A(x; y)$ със свойството

$$A(x; y) = (Ax; y).$$

За да установим това, достатъчно е в извеждането на теорема 1 да заместим числата $\overline{\eta}_1, \overline{\eta}_2, \ldots, \overline{\eta}_n$ с $\eta_1, \eta_2, \ldots, \eta_n$, понеже тези числа са реални. Също така се вижда, че на всяко линейно преобразуване A отговаря линейно преобразуване A^*, и то еднозначно, което притежава свойството

$$(Ax, y)=(x, A^*y).$$

Преобразуването A^* се нарича спрегнато на A. Ако $A^*=A$, то преобразуването A се нарича спрегнато на себе си. Необходимо и достатъчно условие, за да имаме $A^*=A$, се състои в това, че формата $A(x; y)$ да бъде симетрична. Собствените значения на оператора A ще удовлетворяват уравнение от вида (6) и следователно са всичките реални. Но тогава и съществуването на собствените вектори от пространството R е установено, като се следва указаният път. По-нататък доказателствата на следващите теореми 2, 3, 4 не претърпяват някакво изменение, т. е. тези теореми остават в сила и за реалните евклидови пространства.

3. **Унитарно преобразуване.** Преобразуването U се нарича унитарно, ако

$$UU^*=U^*U=E, \tag{7}$$

т. е. $U^*=U^{-1}$. Това линейно преобразуване е свързано с просто геометрично свойство. Именно имаме следното предложение:

6. Всяко унитарно преобразуване в n-мерното евклидово пространство R запазва скаларното произведение, т. е. за всеки два вектора x и y от R имаме

$$(Ux, Uy)=(x, y). \tag{8}$$

Обратно, всяко линейно преобразуване, което запазва скаларното произведение, е унитарно.

Действително, ако $U^*U=E$, то

$$(Ux, Uy)=(x, U^*Uy)=(x, y).$$

Обратно, ако за всеки вектор x и y имаме

$$(Ux, Uy)=(x, y),$$

то

$$(U^*Ux, y)=(x, y)$$

или

$$(U^*Ux, y)=(Ex, y).$$

От равенството на двете билинейни форми в това равенство следва, че $U^*U=E$, т. е. U е унитарно преобразуване.

Ако $x=y$, то имаме $(Ux, Ux)=(x, x)$, т. е. унитарното преобразуване не променя дължините на векторите от R.

Нека $e_1, e_2, \ldots, e_n$ е един ортогонален и нормиран базис на пространството R. Матрицата на U в този базис да бъде

$$\left\| \begin{matrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1} & a_{n2} \ldots a_{nn} \end{matrix} \right\|,$$

т. е.

$$U(e_i) = a_{1i} e_1 + a_{2i} e_2 + \ldots + a_{ni} e_n, \quad 1 \leqq i \leqq n.$$

Съгласно с равенството (8) ще имаме

$$(9) \qquad (Ue_p, Ue_q) = (e_p, e_p) = \begin{cases} 1, & p = q, \\ 0, & p \neq q. \end{cases}$$

Предните равенства могат да се напишат така:

$$(10) \qquad \sum_{i=1}^{n} a_{ip} \bar{a}_{iq} = \begin{cases} 1, & p = q, \\ 0, & p \neq q. \end{cases}$$

Обратно, лесно се вижда, че от равенствата (10) или еквивалентните им равенства (9) следва, че $(Ux, Uy) = (x, y)$, т. е. U е унитарно преобразуване. Също от последните равенства следва непосредствено предложението: необходимо и достатъчно условие, щото линейното преобразуване U да бъде унитарно, се състои в това, че U да трансформира някой нормиран ортогонален базис в също нормиран ортогонален базис.

Ще установим подобни теореми на предните за унитарните преобразувания. Собствените значения на унитарните преобразувания имат модули, равни на единица.

Ако x е собственият вектор на унитарното преобразуване, U и λ е съответното собствено значение, то имаме

$$Ux = \lambda x, \quad x \neq 0.$$

Но тогава ще имаме

$$(x, x) = (Ux, Ux) = (\lambda x, \lambda x) = |\lambda|^2 (x, x),$$

от които равенства следва, $|\lambda|^2 = 1$, т. е. $|\lambda| = 1$.

Нека U е унитарно линейно преобразуване в n-мерното пространство R и α — собственият му вектор. Тогава $(n-1)$-мерното подпространство R_1, състоящо се от ортогоналните на α_1 вектори, е инвариантно относно R.

Нека $x \in R_1$, т. е. $(x, \alpha) = 0$. Имаме

$$(Ux, U\alpha) = (U^* Ux, \alpha) = (x, \alpha) = 0.$$

Понеже $U\alpha = \lambda\alpha$, то $\bar{\lambda}(Ux, e) = 0$. По предното предложение $\lambda \neq 0$ и следователно $(Ux, \alpha) = 0$, т. е. Ux принадлежи на R_1.

Нека U е унитарно линейно преобразуване в n-мерното евклидово пространство R. Тогава съществуват n взаимно ортогонални собствени вектори на преобразуването A и съответстващите им собствени значения по модул са равни на единица.

Преобразуването U като линейно притежава поне един собствен вектор α_1 и R. Пространството R_1, състоящо се от всички вектори от R, ортогонални на вектора α_1, е $(n-1)$-мерно подпространство R_1 на R_2, инвариантно относно A. На това основание в R_1 съществува поне един собствен вектор α_2 на преобразуването A. Нека R_2 е инвариантното подпространство, съставено от всичките вектори, принадлежащи на R_1 и ортогонални на α_2. В R_2 има собствен вектор α_3 на преобразуването A и т. н. Продължавайки така, ние намираме n собствени вектора на преобразуването A, които са два по два ортогонални. Съответните им собствени значения са с модули, равни на единица. Ако означим с $\lambda_1, \lambda_2, \ldots, \lambda_n$ собствените значения, то ще имаме

$$U\alpha_1 = \lambda_1 \alpha_1,$$
$$U\alpha_2 = \lambda_2 \alpha_2,$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$U\alpha_n = \lambda_n \alpha_n.$$

Оттук веднага се вижда, че матрицата на преобразуването A ще има формата

$$(11) \qquad \begin{Vmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ 0 & \lambda_2 & 0 \ldots 0 \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ 0 & 0 & 0 \ldots \lambda_n \end{Vmatrix}.$$

Можем да считаме, че векторите $\alpha_1, \alpha_2, \ldots, \alpha_n$ са нормирани. С предните разглеждания установихме следното предложение:

За всяко унитарно преобразуване U в n-мерното евклидово пространство R съществува нормиран ортогонален базис, при който матрицата на преобразуването A е диагонална, т. е. има формата (11). При това числата $\alpha_1, \alpha_2, \ldots, \alpha_n$ имат модули, равни на единица.

4. Екстремално свойство на собствените значения. Нека A е линейно преобразуване в n-мерното евклидово реално пространство R, което е спрегнато на себе си. Да намерим екстремните стойности на съответстващата квадратична форма $A(x; x) = (Ax; x)$ върху единичната сфера $(x, x) = 1$ в пространството R. Ако с $\xi_1, \xi_2, \ldots, \xi_n$ означим координатите на вектора x при нормиран ортогонален базис от R, то условието $(x, x) = 1$ е следното:

$$(12) \qquad \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 = 1.$$

Следователно трябва да намерим екстремните стойности на формата

$$A(x;\ x)=\sum_{i,k=1}^{n} a_{ik}\,\xi_i\,\xi_k$$

при предположение, че променливите ξ_i $1\leqq i\leqq n$ удовлетворяват на условието (2). По метода на Лагранж трябва да приравним на нула частните производни от първи ред на функцията

$$F=\sum_{i,k=1}^{n} a_{ik}\,\xi_i\,\xi_k-\lambda\sum_{i=1}^{n}\xi_i^2.$$

Така получаваме уравненията

$$2\sum_{k=1}^{n} a_{ik}\,\xi_k-2\,\lambda\,\xi_i=0,\quad 1\leqq i\leqq n$$

или

$$\begin{aligned}
&(a_{11}-\lambda)\,\xi_1+\ a_{12}\,\xi_2+\ \ldots\ +a_{1n}\,\xi_n=0,\\
&a_{21}\,\xi_1+(a_{22}-\lambda)\,\xi_2+\ \ldots\ +a_{2n}\,\xi_n=0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{n1}\,\xi_1+\ \ a_{n2}\,\xi_2+\ \ldots\ +(a_{nn}-\lambda)\,\xi_n=0.
\end{aligned}$$

Но това уравнение е точно уравнението за определяне на собствените значения на оператора A и на собствените вектори за същия оператор. Като означим тогава с $\lambda_1, \lambda_2, \ldots, \lambda_n$ въпросните собствени значения и с $e_1, e_2, \ldots, e_n$ съответните собствени вектори, ще имаме $Ae_p=\lambda_p e_p$, отгдето получаваме $A(e_p;\ e_p)=(Ae_p;\ e_p)=\lambda_p(e_p,\ e_p)=\lambda_p$.

Следователно най-малкото собствено значение на оператора A е равно на минимума на квадратичната форма по единичната сфера $(x,\ x)=1$ и най-голямото собствено значение на същия оператор е равно на максимума на квадратната форма по същата сфера.

Да означим с λ_1 най-малкото собствено значение, т. е. минимума на формата $A(x;\ x)$, и с e_1 — съответния собствен вектор. За да намерим следващото собствено значение на оператора A, нека разглеждаме всички вектори от пространството R, които са ортогонални на вектора e_1. Както видяхме, тези вектори образуват подпространство R_1 от $(n-1)$ измерение, което е инвариантно относно A. Като търсим минимума на формата $A(x;\ x)$ при условието $(x,\ x)=1$ в това подпространство, намираме нов собствен вектор e_2 и собствено значение λ_2. Понеже минимумът на $A(x;\ x)$ в подпространството R_1 не може очевидно да надминава минимума на същата функция в пространството R, то следва, че $\lambda_2\geqq\lambda_1$. Векторите от R_1, които са ортогонални на e_2, ще образуват ново инвариантно спрямо A подпространство R_3 на R, т. е. подпространство на R, съставено от всички вектори от R, които са ортогонални на векторите e_1 и e_2. Следващото собствено значение и собствен вектор ще намерим, като решим същата задача за това $(n-2)$-мерно подпространство и т. н.

Глава VIII

Каноничен вид на линейните преобразувания

1. Нормална форма на линейните преобразувания. Ще разгледаме сега въпроса за намиране на най-простата форма на дадено линейно преобразуване. Предварително ще изясним някои специални случаи на такива преобразувания, които ще ни трябват по-нататък при излагане на общия резултат. Нека A е линеен оператор, който притежава следното свойство: съществуват p вектора e_i от n-мерното пространство R, за които имаме

$$(1) \qquad Ae_1=\lambda_1 e_1,\ Ae_2=e_1+\lambda_1 e_2,\ \ldots,\ Ae_p=e_{p-1}+\lambda_1 e_p,$$

гдето λ_1 е комплексно число. Предполагаме освен това, че векторите $e_1, e_2, \ldots, e_p$ са линейно независими. От първото равенство на (1) следва, че λ_1 е собствено значение и e_1 е собствен вектор на оператора A. Ще докажем, че в подпространството R_1, създадено от векторите $e_1, e_2, \ldots, e_p$, операторът A няма други собствени вектори освен векторите ke_1, гдето k е произволно число. Нека допуснем, че векторът e от R_1 е собствен на оператора A. Понеже e принадлежи на R_1, то ще имаме

$$e=b_1 e_1+b_2 e_2+\ldots+b_p e_p,$$

гдето поне едно от числата $b_1, b_2, \ldots, b_p$ е отлично от нула. Тогава за някое число λ трябва да имаме

$$A(e)=A(b_1 e_1+b_2 e_2+\ldots+b_p e_p)=\lambda(b_1 e_1+b_2 e_2+\ldots+b_p e_p).$$

Като заместим стойностите на векторите $A(e_1), A(e_2), \ldots, A(e_p)$ от (1), ще получим равенството

$$b_1\lambda_1 e_1+b_2(e_1+\lambda_1 e_2)+\ldots+b_p(e_{p-1}+\lambda_1 e_p)=$$
$$=\lambda b_1 e_1+\lambda b_2 e_2+\ldots.+\lambda b_p e_p.$$

Следствие линейната независимост на векторите $e_1, e_2, \ldots, e_p$ коефициентите им в двете части на горното равенство трябва да бъдат равни, т. е.

$$\begin{aligned}
b_1\lambda_1+b_2&=\lambda b_1,\\
b_2\lambda_1+b_3&=\lambda b_2,\\
&\cdots\cdots\cdots\\
b_{p-1}\lambda_1+b_p&=\lambda b_{p-1},\\
b_p\lambda_1\qquad\quad&=\lambda b_p.
\end{aligned}$$

Да предположим, че $\lambda \neq \lambda_1$. Тогава от последното равенство следва, че $b_p=0$ и последователно от останалите равенства получаваме $b_{p-1}=0$, $b_{p-2}=0, \ldots, b_1=0$. Следователно $\lambda=\lambda_1$. Но тогава от първото уравнение получаваме $b_2=0$, от второто $b_3=0$ и т. н. до $b_p=0$. Но тогава собственият вектор e е равен на $b_1 e_1$ и предложението е установено.

Матрицата на преобразуването A относно векторите $e_1, e_2, \ldots, e_p$ ще е равна на

$$(2)\qquad \left\| \begin{array}{ccccc} \lambda_1 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_1 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_1 1 \\ 0 & 0 & 0 & \ldots & 0\lambda_1 \end{array} \right\|.$$

Матриците от тази форма се наричат клетки. Ние ще видим, че всяко линейно преобразуване на пространството R може да се сведе в такава форма, че матрицата му да има формата

$$(3)\qquad \left\| \begin{array}{ccccc} A_1 & & & & \\ & A_2 & & & \\ & & A_3 & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & \cdot \\ & & & & A_m \end{array} \right\|,$$

гдето $A_1, A_2, A_3, \ldots, A_m$ са клетки и елементите, които не принадлежат на тях, са равни на нула. Именно имаме следната теорема на Жордан:

Нека в линейното комплексно n-мерно пространство е зададен линейният оператор A. Да предположим, че операторът A има m линейно независими собствени вектори $e_1, g_1, \ldots, t_1$, съответстващи на собствените значения $\lambda_1, \lambda_2, \ldots, \lambda_m$. Тогава съществува базис, състоящ се от m групи от вектори:

$$e_1, e_2, \ldots, e_p;\ g_1, g_2, \ldots, g_q;\ \ldots;\ t_1, t_2, \ldots, t_s,$$

за които операторът A има следния вид:

$$(4)\qquad \begin{array}{l} Ae_1 = \lambda_1 e_1,\ Ae_2 = e_1 + \lambda_1 e_2, \ldots, Ae_p = e_{p-1} + \lambda_1 e_p, \\ Ag_1 = \lambda_2 g_1,\ Ag_2 = g_1 + \lambda_2 g_2, \ldots, Ag_q = g_{q-1} + \lambda_2 g_q, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ At_1 = \lambda_m t_1,\ At_2 = t_1 + \lambda_m t_2, \ldots, At_s = t_{s-1} + \lambda_m t_s. \end{array}$$

Оттук се вижда, че матрицата на оператора H има точно формата (3), гдето клетките A_i, $i = 1, 2, 3, \ldots, m$, се дават с

$$A_i = \left\| \begin{array}{ccccc} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 0\lambda_i \end{array} \right\|.$$

Ще отбележим, че с матрицата A алгебричните операции се прилагат доста просто. Така лесно се вижда, че

$$A^2=\begin{Vmatrix} A_1^2 & & & & \\ & A_2^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & A_m^2 \end{Vmatrix}$$

и въобще за кое да е цяло положително число k ще имаме

$$A^k=\begin{Vmatrix} A_1^k & & & & \\ & A_2^k & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & A_m^k \end{Vmatrix},$$

гдето неписаните елементи са нули. За произволен полином

$$P(x)=a_0+a_1x+\ldots+a_nx^k$$

ще имаме

$$P(A)=\begin{Vmatrix} P(A_1) & & & & \\ & P(A_2) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & P(A_m) \end{Vmatrix}.$$

2. **Доказателство**[1] **на теоремата на Жордан.** При доказване на теоремата на Жордан ще използуваме метода на индукцията от n към $n+1$, т. е. предполагаме, че теоремата е установена за n-мерни пространства и ще покажем, че тя е вярна и за $(n+1)$-мерни пространства. Предварително ще докажем едно помощно предложение:

За всяко линейно преобразуване A в едно n-мерно пространство R съществува поне едно $(n-1)$-мерно инвариантно подпространство R'.

Спрегнатият оператор A^* има поне един собствен вектор e:

$$A^*e=\lambda e.$$

Множеството вектори x от пространството R, които са ортогонални на вектора e, т. е. за което $(x, e)=0$, образува $(n-1)$-мерно подпростран-

---

[1] Изложеното доказателство принадлежи на И. Г. Петровский и И. М. Гелбфанд (И. М. Гелбфанд — Лекции по линейной алгебре, стр. 162).

ство R' на R. Трябва да се докаже, че R' е инвариантно спрямо A. Действително, ако x е произволен вектор от R', то имаме $(x, e)=0$. Но тогава следва, че

$$(Ax, e)=(x, A^* e)=(x, \lambda e)=0,$$

т. е. $(Ax, e)=0$. Това равенство показва, че векторът Ax е ортогонален на e и следователно принадлежи на R'.

Нека A е произволно линейно преобразуване в едно $(n+1)$-мерно пространство R. По предложението съществува подпространство R' на R, което е инвариантно относно A. За пространството R' предполагаме, че теоремата е установена. Тогава в R' има базис, при който линейното преобразуване A има нормална форма, която да бъде дадена с равенствата (4). В пространството R очевидно има вектори, които не зависят линейно от векторите

$$(5) \qquad e_1, e_2, \ldots, e_p;\ g_1, g_2, \ldots, g_q;\ \ldots;\ t_1, t_2, \ldots, t_s.$$

Нека e е един такъв вектор. Като прибавим e към векторите (5), получаваме $n+1$ вектора, които ще бъдат базис на пространството R. Тогава векторът Ae ще има формата

$$(6) \quad Ae=\alpha_1 e_1+\ldots+\alpha_p e_p+\beta_1 g_1+\ldots+\beta_q g_q+\ldots+\varepsilon_1 t_1+\ldots+\varepsilon_s t_s+\tau e,$$

където $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \ldots, \varepsilon_1, \ldots, \varepsilon_s, \tau$ са комплексни числа. Можем да предположим, че $\tau=0$. Действително, ако A има нормална форма в някой базис, то и операторът $A-\tau E$ има нормална форма в същия базис. Но тогава, ако $\tau\neq 0$, вместо оператора A разглеждаме оператора $A-\tau E$, т. е. прехвърляме τe в лявата част на (6), като оператора $A-\tau E$ означаваме за простота пак с A. Следователно ще имаме

$$(7) \quad Ae=\alpha_1 e_1+\ldots+\alpha_p e_p+\beta_1 g_1+\ldots+\beta_q g_q+\ldots+\varepsilon_1 t_1+\ldots+\varepsilon_s t_s.$$

Да въведем сега вместо вектора e друг вектор e', и то по такъв начин избран, че Ae' да добие по възможност най-прост вид. Векторът e' избираме така:

$$(8) \quad e'=e-(k_1 e_1+\ldots+k_p e_p+\mu_1 g_1+\ldots+\mu_q g_q+\ldots+\omega_1 t_1+\ldots+\omega_s t_s).$$

За Ae' на основание на (6) ще имаме

$$Ae'=Ae-A(k_1 e_1+\ldots+k_p e_p)-A(\mu_1 g_1+\ldots+\mu_q g_q)-\ldots-$$

$$(9) \quad -A(\omega_1 t_1+\ldots+\omega_s t_s)=\alpha_1 e_1+\ldots+\alpha_p e_p+\beta_1 g_1+\ldots+\beta_q g_q+\ldots+$$

$$+\varepsilon_1 t_1+\ldots+\varepsilon_s t_s-A(k_1 e_1+\ldots+k_p e_p)-A(\mu_1 g_1+\ldots+$$

$$+\mu_q g_q)-\ldots-A(\omega_1 t_1+\ldots+\omega_s t_s).$$

Коефициентите $k_1, \ldots, k_p, \mu_1, \ldots, \mu_q, \ldots, \omega_1, \ldots, \omega_s$ са произволни числа. Ние сега ще ги подберем така, че в дясната част на равенството (9) да останат по възможност най-малко събираеми. За всяка група вектори от базиса, в който A има нормална форма, отговаря по едно собствено

значение на този оператор. Нека допуснем отначало, че всичките собствени значения

$$\lambda_1,\ \lambda_2,\ldots,\lambda_m$$

са отлични от нула. Ще покажем, че можем така да подберем числата $k_1,\ldots,k_p,\mu_1,\ldots,\mu_q,\ldots,\omega_1,\ldots,\omega_s$, че коефициентите на векторите $e_1,\ldots,e_p,g_1,\ldots,g_q,\ldots,t_1,\ldots,t_s$ да бъдат равни на нула. Нека разгледаме една група, в която влизат векторите $e_1,e_2,\ldots,e_p$. Очевидно в сумата в (9) тези вектори влизат само в членовете

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_p e_p-A(k_1 e_1+k_2 e_2+\ldots+k_p e_p).$$

Но като използуваме първата група формули от (4), тази сума става

$$(\alpha_1-k_1\lambda_1-k_2)e_1+(\alpha_2-k_2\lambda_1-k_3)e_2+\ldots+$$
$$+(\alpha_{p-1}-k_{p-1}\lambda_1-k_p)e_{p-1}+(\alpha_p-k_p\lambda_1)e_p.$$

Да приравним на нула коефициентите на векторите $e_1,e_2,\ldots,e_p$. Получаваме уравненията:

$$\begin{aligned}
&\alpha_1-k_1\lambda_1-k_2=0,\\
&\alpha_2-k_2\lambda_1-k_3=0,\\
&\cdots\cdots\cdots\\
&\alpha_{p-1}-k_{p-1}\lambda_1-k_p=0,\\
&\alpha_p-k_p\lambda_1=0,
\end{aligned}$$

които решаваме лесно и направо. Така от последното намираме k_p, от предпоследното намираме k_{p-1} и т. н. до k_1. Напълно подобно постъпваме с другите групи и равенството (9) при подбраните така числа $k,\mu,\ldots,\omega$ става $Ae'=0$. Като прибавим този вектор e' към разгледания базис, получаваме базиса

$$e',e_1,e_2,\ldots,e_p,g_1,g_2,\ldots,g_1,\ldots,t_1,t_2,\ldots,t_s$$

на $n+1$-мерното пространство R, в който преобразуването има нормална форма. Векторът e' образува нова група със собствено значение, равно на нула. Ако се върнем към първоначалното преобразуване, собственото значение вместо нулата ще бъде числото τ.

Да разгледаме сега втория случай, когато някои от собствените значения на оператора A в подпространството R' са равни на нула за някои от групите в представянето (4). За членовете в (9), които се отнасят за групи със собствени значения, отлични от нула, непосредствено се прилага горното им свеждане към нула при подходящо подбиране на съответните числа от числата $k,\ \mu,\ldots,\omega$. Да допуснем, че след направеното редуциране в дясната част на (9) останат три групи събираеми $e_1,\ e_2,\ldots,e_p;\ g_1,\ g_2,\ldots,g_q;\ h_1,h_2,\ldots,h_r$ със собствени значения, равни на нула, т. е. $\lambda_1=\lambda_2=\lambda_3=0$. Тогава равенството (9) става

$$\text{(10)}\qquad Ae'=\alpha_1 e_1+\ldots+\alpha_p e_p+\beta_1 g_1+\ldots+\beta_q g_q+\gamma_1 h_1+\ldots+\gamma_r h_r-$$
$$-A(k_1 e_1+\ldots+k_p e_p)-A(\mu_1 g_1+\ldots+\mu_q g_q)-A(\nu_1 h_1+\ldots+\nu_r h_r).$$

Векторите $e_1, e_2, \ldots, e_p$ влизат само в членовете

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_p e_p-A(k_1 e_1+k_2 e_2+\ldots+k_p e_p). \tag{11}$$

Понеже $\lambda_1=0$, то равенствата (4) ни дават

$$Ae_1=0,\ Ae_2=e_1,\ldots,Ae_p=e_{p-1}.$$

Следователно разликата (11) ще бъде равна на

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_p e_p-k_2 e_1-k_3 e_2-\ldots-k_p e_{p-1}$$

и при избора $k_2=\alpha_1,\ k_3=\alpha_2,\ldots,k_p=\alpha_{p-1}$ се свежда на единствения член $\alpha_p e_p$. Като постъпим така и за групите $g_1, g_2, \ldots, g_q$ и $h_1, h_2, \ldots, h_r$, получаваме за вектора e'.

$$Ae'=\alpha_p e_p+\beta_q g_q+\gamma_r h_r.$$

Ако случайно имаме $\alpha_p=\beta_q=\gamma_r=0$, то за вектора e' ще имаме $Ae'=0$ и, както по-горе, с прибавяне на вектора e' към базиса на подпространството R' получаваме базис за пространството R, при който A има нормална форма.

Нека сега поне едно от числата $\alpha_p, \beta_q, \gamma_r$ е отлично от нула. В този случай за разлика от разгледания по-рано става нужда от изменение на някои вектори от базиса в R'. Да наредим числата p, q, r по намаляваща стойност и нека $p>q>r$. Тогава построяваме нова група, като започваме с вектора e'. За тази цел полагаме $e'_{p+1}=e'$, $e'_p=Ae'_{p+1}$, $e'_{p-1}=Ae'_p, \ldots, e'_1=e'_2$. Ще имаме тогава

$$e'_{p+1}=e'=\alpha_p e_p+\beta_q g_q+\gamma_r h_r,$$

$$e'_p=Ae'_{p+1}=\alpha_p e_{p-1}+\beta_q g_{q-1}+\gamma_r h_{r-1},$$

. . . . . .

$$e'_{p-r+2}=Ae'_{p-r+3}=\alpha_p e_{p-r+1}+\beta_q g_{q-r+1}+\gamma_r h_1,$$

$$e'_{p-r+1}=Ae'_{p-r+2}=\alpha_p e_{p-r}+\beta_q g_{q-r},$$

. . . . . .

$$e_1'=Ae'_2=a_p e_1.$$

Заместваме сега векторите $e', e_1, e_2, \ldots, e_p$ от базиса с векторите

$$e_1', e_2', \ldots, e_p', e_{p'+1},$$

като останалите базисни вектори оставяме непроменени. Така ще получим нова нормална форма на преобразуването, като първата клетка се е увеличила с един елемент. Така теоремата е установена напълно.

Намирането на Жордановата форма на линейното преобразуване се получава с използуване на матрицата, която дава характеристичното уравнение, и се установява еднозначността на нормалното представяне.

3. Намиране на Жордановата нормална форма. Ще разгледаме сега начина за намирането на Жордановата нормална форма на едно линейно преобразуване, като едновременно с това ще установим

единствеността на това представяне. Ако една матрица $\mathbf{A}_1$ е получена от матрицата $\mathbf{A}$ с формулата

$$\mathbf{A}_1 = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}, \tag{12}$$

гдето $\mathbf{C}$ е също матрица, то казваме, че матрицата $\mathbf{A}_1$ е получена от матрицата $\mathbf{A}$ с преобразуване с матрицата $\mathbf{C}$ или по-късо матрицата $\mathbf{A}_1$ е подобна на матрицата $\mathbf{A}$. Разбира се, предполага се, че матриците са квадратни и $\mathbf{C}$ не е изродена матрица, т. е. съответната ѝ детерминанта е отлична от нула.

От равенството (12) получаваме лесно формулата

$$\mathbf{A} = \mathbf{C}_1^{-1}\mathbf{A}_1\mathbf{C}_1,\ \mathbf{C}_1 = \mathbf{C}^{-1},$$

т. е. обратно, матрицата $\mathbf{A}$ е подобна на матрицата $\mathbf{A}_1$.

Ако две матрици $\mathbf{A}_1$ и $\mathbf{A}_2$ са подобни на една и съща матрица $\mathbf{A}$, то те са подобни помежду си. Действително по условие имаме

$$\mathbf{A} = \mathbf{C}_1^{-1}\mathbf{A}_1\mathbf{C}_1,\ \mathbf{A} = \mathbf{C}_2^{-1}\mathbf{A}\mathbf{C}_2.$$

Тогава ще имаме

$$\mathbf{C}_1^{-1}\mathbf{A}_1\mathbf{C}_1 = \mathbf{C}_2^{-1}\mathbf{A}_2\mathbf{C}_2$$

или

$$\mathbf{A}_1 = \mathbf{C}_1\mathbf{C}_2^{-1}\mathbf{A}_2\mathbf{C}_2\mathbf{C}_1^{-1}.$$

Като поставим $\mathbf{C}_2\mathbf{C}_1^{-1} = \mathbf{C}$, то предното равенство се свежда на следното:

$$\mathbf{A}_1 = \mathbf{C}^{-1}\mathbf{A}\mathbf{C},$$

т. е. $\mathbf{A}_1$ и $\mathbf{A}_2$ са подобни, което трябваше да се докаже.

Ако $\mathbf{A}$ е матрицата на един линеен оператор A, то при смяна на базиса на пространството матрицата $\mathbf{A}$ преминава в подобната матрица $\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$, гдето $\mathbf{C}$ е матрицата на преминаването от първоначалния базис към новия такъв. Ние видяхме, че характеристичният полином на матрицата $\mathbf{A}$, т. е. детерминантата $D_n(\lambda) = |\mathbf{A} - \lambda\mathbf{E}|$ от матрицата $\mathbf{A} - \lambda\mathbf{E}$, не се променя. Тази детерминанта е един инвариант при преминаване от матрицата $\mathbf{A}$ към подобна на нея матрица. Задачата ни ще бъде да намерим една пълна система от такива инварианти, т. е. функции от елементите на една матрица, които не си променят стойността при сменяване на матрицата с подобни матрици и които еднозначно определят дадената матрица.

На матрицата $\mathbf{A} - \lambda\mathbf{E}$ миньорите от даден ред k, $1 \leqq k \leqq n$, са полиноми на λ. Да означим с $D_k(\lambda)$ общия им най-голям делител. Ще докажем, че $D_k(\lambda)$ са инварианти.

Непосредствено с развитие на детерминантата $D_p(\lambda)$ се вижда, че $D_p(\lambda)$ се дели на $D_{p-1}(\lambda)$, $1 < p \leqq n$. Ще установим няколко предложения:

1. Нека $\mathbf{B}$ е неизродена матрица. Тогава общият най-голям делител на миньорите от k-ти ред на матрицата $\mathbf{B}(\mathbf{A} - \lambda\mathbf{E})$ съвпада с този на миньорите от k-ти ред на матрицата $\mathbf{A} - \lambda\mathbf{E}$. Същото е в сила и за матрицата $(\mathbf{A} - \lambda\mathbf{E})\mathbf{B}$.

Действително да означим с a_{ij} елементите на матрицата $\mathbf{A}-\lambda\mathbf{E}$ и с a'_{ij} тези на матрицата $\mathbf{B}(\mathbf{A}-\lambda\mathbf{E})$. Тогава ще имаме

$$a'_{ij}=\sum_{s=1}^{n} b_{is}\,a_{sj},$$

т. е. a'_{ij} са линейни функции на елементите на матрицата $\mathbf{A}-\lambda\mathbf{E}$ с коефициенти, които не зависят от λ. Но тогава миньорите на матрицата $\mathbf{B}(\mathbf{A}-\lambda\mathbf{E})$ ще бъдат суми от миньорите на матрицата $\mathbf{A}-\lambda\mathbf{E}$ с коефициенти, които не зависят от λ. Следователно всеки общ делител на миньорите от k-ти ред на матрицата $\mathbf{A}-\lambda\mathbf{E}$ ще бъде делител на миньорите от същия ред, които принадлежат на матрицата $\mathbf{B}(\mathbf{A}-\lambda\mathbf{E})$. Но същото заключение е вярно и в обратно направление, понеже матрицата $\mathbf{A}-\lambda\mathbf{E}$ се получава от матрицата $\mathbf{B}(\mathbf{A}-\lambda\mathbf{E})$ с умножение на $\mathbf{B}^{-1}$. С това е установено горното предложение.

2. За подобните матрици полиномите $D_k(\lambda)$ съвпадат.

Нека $\mathbf{A}$ и $\mathbf{A}'=\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ са две подобни матрици. Тогава прилагаме предното предложение за матриците

$$\mathbf{A}-\lambda\mathbf{E},\quad (\mathbf{A}-\lambda\mathbf{E})\,\mathbf{C},$$

$$\mathbf{C}^{-1}(\mathbf{A}-\lambda\mathbf{E}),\quad \mathbf{C}^{-1}(\mathbf{A}-\lambda\mathbf{E})\,\mathbf{C}=\mathbf{A}'-\lambda\mathbf{E}.$$

Понеже при преминаване към нов базис матрицата на линейния оператор се замества с подобен на него оператор, то от предното предложение следва, че най-големият делител $D_k(\lambda)$ на миньорите от k-ти ред на матрицата $(\mathbf{A}-\lambda\mathbf{E})$, гдето $\mathbf{A}$ е матрицата на преобразуването A в някой базис, не зависи от базиса.

Нека сега базисът е така избран, че матрицата на оператора A да има нормалната форма на Жордан. Да намерим полиномите $D_k(\lambda)$. Отначало ще разгледаме случая, когато матрицата се състои само от една клетка, т. е. има вида

$$\begin{Vmatrix} \lambda_0 & 1 & 0 & 0 \dots 0 \\ 0 & \lambda_0 & 1 & 0 \dots 0 \\ \cdot & \cdot & \cdot & \cdot \quad \cdot \\ 0 & 0 & 0 & 0 \dots \lambda_0 \end{Vmatrix}.$$

Съответната детерминанта е

$$(-1)^n D_n(\lambda)=\begin{Vmatrix} \lambda_0-\lambda & 1 & 0 \dots 0 \\ 0 & \lambda_0-\lambda & 1 \dots 0 \\ \cdot & \cdot & \cdot \quad \cdot \\ 0 & 0 & 0 \dots \lambda_0-\lambda \end{Vmatrix}=(\lambda_0-\lambda)^n.$$

Като зачеркнем първия стълб и последния ред, виждаме веднага, че $D_{n-1}(\lambda)=1$ и по подобен начин, че $D_{n-2}(\lambda)=D_{n-3}(\lambda)=\dots=D_1(\lambda)=1$.

За да изследваме общия случай, т. е. за матрици от няколко клетки, ще установим едно помощно предложение.

Нека матрицата **B** има формата

$$\left\| \begin{matrix} \mathbf{B}_1 & 0 \\ 0 & \mathbf{B}_2 \end{matrix} \right\|,$$

гдето $\mathbf{B}_1$ и $\mathbf{B}_2$ са матрици от редове n_1 и n_2. Тогава отличните от нула миньори от ред m на матрицата **B** имат вида

$$\Delta_m = \Delta'_{m_1} \Delta''_{m_2}, \quad m_1 + m_2 = m,$$

гдето Δ'_{m_1} са миньори от m_1-ти ред на матрицата $\mathbf{B}_1$, а Δ''_{m_2} са миньори от m_2-ти ред на матрицата $\mathbf{B}_2$.

Предложението следва от развитието по правилото на Лаплас като тези редове от първите n_1 реда, които влизат в състава на дадения миньор. Тогава даденият миньор или е равен на нула, или е от горния вид.

Нека матрицата **A** има Жордановата нормална форма, която да бъде дадена с формулите (4). Предполагаме, че в матрицата **A** има p клетки, отговарящи на собственото значение λ_1, q клетки, отговарящи на собственото значение λ_2, и т. н. Да означим с $n_1, n_2, \ldots, n_p$ $(n_1 \geqq n_2 \geqq \ldots \geqq n_p)$ реда на клетките, отговарящи на λ_1, с $m_1, m_2, \ldots, m_q$ $(m_1 \geqq m_2 \geqq \ldots \geqq m_q)$ този на клетките, отговарящи на λ_2, и т. н. Матрицата $\mathbf{C} = \mathbf{A} - \lambda \mathbf{E}$ ще се състои от отделни клетки $\mathbf{C}_i$ от вида например

$$\mathbf{C}_1 = \left\| \begin{matrix} \lambda_1 - \lambda & 1 & 0 \ldots & 0 \\ 0 \cdot & \lambda_1 - \lambda & 1 \ldots & 0 \\ \cdot & \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \\ 0 & 0 & 0 \ldots & \lambda_1 - \lambda \end{matrix} \right\|$$

и за детерминантата $D_n(\lambda)$ ще имаме

$$D_n(\lambda) = (\lambda - \lambda_1)^{n_1 + n_2 + \cdots + n_p} (\lambda - \lambda_2)^{m_1 + m_2 + \cdots + m_q} \ldots$$

като произведение от детерминанти от матрици, имащи предния вид. Понеже $D_{n-1}(\gamma)$ е делител на $D_n(\lambda)$, то този полином ще има същите множители, но евентуално в други степени. Да намерим степента, в която влиза множителят $(\lambda - \lambda_1)$. Ще трябва да разгледаме миньорите на матрицата $\mathbf{C} = \mathbf{A} - \lambda \mathbf{E}$ от ред $n-1$, които са отлични от нула. Всеки такъв миньор ще има формата

$$\Delta_{n-1} = \Delta^{(1)}_{i_1} \Delta^{(2)}_{i_2} \ldots \Delta^{(k)}_{i_k},$$

като $i_1 + i_2 + \ldots + i_k = n-1$ и $\Delta^{(s)}_{i_s}$ са съответно миньори от ред i_s, принадлежащи на матрицата $\mathbf{C}_s$. Понеже сумата от редовете на миньорите $\Delta^{(s)}_{i_s}$ е с единица по-малка от реда на $D_n(\lambda)$, то на един и само на един от тези миньори редът ще бъде с единица по-малък от реда на съответната матрица $\mathbf{C}_i$. Но това ни показва, че ние можем така да изберем Δ_{n-1}, че някой от миньорите $\Delta^{(s)}_{i_s}$ да стане единица, като не про-

меняме останалите детерминанти на съответните клетки. Следователно, за да получим миньора, съдържащ $\lambda-\lambda_1$ в най-ниска степен, достатъчно е да зачеркнем един ред и стълб в клетката, отговаряща на λ_1 и имаща най-голям ред n_1. Така виждаме, че най-големият общ делител $D_{n-1}(\lambda)$ на миньорите от $n-1$-ви ред съдържа $\lambda-\lambda_1$ в степен $n_2+n_3+\ldots+n_p$. По подобен начин се убеждаваме, че между миньорите от $n-2$-ри ред най-ниска степен на $\lambda-\lambda_1$ съдържа миньорът Δ_{n-2}, получен със зачеркване на ред и стълб в клетките, съответствуващи на λ_1 и имащи редове n_1 и n_2. Следователно $D_{n-2}(\lambda)$ съдържа $\lambda-\lambda_1$ в степен $n_3+n_4+\ldots+n_p$. Също $D_{n-3}(\lambda)$ ще съдържа $\lambda-\lambda_1$ в степен $n_4+\ldots+n_p$ и т. н., като $D_1(\lambda)$ не съдържа $\lambda-\lambda_1$. Непосредствено се пренасят предните резултати за множителите $\lambda-\lambda_2$, $\lambda-\lambda_3$,... Така доказахме следното предложение:

Нека матрицата на линейния оператор A има Жорданова нормална форма, в която има p клетки от редове $n_1, n_2, \ldots, n_p (n_1 \geqq n_2 \geqq \ldots \geqq n_p)$, отговарящи на собственото значение λ_1, q клетки от редове $m_1, m_2, \ldots, m_q (m_1 \geqq m_2 \geqq \ldots \geqq m_q)$, отговарящи на собственото значение λ_2, и т. н. Тогава имаме

$$D_n(\lambda)=(\lambda-\lambda_1)^{n_1+n_2+\ldots+n_p}(\lambda-\lambda_2)^{m_1+m_2+\ldots+m_q}\ldots,$$

$$D_{n-1}(\lambda)=(\lambda-\lambda_1)^{n_2+\ldots+n_p}(\lambda-\lambda_2)^{m_2+\ldots+m_q}\ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

Ако въведем полиномите

$$E_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)},\ 2\leqq k\leqq n,$$

то за тях получаваме

$$E_n(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{m_1}\ldots,$$

$$E_{n-1}(\lambda)=(\lambda-\lambda_1)^{n_2}(\lambda-\lambda_2)^{m_2}\ldots,$$

$$\ldots\ldots\ldots\ldots\ldots$$

Тези полиноми се наричат **инвариантни множители**. Очевидно тези множители определят еднозначно Жордановата нормална форма. Собствените значения са корените на уравненията $E_n(\lambda)=0$. Редовете на клетките, отговарящи за собственото значение λ_1, са равни на кратността на корена λ_1 в уравненията $E_n(\lambda)=0$, $E_{n-1}(\lambda)=0,\ldots$ Оттук следва също важният факт: **за да има базис, при който матрицата на оператора е диагонална, необходимо и достатъчно е, щото инвариантните множители на тази матрица да имат само прости нули.** Именно в случая всички клетки трябва да са от първи ред.

ЧАСТ IV

ГЛАВНИ СВОЙСТВА НА АЛГЕБРИЧНИТЕ УРАВНЕНИЯ

Глава I

Основна теорема на алгебрата и непосредствени следствия

1. **Теорема на Даламбер за съществуване на корен**[1]. Всяко алгебрично уравнение

$$f(z)=a_0 z^n+a_1 z^{n-1}+\ldots+a_n=0,$$

в което коефициентите са произволни реални или комплексни числа, има поне един корен $z=a+bi$. Като следствие се извежда, че всяко уравнение има толкова корени, колкото е неговата степен.

За да установим теоремата, трябва да докажем, че има комплексно число z_0, за което $f(z_0)=0$ или, което е все едно, $|f(z_0)|=0$. Нека означим с

$$u=|f(z)|;$$

ако поставим $z=x+iy$, u е непрекъсната функция на двете реални променливи x и y във всяка крайна област. Действително за две точки z_1 и z_2 имаме

$$||f(z_1)|-|f(z_2)||\leqq|f(z_1)-f(z_2)|.$$

От това неравенство следва, че $|f(z)|$ е непрекъсната функция за всяка точка $z_1=x_1+iy_1$. Именно, понеже $f(z)$ е непрекъсната функция на комплексното променливо z, то на всяко произволно малко число $\varepsilon>0$ отговаря малко $\delta>0$, така че за всички комплексни числа z_2, за които $|z_2-z_1|<\delta$, да имаме $|f(z_1)-f(z_2)|<\varepsilon$. Но тогава ще имаме също, че $||f(z_1)|-|f(z_2)||<\varepsilon$. По-просто казано, когато $z_2\to z_1$ по произволен начин, $f(z_2)\to f(z_1)$ и от неравенството следва, че и $|f(z_2)|\to|f(z_1)|$.

По една известна теорема u като непрекъсната функция на x и y във всяка крайна област приема за една поне точка минималната си

---

[1] Строго доказателство дава пръв Gauss (1799). По-рано тази теорема е била доказана от D'Alembert, но с някои непълноти. Доказателството, което излагаме, произлиза от Cauchy (1821).

стойност $\mu \geqq 0$. От предложението на стр. 12 следва, че има число $R > 0$ такова, че за всичките z, на които $|z| \geqq R$, ще имаме

$$|f(z)| > |a_0| \, |z|^n \frac{1}{2} \left(\text{положили сме там } \varepsilon = \frac{1}{2} \right).$$

Нека $R_1 \geqq R$ е избрано така голямо, че $|a_0| R_1^n > 2|a_n|$. Понеже по окръжността $|z| = R_1$ ще имаме $|f(0)| = |a_n| < |f(z)|$, то в кръга $|z| \leqq R_1$ ще взема минималната си стойност $\mu = |f(z_0)|$ за една вътрешна точка z_0.

Ако $\mu = 0$, то теоремата е доказана, тъй като $f(z_0) = 0$. Ако допуснем, че $\mu > 0$, то въз основа на една лема на Коши, която ще докажем, следва, че има число $z_0 + h$, за което $|f(z_0+h)| < |f(z_0)| = \mu$, което обаче противоречи на условието, че μ е минимум.

Лема на Коши. Ако числото $f(z_0) \neq 0$, то можем да намерим такова число h, че да имаме

$$|f(z_0+h)| < |f(z_0)|.$$

Да развием $f(z_0+h)$ по степените на h според формулата на Тейлор:

$$f(z_0+h) = f(z_0) + \frac{h}{1!} f'(z_0) + \frac{h^2}{2!} f''(z_0) + \ldots + \frac{h^n}{n!} f^{(n)}(z_0).$$

Може да се случи, че $f'(z_0)$ да бъде равно на нула, така че за общност да предположим, че първият коефициент след $f(z_0)$, който не е нула, е $f^{(m)}(z_0)$.

Като разделим с $f(z_0) \neq 0$, ще имаме

$$\frac{f(z_0+h)}{f(z_0)} = 1 + \frac{h^m}{m!} \frac{f^{(m)}(z_0)}{f(z_0)} + \ldots + \frac{h^n}{n!} \frac{f^{(n)}(z_0)}{f(z_0)}.$$

Даваме на коефициентите на h тригонометрична форма, както и на самото h, като полагаме

$$h = \rho(\cos\varphi + i \sin\varphi),$$

$$\frac{f^{(m)}(z_0)}{m! f(z_0)} = R_m(\cos\varphi_m + i \sin\varphi_m),$$

. . . . . .

$$\frac{f^{(n)}(z_0)}{n! f(z_0)} = R_n(\cos\varphi_n + i \sin\varphi_n).$$

Ще получим

$$(1) \qquad \frac{f(z_0+h)}{f(z_0)} = 1 + R_m \rho^m [\cos(m\varphi + \varphi_m) + i \sin(m\varphi + \varphi_m)] + \ldots +$$

$$+ \rho^n R_n [\cos(n\varphi + \varphi_n) + i \sin(n\varphi + \varphi_n)].$$

Да изберем φ така, че $m\varphi+\varphi_m=\pi$, т. е.

$$\varphi=\frac{\pi-\varphi_m}{m},$$

и да дадем на ρ достатъчно малки стойности, така че

$$R_m\rho^m<1.$$

Първите два члена вдясно на (1) ще станат

$$1-R_m\rho^m>0.$$

Ако сега вземем модулите на двете части и приложим теоремата за модула на сума на няколко комплексни числа, първото от които е $1-R_m\rho^m$, получаваме

$$\left|\frac{f(z_0+h)}{f(z_0)}\right|\leqq 1-R_m\rho^m+R_{m+1}\rho^{m+1}+\ldots+R_n\rho^n$$

или

$$\left|\frac{f(z_0+h)}{f(z_0)}\right|\leqq 1-\rho^m(R_m-R_{m+1}\rho-\ldots-R_n\rho^{n-m}).$$

Очевидно е, че изразът в скобите може да стане произволно близък до R_m, стига да вземем ρ достатъчно малко. Тогава дясната част става произволно близка до $1-R_m\rho^m$, т. е. до едно число, по-малко от 1, с което лемата е доказана. Но тогава, както вече споменахме, следва теоремата на Даламбер.

2. **Разлагане в линейни множители.** Видяхме, че уравнението

$$f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n=0$$

има поне един корен z_1. Да разделим $f(z)$ със $z-z_1$ и нека частното, което е полином от $n-1$-ва степен, да бъде $f_1(z)$, а остатъкът R, който е постоянно число, независещо от z. Ще имаме

$$f(z)=(z-z_1)f_1(z)+R.$$

Ако поставим $z=z_1$, понеже $f(z_1)=0$, получаваме $R=0$, т. е.

$$f(z)=(z-z_1)f_1(z). \tag{2}$$

Обаче по същата теорема на Даламбер уравнението

$$f_1(z)=0$$

ще има поне един корен z_2. На същото основание можем да напишем

$$f_1(z)=(z-z_2)f_2(z). \tag{3}$$

Като продължаваме така, ще получим последователно

$$\left|\begin{array}{l} f_2(z)=(z-z_3)f_3(z), \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ f_{n-1}(z)=(z-z_n)f_n(z), \end{array}\right. \tag{4}$$

гдето $f_n(z)$ е константа, която е равна на a_0. Ако заместим последователно полиномите $f_i(z)$ в (2), (3), (4), получаваме

$$f(z)=a_0(z-z_1)(z-z_2)\ldots(z-z_n). \tag{5}$$

Всеки полином от n-та степен се разлага на произведението от n линейни множителя.

Понеже $f(z)$ се анулира, ако поне един от множителите се анулира, то $z_1, z_2, \ldots, z_n$ са всичките корени на $f(z)=0$.

Всяко уравнение има толкова корена, колкото е степента му. От тези корени z_i някои могат да са равни. Така нека $z_1=z_2$, тогава в $f(z)$ множителят $z-z_1$ влиза в квадрат. Наричаме корена z_1 двукратен. Ако въобще множителят $z-z_1$ се повтаря k пъти в (5), т. е. влиза в k-та степен, то коренът z_1 се нарича k-кратен. Двукратните, трикратните и повече кратни корени се наричат **многократни**. Останалите корени се наричат **прости**.

Читателят лесно ще се убеди, че всеки полином се разлага по единствен начин на линейни множители, т. е. от $a_0(z-z_1)\ldots(z-z_p)=b_0(z-u_1)\ldots(z-u_q)$ следва, че $p=q$, $a_0=b_0$, $z_k=u_k$, $k=1, 2, \ldots, p$.

Едно интересно следствие от разлагането (5) е следното: нека допуснем, че полиномът от n-та степен $f(z)$ се анулира за повече от n различни стойности на z. Ако тези стойности са $z_1, z_2, \ldots, z_n, z_{n+1}, \ldots$, то бихме заключили от разлагането (5), че $a_0=0$. Но тогава полиномът се редуцира на такъв от $n-1$-ва степен; така получаваме на същото основание, че $a_1=0$ и т. н., т. е. всички коефициенти на $f(z)$ трябва да бъдат равни на нула. Следователно имаме:

Ако един полином $f(z)$ от n-та степен се анулира за повече от n различни значения на z, то трябва всички негови коефициенти да бъдат равни на нула и уравнението $f(z)=0$ е едно тъждествено уравнение, което се удовлетворява за всяко z.

Оттук следва непосредствено: ако два полинома от n-та степен получават еднакви стойности за повече от n значения на променливото, то те са идентично равни, понеже разликата им е полином от n-та степен, който се анулира за повече от n значения на променливото и следователно е тъждествено равен на нула.

От горните изводи получаваме лесно общата форма на алгебричното уравнение, на което са дадени корените $z_1, z_2, \ldots, z_n$. Тя ще бъде

$$A(z-z_1)(z-z_2)\ldots(z-z_n)=0,$$

гдето A е произволна константа.

Ако във формулата (5) наредим дясната част по степените на z и приравним коефициентите в двете части, ще получим следните връзки между корените и коефициентите:

(6)
$$\begin{cases} z_1+z_2+\ \ldots\ +z_n=-\dfrac{a_1}{a_0}, \\ z_1z_2+\ldots+z_1z_n+\ldots+z_{n-1}z_n=\dfrac{a_2}{a_0}, \\ z_1z_2z_3+\ \ldots\ +z_{n-2}z_{n-1}z_n=-\dfrac{a_3}{a_0}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ z_1z_2\ldots z_n=(-1)^n\ \dfrac{a_n}{a_0}. \end{cases}$$

Сумата от корените на едно уравнение е равна на отношението на втория коефициент, взет с обратен знак, към първия коефициент. Сумата от произведението на корените, взети по два както комбинациите, е равна на третия коефициент, разделен с първия. Въобще сумата на произведенията на корените по k, взети както комбинациите от k-ти клас, е равна на $k+1$-ия коефициент, умножен с $(-1)^k$ и разделен с първия.

Чрез формулите (6), когато между корените са дадени някои допълнителни връзки, можем в някои случаи да решим уравнението. Обаче само от тях с елиминиране не може да се реши уравнението, понеже получаваме пак същото уравнение.

3. **Уравнения с реални коефициенти.** Ако коефициентите на уравнението са реални числа и то има комплексния корен $\alpha+\beta_i$, тогава то допуска също за корен конюгованото число $\alpha-\beta i$.

Нека е дадено уравнението

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0,$$

на което коефициентите $a_0, a_1, \ldots, a_n$ са реални числа. Нека $x=\alpha+i\beta$ е един комплексен корен. Ако поставим тази стойност на x в лявата част на уравнението, получаваме алгебричния анализ

$$P(\alpha,\ \beta)+iQ(\alpha,\ \beta),$$

в който P и Q са полиноми на α, β; P съдържа само четни степени на β, а Q — само нечетни, понеже четните степени на i са ± 1, а нечетните $\pm i$. Следователно

$$P(\alpha,-\beta)=P(\alpha,\ \beta),\quad Q(\alpha,\ \beta)=-Q(\alpha,\ -\beta).$$

Понеже $\alpha+i\beta$ е корен, то $P(\alpha,\ \beta)+iQ(\alpha,\ \beta)=0$, т. е. $P(\alpha,\ \beta)=Q(\alpha,\ \beta)=0$. Ако в уравнението поставим вместо x числото $\alpha-\beta i$, то лявата му част се обръща в

$$P(\alpha,\ \beta)-iQ(\alpha,\ \beta)=0,$$

което показва, че $\alpha-i\beta$ е също корен.

Следователно имагинерните корени са два по два конюговани и броят им е четно число. Ако полиномът $f(x)$ съдържа множителя

$$x-(\alpha+i\beta),$$

то той ще съдържа и множителя

$$x-(\alpha-i\beta).$$

Но произведението на тези два множителя е

$$(x-\alpha)^2+\beta^2=x^2-2\alpha x+\alpha^2+\beta^2.$$

Следователно една цяла рационална функция с реални коефициенти може да се представи като произведение на реални множители от първа и втора степен.

4. **Най-голям общ делител на два полинома.** За един полином A се казва, че се дели на друг полином B, ако съществува трети полином C така, че да имаме

$$A=BC$$

за всяко x. Под общ делител на два полинома A и B разбираме трети полином най-малко от първа степен, който едновременно дели и двата полинома. Общият делител на A и B, който е от най-висока степен, се нарича общ най-голям делител. Ясно е, че ако D е общият най-голям делител на A и B, то ще имаме

$$A=A_1D,\quad B=B_1D,$$

гдето A_1 и B_1 са два полинома без общ делител или, както се казва, взаимно прости.

Ако можем да разложим A и B като произведения на линейни множители, то D се намира веднага. Именно нека $\alpha_1, \alpha_2, \ldots, \alpha_n$ са корените на уравнението $A=0$, а $\beta_1, \beta_2, \ldots, \beta_m$ са тези на $B=0$. Тогава от

$$A=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

$$B=b_0(x-\beta_1)\ (x-\beta_2)\ldots(x-\beta_n)$$

заключаваме веднага, че D е произведение от тези биномни множители, които едновременно се съдържат и в двете десни части. Оттук следва, че уравнението $D=0$ дава общите корени на уравненията $A=0$ и $B=0$.

Общият най-голям делител D можем също да намерим посредством последователно деление, именно като приложим Евклидовия алгоритъм, който се използува при намиране на най-голям общ делител на две числа. Ако B е от по-ниска степен от A или най-много равна на степента A, то разделяме A на B. Нека Q е частното, а R_1 остатъкът от делението; ще имаме

$$A=BQ+R_1.$$

Всеки общ делител на A и B ще дели и R_1. Обратно, всеки общ делител на B и R_1 е такъв и на A и B.

Остатъкът R_1 е от по-ниска степен от тази на B. Разделяме B с R_1 и нека Q_1 е частното, а R_2 е остатъкът от делението, то ще имаме

$$B=R_1 Q_1+R_2.$$

Всеки общ делител на B и R_1 ще дели R_2 и, обратно, всеки общ делител на R_1 и R_2 ще бъде такъв и на B и R_1. Продължаваме така делението и ще получим една редица от тъждества:

$$\begin{aligned}
&A=BQ+R_1,\\
&B=R_1 Q_1+R_2,\\
&R_1=R_2 Q_2+R_3,\\
&\cdots\cdots\cdots\cdots\\
&R_{\nu-2}=R_{\nu-1} Q_{\nu-1}+R_\nu.
\end{aligned} \tag{7}$$

Степените на полиномите R_i намаляват и така очевидно ще достигнем до остатък R_ν, който не съдържа x, т. е. е постоянно число. Видяхме, че всеки общ делител на A и B ще дели R_1, R_2, ... до R_ν. Ако $R_\nu \neq 0$, то следва, че A и B няма да имат общ делител. Ако $R_\nu=0$, то $R_{\nu-1}$ ще бъде общият най-голям делител на A и B. Действително $R_{\nu-1}$ дели $R_{\nu-2}$ и следователно ще дели $R_{\nu-3}$, $R_{\nu-4}$, ..., B и A. От друга страна, всеки общ делител на A и B дели R_1, R_2, ... до $R_{\nu-1}$.

Следователно имаме правилото: ако при последователното деление последният остатък, който е постоянно число, е отличен от нула, то полиномите са взаимно прости. Ако този остатък е равен на нула, общият най-голям делител е равен на предпоследния остатък.

Ако във формулите (7) определим R_1 от първото уравнение и го заместим във второто, след това получените стойности за R_1 и R_2 заместим в третото и т. н., то ще получим.

$$\begin{aligned}
&R_1=A-BQ,\\
&R_2=-AQ_1+(1+QQ_1)B,\\
&R_3=A(1+Q_1Q_2)-B(Q+Q_2+QQ_1Q_2),\\
&\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

Така ще достигнем до уравнение от вида

$$R_\nu=XA-YB,$$

където X, Y са цели рационални функции на Q, Q_1, Q_2, ..., т. е. полиноми. Оттук получаваме:

Ако A и B са полиноми на x, то могат да се намерят два полинома X, Y на x така, че

$$XA-YB=\text{const.}$$

Тази константа е нула, ако A и B имат общ делител. Тя е отлична от нула, ако A и B са взаимно прости. Оче-

видно във втория случай константата може да считаме равна на единица, като разделим с нея и я вмъкнем в изразите на X и Y. Също от горното следва, че и на общия най-голям делител $R_{\nu-1}$ може да се даде същата форма:

$$R_{\nu-1}=X_1 A-Y_1 B,$$

гдето X_1, Y_1 са полиноми на x, които са, както веднага се вижда, взаимно прости.

5. Отделяне на многократните корени. Нека на уравнението $f(x)=0$ коренът a е m-кратен. Тогава, понеже лявата част $f(x)$ съдържа множителя $(x-a)^m$, то ще имаме

$$f(x)=(x-a)^m \varphi(x),$$

гдето полиномът $\varphi(x)$ съдържа биномните множители, които се отнасят за другите корени на уравнението, т. е. $\varphi(a) \neq 0$. Като диференцираме, получаваме

$$f'(x)=(x-a)^{m-1}[m\varphi(x)+(x-a)\varphi'(x)]. \tag{8}$$

Изразът в скобите при $x=a$ взема стойност

$$m\varphi(a) \neq 0.$$

Следователно, ако a е m-кратен корен на $f(x)=0$, то той е $m-1$-кратен на уравнението $f'(x)=0$. Всеки прост корен на първото уравнение не е корен на второто. Понеже $f''(x)$ е първа производна на $f'(x)$, a ще бъде $m-2$-кратен корен на $f''(x)=0$; така също ще е $m-3$-кратен на $f'''(x)=0,\ldots,$ прост на $f^{(m-1)}(x)=0$. Значи, ако a е m-кратен корен на $f(x)=0$, то той е последователно $m-1$-кратен, $m-2$-кратен, ..., прост корен съответно на

$$f'(x),\ f''(x),\ldots,\ f^{(m-1)}(x)$$

и $f^{(m)}(a) \neq 0$.

От това следва, че ако $\alpha+i\beta$ е m-кратен корен на уравнението $f(x)=0$, което е с реални коефициенти, то $\alpha-i\beta$ е също така m-кратен корен.

От (8) се вижда, че на m-кратен корен a отговаря общият множител $(x-a)^{m-1}$ на $f(x)$ и $f'(x)$ и по-висока степен на $(x-a)$ не може едновременно да дели $f(x)$ и $f'(x)$. В най-големия общ делител на $f(x)$ и $f'(x)$ $x-a$ ще влизат в $m-1$-ва степен. Така например, ако

$$f(x)=(x-a)^m (x-b)^n) \varphi(x),$$

гдето във $\varphi(x)$ влизат само биномните множители за простите корени на $f(x)=0$, ще имаме за общия най-голям делител на $f(x)$ и $f'(x)$

$$D=(x-a)^{m-1}(x-b)^{n-1}.$$

Ако разделим $f(x)$ с D, получаваме

$$\frac{f(x)}{D}=(x-a)(x-b)\varphi(x).$$

Уравнението $\frac{f(x)}{D}=0$ има същите корени като уравнението $f(x)=0$, само че всичките са прости.

Посредством деление (метода на Хермит) може да се определи произведението на всички прости биномни множители, също произведението на всички двойни множители и т. н. на даденото уравнение $f(x)=0$, без да се пресмятат корените. Действително нека X_1 е произведението на простите линейни множители на $f(x)$, X_2—произведението на двойните, X_3—това на тройните и т. н. Тогава

$$f(x)=X_1 X_2^2 X_3^3 X_4^4 \ldots$$

Общият най-голям делител на $f(x)$ и $f'(x)$ ще бъде

$$D_1=X_2 X_3^2 X_4^3 \ldots$$

На същото основание общият най-голям делител на D_1 и D_1' ще бъде

$$D_2=X_3 X_4^2 \ldots,$$

общият най-голям делител на D_2 и D_2' ще бъде

$$D_3=X_4 \ldots$$

Продължаваме така, докато достигнем до полином D, който е взаимно прост с производната си D'.

Ако сега разделим последното, получаваме

$$f_1(x)=\frac{f(x)}{D_1}=X_1 X_2 X_3 X_4 \ldots,$$

$$f_2(x)=\frac{D_1}{D_2}=X_2 X_3 X_4 \ldots,$$

$$f_3(x)=\frac{D_2}{D_3}=X_3 X_4 \ldots,$$

$$f_4(x)=\frac{D_3}{D_4} X_4 \ldots,$$

. . . . . .

Оттук получаваме

$$X_1=\frac{f_1(x)}{f_2(x)}, \quad X_2=\frac{f_2(x)}{f_3(x)}, \quad X_3=\frac{f_3(x)}{f_4(x)}, \ldots$$

Така уравненията

$$X_1=0, \; X_2=0, \; X_3=0, \ldots$$

дават всички корени на уравнението $f(x)=0$ (всеки корен, взет като прост).

Да вземем един пример:

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

Получаваме

$$D_1 = x - 1,$$

отгдето

$$f(x) = X_1 X_2^2,\ D_1 = X_2,$$

$$f_1(x) = \frac{f(x)}{D_1} = X_1 X_2 = x^2 + 2x - 3,\ f_2(x) = D_1 = X_2,$$

$$X_1 = \frac{f_1(x)}{f_2(x)} = x + 3.$$

Следователно имаме

$$x^3 + x^2 - 5x + 3 = (x-1)^2 (x+3).$$

Друг пример:

$$f(x) = x^7 - 2x^6 - 2x^5 + 5x^4 + x^3 - 4x^2 + 1 = 0.$$

Получаваме

$$D_1 = x^3 - x^2 - x + 1,$$

$$D_2 = x - 1.$$

Понеже D_2 няма общ множител с D_2', то ще имаме

$$f(x) = X_1 X_2^2 X_3^3,\ D_1 = X_2 X_3^2,\ D_2 = X_3$$

и следователно

$$f_1(x) = \frac{f(x)}{D_1} = X_1 X_2 X_3 = x^4 - x^3 - 2x^2 + x + 1,$$

$$f_2(x) = \frac{D_1}{D_2} = X_2 X_3 = x^2 - 1,\quad f_3(x) = D_2,$$

$$\frac{f_1(x)}{f_2(x)} = X_1 = x^2 - x - 1,\ \frac{f_2(x)}{f_3(x)} = X_2 = x + 1,\quad D_2 = X_3 = x - 1.$$

Даденото уравнение е

$$(x^2 - x - 1)(x+1)^2 (x-1)^3 = 0.$$

Първият множител дава двата прости корена

$$\frac{1 \pm \sqrt{5}}{2},$$

а другите са: — 1 двукратен, 1 трикратен.

Глава II

Интерполация

1. **Формула на Лагранж.** Много въпроси от приложната математика водят до следната задача: да се намери полином $f(x)$ най-много от $(n-1)$-ва степен, който за n различни стойности $x_1, x_2, \ldots, x_n$ на x да приема n дадени стойности $y_1, y_2, \ldots, y_n$, т. е.

$$f(x_1)=y_1 \quad f(x_2)=y_2, \ldots, f(x_n)=y_n. \tag{1}$$

Ако търсеният полином е

$$f(x)=a_0x^{n-1}+a_1x^{n-2}+\ \ldots+a_{n-1},$$

то горните условия дават

$$\begin{cases} a_0x_1^{n-1}+a_1x_1^{n-2}+\ldots+a_{n-1}=y_1, \\ a_0x_2^{n-1}+a_1x_2^{n-2}+\ldots+a_{n-1}=y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0x_n^{n-1}+a_1x_n^{n-2}+\ldots+a_{n-1}=y_n. \end{cases} \tag{2}$$

В тези уравнения неизвестни са $a_0, a_1, \ldots, a_{n-1}$. Понеже детерминантата пред неизвестните е една детерминанта на Вандермонд, която е отлична от нула, понеже се разлага на произведения на разликите от числата x_i, то системата има само едно определено решение, което можем лесно да намерим. Обаче ние ще дадем друго решение, което води до по-прости пресмятания, с извеждане на формулите на Lagrange и Newton.

Нека $\varphi_k(x)$ е полином от $(n-1)$-ва степен, който се анулира за $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$, а при $x=x_k$ приема стойност 1. Следователно той ще има формата

$$\varphi_k(x)=A(x-x_1)\ (x-x_2)\ldots\ (x-x_{k-1})\ (x-x_{k+1})\ldots(x-x_n).$$

Понеже $\varphi_k(x_k)=1$, то получаваме

$$A=\frac{1}{(x_k-x_1)\ldots(x_k-x_{k-1})\ (x_k-x_{k+1})\ldots\ (x_k-x_n)},$$

$$\varphi_k(x)=\frac{(x-x_1)\ldots(x-x_{k-1})\ (x-x_{k+1})\ldots(x-x_n)}{(x_k-x_1)\ldots(x_k-x_{k-1})(x_k-x_{k+1})\ldots(x_k-x_n)}.$$

Тогава полиномът

$$f(x)=y_1\varphi_1(x)+y_2\varphi_2(x)+\ \ldots+y_n\varphi_n(x)$$

дава решението на търсената задача. Действително при $x=x_1, x_2, \ldots, x_n$ той приема съответно стойностите $y_1, y_2, \ldots, y_n$. Ако поставим

$$F(x)=(x-x_1)\ldots(x-x_k)\ldots(x-x_n)=\prod_{i=1}^{n}(x-x_i),$$

то с диференциране и заместване x с x_k се получава

$$F'(x_k)=(x_k-x_1)\ldots(x_k-x_{k-1})\,(x_k-x_{k+1})\ldots(x_k-x_n).$$

Следователно имаме

$$\varphi_k(x)=\frac{F(x)}{(x-x_k)\,F'(x_k)},$$

$$f(x)=\sum_{k=1}^{n} y_k\frac{F(x)}{(x-x_k)\,F'(x_k)},$$

която е интерполационната формула на Лагранж.

2. **Формула на Нютон.** Да образуваме израза

$$f(x_1,\ x_2)=\frac{f(x_1)-f(x_2)}{x_1-x_2},$$

който се нарича първа интерполираща функция на $f(x)$. Изразът

$$f(x_1,\ x_2,\ x_3)=\frac{f(x_1,\ x_2)-f(x_1,\ x_3)}{x_2-x_3},$$

който при постоянно x_1 представлява първа интерполираща функция на $f(x_1,\ x)$, се нарича втора интерполираща функция на $f(x)$. Следвайки така, получаваме

$$f(x_1, x_2, \ldots, x_n)=\frac{f(x_1,\ x_2, \ldots, x_{n-2},\ x_{n-1})-f\,(x_1,\ x_2, \ldots, x_{n-2},\ x_n)}{x_{n-1}-x_n},$$

която е n-тата интерполираща функция. Ако означим с

$$f(x_1)=y_1,\quad f(x_2)=y_2, \ldots,\ f(x_n)=y_n,$$

то намираме последователно

$$(3)\quad \begin{cases} f(x_1,\ x_2)=\dfrac{y_1}{x_1-x_2}+\dfrac{y_2}{x_2-x_1}, \\ f(x_1,\ x_2,\ x_3)= \\ \dfrac{y_1}{(x_1-x_2)\,(x_1-x_3)}+\dfrac{y_2}{(x_2-x_1)\,(x_2-x_3)}+\dfrac{y_3}{(x_3-x_1)\,(x_3-x_2)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(x_1, x_2, \ldots, x_n)=\dfrac{y_1}{(x_1-x_2)\ldots(x_1-x_n)}+\ldots+\dfrac{y_n}{(x_n-x_1)\ldots(x_n-x_{n-1})}. \end{cases}$$

От дефиницията на $f(x_1,\ x_2)$ веднага следва при $x_2=x$

$$f(x)=f(x_1)+(x-x_1)\,f(x_1,\ x)$$

и аналогично

$$f(x_1,\ x)=f(x_1,\ x_2)+(x-x_2)\,f(x_1,\ x_2,\ x),$$

като заместим, получаваме

$$f(x)=f(x_1)+(x-x_1)f(x_1,\ x_2)+(x-x_1)\,(x-x_2)f(x_1,\ x_2,\ x).$$

Ако продължаваме по този начин, ще получим въобще

$$\text{(4)}\qquad f(x)=f(x_1)+(x-x_1)\,f(x_1,\ x_2)+$$
$$+(x-x_1)\,(x-x_2)\,f(x_1,\ x_2,\ x_3)+\ldots+$$
$$+(x-x_1)\ (x-x_2)\ldots(x-x_{n-1})\ f(x_1,\ x_2,\ldots,\ x_{n-1},\ x).$$

Ако полиномът $f(x)$ трябва да бъде от $(n-1)$-ва степен, то понеже $f(x_1,x)$ е равно на частното на $f(x)-f(x_1)$ с $x-x_1$, $f(x_1,\ x)$ е полином от $n-2$-ра степен. Оттук веднага следва, че полиномът $f(x_1,\ x_2,\ x)$ е от $n-3$-та степен и т. н. и най-после полиномът $f(x_1, x_2, \ldots, x_{n-1},\ x)$ е равен на константа, т. е.

$$f(x_1,\ x_2,\ldots,\ x_{n-1},\ x)=f(x_1,\ x_2,\ldots,\ x_{n-1},\ x_n).$$

Следователно горната формула (4) става

$$f(x)=f(x_1)+(x-x_1)\ f(x_1,\ x_2)+(x-x_1)(x-x_2)\ f(x_1,\ x_2,\ x_3)+\ldots+$$
$$+(x-x_1)\ldots(x-x_{n-1})f(x_1,\ x_2,\ldots,x_{n-1},\ x_n),$$

което представлява формулата на Нютон. Оттук полиномът $f(x)$ от степен $n-1$, който за n различни стойности $x_1, x_2, \ldots, x_n$ на x приема n дадени стойности $y_1, y_2, \ldots, y_n$, може да се представи така:

$$\text{(4')}\qquad f(x)=f(x_1)+\alpha_1(x-x_1)+\alpha_2(x-x_1)(x-x_2)+$$
$$+\alpha_3(x-x_1)(x-x_2)\ (x-x_3)+\ldots+$$
$$+\alpha_{n-1}(x-x_1)\,(x-x_2)\ldots(x-x_{n-1}),$$

гдето числата α са дадени с формулите

$$\alpha_k=\frac{y_1}{(x_1-x_2)\ldots(x_1-x_{k+1})}+\frac{y_2}{(x_2-x_1)\ldots(x_2-x_{k+1})}+\ldots+$$
$$+\frac{y_{k+1}}{(x_{k+1}-x_1)\ldots(x_{k+1}-x_k)}\quad(k=1,\ 2,\ 3,\ \ldots,\ n-1).$$

3. **Разлики.** Нека е дадена редицата

$$a_0,\ a_1,\ a_2,\ldots,\ a_n,\ldots$$

от променливи или числа. Означаваме тогава с

$$\Delta a_k=a_{k+1}-a_k,$$

гдето знакът Δ е символ за означаване на разлика. Така получаваме първите разлики

$$\Delta a_0,\ \Delta a_1,\ \Delta a_2, \ldots$$

От тези числа можем да образуваме разлики, които наричаме втори разлики на дадената редица:

$$\Delta^2 a_0,\ \Delta^2 a_1,\ \Delta^2 a_2, \ldots,$$

гдето

$$\Delta^2 a_k = \Delta a_{k+1} - \Delta a_k = \Delta(\Delta a_k).$$

Продължавайки така, можем да получим разлики от кой да е ред, като имаме последователно

$$\Delta^n a_k = \Delta(\Delta^{n-1} a_k) = \Delta^{n-1} a_{k+1} - \Delta^{n-1} a_k.$$

Очевидно имаме

$$\Delta^{n+m} a_k = \Delta^n(\Delta^m a_k),$$

т. е. символът Δ удовлетворява на релацията

$$\Delta^n \Delta^m = \Delta^{m+n}.$$

Лесно е да намерим израз на $\Delta^n a_k$ посредством елементите a_m. Действително от дефиницията имаме

$$\Delta a_k = a_{k+1} - a_k,$$

$$\Delta^2 a_k = \Delta a_{k+1} - \Delta a_k = a_{k+2} - 2a_{k+1} + a_k,$$

$$\Delta^3 a_k = \Delta^2 a_{k+1} - \Delta^2 a_k = a_{k+3} - 3a_{k+2} + 3a_{k+1} - a_k,$$

$$\Delta^4 a_k = \Delta^3 a_{k+1} - \Delta^3 a_k = a_{k+4} - 4a_{k+3} + 6a_{k+2} - 4a_{k+1} + a_k$$

и оттук въобще

$$\Delta^n a_k = a_{k+n} - \binom{n}{1} a_{k+n-1} + \binom{n}{2} a_{k+n-2} - \ldots + (-1)^n a_k. \tag{5}$$

Тази формула може лесно да се докаже по индуктивен път от n към $n+1$, като се използува свойството на биномните коефициенти. Действително

$$\Delta^{n+1} a_k = \Delta^n a_{k+1} - \Delta^n a_k + a_{k+n+1} - \left[\binom{n}{1} + \binom{n}{0}\right] a_{k+n} +$$

$$+ \left[\binom{n}{2} + \binom{n}{1}\right] a_{k+n-1} - \ldots + (-1)^{n+1} a_k = a_{k+n+1} -$$

$$- \binom{n+1}{1} a_{k+n} + \binom{n+1}{2} a_{k+n-1} - \ldots + (-1)^{n+1} a_k,$$

понеже

$$\binom{n}{1} + \binom{n}{0} = \binom{n+1}{1},\ \binom{n}{2} + \binom{n}{1} = \binom{n+1}{2}, \ldots, \binom{n}{\mu} + \binom{n}{\mu-1} = \binom{n+1}{\mu}.$$

Обратно, имаме

$$a_{k+1}=a_k+\Delta a_k,$$

$$a_{k+2}=a_{k+1}+\Delta a_{k+1}=a_k+2\Delta a_k+\Delta^2 a_k,$$

$$a_{k+3}=a_{k+2}+\Delta a_{k+2}=a_k+3\Delta a_k+3\Delta^2 a_k+\Delta^3 a_k$$

и изобщо

$$(6)\qquad a_{k+n}=a_k+\binom{n}{1}\Delta a_k+\binom{n}{2}\Delta^2 a_k+\binom{n}{3}\Delta^3 a_k+\ldots+\Delta^n a_k.$$

Верността на тази формула се установява, както на горната.

За дадена функция $f(x)$ при фиксирано h си образуваме разликите между функционните стойности, които означаваме с

$$\Delta f(x)=f(x+h)-f(x),$$

$$\Delta^2 f(x)=\Delta(\Delta f(x))=\Delta f(x+h)-\Delta f(x),\quad \Delta^3 f(x)=\Delta(\Delta^2 f(x)),\ldots$$

Тогава във формулата на Нютон (4′) да поставим

$$x_2=x_1+h,\quad x_3=x_2+h=x_1+2h,\ldots,x_n=x_1+(n-1)h,$$

$$y_1=f(x_1),\ y_2=f(x_1+h),y_3=f(x_1+2h),\ldots,y_n=f(x_1+(n-1)h).$$

За $\alpha_1,\ \alpha_2,\ldots,\alpha_{n-1}$ получаваме

$$\alpha_1=\frac{y_2-y_1}{h}=\frac{\Delta y_1}{h}=\frac{\Delta f(x_1)}{h},$$

$$\alpha_2=\frac{1}{2!\,h^2}(y_3-2y_2+y_1)=\frac{\Delta^2 y_1}{2!\,h^2}=\frac{\Delta^2 f(x_1)}{2!\,h^2},$$

$$\alpha_3=\frac{1}{3!\,h^3}(y_4-3y_3+3y_2-y_1)=\frac{\Delta^3\ y_1}{3!\,h^3}=\frac{\Delta^3 f(x_1)}{3!\,h^3},$$

. . . . . .

$$\alpha_{n-1}=\frac{1}{(n-1)!\,h^{n-1}}\Big[y_n-\binom{n-1}{1}y_{n-1}+\binom{n-1}{2}y_{n-2}-\ldots+$$

$$+(-1)^{n-1}y_1\Big]=\frac{\Delta^{n-1}\ y_1}{(n-1)!\,h^{n-1}}=\frac{\Delta^{n-1}f(x_1)}{(n-1)!\,h^{n-1}}.$$

Формулата взема следния по-прост вид, като се постави $x=x_1+zh$,

$$f(x_1+zh)=f(x_1)+\frac{z}{1!}\Delta f(x_1)+\frac{z(z-1)}{2!}\Delta^2 f(x_1)+\ldots+$$

$$+\frac{z(z-1)\ldots(z-n+2)}{(n-1)!}\Delta^{n-1}f(x_1),$$

която има голямо приложение.

Глава III

Симетрични функции

1. **Прости симетрични функции.** Една функция на променливите $x_1, x_2, \ldots, x_n$ се нарича симетрична, ако не си изменя стойността, когато разместваме променливите помежду им по всевъзможни начини. Отначало ще разгледаме само цели рационални симетрични функции. Въобще симетричната функция не е хомогенна. Но лесно е да се види, че тя се разпада на сума от хомогенни симетрични функции.

Така функцията

$$x_1^2+x_2^2+x_1+x_2$$

е симетрична и сума от две хомогенни симетрични функции. Ние ще докажем нещо повече. Именно, че всяка симетрична функция е сума от тъй наречените прости симетрични функции. Под последното име разбираме хомогенна симетрична функция, в която във всеки член влизат еднакви показатели в степените. За да докажем горното предложение, нека

$$Cx_1^{a_1}\ x_2^{a_2}\ldots x_k^{a_k} \tag{1}$$

е един произволно взет член в дадената симетрична функция

$$S=f(x_1,\ x_2,\ldots,\ x_n).$$

Тогава, понеже S не трябва да си изменя стойността, когато променяме $x_1, x_2, \ldots, x_n$ помежду им по всевъзможни начини, функцията S трябва да има за членове всички такива, които се получават, като в (1) разместваме променливите $x_1, x_2, \ldots, x_n$, т. е. S съдържа простата симетрична функция

$$S'=C\sum x_1^{a_1}\ x_2^{a_2}\ldots x_k^{a_k}.$$

Но тогава $S_1=S-S'$ е пак симетрична функция, само че с по-малък брой членове от S. Продължавайки така, ясно е, че постепенно с изваждане на прости симетрични функции ще изчерпим всички членове в S, откъдето следва и предложението.

Според броя на променливите, които влизат във всеки член на простите симетрични функции, последните се делят на едноформени или степенни сборове, двуформени, триформени и т. н., когато съответно този брой или единица, или две, или три и т. н. Степенните сборове бележим с

$$S_i=x_1^i+x_2^i+\ldots+x_n^i=\sum x_1^i.$$

Двуформените функции имат вида

$$\sum x_1^{\alpha}\ x_2^{\beta}.$$

Ако $\alpha \neq \beta$, броят на членовете очевидно ще бъде равен на броя на вариациите от n елемента по два, т. е. на

$$n(n-1).$$

Ако $\alpha = \beta$, този брой се намалява наполовина, т. е. е $\frac{n(n-1)}{2}$. Триформената функция

$$\sum x_1^{\alpha} x_2^{\beta} x_3^{\gamma}$$

при $\alpha \neq \beta$, $\beta \neq \gamma$, $\beta \neq \alpha$ има $n(n-1)(n-2)$ члена. Ако два от показателите α, β, γ са равни помежду си, то тя има $\frac{n(n-1)(n-2)}{2}$ члена и най-сетне, ако $\alpha = \beta = \gamma$, броят на членовете е

$$\frac{n(n-1)(n-2)}{6}.$$

Изобщо, ако показателите в една m-формена симетрична функция са все различни помежду си, то броят на членовете ѝ ще бъде

$$n(n-1)\ldots(n-m+1).$$

Обаче, ако α_1 показатели са равни помежду си, α_2 други също са равни помежду си, но отлични от първите и т. н., то броят на членовете ще бъде равен на

$$\frac{n(n-1)\ldots(n-m+1)}{\alpha_1 ! \alpha_2 ! \alpha_3 ! \ldots}.$$

По-рано видяхме, че ако $x_1, x_2, \ldots, x_n$ са корените на уравнението

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0,$$

коефициентите му се изразяват посредством корените със следните релации:

$$x_1 + x_2 + \ldots + x_n = -\frac{a_1}{a_0},$$

$$x_1 x_2 + \ldots + x_{n-1} x_n = \frac{a_2}{a_0},$$

. . . . . .

$$x_1 x_2 \ldots x_n = (-1)^n \frac{a_n}{a_0}.$$

Левите части на тези релации са прости симетрични функции, които се наричат елементарни симетрични функции.

2. **Степенни сборове.** Една основна теорема в алгебрата е следната: Всяка рационална симетрична функция може да се изрази рационално чрез елементарните симетрични функции. Всяка цяла рационална симетрична функция

може да се изрази като цяла рационална функция от елементарните симетрични функции.

Отначало ще докажем валидността на тази теорема за степенните сборове, като изведем тъй наречените формули на Нютон. Нека

$$f(x)=x^n+a_1 x^{n-1}+a_2 x^{n-2}+\ldots+a_{n-1}x+a_n=0$$

е уравнение с корени $x_1, x_2, \ldots, x_n$ Да диференцираме релацията

$$f(x)=(x-x_1)(x-x_2)\ldots(x-x_n);$$

получаваме

$$f'(x)=(x-x_2)\ldots(x-x_n)+(x-x_1)(x-x_3)\ldots(x-x_n)+\ldots+$$
$$+(x-x_1)(x-x_2)\ldots(x-x_{n-1}).$$

Във всеки член от всички биномни множители, съответствуващи на корените, липсва по един такъв. Но това равенство може да се напише така:

$$f'(x)=\frac{f(x)}{x-x_1}+\frac{f(x)}{x-x_2}+\ldots+\frac{f(x)}{x-x_n}=\sum_{k=1}^{n}\frac{f(x)}{x-x_k}.$$

Общият член на тази сума преработваме така:

$$\frac{f(x)}{x-x_k}=\frac{f(x)-f(x_k)}{x-x_k}=$$
$$=\frac{1}{x-x_k}\,[x^n-x_k{}^n+a_1(x^{n-1}-x_k{}^{n-1})+a_2(x^{n-2}-x_k{}^{n-2})+\ldots+$$
$$+a_{n-1}(x-x_k)]=x^{n-1}+(x_k+a_1)x^{n-2}+(x_k{}^2+a_1x_k+a_2)x^{n-3}+\ldots+$$
$$+(x_k{}^{n-1}+a_1x_k{}^{n-2}+\ldots+a_{n-1}).$$

Ако сега поставим $k=1, 2, 3, \ldots, n$ и съберем, ще получим

$$f'(x)=nx^{n-1}+(S_1+na_1)x^{n-2}+(S_2+a_1S_1+na_2)x^{n-3}+\ldots+$$
$$+(S_{n-1}+a_1S_{n-2}+\ldots+a_{n-2}S_1+na_{n-1}).$$

От друга страна, имаме

$$f'(x)=nx^{n-1}+(n-1)a_1x^{n-2}+(n-2)a_2x^{n-3}+\ldots+a_{n-1}.$$

От сравнението на тези два израза за $f'(x)$ получаваме формулите на Нютон:

$$\text{(2)}\qquad\begin{aligned}
&S_1 &&+a_1=0,\\
&S_2 &&+a_1S_1 &&+2a_2=0,\\
&S_3 &&+a_1S_2 &&+a_2S_1 &&+3a_3=0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&S_i &&+a_1S_{i-1} &&+a_2S_{i-2}+\ldots+a_{i-1}S_1+ia_i=0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&S_{n-1}+a_1S_{n-2}+a_2S_{n-3}+\ldots+(n-1)a_{n-1}=0.
\end{aligned}$$

Можем лесно да получим S_n, $S_{n+1}, \ldots$, като в уравненията

$$f(x)=0,\ xf(x)=0,\ x^2 f(x)=0 \ldots$$

поставим $x=x_1, x_2, x_3, x_4, \ldots, x_n$ и съберем получените равенства. Така получаваме формули, които са сходни с (2) и могат да бъдат образувани като тях, като поставяме $a_{n+1}, a_{n+2}, \ldots$ равни на нула. Тези формули са:

$$\begin{aligned}
&S_n + a_1 S_{n-1} + a_2 S_{n-2} + \ldots + a_{n-1} S_1 + n\ a_n = 0,\\
(2')\qquad &S_{n+1} + a_1 S_n + a_2 S_{n-1} + \ldots + a_{n-1} S_2 + a_n S_1 = 0,\\
&S_{n+2} + a_1 S_{n-1} + a_2 S_n + \ldots + a_{n-1} S_3 + a_n S_2 = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

Ще забележа, че тези формули очевидно произтичат от (2), понеже анулирането на някои корени е еквивалентно с анулиране на коефициенти от края на лявата част на уравнението.

От (2) и (2′) можем лесно да пресметнем S_i:

$$\begin{aligned}
&S_1 = -a_1,\\
&S_2 = a_1^2 - 2a_2,\\
(3)\qquad &S_3 = -a_1^3 + 3a_1 a_2 - 3a_3,\\
&S_4 = a_1^4 - 4a_1^2 a_2 + 4a_1 a_3 + 2a_2^2 - 4a_4,\\
&\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

Можем да изразим S_i с детерминанта. Именно от (2) и (2′) имаме

$$(4)\qquad S_i = -\begin{vmatrix} 1 & 0 & 0 & \ldots & a_1 \\ a_1 & 1 & 0 & \ldots & 2a_2 \\ a_2 & a_1 & 1 & \ldots & 3a_3 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{i-1} & a_{i-2} & a_{i-3} & \ldots & ia_i \end{vmatrix}$$

От (3) или (4) следва, че степенните сборове са цели рационални функции на коефициентите на уравнението, в което $a_0=1$. Също е ясно, че коефициентите на тези цели рационални функции са цели числа.

3. **Друг метод за пресмятане на степенните сборове.** За пресмятане на степенните сборове можем да постъпим малко по-иначе. Нека всеки член от получената формула в предния параграф

$$\frac{f'(x)}{f(x)} = \frac{1}{x-x_1} + \frac{1}{x-x_2} + \ldots + \frac{1}{x-x_n}$$

развием по степените на $\frac{1}{x}$ при $|x|>|x_k|$, $k=1, 2, \ldots, n$,

$$\frac{1}{x-x_k}=\frac{1}{x\left(1-\frac{x_k}{x}\right)}=\frac{1}{x}+\frac{x_k}{x^2}+\frac{x_k^2}{x^3}+\ldots$$

получаваме

$$\frac{f'(x)}{f(x)}=\frac{n}{x}+\frac{S_1}{x^2}+\frac{S_2}{x^3}+\ldots$$

или

$$\frac{xf'(x)}{f(x)}=n+\frac{S_1}{x}+\frac{S_2}{x^2}+\frac{S_3}{x^3}+\ldots$$

Следователно, ако наредим частното $\frac{xf'(x)}{f(x)}$ по падащи степени на x, коефициентите дават степенните сборове. Ако умножим това тъждество с $f(x)$ и приравним коефициентите на степените на x в двете части, ще получим отново формулите (2) и (2′).

По този начин могат да се смятат също степенните сборове с отрицателен показател:

$$S_{-i}=\frac{1}{x_1^i}+\frac{1}{x_2^i}+\ldots+\frac{1}{x_k^i}.$$

Действително имаме

$$\frac{1}{x-x_k}=-\left(\frac{1}{x_k}+\frac{x}{x_k^2}+\frac{x^2}{x_k^3}+\ldots\right),$$

гдето редът е сходящ при $|x|$ по-малко от $|x_k|$. Следователно

$$-\frac{f'(x)}{f(x)}=S_{-1}+xS_{-2}+x^2S_{-3}+\ldots,$$

т. е. ако наредим $-\frac{f'(x)}{f(x)}$ по растящи степени на x, коефициентът на $x^{\lambda-1}$ е $S_{-\lambda}(\lambda>0)$.

Може S_{-i} да се сметнат и с формулите на Нютон. Действително това са степенни сборове с положителен показател за уравнението $f\left(\frac{1}{x}\right)=0$, на което корените са $\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}$.

4. Формули на Варинг за степенните сборове. Тези формули ни дават директни изрази на степенните сборове посредством коефициентите и обратно. Можем да ги получим чисто алгебрически, но значително се опростява извеждането, ако си послужим с развитие в редове. Нека уравнението

$$f(x)=x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0$$

има за корени $x_1, x_2, \ldots, x_n$. Тогава от

$$(x-x_1)(x-x_2)\ldots(x-x_n)=x^n+a_1x^{n-1}+\ldots+a_n$$

имаме

$$\log\left(1-\frac{x_1}{x}\right)+\log\left(1-\frac{x_2}{x}\right)+\ldots+\log\left(1-\frac{x_n}{x}\right)=$$

$$=\log\left(1+\frac{a_1}{x}+\ldots+\frac{a_n}{x^n}\right).$$

Нека вземем $|x|$ така голям, че $\left|\frac{x_i}{x}\right|<1$, и числото

$$\alpha=\frac{a_1}{x}+\frac{a_2}{x^2}+\ldots+\frac{a_n}{x^n}$$

да е по-малко от 1 по абсолютна стойност. Като развием в редове, ще имаме

$$-\frac{S_1}{x}-\frac{S_2}{2x^2}-\frac{S_3}{3x^3}-\ldots=\alpha-\frac{\alpha^2}{2}+\frac{\alpha^3}{3}-\frac{\alpha^4}{4}+\ldots$$

Ако вдясно развием члена

$$\frac{(-1)^{k+1}}{k}\left(\frac{a_1}{x}+\frac{a_2}{x^2}+\ldots+\frac{a_n}{x^n}\right)^k,$$

то общият член по формулата за бинома ще бъде

$$\frac{(\lambda_1+\lambda_2+\ldots+\lambda_n)!}{\lambda_1!\,\lambda_2!\ldots\lambda_n!}\;a_1^{\lambda_1}a_2^{\lambda_2}\ldots a_n^{\lambda_n}x^{-(\lambda_1+2\lambda_2+\ldots+n\lambda_n)},$$

гдето

$$\lambda_1+\lambda_2+\ldots+\lambda_n=k.$$

За да получим коефициента на x^{-p} вдясно, ще трябва да умножим този израз с $\frac{(-1)^{k+1}}{k}$ и сумираме за всички цели λ, положителни или нула, за които

$$\lambda_1+2\lambda_2+\ldots+n\lambda_n=p.$$

Така със сравняването коефициентите в двете части получаваме формулата на Варинг:

$$S_p=\sum(-1)^{\lambda_1+\ldots+\lambda_n}\frac{p\,(\lambda_1+\ldots+\lambda_n-1)!}{\lambda_1!\,\lambda_2!\ldots\lambda_n!}\;a_1^{\lambda_1}a_2^{\lambda_2}\ldots a_n^{n},$$

гдето сумирането е разпростряно върху всички цели положителни или нули числа λ, за които $\lambda_1+2\lambda_2+\ldots+n\lambda_i=p$.

Обратно, от уравнението

$$1+\frac{a_1}{x}+\ldots+\frac{a_n}{x^n}=e^{-\left(\frac{s_1}{x}+\frac{s_2}{2x^2}+\ldots\right)}=1-\frac{\beta}{1!}+\frac{\beta^2}{2!}-\frac{\beta^3}{3!}+\ldots,$$

гдето

$$\beta=\frac{S_1}{x}+\frac{S_2}{2x^2}+\frac{S_3}{3x^3}+\ldots,$$

по същия начин получаваме

$$a_p=\sum\frac{(-1)^{\mu_1+\ldots+\mu_p}}{\mu_1!\,\mu_2!\ldots\mu_p!}\left(\frac{S_1}{1}\right)^{\mu_1}\left(\frac{S_2}{2}\right)^{\mu_2}\ldots\left(\frac{S_p}{p}\right)^{\mu_p},$$

гдето сумирането е за

$$\mu_1+2\,\mu_2+\ldots+p\,\mu_p=p.$$

5. Пресмятане на простите симетрични функции. За да пресметнем двуформената функция

$$\sum x_1^{\alpha}x_2^{\beta},$$

умножаваме степенния сбор S_α с S_β. Ясно е, че при развитие на произведението $\sum x_1^{\alpha}\sum x_1^{\beta}$ ще се явят членове от форма $x_1^{\alpha+\beta}$ и $x_1^{\alpha}x_2^{\beta}$, т. е. ще имаме

$$\sum x_1^{\alpha}\sum x_2^{\beta}=\sum x_1^{\alpha+\beta}+\sum x_1^{\alpha}x_2^{\beta}.$$

Можем също да направим една лесна проверка с преброяване на членовете. Разгледаното произведение вляво има n^2 члена, а вдясно при $\alpha\neq\beta$ има също

$$n+n(n-1)=n^2.$$

От горната релация получаваме

$$\sum x_1^{\alpha}x_2^{\beta}=S_\alpha S_\beta-S_{\alpha+\beta}. \tag{5}$$

Ако $\alpha=\beta$, то вляво членовете стават два по два равни така, че

$$\sum x_1^{\alpha}x_2^{\alpha}=\frac{1}{2}\,(S_\alpha^{\,2}-S_{2\alpha}).$$

Триформените функции получаваме по подобен начин. Именно за да пресметнем $\sum x_1^{\alpha}x_2^{\beta}x_3^{\gamma}$, умножаваме двуформената функция $\sum x_1^{\alpha}x_2^{\beta}$ със степенния сбор $\sum x_1^{\gamma}$. При това умножение се получават членове от формата

$$x_1^{\alpha}x_2^{\beta}x_3^{\gamma},\quad x_1^{\alpha+\gamma}x_2^{\beta},\quad x_1^{\beta+\gamma}x_2^{\alpha},$$

т. е. ще имаме

$$\sum x_1^{\alpha}x_2^{\beta}\sum x_1^{\gamma}=\sum x_1^{\alpha}x_2^{\beta}x_3^{\gamma}+\sum x_1^{\alpha+\gamma}x_2^{\beta}+\sum x_1^{\alpha}x_2^{\beta+\gamma}. \tag{6}$$

Можем и да преброим членовете. Вляво имаме $n^2(n-1)$ члена ($\alpha\neq\beta$), а вдясно

$$n(n-1)\,(n-2)+2\,n(n-1)=n^2(n-1).$$

Уравнението (6) може на основание на (5) да се напише така:

$$(S_\alpha S_\beta - S_{\alpha+\beta}) S_\gamma = \sum x_1^{\alpha} x_2^{\beta} x_2^{\gamma} + S_{\alpha+\gamma} S_\beta - S_{\alpha+\beta+\gamma} + \\ + S_\alpha S_{\beta+\gamma} - S_{\alpha+\beta+\gamma},$$

отгдето имаме

$$\sum x_1^{\alpha} x_2^{\beta} x_3^{\gamma} = S_\alpha S_\beta S_\gamma - S_{\alpha+\gamma} S_\beta - S_{\alpha+\beta} S_\gamma - S_{\beta+\gamma} S_\alpha + 2 S_{\alpha+\beta+\gamma}$$

в случая, че $\alpha \neq \beta$, $\beta \neq \gamma$, $\alpha \neq \gamma$. Ако поставим два индекса от α, β, γ равни помежду си, напр. $\alpha = \beta \neq \gamma$, то ще имаме

$$\sum x_1^{\alpha} x_2^{\alpha} x_3^{\gamma} = \frac{1}{2} (S_\alpha^2 S_\gamma - 2 S_{\alpha+\gamma} S_\alpha - S_{2\alpha} S_\gamma + 2 S_{2\alpha+\gamma})$$

и при $\alpha = \beta = \gamma$

$$\sum x_1^{\alpha} x_2^{\alpha} x_3^{\alpha} = \frac{1}{6} (S_\alpha^3 - 3 S_\alpha S_{2\alpha} + 2 S_{3\alpha}).$$

Продължавайки така, можем да изразим коя да е проста симетрична функция като цяла рационална функция на степенните сборове. Понеже последните са цели рационални функции от коефициентите на уравнението, то така получаваме, че простите симетрични функции са цели рационални функции от коефициентите ($a_0 = 1$).

Оттук веднага следва изказаната теорема, че всяка цяла рационална и симетрична функция от корените на уравнението е цяла рационална функция от коефициентите му при $a_0 = 1$.

6. **Теорема на Бриоски и Кейли за степента и теглото на симетричните функции.** Нека е дадено уравнението

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n = 0,$$

на което корените да бъдат $x_1, x_2, \ldots, x_n$. Ако

$$S = \varphi(x_1, x_2, \ldots, x_n)$$

е цяла рационална симетрична функция от корените на уравнението, то, както вече видяхме, тя се изразява като цяла рационална функция от коефициентите му. Следователно можем да пишем

$$S = \sum C a_1^{\alpha_1} a_2^{\alpha_2}, \ldots, a_n^{\alpha_n},$$

гдето числата α_i са положителни цели числа, като някои от тях могат да са равни на нула. Този израз ще наричаме коефициентен израз или коефициентна функция. Числото

$$\alpha_1 + 2\alpha_2 + \ldots + n\alpha_n$$

се нарича тегло на написания член. Ще докажем следната теорема на Бриоски и Кейли. Степента на коефициентната функция е равна на най-големия показател, в който влиза

всеки корен в симетричната функция, а в случай, че последната функция е хомогенна, всички членове на коефициентната функция имат едно и също тегло, равно на степента на симетричната функция.

Действително от връзките между корените и коефициентите на уравнението се вижда, че всеки коефициент от втория нататък е линейна функция на всеки корен, например x_1, или можем да пишем

$$a_i = M_i x_1 + N_i, \quad i = 1, 2, \ldots, n,$$

гдето M_i, N_i не зависят от x_1. Но тогава от (7) имаме

$$\varphi(x_1, x_2, \ldots, x_n) =$$

$$= \sum C (M_1 x_1 + N_1)^{\alpha_1} (M_2 x_1 + N_2)^{\alpha_2} \ldots (M_n x_1 + N_n)^{\alpha_n}.$$

Степента на x_1 в написания член вдясно е

$$\alpha_1 + \alpha_2 + \ldots + \alpha_n,$$

т. е. равна на степента на самия член. Най-голямото число от тези числа представлява степента на коефициентната функция и е най-голямата степен, в която влиза x_1 в симетричната функция. Първата част на теоремата така е доказана.

За да докажем втората част, постъпваме така: трансформираме даденото уравнение, като умножим корените му с λ. Новото уравнение ще има корени:

$$y_1 = \lambda x_1, \quad y_2 = \lambda x_2, \ldots, y_n = \lambda x_n,$$

или изобщо $y = \lambda x$, гдето x е кой да е корен на даденото уравнение, а y — съответният му корен на трансформираното. Трябва значи вместо x да поставим равното му $\frac{y}{\lambda}$ и така за y получаваме

$$(8) \qquad y^n + a_1 \lambda y^{n-1} + a_2 \lambda^2 y^{n-2} + \ldots + a_{n-1} \lambda^{n-1} y + a_n \lambda^n = 0.$$

Ако степента на хомогенната симетрична функция е m, то

$$\varphi(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) = \lambda^m \varphi(x_1, x_2, \ldots, x_n).$$

Уравнението (7), приложено за (8), дава

$$\varphi(y_1, y_2, \ldots, y_n) = \sum C (a_1 \lambda)^{\alpha_1} (a_2 \lambda^2)^{\alpha_2} \ldots (a_n \lambda^n)^{\alpha_n}.$$

Понеже

$$\varphi(y_1, y_2, \ldots, y_n) = \varphi(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) =$$

$$= \lambda^m \varphi(x_1, x_2, \ldots, x_n) = \lambda^m \sum C a_1^{\alpha_1} a_2^{\alpha_2} \ldots a_n^{\alpha_n},$$

то трябва да имаме за всяко λ

$$\lambda^m \sum C a_1^{\alpha_1} a_2^{\alpha_2} \ldots a_n^{\alpha_n} = \sum C \lambda^{\alpha_1 + 2\alpha_2 + \ldots + n\alpha_n} a_1^{\alpha_1} \ldots a_n^{\alpha_n}.$$

Двете части съдържат само цели положителни степени на λ, които очевидно трябва да бъдат равни помежду си, т. е.

$$\alpha_1+2\alpha_2+\ldots+n\alpha_n=m,$$

което трябваше да се докаже.

Тези теореми ни дават възможност да пресмятаме симетричните функции. Така например нека е дадена симетричната функция

$$E=\sum x_1^3 x_2.$$

Коефициентният ѝ израз ще бъде от степен 3 и тегло 4. Така че той ще има формата

$$E=Aa_1^2 a_2+Ba_1 a_3+Ca_2^2+Da_4,$$

гдето A, B, C, D са числа, подлежащи на определяне (ако уравнението е от по-ниска степен от 4, то поставяме $a_4=0$ или $a_4=a_3=0$ и т. н.). За уравнението

$$x^2-1=0$$

имаме

$$x_1=1,\quad x_2=-1,\quad E=-2,$$

отгдето получаваме

$$-2=C.$$

Уравнението

$$(x-1)^2=x^2-2x+1=0,$$

на което корените са $x_1=x_2=1$, дава

$$2=4A+C,\quad A=1.$$

От уравнението

$$(x-1)^3=x^3-3x^2+3x-1=0,$$

$x_1=x_2=x_3=1$ имаме

$$6=27A+3B+9C,\ B=-1;$$

и най-сетне от

$$(x-1)^4=x^4-4x^3+6x^2-4x+1=0,$$

$x_1=x_2=x_3=x_4=1$ получаваме

$$12=96A+16B+36C+D,\quad D=4,$$

така че

$$\sum x_1^3 x_2=a_1^2 a_2-a_1a_3-2a_2^2+4a_4.$$

7. Метод на Коши за пресмятане на симетричните функции. Един друг метод за пресмятане на симетричните функции, който едновременно дава и главната теорема за тези функции, е методът на Коши.

Нека

$$f(x)=x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0$$

е алгебрично уравнение с корени $x_1, x_2, \ldots, x_n$. Нека $U=U(x_1, x_2, \ldots, x_n)$ е симетрична функция от корените на уравнението. Да допуснем, че по някакъв начин сме елиминирали корените $x_2, x_3, \ldots, x_n$ така, че U се е обърнала в една цяла рационална функция на корена x_1:

$$U=\varphi(x_1)=p_0 x_1^m+p_1 x_1^{m-1}+\ldots+p_{m-1}x_1+p_m.$$

Да разделим $\varphi(x)$ с $f(x)$ и нека Q е частното от делението, а

$$R(x)=q_0x^{n-1}+q_1 x^{n-2}+\ldots+q_{n-2}x+q_{n-1}$$

е остатъкът от делението, т. е.

$$\varphi(x)=f(x)Q+R(x).$$

При $x=x_k (k=1, 2, \ldots, n)$ имаме поради $f(x_k)=0$

$$\varphi(x_k)=R(x_k).$$

Понеже функцията U е симетрична, то

$$\varphi(x_1)=\varphi(x_2)=\ldots=\varphi(x_n),$$

така че уравнението

$$R(x)-U=q_0x^{n-1}+q_1 x^{n-2}+\ldots+q_{n-2}x+(q_{n-1}-U)=0$$

ще има n корена $x_1, x_2, \ldots, x_n$. Следователно всичките му коефициенти са равни на нула, т. е.

$$q_0=q_1=\ldots=q_{n-2},\ q_{n-1}=U.$$

Значи, ако разделим $\varphi(x)$ с $f(x)$, остатъкът не ще зависи от x и ще дава стойността на симетричната функция. Именно на това се базира методът на Коши. Да образуваме уравненията

$$\frac{f(x)}{x-x_1}=f_1(x)=0, \frac{f_1(x)}{x-x_2}=f_2(x)=0, \ldots, \frac{f_{n-2}(x)}{x-x_{n-1}}=f_{n-1}(x)=0,$$

на които корените съответно са $x_2, x_3, \ldots, x_n$ на първото, $x_3, x_4, \ldots, x_n$ на второто и т. н. най-сетне x_n на последното. От последното очевидно ще имаме

$$x_n=-a_1-x_1-x_2-\ldots-x_{n-1}.$$

Като заместим тази стойност на x_n в израза U на симетричната функция, за която предполагаме, че е цяла рационална, ще останат да фигурират само корените

$$x_1, x_2, \ldots\ x_{n-1},$$

т. е.

$$U=U(x_1, x_2, \ldots, x_{n-2}, x_{n-1}).$$

Функцията U е симетрична спрямо корените x_{n-1}, x_n на уравнението $f_{n-2}(x)=0$, коефициентите на което съдържат $x_1, x_2, \ldots, x_{n-2}$, и U

съдържа само единия от тях. Следователно според горната лема, ако разделим

$$U(x_1, x_2, \ldots, x_{n-2}, x)$$

с $f_{n-2}(x)$, остатъкът няма да зависи от x и ще дава стойността на U, т. е. ще имаме

$$U = U(x_1, x_2, \ldots, x_{n-3}, x_{n-2}),$$

която е пак симетрична спрямо корените на уравнението

$$f_{n-3}(x) = 0$$

и съдържа само един от тях x_{n-2}. Коефициентите на това уравнение са цели рационални функции на корените $x_1, x_2, \ldots, x_{n-3}$. По лемата, ако разделим

$$U(x_1, x_2, \ldots, x_{n-3}, x)$$

с $f_{n-3}(x)$, остатъкът няма да зависи от x и ще даде

$$U = U(x_1, x_2, \ldots, x_{n-3}).$$

Ясно е, че продължавайки така, ще можем да елиминираме всички корени в U.

Понеже при деленията коефициентът пред най-високата степен на x е равен на единица, то следва главната теорема, че всяка цяла рационална симетрична функция от корените на уравнението е цяла рационална функция на коефициентите.

Като пример да разгледаме симетричната функция

$$U = (x_1 + x_2)(x_1 + x_3)(x_2 + x_3)$$

от корените на уравнението

$$f(x) = x^3 + a_1 x^2 + a_2 x + a_3 = 0.$$

Тогава имаме

$$\frac{f(x)}{x - x_1} = f_1(x) = x^2 + (x_1 + a_1)x + x_1^2 + a_1 x_1 + a_2,$$

$f_2(x) = \frac{f_1(x)}{x - x_2} = x + x_1 + x_2 + a_1 . f_1(x)$ има за корени $x_2, x_3, f_2(x) = 0$ има корена $x_3 = -(x_1 + x_2 + a)$.

Като заместим x_3 в U, получаваме

$$U = (x_1 + x_2)(a_1 + x_1)(a_1 + x_2) = (a_1 + x_1)[x_2^2 + (a_1 + x_1)x_2 + a_1 x_1].$$

Ако разделим с $f_1(x_2)$, остатъкът дава U и не зависи от x_2:

$$U = -x_1^3 - a_1 x_1^2 - a_2 x_1 - a_1 a_2.$$

Остатъкът от делението на U с $f(x_1)$ дава търсения израз за U:

$$U=a_3-a_1a_2.$$

8 Друг начин за пресмятане на симетричните функции. Една произволна цяла рационална функция (полином) $f(x_1, x_2,\ldots, x_n)$ от променливите $x_1, x_2,\ldots, x_n$ е сума от членове от вида

$$ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n},$$

гдето $\alpha_1, \alpha_2,\ldots, \alpha_n$ са цели неотрицателни числа. След привеждане на подобните членове (т. е. тези с еднакви показатели α_i) тя приема тъй наречения нормален вид. Ще разгледаме сега едно нареждане на членовете ѝ, което ще използуваме тутакси. Едно произведение от степени $x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$ наричаме „по-високо“ от друго $x_1^{\beta_1}x_2^{\beta_2}\ldots x_n^{\beta_n}$, ако първата от разликите $\alpha_1-\beta_1$, α_2 $\beta_2,\ldots,\alpha_n-\beta_n$, която е отлична от нула, е положителна. Нека в нормалното представяне на полинома членовете са така наредени, че съответствуващите степенни произведения да следват на намаляваща височина. Тогава стоящият на първо място член $ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$ се нарича най-висок член на полинома $f(x_1, x_2,\ldots, x_n)$. Лесно се вижда, че най-високият член в произведението от полиномите $f(x_1, x_2,\ldots, x_n)$ и $\varphi(x_1, x_2,\ldots, x_n)$ е равен на произведението от техните най-високи членове.

Нека сега полиномът $f(x_1, x_2,\ldots, x_n)$ е симетрична функция на $x_1, x_2,\ldots, x_n$, на които променливи да означим с $A_1, A_2,\ldots, A_n$ елементарните симетрични функции. Нека $ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$ е най-високият член в $f(x_1, x_2,\ldots, x_n)$. Тогава ще имаме

$$\alpha_1\geqq\alpha_2\geqq\alpha_3\geqq\ldots\geqq\alpha_n.$$

Действително, ако например $\alpha_2>\alpha_1$, то функцията би съдържала члена $ax_1^{\alpha_2}x_2^{\alpha_1}\ldots x_n^{\alpha_n}$, който е по-висок от члена $ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$. От друга страна, $A_1^{\alpha_1-\alpha_2}A_2^{\alpha_2-\alpha_3}\ldots A_{n-1}^{\alpha_{n-1}-\alpha_n}A_n^{\alpha_n}$ е рационална симетрична функция на $x_1, x_2,\ldots, x_n$ с най-висок член $ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$. Тогава функцията

$$f_1(x_1, x_2,\ldots, x_n)=f(x_1, x_2,\ldots, x_n)-$$

$$-aA_1^{\alpha_1-\alpha_2}A_2^{\alpha_2-\alpha_3}\ldots A_{n-1}^{\alpha_{n-1}-\alpha_n}A_n^{\alpha_n}$$

ще бъде цяла рационална симетрична, на която най-високият член $bx_1^{\beta_1}x_2^{\beta_2}\ldots x_n^{\beta_n}$ ще бъде по-нисък от члена $ax_1^{\alpha_1}x^{\alpha_2}\ldots x_n^{\alpha_n}$. Подобно функцията

$$f_2(x_1, x_2,\ldots, x_n)=f_1(x_1, x_2,\ldots, x_n)-$$

$$-bA_1^{\beta_1-\beta_2}A_2^{\beta_2-\beta_3}\ldots A_{n-1}^{\beta_{n-1}-\beta_n}A_n^{\beta_n}$$

ще бъде с по-нисък най-висок член. Продължавайки така, понеже очевидно функцията $f(x_1,\ldots, x_n)$ има краен брой членове, ще получим търсеното представяне на дадената функция във форма на полином $P(A_1, A_2,\ldots, A_q)$ от елементарните симетрични функции $A_1, A_2,\ldots, A_n$,

което е и едно ново доказателство на тази основна теорема за симетричните функции.

Накрая ще направим забележката, че представянето на функцията $f(x_1, x_2, \ldots, x_n)$ посредством на A_1, $A_2, \ldots, A_n$ е еднозначно. Действително, ако имаме

$$f(x_1, x_2, \ldots, x_n) = P(A_1, A_2, \ldots, A_n) = Q(A_1, A_2, \ldots, A_n),$$

то би следвало, че

$$R(A_1, A_2, \ldots, A_n) = P(A_1, A_2, \ldots, A_n) -$$

$$-Q(A_1, A_2, \ldots, A_n) = R_1(x_1, x_2, \ldots, x_n)$$

е идентично равен на нула спрямо неизвестните $x_1, x_2, \ldots, x_n$. Но тогава и $R(A_1, A_2 \ldots, A_n)$ относно неизвестните $A_1, A_2, \ldots, A_n$ трябва да бъде идентично равен на нула, понеже, както лесно се вижда, две различни произведения на степени $A_1^{p_1} A_2^{p_2} \ldots A_n^{p_n}$ и $A_1^{q_1} A_2^{q_2} \ldots A_n^{q_n}$ имат като полиноми на $x_1, x_2, \ldots, x_n$ различни най-високи членове.

9. Дробни рационални симетрични функции. Нека

$$f(x_1, x_2, \ldots, x_n) = \frac{P}{Q}$$

е една дробна рационална симетрична функция. P и Q са полиноми на променливите. Ние ще покажем, че такава функция може да се представи като отношение на две цели рационални симетрични функции. От факта, че f е симетрична функция, не следва, че P и Q са също симетрични. Така

$$\frac{x_1^2 - x_2^2}{x_1^3 - x_2^3}$$

е симетрична, но нито числителят, нито знаменателят са симетрични. Обаче ако съкратим множителя $x_1 - x_4$, тя става отношение на две симетрични:

$$\frac{x_1 + x_2}{x_1^2 + x_2^2 + x_1 x_2}.$$

Ако P или Q е симетрична функция, то другата от тях очевидно е симетрична. Нека допуснем, че Q не е симетрична и нека $Q, Q_1, Q_2, \ldots, Q_{p-1}$ са стойностите, които взема тя, когато върху $x_1, x_2, \ldots, x_n$ извършим всички възможни разместявания. Да умножим числителя и знаменателя на f с $Q_1, Q_2, \ldots, Q_{p-1}$, т. е.

$$f(x_1, x_2, \ldots, x_n) = \frac{PQ_1Q_2 \ldots Q_{p-1}}{QQ_1Q_2 \ldots Q_{p-1}}.$$

Знаменателят $QQ_1 \ldots Q_{p-1}$ е очевидно симетрична функция. Но тогава веднага следва, че и числителят

$$PQ_1 Q_2 \ldots Q_{p-1}$$

трябва да бъде симетрична функция, с което предложението е доказано.

10. Рационални функции от корените на уравнението. Нека $f(x)=0$ е алгебрическо уравнение с корени $x_1, x_2, \ldots, x_n$ и нека $v=\frac{\varphi(x_1)}{\psi(x_1)}$ е една дробна рационална функция от един корен на уравнението Като умножим числителя и знаменателя ѝ с $\psi(x_2)\ldots\psi(x_n)$, ще имаме

$$v=\frac{\varphi(x_1)\psi(x_2)\psi(x_3)\ldots\psi(x_n)}{\psi(x_1)\psi(x_2)\psi(x_3)\ldots\psi(x_n)}.$$

Произведението $\psi(x_1)\psi(x_2)\ldots\psi(x_n)$ е симетрична функция на корените за уравнението $f(x)=0$, следователно е известно като рационална функция от коефициентите му. Произведението $\psi(x_2)\ldots\psi(x_n)$ е цяла рационална симетрична функция от корените на уравнението

$$\frac{f(x)}{x-x_1}=0, \tag{9}$$

коефициентите на което са цели рационални функции на x_1. Следователно това произведение като цяла рационална функция на коефициентите на уравнението (9) ще бъде цяла рационална функция на x_1, т. е.

$$v=v(x_1)=r_0x_1^m+r_1x_1^{m-1}+\ldots+r_m.$$

Ако $m\geqq n$, разделяме $v(x)$ с $f(x)$ и нека $f_1(x)$ е остатъкът от делението, а $Q(x)$ частното:

$$v(x)=f(x)Q(x)+f_1(x).$$

При $x=x_1$, понеже $f(x_1)=0$, имаме $v(x_1)=f_1(x_1)$, гдето $f_1(x)$ е полином най-много от $(n-1)$-ва степен. Така получаваме следния резултат: всяка дробна рационална функция от един корен на едно уравнение може да се представи като цяла рационална функция от същия корен, степента на която е най-много с единица по-малка от степента на уравнението. По-общо нека

$$u=\frac{\varphi(x_1, x_2, \ldots, x_k)}{\psi(x_1, x_2, \ldots, x_k)}$$

е дробна рационална функция на няколко корена на

$$f(x)=0.$$

По изложения метод тя може да се обърне в цяла рационална на един корен, например x_1,

$$u=p_0x_1^m+p_1x_1^{m-1}+\ldots+p_m,$$

гдето коефициентите ще бъдат изобщо дробни рационални функции на $x_2, x_3, \ldots, x_k$. По същия начин $p_0, p_1, \ldots, p_m$ можем да представим като цели рационални функции на x_2 с коефициенти, които са изобщо дробни рационални функции на $x_3, x_4, \ldots, x_k$. Като продължаваме така, очевидно можем да представим u като цяла рационална функция на $x_1, x_2, \ldots, x_k$.

Глава IV

Елиминация

1. Елиминация посредством симетрични функции. Дадени са две уравнения:

$$(1)\qquad \begin{cases} f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n=0, \\ \varphi(x)=b_0x^m+b_1x^{m-1}+b_2x^{m-2}+\ldots+b_m=0, \end{cases}$$

гдето a_i, b_k са произволни реални или комплексни числа. Ако тези уравнения имат общ корен, то между коефициентите им трябва да съществува някоя връзка. Търсенето на тази връзка е предмет на елиминацията. Нека с $\alpha_1, \alpha_2, \ldots, \alpha_n$ означим корените на първото уравнение и с $\beta_1, \beta_2, \ldots, \beta_m$ — корените на второто. Тогава да разгледаме произведението

$$\begin{aligned} \Pi=&(\alpha_1-\beta_1)(\alpha_1-\beta_2)\ldots(\alpha_1-\beta_m). \\ &(\alpha_2-\beta_1)(\alpha_2-\beta_2)\ldots(\alpha_n-\beta_m). \\ &\ldots\ldots\ldots\ldots \\ &(\alpha_n-\beta_1)(\alpha_n-\beta_2)\ldots(\alpha_n-\beta_m). \end{aligned}$$

Ако уравненията имат един общ корен, т. е. някой α_i е равно на някой β_k, то очевидно $\Pi=0$. Обратно, ако е изпълнено последното условие, някоя разлика трябва да бъде равна на нула, т. е. уравненията ще имат общ корен. Следователно анулирането на Π е еквивалентно, двете уравнения да имат общ корен. Но Π, както лесно се вижда, е цяла рационална симетрична функция на корените на първото и второто уравнение от степен m спрямо корените α и степен n спрямо β. Следователно Π се изразява като цяла рационална функция от $\frac{a_1}{a_0}, \ldots, \frac{a_n}{a_0}$ от степен m и от $\frac{b_1}{b_0}, \ldots, \frac{b_m}{b_0}$ от степен n. Тогава

$$R=a_0^m b_0^n\ \Pi$$

ще представлява цяла рационална функция от всички коефициенти на уравненията (1). Изразът R, анулирането на който е еквивалентно на условието двете уравнения (1) да имат общ корен, се нарича *резултанта или елиминанта*.

Така виждаме, че елиминантата е цяла рационална функция от коефициентите на уравненията, степента на която спрямо коефициентите на първото уравнение е равна на степента на второто и спрямо коефициентите на второто е равна на степента на първото уравнение. Сега ще установим, че *елиминантата е изобарна функция с тегло, равно на произведението на степените на двете уравнения*. Действително нека

$$(2)\qquad R=\sum C a_0^{\lambda_0} a_1^{\lambda_1} a_2^{\lambda_2}\ldots b_0^{\mu_0} b^{\mu_1} b^{\mu_2}\ldots$$

и да трансформираме дадените уравнения (1), като умножим корените им с едно произволно число k.

Новите уравнения ще бъдат

$$a_0x^n+a_1kx^{n-1}+a_2k^2x^{n-2}+\ldots+a_n k^n=0,$$

$$b_0x^m+b_1kx^{m-1}+b_2k^2x^{m-2}+\ldots+b_mk^m=0.$$

Произведението П се умножава с множителя k^{mn}, така че за (2) на новите уравнения ще имаме

$$k^{mn}\sum Ca_0^{\lambda_0}a_1^{\lambda_1}a_2^{\lambda_2}\ldots b_0^{\mu_0}b_1^{\mu_1}b_2^{\mu_2}\ldots=$$

$$=\sum Ck^{\lambda_1+2\lambda_2+\ldots+\mu_1+2\mu_2+\ldots}a_0^{\lambda_0}a_1^{\lambda_1}a_2^{\lambda_2}\ldots b_0^{\mu_0}a_1^{\mu_1}a_2^{\mu_2}\ldots,$$

отгдето следва, че $\lambda_1+2\lambda_2+\ldots+\mu_1+2\mu_2+\ldots=mn$, което установява изказаната теорема за теглото.

На елиминантата R можем да дадем друга форма. Именно като вземем предвид, че

$$f(x)=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

$$\varphi(x)=b_0(x-\beta_1)(x-\beta_2)\ldots(x-\beta_m),$$

получаваме

$$(3)\qquad R=a_0^m\varphi(\alpha_1)\varphi(\alpha_2)\ldots\varphi(\alpha_n)=(-1)^{mn}b_0^n f(\beta_1)f(\beta_2)\ldots f(\beta_m).$$

От тези изрази отново лесно се вижда горното свойство за степента и теглото на R.

За пример нека

$$f(x)=a_0x^2+a_1x+a_2=0,$$

$$\varphi(x)=b_0x^2+b_1x+b_2=0$$

са две дадени квадратни уравнения. Тогава, ако α_1, α_2 са корените на първото уравнение, ще имаме по (3)

$$R=a_0^2\varphi(\alpha_1)\varphi(\alpha_2)=$$

$$=a_0^2[b_0^2\alpha_1^2\alpha_2^2+b_0b_1\alpha_1\alpha_2(\alpha_1+\alpha_2)+b_0b_2(\alpha_1^2+\alpha_2^2)+$$

$$+b_1^2\alpha_1\alpha_2+b_1b_2(\alpha_1+\alpha_2)+b_2^2].$$

Или като пресметнем фигуриращите симетрични функции на α_1, α_2, получаваме

$$R=a_2^2b_0^2-a_1a_2b_0b_1+b_0b_2(a_1^2-2a_0a_2)+$$

$$+a_0a_2b_1^2-a_0a_1b_1b_2+a_0^2b_2^2=$$

$$=(a_2b_0-a_0b_2)^2+(a_1b_2-a_2b_1)(a_1b_0-a_0b_1).$$

Случай на няколко общи корена. Видяхме, че ако за двете уравнения (1) означим с R израза

$$R=a_0^mb_0^n\Pi(\alpha_i-b_k),\quad i=1,2,\ldots,n;\quad k=1,2,\ldots,m,$$

то анулирането на R дава условието уравненията да имат общ корен. Да предположим, че $R=0$ и нека един общ корен да означим с x_1, т. е. $x_1=\alpha_1=\beta_1$.

Нека диференцираме израза

$$R=a_0{}^m\varphi(\alpha_1)\varphi(\alpha_2)\ldots\varphi(\alpha_n)$$

последователно, като гледаме на коефициентите на второто уравнение $\varphi(x)=0$ като на независими променливи. Като вземем под внимание, че

$$\varphi(\alpha_k)=b_0\alpha_k{}^m+b_1\alpha_k{}^{m-1}+\ldots+b_i\alpha_k{}^{m-i}+\ldots+b_m,$$

то

$$\frac{\partial\varphi(\alpha_k)}{\partial b_i}=\alpha_k{}^{m-i};$$

ще имаме по правилата за диференциране

$$\frac{\partial R}{\partial b_i}=a_0{}^m\alpha_1{}^{m-i}\varphi(\alpha_2)\ldots\varphi(\alpha_n)+a_0{}^m\alpha_2{}^{m-1}\varphi(\alpha_1)\varphi(\alpha_3)\ldots\varphi(\alpha_n)+$$
$$+\ldots+a_0{}^m\varphi(\alpha_1)\ldots\varphi(\alpha_{n-1})\alpha_n{}^{m-i}.$$

Понеже $\alpha_1=x_1$ е общ корен, то $\varphi(\alpha_1)=0$ и следователно

$$\frac{\partial R}{\partial b_i}=a_0{}^m x_1{}^{m-i}\varphi(\alpha_2)\ldots\varphi(\alpha_n). \tag{4}$$

Частната производна спрямо b_{i-1} ще бъде

$$\frac{\partial R}{\partial b_{i-1}}=a_0{}^m x_1{}^{m-i+1}\varphi(\alpha_2)\ldots\varphi(\alpha_n).$$

От тези две равенства получаваме

$$x_1=\frac{\partial R}{\partial b_{i-1}}:\frac{\partial R}{\partial b_i},$$

т. е. общият корен се дава от уравнението

$$\frac{\partial R}{\partial b_{i-1}}\,x-\frac{\partial R}{\partial b_i}=0$$

или символично

$$\left(\frac{\partial}{\partial b_{i-1}}\,x-\frac{\partial}{\partial b_i}\right)R=0. \tag{5}$$

Когато $R=0$, видяхме, че уравненията имат поне един общ корен x_1. За да имат два общи корена, нужно е допълнително условие. Така от (4) се вижда, че ако имат втори общ корен $x_2=\alpha_2$, то понеже $\varphi(\alpha_2)=0$, частната производна $\frac{\partial R}{\partial b_i}=0$. И обратно, ако това условие е изпълнено, то от същото уравнение следва, че поне единият множител във

$\varphi(\alpha_2),\ldots,\varphi(\alpha_n)$ е равен на нула, т. е. системата (1) има поне още един общ корен. Така че получаваме:

Необходимо и достатъчно условие, щото уравненията да имат само един общ корен, е

$$R=0,\ \frac{\partial R}{\partial b_i}\neq 0$$

за един кой да е индекс i от 0 до m. Общият корен се дава от уравнението (5). Ако освен $R=0$ и $\frac{\partial R}{\partial b_i}=0$, то уравненията (1) имат най-малко два общи корена.

Лесно е да намерим условия за повече общи корени. За тази цел диференцираме $\frac{\partial R}{\partial b_i}$ пак частно спрямо b_i и b_{i-1}. Получаваме

$$\frac{\partial^2 R}{\partial b_i^2}=$$

$$=a_0^m\left[2\alpha_1^{m-i}\alpha_2^{m-i}\varphi(\alpha_3)\ldots\varphi(\alpha_n)+2\alpha_1^{m-i}\alpha_3^{m-i}\varphi(\alpha_2)\ldots\varphi(\alpha_n)+\ldots\right].$$

Ако (1) имат поне два общи корена $\alpha_1=x_1$, $\alpha_2=x_2$, понеже $\varphi(\alpha_1)=\varphi(\alpha_2)=0$,

$$\frac{\partial^2 R}{\partial b_i^2}=2a_0^m x_1^{m-i}x_2^{m-i}\varphi(\alpha_3)\ldots\varphi(a_n). \tag{6}$$

От този израз се вижда, че ако уравненията (1) имат само два общи корена, то $\frac{\partial^2 R}{\partial b_i^2}\neq 0$. Ако обаче имат повече общи корени, то трябва тая производна да бъде равна на нула.

Да допуснем, че $\frac{\partial^2 R}{\partial b_i^2}\neq 0$; търсим производната

$$\frac{\partial^2 R}{\partial b_i\partial b_{i-1}}= \tag{7}$$

$$=a_0^m\alpha_2^{m-i}\alpha_1^{m-i+1}\varphi(\alpha_3)\ldots\varphi(\alpha_n)+a_0^m\alpha_1^{m-i+1}\alpha_2^{m-i}\varphi(\alpha_3)\ldots\varphi(a_n)+\ldots$$

$$+a_0^m\varphi(\alpha_1)\ldots\varphi(\alpha_{n-2})\alpha_{n-i}^{m-i}\alpha_n^{m-i+1},$$

която, понеже $\varphi(\alpha_1)=\varphi(\alpha_2)=0$, $\alpha_1=x_1$, $\alpha_2=x_2$, се свежда на

$$\frac{\partial^2 R}{\partial b_{i-1}\partial b_i}=a_0^m(x_1x_2)^{m-i}(x_1+x_2)\varphi(\alpha_3)\ldots\varphi(\alpha_n).$$

От (6), като вместо i поставим $i-1$, имаме

$$\frac{\partial^2 R}{\partial b_{i-1}^2}=2a_0^m x_1^{m-i+1}x_2^{m-i+1}\varphi(\alpha_3)\ldots\varphi(\alpha_n). \tag{8}$$

От формулите (6), (7), (8) лесно се получава квадратното уравнение, което дава общите корени x_1, x_2. Действително, ако разделим (8) на (6), получаваме

$$x_1 x_2 = \frac{\partial^2 R}{\partial b_{i-1}^2} : \frac{\partial^2 R}{\partial b_i^2}.$$

Ако разделим (7) на (6), получаваме

$$x_1 + x_2 = 2 \frac{\partial^2 R}{\partial b_i \partial b_{i-1}} : \frac{\partial^2 R}{\partial b_i^2}.$$

Тези две релации показват, че x_1 и x_2 са корени на уравнението

$$\frac{\partial^2 R}{\partial b_i^2} x^2 - 2 \frac{\partial^2 R}{\partial b_i \partial b_{i-1}} x + \frac{\partial^2 R}{\partial b_{i-1}^2} = 0,$$

което символично може да се напише така:

$$\left(\frac{\partial}{\partial b_i} x - \frac{\partial}{\partial b_{i-1}}\right)^2 R = 0. \tag{9}$$

Така доказахме, че за да имат уравненията (1) само два общи корена, необходимо и достатъчно е за кое да е b_i да имаме

$$R = 0, \ \frac{\partial R}{\partial b_i} = 0, \ \frac{\partial^2 R}{\partial b_i^2} \neq 0.$$

Общите корени се дават с уравнението (9).

Продължавайки така, лесно се установява общата теорема: Необходимо и достатъчно условие уравненията (1) да имат k общи корена се състои в равенствата

$$R = 0, \ \frac{\partial R}{\partial b_i} = 0, \ \frac{\partial^2 R}{\partial b_i^2} = 0, \ldots, \frac{\partial^{k-1} R}{\partial b_i^{k-1}} = 0, \ \frac{\partial^k R}{\partial b_i^k} \neq 0.$$

Общите корени се дават от уравнението

$$\left(\frac{\partial}{\partial b_i} x - \frac{\partial}{\partial b_{i-1}}\right)^k R = 0.$$

От изразите (3) на резултантата R е очевидно, че тя съдържа членове от вида

$$a_0^m b_0^n \alpha_1^m \alpha_2^m \ldots \alpha_n^m, \quad (-1)^{mn} b_0^n a_0^m \beta_1^n \beta_2^n \ldots \beta_m^n,$$

т. е. можем да пишем

$$R = (-1)^{mn} a_n^m b_0^n + a_0^m b_m^n + \ldots$$

От този израз се вижда, че R не съдържа като множители нито a_0, нито b_0. Лесно е да се види, че R е неразложима, т. е. не

се разлага на произведение на цели рационални функции на коефициентите

$$a_0, a_1, \ldots, a_n \quad \text{и} \quad b_0, b_1, \ldots, b_m.$$

Действително, ако допуснем обратното, то ще имаме поне един неразложим множител, който ще се анулира например, когато $\alpha_1 = \beta_1$. Но тогава при възстановяване на α и β ще трябва този множител да се дели на $\alpha_1 - \beta_1$ и понеже той е симетрична функция на α и β, ще трябва значи да се дели на всички разлики $\alpha_i - \beta_k$, което е противоречие.

2. **Елиминиране посредством търсене на най-голям общ делител.** Нека означим с $D(x)$ общия най-голям делител на левите части $f(x)$ и $\varphi(x)$ на уравненията (1). Тогава общите корени на двете уравнения $f(x)=0$, $\varphi(x)=0$ очевидно ще бъдат дадени с уравнението $D(x)=0$. Следователно, за да имат двете уравнения (1) q общи корена, трябва общият им най-голям делител да бъде от степен q.

Нека степента n на $f(x)$ да бъде най-малко равна на степента m на $\varphi(x)$. Разделяме $f(x)$ с $\varphi(x)$ и нека полученото частно е Q, а остатъкът R_1, т. е.

$$f(x) = \varphi(x)\, Q + R_1.$$

Прилагаме по-нататък това деление и нека $R_2, \ldots, R_\nu$ са получените остатъци, а $Q_1, \ldots, Q_\nu$ са частните от делението. Така ще получим

$$\begin{aligned}
f(x) &= \varphi(x)\, Q + R_1, \\
\varphi(x) &= R_1\, Q_1 + R_2, \\
R_1 &= R_2\, Q_2 + R_3, \\
&\cdots\cdots\cdots\cdots \\
R_{\nu-2} &= R_{\nu-1}\, Q_{\nu-1} + R_\nu, \\
R_{\nu-1} &= R_\nu\, Q_\nu.
\end{aligned}$$

Ако считаме коефициентите на $f(x)$ и $\varphi(x)$ за произволни, то степените на $R_1, R_2, \ldots, R_\nu$ намаляват с единица, започвайки от $m-1$, така че степента на R_ν ще бъде $m-\nu$. Общият най-голям делител ще бъде R_ν. За да намерим условията двете уравнения (1) да имат q общи корена, трябва да напишем, че остатъкът R_{m-q+1} е тъждествено равен на нула, т. е. ако

$$R_{m-q+1} = l_0\, x^{q-1} + l_1 x^{q-2} + \ldots + l_{q-1},$$

то условията ще бъдат

$$l_0 = 0,\ l_1 = 0,\ l_2 = 0, \ldots, l_{q-1} = 0.$$

3. **Метод на Ойлер.** Двете дадени уравнения са

$$\begin{aligned}
f(x) &= a_0\, x^n + a_1\, x^{n-1} + \ldots + a_{n-1}\, x + a_n = 0, \\
\varphi(x) &= b_0\, x^m + b_1\, x^{m-1} + \ldots + b_{m-1}\, x + b_m = 0.
\end{aligned} \tag{1}$$

Да допуснем, че те имат поне един общ корен x_1. Тогава можем да пишем

$$f(x)=(x-x_1)f_1(x), \qquad \varphi(x)=(x-x_1)\varphi_1(x), \tag{10}$$

гдето $f_1(x)$ и $\varphi_1(x)$ са полиноми съответно от $n-1$-ва и $m-1$-ва степен. От (10), като изключим $x-x_1$, получаваме

$$f(x)\varphi_1(x)-f_1(x)\varphi(x)=0.$$

Следователно, ако уравненията (1) имат общ корен, то трябва да съществуват два полинома $M(x)$ и $N(x)$ съответно от степен $m-1$, $n-1$, така че да имаме тъждествено[1]

$$f(x)M(x)+\varphi(x)N(x)=0. \tag{11}$$

Достатъчно е очевидно да вземем $M(x)=\varphi_1(x)$, $N(x)=-f_1(x)$. Обратно, нека (11) да съществува при горното предположение за степените на полиномите $M(x)$ и $N(x)$. Ако в (11) поставим вместо x всички n корена x_i, $i=1, 2, \ldots, n$, на уравнението $f(x)=0$, то получаваме

$$\varphi(x_i)N(x_i)=0.$$

Но понеже степента на $N(x)$ е $n-1$, то ще трябва поне едно от $\varphi(x_i)$ да бъде равно на нула, с което е доказано, че двете уравнения (1) имат поне един общ корен. Следователно необходимо и достатъчно условие уравненията (1) да имат общ корен се състои в съществуването на два полинома:

$$M(x)=p_0x^{m-1}+p_1x^{m-2}+\ldots+p_{m-1},$$
$$N(x)=q_0x^{n-1}+q_1x^{n-2}+\ldots+q_{n-1},$$

които да удовлетворяват на тъждеството

$$f(x)M(x)+\varphi(x)N(x)=0.$$

Като напишем, че всички коефициенти на полинома в лявата част са равни на нула, получаваме за неизвестните p и q, броят на които е $n+m$, същия брой линейни хомогенни уравнения:

$$\begin{aligned}
&a_0p_0 && +b_0q_0 && =0,\\
&a_1p_0+a_0p_1 && +b_1q_0+b_0q_1 && =0,\\
&a_2p_0+a_1p_1+a_2p_2 && +b_2q_0+b_1q_1+b_0q_2 && =0,\\
&\cdots\cdots\cdots\cdots && \cdots\cdots\cdots\cdots && \cdots\cdots\\
&a_np_{m-1} && +b_mq_{n-1} && =0.
\end{aligned}$$

По известната теория, за да има тази система решения, не всички равни на нула, необходимо и достатъчно е детерминантата от коефициентите да бъде равна на нула. Следователно тази детерминанта

---

[1] Същото заключение получаваме по следния начин: полиномът $f(x)$ трябва да дели произведението $\varphi(x)N(x)$ и ако $f(x)$ няма общ делител с $\varphi(x)$, то би следвало, че $f(x)$ трябва да дели $N(x)$, което е невъзможо.

(12) $$R=\begin{vmatrix} a_0 & 0 & 0 & \dots & b_0 & \dots & \\ a_1 & a_0 & 0 & \dots & b_1 & b_0 & \dots \\ a_2 & a_1 & a_0 & \dots & b_2 & b_1 & b_0 \dots \\ & \dots & \dots & \dots & \dots & \dots & \\ & \dots & a_n & \dots & \dots & b_m & \end{vmatrix}$$

е търсената резултанта.

4. Метод на Силвестър. Този метод в същност е почти идентичен с метода на Ойлер, но води по-просто до извеждането на резултантата. От дадените уравнения $f(x)=0$, $\varphi(x)=0$ получаваме нови, като първото го умножаваме с $x^0, x, x^2, \dots, x^{m-1}$, а второто с $x^0, x, x^2, \dots, x^{n-1}$. Така получаваме $n+m$ уравнения:

(13)
$$\left\{\begin{array}{r} f(x)=0 \\ x f(x)=0 \\ \\ \\ x^{m-1} f(x)=0 \\ \varphi(x)=0 \\ x\varphi(x)=0 \\ \dots \\ x^{n-1}\varphi(x)=0. \end{array}\right.$$

Ако дадените уравнения имат общ корен, всички тези уравнения (13) ще имат общ корен. Но уравненията (13) представляват $n+m$ линейни уравнения спрямо неизвестните $x, x^2, \dots, x^{n+m-1}$, на които броят е $n+m-1$. За да имат общо рещение, трябва детерминантата от коефициентите пред неизвестните им да бъде равна на нула. Тази детерминанта R е

$$R=\begin{vmatrix} 0 & 0 \dots 0 & a_0 & a_1 \dots a_{n-1} & a_n \\ 0 & 0 \dots a_0 & a_1 & a_2 \dots a_n & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_0 & a_1 \dots & a_n & 0 \dots 0 & 0 \\ 0 & 0 \dots & b_0 & b_1 \dots b_{m-1} & b_m \\ 0 & 0 \dots b_0 & b_1 & b_2 \dots b_m & 0 \\ \dots & \dots & \dots & \dots & \dots \\ b_0 & b_1 \dots & b_m & 0 \dots 0 & 0 \end{vmatrix}$$

Тя може да се отличава от детерминантата на Ойлер само по знака, защото лесно се получава от нея само с разместване на редове и стълбове. Но в същност показахме, че ако дадените уравнения (1) имат общ корен, детерминантата R е равна на нула. Обратно, ако $R=0$, уравненията ще имат общ корен, понеже детерминантата R е по знак равна на детерминантата на Ойлер. Това ще докажем и директно. Именно да умножим първия стълб в R с x^{n+m-1}, втория с x^{n+m-2} и т. н. до предпоследния с x и получените произведения да прибавим към последния стълб. Така получаваме

$$R=\begin{vmatrix} 0 & 0 \ldots 0 & a_0 & a_1 \ldots a_{n-1} & f(x) \\ 0 & 0 \ldots a_0 & a_1 & a_2 \ldots a_n & xf(x) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_0 & a_1 \ldots \ldots a_n & 0 \ldots & & x^{m-1} f(x) \\ 0 & 0 \ldots \ldots b_0 & b_1 & \ldots b_{m-1} & \varphi(x) \\ 0 & 0 \ldots b_0 & b_1 & b_2 \ldots b_m & x\varphi(x) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_0 & b_1 \ldots \ldots & b_m 0 \ldots 0 & & x^{n-1} \varphi(x) \end{vmatrix}.$$

Ако развием по елементите на последния стълб, получаваме тъждеството

$$R=f(x)\ \varphi_1(x)+\varphi(x)\ f_1(x),$$

където $f_1(x)$ и $\varphi_1(x)$ са полиноми съответно от $n-1$-и и $m-1$-ва степен. Ако $R=0$, то като разсъждаваме, както в метода на Ойлер, заключаваме, че уравненията ще имат поне един общ корен.

С метода на Силвестър лесно можем да намерим условията уравненията да имат повече от един корен. За да илюстрираме това, ние ще разгледаме две уравнения от четвърта и трета степен:

$$f(x)=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4=0,$$

(14)

$$\varphi(x)=b_0+b_1x+b_2x^2+b_3x^3=0.$$

Резултатът, който ще получим, се отнася за какви да е степени. Резултантата R ще бъде

$$R=\begin{vmatrix} 0 & 0 & a_0 & a_1 & a_2 & a_3 & a_4 \\ 0 & a_0 & a_1 & a_2 & a_3 & a_4 & 0 \\ a_0 & a_1 & a_2 & a_3 & a_4 & 0 & 0 \\ b_0 & b_1 & b_2 & b_3 & 0 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 & 0 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & 0 & 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix}.$$

Нека в матрицата на R зачеркнем първия и последен хоризонтален ред и първия и последен стълб и от останалите елементи да образуваме детерминанта R_1, редът на която ще бъде очевидно 5. По същия начин от R_1 със зачеркване на първия и последния ред и стълб получаваме детерминанта R_2 и т. н. Тогава ще докажем следното предложение: За да имат уравненията (14) само един общ корен, необходимо и достатъчно е да имаме $R=0$, $R_1\neq 0$. За да имат два общи корена, необходимо и достатъчно е да имаме $R=0$, $R_1=0$, $R_2\neq 0$ и т. н.

Действително нека $R=0$ и нека $-\mu$ е един общ корен. Тогава ще можем да пишем

$$f(x)=(x+\mu)(\alpha_0+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3),$$
$$\varphi(x)=(x+\mu)(\beta_0+\beta_1 x+\beta_2 x^2), \tag{15}$$

отгдето получаваме

$$\begin{aligned} a_0&=\mu\alpha_0, & b_0&=\mu\beta_0,\\ a_1&=\alpha_0+\mu\alpha_1, & b_1&=\beta_0+\mu\beta_1,\\ a_2&=\alpha_1+\mu\alpha_2, & b_2&=\beta_1+\mu\beta_2,\\ a_3&=\alpha_2+\mu\alpha_3, & b_3&=\beta_2.\\ a_4&=\alpha_3, \end{aligned} \tag{16}$$

Другите корени освен $-\mu$ на уравненията (14) ще бъдат корени на уравненията

$$\alpha_0+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3=0,$$
$$\beta_0+\beta_1 x+\beta_2 x^2=0. \tag{17}$$

Ако дадените уравнения (14) имат повече от един общ корен, то тези уравнения трябва да имат поне един общ корен и ако имат само един общ такъв, то дадената система ще има само два общи корена. Но детерминантата на (17) е следната:

$$D=\begin{vmatrix} 0 & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3\\ \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & 0\\ \beta_0 & \beta_1 & \beta_2 & 0 & 0\\ 0 & \beta_0 & \beta_1 & \beta_2 & 0\\ 0 & 0 & \beta_0 & \beta_1 & \beta_2 \end{vmatrix}.$$

Ще докажем, че $D=R_1$. Действително, ако втория стълб на D умножим с μ и го прибавим към първия, третия стълб умножим с μ и го прибавим към втория, също правим с четвъртия и петия стълб, то като използуваме равенствата (16), виждаме, че D става равна на

$$\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 \\ a_1 & a_2 & a_3 & a_4 & 0 \\ b_1 & b_2 & b_3 & 0 & 0 \\ b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix},$$

която детерминанта е R_1. Ако следователно $D=R_1\neq 0$, то уравненията (17) нямат общ корен, т. е. дадените уравнения (14) имат само един общ корен. Ако обаче $R_1=0$. то (17) имат поне един общ корен, т. е. уравненията (14) имат поне два общи корена и т. н.

5. Метод на Безу. Друг метод, с който резултантата се получава във форма на детерминанта, от ред, равен на по-голямата от степените на двете уравнения, е този на Безу. Да разгледаме най-напред две уравнения от една и съща степен:

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_n=0,$$

(18)

$$\varphi(x)=b_0 x^n+b_1 x^{n-1}+\ldots+b_n=0.$$

От тези уравнения ще получим n уравнения от $n-1$-ва степен, които, ако дадените имат общ корен, също ще имат общ корен. За тая цел приравняваме коефициентите пред x^n, като умножим първото уравнение с b_0, а второто с a_0 и ги изваждаме. От първите два члена изваждаме x^{n-1} пред скоби, като с умножение с останалите изрази в скобите приравняваме коефициентите пред тази степен и изваждаме уравненията и т. н. Следователно ще образуваме следните уравнения от степен $n-1$:

$$\begin{gathered} b_0 f(x)-a_0\varphi(x)=0, \\ (b_0 x+b_1) f(x)-(a_0 x+a_1)\varphi(x)=0, \\ (b_0 x^2+b_1 x+b_2) f(x)-(a_0 x^2+a_1 x+a_2)\varphi(x)=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ (b_0 x^{n-1}=b_1 x^{n-2}+\ldots+b_{n-1}) f(x)- \\ -(a_0 x^{n-1}+a_1 x^{n-2}+\ldots+a_{n-1})\varphi(x)=0, \end{gathered}$$

които може съкратено да напишем така:

$$\begin{gathered} A_{10} x^{n-1}+A_{11} x^{n-2}+\ldots+A_{1,\,n-1}=0, \\ A_{20} x^{n-1}+A_{21} x^{n-2}+\ldots+A_{2,\,n-1}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ A_{n0} x^{n-1}+A_{n1} x^{n-2}+\ldots+A_{n,\,n-1}=0. \end{gathered}$$

За да имат общо решение, понеже на $x, x^2, \ldots, x^{n-1}$ можем да гледаме като на неизвестни, трябва детерминантата от коефициентите да бъде равна на нула. Така че резултантата ще бъде

$$R=\begin{vmatrix} A_{10} & A_{11} \ldots A_{1,n-1} \\ A_{20} & A_{21} \ldots A_{2,n-1} \\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ A_{n0} & A_{n1} \ldots A_{n,n-1} \end{vmatrix}.$$

A_{ik} са от първа степен спрямо коефициентите a и b така, че R, както трябва да се очаква, е от n-та степен спрямо коефициентите a и спрямо b.

Нека сега степените на уравненията да бъдат различни:

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0,$$

(19)

$$\varphi(x)=b_0x^m+b_1x^{m-1}+\ldots+b_m=0,$$

и да допуснем, че $n>m$. Тогава умножаваме второто уравнение с x^{n-m} и така получаваме заедно с първото две уравнения с две неизвестни от равни степени. Ако работим както по-горе, ще получим m уравнения от $n-1$-ва степен:

$$\begin{aligned} &A_{10}x^{n-1}+A_{11}x^{n-2}+\ldots+A_{1,n-1}=0, \\ &A_{20}x^{n-1}+A_{21}x^{n-2}+\ldots+A_{2,n-1}=0, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &A_{m0}x^{n-1}+A_{m1}x^{n-2}+\ldots+A_{m,n-1}=0. \end{aligned} \tag{20}$$

Към тези уравнения прибавяме следните:

$$\begin{aligned} &x^{n-m-1}\varphi(x)=b_0x^{n-1}+b_1x^{n-2}+\ldots+b_mx^{n-m-1}=0, \\ &x^{n-m-2}\varphi(x)=b_0x^{n-2}+b_1x^{n-3}+\ldots+b_mx^{n-m-2}=0, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &\varphi(x)=b_0x^m+b_1x^{m-1}+\ldots+b_m=0. \end{aligned} \tag{20'}$$

Така получаваме n уравнения, именно системите (20) и (20′) от $n-1$-ва степен. Резултантата ще бъде

$$R=\begin{vmatrix} A_{10} & A_{11} & A_{12} \ldots A_{1,n-1} \\ A_{20} & A_{21} & A_{22} \ldots A_{2,n-1} \\ \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ A_{m0} & A_{m1} & A_{m2} \ldots A_{m,n-1} \\ b_0 & b_1 & b_2\ \ldots b_m \ldots \\ \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \ldots\ b_0 & b_1 & \ldots b_m \end{vmatrix}.$$

6. Подробно изследване на метода на Безу. Нека са дадени уравненията

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0,$$

(21)

$$\varphi(x)=b_0x^m+b_1x^{m-1}+\ldots+b_{m-1}x+b_m=0$$

и да предположим, че $n\geqq m$. Да поставим

$$f_k(x)=a_0x^k+a_1x^{k-1}+\ldots+a_k, \quad k=0,\ 1,\ldots,n,$$

$$\varphi_k(x)=b_0x^k+b_1x^{k-1}+\ldots+b_k, \quad k=0,\ 1,\ 2,\ldots,m,$$

така че $f_n(x)$ и $\varphi_m(x)$ съвпадат съответно с $f(x)$ и $\varphi(x)$. Тогава имаме

$$f(x)=x^k f_{n-k}(x)+a_{n-k+1}x^{k-1}+\ldots+a_{n-1}x+a_n,$$

$$\varphi(x)=x^k\varphi_{m-k}(x)+b_{m-k+1}x^{k-1}+\ldots+b_{m-1}x+b_m.$$

По метода трябва да образуваме полиномите

(22)
$$h_k(x)=\varphi(x)f_{n-k}(x)-f(x)\varphi_{m-k}(x)=$$
$$=c_{k1}x^{n-1}+c_{k2}x^{n-2}+\ldots+c_{kn},$$

$k=1,\ 2,\ldots,m$. Ако $n>m$, то прибавяме полиномите

$$h_k(x)=c_{k1}x^{n-1}+c_{k2}x^{n-2}+\ldots+c_{kn}, \quad k=m+1,\ m+2,\ldots,n,$$

дадени с

(22′)
$$h_k(x)=x^{n-k}\varphi(x), \quad h_n(x)=\varphi(x).$$

Ако на променливото x дадем тегло 1, то $f_{n-k}(x)$ и $\varphi_{m-k}(x)$ са изобарни функции спрямо a, b, x с тегло съответно $(n-k)$ и $(m-k)$. Следователно $h_k(x)$ за $1\leqq k\leqq m$ е изобарна с тегло $m+n-k$, отгдето се вижда, че коефициентът c_{ki}, който е билинеарна функция на a и b, е изобарна с тегло $m+i-k$.

Понеже и при $k>m$, $h_k(x)$ има също тегло $m+n-k$, то и c_{ki} при $k>m$ има тегло $m+i-k$.

От (22) се вижда, че всеки общ делител на $f(x)$ и $\varphi(x)$ е общ делител на полиномите $h_k(x)$. Ако $n>m$, то от равенствата

$$h_m(x)=\varphi(x)f_{n-m}(x)-b_0f(x),\ h_n(x)=\varphi(x)$$

и обратните

$$b_0f(x)=h_n(x)f_{n-m}(x)-h_m(x), \quad \varphi(x)=h_n(x)$$

следва, че общият най-голям делител на всички полиноми $h_k(x)$ също е общ делител на $f(x)$ и $\varphi(x)$. При $n=m$ се получава лесно

(23)
$$\begin{cases} a_0\varphi(x)-b_0f(x)=h_n(x), \\ a_1\varphi(x)-b_1f(x)=h_{n-1}(x)-xh_n(x), \\ a_2\varphi(x)-b_2f(x)=h_{n-2}(x)-xh_{n-1}(x), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n-1}\varphi(x)-b_{n-1}f(x)=h_1(x)-xh_2(x), \\ a_n\varphi(x)-b_nf(x)=-xh_1(x). \end{cases}$$

Ако a и b са пропорционални, то $f(x)$ и $\varphi(x)$ се отличават само с един постоянен множител и всички $h(x)$ са идентично равни на нула. Този случай може да се отстрани от разглеждане, понеже уравненията (21) представляват тогава само едно. Следователно $h_n(x)$ например не е равен идентично на нула. От системата (23) ще има поне две уравнения, които могат да бъдат решени спрямо $f(x)$ и $\varphi(x)$. Тогава от тези решения следва, че най-големият общ делител на полиномите $h_k(x)$ дели $f(x)$ и $\varphi(x)$. С това е установено предложението: общият най-голям делител на $f(x)$ и $\varphi(x)$ е също общ най-голям делител на n-те полинома $h_k(x)$.

Ако $f(x)$ и $\varphi(x)$ не са взаимно прости, т. е. уравненията (21) имат общ корен, то n-те линейни и хомогенни спрямо $x^{n-1}, x^{n-2}, \ldots, x, x^0=1$ уравнения

$$c_{k1}x^{n-1}+c_{k2}x^{n-2}+\ldots+c_{kn}=0, \quad k=1, 2, \ldots, n$$

ще имат общо решение. Детерминантата им е

$$(24) \qquad C=\begin{vmatrix} c_{11} & c_{12}\ldots c_{1n} \\ c_{21} & c_{22}\ldots c_{2n} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ c_{n1} & c_{n2}\ldots c_{nn} \end{vmatrix}=\sum \pm c_{1i}, c_{2i_2}\ldots c_{ni_n}.$$

Числата $c_{1i}, c_{2i}, \ldots, c_{ni}$ са хомогенни и билинеарни спрямо a, b, а $c_{m+1,i}, \ldots, c_{ni}$ са линейни и хомогенни само спрямо b. Следователно C е хомогенна спрямо a от степен m и хомогенна спрямо b от степен n. Теглото на c_{ki} спрямо a и b беше $(m+i-k)$. Тогава теглото на всеки член в (24) ще бъде

$$\tau=\sum_{\nu=1}^{n}(m+i_\nu-\nu),$$

където $i_1, i_2, \ldots, i_n$ са числата от 1 до n, но може в друг ред, т. е. $\tau=nm$. Следователно C е изобарна спрямо a и b с тегло mn. Детерминантата C като функция на a и b не е идентично равна на нула, защото в противен случай би трябвало да е идентично равна на нула и когато поставим

$$a_1=a_2=\ldots a_{n-1}=0, \quad b_1=b_2=\ldots=b_{m-1}=0.$$

За така получените специални функции $f(x)$ и $\varphi(x)$ ще имаме при $n>m$

$$c_{kk} = a_0b_m, \quad c_{k,n-m+k}=-a_nb_0 \quad \text{за } k\leqq m,$$

$$c_{k,k-m}=b_0, \qquad c_{kk}=b_m \qquad \text{за } k>m,$$

а всички други c са равни на нула. При $n=m$ само членовете c_{kk} са отлични от нула и всички са равни на $a_0b_n-a_nb_0$. От горното се вижда, че в C ще влизат следните членове:

$$(25) \qquad C=a_0{}^m b_m{}^n+(-1)^{mn}a_n{}^m b_0{}^n+\ldots$$

Детерминантата C е цяла рационална функция на a и b със степен и тегло както R. Но C се анулира, когато $f(x)=0$ и $\varphi(x)=0$ имат общи корени. Тогава с въвеждането на корените на тези уравнения в израза на C, както по-рано се вижда, че C се дели на R и понеже степените им са еднакви, то C се отличава от R с един постоянен множител, който е равен, както това се вижда от (20) и (25), на $(-1)^{mn}$. Следователно резултантата R се изразява с $(-1)^{mn}\,C$ като детерминанта от n-ти ред.

Видяхме, че необходимо и достатъчно условие уравненията (21) да имат общ корен е R да бъде равна на нула. Лесно е да решим въпроса за условията, когато (21) имат няколко общи корена. Именно имаме теоремата:

Детерминантата C има ранг r, ако функциите $f(x)$ и $\varphi(x)$ имат най-голям общ делител $\psi(x)$ от степен $(n-r)$.

Ако $r=n$, $f(x)$ и $\varphi(x)$ са взаимно прости и тогава действително $R\neq 0$. Ако $r=0$, то $f(x)$ и $\varphi(x)$ са с изключение на един постоянен множител равни и всички c_{kl} са равни на нула, т. е. рангът на C е нула. Следователно остава да предположим, че $1\leqq r\leqq n-1$. Да допуснем, че $f(x)$ и $\varphi(x)$ имат общ най-голям делител $\psi(x)$ от степен $(n-r)$

Както видяхме, при теоремата на Руше една система от ν хомогенни линейни уравнения се удовлетворява от една система стойности на неизвестните, от които една поне е отлична от нула, ако броят на уравненията е $<\nu$. Ако броят на уравненията е ν, това е само тогава възможно, когато детерминантата от коефициентите е равна на нула. Следователно могат да се намерят $(r+1)$ числа $\lambda_1, \lambda_2, \ldots, \lambda_{r+1}$ не всички нули, така че да имаме уравненията:

$$(26)\qquad \begin{cases} c_{11}\ \lambda_1+c_{21}\lambda_2+\ldots+c_{r+1,1}\lambda_{r+1}=0, \\ c_{12}\ \lambda_1+c_{22}\lambda_2+\ldots+c_{r+1,2}\lambda_{r+1}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ c_{1r}\ \lambda_1+c_{2r}\lambda_2+\ldots+c_{r+1,r}\lambda_{r+1}=0. \end{cases}$$

Тогава функцията

$$\lambda_1 h_1(x)+\lambda_2 h_2(x)+\ldots+\lambda_{r+1}h_{r+1}(x)$$

има по-ниска степен от $(n-r)$ и понеже трябва да се дели на полинома $\psi(x)$ от $(n-r)$-та степен, то тя трябва да бъде идентично равна на нула. Следователно ще имаме още следните $(n-r)$ уравнения:

$$(27)\qquad \begin{cases} c_{1,r+1}\ \lambda_1+c_{2,r+1}\ \lambda_2+\ldots+c_{r+1,\,r+1}\ \lambda_{r+1}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{1n}\lambda_1\ +c_{2n}\lambda_2\ +\ldots+c_{r+1,n}\lambda_{r+1}\ =0. \end{cases}$$

Кои да е $(r+1)$ уравнения от (26) и (27) са удовлетворени с $(r+1)$ числа $\lambda_1, \lambda_2, \ldots, \lambda_{r+1}$, които не са всички нули, така че детер-

минантата на тези уравнения трябва да се анулира. Следователно всички детерминанти от ред $(r+1)$ на матрицата

$$\left\|\begin{matrix} c_{11} & c_{12} & \dots c_{1n} \\ c_{21} & c_{22} & \dots c_{2n} \\ \dots & \dots & \dots \\ c_{r+1,1} & c_{r+1,2} & \dots c_{r+1,n} \end{matrix}\right\| \tag{28}$$

се анулират.

Същите разсъждения се отнасят и за коя да е матрица с $(r+1)$ хоризонтални реда, образувана от C. Следователно рангът на C е най-много равен на r. Нека сега да допуснем, че рангът на C е r. Тогава подобно на по-рано могат да се намерят $(r+1)$ числа $\mu_n, \mu_{n-1}, \dots, \mu_{n-r}$, не всички нули, за които полиномът

$$\mu_n h_n(x)+\mu_{n-1} h_{n-1}(x)+ \dots +\mu_{n-r} h_{n-r}(x) \tag{29}$$

да е идентично равен на нула. Действително по теоремата на Руше, понеже рангът на системата

$$\begin{aligned} &\mu_n c_{n,1}+\mu_{n-1} c_{n-1,1}+ \dots +\mu_{n-r} c_{n-r,1}=0, \\ &\mu_n c_{n2}+\mu_{n-1} c_{n-1,2}+ \dots +\mu_{n-r} c_{n-r,2}=0, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &\mu_n c_{nn}+\mu_{n-1} c_{n-1,n}+ \dots +\mu_{n-r} c_{n-r,n}=0 \end{aligned}$$

е $< r+1$, то съществуват $(r+1)$ числа μ, не всички равни на нула, които я удовлетворяват. Ако тогава в идентично равната на нула функция (29) поставим изразите на $h(x)$ от (22) и (22′), то ще получим тъждеството

$$\varphi(x) g(x)=f(x) l(x). \tag{30}$$

Трябва да има поне едно число μ, отлично от нула, с индекс $\leqq m$, отгдето следва $n-r \leqq m$. Действително, ако няма такова число, то за получаването на (30) ще имат значение само уравненията (22′), т. е. $f(x)$ не ще фигурира. Следователно би следвало, че $l(x)=0$, отгдето следва, че и $g(x)$ е тъждествено равна на нула, което обаче не е възможно, понеже имаме

$$g(x)=\mu_n+\mu_{n-1} x+\mu_{n-2} x^2+ \dots$$

От това следва, че ще имаме

$$l(x)=\mu_m \varphi_0(x)+\mu_{m-1} \varphi_1(x)+ \dots +\mu_{n-r} \varphi_{m-n+r}(x),$$

гдето не всички μ са равни на нула. Ако $l(x)$ е идентично равен на нула, то полиномите

$$\varphi_{m-n+r}(x), \dots, \varphi_0(x)$$

биха били линейно зависими, което не е възможно, понеже детерминантата от коефициентите им е равна на $b_0^{m-n+r+1}$, следователно е

отлична от нула. С това е доказано, че полиномът $l(x)$ не е идентично равен на нула и е от степен най-много равна на $(m-n+r)$ и от (30) се получава, че $g(x)$ е неидентично равен на нула полином от степен, най-много равна на r. Понеже от (30) следва, че $g(x)\varphi(x)$ се дели на $f(x)$, то $f(x)$ и $\varphi(x)$ ще имат един общ делител най-малко от $(n-r)$-та степен. Общ делител от степен, по-висока от $(n-r)$, полиномите $f(x)$ и $\varphi(x)$ не могат да имат, понеже по доказаната част на теоремата би следвало, че рангът на C е по-малък от r. С това теоремата е напълно доказана.

От горните разглеждания лесно се получава общият най-голям делител $\psi(x)$ на $f(x)$ и $\varphi(x)$. От последните r реда на C образуваната матрица има ранг r. Действително, ако няма този ранг, то могат да се намерят r числа $\mu_n, \mu_{n-1}, \ldots, \mu_{n-r+1}$, не всички нули, така че полиномът

$$\mu_n h_n(x)+\mu_{n-1}h_{n-1}(x)+\ldots+\mu_{n-r+1}h_{n-r+1}(x)$$

да е идентично равен на нула. Но тогава може да се докаже както по-горе, че $f(x)$ и $\varphi(x)$ ще имат още един общ делител най-малко от $(n-r+1)$-ва степен, което противоречи на факта, че $\psi(x)$ е от $(n-r)$-та степен. Лесно е да се докаже, че не всички детерминанти от r-ти ред, образувани от матрицата

$$(31)\qquad \left\|\begin{array}{cccc} c_{n-r+1,1} & c_{n-r+1,2} & \ldots & c_{n-r+1,n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n-1,1} & c_{n-1,2} & \ldots & c_{n-1,n} \\ c_{n,1} & c_{n,2} & \ldots & c_{n,n} \end{array}\right\|,$$

като вземем първите $(r-1)$ стълба и един кой да е стълб от следващите $(n-r+1)$, са равни на нула. Могат действително да се намерят r числа $k_1, k_2, \ldots, k_r$, не всички нули, които удовлетворяват първите $(r-1)$ уравнения на системата

$$(32)\qquad k_1 c_{n-r+1,i}+\ldots+k_{r-1}c_{n-1,i}+k_r c_{n,i}=0,\quad i=1, 2, \ldots, n.$$

Ако всичките поменати детерминанти са равни на нула, то останалите уравнения (32) ще бъдат пак удовлетворени със същите числа k. Но тогава рангът на матрицата (31) ще бъде $< r$.

Общият най-голям делител $\psi(x)$ на $f(x)$ и $\varphi(x)$ се дава със следната формула:

$$(33)\qquad \psi(x)=\left|\begin{array}{cccc} c_{n-r+1,1} & c_{n-r+1,2} & \ldots & h_{n-r+1}(x) \\ \cdot & \cdot & \cdot & \cdot \\ c_{n-1,1} & c_{n-1,2} & \ldots & h_{n-1}(x) \\ c_{n,1} & c_{n,2} & \ldots & h_n(x) \end{array}\right|.$$

Действително, ако развием тази детерминанта по елементите на последния стълб и наредим по падащи степени на x, то отначало всички

членове със степени x^{n-1}, $x^{n-2}, \ldots, x^{n-r+1}$ ще имат коефициенти, равни на нула, понеже се представят като детерминанти с равни стълбове. Другите членове обаче не всички ще изчезнат, понеже имат за коефициенти разглежданите $(n-r+1)$ детерминанти на матрицата (31). Следователно детерминантата (33) представлява един полином от степен най-много равна на $(n-r)$, който не е идентично равен на нула. Но понеже $h_{n-r+1}(x), \ldots, h_{n-1}(x), h_n(x)$ се делят на $\psi(x)$, то и детерминантата се дели на $\psi(x)$, т. е. съвпада с този полином (като вземем под внимание, че общият най-голям делител е напълно определен до една константа).

7. Дискриминанта. В този параграф ще се занимаваме с търсене на условието едно дадено уравнение да има многократни корени. Очевидно за това е необходимо и достатъчно уравнението да има общ корен с производното си уравнение така, че този въпрос се свежда на по-раншния за елиминация. Обаче тук ще третираме тази задача директно. Под дискриминанта разбираме една такава цяла рационална функция от коефициентите на уравнението, анулирането на която е необходимо и достатъчно, за да има даденото уравнение многократни корени. От възможните такива изрази ние ще изберем най-простия и него собствено ще наричаме дискриминанта. Нека

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0 \tag{34}$$

е даденото уравнение с корени $\alpha_1, \alpha_2, \ldots, \alpha_n$. Да образуваме произведението от разликите им:

$$\begin{aligned} P=(\alpha_1-\alpha_2)\ (\alpha_1-\alpha_3)\ldots(\alpha_1-\alpha_n).\\ (\alpha_2-\alpha_3)\ldots(\alpha_2-\alpha_n).\\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\ (\alpha_{n-1}-\alpha_n). \end{aligned}$$

Очевидно $P=0$ само тогава, когато (34) има равни корени. Произведението $P=\prod_{i<j}^{1\ldots n}(\alpha_i-\alpha_j)$ не се променя или си мени само знака, когато разместваме корените помежду им по всевъзможни начини, понеже едно такова разместване само променя знака на разликите или го запазва. Така че P не е симетрична функция на корените, но взема само две стойности P и $-P$. Такава функция се нарича алтернативна. Обаче $P^2=\prod_{i<j}(\alpha_i-\alpha_j)^2$ е симетрична и по теоремата за степента коефициентният ѝ израз спрямо $\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, \frac{a_n}{a_0}$ е от степен $2n-2$.

Изразът

$$D=a_0^{2n-2}P^2,$$

който е цяла рационална функция от степен $2n-2$ от коефициентите на уравнението, се нарича дискриминанта. Теглото ѝ ще бъде равно на $n(n-1)$.

Можем на D да дадем друг вид. Именно като всеки множител $(\alpha_i-\alpha_j)^2$ представим така:

$$-(\alpha_i-\alpha_j)\ (\alpha_j-\alpha_i),$$

получаваме

$$\begin{aligned} D=a_0^{n-2}(-1)^{\frac{n(n-1)}{2}}\ & a_0(\alpha_1-\alpha_2)\ (\alpha_1-\alpha_3)\ldots(\alpha_1-\alpha_n).\\ & a_0(\alpha_2-\alpha_1)\ (\alpha_2-\alpha_3)\ldots(\alpha_2-\alpha_n).\\ & \cdots\cdots\cdots\cdots\\ & a_0(\alpha_n-\alpha_1)\ (\alpha_n-\alpha_2)\ldots(\alpha_n-\alpha_{n-1}). \end{aligned}$$

Но от

$$f(x)=a_0(x-\alpha_1)\ (x-\alpha_2)\ldots(x-\alpha_n)$$

с диференциране и заместване лесно получаваме

$$f'(\alpha_1)=a_0(\alpha_1-\alpha_2)\ (\alpha_1-\alpha_3)\ldots(\alpha_1-\alpha_n)$$

и аналогични формули за останалите $f(\alpha_i)$. Така ще имаме

$$D=(-1)^{\frac{n(n-1)}{2}}\ a_0^{n-2}\ f'(\alpha_1)f'(\alpha_2)\ldots f'(\alpha_n).$$

Можем да дадем на дискриминантата форма на детерминанта. Именно по-рано установихме, че детерминантата на Вандермонд

$$\Delta=\begin{vmatrix} 1 & 1 & \ldots & 1\\ \alpha_1 & \alpha_2 & \ldots & \alpha_n\\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_1^{n-1} & \alpha_2^{n-1} & \ldots & \alpha_n^{n-1} \end{vmatrix}$$

е равна на произведението $\prod_{i>k}^{1\ldots n}(\alpha_i-\alpha_k)$. Следователно, за да получим D, трябва да повдигнем тази детерминанта в квадрат и да умножим с a_0^{2n-2}. Ако с c_{ik} означим общия елемент в Δ^2, то като извършваме умножението по редове, ще имаме

$$c_{ik}=\alpha_1^{i+k-2}+\alpha_2^{i+k-2}+\ldots+\alpha_n^{i+k-2}=S_{i+k-2}.$$

Получаваме формулата

$$D=a_0^{2n-2}\begin{vmatrix} S_0 & S_1 & S_2 & \ldots & S_{n-1}\\ S_1 & S_2 & S_3 & \ldots & S_n\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ S_{n-1} & S_n & S_{n+1} & \ldots & S_{2n-2} \end{vmatrix}.$$

8. Връзка между дискриминантата и резултантата. Нека с R означим резултантата на $f(x)$ и $f'(x)$. Тогава ще имаме

$$R=a_0^{n-1}\ f'(\alpha_1)\ f'(\alpha_2)\dots f'(\alpha_n).$$

Оттук следва, че

$$R=(-1)^{\frac{n(n-1)}{2}}\ a_0 D.$$

Значи R и D се отличават само с множителя $\pm a_0$. Да пресметнем някои дискриминанти. За квадратното уравнение

(35)
$$a_0 x^2+a_1 x+a_2=0.$$

имаме

$$D=a_0^2(\alpha_1-\alpha_2)^2=a_0^2[(\alpha_1+\alpha_2)^2-4\alpha_1\alpha_2]=a_1^2-4a_0 a_2.$$

За кубичното уравнение

(36)
$$f(x)=x^3+px+q=0$$

имаме

$$D=-f'(\alpha_1)f'(\alpha_2)f'(\alpha_3)$$

или

$$-D=27(\alpha_1\alpha_2\alpha_3)^2+9p(\alpha_1^2\alpha_2^2+\alpha_1^2\alpha_3^2+\alpha_2^2\alpha_3^2)+$$
$$+p^3+3p^2(\alpha_1^2+\alpha_2^2+\alpha_3^2),$$

което дава

$$D=-27q^2-4p^3.$$

За тези уравнения лесно можем само въз основа на знака на D да съдим за реалността на корените им, предполагайки, че коефициентите им са реални числа. Така за (35), ако корените са реални, очевидно е, че $D>0$ и обратно.

За (36), понеже $D=(\alpha_1-\alpha_2)^2\ (\alpha_1-\alpha_3)^2\ (\alpha_2-\alpha_3)^2$, ако корените са реални, то $D>0$. Ако обаче единият корен е реален, например α_1, а другите два са имагинерни, то тогава, ако поставим

$$\alpha_1-\alpha_2=\gamma+i\delta,$$
$$\alpha_1-\alpha_3=\gamma-i\delta,$$

ще имаме $D=-4\delta^2(\gamma^2+\delta)^2<0$. Значи, ако $D>0$, трите корена са реални и ако $D<0$, уравнението има два имагинерни корена и един реален.

9. Две уравнения с две неизвестни. Теорема на Безу. Нека са дадени уравненията

$$f(x,\ y)=0,$$

(37)

$$\varphi(x,\ y)=0$$

с две неизвестни x, y. Под решения наричаме такива две числа x_1, y_1, които, заместени съответно вместо x и y в уравненията (37), ги обръщат в тъждество, т. е. имаме

$$f(x_1,\ y_1)=0,$$
$$\varphi(x_1,\ y_1)=0.$$

От това следва, че уравненията (37) при $x=x_1$ имат за y общ корен y_1. Следователно x_1 ще анулира резултантата $R(x)$, която се получава, като елиминираме y, т. е. x_1 е корен на уравнението

$$R(x)=0. \tag{38}$$

Обратно, ако x_1 е корен на (38), то двете уравнения (37) ще имат общо решение. Така виждаме, че решенията на (47) ще бъдат

$$(x_1, y_1), (x_2, y_2)\ldots(x_k, y_k),$$

гдето $x_1, x_2, \ldots, x_k$ са корените на (38), а $y_1, y_2, \ldots, y_k$ са съответствуващите общи корени на уравненията (37), които получаваме от (37) като поставим $x=x_1, x_2, \ldots, x_k$.

Подобно можехме да елиминираме x и щяхме да получим едно уравнение за y:

$$R_1(y)=0.$$

Нека степента на f е n, а тази на φ е m. Тогава можем да пишем уравненията (37) така:

$$\begin{cases} a_0 y^n + a_1 y^{n-1} + \ldots + a_n = 0, \\ b_0 y^m + b_1 y^{m-1} + \ldots + b_m = 0, \end{cases} \tag{39}$$

гдето

$$a_1=\alpha_0 x+\alpha_1, \qquad b_1=\alpha_0' x+\alpha_1',$$

$$a_2=\beta_0 x^2+\beta_1 x+\beta_2, \qquad b_2=\beta_0' x^2+\beta_1' x+\beta_2',$$

$$\ldots\ldots\ldots\ldots \qquad \ldots\ldots\ldots\ldots$$

$$a_n=\delta_0 x^n+\delta_1 x^{n-1}+\ldots+\delta_n. \qquad b_m=\gamma_0' x^m+\gamma_1' x^{m-1}+\ldots+\gamma'_m.$$

Коефициентите a_i, b_i са полиноми на x от степен i. Ако уравнението е числено, то някои от числата $\alpha, \beta, \ldots, \alpha', \beta' \ldots$ могат да бъдат нули така, че степените на a_i, b_i спрямо x да бъдат по-малки от i. Отначало ще предполагаме, че $\alpha, \beta, \ldots, \delta, \alpha', \beta', \ldots$ са произволни параметри, между които няма никаква зависимост. Известно ни е, че резултантата на (39) е цяла рационална функция на коефициентите им с тегло mn. Така че ще можем да напишем

$$R(x)=\sum C a_0^{\lambda_0} a_1^{\lambda_1} a_2^{\lambda_2} \ldots b_0^{\mu_0} b_1^{\mu_1} b_2^{\mu_2} \ldots$$

гдето за всеки член имаме

$$\lambda_1+2\lambda_2+3\lambda_3+\ldots+\mu_1+2\mu_2+3\mu_3+\ldots=mn.$$

Следователно степента на $R(x)$ е равна на mn. Така получаваме теоремата на Безу. Резултантата на две в общ вид уравнения с две неизвестни е от степен, равна на произведението на степените на уравненията. Ако коефициентите са частни числа, то тъй като степените на a или b спрямо x могат да станат по-малки, или ако между коефициентите има някоя връзка, то може степента на резултантата да се намали. Въобще за каквито

и да е уравнения имаме: степента на резултантата на две уравнения е най-много равна на произведението от степените им.

Тука остава неопределен случаят, когато (37) имат общ делител $\psi(x, y)$. Тогава системата има безкрайно много решения, дадени с $\psi(x, y)=0$.

10. Подробно изследване на въпроса. Нека

$$f(x, y)=\sum_{i,k} a_{ik} x^i y^k; \quad \varphi(x, y)=\sum_{i,k} b_{ik} x^i y^k \tag{40}$$

и върху x и y да извършим една линейна субституция:

$$x=\alpha x'+\beta y', \quad y=\gamma x'+\delta y', \quad \alpha\delta-\beta\gamma\neq 0.$$

Полиномите f и φ преминават в полиноми на x', y', които са взаимно прости, т. е. нямат общ делител $\tau(x', y')$, ако дадените са такива. Коефициентите пред x'^n и x'^m в новите полиноми са

$$a'_{n0}=\sum_{i.k} a_{ik}\alpha^i\gamma^k, \quad b'_{m0}\sum_{i,k} b_{ik}\alpha^i\gamma^k,$$

гдето сумирането е разпростряно върху членове, които са от най-висока степен n и m. Понеже такива членове сигурно фигурират в (40), то a'_{n0} и b'_{m0} са полиноми на α и γ, които не са тъждествено равни на нула. Но тогава могат да се намерят стойности на α и γ така, че a'_{n0} и b'_{m0} да са отлични от нула. С това е доказано, че нямаме никакво ограничение, ако допуснем, че в $f(x, y)$ и $\varphi(x, y)$ коефициентите на x^n и x^m са отлични от нула.

Да наредим полиномите f и φ по падащите степени на x:

$$f(x, y)=a_0 x^n+a_1 x^{n-1}+\ldots+a_n,$$

$$\varphi(x, y)=b_0 x^m+b_1 x^{m-1}+\ldots+b^m,$$

гдето $a_0\neq 0$, $b_0\neq 0$; a_i, b_i са съответно полиноми на y най-много от степен i. Както видяхме, резултантата $R(y)$ е най-много от степен mn. Ще установим следното предложение: Ако за корена y_0 на $R(y)=0$ полиномите $f(x, y_0)$ и $\varphi(x, y_0)$ имат един общ множител от степен ν, то y_0 е най-малко ν кратен корен на $R(y)=0$. Уравнението $R(y)=0$ се получава по разните методи на елиминацията. Ние ще изберем този на Безу, като ще пишем $R(x)$ във форма на детерминантата (24) от § 6, като сме разместили редовете:

$$R=\begin{vmatrix} c_{n1} & c_{n2} & \ldots c_{nn} \\ c_{n-1,1} & c_{n-1,2} & \ldots c_{n-1,n} \\ \cdot & \cdot & \cdot \\ c_{11} & c_{12} & \ldots c_{1n} \end{vmatrix}.$$

Условието $f(x, y_0)$ и $\varphi(x, y_0)$ да имат общ най-голям делител от степен $n-r=\nu$, както видяхме, се състои в това, че R има ранг r при $y=y_0$ и че главният миньор

$$M=\begin{vmatrix} c_{n1} & c_{n2} & \dots c_{nr} \\ c_{n-1,1} & c_{n-1,2} & \dots c_{n-1,r} \\ \dots & \dots & \dots \\ c_{n-r+1,1} & c_{n-r+1,2} & \dots c_{n-r+1,r} \end{vmatrix} \neq 0$$

при $y=y_0$. От тази детерминанта да получим друга от ред $(r+1)$, като прибавим първите r елемента на $(r+i)$-ия ред на R и първите r елемента на $(r+k)$-ия стълб, като на последно място се прибави $c_{n-r-i+1,\,r+k}$, която да означим с b_{ik}. Тази детерминанта, понеже е образувана от матрицата на R и е от ред $(r+1)$, ще се анулира за $y=y_0$, т. е. ще се дели на $y-y_0$. От получените $(n-r)^2=\nu^2$ детерминанти b_{ik} образуваме детерминантата

$$B=\begin{vmatrix} b_{11} & b_{12} \dots b_{1\nu} \\ b_{21} & b_{22} \dots b_{2\nu} \\ \dots & \dots \\ b_{\nu 1} & b_{\nu 2} \dots b_{\nu\nu} \end{vmatrix},$$

на която стойността по теоремата на Силвестър ще бъде

$$B=R.M^{\nu-1}.$$

Оттук, понеже B съдържа множителя $(y-y_0)$ най-малко в ν-та степен, то следва, че и R го съдържа най-малко в ν-та степен, понеже M при $y=y_0$ е отлично от нула, с което предложението е доказано. От това следва, че на всеки ν-кратен корен y_0 на $R(y)=0$ уравненията

$$f(x, y)=0,\ \varphi(x, y)=0$$

могат да имат най-много ν общи корена. Оттук следва и строго доказателство на теоремата на Безу, че броят на общите решения е най-много равен на nm. На подобно изследване на този въпрос по-нататък няма да се спираме. Ще отбележа, че теоремата лесно се обобщава за повече уравнения.

11. Двузначни функции. В теорията на дискриминантата си послужихме с произведението P, което си изменя само знака при извършване на една транспозиция. Тук ще обобщим тези разглеждания. Една рационална функция $f(x_1, x_2, \dots, x_n)$ се нарича **алтернативна**, ако променя само знака си при извършване на една транспозиция върху променливите. Нека $x_{\mu_1}, x_{\mu_2}, \dots, x_{\mu_n}$ е една пермутация на

$$x_1\ x_2 \dots x_n.$$

Тогава под субституция

$$S=\begin{pmatrix} x_1 & x_2 \dots x_n \\ x_{\mu_1} & x_{\mu_2} \dots x_{\mu_n} \end{pmatrix}$$

разбираме заместването на x_1 с x_{μ_1}, x_2 с $x_{\mu_2}, \dots, x_n$ с x_{μ_n}. Следователно от n елемента имаме $n!$ субституции. Под циклична субституция на $x_1, x_2, \dots, x_k$ разбираме тази, която замества x_1 с x_2, x_2 с $x_3, \dots, x_{k-1}$ с x_k и x_k с x_1. Нея я бележим така:

$$(x_1\ x_2 \dots x_k).$$

Нека субституцията S замества x_α с

$$x_\beta,\ x_\beta \text{ с } x_\lambda, \dots, x_\delta \text{ с } x_\varepsilon,\ x_\varepsilon \text{ с } x_\alpha.$$

Ако така се изчерпват всички елементи, очевидно субституцията S представлява една циклична субституция. Ако не се изчерпват всички елементи, то вземаме нов елемент и се получава подобна нова циклична субституция и т. н., докато се изчерпят всички елементи. По такъв начин се вижда, че S се разпада на циклични субституции.

Но субституцията

$$(x_1 x_2 \dots x_{k-1}\ x_k)$$

е равносилна на последователното извършване на транспозициите

$$(x_1 x_2),\ (x_1 x_3), \dots, (x_1 x_k)$$

или, както се казва, е равна на тяхното произведение.

От всичко това е ясно, че всяка субституция е еквивалентна на няколко транспозиции, което впрочем се вижда и директно. Значи една алтернативна функция при прилагане на една субституция или си запазва същата стойност, или само си променя знака. Лесно е да докажем, че ако f е цяла рационална алтернативна функция, то тя има вида

$$PU,$$

където U е симетрична цяла рационална функция. Действително от условието имаме, че

$$f(x_1, \dots, x_\alpha, \dots, x_\beta, \dots) = -f(x_1, \dots, x_\beta, \dots, x_\alpha, \dots).$$

Оттук, като гледаме на x_α като на променливо, при $x_\alpha = x_\beta$ получаваме, че полиномът f се анулира. Следователно f трябва да се дели на разликата $x_\alpha - x_\beta$. Понеже тази разлика е произволна, то следва, че f се дели на произведението P на всички такива разлики. Можем да пишем

$$f = PU,$$

където U е една цяла рационална функция на променливите $x_1, x_2, \dots, x_n$. Ако в това тъждество извършим една транспозиция върху променливите, то f се изменя в $-f$, P в $-P$, така че U си запазва стойността.

Оттук следва, че U си запазва стойността при всички субституции, т. е. U е симетрична функция.

Една функция е двузначна, щом получава само две стойности при извършване на транспозициите. Нека стойностите ѝ са φ_1 и φ_2, като при извършване на една транспозиция φ_1 минава във φ_2, а последното—във φ_1. Очевидно функцията

$$\frac{\varphi_1+\varphi_2}{2}=\psi$$

е симетрична, а функцията

$$\frac{\varphi_1-\varphi_2}{2}=\psi_1=\eta P$$

е алтернативна. От това следва, че общата форма на една двузначна цяла рационална функция е

$$\psi+P\eta,$$

гдето ψ и η са симетрични цели рационални функции.

Глава V

Трансформация на уравненията

1. Прости случаи. Под трансформация на уравненията разбираме действието, с което от дадено уравнение се получава ново, на което корените са различните стойности на една рационална функция от корените на даденото уравнение. Ние ще разгледаме отначало случаите, когато всеки корен на трансформираното уравнение зависи само от един съответен корен на даденото, и след това случая при зависимост от повече от един.

Нека е дадено уравнение

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0$$

с корени $x_1, x_2, \ldots, x_n$.

Да намерим първо уравнение с k пъти по-големи корени, т. е. за което на корен x_i на даденото отговаря корен $y_i=kx_i$, $i=1,\ldots,n$. Търсеното уравнение ще бъде

$$f\left(\frac{y}{k}\right)=0$$

или

$$a_0y^n+a_1ky^{n-1}+\ldots+a_nk^n=0.$$

Друга проста трансформация е следната: да се намери уравнение, на което корените да бъдат с числото α по-малки, т. е. трансформацията е

$$y=x-\alpha.$$

Търсеното уравнение очевидно ще бъде

$$f(y+\alpha)=$$

$$=f(\alpha)+\frac{y}{1!}f'(\alpha)+\frac{y^2}{2!}f''(\alpha)+\ldots+\frac{y^{n-1}}{(n-1)!}f^{(n-1)}(\alpha)+\frac{y^n}{n!}f^{(n)}(\alpha)=0.$$

Можем да изберем α така, че да липсва членът с y^{n-1}. За тази цел трябва $f^{(n-1)}(\alpha)=0$ или

$$n!\,a_0\alpha+(n-1)!\,a_1=0, \quad \alpha=-\frac{a_1}{na_0}.$$

Ако искаме да липсва членът с y^{n-2}, то трябва да решим квадратното уравнение $f^{(n-2)}(\alpha)=0$.

Друга трансформация е да се намери уравнение, на което корените са реципрочните стойности на корените на даденото уравнение. Следователно имаме

$$y=\frac{1}{x}.$$

Трансформираното уравнение ще бъде

$$y^n f\left(\frac{1}{y}\right)=0.$$

Да разгледаме сега общата дробна линейна трансформация

(1) $$y=\frac{ax+b}{cx+d},$$

$$ad-bc\neq 0.$$

Понеже оттук имаме

$$x=\frac{b-dy}{cy-a},$$

то трансформираното уравнение ще бъде

$$(cy-a)^n f\left(\frac{b-dy}{cy-a}\right)=0.$$

Трансформацията (1) може да се извърши с прилагане последователно на по-раншните трансформации, именно полагаме

$$y=\frac{a}{c}+\frac{bc-ad}{c(cx+d)}=\frac{a}{c}+x',$$

$$x'=\frac{bc-ad}{c}\frac{1}{x''},$$

$$x''=cx+d.$$

2. Правило на Хорнер. На Хорнер се дължи едно по-практично извършване на трансформацията $y=x-\alpha$.

Предварително да разгледаме делението на един полином с бинома $x-\alpha$. Нека

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_{n-1} x+a_n$$

и $Q(x)$ е частното от делението с $x-\alpha$, а R—остатъкът, гдето

$$Q(x)=b_0 x^{n-1}+b_1 x^{n-2}+\ldots+b_{n-1}.$$

Тогава с приравняване на коефициентите в

$$f(x)=(x-\alpha)\,Q(x)+R$$

получаваме равенствата:

$$\begin{gathered}a_0=b_0,\\ a_1=b_1-\alpha b_0,\\ a_2=b_2-\alpha b_1,\\ \ldots\ldots\\ a_{n-1}=b_{n-1}-\alpha b_{n-2},\\ a_n=R-\alpha b_{n-1},\end{gathered}$$

които дават

$$b_0=a_0,\quad b_1=a_1+\alpha b_0,\quad b_2=a_2+\alpha b_1,\ldots,\ R=a_n+\alpha b_{n-1}.$$

Така че делението можем да представим по-прегледно, като в една редица наредим коефициентите на $f(x)$, а в друга редица, под първата, коефициентите на частното и остатъка. Ще имаме таблицата

$$\alpha\ \Big|\ \begin{array}{c|cccc}a_0 & a_1 & a_2 \ldots & a_{n-1} & a_n\\ \hline b_0 & b_1 & b_2 \ldots & b_{n-1} & R\end{array},$$

гдето имаме $b_0=a_0,\ b_1=a_1+\alpha b_0,\ldots,$ т. е. всяко b е равно на сбора от съответното a и произведението на α с предшествуващото b. Същото важи за остатъка R. За пример да разделим

$$x^4+2x^2-3x+1$$

с $x+2$. В случая $\alpha=-2$ и ще имаме

$$\begin{array}{r|cccc}-2 & 1 & 0 & 2- & 3 & 1\\ \hline & 1 & -2 & 6- & 15 & 31\end{array}.$$

Частното е $x^3-2x^2+6x-15$, а остатъкът 31.

Ще изложим след тези бележки метода на Хорнер. Нека трансформираното уравнение да бъде

$$f(\alpha+y)=q_0+q_1 y+q_2 y^2+\ldots+q_n y^n=0. \tag{2}$$

Коефициентите q се свеждат на стойностите на производните на $f(x)$ за $x=\alpha$, но пресмятанията така са доста сложни. Начинът на Хорнер

се състои в това, че тези коефициенти ги получаваме като остатъци от деление с $x-\alpha$. Именно, ако в (2) поставим

$$\alpha+y=x,$$

то получаваме

$$f(x)=q_0+q_1(x-\alpha)+q_2(x-\alpha)^2+\ldots+q_n(x-\alpha)^n. \tag{3}$$

От това равенство е ясно, че q_0 е остатъкът от делението на $f(x)$ с $(x-\alpha)$, а частното е полиномът

$$f_1(x)=q_1+q_2(x-\alpha)+\ldots+q_n(x-\alpha)^{n-1}, \tag{4}$$

$$f(x)=q_0+(x-\alpha)f_1(x).$$

От (4) също се вижда, че q_1 е остатъкът от делението на частното $f_1(x)$ с $x-\alpha$, а частното от това деление е полиномът

$$f_2(x)=q_2+(x-\alpha)q_3+\ldots+(x-\alpha)^{n-2}q_n.$$

Продължавайки така, виждаме, че коефициентите q на трансформираното уравнение (2) се получават като остатъци от деление с $x-\alpha$.

За пример нека от уравнението

$$x^4-2x^3-3x^2+2x+5=0$$

да получим ново, на което корените са с две по-малки. Методът на Хорнер ни води до следните операции:

	1	−2	−3	2	5
2	1	0	−3	−4	−3
	1	2	1	−2	
	1	4	9		
	1	6			

Трансформираното уравнение е

$$y^4+6y^3+9y^2-2y-3=0.$$

3. Трансформация на Чирнхаус (дадена в 1683 г.). Ще разгледаме сега една трансформация, при която корените на трансформационното уравнение са стойностите на една рационална функция на едно променливо, като даваме на него стойности, равни на корените на даденото уравнение. Както видяхме обаче, всяка дробна рационална функция от един корен на уравнението се изразява като цяла рационална функция от същия корен така, че ние можем да се ограничим на случая, когато корените на трансформираното уравнение y са стойностите на една цяла рационална функция от един корен на даденото уравнение или ако с

$$y_1, y_2, \ldots, y_n$$

означим корените на трансформираното уравнение, ще имаме

$$y=p_0+p_1x+p_2x^2+\ldots+p_rx^r,$$

гдето можем да предположим, че $r\leqq n-1$, понеже степените от n нагоре, както вече видяхме, могат от $f(x)=0$, $xf(x)=0,\ldots$ да се изразят посредством по-ниски степени от n. Същото постигаме с деление на полинома y с $f(x)$. Но тогава ще имаме релациите

$$\begin{aligned}
y &= p_0+p_1x+\ldots+p_rx^r,\\
y^2 &= q_0+q_1x+\ldots+q_{n-1}x^{n-1},\\
y^3 &= r_0+r_1x+\ldots+r_{n-1}x^{n-1},\\
&\cdots\cdots\cdots\cdots\\
y^n &= t_0+t_1x+\ldots+t_{n-1}x^{n-1},
\end{aligned}\tag{5}$$

гдето q са хомогенни функции от втора степен на p, r са хомогенни функции от трета степен на p и т. н. Нека с S_1, $S_2,\ldots$ означим степенните сборове за уравнението $f(x)=0$:

$$S_i=x_1^i+x_2^i+\ldots+x_n^i,$$

а със σ_1, $\sigma_2,\ldots$ — степенните сборове за трансформираното уравнение

$$y^n+A_1y^{n-1}+A_2y^{n-2}+\ldots+A_n=0,$$

$$\sigma_i=y_1^i+y_2^i+\ldots+y_n^i.$$

Ако във формулите (5) поставим

$$x=x_1,\ x_2,\ldots,\ x_n$$

и ги съберем, ще получим

$$\begin{aligned}
\sigma_1 &= np_0+p_1S_1+\ldots+p_rS_r,\\
\sigma_2 &= nq_0+q_1S_1+\ldots+q_{n-1}S_{n-1},\\
\sigma_3 &= nr_0+r_1S_1+\ldots+r_{n-1}S_{n-1},\\
&\cdots\cdots\cdots\cdots\\
\sigma_n &= nt_0+t_1S_1+\ldots+t_{n-1}S_{n-1}.
\end{aligned}\tag{6}$$

Като пресметнем от тези формули σ_1, $\sigma_2,\ldots,\sigma_n$ и ги заместим във формулите на Нютон:

$$\begin{aligned}
&\sigma_1+A_1=0.\\
&\sigma_2+A_1\sigma_1+2A_2=0,\\
&\sigma_3+A_1\sigma_2+A_2\sigma_1+3A_3=0,\\
&\cdots\cdots\cdots\cdots
\end{aligned}\tag{7}$$

ще получим търсените коефициенти A_1, $A_2,\ldots,A_n$.

Ще получим сега трансформираното уравнение в детерминантна форма. С умножаване на първото уравнение от (5) с корена x и сте-

пените му $x^2, \ldots, x^{n-1}$ и освобождаване на по-високите от $(n-1)$ степени, както преди, получаваме n-те уравнения:

$$y = p_0 + p_1 x + \ldots + p_r x^r,$$

$$xy = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1} x^{n-1},$$

$$x^2 y = \beta_0 + \beta_1 x + \ldots + \beta_{n-1} x^{n-1},$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$x^{n-1} y = \omega_0 + \omega_1 x + \ldots + \omega_{n-1} x^{n-1}.$$

Като елиминираме $x, x^2, \ldots, x^{n-1}$, за y получаваме уравнението

$$\begin{vmatrix} p_0 - y & p_1 & p_2 & \ldots p_r\, 0 \ldots \\ \alpha_0 & \alpha_1 - y & \alpha_2 & \ldots\ldots \alpha_{n-1} \\ \beta_0 & \beta_1 & \beta_2 - y & \ldots\ldots \beta_{n-1} \\ \cdots & \cdots & \cdots & \cdots\cdots \\ \omega_0 & \omega_1 & \omega_2 & \ldots\ldots \omega_{n-1} - y \end{vmatrix} = 0,$$

което е трансформираното уравнение.

4. Някои приложения Посредством трансформацията на Чирнхаус можем с подбиране на параметрите p да получим уравнение, в което липсват някои членове. Така ще покажем, че уравненията от трета и четвърта степен можем да решим алгебрически, като ги сведем на непълни уравнения, решими лесно.

Така уравнението от трета степен

$$a_0 x^3 + a_1 x^2 + a_2 x + a_3 = 0$$

с трансформацията

$$y = p_0 + p_1 x + p_2 x^2$$

ще се сведе към непълното уравнение

$$y^3 + A_3 = 0,$$

стига p_0, p_1, p_2 да подберем така, че

$$A_1 = A_2 = 0,$$

или, което е еквивалентно, както веднага се вижда от (7), на

(8) $$\sigma_1 = \sigma_2 = 0.$$

Последните уравнения са от първа и втора степен спрямо p_0, p_1, p_2, така че могат да бъдат решени (едното p, например p_0, можем да положим 1). Като решим (7) и (8) от

$$y = p_0 + p_1 x + p_2 x^2,$$

$$y^2 = q_0 + q_1 x - q_2 x^2$$

с елиминиране на x^2, получаваме корена x:

$$x=\alpha_0+\alpha_1 y+\alpha_2 y^2,$$

с което даденото кубично уравнение се решава.

Уравнението от четвърта степен

$$a_0x^4+a_1x^3+a_2x^2+a_2x+a_4=0$$

посредством трансформацията

$$y=p_0+p_1 x+p_2x^2$$

се обръща в биквадратното уравнение

$$y^4+A_2y^2+A_4=0,$$

стига p_0, p_1, p_2 да удовлетворяват условията

$$A_1=0, \quad A_3=0$$

или на еквивалентните уравнения

$$\sigma_1=0, \ \sigma_3=0.$$

Първото е от първа степен, а второто — от трета.

5. Трансформация на Жерар (дадена в 1834 г.). Видяхме, че посредством решението на линейно и квадратно уравнение можем в трансформираното уравнение да анулираме втория и третия член. Можем да си поставим въпроса за анулирането и на четвъртия коефициент A_3. От формулите (7) се вижда, че за да имаме

$$A_1=0, \quad A_2=0, \quad A_3=0,$$

трябва

$$\sigma_1=0, \quad \sigma_2=0, \quad \sigma_3=0.$$

Спрямо p първото уравнение е линейно, второто квадратно и третото кубично, така че, ако елиминираме директно два от параметрите p, ще получим едно уравнение от шеста степен. На Жерар се дължи едно видоизменение на това решение, при което се използуват само уравнения от първа степен и едно кубично. Нека трансформацията на Чирнхаус е

$$y=p_0+p_1x+p_2x^2+p_3x^3+p_4x^4.$$

Както споменахме, σ_2 е хомогенна функция от втора степен спрямо p_0, p_1, p_2, p_3, p_4 и като такава може да се представи като сума от квадрати на линейни функции:

$$\sigma_2=\lambda_1P_1^2+\lambda_2P_2^2+\lambda_3P_3^2+\lambda_4P_4^2.$$

Ако подберем числата p така, че

$$\lambda_1P_1^2+\lambda_2P_2^2=0, \quad \lambda_3P_3^2+\lambda_4P_4^2=0,$$

то очевидно $\sigma_2=0$. Горните уравнения са удовлетворени, ако са изпълнени линейните уравнения

$$(9) \qquad \sqrt{\lambda_1}P_1=\sqrt{-\lambda_2}P_2, \quad \sqrt{\lambda_3}P_3=\sqrt{-\lambda_4}\,P_4.$$

Така че уравненията $\sigma_1=0$, $\sigma_2=0$, $\sigma_3=0$ се заместват с уравненията $\sigma_1=0$, $\sigma_3=0$ и (9). От тези хомогенни уравнения три са от първа степен, а едно от трета. Избирайки едното p, например $p_0=1$, то другите p_1, p_2, p_3, p_4 се напълно определят. Така по този начин на уравнението от пета степен можем да дадем формата

$$y^5+A_4y+A_5=0.$$

Като направим субституцията $y=z^4/A_4$, то приема вида

$$z^5+z+q=0,$$

към което значи може да се сведе всяко уравнение от пета степен.

6. Обща трансформация. Тази трансформация се състои в следното от даденото уравнение

$$f(x)=0$$

с корени x_1, $x_2,\ldots,x_n$ да се получи ново, на което корените са стойностите, които взема една рационална функция

$$\varphi(x_1, x_2,\ldots, x_k)$$

от корените, когато последните разместваме по всевъзможни начини. Ако y_1, $y_2,\ldots,y_N$ са значенията на тази функция и уравнението, на което са корени, е

$$y^N+A_1y^{N-1}+\ldots+A_N=0,$$

то задачата се състои в пресмятане на коефициентите A. Отначало да видим каква е степента N на това уравнение. Ако φ не е симетрична спрямо никоя група корени, то очевидно броят N на значенията ѝ ще бъде равен на броя на вариациите на n елемента в k-ти клас, т. е.

$$N=n(n-1)\ldots(n-k+1).$$

Ако φ е симетрична спрямо една група от α_1 корена, друга група от α_2 корена и т. н., то

$$N=\frac{n(n-1)(n-2)\ldots(n-k+1)}{\alpha_1!\,\alpha_2!\,\alpha_3!\ldots}.$$

За коефициентите имаме

$$\begin{aligned} -A_1&=y_1+y_2+\ldots+y_N,\\ A_2&=y_1y_2\ +\ldots+y_{N-1}y_N,\\ &\ldots\ldots\ldots\ldots\\ (-1)^NA_N&=y_1y_2\ldots y_N. \end{aligned}$$

Десните части са рационални функции на корените $x_1, x_2, \ldots, x_n$, които са освен това и симетрични, понеже, като разместваме тези корени, количествата $y_1, y_2, \ldots, y_N$ само се разместват помежду им. Следователно ще можем да изразим коефициентите A рационално посредством коефициентите на даденото уравнение.

Да вземем един пример: за уравнението

$$x^3+px+q=0$$

с корени x_1, x_2, x_3 да се намери ново, на което корените да бъдат стойностите на функцията x_1x_2, т. е.

$$y_1=x_1x_2, \quad y_2=x_1x_3, \quad y_3=x_2x_3.$$

Ако търсеното уравнение е

$$y^3+A_1y^2+A_2y+A_3=0,$$

то ще имаме

$$-A_1=y_1+y_2+y_3=\sum x_1x_2=p,$$

$$A_2=y_1y_2+y_1y_3+y_2y_3=x_1x_2x_3\sum x_1=0,$$

$$-A_3=y_1y_2y_3=(x_1x_2x_3)^2=q^2,$$

така че трансформираното уравнение ще бъде

$$y^3-py^2-q^2=0. \tag{10}$$

Това уравнение можем да получим по малко по-друг начин. От релацията $x_1x_2x_3=-q$ имаме

$$y_1=x_1x_2=-\frac{q}{x_3},$$

така че горната трансформация се свежда на

$$y_1=-\frac{q}{x}.$$

Като заместим в даденото уравнение вместо x равното му $-\frac{q}{y}$, ще получим (10).

7. **Уравнение на квадрата от разликите на корените.** Дадено е уравнението

$$f(x)=0$$

с корени $x_1, x_2, \ldots, x_n$. Ще намерим уравнение, на което корените са квадрати от разликите на корените $x_1, x_2, \ldots, x_n$, т. е. стойностите, които взема функцията

$$(x_1-x_2)^2.$$

Степента на това уравнение ще бъде

$$m=\frac{n(n-1)}{2}$$

и нека корените му бъдат

$$y_1, y_2, \cdots, y_m.$$

Трансформираното уравнение ще получим по метода на Лагранж, като пресметнем степенните сборове на корените му. Да образуваме полинома

$$\varphi(x)=(x-x_1)^{2k}+(x-x_2)^{2k}+\ldots+(x-x_n)^{2k}. \tag{11}$$

По бинома на Нютон имаме

$$(x-x_i)^{2k}=$$

$$=x^{2k}-\binom{2k}{1}x^{2k-1}x_i+\binom{2k}{2}x^{2k-2}x_i^2-\ldots-\binom{2k}{1}xx_i^{2k-1}+x_i^{2k},$$

в която формула поставяме $i=1, 2, \ldots, n$ и събираме:

$$\varphi(x)= \tag{12}$$

$$=nx^{2k}-\binom{2k}{1}S_1x^{2k-1}+\binom{2k}{2}S_2x^{2k-2}-\ldots-\binom{2k}{1}S_{2k-1}x+S_{2k},$$

гдето

$$S_p=x_1^p+x_2^p+\ldots+x_n^p.$$

Да означим с b_p степенния сбор

$$b_p=y_1^p+y_2^p+\ldots+y_m^p$$

на трансформираното уравнение. Тогава, ако в (11) поставим $x=x_1$, $x_2, \ldots, x_n$ и съберем, ще получим

$$2\sum_{\alpha<\beta}^{1\ldots n}(x_\alpha-x_\beta)^2=\varphi(x_1)+\varphi(x_2)+\ldots+\varphi(x_n)$$

или

$$2b_k=\sum_{i=1}^{n}\varphi(x_i).$$

Дясната част на това равенство получаваме лесно, като в (12) поставим $x=x_1, x_2, \ldots, x_n$ и съберем. Така ще имаме

$$2b_k=nS_{2k}-\binom{2k}{1}S_1S_{2k-1}+\binom{2k}{2}S_2S_{2k-2}-\ldots+nS_{2k}$$

или като групираме равните членове от двата края,

$$b_k=nS_{2k}-\binom{2k}{1}S_1S_{2k-1}+\binom{2k}{2}S_2S_{2k-2}-\ldots+\frac{(-1)^k}{2}\binom{2k}{k}S_k^2.$$

По тази формула пресмятаме степенните сборове

$$b_1, b_2, \ldots, b_m$$

и оттам коефициентите на трансформираното уравнение.

По напълно подобен начин се получава уравнение, на което корените са квадрати от сумите по два от корените на даденото уравнение.

ЧАСТ V

ЧИСЛЕНО РЕШАВАНЕ НА УРАВНЕНИЯТА

Глава I

Граници на корените и отделяне на рационалните корени

1. Дефиниция. В тази част ще се занимаваме главно с уравнения, на които коефициентите са реални числа. Под горна граница на реалните корени разбираме число, което е по-голямо от най-големия реален корен. Долна граница е число, по-малко от най-малкия корен. Очевидно, ако уравнението има положителни корени, то можем да търсим долна граница на положителните корени. Аналогично ще имаме горна граница на отрицателните корени.

2. Правило на Лагранж. Нека е дадено уравнението

$$f(x)=x^n+a_1x^{n-1}+\ldots+a_n=0 \tag{1}$$

и нека α е абсолютната стойност на най-големия по абсолютна стойност отрицателен коефициент. Тогава, ако m е разликата от степента на уравнението и степента на първия отрицателен член, числото

$$L=1+\sqrt[m]{\alpha}$$

е горна граница.

За да установим теоремата, достатъчно е да докажем, че при $x>L$, $f(x)>0$.

Ако в (1) изпуснем всички положителни членове от втория до първия отрицателен и коефициентите на следващите заместим с $-$, то при $x>0$ така намаляваме лявата част, т. е. ще имаме

$$f(x)\geqq x^n-\alpha(x^{n-m}+x^{n-m-1}+\ldots+1)=x^n-\alpha\frac{x^{n-m+1}-1}{x-1}.$$

Оттук ще получим, че $f(x)>0$, ако изберем x така, че

$$x^n>\alpha\frac{x^{n-m+1}-1}{x-1}$$

или още повече, като изберем $x>1$, ако

$$x^n\geqq\alpha\frac{x^{n-m+1}}{x-1},\quad x^{m-1}(x-1)\geqq\alpha.$$

Последното неравенство е сигурно удовлетворено, ако

$$(x-1)^{m-1}(x-1) \geqq \alpha,\ (x-1)^m \geqq \alpha,$$

отгдето

$$x \geqq 1+\sqrt[m]{\alpha}=L,$$

с което правилото е доказано.

3. Правило на Нютон. Ако едно число L обръща $f(x)$ и всички производни до n-тата в положителни, то е горна граница на реалните корени на уравнението (1).

Действително, полагайки $x=L+y$, имаме

$$f(x)=f(L)+\frac{y}{1!}f'(L)+\frac{y^2}{2!}f''(L)+\ldots+\frac{y^n}{n!}f^{(n)}(L).$$

Оттук се вижда, че при $x \geqq L$, т. е. $y \geqq 0$, имаме

$$f(x) > 0.$$

с което правилото е доказано.

При прилагане на това правило постъпваме така: отначало намираме число x_1, което обръща производната $f^{(n-1)}(x)$ в положителна или нула евентуално. Понеже тази производна е от първа степен, то вземаме число x_1, равно или по-голямо от корена $f^{(n-1)}(x)=0$. След това изпитваме дали полиномът $f^{(n-2)}(x)$ е положителен за $x=x_1$ и в случай, че е отрицателен, повишаваме x_1, докато получим положителна стойност или нула. Така продължаваме до $f(x)$.

4. Правило на Лагер. Ако всичките коефициенти на частното от делението на $f(x)$ с $x-L(L>0)$ и остатъкът са положителни, то L е горна граница.

Действително от

$$\frac{f(x)}{x-L}=F(x)+\frac{R}{x-L},\ \text{т. е. } f(x)=(x-L)F(x)+R,$$

понеже $F(x)>0$ при $x>0$, ще имаме при $x \geqq L$

$$f(x) > 0.$$

Ако означим с

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n,$$
$$f_0(x)=a_0,$$
$$f_1(x)=a_0x+a_1,$$
$$f_2(x)=a_0x^2+a_1x+a_2,$$
$$\ldots\ldots\ldots\ldots$$
$$f_{n-1}(x)=a_0x^{n-1}+a_1x^{n-1}+\ldots+a_{n-1},$$
$$f_n(x)=f(x),$$

то, както видяхме, частното може да се напише така:

$$F(x)=f_0(L)\,x^{n-1}+f_1(L)\,x^{n-2}+\ldots+f_{n-1}(L),\ R=f(L).$$

Полиномите $f_0, f_1, \ldots$ ще наричаме полиноми на Лагер. Те са свързани с простите релации

$$f_1(x)=x\,f_0(x)+a_1,\quad f_2(x)=xf_1(x)+a_2,\quad f_3(x)=xf_2(x)+a_3,\ldots,$$

така че лесно се пресмятат стойностите им за дадено x. Правилото значи се състои в намиране на положително число, което обръща всички полиноми на Лагер в положителни, евентуално в нула. Това правило прилича много на правилото на Нютон, но си служи с прости полиноми.

5. Метод на Коши. Нека $\alpha_r, \alpha_s, \alpha_t, \ldots$ са абсолютните стойности на отрицателните коефициенти в уравнението (1) и k е броят им. Тогава най-голямото от числата

$$(k\alpha_r)^{\frac{1}{r}},\ (k\alpha_s)^{\frac{1}{s}},\ (k\alpha_t)^{\frac{1}{t}},\ldots \tag{2}$$

е горна граница на положителните корени.

Действително, ако g е число, по-голямо от числата (2), ще имаме

$$g^r > k\alpha_r,\ g^s > k\alpha_s,\ g^t > k\alpha_t \ldots$$

и следователно

$$g^n > k\alpha_r g^{n-r},\ g^n > k\alpha_s g^{n-s},\ g^n > s\alpha_t g^{n-t},\ldots$$

Ако съберем тези неравенства и вземем под внимание, че броят им е k, ще получим

$$g^n > \alpha_r g^{n-r}+\alpha_s g^{n-s}+\alpha_t g^{n-t}+\ldots,$$

т. е. първият член в $f(x)$ при $x=g$ е по-голям от сумата на всички отрицателни членове. Но тогава $f(x)>0$ за всяко x, по-голямо от числата (2), с което правилото е установено.

6. Метод на групиране. Предварително ще установим едно предложение, на което се основава този метод. Нека $f(x)$ и $\varphi(x)$ са два полинома с положителни коефициенти—такива, че най-високата степен m на x във $\varphi(x)$ да бъде равна или по-малка от най-ниската степен в $f(x)$. Ще докажем, че ако за едно положително число l полиномът

$$F(x)=f(x)-\varphi(x)$$

е положителен (или нула), то при $x>l$, $F(x)>0$. За тази цел да разгледаме частното

$$\frac{F(x)}{x^m}=\frac{f(x)}{x^m}-\frac{\varphi(x)}{x^m}.$$

Полиномът $\frac{f(x)}{x^m}$ е нареден по положителни степени на x и следователно, когато x расте, той расте също $(x>0)$. Полиномът $\frac{\varphi(x)}{x^m}$

съдържа положителни степени на $\frac{1}{x}$ и когато x расте, той намалява. От това следва, че когато x расте, оставайки положително, $\frac{F(x)}{x^m}$ расте, с което предложението става очевидно.

Тогава лявата част на едно уравнение представяме като сума от полиноми от формата на $F(x)$. Ако всички тези полиноми при едно значение l на x са положителни или нула, то l ще бъде горна граница на положителните корени.

Пример. За уравнението

$$f(x)=x^5-10x^4+15x^3-4x^2-16x+390=0 \tag{3}$$

да намерим с различните методи горната граница. Правилото на Лагранж дава

$$L=1+16=17.$$

Правилото на Коши

$$L=\text{по-голямото от числата (2.10)}, \sqrt[4]{2.16}=20.$$

За да приложим метода на Нютон, образуваме производните, разделени със съответните факториели:

$$f=x^5-10x^4+15x^3+4x^2-16x+390,$$

$$\frac{f'}{1!}=5x^4-40x^3+45x^2+8x-16,$$

$$\frac{f''}{2!}=10x^3-60x^2+45x+4,$$

$$\frac{f'''}{3!}=10x^2-40x+15,$$

$$\frac{f^{IV}}{4!}=5x-10,$$

$$\frac{f^{V}}{5!}=1.$$

Производната f^{IV} при $x=2$ се анулира. Производната f''' при $x=2$ е отрицателна, при $x=4$ — положителна. Производната f'' при $x=6$ е положителна, f' при $x=7$, а f при $x=8$ са положителни. Следователно методът на Нютон дава

$$L=8.$$

По метода на Лагер трябва да образуваме полиномите

$$f_0=1,$$

$$f_1=x-10,$$

$$f_2=xf_1+15,$$

$$f_3 = xf_2 + 4,$$

$$f_4 = xf_3 - 16,$$

$$f_5 = xf_4 + 390 = f.$$

При $x=10$, $f_1=0$, $f_2=15$, $f_3=154$, $f_4=1524$, $f_5=153\,630$. Следователно този метод дава

$$L=10.$$

По метода на групиране лявата част трябва да разделим на полиноми по следния начин:

$$(x-10)x^4+x(15x^2+4x-16)+390=0,$$

отгдето се вижда, че по този метод имаме

$$L=10.$$

7. Долна граница на реалните корени. Лесно е да намерим долната граница на отрицателните корени, т. е. число, по-малко от най-малкия отрицателен корен. За това намираме горната граница L на положителните корени на $f(-x)=0$. Тогава $-L$ очевидно ще е долната граница на отрицателните корени на уравнението $f(x)=0$.

Също лесно можем да намерим долната граница l' на положителните корени, т. е. положително число, по-малко от най-малкия положителен корен. Именно, ако L' е горната граница на положителните корени на $f\left(\frac{1}{x}\right)=0$, то l' ще бъде равно на $\frac{1}{L'}$, понеже от $\frac{1}{x}<L'$ следва $x>\frac{1}{L'}=l'$.

Горната граница пък на отрицателните корени на $f(x)=0$ е долна граница на положителните корени на $f(-x)=0$, взета с обратен знак.

Так за примера, който взехме, имаме

$$x^5 f\left(\frac{1}{x}\right)=390\,x^5-16\,x^4+4x^3+15x^2-10x+1=0.$$

По правилото на Коши по-голямото от числата

$$\frac{2\,.\,16}{390},\quad \sqrt[4]{\frac{20}{390}}$$

е горна граница L' на положителните корени на това уравнение. Може да се приеме $L'=\frac{1}{2}$, понеже

$$\frac{32}{390}<\frac{1}{2},\quad \sqrt[4]{\frac{2}{39}}<\frac{1}{2},$$

отгдето долната граница на положителните корени на (3) е равна на 2.

8. Цели рационални корени. Ако едно уравнение с рационални коефициенти има рационален корен, то, както ще видим, можем лесно да го намерим. Отначало ще установим следното предложение: ако уравнението с цели коефициенти

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n=0 \tag{4}$$

при $a_0=1$ има рационален корен, то той е цял. Действително нека $x=\frac{p}{q}$ е един такъв корен, гдето можем да предполагаме, че p и q са взаимно прости. Тогава от

$$\frac{p^n}{q^n}+a_1\frac{p^{n-1}}{q^{n-1}}+a_2\frac{p^{n-2}}{q^{n-2}}+\ldots+a_n=0$$

с умножаване на q^{n-1} получаваме

$$\frac{p^n}{q}+a_1p^{n-1}+a_2p^{n-2}q+\ldots+a_nq^{n-1}=0,$$

отгдето следва, че $\frac{p^n}{q}$ е цяло число. Понеже p и q са взаимно прости, то това очевидно е възможно само тогава, когато $q=\pm1$, т. е. x е цяло число.

Нека x е един корен на (4). От уравнението с разделяне на x получаваме

$$a_0x^{n-1}+a_1x^{n-2}+\ldots+a_{n-1}+\frac{a_n}{x}=0,$$

което показва, че x дели a_n. Следователно целите корени са делители на свободния член.

За да намерим целите корени на едно уравнение $f(x)=0$, то изпитваме кои делители на свободния член го удовлетворяват. Разбира се, ограничаваме се на такива делители, които се намират между границите на реалните корени. Ако уравнението е от висока степен, по-удобно е, преди да проверим дали един делител го удовлетворява, да приложим следващите правила. Именно нека α е един цял корен на $f(x)=0$. Тогава полиномът

$$Q(x)=\frac{f(x)}{x-\alpha}$$

ще бъде с цели коефициенти, следователно числата

$$\frac{f(1)}{\alpha-1}=-Q(1),\ \frac{f(-1)}{\alpha+1}=-Q(-1)$$

трябва да бъдат цели. Значи, за да бъде едно цяло число α корен на уравнението $f(x)=0$ с цели коефициенти, трябва $\alpha-1$ да дели $f(1)$ и $\alpha+1$ да дели $f(-1)$. Ако едно поне от тези условия не е изпълнено, α не е корен на уравнението.

Можем да постъпваме другояче, като изчисляваме коефициентите на частното, но започвайки от последния. Нека

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0, \tag{5}$$

α е един цял корен и нека поставим

$$\frac{f(x)}{x-\alpha}=b_0x^{n-1}+b_1x^{n-2}+\ldots+b_{n-1}.$$

Оттук, ако сравним коефициентите, получаваме равенствата

$$a_n+\alpha b_{n-1}=0,\quad a_{n-1}+\alpha b_{n-2}=b_{n-1},$$

$$a_{n-2}+\alpha b_{n-3}=b_{n-2},\ldots,a_1+\alpha b_0=b_1,\ a_0=b_0,$$

от които имаме

$$b_{n-1}=-\frac{a_n}{\alpha},\quad b_{n-2}=\frac{b_{n-1}-a_{n-1}}{\alpha}, \tag{6}$$

$$b_{n-3}=\frac{b_{n-2}-a_{n-2}}{\alpha},\ldots,\ b_0=\frac{b_1-a_1}{\alpha}=a_0.$$

Ако α е корен, то трябва числата b_i да бъдат цели и $b_0=a_0$. Ако едно поне е дробно, то очевидно α не може да бъде корен на уравнението.

За пример да вземем уравнението

$$x^5-5x^4-28x^3+280x^2-704x+560=0. \tag{7}$$

Понеже $560=2^4.5.7$, то делителите на 560 ще бъдат числа, образувани с прости делители 2, 5, 7. Да намерим границите на корените. Лявата част може да се напише

$$x^3(x^2-5x-28)+x(280x-704)+560=0,$$

отгдето се вижда, че 9 е горна граница, понеже двата полинома в скобите са положителни. Също се вижда, че -8 е долна граница на отрицателните корени. Следователно възможни рационални корени са

$$\pm1,\quad \pm2,\quad \pm4,\quad \pm5,\quad \pm7,\ 8.$$

Имаме

за $\alpha=8, b_4=-\frac{560}{8}=-70, b_3=\frac{-70+704}{8}=\frac{634}{8}$ не е цяло число,

„ $\alpha=7, b_4=\frac{-560}{7}=-80,\ b_3=\frac{-80+704}{4}$ не е цяло число,

„ $\alpha=-7, b_4=\frac{560}{7}=80,\quad b_3=\frac{80+704}{-7}=-112,$

$$b_2=\frac{-112-280}{-7}=56,\ b_1=\frac{56+28}{-7}=-12,\ b_0=\frac{-12+5}{-7}=1=a_0:$$

Като отстраним корена -7, получаваме

$$x^4-12x^3+57x^2-112x+80=0.$$

-7 не може да бъде повече корен, понеже не дели 80. Изпитваме $\alpha=5$, получаваме

$$b_3=-\frac{80}{5}=-16, \quad b_2=\frac{-16+112}{5}$$

не е цяло.

Също се вижда, че -5, ± 4 не са корени. За $\alpha=2$ имаме

$$b_3=-\frac{80}{2}=-40, \quad b_2=\frac{-40+112}{2}=36,$$

$$b_1=\frac{36-56}{2}=-10, \quad b_0=\frac{-10+12}{2}=1=a_0.$$

Като отстраним корена 2, получаваме

$$x^3-10x^2+36x-40=0.$$

Изпитваме отново 2, понеже дели 40:

$$b_2=\frac{40}{2}=20, \quad b_1=\frac{20-36}{2}=-8, \quad b_0=\frac{-8+10}{2}=1=a_0.$$

Остава уравнението

$$x^2-8x+20=0,$$

което няма рационални корени. Даденото уравнение (7) се разлага така:

$$(x-2)^2\ (x+7)\ (x^2-8x+20)=0.$$

9. Дробни рационални корени. Нека уравнението

$$a_0x^n+a_1x^{n-1}+\ldots+a_n=0 \tag{8}$$

има дробния рационален корен $x=\frac{p}{q}$, гдето p и q са взаимно прости числа. Тогава от условието, че x е корен на уравнението, имаме

$$a_0\frac{p^n}{q^n}+a_1\frac{p^{n-1}}{q^{n-1}}+\ldots+a_{n-1}\frac{p}{q}+a_n=0.$$

Като умножим с q^{n-1}, получаваме

$$\frac{a_0p^n}{q}+a_1p^{n-1}+\ldots+a_{n-1}pq^{n-2}+a_nq^{n-1}=0. \tag{9}$$

от което следва, че $\frac{a_0 p^n}{q}$ е цяло число. Понеже p и q са взаимно прости, трябва a_0 да се дели на q. Ако умножим (9) с q и разделим с p, получаваме

$$a_0 p^{n-1} + a_1 p^{n-2} q + \ldots + a_{n-1} q^{n-1} + \frac{a_n q^n}{p} = 0,$$

отгдето аналогично следва, че p дели a_n.

Следователно числителят на дробния корен дели свободния член, а знаменателят дели коефициента пред най-високата степен.

Обаче търсенето на дробните рационални корени може да се сведе на търсене на цели корени, като трансформираме даденото уравнение в друго, в което коефициентът $a_0 = 1$.

Полагаме

$$x = \frac{y}{a_0},$$

тогава уравнението (8) се трансформира в

$$y^n + a_1 y^{n-1} + a_0 a_2 y^{n-2} + \ldots + a_0{}^{n-1} a_n = 0. \tag{10}$$

На рационални корени на (8) ще отговарят рационалните корени на (10), т. е. целите корени.

Глава II

Брой на корените в един интервал

1. Метод на субституциите. Бидейки дадено едно уравнение с реални коефициенти

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0, \tag{1}$$

явява се въпросът за определяне броя на корените, които се намират в един интервал. Преди да изложим точното решение на този въпрос, ще вземем някои прости теореми, които почти непосредствено ни дават граници за този брой корени. Една такава най-проста теорема е следната:

Ако за две числа a и b полиномът $f(x)$ има еднакви знаци, уравнението (1) има четен брой корени в интервала (a, b). Ако знаците на $f(x)$ са различни, то (1) има нечетен брой корени в същия интервал. Тази теорема е очевидна, като се следи непрекъснатото изменение на $f(x)$, когато x варира от a до b. Но теоремата може да се установи алгебрически, като се използува разлагането на $f(x)$ в биномни множители. Именно нека корените в интервала (a, b) да бъдат

$$a < \alpha_1 \leqq \alpha_2 \leqq \ldots \leqq \alpha_m < b.$$

Тогава ще имаме

$$f(x) = (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_m) \varphi(x),$$

гдето $\varphi(x)$ съдържа биномни множители от вида $x-\beta$, гдето β е вън от интервала (a, b), т. е.

$$\frac{a-\beta}{b-\beta} > 0,$$

и други положителни при реално x квадратни множители $(x-\mu)^2+\nu^2$, отговарящи на имагинерните корени, така че ще имаме

$$\frac{\varphi(a)}{\varphi(b)} > 0.$$

Следователно от

$$\frac{f(a)}{f(b)} = \frac{a-\alpha_1}{b-\alpha_1} \cdot \frac{a-\alpha_2}{b-\alpha_2} \cdots \frac{a-\alpha_m}{b-\alpha_m} \cdot \frac{\varphi(a)}{\varphi(b)}$$

следва, че знакът на $\frac{f(a)}{f(b)}$ се дава с $(-1)^m$, с което теоремата е установена, понеже при m четно имаме

$$\frac{f(a)}{f(b)} > 0,$$

и обратно, и при m нечетно $\frac{f(a)}{f(b)} < 0$ и обратно.

Тази теорема ни дава само четността или нечетността на броя на корените в (a, b), но има случаи, гдето тя напълно решава въпроса. Така нека за уравнението $f(x)=0$ от n-та степен сме намерили $n+1$ числа $\mu_1, \mu_2, \ldots, \mu_{n+1}$, така че за всеки две последователни числа $f(x)$ да има обратни знаци. Тогава във всеки интервал (μ_i, μ_{i+1}) ще имаме поне по един корен и понеже броят на тези интервали е n, то следва, че във всеки интервал ще има само по един корен, т. е. както казваме, корените са отделени.

2. Теорема на Рол[1]. Между два последователни реални корена на едно уравнение производното уравнение има нечетен брой корени, значи най-малко един.

Ще докажем една лема, от която произнесената теорема следва лесно.

Пред всеки реален корен на $f(x)=0$, близко до него $f(x)$ и $f'(x)$ имат обратни знаци, а след корена те имат еднакви знаци. Именно нека α е корен на $f(x)=0$, който за общност да предположим, че е m-кратен, т. е.

$$f(\alpha)=0,\ f'(\alpha)=0, \ldots\ f^{(m-1)}(\alpha)=0,\ f^{(m)}(\alpha) \neq 0.$$

Тогава, като развием числителя и знаменателя в отношението

$$\frac{f(\alpha+h)}{f'(\alpha+h)}$$

[1] Френски математик, живял във времето на Newton. Теоремата е дал в съчинението си „Traité d'Algèbre" (1690).

според формулата на Тейлор по степените на h ще имаме, след като сме съкратили на h^{m-1},

$$\frac{f(\alpha+h)}{f'(\alpha+h)}=h\frac{\frac{f^{(m)}(\alpha)}{m!}+\frac{hf^{(m+1)}(\alpha)}{(m+1)!}+\cdots}{\frac{f^{(m)}(\ )}{(m-1)!}+h\frac{f^{(m+1)}(\alpha)}{m!}+\cdots}.$$

На основание на основните свойства на полиномите можем да вземем h достатъчно малко по абсолютна стойност така, че знаците на числителя и знаменателя да се дават от първите членове. Но понеже тези първи членове са с еднакви знаци, то дробта е положителна. Следователно знакът на $\frac{f(\alpha+h)}{f'(\alpha+h)}$ се дава от h при достатъчно малко $|h|$. Ако $h<0$, т. е. $\alpha+h$ се намира пред корена α, отношението е отрицателно, ако $h>0$, т. е. след корена, отношението е положително.

Нека a и b са два последователни реални корена на уравнението

$$f(x)=0. \tag{2}$$

Нека $h>0$ е достатъчно малко, така че в интервалите $(a,a+h)$, $(b-h,\ b)$ уравнението

$$f'(x)=0 \tag{3}$$

да няма корени освен евентуално a и b и предната лема да е приложима. Тогава, като вземем h достатъчно малко, според лемата имаме

$$f(a+h)f'(a+h)>0,$$

$$f(b-h)f'(b-h)<0,$$

а понеже между a и b $f(x)$ не се анулира,

$$f(a+h)f(b-h)>0.$$

Като умножим тези три неравенства, получаваме

$$f'(a+h)f'(b-h)<0,$$

което по теоремата за субституциите показва, че в интервала $(a+h,\ b-h)$, т. е. вътре в интервала $(a,\ b)$, уравнението (3) има нечетен брой реални корени. Обаче във верността на теоремата може да се убедим с геометрическо представяне. Действително кривата $y=f(x)$ започва от точка с ординатата $y=0\,(x=a)$ и завършва с ординатата пак нула, абсциса $x=b$. Ако y отначало расте, след това трябва да се намалява и ще има един максимум. Може отново y да почне да расте, т. е. да мине през минимум, но след това обязателно трябва да почне да намалява, т. е. да мине през максимум. Така се убеждаваме, че общият брой на максимумите и минимумите е нечетен. Понеже при всеки максимум или минимум $f'(x)$ се анулира, като запазва или променя знака си според това, дали коренът ѝ е четно или нечетно кратен, то теоремата на Rolle става очевидна.

Едно важно следствие от теоремата на Rolle е следното: между два последователни реални корена на производното уравнение даденото може да има най-много един реален корен. Действително нека α и β са два такива корена на (3). Да допуснем, че даденото уравнение (2) има повече от един корен в интервала (α, β) и нека a, b са два негови такива корена. Но тогава по теоремата на Rolle следва, че производното уравнение (3) има поне един корен в (a, b), т. е. α и β не са последователни корени на него, което противоречи на условието. Също се доказва, че даденото уравнение (2) може да има най-много един реален корен, по-малък от най-малкия реален корен на производното уравнение (3), и най-много един, по-голям от най-големия корен на (3).

Това следствие от теоремата ни помага за отделянето на корените на уравнението (2) в случай, че можем да решим уравнението (3). Действително нека $\alpha_1, \alpha_2, \ldots, \alpha_k$ са реалните корени на (3), наредени по растяща големина. Да разгледаме интервалите

$$(-\infty, \alpha_1), (\alpha_1, \alpha_2), (\alpha_2, \alpha_3), \ldots, (\alpha_{k-1}, \alpha_k), (\alpha_k, \infty).$$

Ако за един кой да е интервал знаците на $f(\lambda)$ за границите му са еднакви, то $f(x)=0$ няма корен в него. Ако знаците на $f(x)$ са различни за границите му, то $f(x)=0$ има един корен в този интервал, който така е отделен.

От теоремата на Rolle следва: ако в един интервал $f(x)=0$ има m реални корена, то $f'(x)=0$ има най-малко $m-1$ реални корена в него. Действително нека въпросните корени на $f(x)=0$ са $a_1, a_2, \ldots, a_k$, $a_p < a_{p+1}$, съответно от кратности $\mu_1, \mu_2, \ldots, \mu_k$, $\mu_1+\mu_2+\ldots+\mu_k=m$. Тогава, понеже $a_1, a_2, \ldots, a_k$ са съответно μ_1-1, $\mu_2-1, \ldots, \mu_k-1$ кратни корени на $f'(x)=0$ и във всеки интервал (a_i, a_{i+1}) последното уравнение има поне по един корен, то общо ще има най-малко

$$\mu_1-1+\mu_2-1+\ldots+\mu_k-1+k-1=m-1$$

корена. В частност, ако даденото уравнение има m реални корена, то производното има поне $m-1$ реални корена. От това следва, че ако всичките корени на едно уравнение са реални, то и всичките корени на производното са реални. Понеже, ако даденото е от n-та степен, производното трябва да има най-малко $n-1$ реални корена, т. е. всичките да бъдат реални. На същото основание следва, че и $f''(x)=0$ ще има само реални корени, също и $f'''(x)=0$ и т. н. Освен това от доказателството на горното следствие от теоремата на Рол е очевидно следното допълнение: Когато всички корени на $f(x)=0$ са реални, то между два последователни корена има само един на $f'(x)=0$ и корените на второто уравнение се намират между най-малкия и най-големия корен на първото. Уравнението $f'(x)=0$ не може да има многократни корени, който не са същевременно корени на $f(x)=0$ в случай, че последното уравнение има само реални корени. Действително иначе биха съществували два последователни корена на $f(x)=0$, които съдържат повече от един корен на $f'(x)=0$.

3. Пример. Да намерим условието уравнението

3) $$f(x)=x^3+px+q=0$$

да има само реални корени.

Производното

$$f'(x)=3x^2+p=0$$

трябва да има само реални корени. Понеже корените му са

$$x_1=-\sqrt{-\frac{p}{3}} \quad x_2=\sqrt{-\frac{p}{3}},$$

то трябва $p<0$. Освен това трябва $f(x)$ в трите интервала

$$(-\infty,\ x_1) \quad (x_1,\ x_2), \quad (x_2,\ \infty)$$

да си променя знака и понеже $f(-\infty)=-\infty$, $f(\infty)=\infty$, трябва $f(x_1)>0$, $f(x_2)<0$, които условия са необходими и достатъчни. Тези условия са значи

$$f(x_1)=\left(-\sqrt{-\frac{p}{3}}\right)^3+p\left(-\sqrt{-\frac{p}{3}}\right)+q>0,$$

$$f(x_2)=\left(\sqrt{-\frac{p}{3}}\right)^3+p\sqrt{-\frac{p}{3}}+q<0$$

или като преработим,

$$-\frac{2p}{3}\sqrt{-\frac{p}{3}}+q>0,$$

(5)

$$\frac{2p}{3}\sqrt{-\frac{p}{3}}+q<0.$$

Ако ги умножим, получаваме

(6) $$q^2+\frac{4p^3}{27}<0 \text{ или } \frac{q^2}{4}+\frac{p^3}{27}<0.$$

Това неравенство следва от (5). Обратно, лесно се вижда, че от (6) следва (5). Действително лявата част на (6) е произведение на левите части на (5) и понеже е отрицателна, то или ще имаме неравенствата (5), или неравенствата

$$-\frac{2p}{3}\sqrt{-\frac{p}{3}}+q<0,$$

$$\frac{2p}{3}\sqrt{-\frac{p}{3}}+q>0.$$

В невъзможността на тези неравенства се убеждаваме веднага, като ги извадим. Получаваме

$$-\frac{4p}{3}\sqrt{-\frac{p}{3}}<0,$$

което противоречи на $p<0$, следващо от (6). Така получихме, че необходимото и достатъчно условие уравнението (4) да има само реални корени е

$$\frac{q^2}{4}+\frac{p^3}{27}<0.$$

По подобен начин се третира по-общото триномно уравнение

$$x^n+px^m+q=0.$$

Полиномът от n-та степен

$$P_n(x)=\frac{1}{2^n n!}\,\frac{d^n[(x^2-1)^n]}{dx^n}$$

се нарича полином на Лежандър. Понеже уравнението $(x^2-1)^n=0$ има всичките си корени реални, от които 1 и -1 като n-кратни, по следствията от теоремата на Рол уравнението

$$P_n(x)=0$$

има само реални корени, които са прости и се намират в интервала $(-1, 1)$.

4. Теорема на Пулен—Хермит. Нека $f(x)$ е полином с реални коефициенти. Ако

$$g(x)=c_0x^n+c_1x^{n-1}+\ldots+c_n,\ c_0\neq 0,\ c_n\neq 0$$

е полином само с реални нули, то полиномът

$$h(x)=c_0f^{(n)}(x)+c_1f^{(n-1)}(x)+\ldots+c_nf(x)$$

има поне толкова реални нули, колкото полиномът $f(x)$. Ако $f(x)$ има само реални нули, то всеки многократен корен на $h(x)=0$ е също многократен на $f(x)$.

Нека $-\alpha_1, -\alpha_2, \ldots, -\alpha_n$ са корените на $g(x)=0$, т. е. имаме

$$g(x)=c_0(x+\alpha_1)\ (x+\alpha_2)\ldots(x+\alpha_n).$$

Отначало ще докажем, че ако α е реално, то уравнението

$$F(x)=\alpha f(x)+f'(x)=0,\quad \alpha\neq 0,$$

има поне толкова реални корени, както $f(x)=0$. Корените на уравнението $f(x)=0$ се наричат нули на полинома $f(x)$. Наричат се и корени на $f(x)$.

Затова ще докажем, че полиномът $F(x)$ между два последователни корена на $f(x)$ има нечетен брой корени. Действително нека a и b са два такива корена от кратности, съответно равни на k и l. Тогава от

$$f(x)=(x-a)^k(x-b)^l\varphi(x),\ \varphi(a)\neq 0,\ \varphi(b)\neq 0,$$

гдето $\varphi(x)$ има постоянен знак в интервала (a, b), получаваме

$$\alpha f(x)+f'(x)=(x-a)^{k-1}(x-b)^{l-1}\Phi(x),$$

гдето

$$\Phi(x)=k(x-b)\varphi(x)+l(x-a)\varphi(x)+(x-a)(x-b)\varphi'(x)+$$
$$+\alpha(x-a)(x-b)\varphi(x).$$

Корените на $F(x)=0$, които са в интервала (a, b), са тези на уравнението $\Phi(x)=0$, които се намират в същия интервал. Понеже

$$\Phi(a)=k(a-b)\varphi(a),\quad \Phi(b)=l(b-a)\varphi(b),$$

то $\Phi(a)$ и $\Phi(b)$ са с противни знаци. Значи по теоремата за субституциите полиномът $\Phi(x)$ ще има между a и b нечетен брой корени, отгдето следва същото за полинома $F(x)$. Нека реалните корени на $f(x)$, наредени по растяща големина, да бъдат

$$x_1,\ x_2,\ldots,x_p$$

със съответни кратности $\mu_1, \mu_2,\ldots,\mu_p$. Тогава същите числа ще бъдат и корени на $F(x)$ от кратности с единица по-малки. Понеже по току-що доказаното $F(x)$ има най-малко по един корен в интервалите $(x_1, x_2), (x_2, x_3),\ldots, (x_{p-1}, x_p)$, то този полином ще има най-малко

$$(\mu_1-1)+(\mu_2-1)+\ldots+(\mu_p-1)+p-1=(\mu_1+\ldots+\mu_p)-1$$

реални корена. Като вземем пред вид, че степените на $f(x)$ и $F(x)$ са еднакви, получаваме горния резултат.

Ако всички корени на $f(x)$ са реални, то всеки многократен корен на $F(x)$ е такъв и на $f(x)$. Действително в този случай, както винаги се вижда по горното, ще имаме между всеки два корена x_i, x_{i+1} само един прост корен на $F(x)$, както и вън от интервала (x_1, x_n). Така че единствените многократни корени на $F(x)$ са $x_1, x_2,\ldots, x$.

Получаваме лесно теоремата на Пулен - Хермит, като приложим доказаното предложение последователно. Именно образуваме полиномите:

$$f_1(x)=\alpha_1 f(x)+f'(x),$$
$$f_2(x)=\alpha_2 f_1(x)+f_1'(x)=\alpha_1\alpha_2 f(x)+(\alpha_1+\alpha_2)f'(x)+f''(x),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f_n(x)=\alpha_n f_{n-1}(x)+f'_{n-1}(x)$$
$$=\alpha_1\alpha_2\ldots\alpha_n f(x)+(\alpha_1\alpha_2\ldots\alpha_{n-1}+\ldots)f'(x)+\ldots+(\alpha_1+\ldots+$$
$$+\alpha_n)f^{(n-1)}(x)+f^{(n)}(x).$$

Ако умножим полинома $f_n(x)$ с c_0, получаваме

$$c_0 f_n(x) = h(x),$$

с което теоремата е установена.

Да разгледаме някои приложения:

I. Нека уравнението

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$$

има само реални корени. Тогава и уравнението

$$a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n = 0$$

ще има само реални корени. Ако положим $f(x) = x^n$, получаваме за $h(x)$ полинома

$$h(x) = n!\, a_0 + (n-1)!\, a_1 x + \ldots + n a_{n-1} x^{n-1} + a_n x^n.$$

Следователно уравнението

$$a_0 + \frac{a_1}{1!} x + \frac{a_2}{2!} x^2 + \ldots + \frac{a_n}{n!} x^n = 0$$

има също само реални корени.

Ако с $Df(x)$ означим производната на $f(x)$, то полиномът $h(x)$ може да се пише така:

$$h(x) = c_0 D^n f(x) + c_1 D^{n-1} f(x) + \ldots + c_n f(x) = g(D) f(x).$$

Нека сега $g(x) = 0$ е уравнение със само реални корени, като $g(0) \neq 0$. Да развием $\frac{1}{g(x)}$ по степените на x:

$$\frac{1}{g(x)} = b_0 + b_1 x + b_2 x^2 + \ldots$$

и нека $f(x)$ е произволен полином с реални коефициенти, степента на който да е равна на n. Тогава полиномът

$$\varphi(x) = b_0 f(x) + b_1 f'(x) + \ldots + b_n f^{(n)}(x)$$

няма повече реални нули от полинома $f(x)$.

Действително имаме

$$\varphi(x) = \frac{1}{g(D)} f(x),$$

откъдето получаваме

$$f(x) = g(D) \varphi(x).$$

Остава тогава да се приложи теоремата на Пулен—Хермит за полинома $\varphi(x)$.

По аналогичен начин ще установим сега една теорема на Лагер, която има приложение за установяване реалността на нулите на важни

класи от полиноми. Предварително ще докажем едно помощно предложение.

Нека полиномът $f(x)$ от n-та степен е с реални коефициенти и α е произволно реално число, лежащо вън от интервала $(-n, 0)$. Тогава полиномът

$$\varphi(x)=\alpha f(x)+xf'(x)$$

има поне толкова реални нули, колкото полинома $f(x)$.

Да предположим отначало, че $f(0)\neq 0$. Нека a и b са две последователни положителни нули на $f(x)$ съответно от кратности p и q, т. е.

$$f(x)=(x-a)^p(x-b)^q g(x),$$

като $g(x)$ има значи постоянен знак в интервала (a, b). За $\varphi(x)$ получаваме

$$\varphi(x)=(x-a)^{p-1}(x-b)^{q-1}h(x),$$

където

$$h(x)=[\alpha(x-a)(x-b)+px(x-b)+qx(x-a)]g(x)+x(x-a)(x-b)g'(x).$$

Понеже

$$h(a)=pa(a-b)g(a),\ h(b)=qb(b-a)g(b),$$

то $h(a)$ и $h(b)$ имат противни знаци и следователно между a и b полиномът $\varphi(x)$ има нечетен брой реални нули. Освен това числата a и b са нули на $\varphi(x)$ съответно от кратности $p-1$ и $q-1$. Ако P е броят на положителните нули на $f(x)$, то от горното заключаваме, че $\varphi(x)$ има поне $P-1$ положителни нули. Нека

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n.$$

За $\varphi(x)$ имаме

$$\varphi(x)=a_0\alpha+a_1(\alpha+1)x+a_2(\alpha+2)x^2+\ldots+a_n(\alpha+n)x^n.$$

Да означим с P_1 броя на положителните нули на полинома $\varphi(x)$. При $x=0$ имаме $f(0)=a_0$ и $\varphi(0)=a_0\alpha$ и при $x=\infty$ знакът на полинома $f(x)$ се дава от a_n и знакът на полинома $\varphi(x)$ е този на $a_n(\alpha+n)$. Понеже числата α и $\alpha+n$ имат еднакъв знак, по теоремата за субституциите заключаваме, че числата P и P_1 са едновременно четни или нечетни. Следователно $P_1\geqq P$, т. е. полиномът $\varphi(x)$ има поне толкова положителни нули като полинома $f(x)$. Подобно виждаме, че $\varphi(x)$ има поне толкова отрицателни нули, колкото полиномът $f(x)$.

Ако $f(0)=0$, нека $x=0$ е m-кратна нула на $f(x)$, т. е. $f(x)=x^mf_1(x)$ като полинома $f_1(x)$, степента на който е $n-m$, не се анулира за $x=0$. Но тогава ще имаме

$$\varphi(x)=\alpha f(x)+xf'(x)=x^m[(\alpha+m)f_1(x)+xf_1'(x)].$$

Понеже числото $\alpha+m$ е извън интервала $(-(n-m), 0)$, то по предното полиномът $(\alpha+m)f_1(x)+xf_1'(x)$ има поне толкова реални нули, колкото полиномът $f_1(x)$ и следователно $\varphi(x)$ има поне толкова реални

нули, колкото $f(x)$. Ще отбележим, че от доказателството следва, че в случая, когато $f(x)$ има само реални нули, всяка многократна нула на $\varphi(x)$ е и такава на $f(x)$.

Нека $F(x)=c(x+\alpha_1)(x+\alpha_2)\ldots(x+\alpha_m)$ е полином със само реални нули, които лежат вън от интервала $(0, n)$. Като приложим последователно горното предложение, получаваме полиномите

$$a_0\alpha_1+a_1(\alpha_1+1)x+\ldots+a_n(\alpha_1+n)x^n,$$

$$a_0\alpha_1\alpha_2+a_1(\alpha_1+1)(\alpha_2+1)x+\ldots+a_n(\alpha_1+n)(\alpha_2+n)x^n,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_0\alpha_1\alpha_2\ldots\alpha_m+a_1(\alpha_1+1)(\alpha_2+1)\ldots(\alpha_m+1)x+\ldots+a_n(\alpha_1+n)\ldots(\alpha_m+n)x^n,$$

които ще имат поне толкова реални нули, колкото има $f(x)$. Но последният полином, умножен с c, е полиномът

$$a_0F(0)+a_1F(1)x+a_2F(2)x^2+\ldots+a_nF(n)x^n. \tag{8}$$

Така получаваме следната теорема на Лагер:

Ако полиномът $F(x)$ има само реални нули, които лежат вън от интервала $(0, n)$, то полиномът (8) *има поне толкова реални нули, колкото полиномът $f(x)$.*

Ако нулите на $f(x)$ са всичките реални, то всяка многократна нула на (8) е и такава на $f(x)$.

5. Теорема на Декарт. Тази теорема ни дава възможност да намерим една горна граница на броя на положителните корени. Нека е дадено едно уравнение с реални коефициенти, като лявата му част е наредена по растящи или намаляващи степени на x. Предполагаме, че членовете с коефициенти нула са изпуснати. Казваме, че между два последователни члена има вариация на знаците или само вариация (промяна), ако тези членове са с обратни знаци, и перманенция, ако същите имат еднакви знаци. Така в уравнението

$$x^4+x^3-3x^2+1=0$$

има две вариации и една перманенция. Теоремата на Декарт гласи:

Броят на положителните корени на едно уравнение е най-много равен на броя на вариациите в лявата му част. Ако този брой е по-малък, разликата между двата броя е четно число.

Доказателството се основава на следната лема на Сегнер:

Ако умножим един реален полином $\varphi(x)$ с бинома $x-c$, $(c>0)$, то броят на вариациите се увеличава с нечетно число.

Действително нека

$$\varphi(x)=ax^n+\ldots+bx^{m+1}-rx^m-\ldots-dx^{l+1}+kx^l+\ldots+gx^{p+1}-$$
$$-hx^p-\ldots-A,$$

като сме записали явно знаците на коефициентите, гдето значи

$$a>0,\ldots,\ b>0,\ r\geqq 0,\ldots,\ d>0,\ k\geqq 0,\ldots,\ g>0,\ h\geqq 0,\ldots,\ A>0.$$

Нека броят на вариациите му да бъде V (в случая, понеже $A>0$, V е очевидно нечетно). Да образуваме полинома

$$\varphi(x)(x-c)=ax^{n+1}\pm\ldots-(r+bc)x^{m+1}\pm \tag{10}$$

$$\pm\ldots+(k+dc)x^{l+1}\pm\ldots-(h+gc)x^{p+1}\pm\ldots+Ac.$$

Ако с V_1 означим броя на вариациите на този полином, то понеже свършва с положителен член, V_1 е четно число, отгдето следва, че V_1-V е нечетно. Ще докажем, че $V_1\geqq V+1$. Действително в (10) от a до $-(r+bc)<0$ имаме поне една вариация, докато в (9) от a до $-r$ имаме една вариация. От $-(r+bc)<0$ до $(k+dc)>0$ в (10) имаме поне една вариация, докато от $-d$ до k в (9) имаме една вариация и т. н., виждаме, че на една група членове с една вариация в (9) съответствуват също група членове с една вариация в (10). Понеже до $-h$ в (9) имаме V вариации, то в (10) до $-(h+gc)$ имаме поне V вариации. Но от този член до Ac имаме поне още една, така че V_1 ще бъде най-малко равно на $V+1$.

Ако полиномът (9) завършваше с положителен член, разсъжденията са същи.

Нека сега на уравнението

$$f(x)=0 \tag{11}$$

положителните корени са $c_1, c_2,\ldots, c_p$. Ако V е броят на вариациите на лявата му част, то ще установим, че $V=p+2\mu$, гдето μ е цяло неотрицателно число. Имаме

$$f(x)=(x-c_1)\ (x-c_2)\ldots(x-c_p)\varphi(x),$$

гдето $\varphi(x)$ има за корени отрицателните и имагинерни корени на $f(x)$. Нека броят на вариациите на полинома $\varphi(x)$ да бъде $V'\geqq 0$. Тогава по лемата на Сегнер полиномът

$$(x-c_1)\varphi(x)$$

ще има поне $V'+1$ вариации. Полиномът

$$(x-c_1)(x-c_2)\varphi(x)$$

ще има поне $V'+2$ вариации. Полиномът

$$(x-c_1)(x-c_2)(x-c_3)\varphi(x)$$

ще има поне $V'+3$ вариации и т. н. Най-после полиномът $f(x)$ ще има поне $V'+p$ вариации, т. е. $V\geqq V'+p\geqq p$.

Ако първият коефициент a_0 и последният a_n в $f(x)$ имат еднакъв знак, то V е четно число, но и p е четно по теоремата на субституциите, понеже $f(0)$ и $f(\infty)$ са с еднакъв знак. Ако $a_0a_n<0$, то V е нечетно, но и p е нечетно. Така се вижда, че $V-p$ е винаги четно число или нула в частност.

От теоремата веднага следва, че броят на отрицателните корени на $f(x)$ е най-много равен на броя на вариациите на $f(-x)$. Ако е по-малък, разликата е четно число.

Така за уравнението

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

броят V на вариациите е две, а броят на вариациите на

$$f(-x)=-x^5-x^3-2x^2-5x+3$$

е единица. Даденото уравнение (12) има или два положителни корена, или нито един и само един отрицателен.

Ако уравнението е пълно, т. е. никой коефициент не е нула, то очевидно вариациите на $f(-x)$ отговарят на перманенции в $f(x)$. Така че в този случай броят на отрицателните корени не надминава броя на перманенциите на даденото уравнение. Ако полиномът $f(x)$ е от n-та степен и означим с V броя на вариациите му, а с V_1 броя на вариациите на $f(-x)$, то ако полиномът е пълен, очевидно $V+V_1=n$. От пълния полином можем да получим непълен, като анулираме някои коефициенти, с което могат да се изгубят вариации, но не да се спечелят. Така че за всеки полином от n-та степен имаме

$$V+V_1\leqq n.$$

На основание на това неравенство ще видим, че има един случай, когато броят на положителните корени е точно равен на броя на вариациите.

Ако всички корени на уравнението

$$f(x)=0$$

са реални, то броят на положителните му корени е точно равен на броя на вариациите на $f(x)$ и броят на отрицателните корени е равен на броя на вариациите на $f(-x)$.

Нека p и q са съответно броят на положителните и отрицателните корени, V и V_1 са броят на вариациите в $f(x)$ и $f(-x)$. Тогава ще имаме

$$p\leqq V,\quad q\leqq V_1,\quad V+V_1\leqq n,$$

гдето n е степента на уравнението. Ако допуснем, че имаме поне в една от първите две релации знака неравенство, напр. $p<V$, то като ги съберем, ще имаме

$$p+q<V+V_1\leqq n,$$

т. е. $p+q<n$, което е невъзможно, понеже всички корени са реални. Остава следователно

$$p=V,\quad q=V_1.$$

6. Доказателство и обобщение на Лагер. Друго едно доказателство на теоремата на Декарт, което позволява и обобщение, е това на Лагер. Нека е дадено уравнението

$$f(x)=ax^n+\ldots+bx^m+cx^l+\ldots+dx^k=0,$$

гдето числата $n,\ldots,m,\ldots,l,\ldots,k$ са положителни и намаляващи. Можем да допуснем даже, че са само реални и намаляващи. Да означим с V броя на вариациите на коефициентите. Ще докажем по индуктивен път, че броят p на положителните корени е най-много равен на V, използувайки теоремата на Рол. Ако $V=0$, то всички коефициенти имат еднакъв знак, следователно и $p=0$. Нека $V>0$ и да допуснем, че теоремата е доказана за полином с $V-1$ вариации. Нека $\mu>0$ и да образуваме функцията

$$y=x^{-\mu}f(x),$$

която има същите положителни корени като $f(x)$. По теоремата на Рол производната

$$y'=x^{-\mu}f'-\mu x^{-\mu-1}f=x^{-\mu-1}(xf'-\mu f)$$

и следователно $f_1=xf'-\mu f$ има поне $p-1$ положителни корена. Впрочем това можем да докажем директно, без да си служим с функцията y, както направихме по-рано. Нека b и c да са с обратни знаци и да изберем μ така, че

$$l<\mu<m;$$

тогава в полинома

$$f_1=(n-\mu)ax^n+\ldots+(m-\mu)bx^m+(l-\mu)cx^l+\ldots+(k-\mu)dx^k$$

коефициентите $(n-\mu)a,\ldots,(m-\mu)b$ имат еднакви знаци с $a,\ldots,b$, а коефициентите $(l-\mu)c,\ldots,(k-\mu)d$ имат обратни знаци на $c,\ldots,d$, но $(m-\mu)b$ и $(l-\mu)c$ са с еднакви знаци. Броят на вариациите на полинома f_1 е равен на $V-1$. Ако с p' означим броя на положителните му корени, то ще имаме

$$p'\leqq V-1 \text{ и } p-1\leqq p',$$

отгдето $p\leqq V$. От доказателството или директно с умножаване на някоя степен на x е очевидно, че могат степените $n,\ldots,m,\ldots,k$ да се предполагат и отрицателни. Нека една функция $f(x)$ е представена с реда на Тейлор

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n+\ldots,$$

който да предположим, че е сходящ за значения на x, за които $|x|<r$, и нека има краен брой V вариации на коефициентите. Тогава, ползувайки се от свойствата на реда на Тейлор, веднага се вижда, че доказателството на Лагер остава в сила за нулите, които са в интервала $(0, r)$. Значи броят p на нулите в интервала $(0, r)$ не надминава броя на вариациите. Ако r е радиусът на сходимост и при $x=r$ редът е разходящ, то лесно е да установим, че $V-p$ е четно число. Действително, понеже по предположение V е крайно, то от известно място всички коефициенти ще имат еднакъв знак; ако той е положителен, $f(x)$ ще клони към $+\infty$, когато $x\to r$, а ако е отрицателен, $f(x)\to-\infty$ при $x\to r$, отгдето следва, понеже $f(0)=a_0$, че ако V е четно, и p е четно, ако V е нечетно, и p е нечетно, т. е. $V-p$ е винаги четно число.

Нека отбележим, че теоремата на Рол е в сила за разглежданите функции. Като приложение ще установим следната теорема на Лагер:

Нека

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0$$

е алгебрично уравнение и а е произволно положително число. Тогава броят на корените му, по-големи от а, не надминава броя V на вариациите на редицата

$$\begin{aligned}
&f_0(a)=a,\\
&f_1(a)=a_0a+a_1,\\
&f_2(a)=a_0a^2+a_1a+a_2,\\
&\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&f_n(a)=f(a)=a_0a^n+a_1a^{n-1}+\ldots a_n
\end{aligned}$$

и разликата между двата въпросни броя е четно число.

Действително

$$\frac{f(x)}{x-a}=f_0(a)\,x^{n-1}+f_1(a)\,x^{n-2}+\ldots+f_{n-1}(a)+\frac{f(a)}{x-a}.$$

Ако $x>a$, то можем да развием $\frac{f(a)}{x-a}$ по степените на $\frac{a}{x}$. Като положим $\frac{1}{x}=y$, дясната част на горното равенство се представя така

$$f_0(a)y^{-n+1}+f_1(a)y^{-n+2}+\ldots+f_{n-1}(a)+f(a)y+af(a)y^2+\ldots$$

и по предния резултат броят на нулите на тази функция в интервала $\left(0,\frac{1}{a}\right)$ не надминава V и разликата между двата броя е четно число.

Понеже на значения на y в $\left(0,\frac{1}{a}\right)$ отговарят значения на x, по-големи от a, то изказаната теорема следва непосредствено.

Ако приложим предната теорема за уравнението

$$x^nf\left(\frac{1}{x}\right)=0,$$

получаваме, че броят на корените на даденото уравнение $f(x)=0$, които лежат в интервала $(0,\ a)$, $a>0$, не надминава броя на вариациите на редицата

$$\begin{aligned}
&a_n,\\
&a_n+a_{n-1}\,a,\\
&a_n+a_{n-1}\,a+a_{n-2}\,a^2,\\
&\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&a_n+a_{n-1}a+a_{n-2}\,a^2+\ldots+a_0a^n
\end{aligned}$$

и разликата между двата броя е четно число.

Теоремата на Декарт се доказва посредством теоремата на Рол и по следния начин. Да означим с p броя на положителните корени на уравнението

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_n=0$$

и с V—броя на вариациите на коефициентите му. Без ограничение можем да предположим, че $a_n a_0 \neq 0$. Видяхме, че по теоремата на субституциите следва, че $V-p$ е четно число. Предполагаме, че теоремата на Декарт е установена за уравненията от степен, по-малка от n и нека p' е броят на положителните корени на $f'(x)=0$ и V' е броят на вариациите на коефициентите му. Очевидно е, че $V' \leqq V$. Съгласно с условието $p' \leqq V'$ и по следствието от теоремата на Рол $p \leqq p'+1$. Следователно

$$p \leqq V'+1 \leqq V+1.$$

Понеже $V-p$ е четно число, то предното неравенство става $p \leqq V$. Но теоремата е очевидно вярна за уравнения от първа степен и така установихме, че тя е вярна за уравнения от произволна степен.

7. Теорема на Бюдан—Фурие.[1] Нека

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_n=0 \tag{13}$$

е дадено уравнение с реални коефициенти. Редицата от производни

$$f(x),\ f'(x),\ f''(x),\ldots,\ f^{(n-1)}(x),\ f^{(n)}(x) \tag{F}$$

се нарича редица на Фурие. Нека α и β са произволни реални числа, $\alpha<\beta$. Теоремата на Бюдан—Фурие гласи: *броят на кореното на уравнението* (13), *които се намират в интервала* (α, β), не надминава броя на загубените вариации от редицата на Фурие (F), когато x расте от α до β. Ако броят на поменатите корени е по-малък, разликата е четно число.

Значи, ако с V_a означим броя на вариациите на редицата

$$f(\alpha),\ f'(\alpha)\ f''(\alpha),\ldots,f^{(n)}(\alpha),$$

с V_β броя на вариациите на

$$f(\beta),\ f'(\beta),\ f''(\beta),\ldots,f^{(n)}(\beta)$$

и с m броя на корените в интервала (α, β), то

$$V_\alpha - V_\beta = m+2\mu,$$

гдето $\mu \geqq 0$ е цяло число.

---

[1] Най-напред открита от Budan в 1811 г. и представена от него в Парижката академия на науките. По-сетне отново е дадена с пълното ѝ развитие от Fourien в 1831 г. в съчинението му „Analyse des équations".

Когато x се мени от α до β, може да стане промяна във вариациите на редицата (F), ако някои функции си променят знака, т. е. (вследствие на непрекъснатостта) се анулират.

Нека x премине през един корен a на уравнението (13), който за общност да предположим, че е k-кратен. Да разгледаме тогава редицата

$$f(x),\ f'(x),\ldots,f^{(k-1)}(x),\ f^{(k)}\ (x).$$

Пред корена a в близко съседство всеки два полинома от тази редица имат обратни знаци, т. е. има в редицата k вариации. След корена a всички са с еднакви знаци. Значи губят се от тази редица k вариации, когато x премине през a.

Нека b да е корен на една междинна функция $f^{(\lambda)}(x)$, който за общност да бъде μ-кратен. Да разгледаме редицата

$$f^{(\lambda-1)}(x),\ f^{(\lambda)}(x),\ f^{(\lambda+1)}(x),\ldots,f^{(\lambda+\mu-1)}(x),\ f^{(\lambda+\mu)}(x). \tag{14}$$

От втората нататък всички функции пред b в близко съседство имат обратни знаци, т. е. образуват μ вариации, а след корена b имат еднакви знаци. Значи губят се μ вариации, когато x премине през b. Ако μ е четно, то $f^{(\lambda)}(x)$ не си променя знака, когато x мине през b, следователно редицата (14) губи μ вариации, т. е. четно число. Ако μ е нечетно, $f^{(\lambda)}(x)$ си променя знака, когато x минава през b. Ако между $f^{\lambda-1)}(x)$ и $f^{(\lambda)}(x)$ е имало вариация, то тя изчезва, ако е нямало, появява се една. Следователно броят на изгубените вариации от (14) е равен на $\mu\pm1$, т. е. четно число.

Така виждаме, че от анулирането на $f(x)$ се губят толкова вариации, колкото е броят на корените в интервала (α, β), а от анулирането на междинна функция се губят четно число вариации, с което теоремата на Бюдан—Фурие е доказана.

Пример. Да се определи броят на корените на уравнението

$$f(x)=x^4-3x^2+x-1=0$$

в интервалите

$$(-2,\ -1),\ (-1,\ 0),\ (0,\ 1),\ (1,\ 2).$$

Имаме

	-2	-1	0	1	2
$f(x)=x^4-3x^2+x-1$	$+$	$-$	$-$	$-$	$+$
$f'(x)=4x^3-6x+1$	$-$	$+$	$+$	$-$	$+$
$f''(x)=12x^2-6$	$+$	$+$	$-$	$+$	$+$
$f'''(x)=24\,x$	$-$	$-$	0	$+$	$+$
$f^{IV}(x)=24$	$+$	$+$	$+$	$+$	$+$
$V=$	4	3	3	1	0

В интервала $(-2, -1)$ уравнението има един корен. В интервала $(0, 1)$ има два или няма корени и в $(1, 2)$ има един корен.

Лесно е да се види, че теоремата на Декарт е следствие от тази теорема. Именно броят на положителните корени е броят на корените в интервала $(0, \infty)$. В редицата на Фурие трябва да поставим $\alpha=0$, $\beta=\infty$. Нека уравнението е следното:

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_n=0,$$

на което производните са

$$\begin{aligned}
f'(x)&=na_0 x^{n-1}+\ldots+a_{n-1},\\
f''(x)&=n(n-1)a_0 x^{n-2}+\ldots+2a_{n-2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
f^{(n)}(x)&=n!\,a_0.
\end{aligned}$$

Полиномите $f(x)$, $f'(x),\ldots,f^{(n)}(x)$ при $x=0$ вземат стойностите

$$a_n,\ a_{n-1},\ 2!\,a_{n-2},\ 3!\,a_{n-3},\ldots,(n-1)!\,a_1,\ n!\,a_0,$$

а при $x=\infty$ имат знака на a_0. Следователно

$$V_0=V,\ V_\infty=0,\ V_0-V_\infty=V,$$

гдето V е броят на вариациите на коефициентите.

Въпросът за броя на корените на уравнението (13) в даден интервал (α, β) може да се сведе до една забележка на Якоби към аналогичния въпрос за брой на положителни корени. Ако поставим

$$y=\frac{x-\alpha}{\beta-x},$$

то на всички стойности на x, лежащи между α и β, ще отговарят положителни стойности на y и на стойностите на x вън от интервала (α, β) ще отговарят отрицателни стойности на y. Следователно броят на корените на уравнението (13), които лежат в интервала (α, β), ще е равен на броя на вариациите на коефициентите на уравнението

$$(1+y)^n f\left(\frac{\alpha+\beta y}{1+y}\right)=0$$

или с четно число по-малък.

8. Метод на Лагранж и Коши. Нека е дадено уравнението

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n=0,$$

като предполагаме отначало, че има само прости корени. Да образуваме уравнението $\varphi(y)=0$ от степен $N=\frac{n(n-1)}{2}$, на което корените са квадратите на разликите $(x_r-x_s)^2$ на корените на даденото уравнение. Ако x_r и x_s са два реални корена, то съответният корен $y_1=(x_r-x_s)^2$ на предното уравнение е положителен. Ако x_r и x_s са имагинерни, но $x_s\neq\bar{x}_r$ или единият от тях е имагинерен, то y е имагинерно число.

Ако обаче $x_s=\overline{x_r}$, т. е. $x_r=p+qi$, $x_s=p-qi$, то $y_1=-4q^2$, т. е. y_1 е отрицателно число. Следователно, ако λ е долната граница на положителните корени на уравнението $\varphi(y)=0$, то разликата на два кои да са реални корена на $f(x)=0$ ще бъде по абсолютна стойност по-голяма от $p=+\sqrt{\lambda}$. Всеки интервал с дължина p може да съдържа най-много един корен на даденото уравнение. Въз основа на това отделянето на реалните корени на даденото уравнение става така: разглеждаме интервалите с дължина

$$l,\ l+p,\ l+2p,\ \ldots$$

от долната граница l на реалните корени до горната граница L на същите корени. Ако $f(x)$ за краищата на един интервал

$$(l+kp,\ l+(k+1)\ p)$$

има различни знаци, то в него има един корен, ако $f(x)$ има еднакви знаци, то в този интервал няма корен на уравнението $f(x)=0$.

Този метод е опростен от Коши, като се избягва получаването на уравнението $\varphi(y)=0$, което води до дълги пресмятания, особено при голямо n. Нека r е една горна граница на модулите на всички корени на даденото уравнение. Така в началото видяхме, че ако α е модулът на най-големия по абсолютна стойност коефициент $a_1, a_2, \ldots, a_n$ и $\alpha_0=|a_0|$, то може да се приеме

$$r=1+\frac{\alpha}{\alpha_0}.$$

Нека D е дискриминантата на даденото уравнение

$$D=a_0^{2n-2}(x_1-x_2)^2(x_1-x_3)^2\ldots(x_{n-1}-x_n)^2.$$

Понеже $|x_i|<r$, ще имаме

$$|D|<a_0^{2n-2}(x_1-x_2)^2(2r)^{n(n-1)-2},$$

като предполагаме, че x_1, x_2 са реални. Оттук имаме

$$|x_1-x_2|>\frac{\sqrt{|D|}}{|a_0|^{n-1}(2r)^{\frac{n(n-1)}{2}-1}}=\rho.$$

Значи разликата на два кои да е реални корена по абсолютна стойност ще бъде по-голяма от ρ.

Особено пригоден става този метод, когато коефициентите a_i са цели числа. Понеже D е цяла рационална функция на $a_0, a_1, \ldots, a_n$ с коефициенти цели числа, то в случай D е цяло число, отлично от нула. Значи $|D|\geqq 1$, откъдето следва

$$|x_1-x_2|>\frac{1}{a_0|^{n-1}(2r)^{\frac{n(n-1)}{2}-1}}=\rho_1,$$

така че ρ в метода може да се замести с ρ_1, като не става нужда да се пресмята D.

Лесно се вижда, че методът на Лагранж остава приложим и когато даденото уравнение има многократни корени, като се вземе под внимание, че на тях отговарят корени нула на $\varphi(y)$, които могат да се отстранят. Обаче приложението на тези методи е свързано с много изпитвания, понеже числата ρ, ρ_1 са малки, което прави метода непрактичен.

9. Теорема на Щурм.[1] Тази теорема ни дава точния брой на корените на едно уравнение в един даден интервал. Нека е дадено уравнението

$$f(x)=0. \tag{20}$$

Върху $f(x)$ и $f'(x)$ прилагаме алгоритма на Евклид за търсене на общ най-голям делител, като при всяко деление вземаме остатъците с обратен знак. Ако така взетите остатъци с обратен знак са

$$R_1,\ R_2,\ \ldots,\ R_k,$$

то ще имаме

$$\begin{aligned} f(x) &= f'(x)Q_1 - R_1, \\ f'(x) &= R_1 Q_2 - R_2, \\ R_1 &= R_2 Q_3 - R_3, \\ &\cdots\cdots\cdots \\ R_{k-2} &= R_{k-1} Q_k - R_k. \end{aligned} \tag{21}$$

Редицата от полиноми

$$f(x),\ f'(x),\ R_1,\ R_2,\ \ldots,R_k \tag{S}$$

е редица на Щурм. Да предположим отначало, че $f(x)=0$ няма многократни корени. Тогава, понеже $f(x)$ и $f'(x)$ нямат общ делител, последният остатък R_k ще бъде константа, отлична от нула. Теоремата на Щурм гласи:

Нека уравнението (2) има само прости корени. Броят на реалните му корени, които се намират в интервали (α, β), $\alpha < \beta$, е точно равен на броя на загубените вариации от редицата (S) на Щурм, когато x се мени от α до β.

Доказателството на теоремата се основава на четири свойства, които притежават функциите на Щурм. Първото свойство е познатата вече лема: пред корен на $f(x)=0$ в близко съседство с него $f(x)$ и $f'(x)$ имат обратни знаци, а след корена — еднакви знаци. Второто свойство е, че две съседни функции не могат да се анулират едновременно. Действително, ако x анулира например R_1 и R_2, то от (21) получаваме последователно $R_3=0$, $R_4=0, \ldots, R_k=0$ за това значение на x, което е невъзможно. Третото свойство е, че ако една междинна функция се анулира, съседните ѝ функции са с обратни знаци. Действително, ако за едно x, $R_i=0$, то от релацията

---

[1] Ch. Sturm, „Memoire sur la résolution des équations numériques", Bulletin de Férussac, Paris (1829).

$$R_{i-1}=R_i\,Q_{i+1}-R_{i+1},$$

имаме $R_{i-1}=-R_{i+1}$. Най-сетне четвъртото свойство е, че последната функция R_k не се анулира, понеже е константа, отлична от нула.

Когато x расте от α до β, промени в броя на вариациите могат да станат само при анулиране на някои функции от (S). Ако x мине през някой корен на $f(x)=0$, то понеже отношението

$$\frac{f(x)}{f'(x)}$$

минава по лемата от отрицателно в положително, редицата (S) ще изгуби една вариация. Значи от анулирането на $f(x)$ редицата ще губи толкова вариации, колкото е броят на корените на (20) в интервала (α, β).

Сега лесно ще видим, че при анулирането на една междинна функция нито се губят, нито печелят вариации. Така нека при $x=b$ се анулира R_i. Съседните функции R_{i-1} и R_{i+1} по второто свойство при $x=b$ са отлични от нула. Можем следователно $\varepsilon>0$ да изберем така малко, че в интервала $(b-\varepsilon,\ b+\varepsilon)$ тѣ да не се анулират. Ако нанесем знаците на тези функции при $x=b-\varepsilon$ и $x=b+\varepsilon$, ще имаме

x	R_{i-1}	R_i	R_{i+1}
$b-\varepsilon$	$+$ ..... $-$	$\pm$	$-$. $+$
$b+\varepsilon$	$+$ ..... $-$	$\pm$	$-$ ..... $+$

Виждаме, че какъвто и да е знакът на R_i, пред и след корена b имаме винаги само една вариация. С това теоремата на Щурм е доказана.

От доказателството се вижда, че при образуване на редицата на Щурм можем да се спрем до функция R, която не се анулира в интервала (α, β).

Пример. Дадено е уравнението

$$f(x)=x^5-2x-3=0.$$

За редицата на Щурм получаваме

$$\begin{aligned} f(x)&=x^5-2x-3,\\ f'(x)&=5x^4-2,\\ R_1&=8x+15,\\ R_2&<0. \end{aligned}$$

При $x=-\infty$ имаме две вариации, а при $x=\infty$ имаме една вариация. Уравнението има 1 реален корен и четири имагинерни. При $x=1$ имаме две вариации, а при $x=2$ само една. Коренът е в интервала $(1, 2)$.

10. Случай на многократни корени. По-рано предположихме, че уравнението (20) има само прости корени. Ако това уравнение има многократни корени, ще видим, че теоремата остава в сила, само че многократните корени трябва да се броят като прости.

Нека R_k е последният остатък, отличен от нула. Тогава този полином представлява общият най-голям делител на $f(x)$ и $f'(x)$. Нека разделим всички функции (S) с R_k и да поставим

(22) $$\frac{f(x)}{R_k}=U,\ \frac{f'(x)}{R_k}=U_1,\ \frac{R_1}{R_k}=U_2,\ldots,\ \frac{R_{k-1}}{R_k}=U_k,\ U_{k+1}=1.$$

Тогава от (21) с разделяне на R_k получаваме

(23) $$\begin{aligned} U &= U_1Q_1-U_2,\\ U_1 &= U_2Q_2-U_3,\\ &\cdots\cdots\cdots\\ U_{k-1} &= U_kQ_k-U_{k+1}. \end{aligned}$$

Оттук лесно е да се види, че редицата от полиноми

(S') $$U,\ U_1,\ U_2,\ldots,U_k,\ U_{k+1}$$

има всичките четири свойства на една редица на Sturm. Действително корените на

$$U=\frac{f(x)}{R_k}=0$$

са все прости, защото това са корените на

$$f(x)=0,$$

като на многократните корени кратността е намалена на единица. Освен това отношението

$$\frac{U}{U_1}=\frac{f(x)}{f'(x)},$$

като се анулира, минава от отрицателно в положително. Другите свойства произлизат аналогично на по-рано от равенствата (23). Следователно броят на корените на $U=0$ в интервала (α, β) ще бъде равен на броя на загубените вариации от редицата (S'), когато x се мени от α до β.

Редицата (S) се получава от редицата (S') с умножение с R_k. Следователно броят на вариациите им е еднакъв. Така получаваме:

Броят на корените на уравнението $f(x)=0$, които се намират в интервала (α, β), е равен на броя на загубените вариации от редицата (S), когато x расте от α до β, при това многократните корени се броят като прости.

11. Обобщение. Теоремата на Щурм, както той е показал, може да се обобщи. Да разгледаме една редица от полиноми

(f) $$f(x),\ f_1(x),\ f_2(x),\ldots,f_m(x),$$

която има всички свойства на редицата (S) с изключение на първото свойство. Именно за пълнота на изложението ще ги повторим. Тези три свойства са: 1) две последователни функции не се анулират едновременно, 2) ако една междинна се анулира, то съседните ѝ са с обратни знаци и 3) последната $f_m(x)$ си запазва знака в разглеждания интервал. Тези три свойства гарантираха това, че при преминаване през корен на междинна функция нито се печели, нито губи вариация.

Промени значи могат да станат само от анулирането на $f(x)$. При това, когато отношението $\frac{f(x)}{f_1(x)}$ се анулира, като минава от отрицателно в положително, ще се изгуби една вариация. Ако същото отношение мине от положително в отрицателно, ще се спечели една вариация и най-сетне може същото отношение при това преминаване да си запази знака, при което нито се губи, нито печели вариация. Така получаваме теоремата:

Разликата между броя на анулирането на отношението $\frac{f(x)}{f_1(x)}$, като минава от отрицателно в положително, и броя на анулирането на същото отношение, минавайки от положително в отрицателно, когато x расте от α до β, е равна на разликата от броя на вариациите на редицата (f) при $x=\alpha$ и този при $x=\beta$.

Оттук като следствие веднага получаваме следното обобщение на теоремата на Щурм:

Броят на корените на уравнението $f(x)=0$, които се намират в интервала (α, β), е най-малко равен на абсолютната стойност на разликата между броя на вариациите на редицата (f) при $x=\alpha$ и броя на вариациите ѝ при $x=\beta$.

Можем лесно да образуваме такива редици (f). Така нека $\varphi(x)$ е полином от степен, по-ниска от тази на $f(x)$, който няма общ делител с този полином. Да образуваме посредством деление

$$f(x)=\varphi(x)\,Q_1-R_1,$$
$$\varphi(x)=R_1\,Q_2-R_2,$$
$$R_{m-2}=R_{m-1}\,Q_m-R_m$$

редицата от полиноми

$$f(x),\ \varphi(x),\ R_1,\ R_2,\ \ldots,\ R_m.$$

Тогава явно е от изложеното, че тази редица е една редица (f) и следователно за нея е валидно горното обобщение.

12. Полиноми на Лежандър. Да означим с $D^m f(x)$ m-тата производна на $f(x)$. Видяхме по-рано с помощта на теоремата на Rolle, че полиномите

$$P_n(x)=\frac{1}{2^n\,n!}\,D^n[(x^2-1)^n],$$

наричани полиноми на Лежандър или сферични функции, имат само реални нули, които се намират в интервала $(-1, 1)$. Тук ще дадем едно ново доказателство на това, използувайки някои свойства на тези полиноми, които са от значение в много други въпроси. От дефиницията имаме

$$P_0(x)=1,\quad P_1(x)=x,\quad P_2(x)=\frac{2}{3}x^2-\frac{1}{2}.$$

Нека диференцираме по формулата на Лайбниц ($\lambda \geqq 2$).

$$D^\lambda[(x^2-1)^\lambda] = D^\lambda[(x^2-1)^{\lambda-1}(x^2-1)] =$$
$$= (x^2-1)\,D^\lambda[(x^2-1)^{\lambda-1}] + 2\lambda x D^{\lambda-1}[(x^2-1)^{\lambda-1}] +$$
$$+ \lambda(\lambda-1)\,D^{\lambda-2}[(x^2-1)^{\lambda-1}].$$

По същото правило имаме, от друга страна,

$$D^\lambda[(x^2-1)^\lambda] = D^{\lambda-1}[\lambda(x^2-1)^{\lambda-1}\,2x] =$$
$$= 2\lambda x D^{\lambda-1}[(x^2-1)^{\lambda-1}] + 2\lambda(\lambda-1)D^{\lambda-2}[(x^2-1)^{\lambda-1}].$$

Ако първата умножим с 2 и извадим втората релация, ще получим

$$D^\lambda[(x^2-1)^\lambda] =$$
$$= 2(x^2-1)\,D^\lambda[(x^2-1)^{\lambda-1}] + 2\lambda x D^{\lambda-1}[(x^2-1)^{\lambda-1}]$$

и оттук, като разделим с $2^\lambda(\lambda-1)!$,

$$\lambda P_\lambda(x) = (x^2-1)\,P'_{\lambda-1}(x) + \lambda x P_{\lambda-1}(x). \tag{24}$$

Тази формула е валидна и за $\lambda = 1$, понеже двете части са равни на x.

По-нататък имаме за $\lambda \geqq 1$

$$D^{\lambda+1}[(x^2-1)^\lambda] = D^\lambda[\lambda(x^2-1)^{\lambda-1}\,.\,2x] =$$
$$= 2\lambda x D^\lambda[(x^2-1)^{\lambda-1}] + 2\lambda^2 D^{\lambda-1}[(x^2-1)^{\lambda-1}$$

и следователно, като разделим с $2^\lambda\lambda!$,

$$P'_\lambda(x) = xP'_{\lambda-1}(x) + \lambda P_{\lambda-1}(x). \tag{25}$$

Ако умножим (25) с x^2-1, а (24) с x и ги извадим, получаваме

$$(x^2-1)\,P'_\lambda(x) = \lambda x P_\lambda(x) - \lambda P_{\lambda-1}(x). \tag{26}$$

Ако сега в (24) поставим $\lambda+1$ вместо λ и заместим $P'_\lambda(x)$ в (26), получаваме

$$(\lambda+1)P_{\lambda+1}(x) = (2\lambda+1)\,xP_\lambda(x) - \lambda P_{\lambda-1}(x). \tag{27}$$

От тази релация става ясно, че полиномите

$$P_n(x),\quad P_{n-1}(x),\ldots,P_1(x),\quad P_0(x) \tag{28}$$

образуват една Щурмова редица. Освен това от (26) имаме

$$P_\lambda(1) = P_{\lambda-1}(1),\quad P_\lambda(-1) = -P_{\lambda-1}(-1),$$

отгдето, като вземем под внимание, че $P_0 = 1$,

$$P_\lambda(1) = 1,\quad P_\lambda(-1) = (-1)^\lambda.$$

Следователно редицата (28) губи n вариации, когато x расте от -1 до 1, с което по обобщената теорема на Щурм се доказва, че кореите на полинома $P_n(x)$ са реални и са в интервала $(-1, 1)$.

Като по-общ пример да разгледаме редицата от полиноми

$$(A) \qquad Q_n(x),\quad Q_{n-1}(x),\quad Q_{n-2}(x),\ldots,Q_0(x),$$

свързани с релациите

$$(B) \quad Q_k(x)=(\alpha_k x+\beta_k)\,Q_{n-1}(x)-\gamma_k\,Q_{k-2}(x),\quad k=1,\ 2,\ldots,n,\ Q_{-1}=0.$$

Ако поставим

$$Q_k(x)=a_{k0}x^k+a_{k1}\,x^{k-1}+\ldots,$$

то от (B) следва, че

$$a_{k0}=\alpha_k\,a_{k-1,\,0}.$$

От последното равенство се вижда, че $a_{n0}=\alpha_1\,\alpha_2\ldots\alpha_n\,a_{00}$ и следователно полиномът $Q_n(x)$ е само тогава точно от n-та степен, ако числата $\alpha_1,\ \alpha_2,\ldots,\alpha_n,\ a_{00}$ са отлични от нула. Ще предполагаме, че това условие е изпълнено. Тогава полиномът $Q_k(x)$, $k=0,\ 1,\ 2,\ldots,n$ ще бъде точно от степен k. От значение за разни въпроси от математиката са полиноми от разгледания тип, при които числата $\gamma_2,\ \gamma_3,\ldots,\gamma_n,$ $a_{00},\ \alpha_1,\ \alpha_2,\ldots,\alpha_n$ са положителни. Например такива са полиномите на Лежандър, както се вижда от равенството (27). От релациите (B) следва, че полиномите (A) образуват Щурмова редица. Ако V_x означава броя на вариациите на редицата за дадено x, то очевидно имаме $V_{-\infty}=n$ и $V_\infty=0$. Тогава съгласно с обобщената теорема на Щурм следва, че полиномът $Q_n(x)$ има само реални нули, както и полиномите $Q_{n-1}(x)$, $Q_{n-2}(x),\ldots$ Освен това отношението $\frac{Q_n(x)}{Q_{n-1}(x)}$ трябва да преминава от отрицателно в положително n пъти, когато x расте от $-\infty$ до ∞. Следователно между всеки две последователни нули на $Q_n(x)$ полиномът $Q_{n-1}(x)$ трябва да си променя знака. Понеже $Q_n(x)$ е от n-та степен и $Q_{n-1}(x)$ от $(n-1)$-ва степен, то лесно заключаваме, че тези два полинома трябва да имат само реални и прости нули, които се взаимно разделят. Ако числата $\gamma_2,\ \gamma_3,\ldots,\gamma_n$ са отрицателни, но числата $a_{00},\ \alpha_1,\ \alpha_2,\ldots,\alpha_n$ са реални и с алтернириращи знаци, то имаме $V_{-\infty}=0$, $V_\infty=n$ и достигаме до същото заключение. Така установихме следното свойство:

Нека полиномите (A) са свързани с релациите (B), където $\gamma_2,\ \gamma_3,\ldots,\gamma_n$ са положителни числа и $a_{00},\ \alpha_1,\ \alpha_2,\ldots,\alpha_n$ са реални числа, които са едновременно положителни или алтернириращи по знак. Тогава полиномите $Q_n(x)$ и $Q_{n-1}(x)$ имат само реални и прости нули, които се взаимно разделят, т. е. между всеки две последователни нули на $Q_n(x)$ има по една нула на $Q_{n-1}(x)$.

Естествено същото свойство е валидно за всеки чифт от полиноми $Q_k(x)$ и $Q_{k-1}(x)$, $k=1,\ 2,\ldots,n$. Предното свойство може да се обърне. Нека числата $\gamma_2,\ \gamma_3,\ldots,\gamma_n$ са положителни и $a_{00},\ \alpha_1,\ \alpha_2,\ldots,\alpha_n$ — реални.

Да предположим, че полиномите $Q_n(x)$ и $Q_{n-1}(x)$ имат само реални и прости корени, които се взаимно разделят. Тогава отношението $\frac{Q_n(x)}{Q_{n-1}(x)}$ ще премине n пъти през нулата от отрицателно в положително или от положително в отрицателно, когато x расте от $-\infty$ до ∞. Съгласно с обобщената теорема на Щурм редицата (A) трябва да губи или да печели n вариации когато x се променя в казаните граници. Това е възможно само тогава, когато или числата $a_{00}, \alpha_1, \alpha_2, \ldots, \alpha_n$ са с еднакъв знак, или алтернират по знак, т. е. всеки две съседни от тях са с противни знаци. Така можем да формулираме следното свойство.

Нека полиномите (A) да удовлетворяват релациите (B), като числата $\gamma_2, \gamma_3, \ldots, \gamma_n$ са положителни. Нека полиномите $Q_n(x)$ и $Q_{n-1}(x)$ са съответно от n-та и $(n-1)$-ва степен и нулите им са реални и прости, като се взаимно разделят. Тогава числата $a_{00}, \alpha_1, \alpha_2, \ldots, \alpha_n$ са всичките с еднакъв знак или с алтернищащи знаци.

Естествено тогава и нулите на полиномите $Q_k(x)$ и $Q_{k-1}(x)$, $k=2, 3, \ldots, n-1$, ще бъдат реални и прости и ще се взаимно разделят.

13. Теорема на Билер—Хермит. Друг пример е теоремата на Билер—Хермит. Нека

$$f(x)=0$$

е уравнение, на което всичките корени се намират от едната страна на реалната ос. Тогава, като извадим i пред скоби, нека

$$f(x)=U+iV.$$

Уравненията $U=0$ и $V=0$ имат само реални корени. Нека

$$f(x)=(x-a_1-ib_1)(x-a_2-ib_2)\ldots(x-a_n-ib_n),$$

гдето числата $b_k>0$, $k=1, 3, \ldots, n$. Да поставим

$$f_p(x)=(x-a_1-ib_1)(x-a_2-ib_2)\ldots(x-a_p-ib_p)=U_p+iV_p.$$

Тогава при $p\geqq 2$ имаме

$$f_{p-1}(x)(x-a_p-ib_p)=f_p(x),$$

т. е.

$$(U_{p-1}+iV_{p-1})(x-a_p-ib_p)=U_p+iV_p.$$

Като приравним реалните и имагинерни части, получаваме

$$U_p=(x-a_p)U_{p-1}+b_p V_{p-1},$$

(29)

$$V_p=-b_pU_{p-1}+(x-a_p)V_{p-1}.$$

Аналогично при $p+1$:

(30) $$U_{p+1}=(x-a_{p+1})U_p+b_{p+1}V_p.$$

Ако елиминираме от (29) и (30) V_{p-1} и V_p, получаваме

(31) $$b_pU_{p+1}=U_p[b_p(x-a_{p+1})+b_{p+1}(x-a_p)]-$$
$$-U_{p-1}b_{p+1}[b_p^2+(x-a_p)^2].$$

Ако положим $U_0=1$, виждаме лесно, че тази релация е в сила и за $p=1$.

Равенствата (31) ни показват, че полиномите

$$U_n=U,\ U_{n-1},\ U_{n-2},\ldots,U_1,\ U_0 \tag{32}$$

образуват една редица на Щурм. Следователно броят на реалните корени на $U=0$ по обобщената Щурмова теорема е най-малко равен на броя на изменението на вариациите на редицата (32), когато x расте от $-\infty$ до $+\infty$. Но полиномите U_p започват с члена x^p, така че очевидно при $x=-\infty$ редицата (32) има n вариации, а при $x=\infty$ нито една. Така че всички корени на $U=0$ ще бъдат реални. Същото доказателство е в сила и за $V=0$. Ако $a_0\neq 1$, то разсъжденията остават същи, само малко се усложнява писането.

По-сетне с директен метод отново ще докажем тази теорема в обобщен вид.

14. Брой на корените посредством квадратични форми. Пръв Хермит е въвел квадратичните форми за определяне броя на реалните корени в един интервал, основавайки се на закона за инерчността.

Нека е дадено уравнението

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0, \tag{33}$$

на което корените са $\alpha_1,\ \alpha_2,\ldots,\alpha_n$ и

$$S_m=\alpha_1^m+\alpha_2^m+\ldots+\alpha_n^m \tag{34}$$

са степенните сборове. Да разгледаме квадратичната форма

$$F(x_1,\ x_2,\ldots,x_n)=\sum_{\lambda=1}^{n}\sum_{\mu=1}^{n}S_{\lambda+\mu-2}\,x_\lambda\,x_\mu.$$

Броят на различните помежду си корени на уравнението (33) е равен на ранга на формата, а броят на различните реални корени е равен на сигнатурата ѝ.

Ако заместим за $S_{\lambda+\mu-2}$ израза (34), получаваме

$$F(x_1,\ x_2,\ldots,x_n)=\sum_{\lambda=1}^{n}\sum_{\mu=1}^{n}\left(\sum_{\nu=1}^{n}\alpha_\nu^{\lambda+\mu-2}\right)x_\lambda\,x_\mu=$$

$$=\sum_{\nu=1}^{n}\left(\sum_{\lambda=1}^{n}\alpha_\nu^{\lambda-1}x_\lambda\sum_{\mu=1}^{n}\alpha_\nu^{\mu-1}\,x_\mu\right)=$$

$$=\sum_{\nu=1}^{n}(x_1+\alpha_\nu\,x_2+\ldots+\alpha_\nu^{n-1}\,x_n)^2.$$

Нека реалните корени са

$$\rho_1,\ \rho_2,\ldots,\rho_k\ (\rho_i\neq\rho_s)$$

съответно от кратности

$$p_1, p_2, \ldots, p_k.$$

Нека различните имагинерни корени са

$$\delta_1, \delta_2, \ldots, \delta_l, \overline{\delta}_1, \overline{\delta}_2, \ldots, \overline{\delta}_l$$

от кратности $q_1, q_2, \ldots, q_l$, т. е. имаме

$$p_1+p_2+\ldots+p_k+2(q_1+\ldots+q_l)=n.$$

Квадратичната форма има вида

$$F(x_1, x_2, \ldots, x_n)=\sum_{s=1}^{k} p_s(x_1+\rho_s x_2+\ldots+\rho_s^{n-1}x_n)^2+$$

$$+\sum_{\lambda=1}^{l} q_\lambda(x_1+\delta_\lambda x_2+\ldots+\delta_\lambda^{n-1}x_n)^2+$$

$$+\sum_{\lambda=1}^{l} q_\lambda(x_1+\delta_\lambda x_2+\ldots+\overline{\delta}_\lambda^{n-1}x_n)^2].$$

Естествено е, че $k+2l\leqq n$. Ако $k+2l<n$, да означим с $g_1, g_2, \ldots, g_{n-k-2l}$ нови $n-k-2l$ реални числа, които помежду си са все различни, както и от ρ_s. Тогава детерминантата

$$\begin{vmatrix} 1 & \ldots & 1 & 1 & 1 & \ldots & 1 & 1 & 1 & \ldots & 1 \\ \rho_1 & \ldots & \rho_k & \delta_1 & \overline{\delta}_1 & \ldots & \delta_l & \overline{\delta}_l & g_1 & \ldots & g_{n-k-2l} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \rho_1^{n-1} & \ldots & \rho_k^{n-1} & \delta_1^{n-1} & \overline{\delta}_1^{n-1} & \ldots & \delta_l^{n-1} & \overline{\delta}_l^{n-1} & g_1^{n-1} & \ldots & g_{n-k-2l}^{n-1} \end{vmatrix}$$

като детерминанта на Вандермонд е отлична от нула. Ако поставим

$$\delta_\lambda=u_\lambda'+iv_\lambda', \qquad \overline{\delta}_\lambda=u_\lambda'-iv_\lambda',$$

$$\ldots\ldots \qquad \ldots\ldots$$

$$\delta_\lambda^{n-1}=u_\lambda^{(n-1)}+iv_\lambda^{(n-1)}, \quad \overline{\delta}_\lambda^{n-1}=u_\lambda^{(n-1)}-iv_\lambda^{(n-1)},$$

то и детерминантата с реални елементи

$$\begin{vmatrix} 1 & \ldots & 1 & 1 & 0 & \ldots & 1 & 0 & 1 & \ldots & 1 \\ \rho_1 & \ldots & \rho_k & u_1' & v_1' & \ldots & u_l' & v_l' & g_1 & \ldots & g_{n-k-2l} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \rho_1^{n-1} & \ldots & \rho_k^{n-1} & u_1^{(n-1)} & v_1^{(n-1)} & \ldots & u_l^{(n-1)} & v_l^{(n-1)} & g_1^{n-1} & \ldots & g_{n-k-2l}^{n-1} \end{vmatrix}$$

като равна на предидущата, умножена с $\left(\frac{i}{2}\right)^l$, е отлична от нула. Следователно реалната трансформация

$$(35)\quad \begin{cases} x_s' = x_1 + \rho_s x_2 + \ldots + \rho_s^{n-1} x_n, & s = 1, 2, \ldots, k, \\ \left.\begin{aligned} x'_{k+2\lambda-1} &= x_1 + u_\lambda' x_2 + \ldots + u_\lambda^{(n-1)} x_n \\ x'_{k+2\lambda} &= v_\lambda' x_2 + \ldots + v_\lambda^{(n-1)} x_n \end{aligned}\right\} & \lambda = 1, 2, \ldots, l, \\ x'_{k+2l+\mu} = x_1 + g_\mu x_2 + \ldots + g_\mu^{n-1} x^n, & \mu = 1, 2, \ldots, n-k-2l \end{cases}$$

има детерминанта, отлична от нула. С тази трансформация формата F се обръща в

$$\sum_{s=1}^{k} p_s x_s'^2 + \sum_{\lambda=1}^{l} q_\lambda \left[(x'_{k+2\lambda-1} + i x'_{k+2\lambda})^2 + (x'_{k+2\lambda-1} - i x_{k+2\lambda})^2\right] =$$

$$= \sum_{s=1}^{k} p_s x_s'^2 + \sum_{\lambda=1}^{l} 2q_\lambda x'^2_{k+2\lambda-1} - \sum_{\lambda=1}^{l} 2q_\lambda x'^2_{k+2\lambda}.$$

Имаме следователно $k+l$ положителни и l отрицателни квадрати. Рангът на квадратичната форма $F(x_1, x_2, \ldots, x_n)$ е

$$(k+l)+l = k+2l$$

и сигнатурата

$$(k+l)-l = k,$$

с което теоремата е доказана.

От тази теорема следва: Необходимото и достатъчно условие, щото корените на (33) да бъдат реални и различни помежду си, се състои в това, че формата $F(x_1, x_2, \ldots, x_n)$ да бъде положително дефинитна.

Понеже дискриминантата F е

$$\begin{vmatrix} S_0 & S_1 & S_2 & \ldots & S_{n-1} \\ S_1 & S_2 & S_3 & \ldots & S_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ S_{n-1} & S_n & S_{n+1} & \ldots & S_{2n-2} \end{vmatrix},$$

то горното условие е еквивалентно с положителността на детерминантите

$$\begin{vmatrix} S_0 & S_1 & S_2 & \ldots & S_{i-1} \\ S_1 & S_2 & S_3 & \ldots & S_i \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ S_{i-1} & S_i & S_{i+1} & \ldots & S_{2n-2} \end{vmatrix}, \quad i = 2, 3, \ldots, n.$$

Този метод може да се обобщи за намиране броя на корените, които се намират между две реални числа a и b. Ако γ е едно произволно реално число, да разгледаме квадратичната форма

$$G_\gamma(x_1\ x_2, \ldots, x_n)=\sum_{\lambda=1}^{n}\sum_{\mu=1}^{n}(\gamma S_{\lambda+\mu-2}-S_{\lambda+\mu-1})\,x_\lambda\,x_\mu,$$

която, както по-рано, веднага се свежда на формата

$$\sum_{\nu=1}^{n}(\gamma-\alpha_\nu)\,(x_1+\alpha_\nu\,x_2+\ldots+\alpha_\nu{}^{n-1}\,x_n)^2$$

или на

$$\sum_{s=1}^{k}p_s(\gamma-\rho_s)(x_1+\rho_s\,x_2+\ldots+\rho_s{}^{n-1}x_n)^2+$$

$$+\sum_{\lambda=1}^{l}q_\lambda(\gamma-\delta_\lambda)(x_1+\delta_\lambda\,x_2+\ldots+\delta_\lambda{}^{n-1}\,x_n)^2+$$

$$+\sum_{\lambda=1}^{l}q_\lambda(\gamma-\bar{\delta}_\lambda)(x_1+\bar{\delta}_\lambda\,x_2+\ldots+\bar{\delta}_\lambda{}^{n-1}\,x_n)^2.$$

С реалната трансформация (35) формата става

$$\sum_{s=1}^{k}p_s(\gamma-\rho_s)\,x_s'^2+\sum_{\lambda=1}^{l}q_\lambda[(\gamma-u_\lambda'-iv_\lambda')\,(x'_{k+2\lambda-1}+ix'_{k+2\lambda})^2+$$

$$+(\gamma-u_\lambda'+iv_\lambda')\,(x'_{k+2\lambda-1}-ix'_{k+2\lambda})^2]=\sum_{s=1}^{k}p_s(\gamma-\rho_s)\,x_s'^2+$$

$$+2\sum_{\lambda=1}^{l}q_\lambda[(\gamma-u_\lambda')\,(x'^2_{k+2\lambda-1}-x'^2_{k+2\lambda})+2v_\lambda'\,x'_{k+2\lambda}\,x'_{k+2\lambda-1}].$$

Ако сега приложим трансформацията (35)

$$x_s'=x_s'',\ s=1,\ 2,\ldots,k,$$

$$\left.\begin{aligned}x'_{k+2\lambda-1}&=x''_{k+2\lambda-1}+(u'_\lambda-\gamma+v'_\lambda)\,x''_{k+2\lambda}\\ x'_{k+2\lambda}&=x''_{k+2\lambda-1}+(u'_\lambda-\gamma-v'_\lambda)\,x''_{k+2\lambda}\end{aligned}\right\}\quad(\lambda=1,\ 2,\ldots,l),$$

$$x'_{k+2\lambda+\mu}=x''_{k+2l+\mu}\quad(\mu=1,\ 2,\ldots,n-k-2l),$$

която има детерминанта, отлична от нула, понеже $v_\lambda' \neq 0$ и тези линейни уравнения може да бъдат решени спрямо x'', то получаваме с малки пресмятания

$$(37)\qquad \sum_{s=1}^{k} p_s(\gamma-\rho_s)x_s''^2+$$

$$+4\sum_{\lambda=1}^{l} q_\lambda \{v'_\lambda x''^2_{k+2\lambda-1} - v'_\lambda[(\gamma-u'_\lambda)^2+v_\lambda'^2]x''^2_{k+2\lambda}\}.$$

Ако с j означим броя на корените ρ_s, които са по-малки от γ и следователно другите $k-j$ са по-големи от γ, предполагайки, че γ не е корен на $f(x)=0$, то формата (36) има $j+l$ положителни и $k-j+l$ отрицателни квадрати. Сигнатурата е равна на

$$(j+l)-(k-j+l)=2j-k.$$

Ако следователно a и b $(a<b)$ са две реални числа, отлични от ρ_s, то да изберем отначало $\gamma=a$, сетне $\gamma=b$. Ако имаме j_1 реални корена, по-малки от a, и j_2, по-малки от b, така че j_1-j_2 корена между a и b, то сигнатурата е съответно равна на $2j_1-k$ и $2j_2-k$. Разликата между сигнатурите е $2(j_2-j_1)$. Следователно имаме теоремата:

Ако σ_γ означава сигнатурата на квадратичната форма

$$\sum_{\lambda,\mu}^{1\ldots n} (\gamma S_{\lambda+\mu-2}-S_{\lambda+\mu-1})x_\lambda x_\mu,$$

то между две числа a и b $(a<b)$ лежат точно $\frac{1}{2}(\sigma_b-\sigma_a)$ рални и различни корени на уравнението (33), като предполагаме, че a и b не са корени.

15. Друга теорема. Нека $f(x)$ е даден полином и $h>0$. Да поставим

$$\Delta f(x)=f(x+h)-f(x),\quad \Delta^2 f(x)=\Delta f(x+h)-\Delta f(x),\ldots$$

$$\Delta^k f(x)=\Delta^{k-1} f(x+h)-\Delta^{k-1} f(x),\ldots,$$

които са разликите на полинома $f(x)$ или с други думи разликите на числата

$$f(x),\ f(x+h),\ f(x+2h),\ldots,f(x+lh),\ldots$$

В отделянето на корените тези разлики могат да заместят производните. Именно имаме следната теорема, дадена от автора[1]:

---

[1] N. Obreschkoff — Jahresbericht der Deutschen Math. Verein. Band 37 (1928), стр. 234—237.

Нека $f(x)=0$ е алгебрическо уравнение от n-та степен с реални коефициенти и V_x да е броят на вариациите на редицата

$$f(x),\quad \Delta f(x),\quad \Delta^2 f(x),\ldots,\quad \Delta^n f(x).$$

Тогава броят на реалните корени на това уравнение в интервала (p_1, q), гдето $q-p_1$ е кратно на h, е най-много равен на V_p-V_q, гдето

$$p=p_1-(n-1)h.$$

За доказателство виж цитираната работа. При $h=0$ получаваме теоремата на Фурие. Прилагането на тази теорема обаче води до по-прости пресмятания от теоремата на Фурие.

Глава III

МЕТОДИ ЗА ПРЕСМЯТАНЕ НА КОРЕНИТЕ

1. Метод на Нютон. Нека a, b са две числа, които отделят един корен на уравнението

(1) $$f(x)=0.$$

Ако този корен е x, ще имаме

$$x_c=a+h,\quad f(a+h)=0$$

или по формулата на Тейлор

(2) $$f(a)+h\,f'(a)+\frac{h^2}{2!}\,f''(a)+\ldots=0.$$

Нека числата a, b се отличават най-много с единица, така че $h<1$. Следователно ние ще намерим една приближена стойност за h, ако в уравнението (2) пренебрегнем степените на h от втора нататък. Приближението ще бъде толкова по-голямо, колкото е по-малко числото h. Лявата част на (2) се редуцира на първите си два члена

$$f(a)+hf'(a)=0,$$

отгдето

$$h_1=-\frac{f(a)}{f'(a)},$$

(3)

$$a_1=a-\frac{f(a)}{f'(a)},$$

гдето a_1 ще бъде приближена стойност за x. По същия начин, излизайки от b, получаваме друга приближена стойност:

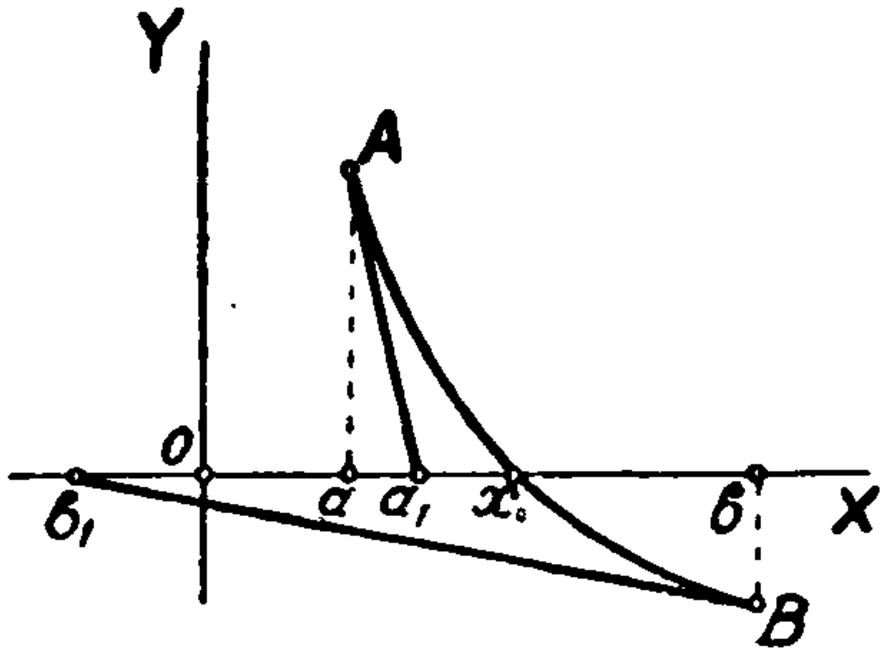
(4) $$b_1=b-\frac{f(b)}{f'(b)}.$$

От тези стойности можем да получим други:

$$a_2 = a_1 - \frac{f(a_1)}{f'(a_1)}, \quad b_2 = b_1 - \frac{f(b_1)}{f'(b_1)}$$

и т. н. Преди да изследваме дали действително се приближаваме така към корена, ще дадем геометричното представяне на метода на Нютон.

Именно ще покажем, че методът се състои в заместване на кривата $y=f(x)$ с нейната тангента в точката с абсциса a или b. Именно в черт. 6 е изобразена кривата $y=f(x)$ между точките A с абсциса a, ордината $f(a)$ и B с координати b, $f(b)$. Уравнението на тангентата T в точката A е



Черт. 6

$$T \equiv y - f(a) = f'(a)(x-a).$$

Пресечната точка на T с абсцисната ос ще бъде дадена с $y=0$:

$$x_1 = a - \frac{f(a)}{f'(a)} = a_1.$$

Аналогично пресечната точка на тангентата в B с абсцисната ос ще бъде дадена с $y=0$:

$$x_1 = b - \frac{f(b)}{f'(b)} = b_1.$$

От чертежа се вижда, че a_1 е по-близко до корена x, но b_1 се отдалечава, така че явява се необходимо едно допълнително изследване.

Нека предположим отначало, че $f'(x)$ не си мени знака в интервала (a, b). Тогава $f(a)$ и $f'(a)$ ще имат обратни знаци, а $f(b)$ и $f'(b)$ ще имат еднакви знаци. Следователно

$$\frac{f(a)}{f'(a)} < 0 \text{ и } a_1 > a,$$

$$\frac{f(b)}{f'(b)} > 0, \; b_1 < b.$$

2. Правило на Фурие. От горе видяхме, че $a_1 > a$, обаче не сме сигурни, че a_1 не надминава x, както например имаме, че $b_1 < a$ в черт. 6. За да бъде a_1 едно истинско приближение, достатъчно е да сме сигурни, че a_1 е по-малко от корена x. Също, за да бъде b_1 истинско приложение , достатъчно е да бъде по-голямо от x.

Нека да предположим, че интервалът (a, b) е така малък, че и $f'(x)$ си запазва знака в него. Ако $x = a+h$ е коренът в интервала (a, b), да развием лявата част на уравнението

$$f(a+h)=0,$$

като се спрем до втория член. Имаме

$$f(a)+hf'(a)+\frac{h^2}{2!}f''(a+\theta h)=0, \qquad 0<\theta<1,$$

отгдето, като решим спрямо h, получаваме

$$h=-\frac{f(a)}{f'(a)}-\frac{h^2 f''(a+\theta h)}{2f'(a)}.$$

За корена x получаваме

$$x=a-\frac{f(a)}{f'(a)}-\frac{h^2 f''(a+\theta h)}{2f'(a)}=a_1-\frac{h^2 f''(a+\theta h)}{2f'(a)}, \tag{5}$$

Оттук се вижда, че условието $x>a_1$ е еквивалентно с

$$\frac{f''(a+\theta h)}{f'(a)}<0.$$

Понеже числото $a+h\theta$ се намира в границите (a, b) и $f''(x)$ има постоянен знак, то горното е еквивалентно с

$$\frac{f''(a)}{f'(a)}<0 \text{ или понеже } \frac{f(a)}{f'(a)}<0, \text{ с}$$

$$\frac{f''(a)}{f(a)}>0.$$

Така получаваме правилото на Фурие: Приближението a_1 е по-точно от a, когато $f(x)$ и $f''(x)$ за $x=a$ имат еднакъв знак. Ако това условие не е изпълнено, не сме сигурни за a_1. Също, ако $f(b)$ и $f''(b)$ имат еднакъв знак, то $x<b_1<b$, т. е. b_1 е по-точно от b. Понеже $f(a)$ и $f(b)$ имат обратни знаци, ясно е, че за едно от числата a и b това условие ще бъде изпълнено.

Лесно е да се види валидността на правилото на Фурие по геометричен път. Така нека разгледаме черт. 6. Кривата е изпъкнала към отрицателната част на Y, $f''(x)$ (както е известно в диференциалното смятане и което впрочем веднага се вижда по това, че ъгловият коефициент на тангентата постоянно расте) е положителна в целия интервал (a, b). Точката a_1, за която

$$f(a)f''(a)>0$$

е по-близко до x, отколкото a. Ако $f''(x)<0$, тогава кривата е изпъкнала към y положително (черт. 7) и b_1, за което

$$f(b)f''(b)>0,$$

е по-точно приближение. Читателят лесно ще разгледа случаите, когато A е под абсцисната ос, а B — над нея.

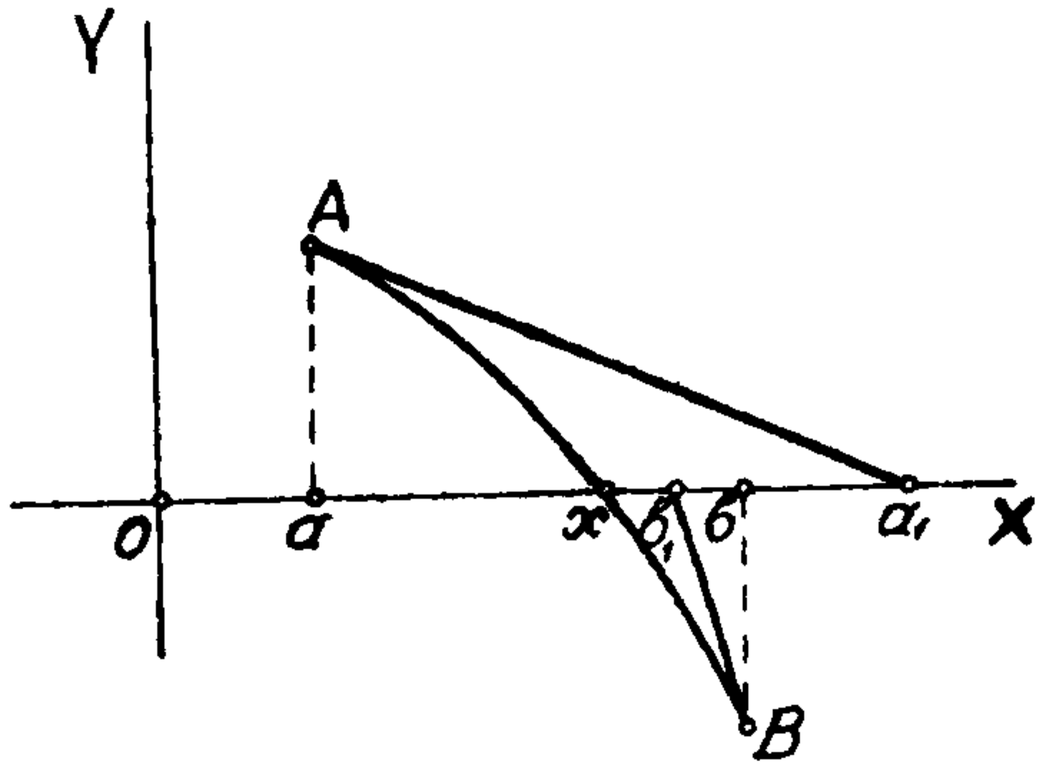
От формулата (5) се получава една горна граница за грешката, която се прави, като вместо корена x се взема числото a_1. Именно нека M е максимумът на $|f''(x)|$ в интервала (a, b), тогава ще имаме

$$|x-a_1| \leqq \frac{h^2|f''(a+h\theta)|}{2|f'(a)|} < \frac{(b-a)^2 M}{2|f'(a)|}.$$

Оттук се вижда, че колкото дължината $b-a$ на интервала (a, b) е по-малка, толкова точността с квадрата ѝ се увеличава. При малко $b-a$ правилото на Фурие може да не се взема под внимание.

За пример нека е дадено уравнението

$$f(x)=x^3+x-3=0.$$



Черт. 7

По теоремата на Декарт това уравнение има един положителен корен, който се намира, както се убеждаваме, със заместване в границите 1,2 и 1,3, т. е.

$$a=1{,}2,\ b=1{,}3.$$

Тогава имаме

$$a_1=a-\frac{f(a)}{f'(a)}=1{,}2+\frac{0{,}072}{5{,}32}=1{,}2135\ldots,$$

$$a_2=a_1-\frac{f(a_1)}{f'(a_1)}=1{,}2135-\frac{0{,}00047856}{5{,}41774675}=$$

$$=1{,}2135-0{,}0000883=1{,}2134117\ldots$$

Ако излезем от горната граница b, то ще имаме

$$b_1=1{,}3-\frac{f(1{,}3)}{f'(1{,}3)}=1{,}3-\frac{0{,}497}{6{,}07}=1{,}22\ldots,$$

$$b_2=1{,}22-\frac{f(1{,}22)}{f'(1{,}22)}=1{,}22-\frac{0{,}035848}{5{,}4652}=1{,}21344\ldots$$

3. Обобщения на метода на Нютон. За получаване на по-добро приближение на корена x на уравнението (1), вместо да пренебрегнем степените на h в (2) от втората нататък, пренебрегваме същите от третата нататък. Така за определяне на h получаваме квадратното уравнение

(6) $$f(a)+hf'(a)+\frac{h^2}{2}f''(a)=0.$$

Ако заместим h с равното му число $x-a$, получаваме за приближена стойност на корена уравнението

$$f(a)+(x-a)f'(a)+\frac{(x-a)^2}{2}f''(a)=0. \tag{7}$$

Последното уравнение показва, че абсцисата на една от пресечните точки на параболата

$$y=f(a)+(x-a)f'(a)+\frac{(x-a)^2}{2}f''(a) \tag{8}$$

с абсцисната ос е приближена стойност на корена x. С непосредствена проверка се вижда, че квадратната функция y и нейната първа и втора производна за $x=a$ приемат стойностите $f(a)$, $f'(a)$ и $f''(a)$, т. е. параболата (8) е оскулачна на кривата $y=f(x)$ в точката $[a, f(a)]$. Следователно геометрическият смисъл на изложения начин за приближено пресмятане на корена x е заместването на кривата $y=f(x)$ с оскулачната парабола към нея в точката $[(a, f(a)]$.

При изложения начин става необходимо да решаваме квадратни уравнения. Ще разгледаме друго обобщение на метода на Нютон, при което се избягва решението на квадратни уравнения и точността на приближенията е от същия порядък. За простота да въведем означенията

$$f(a)=b_0,\quad f'(a)=b_1,\quad \frac{f''(a)}{2!}=b_2,\quad \frac{f'''(a)}{3!}=b_3,\ldots$$

Тогава уравнението (2) може да се пише така:

$$b_0+b_1h+b_2h^2+b_3h^3+b_4h^4+\ldots=0. \tag{9}$$

Ако умножим това равенство с h, получаваме

$$b_0h+b_1h^2+b_2h^3+b_3h^4+b_4h^5+\ldots=0. \tag{10}$$

Като от първото уравнение, умножено с b_1, извадим второто, умножено с b_2, получаваме уравнението

$$b_0b_1+(b_1^2-b_0b_2)h+(b_3b_1-b_2^2)h^3+(b_4b_1-b_3b_2)h^4+\ldots=0,$$

което с пренебрегване на h^3, h^4,... става

$$b_0b_1+(b_1^2-b_0b_2)h=0.$$

Оттук за h получаваме приближената стойност

$$h=-\frac{b_0b_1}{b_1^2-b_0b_2}$$

и следователно за корена x ще имаме приближената формула

$$a_1=a-\frac{f(a)f'(a)}{f'^2(a)-\frac{1}{2}f(a)f''(a)}. \tag{11}$$

За пример да разгледаме уравнението

$$f(x)=x^3+x-3=0,$$

което има един реален корен в интервала (1,2). За него по правилото на Нютон (при $a=1$) получаваме приближението

$$a_1'=1-\frac{f(1)}{f'(1)}=1-\frac{-1}{4}=1\,25.$$

Уравнението (6) става

$$3h^2+4h-1=0,$$

от което за същия корен получаваме

$$a_1''=1+h=1+\frac{-2+\sqrt{7}}{3}=1{,}215\ldots$$

и по формулата (10) намираме приближената стойност на въпросния корен

$$a_1'''=1-\frac{f(1)f'(1)}{f'^2(1)-\frac{1}{2}f(1)f''(1)}=1{,}2105\ldots$$

Като вземем пред вид пресметнатата преди по-точна стойност 1,2134... на корена, виждаме, че приближението a_1'' от трите приближения a_1', a_1'' и a_1''' е най-добро. Въобще приближенията с формулата (11) и посредством квадратното уравнение (8) са по-добри от Нютоновото приближение. В някои случаи, на които не ще се спираме, имаме отклонение от предното твърдение и даже може оскулачната парабола да не пресича абсцисната ос. Ще отбележим, че по следвания метод могат да се получат по-добри приближения от тези, дадени с правилото (11). Именно умножаваме (10) с h

$$b_0h^2+b_1h^3+b_2h^4+b_3h^5+\ldots=0. \tag{12}$$

От уравненията (9),(10) и (12) елиминираме h^2 и h^3, като към (9) прибавим уравненията (10) и (12), умножени съответно с числа λ и μ, които удовлетворяват уравненията

$$\begin{aligned} b_2+\lambda b_1+\mu b_0&=0;\\ b_3+\lambda b_2+\mu b_1&=0. \end{aligned} \tag{13}$$

Така получаваме уравнението

$$b_0+(b_1+\lambda b_0)h+(b_4+\lambda b_3+\mu b_2)h^4+(b_5+\lambda b_4+\mu b_3)h^5+\ldots=0,$$

от което с пренебрегване на степените на h от четвъртата нататък получаваме

$$b_0+(b_1+\lambda b_0)h=0. \tag{14}$$

От (13) (14) за h получаваме

$$h=-\frac{b_0(b_1^2-b_0b_2)}{b_1^3-2b_0b_1b_2+b_0^2b_3}. \tag{15}$$

Новата приближена стойност $a+h$ на корена е по-добра от предните. Трябва да се отбележи, че по следвания път получаваме все по-добри приближени формули, но те стават все по-сложни и непригодни за числени пресмятания. Когато интервалът (a, b), в който лежи коренът x, е малък, най-добре е да използуваме метода на Нютон за по-нататъшни приближения на x. Както видяхме, грешката при последователните приближения с метода на Нютон намалява твърде бързо и освен това изчисленията с този метод са по-прости. При получаване на първите приближения и немалък интервал (a, b) приближенията на Нютон са бавни и по-удобно е да прилагаме уравнението (7) или формулата (11) и даже по-прецизната формула (15).

При известни условия (например при изпълнение на правилото на Фурие) последователните приближения

$$a_{n+1}=a_n-\frac{f(a_n)}{f'(a_n)}, \quad n=0, 1, 2, \ldots, \quad a_0=a$$

клонят към определена граница x_1, която е корен на уравнението, понеже за нея имаме ,

$$x_1=x_1-\frac{f(x_1)}{f'(x_1)},$$

т. е. $f(x_1)=0$. Ще изучим сега реда на приближението x_1-a_{n+1} спрямо предшествуващото приближение x_1-a_n. Без ограничение можем да приемем, че коренът x_1 е прост. Именно многократните корени, както знаем, се лесно отделят. Тогава за $f(x)$ имаме

$$f(x)=(x-x_1)\varphi(x), \tag{16}$$

където функцията $\varphi(x)$ не се анулира за $x=x_1$. Като използуваме (16) и равенството

$$f'(x)=(x-x_1)\varphi'(x)+\varphi(x) \tag{17}$$

за разликата x_1-a_{n+1}, получаваме

$$x_1-a_{n+1}=x_1-a_n+\frac{f(a_n)}{f'(a_n)}=$$

$$=\frac{(x_1-a_n)\varphi(a_n)+(x_1-a_n)[(a_n-x_1)\varphi'(a_n)+\varphi(a_n)]}{f'(a_n)}=-(x_1-a_n)^2\frac{\varphi'(a_n)}{f'(a_n)}.$$

Оттук при $n\to\infty$ получаваме

$$\lim_{n\to\infty}\frac{x_1-a_{n+1}}{(x_1-a_n)^2}=-\frac{\varphi'(x_1)}{f'(x_1)}=-\frac{\varphi'(x_1)}{\varphi(x_1)}.$$

Последното гранично равенство показва, че при всяко следващо приближение грешката е от втори ред спрямо грешката от предхождащото приближение. За пълнота да разгледаме и случая, когато коренът x_1 е p-кратен. Тогава

$$f(x)=(x-x_1)^p g(x),$$

като $g(x_1) \neq 0$. Следвайки същия път, получаваме граничното равенство[1]

$$\lim_{n\to\infty} \frac{x_1 - a_{n+1}}{x_1 - a_n} = \frac{p-1}{p},$$

което показва, че приближението на Нютон в случая $p > 1$ не е така добро. За да получим приближение от по-висок от първи ред, заместваме приближенията със следните[1]

$$a_{n+1} = a_n - p\,\frac{f(a_n)}{f'(a_n)}.$$

С аналогични пресмятания получаваме граничното равенство

$$\lim_{n\to\infty} \frac{x_1 - a_{n+1}}{(x_1 - a_n)^2} = -\frac{\varphi'(x_1)}{p\varphi(x_1)},$$

което показва същия порядък на приближение, както в случая на $p = 1$, т. е. при прост корен.

Да изследваме сега приближенията с формулата (11). Следователно ще трябва да образуваме редицата от числа (приближения)

$$a_{n+1} = a_n - \frac{f(a_n) f'(a_n)}{f'^2(a_n) - \frac{1}{2} f(a_n) f''(a_n)} \tag{18}$$

и да изследваме порядъка на $x_1 - a_{n+1}$ спрямо този на $x_1 - a_n$. Като вземем пред вид формулите (16) (17) и формулата

$$f''(x) = (x - x_1)\varphi''(x) + 2\varphi'(x)$$

(получената от (17) с диференциране), намираме с някои преобразувания следната формула:

$$x_1 - a_{n+1} = (x_1 - a_n)^3 \frac{\varphi'^2(a_n) - \frac{1}{2}\varphi(a_n)\varphi''(a_n)}{f'^2(a_n) - \frac{1}{2} f(a_n) f''(a_n)}.$$

Оттук получаваме граничното равенство

$$\lim_{n\to\infty} \frac{x_1 - a_{n+1}}{(x_1 - a_n)^3} = \frac{\varphi'^2(x_1) - \frac{1}{2}\varphi(x_1)\varphi''(x_1)}{\varphi^2(x_1)},$$

от което се вижда, че грешката $x_1 - a_{n+1}$ е от трети ред спрямо грешката $x_1 - a_n$. Следователно приближенията към корена с редицата (18) са по-бързи.

---

[1] E. Bodewig, Journal für d. reine u. angew. Math., 1949.

4. Regula falsi. Този метод се състои в това, че кривата $y=f(x)$ се замества в интервала (a, b) за x, в който е отделен един корен, със секантата. Нека A е точката $(a, f(a))$, $B(b, f(b))$. Тогава секантата AB има уравнение

$$y-f(a)=\frac{f(b)-f(a)}{b-a}(x-a). \tag{19}$$

Пресечната точка на тази секанта с абсцисната ос се дава с $y=0$ и от (19) имаме за абсцисата ѝ

$$x_1=a-\frac{f(a)}{f(b)-f(a)}(b-a). \tag{20}$$

Приближението е очевидно по-голямо, ако интервалът е по-малък.

Този метод, обяснен алгебрически, се състои в това, че в интервала (a, b) заместваме $f(x)$ с линейна функция. Нека коренът x, лежащ между a и b, е

$$x=a+h, \quad x=b-k.$$

По формулата на Тейлор имаме

$$f(a)=f(x-h)=f(x)-hf'(x)+\frac{h^2}{2!}f''(x)-\ldots,$$

$$f(b)=f(x+k)=f(x)+kf'(x)+\frac{k^2}{2!}f''(x)+\ldots$$

Като пренебрегнем степените от квадрат нагоре на h и k, получаваме приблизително

$$f(a)=-hf'(x),$$

$$f(b)=kf'(x),$$

отгдето имаме

$$\frac{f(a)}{f(b)}=-\frac{h}{k}=\frac{x-a}{x-b}.$$

Ако решим последното уравнение спрямо x, получаваме формулата (20).

Пример. Уравнението

$$f(x)=x^3+x-5=0$$

има един корен между 1 и 2. Ако поставим $a=1$, $b=2$, то (7) дава

$$x_1=1+\frac{-3}{-3-5}=1\frac{3}{8}.$$

Да вземем приближеното 1,4 и поставим

$$a_1=1{,}4, \quad f(1{,}4)=-0{,}86,$$

$$b_1=2, \quad f(2)=+5,$$

коренът е в интервала (1,4, 2). По (7) имаме

$$x_2 = 1{,}4 + \frac{0{,}6 \times 0{,}86}{0{,}86+5} = 1{,}49.$$

Да вземем приближено

$$a_2 = 1{,}4, \quad f(1{,}4) = -0{,}86,$$
$$b_2 = 1{,}5, \quad f(1{,}5) = -0{,}125,$$

то по (20)

$$x_3 = 1{,}5170.$$

Нека

$$a_3 = 1{,}5, \quad f(1{,}5) = -0{,}125,$$
$$b_3 = 1{,}517, \quad f(1{,}517) = +0{,}008,$$

отгдето по (20) имаме приближението

$$x_4 = 1{,}51598.$$

Имаме

$$f(1{,}51598) = -0{,}000002\ldots,$$

а $f(1{,}516) > 0$, с което се вижда, че коренът е в интервала (1,51598, 1,516) и можем да напишем с точност до третия десетичен знак $x = 1{,}515\ldots$

5. Метод на Хорнер. Този метод е подобен на метода на Нютон, но с него се пресмятат последователно десетичните знаци на търсения корен, като се използува делението на Хорнер, с което се запознахме при трансформацията на уравненията. Нека на уравнението

(21) $$f(x) = 0$$

е отделен един корен в интервала $(a, a+1)$, гдето a е цяло число. Намаляваме корените на това уравнение с a по метода на Хорнер. Търсеното уравнение ще бъде

(22) $$f_1(x) = f(a+x) = f(a) + xf'(a) + \ldots = 0.$$

Това уравнение ще има един корен в интервала (0,1), който получаваме по метода на Нютон, като редуцираме уравнението (22) на линейно:

$$x' = -\frac{f(a)}{f'(x)} = 0, a_1 a_2' \ldots$$

Намаляваме корените на (22) с $0, a_1$ и получаваме ново уравнение

$$f_2(x) = f_1(0, a_1 + x) = f_1(0, a_1) + xf_1'(0, a_1) + \ldots = 0.$$

Ако a_1 е точният пръв десетичен знак на корена x, то това уравнение ще има един корен в интервала

$$(0, a_1, \quad 0, a_1 + 1),$$

за който получаваме приближената стойност

$$x'' = -\frac{f_1(0,\ a_1)}{f_1'(0,\ a_1)} = 0{,}0\,a_2\,a_3'' \ldots$$

Тогава намаляваме корените на това уравнение с $0{,}0\,a_2$ и т. н. Така за корена x на (21) получаваме

$$x = a,\, a_1\, a_2\, a_3 \ldots$$

При извършване обаче на тези операции можем да се натъкнем на следните случаи. Може да се случи, че x' да излезе по-голямо от 1. Понеже сме сигурни, че x се намира между a и $(a+1)$, то опитваме тогава $a_1 = 9$. Ако обаче свободният член $f_1(0,\ a_1) = f(a, a_1)$ в $f_2(x)$ има обратен знак на $f(a)$, то ясно е, че сме минали корена, така че намаляваме a_1 на 8, пак изпитваме, докато не се яви този случай. Същото важи и за x'', което трябва да излезе число $< \frac{1}{10}$. Ако $x'' > \frac{1}{10}$, вземаме за $a_2 = 9$ и следим за свободния член на новото трансформирано уравнение. Може да се случи, че сме взели десетичен знак за x, помалък от истинския. Тогава ще забележим, че корекцията, която се получава в новото трансформирано уравнение, ще бъде от същия ред големина, т. е. ако тази цифра е стотица, тя ще бъде пак стотица, а не хилядна. Тогава увеличаваме десетичния знак с една такава единица и продължаваме изпитванията.

Така нека за уравнението

$$x^3 + x - 3 = 0$$

приложим метода на Хорнер. Видяхме, че това уравнение има един корен в интервала (1,2, 1,3), който да пресметнем. Намаляваме корените му с 1,2 с правилото на Хорнер и за по-голяма простота произведенията на 1,2 с коефициентите на частното ги нанасяме на самата схема:

	1	0	1	−3
		1,2	1,44	2,928
1,2		1,2	2,44	−0,072
		1,2	2,88	
		2,4	5,32	
		1,2		
		3,6		

Трансформираното уравнение е

$$x^3 + 3{,}6\,x^2 + 5{,}32\,x - 0{,}072 = 0,$$

което, редуцирано на последните два члена, дава

$$x = \frac{0{,}072}{5{,}32} = 0{,}01.$$

Намаляваме корените на това уравнение с 0,01:

	1	3,5	5,32	—0,072
		0.01	0,0361	—0,053561
0,01		3,61	5,3561	—0,018439
		0,01	0,0362	
		3,62	5,3923	
		0,01		
		3,63		

Трансформираното уравнение е

$$x^3+3{,}63\,x^2+5{,}3923\,x-0{,}018439=0,$$

което дава

$$x=\frac{0{,}018439}{5{,}3923}=0{,}003.$$

Намаляваме корените му с 0,003:

	1	3,63	5,3923	—0,018439
		0,003	0,010899	0,016209597
0,003		3,633	5,403199	—0,002229403
		0,003	0,010908	
		3,636	5,414107	
		0,003		
		3,639		

Получаваме

$$x=\frac{0{,}002229403}{5{,}414107}=0{,}0004.$$

Търсеният корен x_1 на даденото уравнение ще бъде

$$x_1=1{,}2134\ldots,$$

където сме сигурни до хилядните.

6. Метод на Лагранж. Един друг метод, който използува тъй наречените верижни дроби, е този на Лагранж. Нека a, $a+1$ са последователни цели числа, които съдържат един корен на уравнението от n-та степен:

(23) $$f(x)=0.$$

Да поставим

$$x=a+\frac{1}{y},$$

то уравнението за y

$$f_1(y)=y^n f\left(a+\frac{1}{y}\right)=y^n f(a)+\frac{y^{n-1}}{1!}f'(a)+\ldots=0$$

ще има един корен, по-голям от 1, понеже за

$$a+1>a+\frac{1}{y}>a$$

следва $\frac{1}{y}<1$, $y>1$. Нека с последователно заместване на числата 1, 2, 3, ... се намери, че коренът е в интервала

$$(b,\ b+1).$$

Поставяме

$$y=b+\frac{1}{z},$$

като за z получаваме уравнението

$$f_2(z)=z^n f_1\left(b+\frac{1}{z}\right)=0.$$

Това уравнение ще има един корен, по-голям от 1, който нека се намира между целите числа

$$c,\ c+1.$$

Полагаме $z=c+\frac{1}{u}$ и намираме уравнението за u:

$$f_3(u)=u^n f_2\left(c+\frac{1}{u}\right)=0,$$

което ще има един положителен корен между целите числа

$$d,\ d+1,$$

и т. н.

За корена x получаваме приближенията

$$x=a+\frac{1}{y}=a+\frac{1}{b+\frac{1}{z}},\quad a+\frac{1}{b+\frac{1}{c+\frac{1}{u}}}=a+\frac{1}{b+\frac{1}{c+\frac{1}{d+\ldots}}},$$

като той се съдържа между числата

$$a,\ a+1,$$

$$a+\frac{1}{b+1}\ \text{и}\ a+\frac{1}{b},$$

$$a+\frac{1}{b+\frac{1}{c}}\ \text{и}\ a+\frac{1}{b+\frac{1}{c+1}}\ \text{и т. н.}$$

За пример нека е дадено уравнението

$$f(x)=x^3+x-3=0,$$

което има един корен между 1 и 2. Полагаме

$$x=1+\frac{1}{y}$$

и получаваме за y уравнението

$$y^3-4y^2-3y-1=0,$$

което има корен между 4 и 5. Полагаме

$$y=4+\frac{1}{z}$$

и получаваме за z:

$$13z^3-13z^2-8z-1=0.$$

Положителният корен е между 1 и 2. Полагаме

$$z=1+\frac{1}{u}$$

и получаваме за u:

$$9u^3-5u^2-26u-13=0,$$

което уравнение има корен между 2 и 3. Полагаме

$$u=2+\frac{1}{v}$$

и получаваме

$$13v^3-62v^2-49v-9=0.$$

Това уравнение има един корен между 5 и 6. Ако поставим

$$v=5+\frac{1}{w},$$

то получаваме за w

$$179w^3-306w^2-133w-13=0,$$

което има корен между 2 и 3 и т. н. Така получаваме за

$$x=1+\cfrac{1}{4+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{5+\cfrac{1}{2+\ldots}}}}}$$

Оттук получаваме приблизителните стойности

$$1,\ \frac{5}{4},\ \frac{6}{5},\ \frac{17}{14},\ \frac{91}{75},\ \frac{199}{164}=1{,}213\ldots$$

По причина на по-дълги изчисления методът на Лагранж е по-непрактичен от останалите методи, но той има голямо теоретично значение.

7. Метод на Лобачевски — Грефе. Методът на Лобачевски дава всички корени на уравнението, без предварително да сме изследвали реалността им или да сме ги отделили и т. н. Принципът на метода се състои в това, че от даденото уравнение се получава ново, на което кореният са достатъчно високи степени на кореният на даденото уравнение, така че степените на малките по абсолютна стойност корени са много малки сравнително степените на големите корени.

Нека даденото уравнение да бъде

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0 \tag{24}$$

с корени $x_1, x_2, \ldots, x_n$. Да образуваме уравнението

$$y^n+A_1y^{n-1}+A_2y^{n-2}+\ldots+A_n=0 \tag{25}$$

с корени $x_1^m, x_2^m, \ldots, x_n^m$. Да предположим отначало, че кореният са всички реални и че

$$|x_1|>|x_2|>|x_3|>\ldots>|x_n|.$$

Но тогава от (25) имаме

$$x_1^m+x_2^m+\ldots+x_n^m=-A_1$$

или

$$x_1^m\left[1+\left(\frac{x_2}{x_1}\right)^m+\ldots+\left(\frac{x_n}{x_1}\right)^m\right]=-A_1.$$

Когато $m\to\infty$, то в средните скоби изразът клони към единица, така че при достатъчно голямо m с приближение може да се пише

$$x_1^m=-A_1. \tag{26}$$

Също от релацията

$$x_1^m x_2^m+x_1^m x_3^m+\ldots+x_{n-1}^m x_n^m=A_2$$

или

$$x_1^m x_2^m\left[1+\left(\frac{x_3}{x_2}\right)^m+\left(\frac{x_3}{x_1}\right)^m\left(\frac{x_4}{x_2}\right)^m+\ldots\right]=A_2$$

се вижда, че при достатъчно голямо m ще имаме с приближение

$$x_1^m x_2^m=A_2. \tag{27}$$

По-нататък подобно ще имаме приблизително

$$x_1^m x_2^m x_3^m=-A_3 \tag{28}$$

и т. н. От тези релации имаме приблизителните стойности за кореният

$$x_1^m=-A_1,\ x_2^m=-\frac{A_2}{A_1},\quad x_3^m=-\frac{A_3}{A_2},\ \ldots,\ x_n^m=-\frac{A_n}{A_{n-1}}. \tag{29}$$

Тук трябва да се изпита кои значения на m-те корена удовлетворяват уравнението (24). Понеже това изпитване при уравнения от по-висока степен и корен, пресметнат с повече десетични знаци, е сложно, то по-удобно е да се изследват със заместване две граници, между които е този корен.

Получаването на уравнението (25) се значително опростява, като изберем m равно на степен на 2.

Но такава една трансформация може да се получи с последователни трансформации, които се състоят в повдигане в квадрат на корените.

Последната проста трансформация може да се извърши така: нека в даденото уравнение да отделим членовете с четните и нечетните степени на x, т. е.

$$f(x)=\varphi(x^2)+x\psi(x^2)=0.$$

Тогава имаме

$$f(-x)=\varphi(x^2)-x\psi(x^2),$$

$$f(x)f(-x)=\varphi^2(x^2)-x^2\psi^2(x^2)=0.$$

Уравнението за $y=x^2$ ще бъде

$$\varphi^2(y)-y\psi^2(y)=0$$

или

$$b_0y^n-b_1y^{n-1}+b_2y^{n-2}-\ldots+(-1)^nb_n=0, \tag{30}$$

гдето

$$b_0=a_0^2,\ b_1=a_1^2-2a_0a_2,\ b_2=a_2^2-2a_1a_3+2a_0a_4,$$

$$b_3=a_3^2-2a_2a_4+2a_2a_5-2a_0a_6,$$

$$b_4=a_4^2-2a_3a_5+2a_2a_6-2a_1a_7+2a_0a_8,\ldots,$$

$$b_{n-1}=a^2{}_{n-1}-2a_na_{n-2},\ b_n=a_n^2.$$

При тази трансформация членовете, които не са квадрати, постепенно ще стават спрямо последните все по-малки релативно. Така в (25), когато m става достатъчно голямо, отношението на A_2^2 например към A_1A_3 по (26), (27) и (28) е близко до

$$\frac{x_1^{2m}\cdot x_2^{2m}}{x_1^m\cdot x_1^m x_2^m x_3^m}=\left(\frac{x_2}{x_3}\right)^m,$$

така че може да стане произволно голямо.

Именно след известно място коефициентите на трансформираното уравнение (30) ще представляват почти квадрати на коефициентите на по-раншното уравнение. Това именно се случва при направеното предположение, че корените са по абсолютна стойност различни.

Ако корените са реални, то понеже трансформираните уравнения имат положителни корени, в левите им части ще има само вариации. Ако се случи това да не е изпълнено, то показва, че даденото уравнение има имагинерни корени.

Да изследваме случая на имагинерни корени. Нека

$$g(\cos\varphi \pm i\sin\varphi)$$

са два конюговани такива корена. Лявата част на трансформираното уравнение (25) ще има множител

$$x^2-2g^m\cos m\varphi\, x+g^{2m} \text{ или } x^2+f_m x+g^{2m},\ f_m=-2g^m\cos m\varphi,$$

който представлява произведение на биномните множители, отговарящи на корените

$$g^m(\cos m\varphi \pm i\sin m\varphi).$$

Ако за едно $m=2^\lambda$, $m\varphi$ е кратно на π, то нататък постоянно ще имаме същото съотношение, така че двата имагинерни корена ще се държат като равни реални корени.

Ако този случай не се яви, то f_m за различни стойности на m ще се колебае между $-2g^m$ и $2g^m$.

Нека например имаме

$$|x_1|>|x_2|>g>|x_3|>\ldots,$$

тогава

$$\begin{aligned}
-A_1&=x_1^m+x_2^m+\ldots=x_1^m(1+\alpha_1),\\
A_2&=x_1^m x_2^m+\ldots=x_1^m x_2^m(1+\alpha_2),\\
-A_3&=2x_1^m x_2^m g^m\cos m\varphi+\ldots,\\
A_4&=x_1^m x_2^m g^{2m}+\ldots=x_1^m x_2^m g^{2m}(1+\alpha_4),\\
-A_5&=x_1^m x_2^m g^{2m} x_3^m(1+\alpha_5),\\
&\cdots\cdots\cdots\cdots
\end{aligned}$$

гдето α_1, α_2,... са числа, които клонят към нула, когато m клони към безкрайност. От горното се вижда, че ако имаме само два имагинерни корена, то в трансформираните уравнения всички коефициенти освен един ще се отнасят, както в случая само на реални корени. Единият коефициент, в нашия случай A_3, понеже зависи от $\cos m\varphi$, ще се колебае, като може да си изменя знака. Другите коефициенти на трансформираните последователно уравнения ще се приближават към квадратите на тези на предшествуващите уравнения.

От горните релации можем да пишем приблизително

$$x_1^m=-A_1,\ x_2^m=-\frac{A_2}{A_1},\ g^{2m}=\frac{A_4}{A_2},\ x_3^m=-\frac{A_5}{A_4},\ldots,$$

така че, като разделяме последователно коефициентите на трансформираното уравнение, като изпускаме члена, който се колебае, получаваме m-тите степени на x_1, x_2, g, x_3,...

Ако имаме още един чифт конюговани имагинерни корени $g_1(\cos\varphi_1 \pm i\sin\varphi_1)$ и нека

$$g>|x_1>g_1>|x_2|>|x_3|>\ldots,$$

то ще имаме за уравнението (12)

$$-A_1 = 2g^m \cos m\varphi + x_1{}^m + \ldots$$
$$A_2 = g^{2m} + 2x_1{}^m g^m \cos m\varphi + \ldots = g^{2m}(1+\alpha_2),$$
$$-A_3 = g^{2m} x_1{}^m (1+\alpha_3),$$
$$A_4 = 2\varsigma^{2m} x_1{}^m g_1{}^m \cos m\varphi_1 + \ldots,$$
$$-A_5 = g^{2m} x_1{}^m g_1{}^{2m} (1+\alpha_5),$$
$$A_6 = g^{2m} x_1{}^m g_1{}^{2m} x_2{}^m (1+\alpha_6),$$
$$-A_7 = g^{2m} x_1{}^m g_1{}^{2m} x_2{}^m x_3{}^m (1+\alpha_7),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

гдето $\alpha_i \to 0$, когато $m \to \infty$. Оттук се вижда, че ще има два коефициента, именно A_1 и A_4, които ще се колебаят освен в случаите, че едно от числата $m\varphi$, $m\varphi_1$ за някое m е кратно на π, в който случай от известно m нататък имагинерните корени ще се държат като реални двойни. От горните релации имаме приблизително

$$g^{2m} = A_2,\ x_1{}^m = -\frac{A_3}{A_2},\ g_1{}^{2m} = \frac{A_5}{A_3},\ x_2{}^m = -\frac{A_6}{A_5},\ x_3{}^m = -\frac{A_7}{A_6}, \ldots$$

От A_1 и A_4 можем да намерим φ и φ_1. По-добре те се определят от коефициентите на даденото уравнение. Така в случай на само два имагинерни корена

$$g(\cos\varphi \pm i\sin\varphi)$$

имаме

$$2g\cos\varphi + x_1 + x_2 + \ldots = -\frac{a_1}{a_0}.$$

Оттук намираме $\cos\varphi$ и следователно и аргумента φ. Ако уравнението има два чифта имагинерни корени

$$g(\cos\varphi \pm i\sin\varphi),\ g_1(\cos\varphi_1 \pm i\sin\varphi_1),$$

използуваме равенствата

$$2g\cos\varphi + 2g_1\cos\varphi_1 + x_1 + x_2 + \ldots = -\frac{a_1}{a_0};$$

$$\frac{2}{g}\cos\varphi + \frac{2}{g_1}\cos\varphi_1 + \frac{1}{x_1} + \frac{1}{x_2} + \ldots = -\frac{a_{n-1}}{a_n}, \tag{31}$$

където $x_1, x_2, \ldots$ са реалните корени, които предполагаме, че сме пресметнали до желаната точност. Равенствата (31) представляват линейни уравнения спрямо $\cos\varphi$ и $\cos\varphi_1$ и лесно можем да намерим стойностите на тези величини. Ако уравнението има три чифта имагинерни корени

$$g(\cos\varphi \pm i\sin\varphi),\ g_1(\cos\varphi_1 \pm i\sin\varphi_1),\ g_2(\cos\varphi_2 \pm i\sin\varphi_2),$$

към уравненията (31) прибавяме подобните уравнения за $\frac{a_2}{a_0}$ и $\frac{a_{n-2}}{a_n}$, които спрямо неизвестните $t_1 = \cos\varphi_1$, $t_2 = \cos\varphi_2$ и $t = \cos\varphi$ са от втора степен. Тогава от (31) изразяваме t_1 и t_2 посредством t и полу-

ченните изрази за тези неизвестни заместваме в последните две уравнения, които ще бъдат така от втора степен спрямо t и ще имат един общ корен за t, който се лесно намира посредством общия най-голям делител на левите части на получените уравнения. Като намерим по този начин t и определим t_1 и t_2, ще получим и аргументите φ, φ_1 и φ_2 на имагинерните корени. Относно имагинерните корени можем да формулираме следното правило. В последователно трансформираните уравнения имаме толкова колебаещи се членове, колкото чифта конюговани имагинерни корени. Ако в крайното уравнение, което искаме да използуваме за приближено пресмятане на корените, премахнем колебаещите се членове и разделим всеки два последователни така члена, то ще получим m-тите степени на квадратите на модулите на имагинерните корени. Квадратите на модулите се получават от коефициентите, които следват тези, които се колебаят.

Когато уравнението има близки по абсолютна стойност реални корени или близки по модул имагинерни корени, приближенията са по-бавни и става нужда да прибягваме до голям брой последователни трансформации. В такъв случай можем да приложим трансформацията $x=y+\lambda$, в която λ е реално подбрано число. Действително, ако x_1 и x_2 са два реални корена като $|x_1|>|x_2|$, то можем по много начини да изберем λ така, че отношението $\left|\frac{x_2-\lambda}{x_1-\lambda}\right|$ да бъде по-малко от $\left|\frac{x_2}{x_1}\right|$. Ако $r(\cos\varphi+i\sin\varphi)$, $0<\varphi<\pi$, и $r_1(\cos\varphi_1+i\sin\varphi_1)$, $0<\varphi_1<\pi$ са два имагинерни корена с почти равни модули или равни модули, но с различни аргументи при произволно реално λ, модулите на съответните корени на уравнението за y ще се различават. Ако $r'(\cos\varphi'+i\sin\varphi')$ е съответният корен на корена $r(\cos\varphi+i\sin\varphi)$, то ще имаме

$$r(\cos\varphi+i\sin\varphi)=r'(\cos\varphi'+i\sin\varphi')+\lambda,$$

откъдето получаваме релацията между r и r':

$$r'^2=r^2+\lambda^2-2r\lambda\cos\varphi.$$

На Бродески и Смейл се дължи едно приложение на горния начин, което позволява да се намират направо аргументите на имагинерните корени, както и реалните корени със знака им. Именно в трансформацията $x=y+\varepsilon$ предполагаме, че ε е твърде малко число, на което във всичките следващи изчисления можем да пренебрегваме степените от втора нататък. Тогава уравнението за y ще бъде

$$f(y+\varepsilon)=f(y)+\varepsilon f'(y)+\frac{\varepsilon^2}{2}f''(y)+\dots$$

$$=f(y)+\varepsilon f'(y)=a_0y^n+(a_1+\varepsilon b_1)y^{n-1}+(a_2+\varepsilon b_2)y^{n-2}+\dots+(a_n+$$

$$+\varepsilon b_n)=0, \tag{32}$$

където за простота сме възели означенията $b_1=na_0$, $b_2=(n-1)a_1$, ..., $b_n=a_{n-1}$). Да означим с

$$c'_0y^n-c'_1y^{n-1}+c'_2y^{n-2}-\dots+(-1)^nc'_n=0 \tag{33}$$

уравнението, на което корените са квадратите на корените на уравнението (32). Тогава за коефициентите му по формулите (30) с пренебрегване на по-високите от първа степен на ε получаваме

$$c'_\nu = a'_\nu + 2\varepsilon b'_\nu,$$

като

(34) $$a'_0 = a_0^2,\ a'_1 = a_1^2 - 2a_0a_2,\ a'_2 = a_2^2 - 2a_1a_3 + 2a_0a_4, \ldots$$

и

$$b_0' = 0,\ b'_1 = a_1 b_1 - a_0 b_2,\ b'_2 = a_2 b_2 - (a_1 b_3 + a_3 b_1) + a_0 b_4,$$

(35) $$b'_3 = a_3 b_3 - (a_2 b_4 + a_4 b_2) + a_1 b_5 + a_5 b_1 - a_0 b_6, \ldots$$

Ако образуваме сега уравнението, на което корените са квадрати на корените на (33), т. е. четвъртите степени на корените на (32), то лесно виждаме, че това уравнение ще има формата

$$c''_0 y^n - c''_1 y^{n-1} + c''_2 y^{n-2} - \ldots + (-1)^n c''_n = 0,\ c''_\nu = a''_\nu + 4\varepsilon b''_\nu,$$

където числата a''_ν се получават от a_ν по формулите (34) и числата b''_ν се получават от b'_ν по формулите (35). След k такива трансформации ще получим уравнение от вида ($m = 2^k$)

(36) $$A_0 y^n - (A_1 + m\varepsilon B_1) y^{n-1} + (A_2 + m\varepsilon B_2) y^{n-2} - \ldots + (-1)^n (A_n + m\varepsilon B_n) = 0,$$

на което корените са $y_1^m, y_2^m, \ldots, y_n^m$, като с $y_1, y_2, \ldots, y_n$ сме означили корените на уравнението (33). При $\varepsilon = 0$ уравнението (36) става

$$A_0 y^n - A_1 y^{n-1} + A_2 y^{n-2} - \ldots + (-1)^n A_n = 0,$$

на което корените са $x_1^m, x_2^m, \ldots, x_n^m$, като с $x_1, x_2, \ldots, x_n$ сме означили корените на даденото уравнение (24).

Да разгледаме сега случая, когато даденото уравнение има само реални корени с различни абсолютни стойности или по-общо има реален корен x, на който абсолютната стойност не е равна на абсолютната стойност на друг реален корен или на модула на имагинерен някой корен. Тогава съгласно с предните изводи ще имаме

$$x^m = \frac{A_{\mu+1}}{A_\mu},$$

(37) $$y^m = \frac{A_{\mu+1} + m\varepsilon B_{\mu+1}}{A_\mu + m\varepsilon B_\mu}.$$

Ако за простота означим с $P_\mu = \frac{B_\mu}{A_\mu}$, то второто равенство от (37) става

(38) $$y^m = \frac{A_{\mu+1}}{A_\mu} \cdot \frac{1 + m\varepsilon P_{\mu+1}}{1 + m\varepsilon P_\mu} = \frac{A_{\mu+1}}{A_\mu}[1 + m\varepsilon(P_{\mu+1} - P_\mu)],$$

като сме пренебрегнали в развитието на дробта $\frac{1+m\varepsilon P_{\mu+1}}{1+m\varepsilon P_\mu}$ степените на ε от втората нататък. Но от равенството $y=x-\varepsilon$ с повдигане на степен m получаваме

$$y^m=(x-\varepsilon)^m=x^m-m\varepsilon x^{m-1}=x^m\left(1-\frac{m\varepsilon}{x}\right). \tag{39}$$

От (38) и (39) получаваме

$$x=-\frac{1}{P_{\mu+1}-P_\mu}. \tag{40}$$

Забележителното на тази формула е, че направо получаваме по нея корена, а не m-тата му степен, както преди. Нека сега $r(\cos\varphi+i\sin\varphi)=x'$ е имагинерен корен на (32), на който модулът е отличен от модулите на другите имагинерни корени (освен, разбира се, модула на конюгования му корен) и от абсолютните стойности на реалните корени. Ако r' е модулът на $x'-\varepsilon$, то по изведените преди формули ще имаме

$$r^{2m}=\frac{A_{\nu+2}}{A_\nu}.$$

$$r'^{2m}=\frac{A_{\nu+2}+m\varepsilon B_{\nu+2}}{A_\nu+m\varepsilon B_\nu}=\frac{A_{\nu+2}}{A_\nu}\cdot\frac{1+m\varepsilon P_{\nu+2}}{1+m\varepsilon P_\nu}.$$

Втората формула приближено може да се пише така:

$$r'^{2m}=\frac{A_{\nu+2}}{A_\nu}[1+m\varepsilon(P_{\nu+2}-P_\nu)]. \tag{41}$$

Но от приближената формула

$$r'^2=r^2-2\varepsilon r\cos\varphi$$

с повдигане в m-та степен получаваме приближената формула

$$r'^{2m}=r^{2m}-2m\varepsilon r^{2m-2}\cos\varphi=r^{2m}\left(1-m\varepsilon\frac{2r\cos\varphi}{r^2}\right). \tag{42}$$

От (41) и (42) следва формулата

$$-2r\cos\varphi=r^2(P_{\nu+2}-P_\nu).$$

Следователно имагинерният корен $r(\cos\varphi+i\sin\varphi)$ се пресмята с формулите

$$r=\sqrt[2m]{\frac{A_{\nu+2}}{A_\nu}},\qquad \cos\varphi=-\frac{r}{2}(P_{\nu+2}-P_\nu). \tag{43}$$

Преимуществото на този начин за пресмятане на имагинерните корени се състои в това, че всеки такъв корен се пресмята поотделно и няма

нужда да прибягваме до сложни изчисления в случай, особено когато уравнението има няколко имагинерни корена. От друга страна обаче пресмятанията се увеличават от необходимостта да намерим числата B_n.

8. Метод с итерация. Един друг метод за пресмятане на корените на алгебричните и трансцендентни уравнения е методът с прилагане на последователни итерации. На даденото уравнение

$$f(x)=0 \tag{44}$$

можем по различен начин да придадем формата

$$x=\varphi(x). \tag{45}$$

Нека x_1 е корен на уравнението (44) и a е приближена негова стойност. Да образуваме редицата от числа

$$a_1=\varphi(a),\ a_2=\varphi(a_1),\ a_3=\varphi(a_2),\ldots \tag{46}$$

Числата от тази редица ще се приближават все повече към корена x_1. Може да се установи именно следната теорема:

Нека $\varphi(x)$ е реален полином или реална функция, която има производна, за която в даден интервал I имаме

$$|\varphi'(x)|<q<1. \tag{47}$$

Предполагаме, че числата $a, a_1, a_2, a_3, \ldots$ принадлежат на интервала I. Тогава редицата $\{a_n\}_1^\infty$ клони към определена граница x' при $n\to\infty$. Границата x' е корен на уравнението $x=\varphi(x)$.

Действително имаме

$$a_2-a_1=\varphi(a_1)-\varphi(a)=(a_1-a)\varphi'(\zeta), \tag{48}$$

като ζ е число между a и a_1. На основание на (47) от (48) имаме

$$|a_2-a_1|<q|a_1-a|. \tag{49}$$

По подобен начин получаваме

$$\begin{aligned} |a_3-a_2|&<q|a_2-a_1|,\\ |a_4-a_3|&<q|a_3-a_2|,\\ &\cdots\cdots\\ |a_n-a_{n-1}|&<q|a_{n-1}-a_{n-2}|. \end{aligned} \tag{50}$$

Като умножим неравенствата (50) и (49), получаваме

$$|a_n-a_{n-1}|<q^{n-2}|a_1-a|. \tag{51}$$

От последното неравенство следва, че членовете на реда

$$|a_1|+|a_2-a_1|+|a_3-a_2|+\ldots \tag{52}$$

са по-малки от съответните членове на реда

$$|a_1|+|a_1-a|q+|a_1-a|q^2+\ldots$$

и следователно редът (52) е сходящ. Но от сходимостта на този ред следва сходимостта на реда

$$a_1+(a_2-a_1)+(a_3-a_2)+\ldots, \tag{53}$$

т. е. сходимостта на парциалната му n-та сума $a_1+(a_2-a_1)+\ldots+$ $+(a_n+a_{n-1})=a_n$. Ако η е сумата на реда (53), т. е. границата на a_n, от равенството

$$a_{n+1}=\varphi(a_n)$$

при $n\to\infty$ получаваме

$$\eta=\varphi(\eta),$$

което равенство показва, че η е корен x' на уравнението $x=\varphi(x)$. Понеже числата $a, a_1, a_2, \ldots$ принадлежат на интервала I (предположен затворен), то η принадлежи на I. В този интервал няма друг корен на (44). Действително, ако τ е друг корен на (44), то от равенствата

$$\eta=\varphi(\eta),\ \tau=\varphi(\tau)$$

с изваждане получаваме

$$\eta-\tau=\varphi(\eta)-\varphi(\tau)$$

и подобно на (49) получаваме неравенството

$$|\eta-\tau|<q|\eta-\tau|,$$

което е невъзможно по причина на ограничението $q<1$.

От (51) получаваме

$$|\eta-a_n|=|(a_{n+1}-a_n)+(a_{n+2}-a_{n+1})+\ldots|<|a_1-a|(q^n+q^{n+1}+\ldots)= \tag{54}$$
$$=|a_1-a|\frac{q^n}{1-q},$$

което неравенство показва бързината на приближенията a_n към корена η. Ще забележим, че от неравенството

$$|a_n-a_1|=|(a_2-a_1)+(a_3-a_2)+\ldots+(a_n-a_{n-1})|<$$
$$<|a_1-a|\frac{1-q^n}{1-q}<|a_1-a|\frac{1}{1-q}$$

се вижда, че за интервала I може да се вземе затвореният интервал

$$a_1-|a_1-a|\frac{1}{1-q}\leqq x\leqq a_1+|a_1-a|\frac{1}{1-q}.$$

Неравенството (54) показва, че приближенията са по-добри при по-малки числа q, т. е. при по-бавно изменящи се функции $\varphi(x)$ около корена η. От друга страна, с граничен преход получаваме

$$\lim_{n\to\infty}\frac{a_{n+1}-\eta}{a_n-\eta}=\lim_{n\to\infty}\frac{\varphi(a_n)-\varphi(\eta)}{a_n-\eta}=\varphi'(\eta)$$

и следователно при $a_n\to\eta$ порядъкът $a_{n+1}-\eta$ е от същия ред на $a_n-\eta$. Ако $\varphi'(\eta)=0$ и $\varphi''(\eta)\neq 0$, то лесно се вижда, че $a_{n+1}-\eta$ е от втори

ред спрямо $a_n-\eta$. За увеличение тогава на бързината на приближенията представяме уравнението (44) във формата

$$x=\frac{\varphi(x)-\lambda x}{1-\lambda}=\psi(x),$$

където λ е число, близко до $\varphi'(\eta)$. Понеже коренът η е неизвестен, то полагаме $\lambda=\varphi'(a)$, където a е някоя приближена стойност на η. Прилагаме тогава въпросните итерации на уравнението

$$x=\psi(x).$$

Като пример да приложим метода за намиране на корена x_1 на уравнението

(55) $$f(x)=x^3+x-3=0,$$

който лежи в интервала (1, 2). Да напишем уравнението (55) във формата

$$x=\sqrt[3]{3-x}.$$

Тук $\varphi(x)=\sqrt[3]{3-x}$ и $\varphi'(x)$ за $1\leqq x\leqq 2$ лежи между $-\frac{1}{3}$ и $-\frac{1}{3\sqrt[3]{4}}$.

Като положим $a=1,2$, получаваме последователно

$$a_1=\sqrt[3]{3-1,2}=\sqrt[3]{1,8}=1,21,\ldots,$$
$$a_2=\sqrt[3]{3-1,21}=\sqrt[3]{1,79}=1,214,\ldots,$$
$$a_3=\sqrt[3]{3-1,214}=\sqrt[3]{1,786}=1,2132\ldots,$$
$$a_4=\sqrt[3]{3-1,2132}=\sqrt[3]{1,7868}=1,2134\ldots,$$
$$a_5=\sqrt[3]{3-1,2134}=\sqrt[3]{1,7866}=1,2134\ldots$$

Приближението a_5 съвпада с a_4 в първите си 4 десетични знака и коренът с точност до четвъртия знак е равен на 1,2134.

Можем даденото уравнение да напишем и във формата

(56) $$x=3-x^3=\varphi(x).$$

Но производната на дясната част $-3x^2$ надминава по абсолютна стойност единицата и ще трябва да дадем другата форма на това уравнение. Именно полагаме $\lambda=\varphi'(1,2)=-4,32$ и следователно ще трябва уравнението (56) да заместим с уравнението

(57) $$x=\frac{3-x^3+4,32\,x}{5,32}=\omega(x).$$

Като изхождаме от приближението $a=1{,}2$, получаваме

$$a_1=\omega(a)=\omega(1{,}2)=1{,}2135,\ \ldots$$

от която приближена стойност може да получим нова, като в $\omega(x)$ заместим x с тази първа приближена стойност. За по-голяма точност отново изменяме уравнението (55), като приемаме за приближена начална стойност за корена числото $a_1=1{,}213$ и полагаме $\lambda=\varphi'(1{,}213)=-4{,}4141$. Тогава уравнението (57) се замества с уравнението

$$x=\frac{3-x^3+4{,}4141x}{5{,}4141}=\omega_1(x),$$

от което за a_2 получаваме новата приближена стойност

$$a_2=\omega_1(1{,}213)=1{,}21341$$

на корена x_1 и т. н.

Ще забележим, че методът на Нютон може да се разглежда като метод от разгледания вид. Именно уравнението

$$f(x)=0 \tag{58}$$

написваме така:

$$x=x-\frac{f(x)}{f'(x)}=\varphi(x). \tag{59}$$

Очевидно всеки прост корен x_1 на (58) е и корен на (59). Освен това от равенството

$$\varphi'(x)=1-\frac{f'(x)}{f'(x)}+\frac{f(x)f''(x)}{f'^2(x)}=\frac{f(x)f''(x)}{f'^2(x)}$$

се вижда, че $\varphi'(x_1)=0$, което е причина на бързата сходимост на последователните приближения с този метод.

Да напишем сега уравнението $f(x)=0$ във формата

$$x=x-cf(x)=\varphi(x), \tag{60}$$

където c е константа. Ако a е едно приближение до корен x_1 на (60), то съгласно с изискването за по-добри приближения да изберем тази константа така, че производната $\varphi'(x)$ за $x=a$ да се анулира, т. е. $1-cf'(a)=0$. За c получаваме $c=1/f'(a)$ и следователно

$$x=x-\frac{f(x)}{f'(a)}.$$

По този метод ще трябва да образуваме последователните приближения

$$a_1=a-\frac{f(a)}{f'(a)},\quad a_2=a_1-\frac{f(a_1)}{f'(a)},\quad a_3=a_2-\frac{f(a_2)}{f'(a)},\quad a_4=a_3-\frac{f(a_3)}{f'(a)},\ \ldots$$

Сходимостта на горната редица е по-бавна от тази на редицата с метода на Нютон, но преимуществото на прилагането на редицата е, че производната е пресметната само за една стойност $x=a$ на променливото x. Лесно се вижда, че разликата $a_{n+1}-x_1$ е от същия ред като разликата a_n-x_1. В заключение трябва да се каже, че методът на Лобачевски — Грефе има преимущество пред другите методи при решение на алгебрични уравнения и методът на Нютон е за предпочитане при решение на трансцендентни уравнения и при пресмятане на реалните корени на алгебричните уравнения. От значение е итерационният метод, когато не се налага пресмятане на производната на дясната част на уравнението.

9. **Решение на система уравнения.** Нека е дадена системата от две уравнения с две неизвестни x и y

$$\begin{aligned} f(x,y)&=0,\\ \varphi(x,y)&=0. \end{aligned} \tag{61}$$

Предполагаме, че уравненията са с реални коефициенти. Ако уравненията са алгебрични, то видяхме, че с елиминиране на едното неизвестно, например y, получаваме за другото неизвестно x едно алгебрично уравнение

$$F(x)=0, \tag{62}$$

на което можем да пресметнем с желана точност реалните корени. След това по познат вече начин можем да намерим съответните приближени стойности на y и така да получим приближени стойности на системите от реални решения на уравненията (61). Ако уравненията (61) са трансцендентни, то не винаги е възможно да елиминираме едно от неизвестните и следователно изложеният начин е неприложим въобще. Също така в алгебричния случай степента на уравнението (62) е доста висока и изчислението на корените му се доста усложнява. Методът на Нютон е почти непосредствено обобщим за приближено решение на системата уравнения (61). Нека a и b са две приближени стойности на една система от решения x_1 и y_1. Да означим с h и l разликите x_1-a и y_1-b. По формулата на Тейлор ще имаме

$$\begin{aligned} f(x_1,y_1)&=f(a+h,\ b+l)=f(a,\ b)+hf_x(a,\ b)+lf_y(a,\ b)+\\ &+\frac{1}{2!}[h^2f_{xx}(a,\ b)+2hlf_{xy}(a,\ b)+l^2f_{yy}(a,\ b)]+\ \ldots=0,\\ \varphi(x_1,\ y_1)&=\varphi(a+h,\ b+l)=\varphi(a,\ b)+h\varphi_x(a,b)+l\varphi_y(a,\ b)+\\ &+\frac{1}{2!}[h^2\varphi_{xx}(a,\ b)+2hl\varphi_{xy}(a,\ b)+l^2\varphi_{yy}(a,\ b)]+\ \ldots=0. \end{aligned}$$

Като пренебрегнем степените на h и l от втората нататък, получаваме приближените уравнения за h и l

$$\begin{aligned} f(a,\ b)+hf_x(a,\ b)+lf_y(a,\ b)&=0,\\ \varphi(a,\ b)+h\varphi_x(a,\ b)+l\varphi_y(a,\ b)&=0, \end{aligned} \tag{63}$$

откъдето намираме приближените стойности h_1 и l_1 на h и l,

$$(64)\qquad h_1=-\frac{\begin{vmatrix} f(a,\ b) & f_y(a,\ b) \\ \varphi(a,\ b), & \varphi_y(a,\ b) \end{vmatrix}}{\begin{vmatrix} f_x(a,\ b) & f_y(a,\ b) \\ \varphi_x(a,\ b) & \varphi_y(a,\ b) \end{vmatrix}},\quad l_1=-\frac{\begin{vmatrix} f_x(a,\ b) & f(a,\ b) \\ \varphi_x(a,\ b) & \varphi(a,\ b) \end{vmatrix}}{\begin{vmatrix} f_x(a,\ b) & f_y(a,\ b) \\ \varphi_x(a,\ b) & \varphi_y(a,\ b) \end{vmatrix}}.$$

Следователно за x_1 и y_1 получаваме приближените стойности

$$a_1=a+h_1,\quad b_1=b+l_1.$$

Ако положим сега $x_1=a_1+h$ и $y_1=b_1+l$ и постъпим по същия начин получаваме приближенията h_2 и l_2 на h и l, които очевидно се получават с формулите (64), където a и b заместваме с a_1 и b_1 съответно. Следователно по този начин получаваме новите приближени стойности

$$a_2=a_1+h_2,\quad b_2=b_1+l_2$$

на решението x_1, y_1. Ако желаем да получим по-големи приближения, продължаваме същия начин на работа. Естествено, ако при получаване на приближенията новите приближения съвпадат с предните до разлика, по-малка по абсолютна стойност от желаната точност, то спираме изчисленията до това приближение.

Първите приближения a и b на решението x_1 и y_1 получаваме обикновено по графичен начин. Именно начертаваме кривите C_1 и C_2, които съответствуват на уравненията (61), и намираме пресечните, точки на кривите, на които точки абсцисите и ординатите позволяват да се получат приближени (поне груби) стойности на решенията на системата уравнения (61).

Решението на системи уравнения с повече от две неизвестни става по напълно аналогичен начин. Така за три уравнения с три неизвестни x, y, z, на които знаем приближени стойности a, b и c за разликите $h=x_1-a$, $l=y_1-b$, $n=z_1-c$ между решенията x_1, y_1 и z_1 и a, b, c, ще получим по метода на Нютон три линейни уравнения, аналогични на уравненията (63).

Намирането на приближени стойности на корените на произволно уравнение с комплексни коефициенти се свежда към същата задача за реалните решения на система от две уравнения с реални коефициенти с две неизвестни. Действително, ако даденото уравнение е

$$(65)\qquad f(z)=a_0z^n+a_1z^{n-1}+a_2z^{n-2}+\ \ldots\ +a_n=0,$$

то със заместване на z с $x+iy$ получаваме

$$P(x,\ y)+iQ(x,\ y)=0,$$

където $P(x,\ y)$ и $Q(x,\ y)$ са полиноми с реални коефициенти. Уравнението (65) е еквивалентно на системата

$$P(x,\ y)=0,$$

$$Q(x,\ y)=0.$$

Можем да пресметнем корените на (65) и направо с метода на Нютон. Именно, ако a е една приближена стойност на един корен x на (65), то по формулата на Нютон

$$a_1 = a - \frac{f(a)}{f'(a)}$$

намираме ново приближение на x и т. н. Очевидно методът се прилага и за трансцендентни уравнения.

Пресмятането на корените на уравнението (65) с произволни комплексни коефициенти може да се сведе към същата задача за уравнения с реални коефициенти. Нека $\alpha_k + i\beta_k$, $k = 1, 2, \ldots, n$, са корените на уравнението

$$f(z) = a_0 z^n + a_1 z^{n-1} + \ldots + \bar{a}_n = 0. \tag{66}$$

Тогава числата $\alpha_k - i\beta_k$, $k = 1, 2, \ldots, n$, ще са корени на уравнението

$$\overline{f}(z) = \overline{a_0} z^n + \overline{a_1} z^{n-1} + \ldots + \overline{a_n} = 0,$$

което получаваме от предното със заместване на всеки коефициент с неговата конюгована стойност. Очевидно уравнението

$$f(z)\overline{f}(z) = 0 \tag{67}$$

ще има само реални коефициенти и корените му са числата $\alpha_k \pm i\beta_k$, $k = 1, 2, \ldots, n$. Степента на уравнението (67) е два пъти по-голяма от тази на (66), но за пресмятане на корените му можем да приложим удобния метод на Лобачевски — Грефе.

Можем и да не прибягваме до уравнението (67), а да приложим метода на Лобачевски — Грефе направо на даденото уравнение (66). Преимуществото в такъв случай е, че степента на уравнението е по-ниска значително, но изчисленията на коефициентите на трансформираните уравнения се усложняват, понеже въобще са имагинерни числа.

10. Графичен метод на Лил. За намиране на приближени стойности на корените на уравнението $f(x) = 0$ си служат с графични методи. Един примитивен такъв начин е да се построи графиката на полинома $f(x)$ и да се намерят абсцисите на пресечните ѝ точки с абсцисната ос. Този начин е очевидно приложим и за трансцендентни уравнения, но изисква доста губене на време и построяване на точки, които нямат значение за решение на уравнението. Един значително удобен начин за такова решение на алгебричните уравнения е този на Лил.

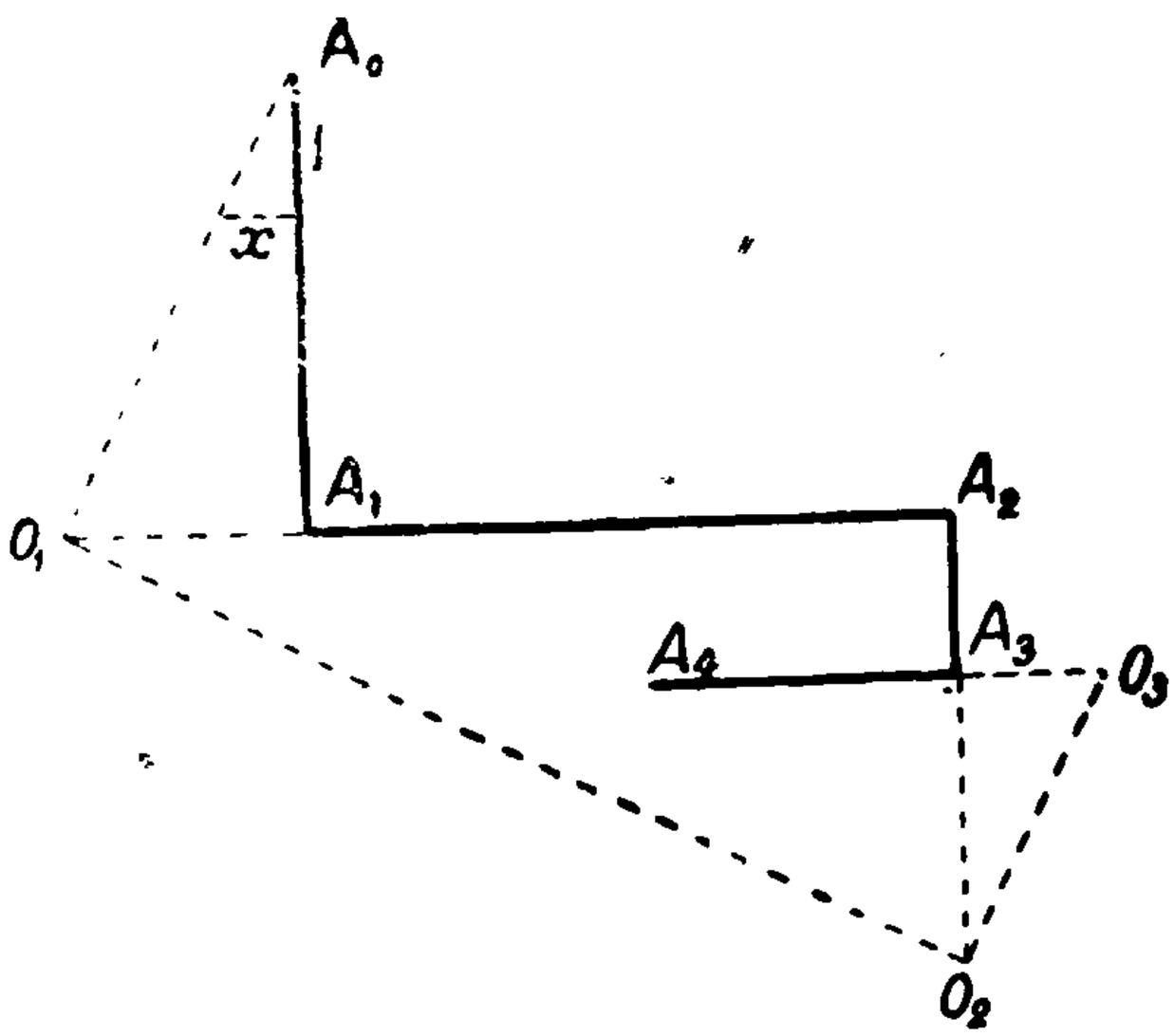
Нека в равнината е даден един мащаб за мерене на отсечките и едно положително направление за въртене, което, както е прието, избираме обратно на движението на часовниковата стрелка. Под AB разбираме насочената отсечка между точките A и B, която си мислим, че е с посока от A към B. Под $|AB|$ разбираме дължината на отсечката. При по-нататъшните ни разглеждания ъглите, които ще използуваме, лежат между $-\frac{\pi}{2}$ и $\frac{\pi}{2}$. Под AOB ще разбираме ъгъла, под който трябва

да се завърти отсечката OA, за да се слее с OB. Разбира се, ъгълът е положителен или отрицателен според това, дали въпросното въртене е в положителна или отрицателна посока.

Нека сега

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ \ldots\ +a_n,\ a_0>0,$$

е произволен полином с реални коефициенти и x е произволно реално число. Построяваме тогава една начупена праволинейна линия $A_0\ A_1\ A_2\ \ldots A_n\ A_{n+1}$, на която съседните отсечки образуват прави ъгли, като следваме следното правило: а) За $i=0,1,2,\ldots,n$ дължината $|A_iA_{i+1}|$ е равна на $|a_i|$. б) Положението на $A_0\ A_1$ вземаме произволно. За $i=1,2,\ldots,n$ отсечката A_iA_{i+1} е перпендикулярна на $A_{i-1}\ A_i$, но направлението ѝ е така избрано, че ъгълът $A_{i+1}\ A_i\ A_{i-1}$ да е равен на $\frac{\pi}{2}$ или $-\frac{\pi}{2}$ според това, дали съответстващите коефициенти a_{i-1} и a_i са с еднакъв или противен знак.



Черт. 8

(Като пример на чертеж 8 е графично представен полиномът $3x^2+4x^2-x+2$.)

. В случай, че $a_{k-1}\neq 0$, но $a_k=a_{k+1}=\ldots=0$, точките A_k, A_{k+1}, A_{k+2}, ... съвпадат. Тогава под отсечката A_kA_{k+1} разбираме естествено перпендикуляра към $A_{k-1}A_k$ в точката A_k, под A_{k+1}, A_{k+2} — перпендикуляра към A_kA_{k+1} в A_{k+1} (която се слива с A_k) и т. н. Направлението на правите A_kA_{k+1}, $A_{k+1}\ A_{k+2}$, ... определяме еднозначно с условието, че A_kA_{k+1}, $A_{k+1}A_{k+2}$, ... преминават съответно в A_kA_{k-1}, $A_{k+1}A_k$, ... посредством въртене на ъгъл $+\frac{\pi}{2}$.

Под знак на отсечката A_kA_{k+1} разбираме знака на a_k, ако $a_k\neq 0$, при $a_{k-1}\neq 0$, $a_k=a_{k+1}=\ldots=0$ знаците на A_kA_{k+1}, $A_{k+1}A_{k+2}$, ... избираме еднакви с този на a_{k-1}.

Трябва сега с полученото графично представяне на полинома $f(x)$ да представим графично стойността му за произволно дадено x. За тази цел определяме точка O_1, лежаща на A_1A_2, така че $\operatorname{tg}(O_1A_0A_1)=x$. Това постигаме лесно, като от A_0 по A_0A_1 нанасяме мярката 1 и от крайната ѝ точка издигаме перпендикуляр към нея, като по него отмерваме x наляво или надясно от A_1A_0 според това, дали $x>0$ или

$x<0$. Ако B е крайната точка на този перпендикуляр, то O_1 е пресечната точка на A_0B с A_1A_2. В точка O_1 прекарваме перпендикулярна права на A_0O_1 и определяме пресечната ѝ точка O_2 с A_2A_3. В тази точка O_2 прекарваме перпендикулярна права на O_1O_2 и определяме пресечната ѝ точка O_3 с A_3A_4 и т. н. Така на всяка права A_kA_{k+1} получаваме по една напълно определена точка O_k. Нека на отсечката O_kA_{k+1} $(k=1, 2, \ldots, n)$, лежаща на правата A_kA_{k+1}, припишем еднакъв или противен знак на този на A_kA_{k+1} според това, дали отсечките A_kA_{k+1} и O_kA_{k+1} са еднакви или противоположни по посока. Да означим тогава с d_k релативната дължина на O_kA_{k+1}. Тогава ще имаме

$$a_0x^k+a_1x^{k-1}+\ \ldots\ +a_k=d_k,\quad k=0,\ 1,\ 2,\ldots,n\quad O_0=A_0.$$

Специално ще имаме $f(x)=d_n$.

При $k=0$ имаме $a_0=d_0$, което е очевидно. При по-нататъшното доказателство използуваме факта, че ъглите $O_kO_{k-1}A_k$ са равни по абсолютна стойност и по знак, т. е.

$$\operatorname{tg}(O_kO_{k-1}A_k)=x,$$

и лежащите на една права отсечка $A_kA_{k+1}A_k$ и A_kO_k имат еднаква или противоположна посока според това, дали ъглите $O_kO_{k-1}A_k$ и $O_{k-1}A_kA_{k+1}$ имат еднакви или противни знаци. Тогава лесно от чертежа получаваме последователно

$$d_1=d_0\operatorname{tg}(O_1A_0A_1)+a_1=a_0x+a_1,$$

$$d_2=d_1\operatorname{tg}(O_2O_1A_2)+a_2=a_0x^2+a_1x+a_2$$

и т. н.

Следователно, като измерим отсечката O_nA_{n+1}, ще получим по графичния метод на Лил стойността d_n на $f(x)$ (на черт. 8 е взето $x=\frac{1}{2}$).

Очевидно е, че x ще бъде корен на уравнението $f(x)=0$, ако при предната конструкция точките O_n и A_{n+1} съвпадат. С няколко опитвания можем така да получим приближени стойности за реалните корени на алгебричните уравнения.

11. Метод на Лагер за уравнения, които имат само реални корени. За уравнения със само реални корени, които можем да предположим за прости, много добро приближение дава следният метод на Лагер, който ще изложим. Нека $f(x)=0$ е уравнение от n-та степен със само реални и прости корени и нека α_k и α_{k+1} са два негови последователни корена. Тогава за всяко произволно число x, лежащо между α_k и α_{k+1}, уравнението

$$[(n-2)f'^2-(n-1)ff'']\ (X-x)^2-2ff'(X-x)-nf^2=0 \tag{68}$$

има един корен между α_k и x и един корен между x и α_{k+1}. От решението на уравнението (68) получаваме за корените му стойностите

$$X_{1,2}=x+\frac{nf}{-f'\pm\sqrt{(n-1)\ [(n-1)f'^2-nff'']}}. \tag{69}$$

Понеже числата X_1-x и X_2-x са с противни знаци, то за X_2 при $X_2>x$ трябва да се вземе радикалът със знак, еднакъв с този на f.

Доказателството[1] се основава на едно помощно неравенство. Нека

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

са произволни реални числа, за които имаме

$$a_1+a_2+\ \ldots\ +a_n=a,$$
$$a_1^2+a_2^2+\ \ldots\ +a_n^2=b.$$

По неравенството на Коши — Буняковски имаме

$$(70)\quad (a-a_1)^2=(a_2+a_3+\ \ldots\ +a_n)^2\leqq(n-1)\ (a_2^2+\ a_3^2+\ \ldots\ +a_n^2)=$$
$$=(n-1)\,(b-a_1^2),$$

отгдето следва, че

$$na_1^2-2a\ a_1+a^2-\ (n-1)\ b\leqq 0.$$

От това неравенство заключаваме, че всяко от числата a_i, $i=1,\ 2,\ \ldots,n$, се съдържа между корените на уравнението

$$nx^2-2ax+a^2-(n-1)\,b=0,$$

т. е. имаме

$$(71)\qquad \frac{a-\sqrt{(n-1)\,(nb-a^2)}}{n}\leqq a_i\leqq\frac{a+\sqrt{(n-1)\,(nb-a^2)}}{n}.$$

Понеже равенство в (70) може да имаме само при

$$a_2=a_3=\ \ldots\ =a_n,$$

то границите в неравенството (71) са достижими само при

$$a_i=\frac{a\mp\sqrt{(n-1)(nb-a^2)}}{n}$$

и

$$a_1=a_2=\ \ldots\ =a_{i-1}=a_{i+1}=\ \ldots\ =a_n=\frac{a\,(n-1)\pm\sqrt{(n-1)\ (nb-a^2)}}{n\ (n-1)}.$$

Правилото на Лагер се получава непосредствено от (71). Именно нека корените на уравнението $f(x)=0$ от n-та степен са всички реални и прости. Ако ги означим с $x_1,\ x_2,\ \ldots\ ,\ x_n$, то имаме

$$\frac{f'(x)}{f(x)}=\sum_{p=1}^{n}\frac{1}{x-x_p},\quad \left(\frac{f'(x)}{f(x)}\right)^2-\frac{f''(x)}{f(x)}=\sum_{p=1}^{n}\frac{1}{(x-x_p)^2}.$$

---

[1] Вж. Н. Обрешков, Годишник на Соф. университет, том 47, 1950/1951, стр. 81.

Остава тогава да поставим в (71)

$$a_i=\frac{1}{x-x_i},\quad a=\frac{f'(x)}{f(x)},\quad b=\left(\frac{f'(x)}{f(x)}\right)^2-\frac{f''(x)}{f(x)}.$$

Нека отбележим, че правилото се приспособява и за случай на уравнения с многократни корени.

От предното извеждане следва също, че в случая, когато x е поголямо от най-големия корен на $f(x)=0$, то X_1 ще бъде приближението на този корен.

За пример да пресметнем с приближение най-големия корен на уравнението

$$T_n(x)=\cos(n \arccos x)=0, \tag{72}$$

гдето $T_n(x)$ е полиномът на Чебишев от n-та степен. Да въведем помощното променливо α, дадено с

$$x=\cos\alpha.$$

Тогава имаме

$$T_n(x)=\cos n\alpha.$$

От този израз се вижда, че корените на уравнението (72) са

$$\cos\frac{\pi}{2n}(2\nu+1),\ \nu=0,\ 1,\ 2,\ \ldots,\ n-1.$$

Следователно всичките са реални и прости, като най-големият е $\cos\frac{\pi}{2n}$.

За $x=1$ намираме лесно

$$T_n(1)=1,\ T_n'(1)=n^2,\ T_n''(1)=\frac{n^2(n^2-1)}{2}.$$

По формула (21) ще имаме следната приблизителна стойност на този корен:

$$\cos\frac{\pi}{2n}=1-\frac{1}{n+(n-1)\sqrt{\frac{2n^2-n}{3}}}.$$

Ако поставим $\frac{1}{n}=x$, тази формула става

$$\cos\frac{\pi x}{2}=1-\frac{x^2}{x+(x-1)\sqrt{\frac{2-x}{3}}},$$

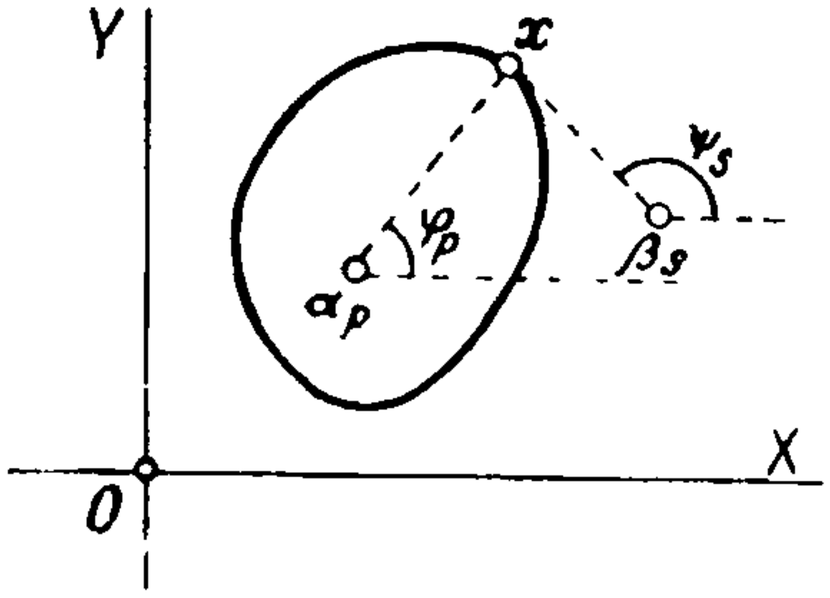
която дава едно много добро приближение за косинуса при всички x, $0<x<1$.

Глава IV

БРОЙ НА КОРЕНИТЕ В ЕДНА ОБЛАСТ

1. Теорема на Коши. С теоремата на Щурм можем да определим броя на реалните корени на едно уравнение, които се намират в един интервал. Пръв Коши е дал метод, с който може да се определи броят на корените на едно уравнение с произволни (реални или имагинерни) коефициенти, които се намират в една дадена област.

Нека C е една затворена крива (черт 9) без двойни точки. Даденото уравнение с прозволни реални или имагинерни коефициенти нека бъде



Чер. 9

$$(1)\quad f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0.$$

Да допуснем, че върху кривата C това уравнение няма корени. Нека да бъдат $\alpha_1, \alpha_2, \ldots, \alpha_k$, а тези, които лежат вън от C, да бъдат

$$\beta_1, \beta_2, \ldots, \beta_m.$$

Да поставим

$$x-\alpha_p=r_p(\cos\varphi_p+i\sin\varphi_p),\quad p=1,\ 2,\ldots,\ k,$$
$$x-\beta_s=\delta_s(\cos\psi_s+i\sin\psi_s),\quad s=1,\ 2,\ldots,\ m,$$
$$f(x)=R(\cos\Phi+i\sin\Phi)=P+iQ,$$

гдето R, Φ са съответно модулът и аргументът на $f(x)$, а P и Q са реалната и имагинерната част. От тъждеството

$$f(x)=a_0(x-a_1)\ldots(x-a_k)(x-\beta_1)\ldots(x-\beta_m)=$$
$$=R(\cos\Phi+i\sin\Phi),$$

ако

$$a_0=r(\cos\alpha+i\sin\alpha),$$

имаме

$$(2)\qquad \Phi=\varphi_1+\varphi_2+\ldots+\varphi_k+\psi_1+\psi_2+\ldots+\psi_m+\alpha.$$

От друга страна, понеже

$$R\cos\Phi=P,\ R\sin\Phi=Q,$$

то имаме

$$(3)\qquad \operatorname{tg}\Phi=\frac{Q}{P}.$$

Нека точката x изписва контура C един път по права посока, т. е. обратно на движението на часовниковата стрелка. Очевидно от чертежа ъглите φ_i, $i=1, 2,\ldots,k$ се увеличават с 2π. Ъглите ψ_i могат да

растат или намаляват до известни граници, но не могат да се изменят на 2π. Именно при извършване на пълната обиколка те приемат своята първоначална стойност. Следователно при едно завъртане на x по C аргументът Φ на $f(x)$ ще се изменя в една или друга посока през време на варирането на x, но при извършване на пълното завъртане ще порасне с $2k\pi$.

Но при едно изменение на ъгъла Φ на 2π, $\operatorname{tg}\Phi$ преминава два пъти през нулата от отрицателен към положителен.

Когато Φ варира, може от една стойност Φ_0 да се увеличи до Φ_1 и да се върне към Φ_0. При това при растенето на Φ на всяко минаване през нулата на $\operatorname{tg}\Phi$ от $-$ в $+$ отговаря при намаляване на Φ минаване от $+$ в $-$. И въобще, когато Φ се връща към първоначалната стойност, то $\operatorname{tg}\Phi$ се анулира еднакво число пъти, минавайки от $-$ в $+$ и обратно. Понеже при извършване на цялата обиколка Φ се увеличава с $2k\pi$, то ако означим с p броя на анулирането на $\operatorname{tg}\Phi$, като минава от $-$ в $+$, с q броя на анулирането, минавайки от $+$ в $-$, ще имаме

$$\frac{p-q}{2}=k.$$

Така получаваме теоремата на Коши. Броят на корените на уравнението

$$f(x)=P+iQ=0,$$

които се намират вътре в един затворен прост контур C, е равен на половината от разликата на броя на анулирането на $\frac{Q}{P}$, като минава от отрицателно в положително, и на броя на анулирането на същото отношение, като минава от положително в отрицателно, когато x изписва един път кривата C.

Под прост контур разбираме крива без многократни точки.

Ако C е съставена от дъги от уникурсални криви, то по тях координатите на x са рационални функции на един параметър t така, че $\frac{Q}{P}$ е отношение на два полинома на t с реални коефициенти, следователно за пресмятане на k може да се приложи обобщената теорема на Щурм.

2. Приложения. С помощта на варирането на аргумента, което лежи в основата на теоремата на Коши, ще установим някои други интересни теореми. Именно нека $f(x)$ и $\varphi(x)$ са два полинома. Тогава имаме следната теорема на Руше. Ако върху C имаме постоянно

$$|f(x)|>|\varphi(x)|,$$

то двете уравнения

$$f(x)=0,\ f(x)+\varphi(x)=0$$

имат еднакъв брой корени в C.

Действително да поставим

$$F(x)=f(x)+\varphi(x)=f(x)\left[1+\frac{\varphi(x)}{f(x)}\right].$$

Нека първото уравнение $f(x)=0$ има p корена в C. Тогава, когато x изпише един път C, видяхме, че аргументът на $f(x)$ ще порасне с $2p\pi$. Точката

$$u=\frac{\varphi(x)}{f(x)},$$

когато x е по C, е в окръжност с радиус, по-малък от 1, понеже $|\varphi(x)|<|f(x)|$. Следователно точката $(1+u)$ ще бъде в една окръжност с радиус <1 с център единица, откъдето следва, че аргументът на $(1+u)$ ще се върне на първоначалната си стойност при изписването на C. Понеже аргументът $F(x)$ е сума от аргументите на $f(x)$ и $(1+u)$, то той ще порасне с $2p\pi$, т. е. уравнението $F(x)=0$ има p корена в C, което трябва да се докаже. Ще направим едно приложение на тази теорема.

Нека a е един корен на уравнението

$$f(x)=0.$$

Да опишем около a окръжност с произволно даден радиус ρ така, че във и върху него да няма друг корен освен a на даденото уравнение. Нека $g(x)$ е полином, който се отличава от $f(x)$ с малко изменение на коефициентите така, че по периферията на построената окръжност C да имаме

$$|f-g|<|f|,$$

което по причина на $f(x)\neq 0$ по C е възможно. Полиномът $g(x)$ ще има толкова нули в C, колкото кратен е коренът a. Понеже ρ може очевидно да се вземе, колкото си щем малко, то така получаваме: корените на едно уравнение са непрекъснати функции на коефициентите му.

Ако една редица от полиноми

$$f_m(x)=a_{0m}x^n+a_{1m}x^{n-1}+\ldots+a_{nm}$$

клони към полинома

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n,$$

т. е.

$$\lim_{m\to\infty}a_{im}=a_i,\ i=0,\ 1,\ 2,\ldots,\ n,$$

то според горното нулите на полиномите $f_m(x)$ ще клонят към нулите на $f(x)$.

Ако нулите на полиномите $f_m(x)$ са реални, то и нулите на $f(x)$ ще бъдат реални.

С варирането на аргумента лесно се доказва следната теорема на Билер—Хермит:

Ако уравнението

$$f(x)=U+iV=0$$

има корени само от едната страна на реалната ос, то корените на уравненията

$$U=0,\ V=0$$

са реални и се взаимно разделят.

Действително нека корените на $f(x)=0$

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$$

лежат над реалната ос. При реално x да поставим

$$x-\alpha_k=\rho_k(\cos\varphi_k+i\sin\varphi_k),\ k=1,\ 2,\ 3,\ \ldots,\ n,$$

$$f(x)=R(\cos\Phi+i\sin\Phi)=U+iV,\ \operatorname{tg}\Phi=\frac{V}{U}.$$

Понеже

$$f(x)=a_0(x-\alpha_1)\ldots(x-\alpha_n),\ \text{то ако}$$

$a_0=r(\cos\alpha+i\sin\alpha)$, ще имаме $\Phi=\alpha+\varphi_1+\varphi_2+\ldots+\varphi_n$.

Ако $\alpha\neq 0,\ \pi$, то полиномите U и V са от n-та степен и когато x варира от $-\infty$ до $+\infty$, понеже всеки φ_i се увеличава от $-\pi$ на 0, аргументът Φ на $f(x)$ пораства от $\alpha-n\pi$ до α.

Но тогава отношението $\frac{V}{U}=\operatorname{tg}\Phi$ ще премине n пъти през нулата от отрицателно в положително. Следователно уравнението $V=0$ ще има n реални корена и между два последователни корена U трябва да си променя знака, т. е. $U=0$ има най-малко $(n-1)$ реални корена, т. е. понеже степента му е n, той ще има само реални корени, които се отделят от корените на U.

Ако $\alpha=0$ или π, тогава степента на V е $n-1$. Тогава същите разсъждения могат да се приложат за анулирането на $\frac{U}{V}=\operatorname{ctg}\Phi$.

Ще установим обратната теорема:

Ако полиномите U и V имат само реални взаимно разделящи се нули, то нулите на полинома $f(x)=U+iV$ лежат от едната страна на реалната ос.

Степените на U и V са или равни, или се различават с единица. Нека U е този полином, на който степента е най-малката. Ако x е нула на $f(x)$, то

$$\frac{U}{V}+i=0.$$

Ако $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$ са нулите на V, то предното уравнение може да се пише така:

(1)
$$C+i+\sum_{k=1}^{n}\frac{U(\alpha_k)}{V'(\alpha_k)}\cdot\frac{1}{x-\alpha_k}=0,$$

гдето C е реално число. Лесно се вижда, че числата

$$\lambda_k = \frac{U(\alpha_k)}{V'(\alpha_k)}, \quad k = 1, 2, \ldots, n,$$

са с еднакви знаци. Действително, ако α_k и α_{k+1} са две последователни нули на $V(x)$, то по условие ще имаме

$$U(\alpha_k)\, U(\alpha_{k+1}) < 0. \tag{2}$$

По теоремата на Рол между тези две нули има една нула на $V'(x)$, т. е.

$$V'(\alpha_k)\ V'(\alpha_{k+1}) < 0. \tag{3}$$

С неравенствата (2) и (3) се доказва твърдението за λ_k.

Очевидно е, че уравнението (1) няма реални корени. Нека $x = \xi + i\eta$ е един негов имагинерен корен. Като приравним на нула имагинерната част на лявата страна на (1), получаваме

$$1 - \eta \sum_{k=1}^{n} \frac{\lambda_k}{(\xi - \alpha_k)^2 + \eta^2} = 0,$$

отгдето е ясно, че $\eta \neq 0$ и че η има знак, еднакъв с този на числата λ_k. Като приложение на теоремата на Руше ще разгледаме два примера.

1. Ако за коефициентите на уравнението

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$$

имаме

$$|a_p| > |a_0| + \ldots + |a_{p-1}| + |a_{p+1}| + \ldots + |a_n|,$$

то в кръга $|x| < 1$ даденото уравнение има точно p корена.

Действително по окръжността $|x| = 1$ имаме

$$|f(x) - a_p x^p| \leqq |a_0| + \ldots + |a_{p-1}| + |a_{p+1}| + \ldots + |a_n| < |a_p| = |a_p x^p|$$

и по теоремата на Руше заключаваме, че броят на корените на $f(x) = 0$ в кръга $|x| < 1$ е равен на броя на корените на уравнението $a_p x^p = 0$ в същия кръг, т. е. е равен на p.

2. Теорема на Пеле. Нека за полинома

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n$$

уравнението

$$F(x) = |a_0| + |a_1| x + \ldots + |a_{p-1}| x^{p-1} - |a_p| x^p + |a_{p+1}| x^{p+1} + \ldots + |a_n| x^n = 0$$

да има два положителни корена α и β, $\alpha < \beta$. Тогава полиномът $f(x)$ има точно p нули в кръга $|x| \leqq \alpha$ и нито една нула във венеца $\alpha < |x| < \beta$.

По теоремата на Декарт уравнението $F(x) = 0$ има два положителни корена или нито един. Правим следователно предположението, че това уравнение има два положителни корена. Нека r е произволно число, удовлетворяващо на условието $\alpha < r < \beta$. Понеже $F(x) > 0$ за

$x=0$ и $x=\infty$, то $F(x)<0$ за $\alpha<x<\beta$. Но тогава от $F(r)<0$ получаваме лесно, че

$$(J)\quad |a_p|r^p>|a_0|+|a_1|r+\ldots+|a_{p-1}|r^{p-1}+|a_{p+1}|r^{p+1}+\ldots+|a_n|r^n.$$

Да означим с $\varphi(x)$ и $\psi(x)$ полиномите

$$\varphi(x)=a_0+a_1x+\ldots+a_{p-1}x^{p-1}+a_{p+1}x^{p+1}+\ldots+a_nx^n,$$

$$\psi(x)=a_px^p.$$

По окръжността $|x|=r$ ще имаме на основание на (J)

$$|\varphi(x)|\leqq|a_0|+|a_1|r+\ldots+|a_{p-1}|r^{p-1}+|a_{p+1}|r^{p+1}+\ldots+|a_n|r^n<$$
$$<|a_p|r^p=|\psi(x)|.$$

По теоремата на Руше уравнението $f(x)=\varphi(x)+\psi(x)=0$ ще има в кръга $|x|<r$ еднакъв брой корени с уравнението $\psi(x)=0$, т. е. $f(x)=0$ ще има точно p корена във въпросния кръг. Но числото r може да се вземе произволно близко до числото α. Следователно полиномът $f(x)$ ще има p нули в кръга $|x|\leqq r$ и няма да има нули във венеца $\alpha<|x|<\beta$.

3. Уравнения, на които всички корени имат отрицателна реална част.[1] Един друг въпрос от значение е да се намерят условията, щото корените на едно уравнение да бъдат с отрицателни реални части.

Отначало ще вземем някои помощни неравенства. Ако $f(x)$ е един полином, то да означим с

$$f^*(x)=\bar{f}(-x),$$

т. е. полиномът, който се получава, като в $f(x)$ заместим всички негови коефициенти с конюгованите им стойности и поставим $-x$ вместо x. Нека уравнението

$$(4)\qquad f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0$$

има само корени с отрицателна реална част, т. е.

$$(5)\qquad f(x)=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

където

$$\alpha_k=\mu_k+i\nu_k,\quad \mu_k<0,\quad k=1,\ 2,\ldots,\ n.$$

Ако с $\mathbf{R}(x)$ означим реалната част на x, то ще докажем, че

$$0\leqq|f(x)|<|f^*(x)|\ \text{ при }\ \mathbf{R}(x)<0,$$

$$(6)\qquad 0\leqq|f^*(x)|<|f(x)|\quad\text{„}\quad \mathbf{R}(x)>0,$$
$$0<|f^*(x)|=|f(x)|\quad\text{„}\quad \mathbf{R}(x)=0.$$

Благодарение на разлагането (5) достатъчно е да докажем това за един биномен множител.

$$\varphi(x)=x-\alpha_k=x-\mu_k-i\nu_k,\quad \varphi^*(x)=-x-\bar{\alpha}_k=-x-\mu_k+i\nu_k.$$

[1] I. Schur — Zeitschrift für angew. Mathematik und Mechanik, 1921, стр. 307—311

Ако поставим $x=y+it$, то веднага получаваме

$$|\varphi^*(x)|^2-|\varphi(x)|^2=[(y+\mu_k)^2+(t-\nu_k)^2]-$$
$$-[(y-\mu_k)^2+(t-\nu_k)^2=4y\mu_k,$$

отгдето неравенствата (6) следват непосредствено. Имаме следното предложение: ако числата α и β са произволни, но $|\alpha|>|\beta|$, то необходимо и достатъчно условие всичките нули на $f(x)$ да бъдат с отрицателни реални части е, щото същото да е в сила за полинома

$$g(x)=\alpha f(x)-\beta f^*(x).$$

Действително, ако $f(x)$ има само нули с отрицателни реални части то по (6) за $\mathbf{R}(x)\geqq 0$ имаме

$$|f(x)|\geqq|f^*(x)|,$$

така че понеже $|\alpha|>|\beta|$, то $|\alpha f(x)|>|\beta f^*(x)|$. Следователно $g(x)\neq 0$ за $\mathbf{R}(x)\geqq 0$. Обратно, нека $g(x)$ да има само нули с отрицателна реална част. Тогава от

$$g(x)=\alpha f(x)-\beta f^*(x)$$

имаме

$$g^*(x)=\bar{\alpha}f^*(x)-\bar{\beta}f(x).$$

От тези две равенства имаме

$$f(x)=\frac{\bar{\alpha}}{|\alpha|^2-|\beta|^2}g(x)+\frac{\beta}{|\alpha|^2-|\beta|^2}g^*(x)$$

и понеже модулът на множителя на $g(x)$ е по-голям от модула на множителя на $g^*(x)$, то по първата част на теоремата $f(x)$ ще има само нули с отрицателна реална част.

Нека ζ е число, на което $\mathbf{R}(\zeta)<0$. Ако $f(x)$ има нули само с отрицателна част, то по (6)

$$|f(\zeta)|<|f^*(\zeta)| \tag{7}$$

и по предложението уравнението

$$f^*(\zeta)f(x)-f(\zeta)f^*(x)=0 \tag{8}$$

ще има само такива корени. Обратно, ако (8) има само корени с отрицателна реална част и (7) е изпълнено, то и $f(x)$ ще има само такива нули. Когато отстраним корена ζ, получаваме теоремата:

Полиномът $f(x)$ има само тогава всичките си нули с отрицателна реална част, когато същото е изпълнено за полинома

$$f_1(x)=\frac{f^*(\zeta)f(x)-f(\zeta)f^*(x)}{x-\zeta}$$

и $|f(\zeta)|<f^*(\zeta)|$, където ζ е произволно число, за което

$$R(\zeta)<0.$$

Така въпросът за решаване дали едно уравнение е само с корени с отрицателни реални части се свежда на неравенството (7) и на уравнение от степен, по-ниска с единица. Така продължаваме, докато достигнем до уравнение от първа степен.

В случай, че уравнението има реални коефициенти, условието реалната част на корените му да бъде винаги отрицателна се значително опростява. Едно просто необходимо условие затова е следното:

1. Ако корените на уравнението с реални коефициенти

$$(1') \qquad f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0, \quad a_0>0$$

лежат наляво от имагинерната ос, то всичките му коефициенти са положителни.

Действително, ако $-\gamma_1, -\gamma_2, \ldots, -\gamma_k$ са реалните корени на (1′) и $-\alpha_\mu \pm i\beta_\mu$, $1 \leqq \mu \leqq s$ са имагинерните му корени, то трябва да имаме $\alpha_\mu > 0$, $\gamma_\nu > 0$, $1 \leqq \mu \leqq s$, $1 \leqq \nu \leqq k$. Лявата част на уравнението (1′) е произведение на a_0 с множителите $x+\gamma_\nu$ и $(x+\alpha_\mu)^2+\beta_\mu^2$, които имат положителни коефициенти и следователно коефициентите на произведението ще бъдат също така положителни.

Ще използуваме нататък и теоремата на Билер—Хермит, която гласи:

2. Ако нулите на полинома $f(x)=U(x)+iV(x)$ лежат от едната страна на реалната ос, то полиномите с реални коефициенти $U(x)$ и $V(x)$ има само реални и прости нули, които се взаимно разделят и обратно.

Нека тогава уравнението (1′) има само корени, които лежат наляво от имагинерната ос. Тогава корените на уравнението

$$F(x)=i^{-n}f(ix)=V(x)+iV_1(x)=0,$$

$$V(x)=a_0x^n-a_2x^{n-2}+a_4x^{n-4}-\ldots, \quad V_1(x)=-a_1x^{n-1}+a_3x^{n-3}-\ldots$$

ще лежат над реалната ос. Съгласно с теорема 2 полиномите $V(x)$ и $V_1(x)$ ще имат само реални прости и взаимно разделящи се нули и, обратно, ако тези полиноми имат само реални прости и взаимно разделящи се нули, като $\frac{a_1}{a_0}>0$, то уравнението $f(x)=0$ ще има само корени наляво от имагинерната ос. Да разделим $V(x)$ с $V_1(x)$. За остатъка, взет с обратен знак, получаваме

$$V_2(x)=\alpha_0x^{n-2}-\alpha_2x^{n-4}+\alpha_4x^{n-6}-\ldots,$$

където

$$\alpha_0=\frac{a_1a_2-a_0a_3}{a_1}, \quad \alpha_2=\frac{a_1a_4-a_0a_5}{a_1}, \quad \alpha_4=\frac{a_1a_6-a_0a_7}{a_1},\ldots$$

Като разделим $V_1(x)$ с $V_2(x)$, получаваме за остатъка с обратен знак полинома

$$V_3(x)=-\beta_0x^{n-3}+\beta_2x^{n-5}-\beta_4x^{n-7}+\ldots,$$

където

$$\beta_0=\frac{\alpha_0 a_3-a_1\alpha_2}{\alpha_0},\quad \beta_2=\frac{\alpha_0 a_5-a_1\alpha_4}{\alpha_0},\quad \beta_4=\frac{\alpha_0 a_7-a_1\alpha_6}{\alpha_0},\ldots$$

Като продължаваме така, получаваме полиноми $V_3(x)$, $V_4(x),\ldots,V_n(x)$, като степените им намаляват с единица. Редицата от полиноми

$$V,\quad V_1,\quad V_2,\quad V_3,\ldots,V_n$$

е очевидно една обобщена редица на Щурм. Ако V_x означава броя на вариациите на тази редица за дадено x, то броят на реалните корени на уравнението $V=0$ е по теоремата на Щурм равен на числото $|V_{-\infty}-V_{\infty}|$ и понеже предното уравнение от степен n има само реални корени, то ще имаме $|V_{-\infty}-V_{\infty}|=n$. Но като имаме пред вид, че по предложение I $V(-\infty)$ и $V_1(-\infty)$ имат еднакъв знак, заключаваме, че $V_{-\infty}=0$ и $V_{\infty}=n$, т. е. коефициентите пред най-високите степени на x в полиномите V, $V_1,\ldots,V_n$ трябва да са отлични от нула и да са алтернативно положителни и отрицателни. Лесно виждаме, че ще трябва да съставим таблицата

(A)

a_0	a_2	a_4	$a_6\ldots$
a_1	a_3	a_5	$a_7\ldots$
α_0	α_2	α_4	$\alpha_6\ldots$
β_0	β_2	β_4	$\beta_6\ldots$
.....			

Тя се образува по следния начин: Елементът, който е на k-тото място в третия ред, се получава, като от произведението на първия елемент от втория ред, с $(k+1)$-вия елемент на първия ред извадим произведението на първия елемент от първия ред с $(k+1)$-вия елемент на втория ред и получената разлика разделим на първия елемент на втория ред. По същия начин четвъртият ред се получава от втория и третия и т. н. В таблицата очевидно първите елементи на първия, третия и т. н. редове представляват коефициентите пред най-високите степени на x в полиномите V, V_2, $V_4,\ldots$, а първите елементи на втория, четвъртия и т. н. редове са съответните коефициенти за полиномите V_1, V_3, $V_5,\ldots$, но взети с обратни знаци. Така установихме следната теорема на Раус:

Необходимо и достатъчно условие уравнението с реални коефициенти

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n=0,\quad a_0\neq 0$$

да има само нули с отрицателна реална част се състои в това че елементите на първия стълб в таблицата (*A*) *да бъдат всичките отлични от нула и с едчакъв зчак.*

На Хурвиц се дължи едно друго решение на разглеждания въпрос, при което условията се дават в детерминантна форма. Теоремата на Хурвиц е следната:

Необходимо и достатъчно условие уравнението с реални коефициенти

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n=0$$

да има само корени с отрицателна реална част се състои в това, че детерминантите

$$\Delta_\lambda^{(n)}=\begin{vmatrix} a_1 & a_0 & 0 & 0 & 0 & 0 & \cdot & \cdot \\ a_3 & a_2 & a_1 & a_0 & 0 & 0 & \cdot & \cdot \\ a_5 & a_4 & a_3 & a_2 & a_1 & a_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{2\lambda-1} & a_{2\lambda-2} & \cdot & \cdot & \cdot & \cdot & \cdot & a_\lambda \end{vmatrix}, \quad \lambda=1,\ 2,\ 3,\ldots,n$$

да бъдат положителни, като се поставя $a_\lambda=0$ *при* $\lambda>n$.

Доказателството на Хурвиц[1] е доста дълго. Ще изложим[2] едно по-просто такова. За късота ще наричаме хурвицов всеки полином, нулите на който лежат наляво от имагинерната ос. Предварително ще установим още две помощни предложения.

3. Ако полиномите $\varphi(x)$ и $\psi(x)$ имат само реални прости и взаимно разделящи се нули, то същото свойство притежават и полиномите $\varphi(x)$ и $F(x)=\varphi(x)+\lambda\psi(x)$, където λ е произволно реално число, отлично от нула.

Доказателството извършваме по употребен вече в подобни въпроси начин. Нека нулите на полинома $\varphi(x)$ са $\alpha_1, \alpha_2, \ldots, \alpha_n$, като предполагаме, че сме ги наредили по растящи стойности. Тогава от равенствата $F(\alpha_k)=\lambda\psi(\alpha_k)$, $k=1, 2, \ldots, n$, следва, че всеки две последователни числа $F(\alpha_k)$ и $F(\alpha_{k+1})$ $1\leqq k\leqq n-1$, са с противни знаци, т. е. във всеки интервал (α_k, α_{k+1}), $1\leqq k\leqq n-1$, полиномът $F(x)$ има нечетен брой нули. Следователно, ако степента на $\psi(x)$ е равна на n или на $n-1$, то нулите на $F(x)$ ще бъдат реални, прости и ще отделят нулите на $\varphi(x)$. Ако степента на $\psi(x)$ е равна на $n+1$, то достигаме до същото заключение, като вземем пред вид, че за $x=-\infty$ и за $x=\infty$ знакът на $F(x)$ съвпада със знака на $\lambda\psi(x)$.

4. Необходимо и достатъчно условие, щото уравнението с реални коефициенти

$$f(x)=\varphi(x^2)+x\psi(x^2)=0,$$

където степените на полиномите $\varphi(x)$ и $\psi(x)$ се отличават най-много с единица, да има само корени от едната страна на имагинерната ос се състои в това, че полиномите $\varphi(x)$ и $\psi(x)$ да имат само реални отрицателни и взаимно разделящи се нули, като при равни степени нулите на полинома $\varphi(x)$ да следват тези на полинома $\psi(\ )$.

Действително нека $f(x)$ има само нули от едната страна на имагинерната ос. Тогава нулите на полинома

$$f(ix)=\varphi(-x^2)+ix\psi(-x^2)$$

---

[1] Hurwitz, Mathematische Annalen, т. 46 (1896).

[2] I. Schur, Zeitschr ft f. angew. Math. und Mechanik, 1921, Lienard, Journal de mathém. pures et appliquées, 1936. N. Obreschkoff. Mathem. Zeitschrift, 1939.

ще лежат от едната страна на реалната ос. По теоремата на Билер — Хермит полиномите $\varphi(-x^2)$ и $x\psi(-x^2)$ ще имат само реални и взаимно разделящи се нули. Следователно $\varphi(x)$ и $\psi(x)$ ще имат само отрицателни и взаимно разделящи се нули, като при равни степени нулите на $\varphi(x)$ ще следват тези на $\psi(x)$. Обратно, ако последните условия за полиномите $\varphi(x)$ и $\psi(x)$ са изпълнени, то по обратната теорема на Билер – Хермит следва, че нулите на $f(ix)$ лежат от едната страна на реалната ос и предложението 3 е така установено. Ще установим сега следното предложение на Шур:

4. Полиномът

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n= \tag{10}$$

$$=\varphi(x^2)+x\psi(x^2),\quad a_0>0,$$

$$\varphi(x)=a_0+a_2x+a_4x^2+\ldots,\quad \psi(x)=a_1+a_3x+a_5x^2\ldots$$

с реални коефициенти е само тогава хурвицов, ако полиномът

$$f_1(x)=\psi(x^2)\left(1-\frac{a_0}{x}\right)+\frac{a_1}{x}\varphi(x^2)= \tag{11}$$

$$=a_1+(a_2a_1-a_0a_3)x+a_3x^2+(a_4a_1-a_0a_5)x^3+\ldots$$

е хурвицов и $a_1>0$.

Нека $f(x)$ е хурвицов полином. Тогава по 1 имаме $a_1>0$, което впрочем се вижда и веднага. По предложение 3 нулите на полиномите $\varphi(x)$ и $\psi(x)$ са реални, отрицателни и се взаимно разделят; ако ги означим с $-\alpha_i$ и $-\beta_i$, ще имаме

$$0<\alpha_1<\beta_1<\alpha_2<\beta_2<\ldots \tag{12}$$

От уравненията

$$\frac{a_2}{a_0}=\frac{1}{\alpha_1}+\frac{1}{\alpha_2}+\ldots,\quad \frac{a_3}{a_1}=\frac{1}{\beta_1}+\frac{1}{\beta_2}+\ldots$$

и неравенствата (12) следва, че

$$\frac{a_2}{a_0}>\frac{a_3}{a_1},\ \text{т. е.}\ a_1a_2-a_0a_3>0.$$

Да разгледаме полинома $g(x)$, дефиниран с

$$xg(x)=a_1\varphi(x)-a_0\psi(x).$$

По предложение 2 полиномите $xg(x)$ и $\psi(x)$ имат само реални взаимно разделящи се нули. Следователно $g(x)$ има само реални и отрицателни нули, които разделят тези на $\psi(x)$. $xg(x^2)$ и $\psi(x^2)$ имат тогава само нули по имагинерната ос, които се взаимно разделят по нея. По обратната теорема на Билер и Хермит нулите на полинома

$$f_1(x)=\psi(x^2)+xg(x^2)=b_0+b_1x+\ldots+b_{n-1}x^{n-1}$$

ще лежат от едната страна на имагинерната ос. Тази страна е лявата, понеже

$$b_0=a_1>0, \quad b_1=a_1a_2-a_0a_3>0.$$

Нека сега, обратно, полиномът (11) е хурвицов и $a_1>0$. Тогава нулите на полиномите $xg(x)$ и $\psi(x)$ са реални и се взаимно разделят, като тези на $\psi(x)$ са отрицателни. Понеже

$$a_1\varphi(x)=xg(x)+a_0\psi(x),$$

то $\varphi(x)$ има само реални нули, които отделят тези на $\psi(x)$. Следователно $\varphi(x)$ има най-много една положителна нула. Но при n четно имаме

$$\varphi(x)=a_0+a_2x+\ldots+a_nx^{\frac{n}{2}},$$

като $b_{n-1}=a_1a_n$. Понеже $b_{n-1}>0$, то следва, че $a_n>0$ и следователно $\varphi(x)$ няма положителна нула. При n нечетно имаме

$$\varphi(x)=a_0+a_2x+\ldots+a_{n-1}x^{\frac{n-1}{2}},$$

$$\psi(x)=a_1+a_3x+\ldots+a_nx^{\frac{n-1}{2}},$$

като

$$b_{n-1}=a_n, \quad b_{n-2}=a_1a_{n-1}-a_0a_n.$$

Понеже по 1 имаме $b_{n-1}>0$, $b_{n-2}>0$, то следва, че $a_n>0$, $a_1a_{n-1}>a_0a_n$, т. е. $a_{n-1}>0$. Следователно и в този случай $\varphi(x)$ няма положителна нула. Така се убеждаваме, че $\varphi(x)$ има само отрицателни реални нули, които следват тези на $\psi(x)$, понеже $b_2=a_2a_1-a_0a_3>0$ дава $\frac{a_2}{a_0}>\frac{a_3}{a_1}$.

От предното следва, че полиномите $x\psi(x)$ и $\varphi(x)$ имат само реални отрицателни нули, които се взаимно разделят. Нулите на $f(x)=\varphi(x^2)+x\psi(x^2)$ ще лежат от едната страна на имагинерната ос, която е лявата, понеже $a_0a_1>0$.

Теоремата на Хурвиц следва лесно от горното предложение по индуктивен път. Допускаме, че теоремата е установена за уравненията от $(n-1)$-ва степен. Нека $\Delta_\lambda^{(n-1)}$ са детерминантите за $f_1(x)$. Пишем ги във формата

$$\Delta_\lambda^{(n-1)}=\frac{1}{a_1}\begin{vmatrix} a_1 & 0 & 0 & \cdot & \cdot \\ a_3 & a_2a_1-a_0a_3 & a_1 & \cdot & \cdot \\ a_5 & a_4a_1-a_0a_5 & a_3 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}.$$

Умножаваме първия стълб с a_0 и го прибавяме към втория, третия стълб с a_0 и го прибавяме към четвъртия и т. н.

Получаваме така

$$\Delta_\lambda^{(n-1)} = a_1^p \, \Delta_\lambda^{(n)}, \; p = \left[\frac{\lambda+1}{2}\right] - 1.$$

Условията

$$\Delta_\lambda^{(n-1)} > 0, \; \lambda = 1, \; 2, \ldots, \; n-1, \; a_1 > 0$$

са еквивалентни с

$$\Delta_\lambda^{(n)} > 0, \; \lambda = 1, 2, \ldots, \; n.$$

4. Брой на корените в една окръжност. Отначало да намерим условието, щото всички корени на едно уравнение да лежат в една окръжност. Очевидно с подходяща линейна трансформация можем да сведем тази окръжност да бъде с център в нулевата точка и радиус, равен на 1, наричана често и единична окръжност. Предварително ще въведем някои означения. Нека $f(x)$ е полином с нули $\alpha_1, \alpha_2, \ldots, \alpha_n$,

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n = a_0 (x-\alpha_1) \ldots (x-\alpha_n). \tag{13}$$

Да означим с $f^*(x)$ полинома

$$f^*(x) = x^n \overline{f}\left(\frac{1}{x}\right) = \overline{a}_n x^n + \overline{a}_{n-1} \, x^{n-1} + \ldots + \overline{a}_0$$

(този полином е различен от съответния, дефиниран в предния параграф). Ще имаме

$$f^*(x) = \overline{a}_0 (1 - \overline{\alpha}_1 x) \; (1 - \overline{\alpha}_2 x) \ldots (1 - \alpha_n x), \tag{14}$$

отгдето се вижда, че нулите му са

$$\frac{1}{\overline{\alpha}_1}, \quad \frac{1}{\overline{\alpha}_2}, \ldots, \frac{1}{\overline{\alpha}_n}.$$

От теорията на комплексните числа веднага се вижда, че това са инверзните точки на $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Лесно е да се види, че по окръжността $|x| = 1$ имаме

$$|f(x)| = |f^*(x)|. \tag{15}$$

Действително, ако $|x| = 1$, то ще имаме

$$\left|\frac{x-\alpha}{1-x\alpha}\right| = 1. \tag{16}$$

Това лесно се проверява директно, понеже $\overline{(x-\alpha)} = \frac{1}{x} - \overline{\alpha} = \frac{1}{x}(1 - x\overline{\alpha})$, но може също да се установи, като се вземе под внимание, че дробната линейна трансформация $\frac{\lambda x + \mu}{\delta x + \nu}$ обръща окръжност в окръжност (в частност права) и вземем под внимание, че при $x = 1, \; -1, \; i,$

$-i$ числата $1-x\bar{\alpha}$ са конюговани на $x-\alpha$, последните, умножени с ± 1 или $\pm i$, така че (16) е за тях изпълнено. Но тогава

$$\left|\frac{f(x)}{f^*(x)}\right|=\left|\frac{x-\alpha_1}{1-x\bar{\alpha}_1}\right|\cdot\cdot\cdot\left|\frac{x-\alpha_n}{1-x\bar{\alpha}_n}\right|\cdot\left|\frac{a_0}{\bar{a}_0}\right|=1 \text{ за } |x|=1.$$

Сега ще можем да установим следната теорема[1] на Шур. За да има уравнението

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0,\ a_0\neq 0$$

само корени в кръга $|x|<1$, необходимо и достатъчно е, че $|a_0|>|a_n|$ и че полиномът

$$f_1(x)=\frac{1}{x}[\bar{a}_0f(x)-a_nf^*(x)] \tag{17}$$

от степен $(n-1)$ да има само нули в кръга $|x|<1$. Действително, ако корените на $f(x)=0$ са в кръга $|x|<1$, то понеже

$$(-1)^n\frac{a_n}{a_0}=\alpha_1\alpha_2\ldots\alpha_n,$$

ще имаме

$$\left|\frac{a_n}{a_0}\right|<1,$$

т. е. $|a_0|>|a_n|$, и по теоремата на Руше, понеже за $|x|=1$,

$$|\bar{a}_0f(x)|>|a_nf^*(x)|,$$

то полиномът

$$\bar{a}_0f(x)-a_nf^*(x)=xf_1(x)$$

ще има n нули в $|x|<1$. Обратното при $|a_0|>|a_n|$ по теоремата на Руше е пак очевидно.

Този въпрос в същност не е различен от въпроса за броя на корените наляво или дясно от имагинерната ос. Така, ако поставим

$$x=\frac{1+y}{1-y}\cdot$$

то на правата $\mathbf{R}(y)=0$ отговаря окръжността $|x|=1$, понеже при $y=it$

$$|x|=\left|\frac{1+it}{1-it}\right|=1,$$

и на половин равнината

$$\mathbf{R}(y)\leqq 0$$

отговаря кръгът $|x|\leqq 1$, като на неравенство в първото отговаря пак неравенство във второто. Така че броят на корените на $f(x)=0$,

[1] I. Schur — Journal f. Math. 148, 1918, стр. 122 — 145. A. Cohn. Math. Zeitschr. 14, 1922, стр. 110—138.

които са вътре в кръга $|x| \leqq 1$, по периферията му или вън от кръга, ще бъде равен съответно на броя на корените на

$$(1-y)^n f\left(\frac{1+y}{1-y}\right)=0,$$

които са наляво от имагинерната ос, по нея или надясно от нея.

Ще разгледаме сега един метод на А. Кон[1] за определяне на броя на нулате на полинома (13), лежащи в една окръжност, която, както споменахме, можем да считаме, че е окръжността K с център $z=0$ и радиус 1. Този метод се състои в четири правила.

1. Нека $|a_0| > |a_n|$. Образуваме

$$\bar{a_0} f(x) - a_n f^*(x) = x f_1(x).$$

Полиномът $f_1(x)$ е най-много от $n-1$-ва степен. Тогава $f(x)$ има една нула повече в окръжността K, отколкото $f_1(x)$.

II. Нека $|a_0| < |a_n|$. Тогава полиномът

$$\bar{a_n} f(x) - a_0 f^*(x) = f_1(x)$$

е от степен най-много $n-1$-ва и има в окръжността K толкова нули колкото полиномът $f(x)$.

III. Нека $a_0 = \varepsilon \bar{a_n}$, $a_1 = \varepsilon a_{n-1}, \ldots,$ $a_{k-1} = \varepsilon \bar{a}_{n-k+1}$, където

$$\varepsilon| = 1,\ k \leqq \left[\frac{n}{2}\right],\ a_k \neq \varepsilon \bar{a}_{n-k}.$$

Ако означим с

$$\frac{a_k - \varepsilon a_{n-k}}{a_0} = b,$$

то прилагаме правилото II за полинома

$$G(x) = \left(x^k + 2\frac{b}{|b|}\right) f(x).$$

Тогава получаваме едно уравнение от n-та степен, което попада под правило I и което има еднакъв брой корени в K с уравнението $f(x)=0$.

Ако няма число k, така че $a_k \neq \varepsilon \bar{a}_{n-k}$, то имаме новия случай:

IV. $$a_\nu = \varepsilon \bar{a}_{n-\nu},\ |\varepsilon| = 1,\quad \nu = 0, 1, 2, \ldots, n.$$

Тогава полиномът

$$f_1(x) = x^{n-1} f'\left(\frac{1}{x}\right)$$

има толкова нули в кръга $|x|<1$, колкото полиномът $f(x)$.

---

[1] Виж цитираната му вече работа.

От горните правила се вижда, че проблемата за определяне броя на нулите на един полином от n-та степен се свежда на тази за полином от $(n-1)$-ва степен.

За определяне на броя на нулите на $f(x)$ по периферията на кръга $|x| \leqq 1$, т. е. по окръжността K, имаме следното правило на Кон:

Ако $f_m(x)$ е първият полином, попадащ под правилото IV, и ако полиномът

$$f_{m+1}(x) = x^{s-1} f'_m\left(\frac{1}{x}\right),$$

гдето s е степента на $f_m(x)$, има l нули в кръга $|x|<1$, то $f(x)$ има $s-2l$ нули с модул 1.

При доказателството използуваме теоремата на Руше: Ако за два полинома $\varphi(x)$ и $\psi(x)$ по K имаме $|\varphi(x)| > |\psi(x)|$, то полиномите $\varphi(x)$ и $\varphi(x)+\varphi(x)$ имат еднакъв брой нули в K.

Полиномите $f(x)$ и $f^*(x)$ имат същите нули $\alpha_1, \alpha_2, \ldots, \alpha_p$ с модул 1. Ако поставим

$$F_1(x) = (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_p),$$

$$f(x) = F_1(x) F(x),$$

то ще имаме

$$f^*(x) = -\bar{\alpha}_1 \bar{\alpha}_2 \ldots \bar{\alpha}_p F_1(x) F^*(x),$$

гдето

$$F^*(x) = x^{n-p} \bar{F}\left(\frac{1}{x}\right).$$

Следователно за $|x|=1$ ще имаме $|F(x)| = |F^*(x)|$. Ако $|\mu|<1$, то по окръжността $K(|x|=1)$ ще имаме

$$|F(x)| > |\mu F^*(x)|.$$

Следователно $F(x)$ и $F(x)+\mu F^*(x)$ ще имат еднакъв брой нули в K. Ако умножим тези полиноми с $F_1(x)$, получаваме, че полиномите

$$f(x),\ f(x)+\lambda f^*(x),\ \lambda = \mu \prod_{s=1}^{p}\left(-\frac{1}{\bar{\alpha}_s}\right)$$

при $|\lambda|<1$ ще имат еднакъв брой нули в K. Оттук непосредствено се получават правилата I и II.

За да установим правилото III, нека означим с $\gamma = \frac{2b}{|b|}$.

Тогава за полинома $G(x)$ имаме

$$G(x) = \varepsilon \bar{a}_n x^{n+k} + \varepsilon \bar{a}_{n-1} x^{n+k-1} + \ldots + \varepsilon \bar{a}_{n-k+1} x^{n+1} +$$

$$+ (\gamma \varepsilon \bar{a}_n + a_n) x^n + \ldots + (\gamma a_{n-k} + a_n) x^k +$$

$$+ \gamma a_{n-k+1} x^{k-1} + \ldots + \gamma a_n.$$

Понеже $|\varepsilon \bar{a}_n| < |\gamma a_n|$, то правило II се прилага и по него трябва да образуваме полинома

$$g(x)=\bar{\gamma}\,\bar{a}_n G(x)-\varepsilon\,\bar{a}_n G^*(x)=$$

$$=\bar{a}_n\{[(\gamma\bar{\gamma}-1)\varepsilon\bar{a}_n+\bar{\gamma}(a_n-\varepsilon\bar{a}_{n-k})]x^n+\ldots+(\gamma\bar{\gamma}-1)a_n\},$$

т. е.

$$\frac{1}{\bar{a}^n}g(x)=(1+2|b|)a_0x^n+\ldots+a_n.$$

Очевидно полиномът $G(x)$ има толкова нули в кръга $|x|<1$, колкото полиномът $f(x)$. Следователно на основание на правилото II полиномът $g(x)$ ще има еднакъв брой нули с полинома $f(x)$ в същия кръг. Понеже $(1+2|b|)|a_0|>|a_n|$, то можем да приложим за $\frac{1}{\bar{a}_n}g(x)$ правилото I.

При случая IV полиномът $f(x)$ ще удовлетворява на равенството

$$\bar{a}_nf(x)=a_0f^*(x).$$

Правилото IV се основава на едно предложение:

За всеки полином $\varphi(x)$ съществува едно произволно малко избрано) число $\delta>0$, такова, че полиномът

$$\varphi_\delta(x)=\varphi(x)-\delta x\varphi'(x)$$

има еднакъв брой нули в кръга K, както $\varphi(x)$.

Ако $\varphi(x)$ няма нули по $|x|=1$, то на основание на непрекъснатостта на нулите на $\varphi_\delta(x)$ спрямо δ (вж. част IV, глава IV, § 2) предложението е очевидно. Остава да се докаже, че при достатъчно малки δ нулите на $\varphi(x)$, лежащи по окръжността K, не могат да навлизат вътре в нея.

Нека α е една m-кратна нула на $\varphi(x)$. Ще имаме

$$\varphi(x),=(x-\alpha)^m\psi(x),\ \psi(\alpha)\neq0$$

и

$$\varphi_\delta(x)=(x-\alpha)^{m-1}h(x),$$

гдето

$$h(x)=[(1-m\delta)\ x-\alpha]\psi(x)-\delta x(x-\alpha)\psi'(x).$$

α е $(m-1)$-кратен корен на $\varphi_\delta(x)=0$. Ще докажем, че на този корен α ще съответствува един друг корен на същото уравнение, който при достатъчно малки $\delta>0$ ще лежи в кръг C_δ с център

$$(1+m\delta)\alpha$$

и с радиус

$$\frac{|\alpha|}{2}\delta,$$

отгдето следва, че той ще има модул, по-голям от този на α.

Да изберем отначало δ така малко, че $\psi(x)$ да се анулира само вън от C_δ, което е възможно, понеже $\psi(\alpha) \neq 0$. Освен това и коренът

$$\frac{\alpha}{1-m\delta} = (1+m\delta)\alpha + \frac{\alpha m^2\delta^2}{1-m\delta}$$

на уравнението

$$[(1-m\delta)x-\alpha]\psi(x)=0$$

да лежи в кръга C_δ. Последното уравнение ще има тогава в C_δ само един корен. За да установим сега, че полиномът $h(x)$ при подбрано δ има само една нула в C_δ, достатъчно е да докажем на основание на теоремата на Руше, че по окръжността C_δ при достатъчно малки $\delta>0$ имаме

$$|[(1-m\delta)x-\alpha]\psi(x)| > |\delta x(x-\alpha)\psi'(x)|.$$

Но по C_δ имаме

$$x=(1+m\delta)\alpha+\gamma\delta,$$

гдето γ е комплексно число с модул, равен на $\frac{|\alpha|}{2}$. Горното неравенство става

$$|(1-m\delta)[(1+m\delta)\alpha+\gamma\delta]-\alpha\}\psi(x)| >$$
$$> \delta^2|(1+m\delta)\alpha+\gamma\delta|(m\alpha+\gamma)|\psi'(x)|$$

или

$$|[\gamma-m(m\alpha+\gamma)\delta]\psi(x)| > \delta|(1+m\delta)\alpha+\gamma\delta](m\alpha+\gamma)\psi'(x)|,$$

което действително при достатъчно малка δ се удовлетворява.

Понеже при δ подходящо малко кръговете C_δ за различните корени α нямат общи точки, $\varphi_\delta(x)$ ще има толкова нули по K и вън от K, колкото има полиномът $\varphi(x)$ по K. Като изберем още δ така малко, че нулите на $\varphi(x)$, лежащи вътре в K, да остават пак там, и вземем под внимание, че степените на тези полиноми са равни, предложението става очевидно.

Нека сега $f(x)=0$ е уравнение, което удовлетворява на условията в IV, т. е.

$$\bar{a}_n f(x)=a_0 f^*(x).$$

По предното предложение има число $\delta>0$, $\delta<\frac{1}{n}$, такова, че уравнението

(8) $$f(x)-\delta x f'(x)=(1-n\delta)a_0 x^n+\ldots+a_n=0$$

има еднакъв брой корени в кръга $|x|<1$, като уравнението

$$f(x)=0.$$

Понеже $|a|=|a_n|$, следователно

$$(1-n\delta)|a_0|<|a_n|,$$

то уравнението пада под случая II. Следователно уравнението

$$\varphi(x)=\bar{a}_n[f(x)-\delta x f'(x)]-(1-n\delta)a_0[f(x)-\delta x f'(x)]^*=0$$

има в кръга на $|x|<1$ същия брой корени, както уравнението

$$f(x)=0.$$

С преобразуване получаваме

$$\varphi(x)=a_0\delta\ (2-n\delta)\, f^*(x).$$

Така установихме следната теорема:

В случая IV уравненията $f(x)=0$ и $f'^*(x)=0$ имат в кръга $|x|<1$ еднакъв брой корени.

Понеже корените на уравнението

$$x^{n-1}f'\left(\frac{1}{x}\right)=0 \tag{19}$$

са конюгованите стойности на корените на

$$f'^*(x)=x^{n-1}\overline{f'}\left(\frac{1}{x}\right)=0, \tag{20}$$

т. е. тези две уравнения имат еднакъв брой корени в кръга $|x|<1$, то в горната теорема може уравнението (20) да се замести с (19) и така получаваме правилото IV.

От доказателствата на правилата I, II, III се вижда лесно, че броят на нулите по окръжността $|x|=1$ не се променя. Нека тогава при прилагането им върху даденото уравнение първото поред уравнение, попадащо към правилото IV, да бъде

$$f_m(x)=0, \tag{21}$$

степента на което да бъде s.

Понеже тогава $f_m{}^*(x)$ до постоянен множител е равен на $f_m(x)$, то ако α е произволен корен на уравнението (21), $\frac{1}{\alpha}$ е пак корен на същото уравнение. Следователно (21) ще има еднакъв брой корени вътре и вън на K. Ако означим с l този общ брой, то броят на корените на (21), лежащи по окръжността $|x|=1$, е равен на

$$t=s-2l.$$

По правилото IV броят на корените на уравнението

$$f_{m+1}(x)=x^{s-1}f_m'\left(\frac{1}{x}\right)=0,$$

които лежат в K, ще бъде равен на l.

Очевидно от правилото на Кон следва теоремата на Шур. Като друго следствие ще изведем следната теорема на Шур: Необходимо и достатъчно условие, щото корените на уравнението

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n=0 \tag{22}$$

да лежат на окръжността $|x|=1$, е да имаме

$$a_\nu=\varepsilon\bar{a}_{n-\nu},\ |\varepsilon|=1,\ \nu=0,\ 1,\ 2,\ldots,\ n, \tag{23}$$

и всеки корен на уравнението

$$f'(x)=0 \tag{24}$$

да лежи в кръга $|x|\leqq 1$ или по периферията му.

Действително, ако уравнението (22) има само корени по $|x|=1$, които да означим с $\alpha_1,\ \alpha_2,\ldots,\ \alpha_n$, то понеже модулите им са равни на 1, корените $\frac{1}{=}=\alpha_\nu,\ \nu=1,\ 2,\ldots,\ n$, на уравнението

$$\alpha_\nu$$

$$f^*(x)=x^n\bar{f}\left(\frac{1}{x}\right)=0$$

съвпадат с тях. Следователно полиномите $f^*(x)$ и $f(x)$ се отличават с един постоянен множител и оттук получаваме равенствата (23). Съгласно с правилата на Кон уравнението

$$x^{n-1}f'\left(\frac{1}{x}\right)=0 \tag{25}$$

не ще има корени с модул, по-малък от 1. Оттук следва, че корените на (24) ще лежат или вътре, или по периферията на K. Обратно, ако равенствата (23) са в сила и уравнението (24) има само корени с модул $\leqq 1$, то по същите правила следва, че (22) има само корени с модул 1, понеже (25) няма корени в кръга $|x|<1$.

5. Уравнение с положителни коефициенти. За такива уравнения ще докажем някои прости теореми. Една такава е следната.[1] Ако в полинома

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n$$

имаме

$$a_0>a_1>a_2>\ldots>a_n>0,$$

то нулите му са в кръга $|x|<1$.

Това предложение лесно се доказва по теоремата на Шур. Именно в случая имаме

$$f_1(x)=a_0^{(1)}x^{n-1}+a_1^{(1)}x^{n-2}+\ldots+a^{(1)}_{n-1},$$

където

$$a_\nu^{(1)}=a_0a_\nu-a_na_{n-\nu},$$

също

$$a_0^{(1)}>a_1^{(1)}>\ldots>a^{(1)}_{n-1}.$$

Така въпросът се свежда на уравнение от първа степен, за което предложението е очевидно.

---

[1] Kakeya. The Tôhoku Math. Journal, 1912. G. Eneström, The Tôhoku Math, Jour- 1920, Stockh. Öfv. L., 1893 г.

Обаче предложението може да се докаже и директно, като за удобство му дадем друга формулировка, като вместо x сме поставили $\frac{1}{x}$. Именно ще докажем: ако в полинома

$$P(x)=a_0+a_1x+\ldots+a_nx^n$$

имаме[1]

$$a_0>a_1>a_2>\ldots>a_n>0,$$

то всичките му нули са вън от окръжността $|x|=1$. Действително имаме

$$(1-x)P(x)=$$

$$=a_0-[(a_0-a_1)x+(a_1-a_2)x^2+\ldots+(a_{n-1}-a_n)x^n+a_nx^{n+1}].$$

Следователно

$$|(1-x)P(x)|\geqq a_0-(a_0-a_1)|x|-(a_1-a_2)|x|^2-\ldots-$$
$$-(a_{n-1}-a_n)|x|^n-a_n|x|^{n+1},$$

гдето знака равенство можем да имаме само тогава, когато $x\geqq 0$. Оттук при $|x|\leqq 1$, $x\neq 0$, $x\neq 1$, ще имаме

$$|(1-x)P(x)|>a_0-(a_0-a_1)-\ldots-(a_{n-1}-a_n)-a_n=0.$$

Понеже освен това $P(0)=a_0\neq 0$, $P(1)\neq 0$, то за

$$|x|\leqq 1,\ P(x)\neq 0.$$

Като следствие ще изведем следното правило: Ако полиномът

$$f(x)=a_0+a_1x+\ldots+a_nx^n$$

има положителни коефициенти, то модулите на нулите се намират между минимума и максимума на числата

$$\frac{a_0}{a_1},\ \frac{a_1}{a_2},\ldots,\ \frac{a_{n-2}}{a_{n-1}},\ \frac{a_{n-1}}{a_n}.$$

Действително, ако направим субституцията $x=\lambda y$ и приложим доказаната теорема за уравнението

$$a_0+a_1\lambda y+a_2\lambda^2y^2+\ldots+a_n\lambda^ny^n=0,$$

ще имаме за корените $|y|<1$ при

$$\frac{a_0}{a_1\lambda}<1,\ldots,\ \frac{a_{n-1}}{a_n\lambda}<1.$$

Следователно $|x|<\lambda$. Ако същите отношения са >1, то ще имаме $|y|>1$, т. е. $|x|>\lambda$.

---

[1] Ако допуснем тук да фигурират и равенства, читателят ще види лесно на основание на следващото доказателство, че полиномът $P(x)$ има нули и по $|x|=1$ само тогава, ако знакът неравенство имаме на равноотстоящи места, $a_{q\nu-1}>a_{q\nu}$, q дели $n+1$, $\nu=1, 2,\ldots,\frac{n+1}{q}$, като в останалите места има само равенство.

6. Решение на тези въпроси с квадратични форми.[1] Същите въпроси се решават и с теорията на квадратичните форми. Отначало ще се занимаваме с определяне на броя на корените в единичната окръжност, която означихме с K. Нека

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n=0$$

да бъде даденото уравнение с произволни комплексни коефициенти. Нека

$$f^*(x)=x^n\overline{f}\left(\frac{1}{x}\right)=\overline{a}_n+\overline{a}_{n-1}x+\ldots+\overline{a}_0x^n$$

и

$$k(f)=\frac{f(x)f^*(y)-f(y)f^*(x)}{x-y}=\sum_{i,k=0}^{n-1}A_{ik}x^iy^k. \tag{26}$$

Ако в (26) заместим x, y с $\frac{1}{x}$, $\frac{1}{y}$, после умножим с x^{n-1} y^{n-1} и вземем на коефициентите конюгованите им стойности, то лявата част не се изменя, а дясната преминава в

$$\sum\overline{A}_{ik}x^{n-1-i}\ y^{n-1-k},$$

така че ще имаме

$$\overline{A}_{ik}=A_{n-1-i,\ n-1-k}.$$

От друга страна, $A_{ik}=A_{ki}$ така, че ако поставим

$$a_{ik}=A_{i,\ n-1-k},$$

то ще имаме

$$\overline{a}_{ik}=\overline{A}_{i,\ n-1-k}=A_{n-1-i,k}=A_{k,n-1-i}=a_{ki}.$$

Ако в (26) заместим x^i с u_i, y^{n-1-k} с $\overline{u}_k$, то ще получим следователно една хермитова форма

$$H(f;\ u_0,\ u_1,\ldots,\ u_{n-1})=\sum u_{ik}u_i\overline{u}_k,$$

която наричаме принадлежаща към $f(x)$ хермитова форма.

Ако имаме $f(x)=f_1(x)f_2(x)$, гдето

$$f_1(x)=b_0+b_1x+\ldots+b_px^p,\ f_2(x)=c_0+c_1x+\ldots+c_qx^q \qquad (p+q=n),$$

то получаваме $f^*(x)=f_1^*(x)f_2^*(x)$ и

$$k(f)=f_1(x)f_1^*(y)k(f_2)+f_2(y)f_2^*(x)k(f_1).$$

Ако поставим

$$k(f_2)=\sum_{i,k=0}^{q-1}C_{i,k}x^iy^k,\ C_{i,q-1-k}=c_{ik},$$

---

[1] Виж работите на I. Schur — Journal f. Mathem. 148 (1918), стр. 125—145, A. Cohn. Math. Zeitschr. 14 (1922), s. 110—138, особено Lienard et Chipart, Journal de Math. 10 (1914), стр. 271—346 и M. Fujiwara, Math. Zeitschr. 24 (1925), стр.161—169.

то ще имаме

$$f_1(x)\,f_1^*(y)\,k(f_2)=$$

$$=\sum C_{ik}\,(b_0\,x^i+\ldots+b_p\,x^{i+p})\,(\overline{b}_p\,y^k+\ldots+\overline{b}_0\,y^{k+p}).$$

Ако заместим x^i с u_i, y^{n-1-k} с $\overline{u}_k$, то дясната част преминава в

$$\sum_{i,k=0}^{q-1} C_{ik}\,w_i\,\overline{w}_{q-1-k}=\sum_{i,k=0}^{q-1} c_{ik}\,w_i\,\overline{w}_k,$$

гдето

$$w_i=b_0\,u_i+b_1\,u_{i+1}+\ldots+b_p\,u_{i+p},$$

която квадратична форма представлява

$$H(f_2;\ w_0,\ w_1,\ldots,\ w_{q-1}).$$

По същия начин се получава със заместване на x^i, y^{n-1-k} с u_i, $\overline{u}_k$, от $f_2(y)\,f_2(x)\,k(f_1)$ хермитовата форма

$$H(f_1;\ v_0,\ v_1,\ldots,\ v_{p-1}),$$

принадлежаща на $f_1(x)$, гдето

$$v_i=\overline{c}_q\,u_i+\overline{c}_{q-1}\,u_{i+1}+\ldots+\overline{c}_0\,u_{i+q}.$$

Така получаваме основната релация

$$H(f;\ u_0,\ u_1,\ldots,\ u_{n-1})=$$

$$=H(f_1;\ v_0,\ v_1,\ldots,\ v_{p-1})+H(f_2;\ w_0,\ w_1,\ldots,\ w_{q-1}).$$

Променливите на брой n

$$v_0,\ v_1,\ldots,\ v_{p-1},\ w_0,\ w_1,\ldots,\ w_{q-1}$$

са n линейни форми на $u_0, u_1,\ldots, u_{n-1}$, на които детерминантата от коефициентите, както лесно се вижда, съвпада с резултантата на $f_1(x)$ и $f_2^*(x)$. Следователно те са линейно независими, ако последните полиноми нямат обща нула.

За $f(x)=a(x-\alpha)$ получаваме $k=|a|^2(1-|\alpha^2|)$, следователно

$$H=|a|^2(1-|\alpha|^2)\,u_0\,\overline{u}_0.$$

Оттук, ако

$$f(x)=a(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

то

$$H(f;\ u_0, u_1,\ldots, u_{n-1})=\sum_{k=1}^{n}(1-|\alpha_k|^2)\,V_k\,\overline{V}_k,$$

гдето V_k са линейни форми на $u_0,\ u_1\ldots,\ u_{n-1}$. Тези форми са линейно независими, ако $f(x)$ и $f^*(x)$ нямат общ корен и само тогава. Това условие е еквивалентно с това, че рангът на формата H е равен на n, понеже дискриминантата на H е отлична само с един постоянен мно-

жител от резултантата на $f(x)$ и $f^*(x)$. Така получаваме теоремата на Шур:

Необходимото и достатъчно условие, щото корените на уравнението

$$f(x)=a_0+a_1x+\ldots+a_nx^n=0$$

да бъдат с модули, по-малки от единица, е хермитовата форма

$$H(f;\ u_0, u_1, \ldots, u_{n-1})=\sum_{\lambda=1}^{n} |\ a_\lambda u_0+a_{\lambda+1}u_1+\ldots+a_nu_{n-\lambda}\ |^2-$$

$$-\sum_{\lambda=1}^{n} |\ \bar{a}_0u_{n-\lambda}+\bar{a}_1u_{n-1-\lambda}+\ldots+\bar{a}_{n-\lambda}u_0\ |^2$$

да бъде положително дефинитна.

Също получаваме и обобщението на Кон: Нека r и s са броевете на положителните и отрицателни квадрати в каноничното представяне на $H(f)$. Ако $r+s=n$, то броят на корените, на които модулът е по-малък от единица, е равен на r, а броят на тези, на които модулът е по-голям от единица, е равен на s.

Сега ще се занимаваме с корените, на които реалните части са отрицателни.

Нека

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n=0$$

е едно алгебрическо уравнение с произволни реални или имагинерни коефициенти и нека сега поставим

$$f^*(x)=\bar{f}(-x)=\bar{a}_0-\bar{a}_1x+\bar{a}_2x^2-\ldots+(-1)^n\bar{a}_nx^n,$$

$$K(f)=\frac{f(x)f^*(y)-f(y)f^*(x)}{x-y}=\sum_{i,k=0}^{n-1}A_{ik}\,x^i\,y^k.$$

Ако заместим x, y с $-x,-y$ и коефициентите с адюнгиранит им стойности, то лявата част остава непроменена, а дясната минава в

$$\sum\ (-1)^{i+k}\bar{A}_{ik}x^iy^k,$$

отгдето следва, че $\bar{A}_{ik}=(-1)^{i+k}A_{ik}$. Ако тогава поставим

$$a_{ik}=(-1)^kA_{ik},$$

то понеже $A_{ik}=A_{ki}$, получаваме $\bar{a}_{ik}=a_{ki}$. Ако в K заместим x^i с u_i, y^k с $(-1)^k\bar{u}_k$, то получаваме една към $f(x)$ принадлежаща хермитова форма, която ще означим

$$H(f;\ u_0,\ u_1, \ldots,\ u_{n-1})=\sum a_{ik}\,u_i\bar{u}_k.$$

С подобни на по-раншните разглеждания установяваме лесно, че ако $f(x)=f_1(x)f_2(x)$, то

$$H(f;u_0, u_1,\ldots,u_{n-1})=H(f_1;v_0, v_1,\ldots,v_{p-1})+H(f_2;w_0, w_1,\ldots,w_{q-1}),$$

гдето

$$v_i=b_0u_i+b_1u_{i+1}+\ \ldots\ +b_pu_{i+p},$$

$$w_k=\bar{c}_0u_k-\bar{c}_1u_{k+1}+\ldots+(-1)^q\bar{c}_qu_{k+q}.$$

Тези линейни форми са само тогава линейно независими, когато $f_1(x)=0$ и $f_2^*(x)=0$ нямат общ корен.

За $f(x)=a(x-\alpha)$ имаме

$$K(f)=-|a|^2(\alpha+\bar{\alpha}),\ H(f)=-|a|^2(\alpha+\bar{\alpha})u_0\bar{u}_0.$$

Оттук получаваме теоремите:

Необходимо и достатъчно условие, щото $f(x)$ да има само нули с отрицателна реална част, се състои в това, че принадлежащата към $f(x)$ хермитова форма да е положително дефинитна.

Нека m и l да е броят на положителните, респективно отрицателните квадрати в каноничното представяне на $H(f)$. Ако $m+l=n$, то броят на корените с отрицателна реална част е m, броят на тези с положителна реална част е l.

За по-нататъшни подробности гледай цитираните работи.

Глава V

Някои теореми за разпределението на корените в равнината на комплексните числа

1. Теорема на Гаус[1]. Ако всички корени на едно уравнение $f(x)=0$ лежат, от една страна, на една права g равнината на комплексните числа, то корените на уравнението $f'(x)=0$ лежат от същата страна на правата g. Ако не всички корени на $f(x)=0$ са върху g, то само тогава има корени на $f'(x)=0$ върху g, когато те са едновременно и корени на $f(x)=0$.

Ако точката x изписва една права в равнината на числа, то точката $\lambda x+\mu$ изписва пак права, която можем да за реалната ос. Но понеже от $x=ay+b$, ако $f(x)=f(ay+b)=\varphi(y$, имаме

$$af'(x)=\varphi'(y),$$

[1] Отначало дадена от Gauss, обаче останала в неизвестност и отново открита от Lucas, Comptes Rendus, 89, 1879.

ясно е, че теоремата е достатъчно да се докаже, когато g се слива с реалната ос.

Нека $x_k=\alpha_k+i\beta_k$, $k=1,2,\ldots,n$, са корените на $f(x)=0$. Тогава корените на производното, които не са такива и на даденото, ще бъдат корени на

$$\frac{f'(x)}{f(x)}=\frac{1}{x-x_1}+\frac{1}{x-x_2}+\ldots+\frac{1}{x-x_n}=0. \tag{1}$$

Ако поставим $x=\zeta+i\eta$, то като приравним към нула имагинерната част на (1), ще получим

$$\sum_{k=1}^{n}\frac{\eta-\beta_k}{(\zeta-\alpha_k)^2+(\eta-\beta_k)^2}=0. \tag{2}$$

По предположение x_k лежат от едната страна на реалната ос, например отгоре или по нея, но има поне един отгоре, т. е.

$$\beta_k\geqq 0,\quad k=1,\ 2,\ \ldots,\ n,$$

и знакът $>$ е валиден поне за един корен x_k. От равенството (2) е очевидно, че не може $\eta_l\leqq 0$, понеже в такъв случай лявата част е отрицателно число.

От тази теорема веднага се получава същото свойство и за нулите на $f''(x)$, $f'''(x)$ и т. н.

Ако един изпъкнал многоъгълник съдържа всички корени на едно уравнение

$$f(x)=0,$$

то той съдържа и всичките корени на

$$f'(x)=0.$$

Това се доказва веднага, като се приложи теоремата, която доказахме, за страните на многоъгълника. В частност оттук следва, че ако всички корени на $f(x)=0$ са реални, то и тези на $f'(x)=0$ са реални, което е едно следствие и на теоремата на Рол. От това се вижда, че теоремата на Гаус представлява едно обобщение на това следствие от теоремата на Рол.

2. Теорема на Лагер. Нека $f(x)$ е полином от n-та степен с произволни комплексни коефициенти и α е едно произволно число, за което $f(\alpha)\neq 0$, $f'(\alpha)\neq 0$. Тогава във всяка окръжност C и вън от нея, която минава през точките α и

$$\alpha-\frac{nf(\alpha)}{f'(\alpha)},$$

има поне една нула на $f(x)$ или всичките нули на този полином лежат върху C.

Теоремата ще докажем лесно, като използуваме едно следствие от теоремата на Гаус. Именно нека y_s е една произволна нула на про-

изводната $f^{(k)}(x)$, която не е нула на $f(x)$. Нека D е произволна права, минаваща през y_s. Да допуснем, че не всички нули на $f(x)$ лежат по D. Тогава от едната и от другата страна на D трябва да има поне по една нула на полинома $f(x)$, понеже, ако допуснем, че има само от едната страна, по теоремата на Гаус това е възможно само ако y_s е нула и на $f(x)$, което противоречи на условието.

Оттук лесно ще докажем[1] теоремата на Лагер в една по-обща форма. Нека поставим

$$x=\alpha+\frac{1}{t}.$$

Тогава, когато t изписва права D_1, x ще изписва окръжност C, която ще мине през α, понеже на безкрайната точка по $D_1\,(t=\infty)$ отговаря точката $x=\alpha$. Нека

$$\varphi(t)=t^n f\left(\alpha+\frac{1}{t}\right)=\sum_{\mu=0}^{n}\frac{f^{(\mu)}(\alpha)}{\mu!}t^{n-\mu}$$

и t_k е една нула на $\varphi^{(k)}(t)$, k е едно от числата

$$1,\ 2,\ \ldots,\ n-1.$$

Нека правата D_1 минава през t_k. Тогава съответната ѝ окръжност C ще минава през точката

$$x_k=\alpha+\frac{1}{t_k}.$$

Според горната бележка уравнението

$$\varphi(t)=0$$

ще има от двете страни на D_1 поне по един корен или всичките ще лежат по нея. Или понеже на полуравнините от страните на D_1 отговарят за x вътрешността и външността на окръжността C, то ще има значи вътре и вън от C поне по един корен на $f(x)=0$ или всичките му корени ще лежат по C. Ако поставим $k=n-1$, уравнението $\varphi^{(n-1)}(t)=0$ дава

$$t_{n-1}=-\frac{f'(\alpha)}{nf(\alpha)},$$

отгдето

$$x_{n-1}=\alpha-\frac{nf(\alpha)}{f'(\alpha)},$$

с което се доказва теоремата на Лагер. С тази теорема ще установим лесно една друга теорема.

3. Теорема на Феер[2]. Всеки полином от формата

$$a_0+a_1x^{\nu_1}+a_2x^{\nu_2}+\ldots+a_kx^{\nu k}\ (0<\nu_1<\nu_2<\ldots<\nu_k),$$

---

[1] Виж статията ми в Годишника на университета, том XXII, 1926, стр. 1—32.

[2] Mat. Annalen, 87 (1907).

в който $a_0\, a_1, \ldots, a_k$ са произволни комплексни числа, $a_1 \neq 0$[2], има поне една нула ζ, за която

$$|\zeta|^{\nu_1} \leqq \binom{\nu_1+k-1}{k-1} \frac{|a_0|}{|a_1|}.$$

Интересното е в тази теорема, че модулът на най-малката по абсолютна стойност нула остава краен, ако числата a_0, a_1, ν_1, k са крайни, каквито и да бъдат другите коефициенти a и степента ν_k.

Доказателството ще получим от следното следствие на теоремата на Лагер: уравнението $f(x)=0$ от n-та степен има поне един корен, по-малък по абсолютна стойност от всичките корени на уравнението

$$xf'(x)-nf(x)=0.$$

Понеже, ако α е кой да е корен на последното уравнение, по теоремата на Лагер трябва да има поне един корен на $f(x)=0$ в окръжността с диаметър $|\alpha|$ и център $\frac{\alpha}{2}$, понеже

$$\alpha - \frac{nf(\alpha)}{f'(\alpha)} = 0,$$

модулът на който корен е очевидно $\leqq |\alpha|$.

Ако $a_0=0$, то има нула $\zeta=0$ и теоремата на Феер е очевидна. Нека $a_0 \neq 0$. При $k=1$ теоремата е очевидна. Ние ще допуснем, че сме я доказали при $k=p-1$ и ще установим валидността ѝ при $k=p$. Нека

$$f_p(x) = a_0 + a_1 x^{\nu_1} + a_2 x^{\nu_2} + \ldots + a_p x^{\nu_p}$$

и ξ_1 е нула на полинома

$$\nu_p f_p(x) - x f'_p(x) = a_0 \nu_p + a_1 (\nu_p - \nu_1) x^{\nu_1} + \ldots + a_{p-1} (\nu_p - \nu_{p-1}) x^{\nu_{p-1}},$$

която по предположението да удовлетворява на неравенството

$$\zeta_1|^{\nu_1} \leqq \binom{\nu_1+p-2}{p-2} \left| \frac{\nu_p a_0}{(\nu_p-\nu_1) a_1} \right|.$$

По горното следствие на теоремата на Лагер $f_p(x)$ ще има поне една нула ζ, за която

$$|\zeta| \leqq |\xi_1|,$$

т. е.

$$|\zeta|^{\nu_1} \leqq \binom{\nu_1+p-2}{p-2} \left| \frac{a_0}{a_1} \right| \frac{\nu_p}{\nu_p-\nu_1}. \tag{3}$$

Но понеже $\nu_p - \nu_1 \geqq p-1$, то

$$\frac{\nu_p}{\nu_p-\nu_1} = 1 + \frac{\nu_1}{\nu_p-\nu_1} \leqq 1 + \frac{\nu_1}{p-1} = \frac{\nu_1+p-1}{p-1}$$

и още

$$\binom{\nu_1+p-2}{p-2}\frac{\nu_p}{\nu_p-\nu_1}\leqq\binom{\nu_1+p-2}{p-2}\frac{\nu_1+p-1}{p-1}=\binom{\nu_1+p-1}{p-1}.$$

Тогава от (3) следва

$$|\zeta|^{\nu_1}\leqq\binom{\nu_1+p-1}{p-1}\left|\frac{a_0}{a_1}\right|,$$

с което е доказано, че теоремата е валидна за $k=p$. При $\nu_1=1$ получаваме, че полиномът

$$a_0+a_1x+a_2x^{\nu_2}+\ldots+a_kx^{\nu_k}$$

има една нула с модул $\leqq k\left|\frac{a_0}{a_1}\right|$. Тази граница се достига за полинома

$$a_0\left(1+\frac{xa_1}{ka_0}\right)^k=a_0+a_1x+\ldots,$$

който има нулата $-\frac{ka_0}{a_1}$.

4. Теорема на Грейс. Под кръгова област ще разбираме кой да е кръг заедно с ограничаващата го окръжност или външната част от равнината на една окръжност заедно със самата окръжност, или една полуравнина заедно с ограничаващата я права. Ще докажем следната теорема:

Нека числата $x_1, x_2,\ldots,x_n$ са в една кръгова област K. Да означим с S_k, $k=1, 2, 3,\ldots, n$, елементарните симетрични функции на $x_1, x_2,\ldots,x_n$, като $S_0=1$.

Уравнението

$$a_0+\binom{n}{1}a_1x+\binom{n}{2}a_2x^2+\ldots+a_nx^n=0, \tag{4}$$

за което е изпълнено условието

$$a_0S_0+a_1S_1+a_2S_2+\ldots+a_nS_n=0, \tag{5}$$

има поне един корен в K.

При $n=1$ теоремата е очевидна. Приемам, че сме я доказали за $n=p-1$. С помощта на теоремата на Лагер ще покажем, че теоремата е вярна и за $n=p$. Да означим с $f(x)$ полинома

$$f(x)=a_0+a_1\binom{p}{1}x+a_2\binom{p}{2}x^2+\ldots+a_px^p.$$

С лесно пресмятане получаваме

$$xf'(x)-pf(x)=-p\left[a_0+\binom{p-1}{1}a_1x+\ldots+a_{p-1}x^{p-1}\right].$$

За числото x_p от теоремата на Лагер

$$x_p = x - p\frac{f(x)}{f'(x)}$$

ще имаме израза

$$x_p = -\frac{a_0 + \binom{p}{1} a_1 x + \ldots + a_{p-1}x^{p-1}}{a_1 + \binom{p}{1} a_2 x + \ldots + a_p x^{p-1}}, \tag{6}$$

откъдето получаваме

$$(a_0 + a_1 x_p) + \binom{p-1}{1}(a_1 + a_2 x_p)\, x + \ldots + (a_{p-1} + a_p x_p)\, x^{p-1} = 0. \tag{7}$$

Това уравнение с неизвестно x има същата форма като уравнението (4), но $n = p-1$ и коефициентите a_μ са заместени с $a_\mu + x_p a_{\mu+1}$. Лесно се вижда, че релацията (5) е изпълнена за уравнението (7), като в нея се поставят съответните коефициенти и елементарните симетрични функции S'_k на числата $x_1, x_2, \ldots, x_{p-1}$. Действително имаме

$$S_0' = 1,\ S_\mu' + x_p S'_{\mu-1} = S_\mu,\ 1 \leqq \mu \leqq p-1,\ S'_{p-1}\, x_p = S_p,$$

така че

$$(a_0 + a_1 x_p)\, S_0' + (a_1 + a_2 x_p)\, S_1' + \ldots + (a_{p-1} + a_p x_p)\, S'_{p-1} =$$

$$= a_0 S_0' + a_1 (S_1' + x_p S_0') + a_2 (S_2' + x_p S_1') + \ldots + a_p x_p S'_{p-1} =$$

$$= a_0 S_0 + a_1 S_1 + a_2 S_2 + \ldots + a_p S_p = 0.$$

Уравнението (7) ще има поне един корен Z в K. Но от (6) при $x = Z$ получаваме

$$Z - p\frac{f(Z)}{f'(Z)} = x_p$$

и както лесно се вижда, има окъжност C, която минава през точките Z и x_p и лежи в K. По теоремата на Лагер уравнението $f(x) = 0$ ще има поне един корен в C и вън от C, т. е. в K, с което теоремата на Грейс е установена.

На горната теорема ще дадем сега по-друга форма. Нека поставим

$$S_{n-k} = (-1)^{n-k}\frac{b_k}{b_n}\binom{n}{k}.$$

Тогава $x_1, x_2, \ldots, x_n$ ще бъдат корени на уравнението

$$b_0 + \binom{n}{1} b_1 x + \binom{n}{2} b_2 x^2 + \ldots + b_n x^n = 0.$$

Следователно теоремата на Грейс може да се формулира така:

Нека са дадени уравненията

$$(8)\qquad A(x)=a_0+\binom{n}{1}a_1x+\binom{n}{2}a_2x^2+\ldots+a_nx^n=0,$$
$$B(x)=b_0+\binom{n}{1}b_1x+\binom{n}{2}b_1x^2+\ldots+b_nx^n=0$$

и нека между коефициентите им имаме релацията

$$a_0b_n-\binom{n}{1}a_1b_{n-1}+\binom{n}{2}a_2b_{n-2}-\ldots+(-1)^n a_nb_0=0.$$

(Казваме, че $A(x)$ и $B(x)$ са аполарни). Ако всички корени на едното уравнение лежат в една кръгова област K, то има в същата област поне един корен на другото уравнение.

Теоремата на Грейс, както ще видим по-нататък, има големи приложения.

5. Теорема за композиране. Нека $A(x)$ и $B(x)$ са два полинома (8) и ζ е корен на уравнението

$$C(x)=a_0b_0+\binom{n}{1}a_1b_1x+\ldots+a_nb_nx^n=0.$$

Нека $a_0b_0\neq 0$ и $a_nb_n\neq 0$ така, че нулите β_ν на $B(x)$ са отлични от нула. Тогава полиномите $A(x)$ и $x^nB\left(-\frac{\zeta}{x}\right)$ са аполарни. Нулите на втория полином са

$$-\frac{\zeta}{\beta_1},-\frac{\zeta}{\beta_2},\ldots,-\frac{\zeta}{\beta_n}.$$

Ако нулите на $A(x)$ са в една кръгова област K, то поне едно от последните числа по теоремата на Грейс ще принадлежи на K, т. е. има число k от K и нула β_g така, че $\zeta=-\beta_gk$. Това остава в сила и като се освободим от ограничението, че $a_0b_0\neq 0$, $a_nb_n\neq 0$. Защото, ако $a_0b_0=0$, то поне една нула на полинома $A(x)$ или $B(x)$ става равна на нула, а при $a_nb_n=0$ поне една става ∞. Така получихме:

Нека

$$A(x)=a_0+\binom{n}{1}a_1x+\binom{n}{2}a_2x^2+\ldots+a_nx^n=0,$$
$$B(x)=b_0+\binom{n}{1}b_1x+\binom{n}{2}b_2x^2+\ldots+b_nx^n-0$$

са две дадени уравнения. Корените на $A(x)=0$ нека принадлежат на една кръгова област K. Нека $\beta_1,\beta_2,\ldots,\beta_n$ са корените на $B(x)=0$. Тогава всеки корен ζ на композираното уравнение

$$C(x)=a_0b_0+\binom{n}{1}a_1b_1x+\binom{n}{2}a_2b_2x^2+\ldots+a_nb_nx^n=0$$

има формата $\zeta=-b_q k$, където q е едно от числата $1, 2, \ldots, n$ и k е подходящо число от K.

Трябва да се отбележи, че при $a_n=0$, понеже единият корен на $A(x)=0$ отива в ∞, то K съдържа безкрайната точка и при $b_n=0$ между корените β се брои и ∞.

От това предложение следват директно различни интересни предложения.

Така, ако нулите на $A(x)$ са в кръга $|x|\leqq r$, а на $B(x)$ са в кръга $|x|\leqq r_1$, то нулите на $C(x)$ са в кръга $|x|\leqq rr_1$. Същото важи за външността на тези кръгове. Ако комбинираме тези две предложения, получаваме: ако нулите на $A(x)$ са по окръжността $|x|=r$, а тези на $B(x)$ по $|x|=r_1$, то нулите на $C(x)$ са по окръжността $|x|=rr_1$, понеже трябва да бъдат и вътре, и вън или върху тази окръжност.

Благодарение на една бележка на Шур може да се отиде по-нататък по следния начин. Нека нулите на $A(x)$ да принадлежат на една половин равнина H, която съдържа нулевата точка, а нулите β на $B(x)$ да са интервалът $(-1, 0)$. Тогава, понеже числата $-\beta_s k$ пак принадлежат на H, то нулите на $C(x)$ принадлежат на H. Оттук, ако нулите на $A(x)$ принадлежат на една изпъкнала област R, която съдържа нулата, и нулите на $B(x)$ са в интервала $(-1, 0)$, то нулите на $C(x)$ пак принадлежат на R, понеже трябва да лежат във всяка половин равнина, която съдържа R. Следователно имаме теоремата:

Нека

$$A(x)=a_0+\binom{n}{1} a_1 x+\binom{n}{2} a_2 x^2+\ldots+a_n x^n=0,$$

$$B(x)=b_0+\binom{n}{1} b_1 x+\binom{n}{2} b_2 x^2+\ldots+b_n x^n=0,$$

$$C(x)=a_0 b_0+\binom{n}{1} a_1 b_1 x+\binom{n}{2} a_2 b_2 x^2+\ldots+a_n b_n x^n=0$$

са три уравнения. Корените на $A(x)=0$ да лежат в една изпъкнала област R, която съдържа началото, а тези на $B(x)=0$ да принадлежат на интервала $(-1, 0)$. Тогава корените на $C(x)=0$ лежат пак в R.

Във всички тези разглеждания, като казваме във за простота, изразяваме, че се разбира освен вътрешността й и контурът на разглежданата област.

От горното предложение, като сведем R на един реален интервал, получаваме:

Ако корените на $A(x)=0$ лежат в интервала $(-a, a)$, тези на $B(x)=0$ са реални и с еднакъв знак и лежат в интервала $(-b, 0)$ рли $(0, b)$, то корените на $C(x)=0$ принадлежат на интервала $(-ab, ab)$.

Значи, ако полиномът $A(x)$ има само реални нули, полиномът $B(x)$ само реални и с еднакъв знак, то полиномът $C(x)$ има само реални нули.

Нека полиномите

$$f(x)=a_0+a_1x+\ldots+a_kx^k,$$
$$\varphi(x)=b_0+b_1x+\ldots+b_lx^l$$

имат само реални нули, от които вторите са с еднакъв знак. Нека n е произволно цяло число, по-голямо от k и l. Като композираме уравненията

$$x^nf\left(\frac{1}{x}\right)=a_kx^{n-k}+\ldots+a_0x^n=0,$$

$$x^n\varphi\left(\frac{1}{x}\right)=b_l\,x^{n-l}+\ldots+b_0x^n=0,$$

по теоремата на Грейс ще получим ново уравнение

$$a_0b_0x^n+\frac{a_1b_1}{\binom{n}{1}}x^{n-1}+\ldots+\frac{a_mb_m}{\binom{n}{m}}x^{n-m}=0,$$

в което $m=\min(k,\ l)$ само с реални нули. Ако заместим x с $\frac{x}{n}$ и умножим с n^n, получаваме

$$a_0b_0x^n+1!\,a_1\,b_1x^{n-1}+\frac{n}{n-1}\,2!\,a_2b_2x^{n-2}+\ldots+$$

$$+\frac{n}{n-1}\cdot\frac{n}{n-2}\cdots\frac{n}{n-m+1}\,m!\,a_mb_m\,x^{n-m}=0,$$

което ще има само реални корени. Тогава по предложението на стр. 310, когато $n\to\infty$, полиномът

$$a_0b_0x^m+1!\,a_1b_1x^{m-1}+\ldots+m!\,a_mb_m$$

ще има само реални нули. Така със смяна на x с $\frac{1}{x}$ получаваме следната теорема на Шур:

Ако корените на

$$a_0+a_1x+\ldots+a_kx^k=0$$

са реални, а тези на

$$b_0+b_1x+\ldots+b_lx^l=0$$

са реални и с еднакъв знак, то корените на

$$a_0b_0+1!\,a_1b_1x+\ldots+m!\,a_mb_mx^m=0,$$

където $m=\min(k,\ l)$ са също реални.

Полученият полином

$$a_k+a_{k-1}x+\ldots+a_0x^k$$

от полинома $f(x)$ със заместване на x с $\frac{1}{x}$ има пак само реални нули. Понеже полиномът

$$(1+x)^k = 1+\binom{k}{1}x + \ldots + x^k$$

има само реални нули, то по теоремата на Шур

$$a_k + ka_{k-1}\,x + \ldots + k!\,a_0 x$$

има само реални нули. Следователно полиномът

$$\frac{a_k}{k!} + \frac{a_{k-1}}{(k-1)!}\,x + \ldots + a_0 x^k$$

или полиномът

$$a_0 + \frac{a_1}{1!}\,x + \ldots + \frac{a_{k-1}}{(k-1)!}\,x^{k-1} + \frac{a_k}{k!}\,x^k$$

има само реални нули. Това получихме и по-рано от една теорема на Хермит. Като приложим теоремата на Шур върху този полином и $\varphi(x)$, получаваме следната теорема на Мало[1]:

Ако корените на уравнението

$$a_0 + a_1 x + \ldots + a_k x^k = 0$$

са реални, а тези на

$$b_0 + b_1\,x + \ldots + b_l x^l$$

са реални и с еднакъв знак, корените на

$$a_0 b_0 + a_1 b_1\,x + \ldots + a_m\,b_m\,x^n = 0,$$

гдето $m = \min(k, l)$ са пак реални.

6. Теорема на Грейс—Хейвуд. Теоремата на Гаус представлява обобщение на едно следствие от теоремата на Рол. По последната $f'(x)$ има поне един реален корен между две числа a и b, за които $f(x)$ взема равни стойности при предположение, че коефициентите на полинома $f(x)$ са реални. Теоремата на Грейс — Хейвуд представлява обобщение на тази теорема за полиноми с комплексни коефициенти. Тя гласи: Един полином от m-та степен взема за $x=-1$ и $x=+1$ равни стойности. Тогава производната му се анулира в кръга с радиус $\operatorname{ctg}\frac{\pi}{m}$ с център началото ($x=0$).

Доказателството се основава на теоремата на Грейс, на която, както лесно се вижда, може да се даде следната форма.[2]

---

[1] Malo — Journal dê Mathém. spéc. 4, 1895.

[2] Вж. G. Szegö-Mat. Zeitchrifi, 13.

Нека $\beta_0, \beta_1, \ldots, \beta_n$ са дадени числа, които не всички са нули. Нека е дадено уравнението

$$\alpha(x)=\alpha_0+\alpha_1 x+\ldots+\alpha_n x^n=0,$$

за което да имаме

$$L\equiv\alpha_0\beta_n+\alpha_1\beta_{n-1}+\ldots+\alpha_n\beta_0=0.$$

Тогава има поне един корен на $\alpha(x)=0$ във всяка кръгова област, която съдържа всичките корени на

$$\beta(z)=\beta_0-\binom{n}{1}\beta_1 z+\binom{n}{2}\beta_2 z^2+\ldots+(-1)^n\beta_n z^n=0.$$

Оттук веднага се вижда, че полиномът $\beta(z)$ се получава от линеарната форма L, ако за образуването на L вместо коефициентите на $\alpha(x)$ се използуват тези на $(x-z)^n$.

Нека $\alpha(x)$ е производната на полинома $a(x)$ от m-та степен в теоремата на Грейс — Хейвуд. Понеже по условие

$$a(-1)=a(1),$$

то ще имаме

$$\int_{-1}^{+1} a'(x)\,dx=0$$

или понеже $a'(x)=\alpha(x)$, $n=m-1$,

$$\int_{-1}^{+1}\alpha(x)\,dx=2\alpha_0+\frac{2}{3}\alpha_2+\frac{2}{5}\alpha_4+\ldots=0.$$

Това е една линеарна релация между коефициентите на полинома $\alpha(x)$. Съответният полином $\beta(z)$ ще бъде

$$\beta(z)=\int_{-1}^{+1}(x-z)^{m-1}\,dx=\frac{(1-z)^m-(-1-z)^m}{m}.$$

С лесно пресмятане намираме, че нулите на последния полином са

$$z_\nu=i\operatorname{ctg}\frac{\nu\pi}{m}\ (\nu=1, 2, \ldots, m-1),$$

от които най-голямата по абсолютна стойност е z_1. Ако опишем около началото една окръжност с радиус

$$\operatorname{ctg}\frac{\pi}{m},$$

всички корени на $\beta(z)=0$ ще бъдат в нея и тогава по теоремата на Грейс следва, че полиномът $\alpha(x)$ ще има поне една нула в нея. Също

във всеки кръг, който минава през точките $\pm i \operatorname{ctg} \frac{\pi}{m}$, ще има поне една нула на

$$\alpha(x) - a'(x).$$

Също и във външността на всяка окръжност, минаваща през две съседни точки:

$$i \operatorname{ctg} \frac{\nu\pi}{m}, \quad i \operatorname{ctg} \frac{(\nu+1)\pi}{m}.$$

Като казваме във, разбираме и периферията на областта, както вече споменахме по-рано.

7. Някои множители в теорията на алгебричните уравнения. По-рано посредством една теорема на Лагер видяхме, че съществуват числа $\gamma_0, \gamma_1, \ldots, \gamma_n$ така, че щом уравнението

$$a_0 - a_1 x - \ldots - a_n x^n = 0 \tag{9}$$

има само реални корени, то и уравнението

$$a_0 \gamma_0 - a_1 \gamma_1 x - \ldots + a_n \gamma_n x^n = 0 \tag{10}$$

има само реални корени. Тук ще си поставим за цел да намерим всички такива числа γ. Именно една редица от числа

$$\gamma_0, \gamma_1, \gamma_2, \ldots, \gamma_n, \ldots \tag{11}$$

наричаме редица множители от първи вид, когато за всяко уравнение (9) със само реални корени уравнението (10) има пак само реални корени. Редицата множители (11) ще наричаме от втори вид, ако за всяко уравнение (9) с реални и с еднакъв знак корени уравнението (10) има само реални корени. Към корените с еднакъв знак се причисляват и тези, които са равни на нула, ако евентуално съществуват. Условията за редиците са дадени от Шур и Полиа[1]. Ние[2] ще ги изведем, като използуваме теоремата на Грейс.

Да решим отначало въпроса за множителите от първи вид. Понеже уравнението

$$(1-x)^n = 1 + \binom{n}{1} x + \binom{n}{2} x^2 - \ldots + x^n = 0$$

има само реални корени, то уравнението

$$\gamma_0 - \binom{n}{1} \gamma_1 x + \binom{n}{2} \gamma_2 x^2 + \ldots - \gamma_n x^n = 0 \tag{12}$$

ще има само реални корени. По теоремата на Декарт обаче имаме, че след първото поред число $\gamma_s \neq 0$ две последователни числа не могат

---

[1] G. Pólya und I. Schur — Crelles Journal, 134 (1914) стр. 89 — 113.
[2] Вж. N. Obreschkoff — Jahresbericht der Deuts. Math. Vereinigung, 35 стр. 301 — 304.

да се анулират, т. е. ако $\gamma_p=0$, $p>s$, то сигурно е, че $\gamma_{p-1}\gamma_{p+1}\neq 0$. Нека предположим, че $\gamma_p=0$, $p>s$. Тогава γ_{p+1} и γ_{p-1} ще са отлични от нула. Но уравнението $x^{p-1}+2x^p+x^{p+1}=0$ има само реални корени и следователно уравнението $\gamma_{p-1}x^{p-1}+\gamma_{p+1}x^{p+1}=0$ трябва да има само реални корени, т. е. $\gamma_{p-1}\ \gamma_{p+1}<0$. Но понеже уравнението $x^{p-1}-x^{p+1}=0$ има само реални корени, то трябва уравнението $\gamma_{p-1}x^{p-1}-\gamma_{p+1}x^{p+1}=0$ да има само реални корени, т. е. $\gamma_{p-1}\ \gamma_{p+1}>0$. От предните резултати следва, че γ_p не може да бъде равно на нула и $\gamma_{p-1}\ \gamma_{p+1}$ трябва да бъде положително число. Следователно за числата γ_μ ще имаме $\gamma_0=\gamma_1=\ldots=\gamma_{s-1}=0$, като следващите са отлични от нула всичките или с еднакъв знак, или с алтериращи знаци, т. е. уравнението (12) трябва да има само реални корени, като отличните от нула корени са с еднакъв знак. Така получаваме теоремата: *Необходимо и достатъчно условие редицата* (11) *да бъде от първи тип е уравненията*

$$\gamma_0+\binom{n}{1}\gamma_1 x+\binom{n}{2}\gamma_2 x^2+\ldots+\binom{n}{n}\gamma_n x^n=0,\quad n=1,\ 2,\ 3,\ldots$$

да имат само реални корени и с еднакъв знак.

Естествено касае се за отличните от нула корени. Нека сега редицата (11) е от втори вид. Тогава следва също, че (12) има само реални корени. Обратно, ако (12) има само реални корени и (9) има само реални корени с еднакъв знак, то по теоремата на Грейс уравнението (10) ще има само реални корени, т. е. редицата множители (11) е от втори вид. *Необходимо и достатъчно условие редицата* (11) *да е от втори вид е уравненията*

$$\gamma_0+\binom{n}{1}\gamma_1 x+\binom{n}{2}\gamma_2 x^2+\ldots+\binom{n}{n}\gamma_n x^n=0,\quad n=2,\ 3,\ldots$$

да имат само реални корени.

8. Други теореми. Теоремата на Будан-Фурие е обобщена от автора[1] на този учебник, а именно.

Нека $f(x)=0$ е уравнение от n-та степен с реални коефициенти и нека V_x означава броя на вариациите на редицата

$$f(x),\ f'(x),\ f''(x),\ldots,f^{(n)}(x).$$

Тогава броят на корените му, които се намират в един симетричен спрямо реалната ос четириъгълник в равнината на комплексното променливо, чиито два противоположни върха са a и b $(a<b)$ със съответни при тях ъгли $\frac{2\pi}{n-V_a}$, $\frac{2\pi}{V_b}$, е или равен на V_a-V_b, или е с четно число по-малък.

[1] Н. Обрешков, Годишник на Соф. университет (1921) и (1927), стр. 177—200 и в Jahresbericht der Deutsch. Math. Vereinigung, 33 (1924), стр. 52—64.

Доказателството[1] се основава на следните, установени от автора предложения:

Нека $f(x)$ е полином от n-та степен с реални коефициенти, броят на вариациите на които е равен на V. Тогава броят на вариациите на коефициентите на полинома

$$f(x)(x^2-2\rho x\cos\varphi+\rho^2),\ \rho>0;\quad 0\leqq\varphi<\frac{\pi}{n+2-V}$$

е равен на $V+2$ или е с четно число по-голям.

Нека $f(x)$ е полином с реални коефициенти, броят на вариациите на които е равен на V. Тогава броят на вариациите на полинома

$$f(x)(x^2+2\rho x\cos\varphi+\rho^2),\quad \rho>0,\quad 0\leqq\varphi<\frac{\pi}{V+2},$$

е равен на V или е с четно число по-малък.

В частния случай при $a=0$ и $b=\infty$ получаваме следното обобщение на теоремата на Декарт:

Нека $f(x)=0$ е уравнение от n-та степен с реални коефициенти, броят на вариациите на които е V. Тогава броят на корените му, на които аргументите се намират между $-\frac{\pi}{n-V}$ и $\frac{\pi}{n-V}$ (границите изключени), е равен на V или с четно число по-малък.

Ще разгледаме още някои теореми за нулите на полиномите. От теоремата на Грейс се получава лесно следното предложение:

Нека n-те комплексни числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ принадлежат на една кръгова област K. Тогава, каквото и да е комплексното число x, съществува едно число α, принадлежащо на K, което зависи от x, но такова, че

$$(1)\qquad (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n)=(x-\alpha)^n.$$

Действително, ако с A означим израза

$$A=(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

то за елементарните симетрични функции $s_1, s_2, \ldots, s_n$ на $\alpha_1, \alpha_2, \ldots, \alpha_n$ ще имаме

$$x^n-A-s_1x^{n-1}+s_2x^{n-2}-\ldots+(-1)^n s_n=0.$$

По теоремата на Грейс уравнението

$$(2)\qquad x^n-A-\binom{n}{1}zx^{n-1}+\binom{n}{2}z^2x^{n-2}-\ldots+(-1)^n z^n=0$$

ще има поне един корен $z=\alpha$ в K. Но (2) при $z=\alpha$ се обръща в (1) и предложението е установено.

---

[1] За по-големи подробности и доказателство виж Н. Обрешков. Задачи и теореми по Висша алгебра, 1960, задачи 51, 52, 53, 54, 55, 57, стр. 58 — 60 и стр. 243—248 и 268—270.

Ще разгледаме едно предложение на доказаното предложение. Нека

$$P(x) = a_0 + a_1 x + \ldots + a_n x^n, \quad a_n \neq 0,$$

е произволен полином, нулите на който лежат в една кръгова област K. Въз основа на предното уравнението

(3) $$P(x) - \lambda = 0$$

може да се напише така:

$$(x - \alpha)^n - \frac{\lambda}{a_n} = 0,$$

гдето α е точка от K. За корените на последното уравнение получаваме

$$x = \alpha + \sqrt[n]{\frac{\lambda}{a_n}}.$$

Следователно корените на уравнението (3) лежат в област, получена от K с транслациите $\sqrt[n]{\frac{\lambda}{a_n}}$.

Теоремата на Гаус—Люка е обобщена от Уолш за нулите на производните на произведение от два полинома, нулите на които лежат в дадени кръгови области. Нека полиномът $f(x)$ от степен n е произведение на двата полинома $f_1(x)$ и $f_2(x)$ от степени p и q $(p + q = n)$. като нулите $\alpha_1, \alpha_2, \ldots, \alpha_p$ на $f_1(x)$ лежат в кръговата област K_1 и нулите $\beta_1, \beta_2, \ldots, \beta_q$ на $f_2(x)$ лежат в кръгова област K_2. Нека x_0 е една нула на $f^{(m)}(x)$. Релацията

$$f^{(m)}(x_0) = 0,$$

в която разглеждаме x_0 и β_j като фиксирани, е симетрична спрямо числата α_s и линейна спрямо тях. Ако заместим всичките α_s с α, то $f(x)$ до постоянен множител се обръща в полинома

$$h(x) = (x - \alpha)^p f_2(x).$$

Следователно по теоремата на Грейс ще има поне една точка α от областта K_1 такава, че m-та производна на $h(x)$ ще се анулира за $x = x_0$, т. е. $h^{(m)}(x_0) = 0$. Ако сега приложим същото разсъждение за последната релация, като оставим α и x_0 фиксирани, получаваме, че съществува точка β от K_2 такава, че m-та производна на полинома

$$\varphi(x) = (x - \alpha)^p (x - \beta)^q$$

се анулира за $x = x_0$. Така получаваме следната теорема на Уолш:

Ако означим с Δ областта, описана от нулите на полинома $\varphi^{(m)}(x)$, когато α и β описват съответно K_1 и K_2, независимо едно от друго, нулите на полинома $f^{(m)}(x)$ ще лежат в Δ.

Лесно можем да получим областта Δ. Ако α и β са произволни комплексни числа, то нулите на полинома $\varphi^{(m)}(x)$ ще лежат на отсечката, която съединява α с β. В това се убеждаваме лесно по теоремата на Гаус — Люка или по теоремата на Рол, като с подходяща линейна трансформация докараме съответните точки на α и β да лежат по реалната ос. По нея нулите на $\varphi^{(m)}(x)$, които са отлични от α и β, са прости. Нека $\lambda_1, \lambda_2, \ldots, \lambda_k$ са отношенията, в които тези нули на същия полином разделят отсечката $\alpha \ldots \beta$. С трансформацията

$$x = \alpha + (\beta - \alpha)\, t$$

лесно се убеждаваме, че числата λ_s са нулите на m-та производна на полинома $t^p (1-t)^q$, които са отлични от нула и единица. Нека γ_1 и γ_2 са центровете на окръжностите C_1 и C_2, като K_1 и K_2 означават в случая кръговете, заградени от тези окръжности, и r_1 и r_2 са радиусите им. Ще установим, че Δ е област, състояща се от всичките кръгове $K^{(s)}$ с центрове

$$\frac{\gamma_1 + \lambda_s \gamma_2}{1 - \lambda_s}$$

и радиуси

$$\frac{r_1 + \lambda_s r_2}{1 + \lambda}, \quad s = 1, 2, \ldots, k.$$

Действително нулите на $\varphi^{(m)}(x)$ са

$$x_s = \frac{\alpha + x_s \beta}{1 + \lambda_s}.$$

Ако α е в K_1 и β в K_2, то за x ще имаме

$$x_s - \frac{\gamma_1 + \lambda_s \gamma_2}{1 + \lambda_s} = \frac{\alpha - \gamma_1}{1 + \lambda_s} + \frac{\lambda_s}{1 + \lambda_s} (\beta - \gamma_2),$$

откъдето получаваме

$$\left| x_s - \frac{\gamma_1 + \lambda_s \gamma_2}{1 + \lambda_s} \right| \leqq \frac{r_1 + \lambda_s r_2}{1 + \lambda_s},$$

т. е. x_s ще лежи в $K^{(s)}$. Лесно можем да видим, че всяка точка от Δ може да се получи по горния начин, но това не е от значение в случая.

Ще установим сега един друг резултат по същество. Нека

$$f(x) = a_0 + a_1 x + \ldots + a_{p-1} x^{p-1} + \ldots + x^n$$

е полином с произволни комплексни коефициенти, като коефициентът пред x^n е равен на 1. Да означим с M_p най-голямото от числата $|a_0|, |a_1|, \ldots, |a_{p-1}|$ и с r — число, по-голямо от 1. Очевидно или полиномът $f(x)$ ще има p нули, на които модулите са по-малки от r, или поне $n+1-p$ нули с модули $\geqq r$. Да предположим, че имаме втория

случай и да разделим $f(x)$ с $\alpha - x$, като α е нула на $f(x)$ с модул $\geqq r$. За коефициентите на частното

$$f_1(x) = a'_0 + a'_1 x + \ldots a'_{p-1} x^{p-1} + \ldots + x^{n-1}$$

получаваме

$$a'_0 = \frac{a_0}{\alpha},$$

$$a'_1 = \frac{a_0}{\alpha^2} + \frac{a_1}{\alpha},$$

. . . . . .

$$a'_k = \frac{a_0}{\alpha^{k+1}} + \frac{a_1}{\alpha^k} + \ldots + \frac{a_k}{\alpha}.$$

. . . . . .

Ако $k \leqq p-1$, то ще имаме

$$|a'_k| < M_p\left(\frac{1}{r} + \frac{1}{r^2} + \ldots + \frac{1}{r^{k+1}} + \ldots\right) = \frac{M_p}{r-1}.$$

Следователно първите p коефициента на $f_1(x)$ имат модули, по-малки от $\frac{M_p}{r-1}$. Едно второ деление, отговарящо на друга нула на $f(x)$ с модул $\geqq r$, дава полином $f_2(x)$, на който първите p коефициента ще имат модули по-малки от $\frac{M_p}{(r-1)^2}$, и. т. н.; след $n-p$ деления ще получим полином $f_{n-p}(x)$ от степен p, на който модулите на първите p коефициента ще бъдат по-малки от $\frac{M_p}{(r-1)^{n-p}}$ и последният коефициент ще е равен на ± 1. По теоремата на стр. 11 модулите на корените на уравнението

$$f_{n-p}(x) = 0$$

ще бъдат по-малки от

$$1 + \frac{M_p}{(r-1)^{n-p}}.$$

Като приравним с r това число и решим полученото уравнение спрямо r, ще имаме

$$r = 1 + \sqrt[q]{M_p}, \quad q = n - p + 1.$$

Така доказахме следната теорема на Монтел:

Полиномът $f(x)$ има поне p нули, на които модулите са по-малки от $1 + \sqrt[q]{M_p}$.

Ако $M_p = 1$, тази граница е равна на 2.

Предната теорема показва, че полиномът $f(x)$ при фиксирани първи p коефициента има поне p нули в един фиксиран кръг около началото, каквито и да бъдат останалите коефициенти и степента n, като се предполага, че коефициентът пред x^n е равен на 1.

ЧАСТ VI

АЛГЕБРИЧЕСКО РЕШЕНИЕ НА УРАВНЕНИЯТА

Глава I

Алгебрическо решение на уравненията от трета и четвърта степен

1. **Уравнения от трета степен.** Под алгебрическо решение на уравнението

$$f(x)=a_0 x^n+a_1 x^{n-1}+\ldots+a_n=0$$

се разбира определянето на корените му посредством краен брой рационални действия и извличане на корени от коефициентите на уравнението.

Както видяхме при изучаване на трансформацията на Чирнхаус, уравненията от трета и четвърта степен са решими алгебрически. Тук ще вземем други директни и по-прости методи.

Непълното кубическо уравнение

$$x^3=A$$

със субституцията $x=y\sqrt[3]{A}$, гдето $\sqrt[3]{A}$ е едно от трите значения на радикала (например, ако A е реално, може да се вземе реалното значение), се свежда на

$$y^3=1.$$

Последното уравнение се разпада на

$$y-1=0, \quad y^2+y+1=0.$$

Корените на второто са

$$\alpha=\frac{-1+i\sqrt{3}}{2}, \quad \beta=\alpha^2=\frac{-1-i\sqrt{3}}{2},$$

така че корените на $x^3=A$ са $\sqrt[3]{A}$, $\alpha\sqrt[3]{A}$, $\alpha^2\sqrt[3]{A}$.

Общото уравнение от трета степен

$$a_0x^3+a_1x^2+a_2x+a_3=0$$

със субституцията $y=x+\frac{a_1}{3a_0}$ се свежда на друго, в което липсва вторият член. Така можем да предполагаме, че даденото кубично уравнение има вида

$$x^3+px+q=0. \tag{1}$$

По метода на Худе[1] това уравнение се решава така: полагаме

$$x = y + z,$$

гдето y и z са нови неизвестни; ще имаме

$$x^3 = y^3 + z^3 + 3yz(y + z)$$

или

$$x^3 - 3zyx - (y^3 + z^3) = 0. \tag{2}$$

Като сравним коефициентите на (1) и (2), получаваме

$$yz = -\frac{p}{3}, \tag{3}$$

$$y^3 + z^3 = -q$$

или

$$y^3 + z^3 = -q,$$

$$y^3 z^3 = -\frac{p^3}{27}.$$

Ако поставим $y^3 = u_1$, $z^3 = u_2$, то последните равенства ни показват, че u_1 и u_2 са корени на уравнението

$$u^2 + qu - \frac{p^3}{27} = 0,$$

$$u_{1,2} = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}},$$

откъдето

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}, \tag{4}$$

която формула се нарича формула на Cardano.

2. Разискване на формулата. Като вземем под внимание, че кубическият корен има три значения, като комбинираме трите значения първия радикал с трите на втория, получаваме девет стойности на x. Но лесно е да видим, че само три от тях са корените на уравнение (1). От (3) виждаме, че радикалите y и z трябва а удовлетворяват на условието $yz = -\frac{p}{3}$. Нека A и B са две стойности на

---

[1] Пръв е решил кубичното уравнение италианският математик Scipione del Ferro в 1515 г. След това решението е било отново открито също от италианските математици Nikolaus Tartaglia в 1535 г. и Geronimo Cardano в 1545 г. По това време Ludovico Ferrari, ученик на Cardano, намира решението на уравнението от четвърта степен. Методът на Hudde е даден в 1639 г. Лобачевски е дал друго решение на уравнението (1) със заместването $x = y - \frac{p}{3y}$.

тях, удовлетворяващи на условието $AB=-\frac{p}{3}$. Значенията на първия радикал са A, αA, $\alpha^2 A$, а на втория: B, αB, $\alpha^2 B$. Очевидно A може да се комбинира само с B, понеже с αB бихме имали $A.\alpha B=-\frac{p}{3}\alpha$, което не е равно на $-\frac{p}{3}$. Също не може да се комбинира с $\alpha^2 B$. С подобни разглеждания става ясно, че трите корена ще бъдат

$$x_1=A+B,$$

$$x^2=\alpha A+\alpha^2 B,$$

$$x_3=\alpha^2 A+\alpha B$$

или като заместим стойностите на α и α^2,

$$x_1=A+B,$$

$$x_2=-\frac{A+B}{2}+i\,\frac{A-B}{2}\sqrt{3}, \tag{5}$$

$$x_3=-\frac{A+B}{2}-i\,\frac{A-B}{2}\sqrt{3}.$$

От тези формули лесно е да изследваме реалността на корените, когато p и q са реални. Ако означим с

$$\Delta=\frac{q^2}{4}+\frac{p^3}{27},$$

то ще имаме три случая.

1) $\Delta>0$. Тогава можем да вземем за A и B реалните им стойности. От (5) се вижда, че уравнението (1) има два имагинерни и един реален корен. Имагинерните корени не могат да бъдат равни помежду си, понеже, ако допуснем, че $x_2=x_3$, то трябва $A=B$, отгдето $A^3=B^3$, което води до $\Delta=0$.

2) $\Delta=0$. Тогава получаваме

$$A=B=\sqrt[3]{-\frac{q}{2}},$$

$$x_1=2\sqrt[3]{-\frac{q}{2}},$$

$$x_2=x_3=-\sqrt[3]{-\frac{q}{2}}.$$

Трите корена са реални, два от които са равни помежду си.

3) $\Delta < 0$. В този случай A и B са имагинерни и конюговани помежду си, т. е.

$$A = g + hi, \quad B = g - hi,$$

$$x_1 = 2g,$$

$$x_2 = -g - h\sqrt{3},$$

$$x_3 = -g + h\sqrt{3}.$$

Трите корена са реални. Формулата на Кардано дава обаче корените в имагинерна форма. Този случай се нарича неразложим (Casus irreducibilis). Доказва[1] се, че не може да се реши уравнението с реални радикали, ако е неразложимо.

Ще дадем тригонометричното решение на уравнението. Радикалът $A = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}}$ може да се пише $\sqrt[3]{-\frac{q}{2} + i\sqrt{-\Delta}}$. Понеже $\Delta < 0$, числото $-\frac{q}{2} + i\sqrt{-\Delta}$ е комплексно и му придаваме тригонометрична форма:

$$-\frac{q}{2} + i\sqrt{-\Delta} = \rho(\cos\varphi + i\sin\varphi),$$

отгдето

$$-\frac{q}{2} = \rho\cos\varphi, \quad \sqrt{-\Delta} = \rho\sin\varphi,$$

$$\rho = \sqrt{\frac{q^2}{4} - \Delta} = \sqrt{-\frac{p^3}{27}}, \quad \cos\varphi = -\frac{q}{2\sqrt{-\frac{p^3}{27}}}.$$

Тогава ще получим

$$x = y + z = \sqrt[3]{\rho(\cos\varphi + i\sin\varphi)} + \sqrt[3]{\rho(\cos\varphi - i\sin\varphi)} =$$

$$= \sqrt[3]{\rho}\left(\cos\frac{\varphi + 2k\pi}{3} + i\sin\frac{\varphi + 2k\pi}{3}\right) +$$

$$+ \sqrt[3]{\rho}\left(\cos\frac{\varphi + 2k_1\pi}{3} - i\sin\frac{\varphi + 2k_1\pi}{3}\right),$$

гдето k и k_1 вземат стойности 0, 1, 2. Произведението на y и z е равно на $-\frac{p}{3}$ по условие, а от формулата получаваме за това произведение стойността

$$-\frac{p}{3}\left(\cos\frac{2(k-k_1)\pi}{3} + i\sin\frac{2(k-k_1)\pi}{3}\right).$$

---

[1] Hölder — Mathematische Annalen, 38, 1891, стр. 307.

Следователно $k=k_1$. Така че трите корена на (1) при $\Delta<0$ ще бъдат дадени с

$$2\sqrt{-\frac{p}{3}}\cos\frac{\varphi+2k\pi}{3},\quad k=0,\ 1,\ 2.$$

Кардановата формула дава често корена в една сложна форма. Така за уравнението

(6) $$x^3+3x-14=0$$

получаваме

$$A=\sqrt[3]{7+\sqrt{50}},\ B=\sqrt[3]{7-\sqrt{50}}.$$

Но

$$\sqrt[3]{7\pm\sqrt{50}}=1\pm\sqrt{2},$$

така че уравнението (6) има рационален корен 2 и останалите корени са $-1\pm i\sqrt{6}$.

3. Решение на уравненията от четвърта степен. Общото уравнение от четвърта степен

$$f(x)=a_0x^4+a_1x^3+a_2x^2+a_3x+a_4=0$$

с трансформацията $x=y-\frac{a_1}{4a_0}$ може да се сведе в друго, в което липсва вторият член. Така че уравненията от четвърта степен, с които ще се занимаваме, са от вида

(7) $$x^4+px^2+qx+r=0.$$

Ще изложим първо метода на Ойлер, който представлява едно разширение на метода на Худе. Въвеждаме три нови неизвестни

$$x=y+z+t.$$

При това ще имаме

$$x^2=y^2+z^2+t^2+2(yz+yt+zt),$$

$$x^4=(y^2+z^2+t^2)^2+4(y^2+z^2+t^2)\ (yz+yt+zt)+$$

$$+4(y^2z^2+y^2t^2+z^2t^2)+8yzt(y+z+t).$$

Като елиминираме от тези уравнения израза $yz+yt+zt$ и вземем под внимание, че $x=y+z+t$, получаваме

$$x^4-2(y^2+z^2+t^2)x^2-8yztx+(y^2+z^2+t^2)^2-$$

$$-4(y^2z^2+y^2t^2+z^2t^2)=0.$$

Като приравним с коефициентите на уравнението (7), ще получим

$$y^2+z^2+t^2=-\frac{p}{2},$$

$$yzt=-\frac{q}{8}, \tag{8}$$

$$y^2z^2+y^2t^2+z^2t^2=\frac{p^2-4r}{16}.$$

Ако заместим второто уравнение с $y^2z^2t^2=\frac{q^2}{65}$, то от (8) става ясно, че числата y^2, z^2, t^2 са корени на кубичното уравнение, наречено **резолвентно**:

$$u^3+\frac{p}{2}\,u^2+\frac{p^2-4r}{16}\,u-\frac{q^2}{64}=0. \tag{9}$$

Ако корените му са u_1, u_2, u_3, то

$$u_1=y^2,\ u_2=z^2,\ u_3=t^2,$$

$$x=\sqrt{u_1}+\sqrt{u_2}+\sqrt{u_3}.$$

Тази формула дава 8 значения на x, а даденото уравнение е от четвърта степен. Обаче от (8) се вижда, че знаците на радикалите трябва така да бъдат подбрани, че

$$(\pm\sqrt{u_1})\ (\pm\sqrt{u_2})\ (\pm\sqrt{u_3})=-\frac{q}{8}.$$

4. Разискване на решението. От (9) се вижда, че това резолвентно уравнение има винаги поне един положителен корен. Действително при $u=0$ лявата част взема отрицателната стойност $-\frac{q^2}{64}$, а при $u=\infty$ — взема положителната стойност $+\infty$, значи тази част ще се анулира поне един път, когато u се мени от 0 до ∞. Това може другояче да се получи от факта, че произведението $u_1u_2u_3$ е равно на $\frac{q^2}{64}$, т. е. е положително. Тогава, ако трите корена са реални, очевидно е, че единият е поне положителен. Ако има два имагинерни, то те са конюговани, т. е. тяхното произведение е положително, и следователно другият корен, който трябва да е реален, е положителен. Нека този положителен корен е u_1. Ще имаме три случая: а) u_1, u_2, u_3 телни; $\sqrt{u_1}$, $\sqrt{u_2}$, $\sqrt{u_3}$ са реални. Тогава за корените на даденото уравнение имаме схемата (A) или (B):

$$(A)\qquad \begin{cases} x_1 = +\sqrt{u_1}+\sqrt{u_2}-\sqrt{u_3}, \\ x_2 = +\sqrt{u_1}-\sqrt{u_2}+\sqrt{u_3}, \\ x_3 = -\sqrt{u_1}+\sqrt{u_2}+\sqrt{u_3}, \\ x_4 = -\sqrt{u_1}-\sqrt{u_2}-\sqrt{u_3}, \end{cases}$$

$$(B)\qquad \begin{cases} x_1 = +\sqrt{u_1}-\sqrt{u_2}-\sqrt{u_3}, \\ x_2 = -\sqrt{u_1}+\sqrt{u_2}-\sqrt{u_3}, \\ x_3 = -\sqrt{u_1}-\sqrt{u_2}+\sqrt{u_3}, \\ x_4 = +\sqrt{u_1}+\sqrt{u_2}+\sqrt{u_3} \end{cases}$$

според това, дали q е положително или отрицателно.

Всички корени на (7) са реални.

b) $u_1 > 0$, u_2, $u_3 < 0$. Следователно $\sqrt{u_2} = hi$, $\sqrt{u_3} = ki$, $h,k > 0$. Корените на уравнението (7) са имагинерни. В частност, когато $k = h$, т. е. $u_2 = u_3$, имаме един двоен реален корен и два имагинерни.

В този случай $(+\sqrt{u_1})\ (+\sqrt{u_2})\ (+\sqrt{u_3}) = -hk\sqrt{u_1} < 0$. Следователно за корените на даденото уравнение ще имаме схемата (B), ако $q > 0$, и схемата (A), ако $q < 0$.

c) $u_1 > 0$, u_2 и u_3 конюгованн имагинерни. Тогава и числата $\sqrt{u_2}$, $\sqrt{u_3}$ са конюговани имагинерни, следователно са от форма $\alpha+\beta i$, $\alpha-\beta i$. Даденото уравнение (7) има два имагинерни и два реални корена. Понеже произведението

$$(+\sqrt{u_1})\ (+\sqrt{u_2})\ (+\sqrt{u_3}) = (\alpha^2+\beta^2)\sqrt{u_1} > 0,$$

то имаме за корените същите случаи, както в a).

От схемата (A) или (B) лесно се вижда, че корените на резолвентното уравнение са цели рационални функции от корените на даденото уравнение, които са трите стойности на една от тях. Именно имаме

$$u_1 = \frac{1}{16}(x_1+x_2-x_3-x_4)^2,$$

$$u_2 = \frac{1}{16}(x_1+x_3-x_2-x_4)^2,$$

$$u_3 = \frac{1}{16}(x_2+x_3-x_1-x_4)^2,$$

които са трите стойности на функцията

$$\frac{1}{16}(x_1+x_2-x_3-x_4)^2,$$

като разместваме корените по всевъзможни начини.

5. Метод на Декарт. Този метод се състои в това, че лявата част на уравнението разлагаме на два квадратни тричлена, коефициентите на които определяме посредством едно кубично уравнение.

Даденото уравнение

$$x^4+px^2+qx+r=0 \tag{10}$$

представяме така:

$$(x^2+ux+v)\ (x^2-ux+v_1)=0;$$

тогава с приравняване на коефициентите имаме

$$v+v_1-u^2=p,$$

$$u\ (v_1-v)=q, \tag{11}$$

$$vv_1=r.$$

Оттук получаваме

$$v+v_1=p+u^2,$$

$$v-v_1=-\frac{q}{u},$$

откъдето

$$v=\frac{1}{2}\left(p+u^2-\frac{q}{u}\right);\ v_1=\frac{1}{2}\left(p+u^2+\frac{q}{u}\right).$$

Като заместим в третото уравнение (11), получаваме

$$\left(p+u^2+\frac{q}{u}\right)\ \left(p+u^2-\frac{q}{u}\right)=4r$$

или

$$u^6+2pu^4+(p^2-4r)\,u^2-q^2=0.$$

Ако положим $u^2=y$, ще имаме

$$y^3+2py^2+(p^2-4r)\,y-q^2=0, \tag{12}$$

което е резолвентното уравнение:

Ако решим уравнението (12), то даденото се разпада на две квадратни уравнения:

$$x^2+ux+\frac{1}{2}\left(p+u^2-\frac{q}{u}\right)=0,$$

(13)

$$x^2-ux+\frac{1}{2}\left(p+u^2+\frac{q}{u}\right)=0.$$

6. Разискване на решението. Нека α_1^2, α_2^2, α_3^3 са корените на (12), тогава

$$\alpha_1^2\ \alpha_2^2\ \alpha_3^2=q^2.$$

Ако извлечем квадратен корен, то понеже уравненията (13) остават същи, като сменим u с $-u$, можем така да подберем знаците, че

$$\alpha_1\ \alpha_2\ \alpha_3=q. \tag{14}$$

Като поставим $u=\alpha_1$ в уравненията (13) и решим отначало първото уравнение, получаваме

$$x_{1,2}=-\frac{\alpha_1}{2}+\frac{1}{2}\sqrt{\alpha_1^2-2p-2\alpha_1^2+\frac{2q}{\alpha_1}}$$

или понеже от (12)

$$\alpha_1^2+\alpha_2^2+\alpha_3^2=-2p,$$

(15) $$x_{1,2}=-\frac{\alpha_1}{2}\pm\frac{1}{2}\sqrt{\alpha_2^2+\alpha_3^2+2\alpha_2\alpha_3}=\frac{-\alpha_1\mp(\alpha_2+\alpha_3)}{2}.$$

Второто уравнение (13) дава по аналогичен начин

(16) $$x_{3,4}=\frac{\alpha_1\pm(\alpha_2-\alpha_3)}{2}.$$

Така получаваме корените x_1, x_2, x_3, x_4 на даденото уравнение.

Да изследваме реалността на корените x_i според реалността на корените на резолвентното уравнение. Понеже свободният член на (12) е отрицателен, то, както по-рано се убеждаваме, че това уравнение ще има винаги поне един положителен корен, който нека означим с $\alpha_1^2=y_1$.

1) Трите корена y_1, y_2, y_3 на резолвентното уравнение са положителни. Тогава, понеже $y_1=\alpha_1^2$, $y_2=\alpha_2^2$, $y_3=\alpha_3^2$, числата α_1, α_2, α_3 са реални и от това следва реалността на корените на даденото уравнение.

2) $y_1>0$, $y_2<0$, $y_3<0$. Тогава за α_2, α_3 можем да пишем

$$\alpha_2=hi,\quad \alpha_3=ki,$$

гдето $-\alpha_1 hk=q$. За корените имаме

$$x_{1,2}=-\frac{\alpha_1}{2}\pm\frac{i}{2}(h+k),$$

$$x_{3,4}=\frac{\alpha_1}{2}\pm\frac{i}{2}(h-k);$$

т. е. всичките са имагинерни. Само в частния случай, когато $h=\pm k$, т. е. $y_2=y_3$, даденото уравнение има реален двоен корен.

2) $y_1>0$, y_2 и y_3 конюговани имагинерни. Тогава

$$\alpha_2=\alpha+i\beta,$$

$$\alpha_3=\alpha-i\beta$$

и понеже $\alpha_1\alpha_2\alpha_3=\alpha_1(\alpha^2+\beta^2)=q$, вземаме $\alpha_1>0$, ако $q>0$, и $\alpha_1<0$, ако $q<0$.

За корените на даденото уравнение получаваме

$$x_{1,2}=-\frac{\alpha_1}{2}\pm\alpha,$$

$$x_{3,4}=\frac{\alpha_1}{2}\pm i\beta,$$

т. е. то има два реални и два имагинерни.

Можем, както при метода на Ойлер, да изразим корените на резолвентното уравнение чрез корените на даденото. От (15) и (16) получаваме лесно

$$\alpha_1 = \frac{1}{2}\ (-x_1 - x_2 + x_3 + x_4),$$

$$\alpha_2 = \frac{1}{2}\ (x_1 - x_2 + x_3 - x_4),$$

$$\alpha_3 = \frac{1}{2}\ (x_1 - x_2 - x_3 + x_4).$$

Следователно корените y_1, y_2, y_3 ще представляват трите стойности, които взема рационалната функция

$$\frac{1}{4}\ (x_1 + x_2 - x_3 - x_4)^2,$$

когато разместваме x_1, x_2, x_3, x_4 помежду им по всевъзможни начини.

7. Общ метод на Лагранж.[1] Лагранж е изследвал решенията, които са били давани по-рано, и намерил, че те могат да се сведат към един общ метод. Именно нека е дадено уравнението

$$f(x) = 0 \tag{17}$$

с корени $x_1, x_2, \ldots, x_n$. За да го решим, търсим една такава рационална функция

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

от корените му, която взема по-малко различни стойности $y_1, y_2, \ldots, y_m$ от степента му, когато разместваме величините x помежду им по всевъзможни начини. Нека уравнението, на което корените са $y_1, y_2, \ldots, y_m$, да бъде

$$y^m + A_1 y^{m-1} + A_2 y^{m-2} + \ldots + A_m = 0.$$

От релациите

$$\begin{aligned} -A_1 &= y_1 + y_2 + \ldots + y_m, \\ A_2 &= y_1 y_2 + \ldots + y_{m-1} y_m, \\ &\ldots\ldots\ldots\ldots \\ (-1)^m A_m &= y_1 y_2 \ldots y_m \end{aligned}$$

виждаме, че коефициентите му A са симетрични рационални функции на корените (17), следователно могат да се пресметнат. Така решението на даденото уравнение ще се сведе към решение на уравнение от по-ниска степен, стига да съществува една такава функция φ.

Нека отначало разгледаме уравнението от трета степен

$$x^3 + px + q = 0 \tag{18}$$

---

[1] Даден от Lagrange в съчинението му „Réflexions sur la résolution algébrique des équations" (1770 и 1771).

с корени x_1, x_2, x_3. Да разгледаме линейната функция

$$y_1 = x_1 + \alpha x_2 + \alpha^2 x_3, \tag{19}$$

гдето α е един имагинерен корен на $x^3 = 1$. Ако в (19) извършим една кръгова субституция, т. е. заместим x_1 с x_2, x_2 с x_3, x_3 с x_1, получаваме

$$x_2 + \alpha x_3 + \alpha^2 x_1 = \alpha^2 (x_1 + \alpha x_2 + \alpha^2 x_3) = \alpha^2 y_1.$$

Същата субституция, приложена върху получения резултат, дава

$$x_3 + \alpha x_1 + \alpha^2 x_2 = \alpha (x_1 + \alpha x_2 + \alpha^2 x_3) = \alpha y_1.$$

Ако сега върху y_1 извършим транспозицията $(x_1 x_2)$, то получаваме новата стойност

$$y_2 = x_2 + \alpha x_1 + \alpha^2 x_3,$$

от която при извършване на кръгова субституция, заместваща x_1, x_2, x_3 съответно с x_3, x_1, x_2, то получаваме нова стойност

$$x_1 + \alpha x_3 + \alpha^2 x_2 = \alpha^2 y_2$$

и при ново прилагане на същата субституция

$$x_3 + \alpha x_2 + \alpha^2 x_1 = \alpha y_2.$$

Понеже от три елемента имаме шест възможни заместявания, то всичките стойности на y_1 са

$$y_1,\ \alpha y_1,\ \alpha^2 y_1,\ y_2,\ \alpha y_2,\ \alpha^2 y_2.$$

От това е ясно, че функцията y_1^3 има само две стойности:

$$y_1^3 = (x_1 + \alpha x_2 + \alpha^2 x_3)^3 = u_1,$$

$$y_2^3 = (x_2 + \alpha x_1 + \alpha^2 x_3)^3 = u_2.$$

Квадратното уравнение, на което корените са u_1 и u_2, нека бъде

$$u^2 - \gamma u + \delta = 0,$$

пдето $\gamma = u_1 + u_2$, $\delta = u_1 u_2$ са симетрични функции на корените x_1, x_2, x_3; пресмятането им ще извършим, като използуваме теоремата за степента и теглото. Функцията

$$\gamma = (x_1 + \alpha x_2 + \alpha^2 x_3)^3 + (x_2 + \alpha x_1 + \alpha^2 x_3)^3$$

е от степен и тегло, равни на три, така че, както лесно се извежда за уравнение от вида (18) (т. е. в което $a_1 = 0$), тя ще бъде

$$\gamma = Aq,$$

гдето A е число, подлежащо на определяне. За да изчислим A, използуваме уравнението $x^3 - 1 = 0$, на което корените са 1, α, α^2 и $q = -1$. За него ще имаме

$$\gamma' = (1 + \alpha + \alpha^2)^3 + (\alpha + \alpha + \alpha)^3 = -A.$$

Понеже $1+\alpha+\alpha^2=0$, $\alpha^3=1$, то $A=-27$, отгдето

$$\gamma=-27\,q.$$

Функцията

$$\delta=(x_1+\alpha x_2+\alpha^2 x_3)^3\ (x_2+\alpha x_1+\alpha^2 x_3)^3$$

е от степен и тегло, равни на шест. Следователно за уравнението (18) тя ще има вида

$$\delta=Mp^3+Kq^2.$$

За уравнението $x^3-1=0$ ще имаме

$$\delta'=(1+\alpha+\alpha^2)^3\,(3\alpha)^3=K,$$

отгдето $K=0$. За уравнението $x^3-x=0$ с корени 0, 1, -1 имаме

$$\delta''=(\alpha-\alpha^2)^3\,(1-\alpha^2)^3=-M,$$

$$-M=\alpha^3\,(1-\alpha)^3\,(1-\alpha^2)^3=(2-\alpha-\alpha^2)^3=27,\ M=-27,$$

$$\delta=-27\,p^3.$$

Резолвентното уравнение ще бъде

$$u^2+27qu-27\,p^3=0. \tag{20}$$

Тогава от

$$x_1+x_2+x_3=0,$$

$$x_1+\alpha x_2+\alpha^2 x_3=\sqrt[3]{u_1},$$

$$\alpha x_1+x_2+\alpha^2 x_3=\sqrt[3]{u_2}$$

получаваме

$$x_1=\frac{\sqrt[3]{u_1}+\alpha^2\sqrt[3]{u_2}}{3},\quad x_2=\frac{\alpha^2\sqrt[3]{u_1}+\sqrt[3]{u_2}}{3},\quad x_3=\frac{\alpha\sqrt[3]{u_1}+\alpha\sqrt[3]{u_2}}{3}. \tag{20'}$$

Можем обаче да получим формула, в която да фигурира само единият радикал, като по такъв начин се избегне многозначността. Действително да умножим

$$(x_1+\alpha x_2+\alpha^2 x_3)(x_1+\alpha^2 x_2+\alpha x_3)=$$

$$=x_1^2+x_2^2+x_3^2+(\alpha+\alpha^2)\ (x_1 x_2+x_1 x_3+x_2 x_3)=$$

$$=-2p+p\,(\alpha+\alpha^2)=-3p.$$

Значи ще имаме

$$x_1+x_2+x_3=0,$$

$$x_1+\alpha x_2+\alpha^2 x_3=\sqrt[3]{u_1},$$

$$x_1+\alpha^2 x_2+\alpha x_3=-\frac{3p}{\sqrt[3]{u_1}}.$$

Оттук получаваме

$$x_1=\frac{1}{3}\sqrt[3]{u_1}-\frac{p}{\sqrt[3]{u_1}},$$

$$x_2=\frac{1}{3}\alpha^2\sqrt[3]{u_1}-\frac{p}{\alpha^2\sqrt[3]{u_1}},$$

$$x_3=\frac{1}{3}\alpha\sqrt[3]{u_1}-\frac{p}{\alpha\sqrt[3]{u_1}},$$

в които формули за $\sqrt[3]{u_1}$ може да се вземе коя да е негова стойност.

Функцията $y_1^3=(x_1+\alpha x_2+\alpha^2 x_3)^3$ се нарича резолвентна функция на Лагранж. Обобщена, тя играе важна роля в алгебрическото решение на уравненията, както в това ще се убедим по-нататък.

За уравнението от четвърта степен

$$x^4+px^2+qx+r=0 \tag{21}$$

трябва да вземем функция, която получава три стойности. Такава функция, както видяхме при разискване на методите на Декарт и Ойлер, е

$$(x_1+x_2-x_3-x_4)^2.$$

Ние ще разгледаме друга такава с три стойности:

$$y_1=x_1x_2+x_3x_4,$$

$$y_2=x_1x_3+x_2x_4,$$

$$y_3=x_1x_4+x_2x_3.$$

Кубичното уравнение, на което y_1, y_2, y_3 са корени, нека бъде

$$y^3+A_1y^2+A_2y+A_3=0.$$

За коефициентите му ще имаме

$$-A_1=y_1+y_2+y_3=\sum x_1x_2=p, \quad A_1=-p,$$

$$A_2=y_1y_2+y_1y_3+y_2y_3=$$

$$=(x_1x_2+x_3x_4)(x_1x_3+x_2x_4)+\ldots=\sum x_1^2x_2x_3$$

или

$$A_2=\sum x_1\sum x_1\,x_2\,x_3-4x_1\,x_2\,x_3\,x_4=-4r,$$

$$-A_3=y_1\,y_2\,y_3=(x_1\,x_2+x_3\,x_4)\,(x_1\,x_3+x_2\,x_4)\,(x_1\,x_4+x_2\,x_3)=$$

$$=\sum x_1^3\,x_2\,x_3\,x_4+\sum x_1^2\,x_2^2\,x_3^2,$$

или

$$-A_3=x_1\,x_2\,x_3\,x_4\sum x_1^2+\Big(\sum x_1\,x_2\,x_3\Big)^2-2\sum x_1^2\,x_2^2\,x_3\,x_4=$$

$$=rS_2+q^2-2x_1\,x_2\,x_3\,x_4\sum x_1\,x_2=q^2-4pr,\quad A_3=4pr-q^2.$$

Резолвентното уравнение ще бъде

$$y^3-py^2-4ry+4pr-q^2=0.$$

Определянето на x_1, x_2, x_3, x_4 може да стане така: ако поставим $x_1\,x_2=t_1$, $x_3\,x_4=t_2$, то ще имаме

$$t_1+t_2=y_1,$$

$$t_1\,t_2=x_1\,x_2\,x_3\,x_4=q.$$

Следва, че t_1, t_2 са корени на квадратното уравнение

$$t^2-y_1\,t+q=0,$$

отгдето

$$(22)\qquad x_1\,x_2=\frac{y_1+\sqrt{y_1^2-4q}}{2},\quad x_3\,x_4=\frac{y_1-\sqrt{y_1^2-4q}}{2}.$$

Аналогично ще получим

$$(23)\qquad x_1\,x_3=\frac{y_2+\sqrt{y_2^2-4q}}{2},\quad x_2\,x_4=\frac{y_2-\sqrt{y_2^2-4q}}{2},$$

$$x_1\,x_4=\frac{y_3+\sqrt{y_3^2-4q}}{2},\quad x_2\,x_3=\frac{y_3-\sqrt{y_3^2-4q}}{2}.$$

Знаците на радикалите са подложени на зависимостта

$$\sqrt{y_1^2-4q}\;\;\sqrt{y_2^2-4q}\;\;\sqrt{y_3^2-4q}=$$

$$=(x_1\,x_2-x_3\,x_4)\,(x_1\,x_3-x_2\,x_4)\,(x_1\,x_4-x_2\,x_3)=$$

$$=x_1\,x_2\,x_3\,x_4\,(x_1^2+x_2^2+x_3^2+x_4^2)-\sum x_1^2\,x_2^2\,x_3^2=-q^2.$$

От тази релация е ясно, че ако променим знака на един радикал, то трябва да променим и знака на другия, при което, както лесно се убеждаваме, отговаря само пермутиране на корените.

Пресмятането по-нататък на самите корени става без нови радикали. Действително, ако $q\neq 0$, имаме

$$-qx_1=(x_1\,x_2\,x_3+x_1\,x_2\,x_4+x_1\,x_3\,x_4+x_2\,x_3\,x_4)\,x_1=$$

$$=(x_1\,x_2)\;(x_1\,x_3)+(x_1\,x_2)(x_1\,x_4)+(x_1\,x_3)\;\;(x_1\,x_4)+r,$$

отгдето, използувайки (22) и (23), веднага се намира x_1, а оттам x_2, x_3, x_4. Ако $q=0$, то уравнението (21) е биквадратно и тогава със заместването $x^2=y$ се свежда на едно квадратно.

За да решим по метода на Лагранж уравненията от пета степен, трябва да намерим рационална функция от пет променливи, която взема четири стойности. Обаче доказва се, че такава функция не съществува.

Ще видим по-нататък, че уравненията от пета степен нагоре по една теорема на Абел са нерешими алгебрически.

Реципрочни уравнения

1. Реципрочни уравнения. Едно уравнение се нарича реципрочно, ако реципрочната стойност на всеки корен е пак корен на уравнението. Ако уравнението

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0 \tag{1}$$

е реципрочно, то на корените му

$$x_1,\ x_2,\ldots,x_n \tag{2}$$

реципрочните стойности

$$\frac{1}{x_1},\ \frac{1}{x_2},\ldots,\frac{1}{x_n} \tag{3}$$

са пак корени на (1), т. е. това са числата (2), само че в друг ред. Числата (3) са корени на уравнението

$$f_1(x)=x^n f\left(\frac{1}{x}\right)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0=0. \tag{4}$$

За да бъде уравнението (1) реципрочно, трябва значи (1) и (4) да имат едни и същи корени, за което е наобходимо и достатъчно коефициентите им да бъдат пропорционални:

$$a_0=qa_n,\quad a_1=qa_{n-1},\quad a_2=aq_{n-2},\ldots,\ a_{n-1}=qa_1,\ a_n=qa_0.$$

Ако умножим първото и последно равенство, получаваме

$$q^2=1,\quad q=\pm1.$$

Следователно ще имаме два вида реципрочни уравнения: първи вид, когато между коефициентите има връзките

$$a_0=a_n,\quad a_1=a_{n-1},\quad a_2=a_{n-2}\ldots,$$

и втори вид, ако

$$a_0=-a_n,\quad a_1=-a_{n-1},\quad a_2=-a_{n-2}\ldots$$

Ясно е, че ще имаме

$$f_1(x)=qf(x), \tag{5}$$

гдето $q=1$ при първия вид и $q=-1$ при втория.

Ако степента на уравнението (1) е нечетно число, то ще има един корен, който е реципрочен на себе си, т. е. $x=\frac{1}{x}$, отгдето се вижда, че то ще има поне един корен, равен на 1 или -1. От (5) или директно може веднага да се изследва в кои случаи този корен е 1 или -1. Действително при n нечетно, при $q=1$ и $x=-1$ получаваме от

$$x^n f\left(\frac{1}{x}\right)=f(x),$$

$$-f(-1)=f(-1),\ f(-1)=0,$$

т. е. -1 е корен на уравнения от първия вид. Ако $q=-1$, то от

$$x^n f\left(\frac{1}{x}\right)=-f(x) \tag{6}$$

при $x=1$ имаме

$$f(1)=-f(1),\quad f(1)=0,$$

т. е. 1 е корен на уравненията от втория вид. Ако n е четно и уравнението е от втори вид, то от (6) при $x=1$ и $x=-1$ имаме

$$f(1)=-f(1),\quad f(-1)=-f(-1),$$

отгдето $f(1)=f(-1)=0$, т. е. ± 1 са корени. Като отстраняваме тези реципрочни на себе си корени (1 и -1), най-сетне очевидно ще достигнем до реципрочно уравнение от четна степен и първи вид. Такова уравнение се нарича редуцирано. Общата форма на такова уравнение ще бъде

$$a_0 x^{2m}+a_1 x^{2m-1}+\ldots+$$

$$+a_{m-1} x^{m+1}+a_m x^m+a_{m-1} x^{m-1}+\ldots+a_1 x+a_0=0.$$

С подходящо заместване ще покажем, че степента му може да се намали два пъти. Действително да разделим на x^m и да групираме членовете му така:

$$a_0\left(x^m+\frac{1}{x^m}\right)+a_1\left(x^{m-1}+\frac{1}{x^{m-1}}\right)+\ldots+$$

$$+a_{m-1}\left(x+\frac{1}{x}\right)+a_m=0.$$

Да вземем за ново неизвестно

$$y=x+\frac{1}{x}$$

и означим с

$$P_i=x^i+\frac{1}{x^i}.$$

Лесно ще видим, че P_i представлява полином на y от степен i, който можем да пресметнем. Ще намерим една рекурентна формула между три последователни P_i. Да умножим P_{i-1} с y:

$$P_{i-1}\,y=\left(x^{i-1}+\frac{1}{x^{i-1}}\right)\left(x+\frac{1}{x}\right)=$$

$$=x^i+\frac{1}{x^i}+x^{i-2}-\frac{1}{x^{i-2}}=P_i+P_{i-2}.$$

Значи

$$P_i=yP_{i-1}-P_{i-2},$$

отгдето, понеже $P_1=y$, $P_2=y^2-2$, $P_0=2$, е очевидно, че P_i е полином от степен i спрямо y. Така получаваме за y едно уравнение от m-та степен:

$$a_0P_m+a_1P_{m-1}+\ldots+a_{m-1}P_1+a_m=0.$$

Ако $y_1, y_2, \ldots, y_m$ са корените на това уравнение, то корените на (1) ще бъдат дадени с квадратните уравнения:

$$x^2-y_1x+1=0,$$

$$x^2-y_2x+1=0,$$

$$\ldots\ldots\ldots$$

$$x^2-y_mx+1=0.$$

Например нека е дадено уравнението

$$x^5+2x^3-2x^2-1=0.$$

То има за корен 1, след отстраняването на който получаваме уравнението

(7) $$x^4+x^3+3x^2+x+1=0,$$

което може да се пише така:

$$\left(x^2+\frac{1}{x^2}\right)+\left(x+\frac{1}{x}\right)+3=0.$$

Като поставим

$$x+\frac{1}{x}=y,$$

получаваме

$$y^2+y+1=0,\quad y_{1,2}=-\frac{1}{2}\pm i\frac{\sqrt{3}}{2}.$$

Корените на (7) са корени на квадратните уравнения

$$x^2-x\left(-\frac{1}{2}\pm i\frac{\sqrt{3}}{2}\right)+1=0.$$

2. Биномни уравнения. Уравненията

8) $$x^n - A = 0$$

се наричат биномни. Ако поставим $x = z\sqrt[n]{A}$, то получаваме

(9) $$z^n - 1 = 0.$$

Корените на това уравнение по формулата на Моавър са

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

$$(k = 0,\ 1,\ 2, \ldots, n-1).$$

На корена z_k конюгованият е равен на z_{n-k}. Действително

$$z_{n-k} = \cos\frac{2\pi(n-k)}{n} + i\sin\frac{2\pi(n-k)}{n} = \cos\frac{2\pi k}{n} - i\sin\frac{2\pi k}{n} = \bar{z}_k.$$

Ако n е нечетно, (9) има само един реален корен $z_0 = 1$, понеже аргументът на z_k е нула и е кратен на π само при $k = 0$ и $k = \frac{n}{2}$ което е невъзможно, понеже n е нечетно. При n четно уравнението (9) има два реални корена:

$$z_0 = 1, \quad z_{\frac{n}{2}} = -1.$$

При $A > 0$ в субституцията $x = z\sqrt[n]{A}$ вземаме реалния n-ти корен Тогава всички корени на (8) са $\alpha\sqrt[n]{A}$, гдето α взема стойности, равни на всички корени на (9). При $A < 0$ и n-нечетно пак е възможна горната реална субституция. Но ако n е четно, то полагаме $x = z\sqrt[n]{-A}$ уравнението (8) се трансформира в

10) $$z^n + 1 = 0.$$

По формулата на Моавър корените на това уравнение са

$$z_k = \cos\frac{(2k+1)\pi}{n} + i\sin\frac{(2k+1)\pi}{n}$$

$$(k = 0,\ 1,\ 2, \ldots, n-1).$$

Оттук лесно се намира броят на реалните корени.

Именно z_k може да е реално само при $\frac{(2k+1)\pi}{n}$ равно на π, понеже $0 < \frac{2k+1}{n} < 2$, т. е. $k = \frac{n-1}{2}$, което е възможно само при n нечетно. Така че (10) има при n нечетно един реален корен — 1 и при n четно всички негови корени са имагинерни, което впрочем и директно

се вижда в този случай, понеже при z реално лявата част на (10) е положителна.

На z_k конюгованият корен е z_{n-k-1}. Действително

$$z_{n-k-1}=\cos\frac{\pi}{n}[2(n-k)-1]+i\sin\frac{\pi}{n}[2(n-k)-1]=$$

$$=\cos\left(2\pi-\frac{2k+1}{n}\pi\right)+i\sin\left(2\pi-\frac{2k+1}{n}\pi\right)=$$

$$=\cos\frac{(2k+1)\pi}{n}-i\sin\frac{(2k+1)\pi}{n}$$

или $z_{n-k-1}=\bar{z}_k$.

Уравненията (9) и (10) са реципрочни. Следователно можем да приложим върху тях теорията на тези уравнения. Отначало нека разгледаме уравненията (9) при n нечетно $=2m+1$. Тогава, като отстраним корена 1 на уравнението

$$x^{2m+1}-1=0,$$

получаваме уравнението

$$x^{2m}+x^{2m-1}+\ldots+x+1=0,$$

от което

$$\left(x^m+\frac{1}{x^m}\right)+\left(x^{m-1}+\frac{1}{x^{m-1}}\right)+\ldots+\left(x+\frac{1}{x}\right)+1=0.$$

Като поставим $x+\frac{1}{x}=y$, получаваме уравнение от m-та степен. Така при $n=3$ уравнението е

$$x^3-1=0.$$

Корените му са

$$1,\ \frac{-1\pm i\sqrt{3}}{2}.$$

При $n=5$

$$x^5-1=0,$$

от което, като отстраним корена 1, получаваме

(11)
$$x^4+x^3+x^2+x+1=0,$$

$$\left(x^2+\frac{1}{x^2}\right)+\left(x+\frac{1}{x}\right)+1=0,$$

$$(y^2-2)+y+1=0,$$

$$y^2+y-1=0,$$

$$y_{1,2}=\frac{-1\pm\sqrt{5}}{2}.$$

Корените на (11) се дават от уравненията

$$x^2 - xy_{1,2} + 1 = 0,$$

$$x_{1,2,3,4} = \frac{-1 \pm \sqrt{5}}{4} \pm \frac{\sqrt{10 \pm 2\sqrt{5}}}{4} i.$$

Уравнението

$$x^{2m+1} + 1 = 0$$

има само реалния корен -1. С премахване на тоя корен то се свежда на уравнението

$$x^{2m} - x^{2m-1} + x^{2m-2} - \ldots + 1 = 0.$$

Обаче това уравнение веднага се свежда на предишното, стига да поставим $x = -y$.

Уравнението

$$x^{2m} - 1 = 0$$

веднага се свежда на уравненията

$$x^m - 1 = 0, \quad x^m + 1 = 0.$$

Така корените на $x^4 - 1 = 0$, което се разлага на $x^2 - 1 = 0$, $x^2 + 1 = 0$, са

$$1, \ -1, \ i, \ -i.$$

На уравнението

$$x^6 - 1 = 0$$

корените ще бъдат

$$\pm 1, \ \frac{-1 \pm i\sqrt{3}}{2}, \ \frac{1 \pm i\sqrt{3}}{2}.$$

Уравнението

$$x^{2m} + 1 = 0$$

няма реален корен. Като го напишем така:

$$x^m + \frac{1}{x^m} = 0,$$

със субституцията

$$y = x + \frac{1}{x}$$

то се свежда на уравнение от m-та степен. Така $x^4 + 1 = 0$ дава

$$x^2 + \frac{1}{x^2} = 0, \quad y^2 - 2 = 0, \quad y_{1,2} = \pm\sqrt{2}.$$

Корените му са $\frac{\pm 1 \pm i}{\sqrt{2}}$.

Глава II

За корените на единицата

1. Примитивни корени. Корените на биномните уравнения

(1) $$x^n=1$$

се наричат корени на единицата. Нека α е един корен на (1) Тогава всички членове на редицата

(1) $$\ldots\alpha^{-3},\ \alpha^{-2},\ 1,\ \alpha,\ \alpha^2,\ \alpha^3,\ldots$$

са пак корени на (1). Действително при k цяло имаме

$$(\alpha^k)^n=(\alpha^n)^k=1.$$

В редицата (2) корените се повтарят; така имаме

$$\alpha^n=1,\ \alpha^{n+1}=\alpha,,\ldots,\alpha^{-1}=\alpha^{n-1},\ \alpha^{-2}=\alpha^{n-2},\ldots$$

Нека μ е най-малката положителна степен, за която

$$\alpha^\mu=1,$$

т. е. α е корен на биномното уравнение

(3) $$x^\mu=1,$$

но не е корен на биномното уравнение от по-ниска степен. Казваме, че α е примитивен корен на (3) или че принадлежи на степен μ.

Ако α е примитивен корен на уравнението (3), то всички негови корени са дадени с редицата

(4) $$\alpha,\ \alpha^2,\ \alpha^3,\ldots,\alpha^{\mu-1},\ \alpha^\mu=1.$$

Действително всички числа от (4) са корени на (3). Две кои да е от тях не могат да бъдат равни помежду си. Действително, ако допуснем, че

$$\alpha^s=\alpha^r,\ s>r,$$

то бихме получили

$$\alpha^{s-r}=1,$$

което не е възможно, понеже $s-r<\mu$, освен ако $s=r$. Понеже броят на числата (4) е μ, то следва оттук, че това са всички корени на (3).

Друго едно важно свойство е следното: ако α е примитивен корен на (3), то степените на биномните уравнения, на които α е корен, се делят на μ.

Действително нека

(5) $$\alpha^m=1$$

и нека q и r са съответно частното и остатъкът от делението на m с μ, т. е.

$$m=q\mu+r,\quad 0\leqq r<\mu.$$

Условието (5) е

$$\alpha^{\mu q+r}=1, \quad \alpha^{r}=1,$$

понеже $\alpha^{\mu q}=1$. Последното е възможно само ако $r=0$, отгдето $m=q\mu$, което трябваше да се докаже. Следователно всичките непримитивни корени на $x^n=1$ ще са примитивни корени на биномни уравнения, степените на които делят n. Така например корените на $x^6=1$ са числата

$$1, \ -1, \ \varepsilon, \ \varepsilon^2, \ -\varepsilon, \ -\varepsilon^2,$$

гдето

$$\varepsilon=\frac{-1+i\sqrt{3}}{2}.$$

Непримитивните са корени на уравненията

$$x=1, \ x^2=1, \ x^3=1,$$

т. е. числата

$$1, \ -1, \ \varepsilon, \ \varepsilon^2.$$

Следователно примитивните корени на $x^6=1$ са

$$-\varepsilon, \ -\varepsilon^2.$$

Нека разгледаме две биномни уравнения:

(6) $$x^n-1=0, \quad x^m-1=0.$$

Те имат поне един общ корен, равен на единица, който принад. лежи на всички биномни уравнения. Общите им корени ще бъдат дадени с анулирането на общия най-голям делител на левите им части. Ще установим, че най-големият общ делител на

$$x^n-1, \quad x^m-1$$

е равен на x^d-1, гдето d е общият най-голям делител на n и m. Да допуснем, че $n \geqq m$, и нека частното и остатъкът от делението на n с m е q_1 и r_1. Тогава от

$$\frac{x^n-1}{x^m-1}=\frac{x^{mq_1+r_1}-1}{x^m-1}=\frac{x^{r_1}(x^{mq_1}-1)}{x^m-1}+\frac{x^{r_1}-1}{x^m-1},$$

понеже

$$\frac{x^{mq_1}-1}{x^m-1}=x^{m(q_1-1)}+\ldots+1,$$

$x^{r_1}-1$ е остатъкът от делението на x^n-1 с x^m-1. Ако по-нататък

$$m=r_1q_2+r_2, \quad 0 \leqq r_2<r_1,$$

то остатъкът от делението на x^m-1 с $x^{r_1}-1$ е равен на $x^{r_2}-1$. Като продължаваме така, ще достигнем до последния остатък x^d-1, който според познатата ни вече теория ще бъде общ най-голям дели-

тел на полиномите x^n-1 и x^m-1. Общите корени на уравненията (6) ще бъдат дадени с

$$x^d=1.$$

Следователно, ако n и m са взаимно прости, $d=1$, т. е. уравненията (6) имат само единицата за общ корен.

2. Някои общи теореми. Ако a и b са взаимно прости числа, то корените на уравнението

(7) $$x^{ab}=1$$

се получават, като умножим всеки корен на

(8) $$x^a=1$$

с всеки корен на

(9) $$x^b=1$$

Примитивните корени на (7) се получават, като умножаваме примитивните корени на (8) и (9).

Нека β е корен на (8), а γ на (9). От

$$\beta^a=1,\ \gamma^b=1$$

получаваме

$$\beta^{ab}=1,\quad \gamma^{ab}=1,\ (\beta\gamma)^{ab}=1.$$

Следователно $\alpha=\beta\gamma$ ще е корен на (7). Така получаваме на брой ab корена на (7). За да докажем, че това са всички негови корени, трябва да установим, че тези произведения са различни помежду си. Действително да допуснем противното, т. е.

$$\beta\gamma=\beta_1\gamma_1.$$

Като повдигнем това равенство в b-та степен и вземем под внимание, че $\gamma^b=\gamma_1{}^b=1$, получаваме

$$\beta^b=\beta_1{}^b,\left(\frac{\beta}{\beta_1}\right)^b=1.$$

Но имаме

$$\left(\frac{\beta}{\beta_1}\right)^a=1,$$

т. е. $\frac{\beta}{\beta_1}$ е общ корен на уравненията (8) и (9). Понеже a и b са взаимно прости, то трябва

$$\frac{\beta}{\beta_1}=1,\ \beta=\beta_1,$$

отгдето $\gamma=\gamma_1$.

Нека β е примитивен корен на (8), а γ — примитивен корен на (9), то $\alpha=\beta\gamma$ е примитивен корен на (7). Ако допуснем обратното, ще има число $s<ab$ такова, че

$$(\beta\gamma)^s=1,\quad \beta^s\gamma^s=1.$$

Ако повдигнем това равенство на степен b, получаваме

$$\beta^{bs}\gamma^{bs}=1, \quad \beta^{bs}=1.$$

Понеже β е примитивен корен на (8), трябва sb да се дели на a, отгдето, понеже a и b са взаимно прости, трябва s да се дели на a. По същия начин получаваме, че s се дели на b, т. е. s е най-малко равно на ab.

Ако β не е примитивен корен на (8), т. е.

$$\beta^k=1, \; k<a,$$

то ще имаме

$$(\beta\gamma)^{kb}=1, \quad kb<ab,$$

т. е. $\beta\gamma$ не е примитивен на (7). Така теоремата е установена напълно

Ясно е, че тази теорема може да се разшири за степени $n=abcd\ldots$, гдето всички множители $a, b, c, d\ldots$ са прости помежду си. Така получаваме теоремата:

Ако $n=p^{\lambda}q^{\mu}r^{\nu}\ldots$, гдето $p, q, r,\ldots$ са простите делители на n, то решението на уравнението

$$x^n-1=0$$

се свежда на решението на уравненията

$$x^{p^{\lambda}}=1, \quad x^{q^{\mu}}=1, \quad x^{r^{\nu}}=1,\ldots$$

Ако β е кой да е корен на първото, γ кой да е корен на второто, δ кой да е корен на третото и т. н., то произведенията на $\beta\gamma\delta\ldots$, броят на които е $p^{\lambda}q^{\mu}r^{\nu}\ldots=n$, дават всички корени на уравнението $x^n=1$.

Ако $\beta, \gamma, \delta,\ldots$ са примитивните корени на своите уравнения, то $\beta\gamma\delta\ldots$ дава всички примитивни корени на $x^n=1$.

Така например корените на

$$x^{12}=1, \text{ т. е. } x^{2^2\cdot 3}=1,$$

получаваме, като умножим корените на уравненията

$$x^3=1, \quad x^4=1.$$

Понеже първото има корени $1, \varepsilon, \varepsilon^2$, от които $\varepsilon, \varepsilon^2$ са примитивни, а второто — корени $1, -1, i, -i$, от които $i, -i$ са примитивни, то всичките корени на $x^{12}=1$ са

$$1, -1, i, -i, \varepsilon, -\varepsilon, \underline{i\varepsilon}, \underline{-i\varepsilon}, \varepsilon^2, -\varepsilon^2, \underline{i\varepsilon^2}, \underline{-i\varepsilon^2},$$

от които подчертаните са примитивни.

Посредством горната теорема виждаме, че решението на биномните уравнения се свежда на решението на уравненията

(10) $$x^{p^{\lambda}}=1,$$

гдето p е просто число. Както ще покажем, решението на това биномно уравнение се свежда на решение на биномни уравнения от степен p.

Действително нека β_1 е корен на уравнението

$$x^p = 1,$$

β_2 на уравнението

$$x^p = \beta_1$$

и най-сетне β_λ е корен на

$$x^p = \beta_{\lambda-1}.$$

Тогава произведението $\beta_1 \beta_2 \dots \beta_\lambda$ е корен на (10). Освен това всичките корени на (10) се получават по този начин. Наистина имаме

$$\begin{aligned}
&(\beta_1 \beta_2 \dots \beta_\lambda)^p = \beta_1 \beta_2 \dots \beta_{\lambda-1},\\
&(\beta_1 \beta_2 \dots \beta_\lambda)^{p^2} = (\beta_1 \beta_2 \dots \beta_{\lambda-1})^p = \beta_1 \beta_2 \dots \beta_{\lambda-2},\\
&\qquad \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \qquad (11)\\
&(\beta_1 \beta_2 \dots \beta_\lambda)^{p^\lambda} = \beta_1^{\lambda} = 1.
\end{aligned}$$

Освен това две такива произведения са различни помежду си, понеже от

$$\beta_1 \beta_2 \dots \beta_\lambda = \beta'_1 \beta'_2 \dots \beta_\lambda'$$

с повдигане в степен p бихме получили

$$\beta_1 \beta_2 \dots \beta_{\lambda-1} = \beta_1' \dots \beta'_{\lambda-1},$$

т. е. $\beta_\lambda = \beta_\lambda'$, и продължавайки така, $\beta_{\lambda-1} = \beta'_{\lambda-1}, \dots, \beta_1 = \beta_1'$. Ако $\beta_1 = 1$, то $\alpha = \beta_1 \beta_2 \dots \beta_\lambda$ не е примитивен корен на (10), понеже от (11) имаме $\alpha^{p^{\lambda-1}} = \beta_1 = 1$. Ако обаче $\beta_1 \neq 1$, то α ще бъде примитивен корен на (10) и така се получават всички примитивни корени на това уравнение.

3. Брой на примитивните корени. Ще изведем този брой по два начина. Отначало да определим броя на примитивните корени на уравнението

$$x^{p^\lambda} = 1, \qquad (12)$$

гдето p е просто число. Всички непримитивни корени на това уравнение ще бъдат корени на уравненията

$$x = 1,\ x^p = 1,\ x^{p^2} = 1, \dots,\ x^{p^{\lambda-1}} = 1.$$

Но корените на тези уравнения са такива на последното от тях. Следователно(12) има на брой

$$p^\lambda - p^{\lambda-1} = p^\lambda \left(1 - \frac{1}{p}\right)$$

примитивни корена.

Нека разгледаме общия случай. Дадено е биномното уравнение

$$x^n = 1,$$

гдето $n=p^\lambda q^\mu r^\nu \ldots$ Както вече показахме, примитивните корени на това уравнение се получават, като умножаваме примитивните корени на уравненията

$$x^{p^\lambda}=1,\ x^{q^\mu}=1,\ x^{r^\nu}=1\ldots,$$

отгдето следва, че броят на примитивните корени на $x^n=1$ ще бъде равен на

$$p^\lambda\left(1-\frac{1}{p}\right)q^\mu\left(1-\frac{1}{q}\right)r^\nu\left(1-\frac{1}{r}\right)\ldots=n\left(1-\frac{1}{p}\right)\left(1-\frac{1}{q}\right)\left(1-\frac{1}{r}\right)\ldots$$

По-рано видяхме, че ако α е примитивен корен на

$$x^n=1,$$

то всичките му корени са

$$\alpha,\ \alpha^2,\ \alpha^3,\ldots,\ \alpha^{n-1},\ \alpha^n=1. \tag{13}$$

Да изследваме колко от тези корени са примитивни. Нека числото α^m принадлежи на степен s, т. е. α^m е примитивен корен на $x^s=1$. Тогава от условието $\alpha^{ms}=1$ следва, че ms трябва да се дели на n. Ако d е общият най-голям делител на m и n, то трябва s да се дели на $\frac{n}{d}$. Най-малкото такова s е равно на $\frac{n}{d}$, т. е. α^m принадлежи на тази степен. Оттук заключаваме, че α^m е само тогава примитивен корен на $x^n=1$, когато m и n са взаимно прости.

Следователно в редицата (13) има толкова примитивни корени, колкото е броят на числата, по-малки и взаимно прости с n. Последното число бележим с $\varphi(n)$ и наричаме индикатор от n. Значи броят на примитивните корени на $x^n=1$ е равен на $\varphi(n)$ или като използуваме горния резултат, получаваме равенството

$$\varphi(n)=n\left(1-\frac{1}{q}\right)\left(1-\frac{1}{q}\right)\left(1-\frac{1}{r}\right)\ldots,$$

което ще докажем и директно в теорията на числата.

4. Уравнение на примитивните корени. Лесно се получава уравнението, на което корените са примитивни корени на

$$x^n=1.$$

Отначало да разгледаме случай $n=p$ просто число. Тогава всички корени с изключение на единицата са примитивни. Следователно за уравнението

$$x^p=1$$

уравнението на примитивните корени е

$$\frac{x^p-1}{x-1}=x^{p-1}+x^{p-2}+\ldots+1=0.$$

На уравнението

$$x^{p^\lambda}=1$$

непримитивните корени ще бъдат корени на уравнението

$$x^{p^{\lambda-1}}=1.$$

Следва, че уравнението на примитивните корени ще бъде

$$\frac{x^{p^\lambda}-1}{x^{p^{\lambda-1}}-1}=x^{(p-1)p^{\lambda-1}}+\ldots+x^{p^{\lambda-1}}+1=0.$$

Нека имаме сега по-общ случай:

$$n=p^\lambda q^\mu,$$

p и q прости числа. Непримитивните корени на

$$x^{p^\lambda q^\mu}=1 \tag{14}$$

ще бъдат корени на уравненията

$$x^{p^\alpha q^\beta}=1,$$

гдето $\alpha=0, 1, 2,\ldots, \lambda$; $\beta=0, 1, 2,\ldots, \mu$; като изключим случая $\alpha=\lambda$, $\beta=\mu$. Но корените на тези уравнения принадлежат или на уравнението

$$x^{p^{\lambda-1}q^\mu}=1, \text{ т. е. } x^{\frac{n}{p}}=1,$$

или на

$$x^{p^\lambda q^{\mu-1}}=1, \text{ т. е. } x^{\frac{n}{q}}=1,$$

или на двете. Последните уравнения имат общи корени, дадени с уравнението

$$x^{\frac{n}{pq}}=1.$$

Така че всичките непримитивни корени на (14) ще бъдат корени на уравнението

$$\frac{(x^{\frac{n}{p}}-1)\,(x^{\frac{n}{q}}-1)}{x^{\frac{n}{pq}}-1}=0.$$

Уравнението на примитивните корени на (14) ще бъде

$$\frac{(x^n-1)\,(x^{\frac{n}{pq}}-1)}{(x^{\frac{n}{p}}-1)\,(x^{\frac{n}{q}}-1)}=0.$$

Следвайки същия път на доказване, получаваме, че уравнението на примитивните корени на

$$x^n=1,\ n=p^\lambda q^\mu r^\nu\ldots$$

ще бъде

$$\frac{(x^n-1)\,(x^{\frac{n}{pq}}-1)\,(x^{\frac{n}{pr}}-1)\,(x^{\frac{n}{qr}}-1)\ldots}{(x^{\frac{n}{p}}-1)\,(x^{\frac{n}{q}}-1)\,(x^{\frac{n}{r}}-1)\ldots(x^{\frac{n}{pqr}}-1)\ldots}=0.$$

Оттук отново можем да намерим броя на примитивните корени, като пресметнем степента на горното уравнение. Именно ще имаме за този брой

$$n\left(1+\frac{1}{pq}+\frac{1}{pr}+\frac{1}{qr}+\ldots-\frac{1}{p}-\frac{1}{q}-\frac{1}{r}-\ldots-\frac{1}{pqr}-\ldots\right)=$$

$$=n\left(1-\frac{1}{p}\right)\left(1-\frac{1}{q}\right)\left(1-\frac{1}{r}\right)\ldots$$

Така за уравнението

$$x^{12}-1=0$$

уравнението на примитивните корени ще бъде

$$\frac{(x^{12}-1)\,(x^2-1)}{(x^6-1)\,(x^4-1)}=x^4-x^2+1=0.$$

5. Степенни сборове за корените на единицата. За биномното уравнение

$$x^n=1$$

степенните сборове се пресмятат лесно. Именно от формулите на Нютон, като вземем под внимание, че

$$a_1=a_2=\ldots=a_{n-1}=0,\ a_n=-1,$$

получаваме

$$S_k=0,\quad 1\leqq k<n,$$

$$S_n=n,\quad S_{n+1}=0,\ldots,\ S_{2n-1}=0,\quad S_{2n}=n,\ldots$$

Изобщо S_μ е равно на нула, ако μ не се дели на n, и е равно на n в противен случай.

Този резултат можем да получим директно така: нека α е примитивен корен на $x^n=1$. Тогава корените му, както видяхме, са

$$\alpha,\ \alpha^2,\ \alpha^3,\ldots,\alpha^{n-1},\ \alpha^n=1.$$

Следователно имаме

$$S_k=1+\alpha^k+\alpha^{2k}+\ldots+\alpha^{(n-1)k}=\frac{\alpha^{nk}-1}{\alpha^k-1}=0,$$

ако k не се дели на n, и равно на n, когато k се дели на n. От горното също следва, че коя да е проста симетрична функция от корените

$$\sum \alpha^p \beta^q \gamma^r \ldots$$

е равна на нула, стига степентай $p+q+r+\ldots$ да не се дели на n. Действително знаем, че тази функция се пресмята посредством степенните сборове с членове от вида $S_i S_k S_l \ldots$, гдето

$$i+k+l+\ldots = p+q+r+\ldots$$

Следователно едно поне от числата $i, k, l, \ldots$ не се дели на n и членовете ще са равни на нула. Същото имаме и за хомогенните симетрични функции, полеже са сума от прости.

ЧАСТ VII

ГРУПИ И ПОЛЕТА

Глава I. ГРУПА

1. Понятие за група. Нека **G** е едно непразно множество от числа, операции, дефинирани по еднакъв начин, или какви да е математически неща, които ще наричаме елементи, броят на които е краен или безкраен. Множеството **G** ще наричаме група, ако са изпълнени следните постулати:

I. Закон за групата. На всеки наредена чифт от два равни или не елемента от G отговаря еднозначно един елемент от нея, който наричаме произведение на взетите два елемента. Това означаваме с формулата $ab=c$.

II. Съдружителен закон. За произведението е в сила равенството

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

III. Елемент единица. Съществува елемент e от **G**, наречен елемент единица или просто единица, за който имаме равенството

$$ae=ea=a$$

за всеки елемент a от **G**.

IV. Обратен (инверзен) елемент. На всеки елемент a съответствува елемент x, който удовлетворява уравнението

$$xa=e.$$

Елементът x се нарича обратен (инверзен) на a и се бележи с a^{-1}. Като се основаваме на постулата I, можем лесно да видим, че съществува само една единица. Защото, ако предположим, че съществува и друга единица, т. е. елемент e_1, удовлетворяващ на равенството $ae_1=e_1a=a$ за всеки елемент a от **G**, би следвало, че $e_1e=ee_1$ представлява елементът e и елементът e_1. По постулата I имаме тогава $e=e_1$.

На a^{-1} обратният елемент е a. Действително от равенството

$$ya^{-1}=e$$

с умножение вдясно с a получаваме

$$ya^{-1}a=ea=a,$$

отгдето следва, че $y=(a^{-1})^{-1}=a$.

Ако броят на елементите на групата е краен, то тя се нарича **крайна**. Тогава броят на елементите ѝ се нарича **ред на групата**. Другите групи се наричат **безкрайни**.

Ако елементите на една група са комутативни два по два, групата се нарича **комутативна или абелева**.

Постулатите III и IV могат да се заместят със следните:

III′. В **G** има поне една (лява) единица e, притежаваща **свойството**

$$ea = a$$

за всеки елемент $a \in \mathbf{G}$.

IV′. За всеки елемент $a \in \mathbf{G}$ съществува поне един (ляв) обратен елемент a^{-1} от **G**, притежаващ **свойството**, че

$$a^{-1}a = e.$$

Действително от предните постулати следва

$$a^{-1}aa^{-1} = ea^{-1} = a^{-1}.$$

Като умножим предното равенство с елемент x, обратен на a^{-1}, получаваме

$$xa^{-1}\,aa^{-1} = xa^{-1}$$

или

$$eaa^{-1} = e.$$

Следователно ще имаме

$$aa^{-1} = e.$$

Това равенство показва, че левият обратен елемент е и десен такъв. От равенството $a \,.\, a^{-1} = e$ следва, че a е обратен елемент на a^{-1}. По-нататък имаме

$$ae = aa^{-1}a = ea = a,$$

т. е. лявата единица е и дясна единица.

Ще установим теореми за възможност на действието деление.

V. Ако a и b са **произволни елементи от G, то уравненията**

$$ax = b,\ ya = b$$

имат решение в G.

Действително, ако умножим първото уравнение вляво с a^{-1}, а второто—вдясно с a^{-1}, получаваме

$$a^{-1}ax = a^{-1}b, \quad yaa^{-1} = ba^{-1},$$

т. е.

$$x = a^{-1}b, \quad y = ba^{-1}.$$

С непосредствено заместване се убеждаваме лесно, че тези елементи удовлетворяват на уравненията.

VI. От $ax=ax'$, както и от $xa=x'a$, следва, че $x=x'$.

Действително, ако умножим вляво първото равенство с a^{-1}, получаваме $x=x'$. Аналогично се третира второто равенство.

В частност оттук следва съществуването на само една единица e (като решение на уравнението $xa=a$) и съществуването само на един обратен елемент (като решение на уравнението $xa=e$).

Ако приемем теорема V за постулат, лесно се вижда, че постулатите III′, IV′ стават теореми, т. е. за дефиниция на група можем да приемем само постулатите I, II и V. Действително нека c е произволен елемент от **G** и да означим с e решението на уравнението $xc=c$, т. е. ще имаме

$$ec=c.$$

Ако a е произволен елемент от **G**, нека решим уравнението

$$cx=a.$$

Тогава ще имаме

$$ea=ecx=cx=a,$$

т. е. e е единица и III е установено.

Постулатът IV се явява в случая като непосредствено следствие от възможността за разрешимост на уравненията.

Ако групата е крайна, постулатът V може да се замести с VI.

Съгласно с VI следва, че $ax \neq ax'$, щом като $x \neq x'$. Също $xa \neq x'a$. Нека тогава $a_1, a_2, \ldots, a_m$ са елементите на **G** и да вземем един кой да е елемент a_i. Тогава произведенията

$$a_ia_1,\ a_ia_2,\ a_ia_3, \ldots,\ a_ia_m$$

са все различни и са елементи от **G**. Понеже броят им е m, то това са всичките елементи на **G**. Подобно

$$a_1a_i,\ a_2a_i,\ a_3a_i,\ a_4a_i, \ldots,\ a_ma_i$$

са всичките елементи на **G**. Всеки елемент a_i ще фигурира по един път в предните две редици и уравненията

$$a_ix=a_j,\ ya_i=a_j$$

ще имат по едно единствено решение.

Ако групата е абелева, обикновено вместо произведение се пише сума и единицата се замества с нула (0). Обратният на a елемент се означава с $-a$ в пълна аналогия с числата.

Така множеството от всичките рационални числа, отлични от нула, е абелева група. Единицата на тази група е числото 1. Числата 1 и -1 образуват също група. Множеството от всички субституции от n елемента образува група, от $n!$-ти ред. Съвкупността от всички завъртвания на равнината около една нейна точка образува група. Ако α е едно завъртване и β е друго такова, то под $\alpha\beta$ разбираме извършването на

завъртването α и след това на β. Очевидно в резултат получаваме завъртване $\gamma=\alpha\beta$, което привежда равнината от началното положение в крайното.

Нека z е комплексно променливо и нека a, b, c, d да бъдат какви да е цели числа, удовлетворяващи на равенството

$$ad-bc=1.$$

Да разгледаме дробната линейна функция

(1) $$z'=\frac{az+b}{cd+d},$$

която привежда точката z в точката z', т. е. на всяка точка z от комплексната равнина ще отговаря точка z' от същата равнина. На късо тази функционна зависимост да означим с

$$z'=S(z).$$

Нека

$$z'=T(z)=\frac{c_1z+b_1}{c_1z+d_1}$$

е друга субституция от разгледаната съвкупност, която може да бъде и еднаква с първата. Прилагайки върху z' субституцията $T(z)$, получаваме

$$z''=T(z')=T[S(z)]=\frac{a_2z+b_2}{c_2z+d_2}=U(z),$$

където a_2, b_2, c_2, d_2 са цели числа, за които получаваме

$$(a_2d_2-b_2c_2)=(aa_1+b_1c)(c_1b+d_1b)-(a_1b+b_1d)(ac_1+ad_1)=$$
$$=(a_1d_1-b_1c_1)(ad-bc)=1.$$

Следователно субституцията $z''=TS(z)$ принадлежи на същата съвкупност.

На $S(z)$ обратната субституция е

$$z=S^{-1}(z')=\frac{dz'-b}{-cz'+d},$$

която принадлежи също на дадената съвкупност. Така виждаме, че съвкупността от субституцията (1) образува група, като единицата на групата е субституцията $z'=z$. Субституцията $U(z)$ е произведение на субституциите $S(z)$ и $T(z)$. При това лесно се вижда, че въобще $ST\neq TS$ и следователно групата не е комутативна. Нека отбележим, че субституциите могат да се номерират в една редица.

Да разгледаме всички субституции от n елемента:

$$S=\begin{pmatrix}1&2&3&\dots&n\\ i_1&i_2&i_3&\dots&i_n\end{pmatrix}.$$

Произведението на две субституции е също субституция от същия вид. Очевидно обратната субституция S^{-1} на S е

$$S^{-1}=\begin{pmatrix} i_1 & i_2 & i_3 & \dots & i_n \\ 1 & 2 & 3 & \dots & n \end{pmatrix}.$$

Следователно тези субституции образуват група от ред $n!$. Единицата на групата е идентичната субституция

$$E=\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ 1 & 2 & 3 & \dots & n \end{pmatrix}.$$

Понеже симетричните функции не се изменят при прилагане на субституциите от групата, то тя се нарича симетричната група.

Съдружително свойство за повече множители. Произведението на три елемента a, b, c от една група G дефинирахме с

$$abc=(ab)\,c=a\,(bc).$$

Подобно по индуктивен път дефинираме произведение на повече елементи:

$$a_1a_2\dots a_na_{n+1}=(a_1a_2\dots a_n)\,a_{n+1}.$$

Ще докажем, че произведението на произволен брой елементи не зависи от реда на умножение на отделните елементи при предположение, че не променяме местата на тези елементи, т. е. съдружителният закон е в сила за произволен брой множители.

Предполагаме, че свойството е доказано за произведения от $n+m$ множителя. Ще докажем, че то е в сила и за произведения от $n+m+1$ множителя, т. е. ще имаме

$$(2)\qquad \begin{aligned}(a_1a_2\dots a_m)\,(a_{m+1}\,a_{m+2}\dots a_{m+n}\,a_{m+n+1})&= \\ =(a_1a_2\dots a_m\,a_{m+1}\dots a_{m+n})\,a_{m+n+1}&.\end{aligned}$$

Но в произведението $a_1\,a_2\dots a_m\,a_{m+1}\dots a_{m+n}$ броят на множителите не надминава $n+m$ и следователно ще имаме

$$a_1a_2\dots a_ma_{m+1}\dots a_{m+n}=(a_1a_2\dots a_m)\;(a_{m+1}\dots a_{m+n}).$$

Тогава произведението (2) може да се пише

$$(a_1a_2\dots a_m)\;[(a_{m+1}\,a_{m+2}\dots a_{m+n})\;a_{m+n+1}\,].$$

Прилагайки съдружителния закон, предният израз става

$$(a_1a_2\dots a_m)\;(a_{m+1}\,a_{m+2}\dots a_{m+n})\,a_{m+n+1}.$$

На основание на същия закон за не повече от $m+n$ множителя предното произведение е равно на

$$(a_1a_2\dots a_ma_{m+1}\,a_{m+2}\dots a_{m+n})\,a_{m+n+1}.$$

С това установихме верността на равенството (2). Понеже съдружителното свойство е вярно (по приемане) за $n=3$, то следва, че същото свойство е вярно и за произволен брой множители.

В случай, че групата **G** е абелева, ще докажем, че произведението на произволен брой елементи не се изменя с разместването помежду им.

Предполагаме, че това свойство сме установили за произведение от n елемента. Ще установим, че свойството остава в сила и за произведение от $n+1$ елемента. Понеже свойството е вярно за два множителя, то следва тогава, че то ще е вярно за произволен брой множители.

Нека имаме произведението от $n+1$ елемента

$$a_1a_2a_3 \ldots a_n a_{n+1}$$

и нека a_p и a_{p+1} са два последователни множителя от него. Ще разгледаме два случая:

1. $p<n$. Тогава a_p и a_{p+1} са множители от произведението

$$a_1a_2a_3 \ldots a_{n-1}a_n.$$

Понеже разместителното свойство е в сила, ще имаме

$$a_1a_2 \ldots a_{p-1}\, a_pa_{p+1}\, a_{p+2} \ldots a_n = a_1a_2 \ldots a_{p-1}a_{p+1}a_pa_{p+2} \ldots a_n,$$

$$a_1a_2 \ldots a_{p-1}a_pa_{p+1}a_{p+2} \ldots a_na_{n+1} = a_1a_2 \ldots a_{p-1}a_{p+1}a_pa_{p+2} \ldots a_na_{n+1}.$$

2. $p=n$. Тогава $a_p=a_n$ и $a_{p+1}=a_{n+1}$. На основание на съдружителното свойство имаме

$$a_1a_2 \ldots a_na_{n+1} = a_1(a_2a_3 \ldots a_na_{n+1})$$

и по предположение можем да пишем

$$a_2a_3 \ldots a_na_{n+1} = a_2a_3 \ldots a_{n+1}a_n.$$

Следователно ще имаме

$$a_1a_2 \ldots a_na_{n+1} = a_1a_2 \ldots a_{n+1}\, a_n.$$

Така доказахме, че произведението на $n\geqq 2$ елемента не се изменя при разместване на два съседни множителя. Лесно се вижда, че произведението не се изменя и при разместване на два кои да е множителя a_r и a_s $(r<s)$ помежду им. Действително лесно се вижда, че разместването на тези два множителя помежду им може да се извърши с няколко размествания на съседни множители във въпросното произведение. Именно да разгледаме произведението

(3) $$a_ra_{r+1}a_{r+2} \ldots a_{s-2}a_{s-1}a_s$$

и да разместим a_r с a_{r+1}, след това a_r с a_{r+2} и т. н. до a_r с a_s. Предното произведение става

$$a_{r+1}\, a_{r+2} \ldots a_{s-2}\, a_{s-1}\, a_s\, a_r.$$

Като разместим a_s с предшествуващия множител a_{s-1}, след това a_s с a_{s-2} и т. н. до a_s с a_{r+1}, стигаме до произведението

(4) $$a_s\, a_{r+1}\, a_{r+2} \ldots a_{s-2}\, a_{s-1}\, a_r,$$

което се получава направо от (3) с разместване на a_r и a_s помежд им. Следователно произведенията ( ) и (4) са равни. Но видяхме рано, че всяко разместване на няколко елемента може да се извърши с последователно разместване на по два елемента. С това установихме, че произведението на n множителя не се променя при произволното им разместване, което трябваше да се докаже.

При адитивното написване на действията при абелевите групи предните свойства се отнасят за сума от няколко събираеми. Доказателството на въпросните свойства е напълно също, като сменим знака на произведение със знака + за сумиране.

Произведението на n еднакви множителя се нарича степен, т. е.

$$\overbrace{a \cdot a \ldots a}^{n \text{ пъти}} = a^n.$$

Отрицателна степен дефинираме както при числата $a^{-n} = (a^{-1})^n$ и приемаме $a^0 = 1$, като с 1 сме означили единицата на групата. За произволни цели числа n и m ще имаме тогава

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

Тези равенства се установяват, както в елементарната алгебра.

2. Подгрупи. Едно множество от поне един елемент, което е подмножество на една група, се нарича комплекс. Ако комплексът от своя страна е група, то той се нарича подгрупа на дадената група. Всяка група е същевременно подгрупа на себе си. Ако подгрупата $\mathbf{G}_1$ на групата $\mathbf{G}$ не се слива с нея, то тя се нарича същинска подгрупа.

Условията един комплекс $\mathbf{G}_1$ да бъде подгрупа на една група $\mathbf{G}$ могат да се опростят. Така условието II е изпълнено за елементите на $\mathbf{G}_1$, понеже то е изпълнено за всички елементи от $\mathbf{G}$. Постулатите III и IV изискват $\mathbf{G}_1$ да съдържа единицата и всеки обратен елемент a^{-1} на елемента a от $\mathbf{G}_1$. От друга страна, изискването на единицата се явява като следствие, понеже $\mathbf{G}_1$, щом съдържа a и a^{-1}, ще съдържа $a \cdot a^{-1} = e$. Следователно ще имаме:

Необходимо и достатъчно условие, щото непразното множество $\mathbf{G}_1$, принадлежащо на дадена група $\mathbf{G}$, да бъде пак група, се състои в следното:

1. $\mathbf{G}_1$ съдържа произведението ab на всеки свой елемент a и b.
2. $\mathbf{G}_1$ съдържа заедно с всеки свой елемент a и обратния му a^{-1}.

Условията 1 и 2 могат да се обединят в едно $\mathbf{G}_1$ съдържа заедно с a и b също ab^{-1}. Действително тогава $\mathbf{G}_1$ съдържа заедно с елемента a също и елемента $aa^{-1} = e$, т. е. единицата, и на основание на това ще съдържа и елемента $ea^{-1} = a^{-1}$. Ако a и b са произволни елементи от $\mathbf{G}_1$, то $\mathbf{G}_1$ съдържа b^{-1} и следователно $a(b^{-1})^{-1} = ab$ ще бъде елемент пак от $\mathbf{G}_1$. Ако множеството $\mathbf{G}_1$ е крайно, то условието 2 се явява излишно, понеже в такъв случай III и IV могат да се заместят с VI, а последното е сигурно изпълнено за $\mathbf{G}_1$, понеже е в сила за $\mathbf{G}$.

Например ако a е елемент от една група **G**, то степените

$$(5) \qquad \ldots, a^{-2}, a^{-1}, a^0, a, a^2, a^2, \ldots$$

образуват подгрупа (наречена циклична) на групата **G**. Действително за два кои да е елемента от редицата (5) имаме

$$a^p a^q = a^{p+q}.$$

Може да имаме два случая относно редицата (5). Или всичките елементи от нея са все различни и тогава цикличната група е безкрайна, или има поне два равни елемента:

$$a^m = a^k, \quad m > k.$$

Но тогава ще имаме $a^{m-k} = e$. Нека n е най-малката положителна степен, за която имаме $a^n = e$. Тогава елементите

$$(6) \qquad a^0 = e, a, a^2, \ldots, a^{n-1}$$

ще бъдат все различни. Защото в противен случай от равенството

$$a^p = a^q, \quad p > q, p, q \leqq n-1,$$

би следвало, че $a^{p-p} = e$, $p-q < n$, което противоречи на предположението относно n. Но тогава ще имаме

$$a^{n+1} = a, a^{n+2} = a^2, \ldots, a^{-1} = a^{n-1}, a^{-2} = a^{n-2}, \ldots$$

и в редицата различни са само елементите от редицата (6), т. е. цикличната група е крайна от ред n.

Субституциите от n-те елемента 1, 2, 3, . . . , n, които са произведение на четен брой транспозиции, образуват група, която е подгрупа на симетричната група, понеже произведението на такива субституции е пак субституция, разлагаща се на четен брой транспозиции. Тази група се нарича алтернативна група. Редът ѝ е равен на $\frac{n!}{2}$.

Нека **G** е една група и **A** и **B** са два произволни комплекса, принадлежащи на **G**. Под произведение **AB** на два комплекса се разбира множеството от всичките произведения ab, където a и b са съответно елементи на **A** и **B**. За всеки три комплекса ще имаме очевидно

$$(\mathbf{AB})\,\mathbf{C} = \mathbf{A}\,(\mathbf{BC}) = \mathbf{ABC}$$

и следователно скобите при умножение могат да се изпускат. В случай, че комплексът **A** е група, ще имаме $\mathbf{AA} = \mathbf{A}$.

Нека **A** и **B** са подгрупи на групата **G**. Да изследваме в кой случай комплексът **AB** е група. Ако a и b са елементи съответно от **A** и **B**, обратният елемент на ab е равен на $b^{-1}a^{-1}$ и трябва да принадлежи също на **AB**. Следователно трябва да имаме равенството

$$\mathbf{AB} = \mathbf{BA}.$$

Това равенство е и достатъчно. Действително от него следва

$$\mathbf{AB}.\mathbf{AB}=\mathbf{A}(\mathbf{BA})\mathbf{B}=\mathbf{A}(\mathbf{AB})\mathbf{B}=(\mathbf{AA})(\mathbf{BB})=\mathbf{AB},$$

което показва, че **AB** е група, която очевидно ще е подгрупа на **G**. Така установихме предложението: произведението на две подгрупи на групата **G** е само тогава също подгрупа, ако дадените подгрупи са комутативни помежду си. В случай, че групата **G** е абелева, следва, че произведението на всеки две подгрупи е пак подгрупа.

Нека **A** е подгрупа на групата **G** и g е един елемент от **G**. Комплексът $\mathbf{A}g$ се нарича десен съседен на **A** или десен съседен клас и комплексът $g\mathbf{A}$ се нарича ляв съседен на **A** или ляв съседен клас на **A**. Ако h е някой друг елемент от **G** или равен на g, то комплексът $g\mathbf{A}h$ се нарича двустранен на **A**. За всяка група **G** очевидно ще имаме

$$g\mathbf{G}=\mathbf{G}g=g\mathbf{G}h,$$

така че групата е на себе си десен, ляв и двойностранен комплекс.

Ще се ограничим в разглежданията по-нататък на десните съседни комплекси. Доказаните предложения се непосредствено пренасят и за левите такива.

Два съседни комплекса $\mathbf{A}g$ и $\mathbf{A}h$ само тогава съвпадат, ако hg^{-1} е елемент от подгрупата **A**. Това се вижда лесно от равенствата

$$\mathbf{A}g=\mathbf{A}h,\quad \mathbf{A}hg^{-1}=\mathbf{A},\quad \mathbf{A}h=\mathbf{A}hg^{-1}\quad g=(\mathbf{A}hg^{-1})g=\mathbf{A}g.$$

Два различни съседни комплекса нямат общи елементи. Действително, ако $\mathbf{A}g$ и $\mathbf{A}h$ имат общ елемент, например $ag=bh$, то оттук получаваме $gh^{-1}=a^{-1}b$ и следователно gh^{-1} трябва да е елемент от **A**. Но от предното предложение следва тогава, че $\mathbf{A}g$ и $\mathbf{A}h$ съвпадат.

Така виждаме, че всеки елемент g принадлежи на един и само един съседен комплекс. Понеже $gg=g$, то въпросният комплекс е точно комплексът $\mathbf{A}g$. По този начин групата се разпада на класи от комплекси и всеки неин елемент принадлежи на една класа (комплекс). Ако подгрупата **A** е крайна, то комплексите имат еднакъв брой елементи, и то толкова, колкото е редът на **A**. Ако **A** е безкрайна, комплексите $\mathbf{A}g$ и $\mathbf{A}h$ като множества са равномощни, понеже между елементите им съществува взаимно еднозначно съответствие.

Броят на различните съседни комплекси по подгрупата **A** от групата **G** се нарича индекс на **A** в **G**. Ако групата **G** е крайна от ред n и групата **A** е от ред m, то за индекса ѝ i ще имаме

$$n=mi.$$

Така установяваме следната теорема на Лагранж:

1. *Редът на всяка подгрупа на дадена група дели реда на групата.* В частност редът на всеки елемент от групата е делител на реда на групата.

Ако $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ са комплекси от една група, то под сума на тези комплекси, която означаваме с $\mathbf{A}_1+\mathbf{A}_2+\ldots+\mathbf{A}_k$, разбираме комплексът, съставен от елементите им, като всеки елемент е взет еднократно.

Нека **G** е една група и **A** е нейна подгрупа. Ако **A** не се слива с **G**, то ще има елемент g_1 от групата **G**, който не принадлежи на **A**. Но комплексът $\mathbf{A}g_1$ е съставен от елементи, различни от елементите на **A**, и принадлежи на групата **G**. Ако така се изчерпват всичките елементи на **G**, то ще имаме $\mathbf{G}=\mathbf{A}+\mathbf{A}g_1$. Ако това не е в сила, ще има елемент g_2 от **G**, който не принадлежи на **A** и $\mathbf{A}g_1$, и ще получим нов комплекс $\mathbf{A}g_2$, различен от първите два и принадлежащ на групата **G**. Като продължаваме така, достигаме до следния резултат:

Ако **G** е крайна група и **A** е нейна подгрупа, то **G** се разлага на сумата от комплекси

$$\mathbf{G}=\mathbf{A}+\mathbf{A}g_1+\mathbf{A}g_2+\ldots+\mathbf{A}g_{m-1}. \tag{7}$$

Тук $g_1, g_2, \ldots, g_{m-1}$ са елементи от **G**, които не принадлежат на **A**. Подобно разлагане имаме и по леви комплекси

$$\mathbf{G}=\mathbf{A}+t_1\mathbf{A}+t_2\mathbf{A}+\ldots+t_{m-1}\mathbf{A}. \tag{8}$$

Елементите $g_1, g_2, \ldots, g_{m-1}$ и $t_1, t_2, \ldots, t_{m-1}$ в (7) и (8) не са еднозначно определени. Обаче съседните комплекси на **A** в (7) и в (8) са еднозначно определени до реда им. Действително, ако имаме ново разлагане

$$\mathbf{G}=\mathbf{A}+\mathbf{A}g_1'+\mathbf{A}g_2'+\ldots+\mathbf{A}g_{m-1}', \tag{9}$$

трябва поне един елемент от $\mathbf{A}g_\mu'$ да влиза в един съседен комплекс $\mathbf{A}g_\nu$. От това следва, че ще имаме $\mathbf{A}g_\mu'=\mathbf{A}g_\nu$ и комплексите в (9) ще съвпадат с комплексите в (7).

От (7) и (8) се получават лесно разлаганията

$$\mathbf{G}=\mathbf{A}+g_1^{-1}\mathbf{A}+g_2^{-1}\mathbf{A}+\ldots+g_{m-1}^{-1}\mathbf{A}, \tag{10}$$

$$\mathbf{G}=\mathbf{A}+\mathbf{A}t_1^{-1}+\mathbf{A}t_2^{-1}+\ldots+\mathbf{A}t_{m-1}^{-1}. \tag{11}$$

Действително десните части на формулите (10) и (11) съдържат само елементи от **G**. Представянията ще бъдат установени, ако докажем, че съседните комплекси са различни помежду си. Наистина, ако имаме

$$g_\mu^{-1}\mathbf{A}=g_\nu^{-1}A,$$

то с умножаване вляво с g_μ получаваме

$$\mathbf{A}=g_\mu g_\nu^{-1}\mathbf{A}.$$

Следователно $g_\mu g_\nu^{-1}$ ще бъде елемент от **A**, отгдето следва, че ще имаме

$$\mathbf{A}=\mathbf{A}\,g_\mu g_\nu^{-1}.$$

Като умножим отдясно с g_ν, получаваме

$$\mathbf{A}g_\nu=\mathbf{A}g_\mu,$$

което противоречи на факта, че съседните комплекси са различни. По същия начин се установява, че

$$\mathbf{A}t_{\mu}^{-1} \neq \mathbf{A}t_{\nu}^{-1}, \quad \mu \neq \nu.$$

Ако **A** е подгрупа на групата **G** и за всеки елемент a от **G** имаме $a\mathbf{A} = \mathbf{A}a$, то **A** се нарича *инвариантна подгрупа на* **G** или *нормален делител.* Очевидно всяка подгупа на една абелева група е нормален делител на нея. Елементите на групата **G**, които са комутативни с всеки неин елемент, образуват подгрупа **A**, която е нормален делител на **G**. Действително нека a и b са два елемента от **A**. От равенствата $a\mathbf{G} = \mathbf{G}a$ и $b\mathbf{G} = \mathbf{G}b$ получаваме

$$ab\mathbf{G} = a(\mathbf{G}b) = \mathbf{G}\,a\,b,$$

т. е. ab принадлежи също на **G**. Следователно **A** е група и понеже за всеки елемент a от **G** ще имаме $a^{-1}\mathbf{A}a = \mathbf{A}$, то **A** е инвариантна подгрупа на **G**, която се нарича и *център* на групата **G**.

Ако $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p$ са комплекси от една група **G**, с $(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p)$ означаваме комплекса, съставен от общите елементи на комплексите $\mathbf{A}_1, \mathbf{A}_2, \ldots \mathbf{A}_p$.

Ако $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots, \mathbf{A}_p$ *са групи, то и* $\mathbf{A}' = (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p)$ *е група.* Действително, ако елементите a и b принадлежат на $\mathbf{A}'$, то и елементът ab ще принадлежи на $\mathbf{A}'$, понеже принадлежи на всичките групи $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p$.

2. *Ако* **A** *е група от ред* g, **B** е *група от ред* h *и ако групата* **(A, B)** *е от ред* k, *то комплексът* **AB** *съдържа точно* $\frac{gh}{k}$ *различни елемента.*

Действително нека

$$\mathbf{A} = \sum_{i=1}^{p} a_i, \qquad \mathbf{B} = \sum_{j=1}^{h} b_j.$$

Комплексът **AB** съдържа елементите

$$a_i b_j, \quad i = 1, 2, \ldots, g; \quad j = 1, 2, \ldots, h.$$

Ако докажем, че тези елементи са по k равни, то теоремата ще бъде установена. От равенството

$$a_i\, b_j = a_q\, b_p$$

с умножаване вляво на a_q^{-1} и вдясно на b_j^{-1} получаваме

$$a_q^{-1}\, a_i = b_p\, b_j^{-1}.$$

Елементът вляво на това равенство принадлежи на **A**, а елементът вдясно на **B**. Следователно ще имаме

$$a_q^{-1}\, a_i = c, \quad b_p\, b_j^{-1} = c,$$

където c е елемент от **(A, B)**. От предните равенства получаваме

$$a_q = a_i\, c^{-1}, \quad b_p = c\, b_j.$$

Обратно, ако c е произволен елемент от (**A**, **B**), то елементите

$$a_i c^{-1}, \quad c b_j$$

принадлежат съответно на **A** и **B** и тяхното произведение е равно на $a_i b_j$. Така съответно на k-те елемента на (**A**, **B**) се получават k такива представяния.

3. *Ако две инвариантни подгрупи* **A** *и* **B** *на* **G** *освен е нямат никой общ елемент, то всеки елемент на* **A** *е комутативен с всеки елемент на* **B**.

Нека a е от **A**, а b—от **B**. Тогава $b^{-1} ab$ е елемент от **A**, следователно и $b^{-1} a^{-1} b$ е пак от нея, понеже е инверзен елемент на предния. Оттук следва, че

$$u = (b^{-1} a^{-1} b) a$$

принадлежи на **A**. Аналогично се вижда, че

$$u = b^{-1} (a^{-1} ba)$$

принадлежи на **B** и понеже

$$(\mathbf{A}, \mathbf{B}) = \mathbf{e},$$

то $u = e$ и тогава от

$$b^{-1} a^{-1} ba = e$$

с умножаване наляво отначало с b и после с a получаваме

$$a^{-1} ba = be = b, \quad ba = ab,$$

с което теоремата е установена.

4. *Ако* **A** *и* **B** *са подгрупи на* **G** *и ако една поне е инвариантна, то* **AB** = **BA** *и този комплекс е група. Ако и двете са инвариантни, то* **AB** *е или равна на* **G**, *или е една инвариантна подгрупа.*

Ако **A** е инвариантна подгрупа на **G** и g е елемент от **G**, то $\mathbf{A}g = g\mathbf{A}$. Понеже това е валидно за всеки елемент g от **G**, ще следва търсената релация **AB** = **BA**. Установихме, че **AB** е група.

Ако **A** и **B** са инвариантни подгрупи и g е елемент от **G**, то ще имаме

$$g^{-1} \mathbf{A} g = \mathbf{A}, \quad g^{-1} \mathbf{B} g = \mathbf{B},$$

откъдето следва

$$g^{-1} (\mathbf{AB})\, g = g^{-1} \mathbf{A} g \,.\, g^{-1} \mathbf{B} g = \mathbf{AB},$$

с което теоремата е напълно доказана.

3. Изоморфизъм и хомоморфизъм. Две групи **G** и **G′** се наричат изоморфни, ако могат да се изобразят еднозначно една в друга, т. е. на всеки елемент a от **G** да отговаря определен елемент a' от **G′**, и то еднозначно обратимо, и ако a' и b' са съответетствуващи елементи от **G′** на елементите a и b от **G**, то на елемента ab да отговаря елементът $a'b'$. Изоморфизмът на групите се означава с $\mathbf{G} \cong \mathbf{G}'$. Ако групите **G** и **G′** съвпадат, изоморфизмът се нарича автоморфизъм.

Следователно автоморфизъм имаме, когато на всеки елемент a от една група **G** отговаря взаимно еднозначно елементът a' също от **G** менатото свойство за отговаряне на произведението на два кои да е елемента. Свойството изоморфизъм е преносимо. Именно лесно се че от изоморфизма на групите **G** и **G'** и на групите **G'** и **G''** следва изоморфизма на групите **G** и **G''**. Автоморфизмът може да се разглежда като еднозначно преобразуване на групата в себе си. Множеството от всички такива преобразувания очевидно ще образува група, наречена група на автоморфизма.

Ще разгледаме един случай на автоморфизъм, имащ приложение в други въпроси. Нека a е един произволен елемент от една група **G**, който ще бъде фиксиран при нататъшните разглеждания. На всеки елемент x от **G** ще съответствува елементът

$$x = a^{-1} xa \tag{1}$$

също от **G**. Лесно е да се види, че така получаваме един автоморфизъм на групата **G**. Първо, уравнението е еднозначно решимо спрямо x:

$$x = a\bar{x}\, a^{-1}.$$

Второ, за кои да е два елемента $\bar{x}$ и $\bar{y}$ от **G** имаме

$$\bar{x}\bar{y} = a^{-1} xa \,.\, a^{-1} ya = a^{-1}(xy)\, a.$$

Елементът $a^{-1} xa$ се нарича получен от x при трансформацията посредством елемента a. x и $a^{-1} xa$ се спрегнати елементи изъм се нарича вътрешен на групата. автоморфизми (ако има такива) наричат външни автоморфизми.

С прилагане на вътрешния автоморфизъм $x \to a^{-1} xa$ $\mathbf{G}_1$ преминава в подгрупата $a^{-1}\mathbf{G}_1 a$, която се нарича спрегната на $\mathbf{G}_1$. Ако групата $a^{-1}\mathbf{G}_1 a$ съвпада с $\mathbf{G}_1$ за всеки елемент $a \in \mathbf{G}$, то ще имаме $a\mathbf{G}_1 = \mathbf{G}_1 a$ за всеки елемент a от **G**. Следователно $\mathbf{G}_1$ ще бъде инвариантна подгрупа на **G**.

Една група $\bar{\mathbf{G}}$ се нарича хомоморфна с групата **G** или хомоморфен образ на тази група, ако на всеки елемент a от **G** отговаря един и само един елемент от $\bar{\mathbf{G}}$ и всеки елемент от $\bar{\mathbf{G}}$ е образ на поне един елемент от **G**. При това на произведението ab на два кои да е елемента от **G** отговаря произведението ab на съответните елементи $\bar{a}$, $\bar{b}$ от $\bar{\mathbf{G}}$. Това съотношение между групите се нарича хомоморфизъм и се отбелязва с $\mathbf{G} \sim \bar{\mathbf{G}}$. Ако групата $\bar{\mathbf{G}}$ принадлежи на **G**, то хомоморфизмът се нарича ендоморфизъм. Ако хомоморфизмът е в сила и в обратна посока, то той се свежда на изоморфизъм. Елементите от **G**, които имат за съответствуващ един кой да е елемент a от $\bar{\mathbf{G}}$, могат да бъдат групирани в една класа. Така групата **G** се разпада на класи.

Класът **e** от групата **G**, който при хомоморфизма има за съответен елемент единицата $\bar{e}$ от групата $\bar{\mathbf{G}}$, е инвариантна подгрупа на **G** и другите класове са съседни комплекси на тази подгрупа.

Наистина нека елементите a и b принадлежат на множеството **e**. Тогава на произведението им ab ще съответствува в групата **G** елементът $\bar{e}^2=\bar{e}$, т. е. ab е също елемент от **e**, което показва, че **e** група, като вземем още пред вид, че a^{-1} преминава в елемента $\bar{e}^{-1}=\bar{e}$. Елементите от всеки съседен ляв комплекс ae преминават в един и същ елемент $\bar{a}\,\bar{e}=\bar{a}$. Нека да разгледаме обратния въпрос. Именно да означим с a' елемент от **G**, който преминава в $\bar{a}$. С x да означим елемента, който удовлетворява на равенството $ax=a'$. Но тогава ще имаме $\bar{a}\bar{x}=\bar{a}$, т. е. $\bar{x}=\bar{e}$. Последното равенство показва, че x е елемент от **e** и следователно a' е елемент от ae.

Така виждаме, че класът от **G**, съответствуващ на елемента $\bar{a}$, е точно левият съседен комплекс $a\mathbf{e}$. По напълно същия начин установяваме, че съответствуващият клас на елемента $\bar{a}$ е десният съседен комплекс $\mathbf{e}a$. От предното заключаваме, че комплексите $a\mathbf{e}$ и $\mathbf{e}a$ са равни, т. е. **e** е инвариантна подгрупа на групата **G**.

4. Факторни групи. Нека **G** е една крайна група и **A** е една инвариантна подгрупа и да имаме

$$\mathbf{G}=\mathbf{A}+\mathbf{A}g_2+\ldots+\mathbf{A}g_r. \tag{1}$$

Лесно ще видим, че понеже **A** е инвариантна подгрупа, произведението на два десни комплекса е пак десен комплекс. Действително имаме

$$(\mathbf{A}g_p)(\mathbf{A}g_q)=\mathbf{A}(g_p\,\mathbf{A})\,g_q=\mathbf{A}(\mathbf{A}g_p)\,g_q=\mathbf{A}\,\mathbf{A}g_p\,g_q=\mathbf{A}g_p\,g_q.$$

Но $g_p g_q$ ще принадлежи на един десен комплекс $\mathbf{A}g_s$:

$$(\mathbf{A}g_p)(\mathbf{A}g_q)=\mathbf{A}g_s. \tag{2}$$

Да означим с

$$U_1=\mathbf{A},\quad U_2=\mathbf{A}g_2,\ldots,U_r=\mathbf{A}g_r$$

десните комплекси, които да разглеждаме като елементи на една съвкупност K. Ще установим, че K е една абстрактна група. По (2) виждаме, че произведението на два елемента от K принадлежи пак на K. Очевидно инвариантната подгрупа **A** е единицата на K и всеки елемент $\mathbf{A}g_i$ има обратен елемент $g_i^{-1}\mathbf{A}$. Остава да се докаже, че постулатът VI е изпълнен. Действително, ако $U_i\neq U_j$, то елементите g_i и g_j принадлежат на различни десни комплекси и същото е вярно за произведенията tg_i и tg_j на елементите g_i и g_j с един произволен елемент t от **G**. Понеже, ако тези две произведения са елементи на един комплекс $\mathbf{A}g_k$, то бихме имали равенствата

$$tg_i=vg_k,\quad tg_j=v'g_k,$$

където v и v' са елементи от **A**. Оттук ще имаме

$$g_l = t^{-1} v g_k = v''(t^{-1} g_k),$$
$$g_j = t^{-1} v' g_k = v'''(t^{-1} g_k),$$

където v'' и v''' са елементи от **A**. Но тогава би следвало, че g_l и g_j са елементи от един и същ десен комплекс, което е противоречие.

Следователно така доказахме, че K е група, която се бележи с **G**/**A** и се нарича **факторна група** на **G** и **A**.

Ако на елементите на комплекса $\mathbf{A}g_1$ направим да съответствува елементът U_1 на **G**/**A**, на елементите $\mathbf{A}g_2$ — елементът U_2 и т. н., то така получаваме едно изобразяване на групата **G** върху групата **G**/**A**, което е 1—μ—**хомоморфно**, като μ е редът на **A**. Действително на елементите от $\mathbf{A}g_i$ и $\mathbf{A}g_j$ съответствуват U_i и U_j и на произведението им по (2) ще съответствува Ug, определен с $g_i g_j = g_q$.

Ако u' е произволен елемент от **A**, *ще установим, че от* (2) *следва и*

$$(3)\qquad \begin{aligned}(\mathbf{A}g_p)(u'g_q) &= \mathbf{A}g_s\\ (u'g_p)(\mathbf{A}g_q) &= \mathbf{A}g_s\end{aligned}$$

Достатъчно е да докажем първото равенство. Имаме

$$(\mathbf{A}g_p)(u'g_q) = (g_p\,\mathbf{A})(u'\,g_q) = g_p(\mathbf{A}u')\,g_q = g_p(\mathbf{AA})\,g_q =$$
$$= (\mathbf{A}g_p)(\mathbf{A}g_q) = \mathbf{A}g_s.$$

5. *Всяка подгрупа на факторната група с индекс i се състои от съседни комплекси, на които елементите образуват една подгрупа на цялата група* **G** *с индекс i.*

Нека подгрупата на факторната група се състои от комплексите **A**, $\mathbf{A}g_2, \ldots \mathbf{A}g_k$. По предположение произведението на два кои да са от тези комплекси е пак такъв комплекс. Значи произведението на два кои да са елемента от комплекса $\mathbf{A}+\mathbf{A}g_2+\ldots+\mathbf{A}g_k$ е пак елемент от него, т. е. той образува група, която е подгрупа на **G** с индекс *i*. Лесно се вижда, че имаме обратно. Всяка подгрупа **H** на **G**, която съдържа една инвариантна подгрупа **A** на **G**, принадлежи на една подгрупа на факторната група **G**/**A**. Имаме по-прецизната теорема:

6. *Всяка инвариантна подгрупа на факторната група* **G**/**A** дава и. пг. (инвариантна подгрупа) на **G**. Обратно, на всяка и. пг. на **G**, която съдържа **A**, отговаря една и. пг. на факторната група **G**/**A**.

Нека съседните комплекси **A**, $\mathbf{A}g_2, \ldots, \mathbf{A}g_k$ да образуват една инвариантна подгрупа на **G**/**A**.

Тогава за произволен съседен комплекс $\mathbf{A}h$ ще имаме

$$\mathbf{A}h^{-1}\mathbf{A}g_i\mathbf{A}h = \mathbf{A}g_j, \qquad i, j = 1, 2, \ldots, k,\ g_1 = e.$$

Понеже $\mathbf{A}h = h\mathbf{A}$, то следва оттук, че комплексът

$$h^{-1}\mathbf{A}g_i h$$

има само елементи от $\mathbf{A}g_j$, т. е. се слива с него. Следователно комплексът

$$\mathbf{A}+\mathbf{A}g_2+\ldots+\mathbf{A}g_k$$

е група, комутативна на всеки елемент от **G**, т. е. е една и. пг.

Обратно, нека

$$\mathbf{R}=\mathbf{A}+\mathbf{A}g_2+\ldots+\mathbf{A}g_k$$

е и. пг. на **G**, която съдържа **A**. Тогава на всеки елемент t от **G** ще имаме

$$t^{-1}\mathbf{R}t=\mathbf{R}$$

и следователно елементите на $t^{-1}\mathbf{A}g_it$ лежат пак в **R**. Трябва да се докаже, че **R**/**A** е и. пг. на **G**/**A**, т. е. за произволно t съседният комплекс $\mathbf{A}t^{-1}\mathbf{A}g_i\mathbf{A}t$ се намира пак в **R**. Понеже $\mathbf{A}t=t\mathbf{A}$, то ще имаме

$$\mathbf{A}t^{-1}\mathbf{A}g_i\mathbf{A}t=\mathbf{AAA}\,t^{-1}g_it=\mathbf{A}\,t^{-1}g_it.$$

Понеже с g_i също $t^{-1}g_it$ принадлежи на **R**, то $\mathbf{A}t^{-1}g_it$ е също един съседен комплекс, който се намира в **R**, с което теоремата се доказва. При дефиниране на факторната група излизахме от едно номериране на съседните комплекси. При друго номериране ще получим факторна група, изоморфна на първата, така че от гледището на абстрактната теория тя е еднаква с нея.

5. Ред на разлагане на една група. Теорема на Жордан — Холдер. Една инвариантна подгрупа **A** на **G** се нарича максимална, ако няма друга инвариантна подгрупа **H** на **G**, така че **A** да бъде инвариантна подгрупа на **H**. Накратко ще пишем, че **A** е м. и. пг. на **G**, а инвариантните подгрупи ще означаваме с и. пг., както вече приехме.

Една и. пг. **A** на **G** е сигурно максимална, ако индексът ѝ е просто число. Защото, ако има и. пг. **B** съдържащи **A**, то индексът на **A** ще бъде равен на ij, гдето i е индексът на **A** спрямо **B**, а j — индексът на **B** спрямо **G**, т. е. няма да бъде просто число.

7. *Ако* **A** *и* **B** *са две различни м. и. пг. на* **G**, *то* $\mathbf{AB}=\mathbf{G}$. Понеже $\mathbf{A}\neq\mathbf{B}$, то има елемент v от **B**, който не се намира в **A**. Но v се намира в **AB** и следователно **A** не е идентично равна на **AB**, но е една истинска подгрупа, и то инвариантна, понеже е такава на **G**. Ако $\mathbf{AB}\neq\mathbf{G}$, то **A** е и. пг. на **AB**, която група по теорема 3 е инвариантна на **G**. Но това противоречи на условието, че **A** е м. и. пг. на **G**.

Една група **G** се нарича проста, ако освен **G** и **E** няма други инвариантни подгрупи. Основавайки се на теорема 6, веднага получаваме теоремата:

8. *Една и. пг.* **A** *на* **G** *е само тогава максимална, когато факторната група* **G**/**A** *е проста.*

Една друга важна теорема е следната:

9. *Ако* **A** *и* **B** *са две различни м. и. пг. на* **G** *и ако поставим* $\mathbf{M}=(\mathbf{A},\mathbf{B})$, то факторните групи **G**/**A** и **B**/**M** са равни помежду си; *също и факторните групи* **G**/**B** *и* **A**/**M**.

Комплексът **M** е група и понеже **A** и **B** са различни, той е една истинска подгрупа на тези две групи. Ако t е елемент от **G** и g е елемент от **M**, то

$$t^{-1}gt$$

е елемент от **G** и **B**, т. е. от **M**, понеже **A** и **B** са инвариантни подгрупи на **G**. С това е установено, че **M** е и. пг. на **G**.

Нека разлагането на **B** по **M** е

(4) $$\mathbf{B}=\mathbf{M}g_1+\mathbf{M}g_2+\ldots+\mathbf{M}g_i,$$

гдето $g_1=e$. Понеже по теорема 7 $\mathbf{AB}=\mathbf{G}$ ще имаме

$$\mathbf{G}=\mathbf{AB}=\mathbf{A}(\mathbf{M}g_1+\mathbf{M}g_2+\ldots+\mathbf{M}g_i)$$

или понеже **M** е подгрупа на **A**, т. е. $\mathbf{AM}=\mathbf{A}$,

$$\mathbf{G}=\mathbf{A}g_1+\mathbf{A}g_2+\ldots+\mathbf{A}g_i.$$

Това е точното разлагане на **G** по подгрупата **A**. Действително, ако два съседни комплекса са равни

$$\mathbf{A}g_p=\mathbf{A}g_q,$$

то $g_p g_q^{-1}$ ще бъде елемент от **A** и по причина на (4) също ще принадлежи на **B**, следователно ще принадлежи и на **M**; но тогова комплексите $\mathbf{M}g_p$ и $\mathbf{M}g_q$ ще бъдат идентични, което противоречи на разлагането (4).

Тогава, ако поставим

(5) $$(\mathbf{M}g_\mu)\ (\mathbf{M}g_\nu)=\mathbf{M}g_r,$$

ще имаме

(6) $$\mathbf{M}g_r=\mathbf{M}(g_\mu\mathbf{M})g_\nu=\mathbf{M}\,(\mathbf{M}g_\mu)g_\nu=(\mathbf{M}\,\mathbf{M})\,(g_\mu\,g_\nu)=\mathbf{M}g_\mu\,g_\nu$$

и с подобно пресмятане, като използуваме (6).,

$$(\mathbf{A}g_\mu)(\mathbf{A}g_\nu)=\mathbf{A}g_\mu\,g_\nu=(\mathbf{AM})\,g_\mu\,g_\nu=\mathbf{A}\,(\mathbf{M}g_\mu\,g_\nu)=\mathbf{AM}g_r=\mathbf{A}g_r.$$

От (5) следва следователно

$$(\mathbf{A}g_\mu)\,(\mathbf{A}g_\nu)=\mathbf{A}g_r$$

и оттук по дефиницията на факторна група е ясно, че

$$\mathbf{G}\,/\,\mathbf{A}=\mathbf{B}\,/\,\mathbf{M}.$$

Аналогично по симетрия

$$\mathbf{G}\,/\,\mathbf{B}=\mathbf{A}\,/\,\mathbf{M}.$$

Нека сега $\mathbf{G}_0$ е една произволна група. Да изберем една коя да е м. и. пг. $\mathbf{G}_1$, на тази отново една м. и. пг. $\mathbf{G}_2$ и т. н., докато достигнем до групата единица **E**. Така ще получим една редица от групи:

(7) $$\mathbf{G}_0,\ \mathbf{G}_1,\ \mathbf{G}_2,\ldots,\ \mathbf{G}_n,$$

всяка от които е м. и. пг. на предшествуващата и последната $\mathbf{G}_n=\mathbf{E}$. Групите (7) се наричат един ред на разлагане на $\mathbf{G}_0$. Принадлежащите факторни групи

(8) $$\mathbf{G}_0\,/\,\mathbf{G}_1,\ \mathbf{G}_1\,/\,\mathbf{G}_2,\ldots,\ \mathbf{G}_{n-1}\,/\,\mathbf{G}_n$$

по теорема 8 са все прости и нека редовете им са равни съответно на

(9) $$i_1,\ i_2,\ldots,i_n.$$

Числото i_s е индексът на $\mathbf{G}_s$ относно $\mathbf{G}_{s-1}$. По тази причина числата (9) се наричат *принадлежащ към реда на разлагане* (7) *индексен ред.*

Аналогично подбирайки и други м. и. пг., може да се получат други редове на разлагане. Обаче за всичките такива разлагания е валидна една основна теорема.

10. Теорема на Жордан-Холдер.[1] *При две произволни разлагания на една група факторните групи в единия ред на разлагане са равни на факторните групи на другия ред на разлагане, само че местата им в редовете на разлагане въобще не са едни и същи. Значи и принадлежащите им индексни редове, абстрахирайки се от мястото им, са еднакви.*

Ще докажем теоремата по индуктивен път, като приемем, че тя е вярна за групите, на които редът съдържа най-много $l-1$ прости делители (равни или не). Теоремата е очевидна, ако редът на групата $\mathbf{G}_0$ е просто число, понеже единствената подгрупа, различна от $\mathbf{G}_0$, може да бъде само $\mathbf{E}$ и така имаме само един ред на разлагане $\mathbf{G}_0$, $\mathbf{E}$.

Нека сега $\mathbf{G}_0$ е една група, на която редът съдържа l прости делители и така да имаме за нея два реда на разлагане:

$$\begin{matrix}\mathbf{G}_0,\ \mathbf{G}_1,\ \mathbf{G}_2,\ \ldots,\ \mathbf{G}_n,\\ \mathbf{G}_0,\ \mathbf{H}_1,\ \mathbf{H}_2,\ \ldots,\ \mathbf{H}_m.\end{matrix} \tag{10}$$

Ако $\mathbf{G}_1=\mathbf{H}_1$, то

$$\begin{matrix}\mathbf{G}_1,\ \mathbf{G}_2,\ \ldots,\ \mathbf{G}_n,\\ \mathbf{G}_1,\ \mathbf{H}_2,\ \ldots,\ \mathbf{H}_m\end{matrix} \tag{11}$$

са два реда на разлагане на $\mathbf{G}_1$ и редът $\mathbf{G}_1$ като делител на реда на $\mathbf{G}_0$ ще съдържа най-много $l-1$ прости делители. По предположение теоремата е вярна за разлагането (11). Оттук следва, че тя ще бъде вярна и за разлагането (10). Остава да се разгледа случаят, когато $\mathbf{G}_1$ и $\mathbf{H}_1$ са различни помежду си. Да поставим тогава $\mathbf{M}=(\mathbf{G}_1,\ \mathbf{H}_1)$. $\mathbf{M}$ е и. пг. на $\mathbf{G}_1$ и $\mathbf{H}_1$ и още по теорема 9

$$\mathbf{G}_0/\mathbf{G}_1=\mathbf{H}_1/\mathbf{M}. \tag{12}$$

Понеже $\mathbf{G}_1$ е м. и. пг. на $\mathbf{G}_0$, то $\mathbf{G}_0/\mathbf{G}_1$ по теоремата 8 е проста група и по (12) $\mathbf{H}_1/\mathbf{M}$ е проста, т. е. $\mathbf{M}$ е м. и. пг. на $\mathbf{H}_1$. Също $\mathbf{M}$ е м и. пг. на $\mathbf{G}_1$.

Да разгледаме едно разлагане на $\mathbf{M}$:

$$\mathbf{M},\ \mathbf{M}_1,\ \mathbf{M}_2,\ \ldots,\ \mathbf{M}_r. \tag{13}$$

Така ще имаме два реда на разлагане на групата $\mathbf{G}_1$:

$$\begin{matrix}\mathbf{G}_1,\ \mathbf{G}_2,\ \mathbf{G}_3,\ \ldots,\ \mathbf{G}_n,\\ \mathbf{G}_1,\ \mathbf{M},\ \mathbf{M}_1,\ \ldots,\ \mathbf{M}_r\end{matrix} \tag{14}$$

---

[1] Втората част на теоремата, която е непосредствено следствие от първата, е дадена от C. Jordan (1870), който е много развил теорията на групите, а първата — от O. Hölder (1897).

и понеже редът на G_1 има най-много $l-1$ прости делители, то отговарящите факторни групи на разлагането (14) са равни, абстрахирайки се от мястото им. Като прибавим в (14) отляво групата G_0, виждаме от това, че факторните групи на реда на разлагане

$$G_0, G_1, G_2, \ldots, G_n,$$

абстрахирайки се от мястото им, са следните:

$$(15) \qquad G_0/G_1, G_1/M, M/M_1, \ldots, M_{r-1}/M_r.$$

Също се получава, че факторните групи на реда за разлагане

$$G_0, H_1, H_2, \ldots, H_m,$$

абстрахирайки се от мястото им, са следните:

$$(16) \qquad G_0/H_1, H_1/M, M/M_1, \ldots, M_{r-1}/M_r.$$

Но групите (15) и (16) съвпадат, като първите две по (12) са разменени, с което теоремата е доказана.

Една група се нарича метациклична, ако индексите на разлагането ѝ са все прости числа.

6. Абелеви групи. Под абелева група G (както вече споменахме) разбираме такава група, на която елементите са комутативни. Ако следователно S и t са два кои да са елемента от нея, то $st=ts$. Следователно в едно произведение можем да разместваме местата на елементите, без то да се измени.

Всяка циклична група е абелева. Има обаче и други абелеви групи. Така например субституционната група

$$E+(12)(34)+(13)(24)+(14)(23)$$

е абелева, без да бъде циклична.

Всяка подгрупа на една абелева група е пак абелева. Това следва непосредствено от дефиницията. Ако A е подгрупа на абелевата група G и s е елемент от G, то ще имаме $As=sA$, което ни показва, че всяка подгрупа на една абелева група е инвариантна подгрупа.

Съществува по-обща теорема:

11. *Всяка крайна абелева група е метациклична, т. е. индексите на разлагането ѝ са все прости числа.*

Нека A да е подгрупа (следователно инвариантна) на G и нека s е елемент от G, който не се съдържа в A. С l да означим най-малкия положителен показател, за който s^l се съдържа в A. Такъв l имаме, понеже, ако k е редът на s, то $s^k=l$ се съдържа в A. Ако l е един прост делител на l, $l=mp$, то елементът

$$s^m=t$$

не се съдържа в A и t^p е най-малката степен на t, за която t^p се съдържа в A. Комплексът

$$H_1=A+At+At^2+\ldots+At^{p-1}$$

е една подгрупа на **G** и **A** е подгрупа на $\mathbf{H}_1$ с индекс p, който е просто число. Следователно **A** е максимална инвариантна подгрупа на $\mathbf{H}_1$.

Ако $\mathbf{H}_1 \neq \mathbf{G}$, то $\mathbf{H}_1$ е подгрупа на **G**, за която можем да приложим горните разсъждения, докато достигнем до групата **G**. Така можем да започнем с подгрупата единица **E** и да получим една редица от групи

$$\mathbf{E},\ \mathbf{H}_1,\ \mathbf{H}_2,\ldots,\ \mathbf{H}_k,$$

от които последната е **G**. Всяка група е максимална инвариантна подгрупа на следващата с индекс просто число, с което теоремата е установена.

Глава II

Пръстен и поле

1. Пръстен. В аритметиката и алгебрата се извършват действия с величини от разнообразен характер: цели числа, рационални числа, хиперкомплексни числа, матрици, полиноми и рационални функции и т. н. В елементарната геометрия се събират отсечки, ъгли, вектори. Явява се необходимостта да се даде общо определение на тези величини и да се изследват действията с тях.

Нека A е едно множество от някакви елементи, които означаваме с a, b, c и т. н. Броят им може да бъде краен или безкраен. Предполагаме, че можем да ги различаваме, като за два кои да е елемента a и b ще имаме съотношение на равенство $a=b$ или на неравенство $a \neq b$. Този постулат за равенство се състои в следните свойства: 1. Той е напълно определен, т. е. за кои да е два елемента a и b имаме $a=b$ или $a \neq b$. 2. Равенството е рефлексивно (възвратно), т. е. всеки елемент a е равен на себе си: $a=a$. 3. Равенството е симетрично (обратимо), т. е. от $a=b$ следва $b=a$. 4. Равенството е транзитивно (преносимо), т. е. от $a=b$, $b=c$ следва $a=c$.

В множеството A са дефинирани две свързвания на елементите му, като на всеки два елемента (равни или не) съответствува един елемент c, наречен сума на елементите a и b, който отбелязваме с $c=a+b$, и един елемент d, наречен произведение на елементите a и b, който отбелязваме с $d=ab$. Множеството се нарича п р ъ с т е н, ако при тези две свързвания (действия) са в сила следните закони:

$$a+b=b+a$$

комутативен (разместителен) закон,

$$(a+b)+c=a+(b+c), \quad (ab)c=a(bc)$$

асоциативен (съдружителен) закон,

$$(a+b)c=ac+bc, \quad c(a+b)=ca+cb$$

дистрибутивен (разпределителен) закон.

Ако за умножението е в сила и комутативният закон, то пръстенът се нарича к о м у т а т и в е н. Ние главно ще се занимаваме с такива пръстени.

От определението за събиране и законите за него е ясно, че пръстенът образува абелева група относно това действие. Тогава можем да приложим доказаните свойства за тези групи непосредствено и за пръстените. Получаваме, че съществува единствен елемент нула 0, притежаващ свойството

$$a+0=a$$

за всеки елемент a от A. Освен това за всеки елемент a съществува противоположен елемент $-a$ със свойството

$$-a+a=0.$$

Уравнението $a+x=b$ е разрешимо, и то еднозначно. Единственото му решение се дава с елемента

$$x=-a+b,$$

който ще означаваме и с $b-a$. Елементът $b-a$ се нарича разлика на елементите b и a. По силата на формулата

$$a-b=a+(-b)$$

всяка разлика се преобразува в сума и следователно за разликите ще имаме същите основни закони, както за сумите, например

$$(a-b)-c=(a-c)-b, \quad -(-a)=a, \quad a-a=0$$

и т. н.

Следвайки същия път на доказателство, както при теорията на групите, като заместим умножението със събирането, лесно виждаме, че асоциативният закон е верен за произволен брой събираеми и за произведение от произволен брой елементи. Също така комутативният закон остава в сила за сумата на произволен брой събираеми. Естествено разглеждаме само краен брой събираеми или множители.

По индуктивен път получаваме лесно равенствата

$$a(b_1+b_2+\ldots+b_n)=ab_1+ab_2+\ldots+ab_n,$$

$$(a_1+a_2+\ldots+a_n)\,b=a_1b+a_2b+\ldots+a_nb$$

и с повторно приложение на предните равенства получаваме правилото за умножение на суми:

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

$$=(a_1b_1+a_1b_2+\ldots+a_1b_m)+(a_2b_1+a_2b_2+\ldots+a_2b_m)+\ldots+$$

$$+(a_nb_1+a_nb_2+\ldots+a_nb_m)=\sum_{i=1,\ k=1}^{n,m} a_i\,b_k.$$

Дистрибутивният закон е в сила и за умножение на разлики, т. е. имаме

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

Това равенство следва от равенствата

$$a(b-c)+ac=a(b-c+c)=ab.$$

Специално имаме

$$a(a-a)=a\,.\,0=a\,.\,a-a\,.\,a=0.$$

Ако един от множителите е равен на нула, то произведението е равно на нула. По-нататък с примери ще видим, че може да намерим пръстени, в които обратното не е вярно, т. е. може да се случи да имаме $ab=0$, като елементите a и b са отлични от нула. В такъв случай наричаме елементите a и b делители на нулата. Елементът a е ляв делител на нулата, а елементът b е десен делител. Ако пръстенът е комутативен, то не правим естествено разлика между двата вида делители, понеже те съвпадат.

Ако в пръстени няма делители на нулата освен самата нула, т. е. ако от равенството $ab=0$ следва, че поне единият от множителите a и b е равен на нула, то пръстенът се нарича пръстен без делители на нулата. Ако освен това той е комутативен, то наричаме го област на цялостност.

Пръстенът може да не притежава единица. Ако един пръстен A притежава лява единица e, т. е. за всеки елемент x от него имаме $ex=x$, и една дясна единица e', $xe'=x$ за всеки елемент x, то тези единици трябва да съвпадат, понеже от $e.e'=e$ и $e'.e=e'$ следва,че $e=e'$.

Тогава всяка друга единица трябва също да съвпада с e. Такива пръстени наричаме пръстени с единица.

Нека A е пръстен с единица, която да означим с e. Ако a е произволен елемент от A и уравнението

$$xa=e$$

има решение x в A, то елементът x означаваме с $a_л^{-1}$ и го наричаме ляв обратен елемент на a. Ако уравнението

$$ay=e$$

има решение y в A, то $y=a_д^{-1}$ наричаме десен обратен елемент на a. В случай, че a притежава ляв и десен обратен елемент, то те съвпадат. В това се убеждаваме от следните равенства:

$$a_л^{-1}=a_л^{-1}\left(aa_д^{-1}\right)=\left(a_л^{-1}a\right)a_д^{-1}=a_д^{-1}.$$

В този случай казваме, че елементът a притежава обратен елемент, който отбелязваме с a^{-1}.

На основание на закона на асоциативност можем да определим степен a^n на всеки елемент за n естествено число и да установим валидността на познатите правила:

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

като последните две са за комутативни пръстени. Ако пръстенът притежава и единица e, то можем да положим $a^0=e$ и горните правила остават в сила при допълнителното предположение, че влизащите там елементи имат обратни, като n, m са произволни цели числа.

Ако n е произволно натурално число, то под na разбираме сумата

$$na=a+a+\ldots+a.$$

Очевидно ще имаме тогава

$$na+ma=(n+m)a, \quad n(a+b)=na+nb,$$

$$n.ma=nm.a \quad n.ab=na.b=a.nb.$$

Ако положим $(-n)a=-na$, то правилата са в сила за произволни цели числа n и m. Въобще числото n не е елемент от пръстена A. Ако обаче A притежава единица e, то na представлява произведение от елементи на A. Действително имаме

$$na=n.ea=ne.a,$$

като ne и a са елементи от A.

Да разгледаме няколко примера на пръстени. Множеството на всичките цели числа образува пръстен, който е комутативен с единица—числото 1, и с нула—числото 0. Множеството на всички четни числа е комутативен пръстен с нула, но не притежава единица. Множеството на всичките реални числа е също пръстен с единица и нула, както и множеството на всичките комплексни числа.

Нека A е множеството от всичките матрици

$$\begin{Vmatrix} a & b \\ c & d \end{Vmatrix},$$

където a, b, c, d са произволни цели числа или по-общо произволни реални или комплексни числа. По познатите ни правила за събиране и умножение на матрици виждаме, че тези множества са пръстени, които са некомутативни. Понеже

$$\begin{Vmatrix} a & 0 \\ 0 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 0 \\ b & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix},$$

то виждаме, че в тях имаме делители на нулата.

Ако елементите на един пръстен са числа — реални или комплексни, то той се нарича числов пръстен. Очевидно всеки числов пръстен няма делители на нулата.

2. Поле. Един пръстен се нарича тяло, ако съдържа поне един елемент, отличен от нула, и ако уравненията

$$ax=b,$$

$$ya=b$$

са решими за всяко $a \neq 0$. Ако освен това пръстенът е и комутативен, то той се нарича поле или област на рационалност. Както при групите от предните постулати следват свойствата:

1. Съществува лява единица e. Действително нека $a \neq 0$ е произволно избран елемент от A и да означим с $y=e$ решението на уравнението $ya=a$. За произволен елемент b от A ще имаме (z решение на $az=b$)

$$eb=eaz=az=b,$$

т. е. съществува лява единица e. Подобно се убеждаваме, че съществува дясна единица, т. е. съществува въобще единица.

2. От предното свойство следва, че за всеки елемент a съществува десен и ляв обратен елемент, т. е. съществува въобще обратен елемент a^{-1}.

Забележка. Тялото няма делители на нулата. Това лесно следва от равенството $ab=0$ с умножение на $a^{-1}\,(a \neq 0)$: $a^{-1}ab = eb = b = 0$.

Уравненията (1) са решими еднозначно. Действително да предположим, че първото уравнение има две решения x и x_1. Тогава от равенството $ax = ax_1$ с умножение на a^{-1} получаваме $x = x_1$. Като умножим първото уравнение от (1) с a^{-1} и второто с a^{-1}, получаваме

$$x = a^{-1}b,$$

$$y = ba^{-1}.$$

Ако тялото е комутативно, т. е. поле, то $a^{-1}b = ba^{-1}$ и вместо тези произведения пишем просто $\frac{b}{a}$.

Множеството от всичките рационални числа очевидно образува поле, което се нарича естествено поле или естествена област на рационалност. При полета, съставени от числа, дефиницията може да се опрости в следната форма: едно множество M от числа се нарича поле, ако съдържа поне едно число, отлично от нула, и сумата, произведението и частното (при делител, отличен от нула) на две кои да е числа от M принадлежат пак на M. Лесно се вижда тогава, че всяко числово поле K съдържа натуралното поле. Действително, ако $a \neq 0$ е число от K, то $\frac{a}{a} = 1$ трябва да бъде елемент от K.

Оттук всяко цяло положително число n ще е елемент от K, както и всяко отрицателно цяло число $-n$ и като следствие и всяко дробно число. Така виждаме, че числовите полета са с безбройно много елементи и всяко поле с краен брой елементи не е следователно числово поле.

Нека сега A е поле с краен брой елементи, на което единицата да бъде e.

Ако n е произволно естествено число, то видяхме, че ne са елементи, които принадлежат също на A. Понеже броят на елементите на A е краен, то следва, че имаме поне два равни елемента:

$$me = pe, \quad m > p.$$

Оттук получаваме равенството $(m-p)\,e = 0$, т. е. съществува естествено число n, за което имаме $ne = 0$. Ще докажем, че числото n е просто при предположение, че n е най-малкото естествено число, за което елементът ne е равен на нула. Ако a е произволен елемент от A, то имаме

$$na = n\,(ea) = (ne)\,a = 0.$$

Следователно за всеки елемент a имаме $na = 0$. Нека сега предположим, че n е съответно число, $n = n_1 n_2$. Но тогава ще имаме $ne = n_1 e n_2 e = 0$. Оттук следва, че поне един от елементите $n_1 e$, $n_2 e$ трябва

да е равен на нула, защото полето няма делители на нула. Но тогава идваме до противоречие с предположението ни за числото n.

Ако за единицата e на едно поле B съществува естествено число n със свойството $ne=0$ и n е най-малкото такова число, то полето се нарича поле с характеристика n. Полето може да бъде с краен или безкраен брой членове. Вземайки пред вид предните извеждания, виждаме, че числото n трябва да е просто и за всеки елемент a от B имаме $na=0$. Очевидно всяко числово поле е с характеристика нула.

Ще разгледаме сега пример на крайно поле. Нека n е произволно естествено число, отлично от 1. Ако m е произволно цяло число, то остатъкът от делението на това число с n ще бъде едно от числата $0, 1, 2, \ldots, n-1$. Да означим с A_k съвкупността от всичките цели числа, които имат при поменатото деление остатък, равен на k, $0 \leqq k \leqq n-1$. Очевидно това са числата от вида $\lambda n+k$, където λ е произволно цяло число. По този начин множеството от всичките цели числа се разделя на класи:

$$A_0, A_1, A_2, \ldots, A_{n-1}, \tag{2}$$

нямащи общи числа. Да означим с P множеството, на което елементите са класите (2). Ще покажем, че P е пръстен. За тази цел ще трябва да дефинираме събиране, изваждане и умножение в P. Нека A_i и A_j са два произволни елемента, равни или не, от (2). Ако a е число от класа A_i и b е число от класа A_j, то числото $a+b$ ще принадлежи на класа A_{i+j}, ако $i+j<n$, или на класа A_{i+j-n}, ако $i+j \geqq n$. На това основание ние дефинираме сума на елементите A_i и A_j от P с $A_i+A_j=A_{i+j}$ при $i+j<n$ и $A_i+A_j=A_{i+j-n}$ при $i+j \geqq n$. Нека сега r е остатъкът от делението на произведението на взетите по-горе числа a и b с числото n. Лесно се вижда, че остатъкът от делението на произведението на кое да е число a' от класа A_i с кое да е число b' от класа A_j е също равен на r. Именно числата a' и b' ще имат формата

$$a'=a+n\lambda, \quad b'=b+n\mu,$$

където λ и μ са цели числа. Но тогава за произведението им получаваме

$$a'b'=ab+n(a\mu+b\lambda+n\lambda\mu),$$

т. е. $a'b'=ab+nu$, където u е цяло число, с което твърдението е установено. Тогава под произведение на елементите A_i и A_j ще разбираме елемента A_r, $A_iA_j=A_r$. Нулевият елемент на P ще бъде очевидно елементът A_0. На A_k противоположният елемент е A_{n-k}. С горните разглеждания показахме, че P е пръстен, който очевидно е и комутативен. Ако n е съставно число, то P има делители на нулата. Действително нека $n=n_1n_2$, $0<n_1, n_2<n$. Тогава съгласно с определението за умножение ще имаме $A_{n_1}A_{n_2}=A_0$, т. е. A_{n_1} и A_{n_2} са делители на нулата A_0 на пръстена P.

Ако числото n е просто, то P е поле.

Нека A_m и A_q са произволни елементи на P и $A_q \neq A_0$. Трябва да покажем, че делението на A_m с A_q е възможно, т. е. съществува такъв елемент A_s от P, за който да имаме $A_q A_s = A_m$. Ако $A_m = A_0$, то и $A_s = A_0$. Нека тогава $A_m \neq A_0$. Понеже $A_q \neq A_0$, то q е едно от числата 1, 2, 3, . . . ,$n-1$. Да разгледаме сега редицата от числата

$$q,\ 2q,\ 3q,\ \ldots,\ (n-1)\,q. \tag{3}$$

Всички тези числа принадлежат на класи, различни от A_0, защото, ако някое от тях би принадлежало на класа A_0, то това число би трябвало да се дели на n, а това е невъзможно, понеже произведението на две естествени числа, по-малки от n, не може да се дели на простото число n. Освен това никои две числа от (3) не могат да принадлежат на една и съща класа, понеже разликата им би се делила на n, което на същото основание като по-горе е невъзможно. Следователно числата от редицата (3) принадлежат на класите A_1, A_2, . . ., A_{n-1}, разбира се, не непременно в същия ред. Но тогава в класа A_m ще има едно число sq от редицата (3) и можем да пишем $A_s A_q = A_m$, т. е. A_s е частното от делението на A_m с A_q. Така доказателството, че пръстенът P е и поле, се привършва.

3. Хомоморфизъм и изоморфизъм. Тези понятия са аналогични на въведените вече понятия в теорията на групите. Нека A и $\bar{A}$ са две множества с дефинирани в тях две композиции — събиране и умножение на елементите им. Казваме, че $\bar{A}$ е хомоморфен образ на A, ако са изпълнени следните условия:

1. На всеки елемент a от A съответствува елемент $\bar{a}$ от $\bar{A}$ и на всеки елемент $\bar{a}$ от $\bar{A}$ съответствува поне един елемент a от A.

2. За всеки елемент a и b от A от съответствието

$$a \longrightarrow \bar{a},\quad b \longrightarrow \bar{b}$$

следват съответствията $a+b \longrightarrow \bar{a}+\bar{b}$ и $ab \longrightarrow \bar{a}\,\bar{b}$. Това изобразяване се нарича хомоморфизъм. Означава се с $A \sim \bar{A}$. Ако на всяка $\bar{a}$ съответствува само по едно a, то изобразяването се нарича изоморфизъм и множествата A и $\bar{A}$ се наричат изоморфни. Означаваме изоморфизма с $A \cong \bar{A}$. Очевидно изоморфизмът на две множества е симетрично свойство и рефлексивно. От друга страна, ако множеството A е изоморфно с множеството A_1 и това множество е изоморфно с множеството A_2, то A е изоморфно множество с A_2, т. е. изоморфизмът е преносимо свойство.

Ако $A \sim \bar{A}$ и A е пръстен, то и $\bar{A}$ е пръстен.

Нека $\bar{a}$, $\bar{b}$, $\bar{c}$ са три кои да е елемента от $\bar{A}$. Да означим с a, b, c кои да е три елемента от A, на които образите в $\bar{A}$ са $\bar{a}$, $\bar{b}$, $\bar{c}$. Понеже A е пръстен, то имаме $a(b+c) = ab + ac$. На основание на условията 1 и 2 за хомоморфизъм получаваме, че $\bar{a}(\bar{b}+\bar{c}) = \bar{a}\,\bar{b} + \bar{a}\bar{c}$. Също така от $a+b = b+a$ следва, че $\bar{a}+\bar{b} = \bar{b}+\bar{a}$ от $a(bc) = (ab)\,c$

следва $\bar{a}(\bar{b}\bar{c})=(\bar{a}\bar{b})\bar{c}$ и т. н. За да решим уравнението $\bar{a}+\bar{x}=\bar{b}$, намираме елементи a и b от A, на които $\bar{a}$ и $\bar{b}$ са образи, и решаваме уравнението $a+x=b$ в A. Образът $\bar{x}$ на x е решение на даденото уравнение $\bar{a}+\bar{x}=\bar{b}$. На елемента нула от A отговаря също нулата от $\bar{A}$. На противоположния елемент — a на a в A отговаря противоположният елемент — $\bar{a}$ на $\bar{a}$ в $\bar{A}$, на единичния елемент e от A отговаря единичният елемент $\bar{e}$ от $\bar{A}$. Тези свойства следват непосредствено от съответствията: 1. От $a+0=a$ следва $\bar{a}+\bar{0}=\bar{a}$. 2. От $a-a=0$ следва $\bar{a}-\bar{a}=\bar{0}$. 3. От $ae=a$ следва $\bar{a}\bar{e}=\bar{a}$.

Ако пръстенът A е комутативен, то и пръстенът A е също такъв. В частния случай на изоморфизъм имаме и следните свойства:

Ако A е област на цялостност, то и $\bar{A}$ е такава област. Ако A е поле, то и $\bar{A}$ е поле.

4. Поле от отношения. Нека R е комутативен пръстен, елементите на който принадлежат на едно тяло A. Тогава уравненията

$$bx=a, \quad yb=a \quad (b\neq 0) \tag{1}$$

ще имат решения в A, които се дават с формулите $x=b^{-1}a$, $y=ab^{-1}$. Но от равенството $ab=ba$ с умножение вляво и дясно с b^{-1}, a^{-1} получаваме $b^{-1}a=ab^{-1}$ и следователно уравненията (1) имат едно единствено решение, което означаваме с $\frac{a}{b}$. Лесно се вижда, че правилата за действията с обикновените дроби остават в сила и за символите $\frac{a}{b}$. Именно ще имаме

$$\text{I. } \frac{a}{b}=\frac{c}{d} \text{ само тогава, когато } ad=bc$$

$$\text{II. } \frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}. \tag{2}$$

$$\text{III. } \frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}. \qquad \text{IV. } \frac{\frac{a}{b}}{\frac{c}{d}}=\frac{ad}{bc}.$$

Доказателството на горните равенства се извършва с Така, като умножим равенството $ab^{-1}=cd^{-1}$ с bd,

$$ab^{-1}bd=cd^{-1}bd, \quad ad=cb.$$

Обратно, от $ad=bc$ с умножение на $b^{-1}d^{-1}$ получаваме

$$adb^{-1}d^{-1}=bcb^{-1}d^{-1} \text{ или } ab^{-1}=cd^{-1}.$$

За второто равенство умножаваме лявата му част с bd и получаваме

$$\left(\frac{a}{b}+\frac{c}{d}\right)bd=\frac{a}{b}\,bd+\frac{c}{d}\,bd,$$

откъдето с очевидни съкращения ще имаме

$$\left(\frac{a}{b}+\frac{c}{d}\right)bd=ad+cb.$$

Остава да умножим двете части на това равенство с $(bd)^{-1}$, за да получим второто равенство от (2). Третото равенство се установява аналогично. Именно с умножение отначало на bd и след това с $(bd)^{-1}$ получаваме

$$\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}.$$

За да докажем последното равенство от (2), вземаме под внимание, че $\frac{d}{c}$ е обратен елемент на $\frac{c}{d}$, понеже по правилото за умножение имаме $\frac{c}{d}\cdot\frac{d}{c}=\frac{cd}{cd}=e$. На основание на това ще имаме

$$\frac{\frac{a}{b}}{\frac{c}{d}}=\frac{a}{b}\cdot\frac{d}{c}=\frac{ad}{bc},$$

което трябваше да се докаже.

От равенствата (2) се вижда, че множеството от всичките отношения $\frac{a}{b}$ образува поле, което е подполе на A. Това поле се нарича поле на отношенията на пръстена R в A.

Така установихме съществуването на едно поле на отношения P при предположение, че комутативният пръстен R е част от едно тяло A. Но може да се случи, че пръстенът R да е част от друго тяло A' и тогава ще получим ново поле на отношения P'. От самото построяване на такива полета се вижда, че те зависят изключително от структурата на пръстена R. Ще докажем по-точно следната теорема:

Ако за комутативния пръстен, непритежаващ делители на нулата, съществуват полета на отношения, то те са изоморфни, т. е. съществува до изоморфизъм единствено поле на отношения.

Да предположим съществуването на две полета на отношения P и P', респективно в телата A и A'. Нека a и b са два произволни елемента от пръстена R и да означим с x, x' отношенията на тези елементи в полетата P и P', т. е. x е решение на уравнението

$$bx=a$$

в тялото A, а x' е решението на същото уравнение в тялото A'. Да поставим сега в съответствие на елемента x елемента x'. Това съответствие е взаимно еднозначно, т. е. $x \longleftrightarrow x'$. Действително нека решението y на уравнението $b_1 y = a_1$ от тялото A да съответствува на същия елемент x'. Но това значи, че предното уравнение има за решение $y = \frac{a_1}{b_1} = x'$ в тялото A'. Следователно в тялото A' отношенията

$$\frac{a}{b}, \quad \frac{a_1}{b_1}$$

ще бъдат равни. Оттук следва, че $ab_1 = a_1 b$, което равенство показва че отношенията $\frac{a}{b}$ и $\frac{a_1}{b_1}$ ще са равни и в тялото A, т. е. $y = x$. Също очевидно е, че на всеки елемент x' от P' може да се намери съответен елемент x от P. Така виждаме, че между полетата на отношения P и P' има взаимно еднозначно съответствие. Остава да се установи че това съответствие е изоморфизъм. Нека α и β са два кои да е елемента от P и съответните уравнения в тялото A, на които те са корени, да бъдат

$$b_1 x = a_1, \quad b_2 x = a_2, \quad b_1 \neq 0, \quad b_2 \neq 0.$$

Да означим с α' и β' решенията на тези уравнения в тялото A'. По предното въвеждане ще имаме съответствията

$$\alpha \longleftrightarrow \alpha', \quad \beta \longleftrightarrow \beta'.$$

Съгласно с равенството II от (2) $\alpha + \beta$ е решение на уравнението

$$b_1 b_2 x = a_1 b_2 + a_2 b_1 \tag{7}$$

от тялото A. Да означим с γ' решението на (7) от A'. Тогава имаме съответствието $\alpha + \beta \longleftrightarrow \gamma'$. Но според равенството II от (2) сумата $\alpha' + \beta'$ трябва да е решение на уравнението (7) в тялото A' и понеже уравнението (7) има само едно решение, следва, че $\gamma' = \alpha' + \beta'$. С това установихме съответствието на сумите $\alpha + \beta$ и $\alpha' + \beta'$. Подобно на основание на третото равенство III от (2) доказваме съответствието $\alpha\beta \longleftrightarrow \alpha'\beta'$. Така установихме, че полетата P и P' са изоморфни.

С предните разглеждания ние установихме съществуване на поле на отношенията при предположение, че даденият пръстен е част от тяло. Това поле е единствено до изоморфизъм. Сега ще установим съществуването на полето независимо от предположението за тялото.

Поле от отношения съществува за всяка област на цялостност.

Тривиалният случай, когато R се състои само от нулата, е очевиден. Да разгледаме множеството M от всички наредени чифтове (a, b) от елементите на R, като $b \neq 0$. Да разложим това множество на подмножества, които ще наричаме класове. За тази цел, ако (a, b) е един произволно взет чифт, то към същия клас причисляваме всички чифтове (a_1, b_1), за които имаме

$$ab_1 = a_1 b. \tag{8}$$

Така въпросният клас (a, b) е напълно определен с чифта (a, b) и ще съпоставим на него дробта $\frac{a}{b}$. В класа K сигурно влиза неговият представител (a, b), понеже равенството (8) се удовлетворява за $a_1=a$, $b_1=b$. От друга страна, всеки чифт (a_1, b_1) от класа K може да се вземе за негов представител. За да установим това, трябва да докажем верността на равенството $\frac{a}{b}=\frac{a_1}{b_1}$. Нека (a_1, b_1) е чифт от класа K. Следователно

$$ab_1=a_1b. \tag{9}$$

Ако (a_2, b_2) е произволен чифт от класа K, то имаме

$$ab_2=a_2b,$$

откъдето с умножение на b_1 получаваме

$$ab_1b_2=a_2b_1b.$$

Като заместим ab_1 с равното му от (9), получаваме $a_1bb_2=a_2b_1b$. Със съкращаване на b ще имаме

$$a_1b_2=a_2b_1.$$

Това равенство показва, че чифтът (a_2, b_2) принадлежи на класа $\frac{a_1}{b_1}$. Следователно всеки чифт от класа $\frac{a}{b}$ принадлежи на класа $\frac{a_1}{b_1}$. Обратно, по подобен начин установяваме, че всеки чифт от класа $\frac{a_1}{b_1}$ принадлежи на класа $\frac{a}{b}$. Следователно класовете $\frac{a}{b}$ и $\frac{a_1}{b_1}$ съвпадат, т. е. имаме $\frac{a}{b}=\frac{a_1}{b_1}$.

Ако един чифт (a_1, b_1) от множеството P не принадлежи на класа $\frac{a}{b}$, то класовете $\frac{a}{b}$ и $\frac{a_1}{b_1}$ ще бъдат различни и не ще имат нито един общ елемент. Действително, ако биха имали един общ чифт (a_2, b_2), то би трябвало да имаме

$$ab_2=a_2b, \tag{10}$$

$$a_1b_2=a_2b_1. \tag{11}$$

Като умножим равенството (10) с a_1, получаваме

$$aa_1b_2=a_1ba_2.$$

Като заместим тук произведението a_1b_2 с равното му от (11), получаваме

$$aa_2b_1=a_1ba_2,$$

откъдето със съкращаване на a_2 имаме

$$ab_1 = a_1 b.$$

Това равенство показва, че чифтът (a_1, b_1) трябва да принадлежи на класа $\frac{a}{b}$, което противоречи на условието. Следователно $\frac{a}{b} = \frac{a_1}{b_1}$ е само тогава вярно, когато $ab_1 = a_1 b$.

Така установихме, че множеството M се разпада на класове $\frac{a}{b}$, нямащи нито един общ елемент два по два. Да означим с P множеството на всички така получени класове. Ще въведем сега действията събиране и умножение за елементите на това множество и ще установим по такъв начин, че то е поле. Именно на основание на равенствата (2) за отношенията сумата и произведението на елементите на P дефинираме с равенствата

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \tag{12}$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}. \tag{13}$$

Първо, символите в десните части на предните равенства имат смисъл. Пръстенът R няма делители на нулата и от $b \neq 0$, $d \neq 0$ следва, че $bd \neq 0$. Второ, десните части на същите равенства не зависят от избора на представителите на класовете. Наистина да вземем вместо (a, b) като представител на класа $\frac{a}{b}$ друг чифт (a', b'). Тогава

$$ab' = ba'. \tag{14}$$

Като умножим това равенство с d, получаваме

$$adb' = a'db.$$

Като прибавим сега към двете части на това равенство произведението bcb', получаваме

$$adb' + bcb' = a'db + b'cb,$$

$$(ad+bc)\, b' = (a'd+b'c)\, b.$$

С умножение на двете части с d ще имаме

$$(ad+bc)\, b'd = (a'd+b'c)\, bd.$$

Последното равенство показва, че

$$\frac{ad+bc}{bd} = \frac{a'd+b'c}{b'd}.$$

Аналогично с умножение на равенството (12) с cd получаваме

$$acb'd = a'cbd,$$

откъдето имаме

$$\frac{ac}{bd} = \frac{a'c}{b'd}.$$

По напълно подобен начин установяваме, че като изберем вместо (c, d) друг представител на класа, то не ще изменим десните части на равенствата (12) и (13).

Лесно е да проверим, че всичките свойства на полето са в сила. Например съчетателният закон следва непосредствено от равенствата

$$\frac{a}{b}+\left(\frac{c}{d}+\frac{g}{h}\right)=\frac{a}{b}+\frac{ch+dg}{dh}=\frac{adh+bch+bdg}{bdg}$$

$$\left(\frac{a}{b}+\frac{c}{d}\right)+\frac{g}{h}=\frac{ad+cb}{bd}+\frac{g}{h}=\frac{adh+cbh+bdg}{bdh}=\frac{a}{b}+\left(\frac{c}{d}+\frac{g}{h}\right).$$

Аналогично се проверяват останалите закони.

Построеното така тяло P е очевидно комутативно, т. е. поле. Трябва да установим, че то съдържа пръстена R. Затова трябва да докажем, че елементите на R са тъждествени с някои дроби от P. Това можем да направим по следния начин: на всеки елемент c от R отнасяме всичките дроби $\frac{cb}{b}$, където $b\neq 0$. Тези дроби съвпадат, т. е. $\frac{cb}{b}=\frac{cb_1}{b_1}$, понеже $(cb)b_1=b(cb_1)$. По този начин на всеки елемент c правим да съответствува една дроб, т. е. един клас. На различни елементи c и c' ще съответствуват и различни дроби. Действително от

$$\frac{cb}{b}=\frac{c'b'}{b'}$$

следва

$$cbb'=bc'b'$$

и със съкращение на bb' (понеже $b\neq 0$, $b'\neq 0$) имаме $c=c'$.

Следователно на елементите c на R съответствуват взаимно еднозначно определени дроби от $\frac{cb}{b}$. Ако между елементите c_1, c_2, c_3 на пръстена имаме връзката $c_1+c_2=c_3$, то при произволни b_1, $b_2\neq 0$ и $b_3=b_1b_2$ ще имаме

$$\frac{c_1b_1}{b_1}+\frac{c_2b_2}{b_2}=\frac{c_1b_1b_2+c_2b_1b_2}{b_1b_2}=\frac{c_3b_3}{b_3}.$$

Ако между c_1, c_2, c_3 имахме връзката $c_3=c_1c_2$, то ще имаме

$$\frac{c_1b_1}{b_1}\cdot\frac{c_2b_2}{b_2}=\frac{c_1c_2b_1b_2}{b_1b_2}=\frac{c_3b_3}{b_3}.$$

От предните равенства виждаме, че между дробите $\frac{c_kb_k}{b_k}$, съответствуващи на елементите c_k на R, съществува същият начин на събиране и умножение, както между самите елементи c_k, т. е. тези дроби образуват област, изоморфна на пръстена R. На основание на това ние можем да заменим дробите $\frac{cb}{b}$ със съответствуващите им еле-

менти c. Така достигнахме до поле, съдържащо дадената област на цялостност.

Например областта на цялостност, състояща се от целите числа се разширява в полето на рационалните числа.

5. Пръстен от полиноми. Нека R е един произволен пръстен. Да разгледаме изрази от вида

$$f(x)=a_0 x^{n_0}+a_1 x^{n_1}+a_2 x^{n_2}+\ldots+a_p x^{n_p},$$

където $n_0, n_1, \ldots, n_p$ са цели неотрицателни числа и $a_0, a_1, a_2, \ldots, a_p$ са елементи от R. Тук символът x не е елемент от пръстена R. Изразът $f(x)$ се нарича полином и x — неизвестно. Групирайки еднаквите степени на x, можем да предположим, че числата $n_0, n_1, n_2, \ldots, n_p$ са различни. Два полинома считаме за равви, ако коефициентите пред еднаквите им степени са равни помежду си. При това членовете с коефициенти нула могат да се пишат или да се изпускат. Така, ако пишем и членовете с коефициенти, евентуално равни на нула, общата форма на полинома $f(x)$ ще бъде

$$f(x)=b_0 x^n+b_1 x^{n-1}+b_2 x^{n-2}+\ldots+b_n.$$

Ще дефинираме сега основните действия с полиномите $f(x)$, като предполагаме неизвестното x комутативно с всеки елемент от R, т. е. $\alpha x=x\alpha$ за $\alpha \in R$. Събирането и умножението на полиномите извършваме както преди. Именно, ако $\varphi(x)=\sum c_\nu x^\nu$ е също полином от разгледания вид, то под сума на полиномите $\varphi(x)$ и $\psi(x)=\sum d_\nu x^\nu$ разбираме полином $g(x)$ с коефициенти

$$g_\nu=c_\nu+d_\nu,$$

където $c_\nu=0$, ако c_ν липсва в полинома $\varphi(x)$ и $d_\nu=0$ при липса на този коефициент в полинома $\psi(x)$. Коефициентите на полинома $\varphi(x)\psi(x)$ се дават с

$$h_\tau=\sum_{\tau+\eta=\nu} c_\tau d_\eta$$

при подобна бележка за липсващите членове. Лесно се вижда, че множеството от всички полиноми образува пръстен. Действително събирането очевидно е комутативно (понеже се свежда към събиране на елементи от пръстена). Единият разпределителен закон за три полинома

$$f(x)=\sum a_\nu x^\nu, \qquad \varphi(x)=\sum b_\nu x^\nu, \qquad \psi(x)=\sum c_\nu x^\nu$$

се свежда на установяване на равенството

$$f(x)[\varphi(x)+\psi(x)]=f(x)\psi(x)+f(x)\psi(x).$$

Верността на това равенство следва от равенствата

$$\sum_{\alpha+\beta=\nu} a_\alpha (b_\beta+c_\beta)=\sum_{\alpha+\beta=\nu} a_{\alpha\beta}+\sum_{\alpha+\beta=\nu} a_\alpha c_\beta.$$

Подобно имаме за другия разпределителен закон. Съдружителният закон

$$[f(x)\varphi(x)]\psi(x)=f(x)[\varphi(x)\psi(x)]$$

е верен, понеже

$$\sum_{\tau+\nu=m}\Big(\sum_{p+q=\nu}a_p b_q\Big)c_\tau=\sum_{p+q+\tau=m}a_p b_q c_\tau=\sum_{p+\mu=m}a_p\Big(\sum_{q+\tau=\mu}b_q c_\tau\Big).$$

Полученият така пръстен се нарича пръстен на полиномите на неизвестното x над пръстена R. Отбелязваме го с $R[x]$. Ако пръстенът R е комутативен, то и пръстенът $R[x]$ е очевидно комутативен. Дефиницията за степен на полином е съща, както преди. Полиномите от нулева степен са от формата ax^0, където a е елемент от R. Но тези полиноми се събират и умножават точно така, както съответните числа от пръстена R и следователно, като съпоставим на ax^0 елемента a от R, то полиномите от нулева степен ax^0 образуват система, изоморфна на пръстена R, и на това основание можем да идентифицираме с елементите на R. Следователно пръстенът $R[x]$ ще съдържа пръстена R. Казваме, че пръстенът $R[x]$ е получен от R с адюнгиране на неизвестното x.

Ако пръстенът R е област на цялостност, то и пръстенът $R[x]$ е област на цялостност.

Нека $f(x)$ и $\varphi(x)$ са два полинома от степени n и m и $a_n b_m \neq 0$. Ако a_n и b_m са коефициентите на x^n и x^m, то коефициентът на x^{n+m} в полинома $f(x)\varphi(x)$ е равен на $a_n b_m$ и следователно не е равен на нула. В пръстена не ще има делители на нулата, т. е. той е област на цялостност.

В случай, че пръстенът R има единица, то формалното определение на сума и произведение на полиноми се свежда на обикновените начини за тези действия. Действително нека означим с 1 единицата на R и елемента $1 . x$ — с x. За степените на x ще имаме $x^\nu=(1x)^\nu= =1^\nu x^\nu=1 x^\nu$ и за произведенията $a_\nu x^\nu$ получаваме $(a_\nu x^0)(1 x^\nu)=a_\nu x^\nu$.

Нека

$$f(x)=a_0+a_1 x+\ldots+a_n x^n$$

е произволен полином от $R[x]$ и α произволен елемент от пръстена R. Тогава елементът от пръстена

$$f(\alpha)=a_0+a_1\alpha+\ldots+a_n\alpha^n,$$

който се получава, като вместо x поставим α в $f(x)$ и извършим означените там действия, се нарича значение на полинома $f(x)$ за $x=\alpha$.

6. Деление на полиноми. Нека P е произволно поле и $P[x]$ е пръстенът, образуван от полиномите

$$f(x)=a_0x^n+a_1 x^{n-1}+\ldots+a_n,$$

на която коефициентите са елементи от P. По-рано разгледахме делимостта на полиномите в областта на реалните и комплексните числа. Тук по напълно аналогичен път ще разгледаме същия въпрос за про-

изволни пръстени $P[x]$. Казваме, че полиномът $f(x)$ от пръстена $P[x]$ се дели на полинома $\varphi(x)$ от $P[x]$, ако съществува полином $\psi(x)$ от $P[x]$, за който имаме

$$f(x)=\varphi(x)\,\psi(x).$$

Полиномът $\varphi(x)$ се нарича делител на полинома $f(x)$. По-нататък ще разгледаме само полиноми от пръстена $P[x]$. Имаме следната обща теорема:

За всеки два произолни полинома $f(x)$ и $g(x)$ могат да се намерят такива два полинома $g(x)$ и $R(x)$, че

$$f(x)=g(x)\,q(x)+R(x). \tag{1}$$

При това степента на $R(x)$ е по-малка от тази на $g(x)$. Полиномите $q(x)$ и $R(x)$, които удовлетворяват на предните условия, са определени еднозначно.

Доказателството е аналогично на известната ни теорема за разглеждания случай по-рано. Нека полиномите са

$$\begin{aligned} f(x)&=a_0x^n+a_1x^{n-1}+\ldots+a_n, \quad a_0\neq 0,\\ g(x)&=b_0x^m+b_1x^{m-1}+\ldots+b_m, \quad b_0\neq 0. \end{aligned} \tag{2}$$

Ако $n<m$, то равенството (1) е изпълнено при $q(x)=0$, $R(x)=f(x)$. Следователно можем да допуснем, че $n\geqq m$. Полиномът

$$f(x)-\frac{a_0}{b_0}x^{n-m}g(x)=f_1(x) \tag{2'}$$

очевидно принадлежи на $P[x]$ и е от степен n_1, по-малка от n. Да означим с $a_0^{(1)}$ коефициента на x^{n_1} в полинома $f_1(x)$. Да предположим, че $n_1\geqq m$. Полиномът

$$f_1(x)-\frac{a_0^{(1)}}{b_0}x^{n_1-m}g(x)=f_2(x) \tag{2''}$$

ще бъде от степен $n_2<n_1$ и ако $a_0^{(2)}$ е коефициентът пред x^{n_2} в него образуваме полинома

$$f_2(x)-\frac{a_0^{(2)}}{b_0}x^{n_2-m}g(x)=f_3(x) \tag{2'''}$$

и т. н. Понеже степените на полиномите $f_1(x)$, $f_2(x)$, $f_3(x)$, ... намаляват, то ще достигнем до такъв полином

$$f_{p-1}(x)-\frac{a_0^{(p-1)}}{b_0}x^{n_{p-1}-m}g(x)=f_p(x), \tag{2^{(p)}}$$

степента на който е по-малка от m, и тогава спираме до този полином. Като съберем равенствата (2'), (2''), (2'''), ..., ($2^{(p)}$), получаваме

$$f(x)-q(x)\,g(x)=R(x),$$

където

$$q(x)=\frac{a_0}{b_0}x^{n-m}+\frac{a_0^{(1)}}{b_0}x^{n_1-m}+\ldots+\frac{a_0^{(p-1)}}{b_0}x^{n_{p-1}-m},\quad R(x)=f_p(x),$$

Очевидно коефициентите на полиномите $q(x)$, $R(x)$ принадлежат на полето P.

Еднозначността на представянето (1) се вижда почти непосредствено. Ако имахме и второ представяне от подобен вид

$$f(x)=g(x)\,q_1(x)+R_1(x), \tag{3}$$

то от (1) и (3) бихме имали

$$g(x)[\,q(x)-q_1(x)\,]=R_1(x)-R(x).$$

Ако $q(x)-q_1(x)\neq 0$, то полиномът в лявата част на това равенство ще има по-висока степен от полинома в дясната част, което е невъзможно.

Полиномът $q(x)$ се нарича частно от делението на $f(x)$ с $g(x)$ и $R(x)$ — остатък от това деление. За да се дели $f(x)$ на $g(x)$, необходимо и достатъчно е остатъкът да бъде равен на нула.

От самия начин на намиране на полиномите $q(x)$ и $R(x)$ се вижда, че в случай на разширение на полето P в поле P' тези полиноми остават същите. Именно всеки полином от пръстена $P[x]$ ще принадлежи и на пръстена $P'[x]$. Следователно, ако полиномът $f(x)$ не се дели на $\varphi(x)$ при разширение на полето P, той също няма да се дели на $\varphi(x)$.

За делимостта на полиномите следват лесно следните предложения:

Всеки полином се дели на произволен, отличен от нула елемент от полето P.

Действително, ако $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n$ е полином от $P[x]$ и $\alpha\neq 0$ елемент от P, то имаме

$$f(x)=\alpha\left(\frac{a_0}{\alpha}\,x^n+\frac{a_1}{\alpha}\,x^{n-1}+\ldots+\frac{a_n}{\alpha}\right)$$

и полиномът в скобите принадлежи на $P[x]$.

Ако полиномът $f(x)$ се дели на полинома $\varphi(x)$, а $\varphi(x)$ се дели на полинома $\psi(x)$, то $f(x)$ се дели на $\psi(x)$.

Следва от равенствата

$$f(x)=\varphi(x)\,q(x),\quad \varphi(x)=\psi(x)\,h(x),$$

$$f(x)=\psi(x)[\,q(x)\,h(x)\,],$$

където gh е полином от $P[x]$.

Ако $f(x)$ и $\varphi(x)$ се делят на $g(x)$, то сумата и разликата им се делят на $g(x)$.

Ако $f(x)$ се дели на $\varphi(x)$ и $\psi(x)$ е произволен полином от $P[x]$, то $f(x)\psi(x)$ се дели на $\varphi(x)$.

Ако полиномите $f_1(x)$, $f_2(x),\ldots,f_n(x)$ се делят на полинома $g(x)$ и $\varphi_1(x)$, $\varphi_2(x),\ldots,\varphi_n(x)$ са произволни полиноми от $P[x]$, то полиномът

$$f_1(x)\varphi_1(x)+f_2(x)\varphi_2(x)+\ldots+f_n(x)\varphi_n(x)$$

се дели на $g(x)$.

Ако $f(x)$ се дели на $\varphi(x)$, то $f(x)$ се дели и на $\alpha\varphi(x)$, където α е произволен елемент от полето P.

Ако полиномите $f(x)$ и $\varphi(x)$ са от еднаква степен и $f(x)$ се дели на $\varphi(x)$, то $f(x)=\alpha\varphi(x)$, където α е елемент от полето P.

Вземаме под внимание, че степента на частното трябва да е нула. Очевидно всеки полином $f(x)$ се дели на $\alpha f(x)$, където $\alpha\neq 0$ е елемент от P и $\alpha f(x)$ се дели на $f(x)$.

Ако полиномът $f(x)$ се дели на полинома $\varphi(x)$ и $\varphi(x)$ се дели на $f(x)$, то $f(x)=\alpha\varphi(x)$, α е елемент от P.

Всеки полином $\varphi(x)$, който дели два полинома $f(x)$ и $\varphi(x)$, се нарича общ техен делител. Ако два полинома нямат друг общ делител освен полиномите от нулева степен, т. е. елементите на полето P, то те се наричат взаимно прости. Най-голям общ делител на два полинома наричаме общия им делител от най-висока степен. За намирането му си служим с алгоритъма на Евклид, който вече познаваме. От него следва, че до полиноми от нулева степен той е еднозначно определен и за най-големия общ делител $D(x)$ на полиномите $f(x)$ и $\varphi(x)$ имаме равенството

$$D(x)=f(x)g(x)+\varphi(x)h(x), \tag{4}$$

където $g(x)$ и $h(x)$ са полиноми от $P[x]$. В частност получаваме:

Полиномите $f(x)$ и $\varphi(x)$ са само тогава взаимно прости, когато могат да се намерят два полинома $g(x)$ и $h(x)$ такива, че да имаме

$$f(x)g(x)+\varphi(x)h(x)=1. \tag{5}$$

Оттук следват предложенията:

Ако полиномът $f(x)$ е взаимно прост с полиномите $\varphi(x)$ и $\psi(x)$, то той е взаимно прост и с произведението им.

Действително съгласно с предположението ще имаме по (5) равенството

$$f(x)g(x)+\varphi(x)h(x)=1, \tag{6}$$

където $g(x)$ и $h(x)$ са полиноми от $P[x]$. С умножение на $\psi(x)$ получаваме

$$f(x)[g(x)\psi(x)]+[\varphi(x)\psi(x)]h(x)=\psi(x),$$

отгдето следва, че всеки общ делител на $f(x)$ и $\varphi(x)\psi(x)$ ще бъде делител на $\psi(x)$. Но по условие полиномите $f(x)$ и $\psi(x)$ са взаимно прости.

Ако произведението на полиномите $f(x)$ и $\psi(x)$ се дели на полинома $\varphi(x)$ и ако $f(x)$ и $\varphi(x)$ са взаимно прости, то $\psi(x)$ се дела на $\varphi(x)$.

Като умножим равенството (6) с $\psi(x)$, получаваме

$$[f(x)\psi(x)]g(x)+\varphi(x)h(x)\psi(x)=\psi(x).$$

Понеже вляво полиномите $f(x)\psi(x)$, $\varphi(x)h(x)\psi(x)$ се делят на $\varphi(x)$, с това следва, че $\psi(x)$ се дели на $\varphi(x)$.

Ако полиномът $f(x)$ се дели на двата взаимно прости помежду си полинома $\varphi(x)$ и $\psi(x)$, то $f(x)$ се дели на тяхното произведение.

Наистина от делимостта на полинома $f(x)$ на $\varphi(x)$ заключаваме, че $f(x)=\varphi(x)h(x)$, където $h(x)$ е полином. Понеже $f(x)$ трябва да се дели $\psi(x)$ и последният полином е взаимно прост с $\varphi(x)$, то трябва $h(x)$ да се дели на $\psi(x)$ и ще имаме $h(x)=\psi(x)q(x)$, $f(x)=q(x)[\varphi(x)\psi(x)]$.

Един полином ще наричаме нормиран, ако коефициентът пред най-високата степен на x е равен на 1. Тогава общият най-голям делител на нормирани полиноми ще бъде еднозначно определен при предположение, че е нормиран.

Можем да дефинираме аналогично най-голям общ делител на няколко полинома $f_1(x), f_2(x), \ldots, f_m(x)$ като общ делител от най-висока степен или като общ делител, който се дели на всеки общ делител на тези полиноми. Лесно се вижда, че най-големият общ делител на полиномите $f_1(x), f_2(x), \ldots, f_m(x)$ е равен (до множител от нулева степен) на най-големия общ делител на полинома $f_m(x)$ и на най-големия общ делител на полиномите $f_1(x), f_2(x), \ldots, f_{m-1}(x)$.

7. Разложимост на полиномите. Един полином $f(x)$ от пръстена $P[x]$ се нарича разложим, ако може да се представи като произведение от два полинома $P[x]$, на които степените са най-малко равни на 1. В противен случай полиномът се нарича неразложим. Така, ако P е естествената област на рационалност, то полиномът

$$x^2-4=(x-2)(x+2)$$

е разложим, а полиномът x^2+1 е неразложим. Полиномът x^2-4x+2 е неразложим, ако P е естествената област на рационалност, но е разложим при предположение, че P е полето от числата $a+b\sqrt{2}$, a и b — рационални числа, понеже

$$x^2-4x+2=(x-2-\sqrt{2})\quad(x-2+\sqrt{2}).$$

Ако полиномът $f(x)$ от $P[x]$ е неразложим и полиномът $F[x]$ от $P[x]$ има общ множител с него от най-малко първа степен, то $F(x)$ се дели на $f(x)$.

Действително съгласно с условието общият най-голям делител $D(x)$ на $f(x)$ и $F(x)$ ще бъде поне от първа степен. Но видяхме от начина на получаването му, че той е полином от пръстена $P[x]$. Понеже $D(x)$ дели неразложимия полином $f(x)$, то $D(x)=\alpha f(x)$, където α е елемент от P. Тогава от делимостта на полинома $F(x)$ на $D(x)$ следва, че $F(x)$ се дели на $f(x)$. Имаме следната основна теорема:

Всеки нормиран полином $f(x)$ от $P[x]$ може да се разложи по единствен начин на неразложими в $P[x]$ нормирани полиноми, т. е.

$$f(x)=p_1(x)p_2(x)\ldots p_k(x). \tag{1}$$

Действително нека $f(x)$ е нормиран полином от $P[x]$. Ако е неразложим, то теоремата е очевидна. Ако $f(x)$ е разложим, то ще

имаме $f(x)=\varphi(x)\psi(x)$, гдето $\varphi(x)$, $\psi(x)$ са полиноми от $P[x]$, които могат да се предположат нормирани. В случай, че някой от полиномите $\varphi(x)$ и $\psi(x)$ е разложим, то продължаваме разлагането, докато представим полинома $f(x)$ в произведение на неразложими нормирани полиноми. До такива сигурно ще достигнем, понеже степените на въпросните полиноми намаляват и полиномите от първа степен, както веднага се вижда, са неразложими. Така първата част на теоремата е установена. За да установим втората част, нека предположим, че имаме и разлагането на неразложими и нормирани полиноми

$$f(x)=q_1(x)q_2(x)\dots q_s(x). \tag{2}$$

Но тогава от (1) и (2) ще имаме равенството

$$p_1(x)p_2(x)\dots p_k(x)=q_1(x)q_2(x)\dots q_s(x). \tag{3}$$

Оттук следва, че например полиномът $q_1(x)$ трябва да дели произведението $p_1(x)p_2(x)\dots p_k(x)$. Ако $q_1(x)$ не съвпада с полинома $p_1(x)$, то той трябва да дели полинома $p_2(x)\dots p_k(x)$. В случай, че $q_1(x)$ не съвпада с $p_2(x)$, то $q(x)$ трябва да дели произведението $p_3(x)\dots p_k(x)$. Продължавайки така, получаваме, че $q_1(x)$ трябва да съвпада поне с един от полиномите $p_1(x)$, $p_2(x),\dots,p_k(x)$. Размества́йки означенията на полиномите $p_l(x)$, можем да приемем, че $q_1(x)=p_1(x)$. Но тогава равенството (3) става

$$p_2(x)p_3(x)\dots p_k(x)=q_2(x)q_3(x)\dots q_s(x). \tag{4}$$

От това равенство, прилагайки горното заключение, получаваме, че полиномът $q_2(x)$ трябва да е равен на някой от полиномите $p_2(x)$, $p_3(x),\dots,p_k(x)$, например $q_2(x)=p_2(x)$. Тогава равенството (4) става

$$p_3(x)p_4(x)\dots p_k(x)=q_3(x)q_4(x)\dots q_s(x).$$

Като продължаваме така, убеждаваме се, че полиномите $q_1(x)$, $q_2(x),\dots,q_s(x)$ трябва да бъдат равни на полиномите $p_1(x)$, $p_2(x),\dots$, $p_k(x)$, с което теоремата е установена напълно.

Могат евентуално някои полиноми в (1) да се повтарят. Следователно за всеки нормиран полином $f(x)$ от $P[x]$ ще имаме представянето

$$f(x)=\varphi_1^{\lambda_1}(x)\varphi_2^{\lambda_2}(x)\dots\varphi_m^{\lambda_m}(x), \tag{5}$$

където $\varphi_1(x)$, $\varphi_2(x),\dots,\varphi_m(x)$ са неразложими нормирани полиноми, от $P[x]$ и λ_1, $\lambda_2,\dots,\lambda_m$ са естествени числа. Като се ограничаваме на различни неразложими полиноми, то представянето (5) относно полиномите $\varphi_i(x)$, $1\leqq i\leqq m$ и числата λ_i, $1\leqq i\leqq m$ е еднозначно. При ненормирани полиноми ще имаме разлагането

$$f(x)=\alpha\psi_1^{\lambda_1}(x)\psi_2^{\lambda_2}(x)\dots\psi_m^{\lambda_m}(x), \tag{6}$$

гдето $\psi_1(x)$, $\psi_2(x),\dots,\psi_m(x)$ са неразложими полиноми, λ_1, $\lambda_2,\dots,\lambda_m$ са естествени числа и α е произволен елемент от P. Представянето

(5) е също еднозначно относно различните полиноми $\varphi_i(x), 1 \leqq i \leqq m$ и числата λ_i, $1 \leqq i \leqq m$.

Числата $\lambda_1, \lambda_2, \ldots, \lambda_m$ са кратностите на полиномите в разлагането на полинома $f(x)$. За многократните неразложими множители имаме следната теорема (полето P се предполага винаги, че е с характеристика 0):

Ако неразложимият полином $p(x)$ е m-кратен множител в разлагането на полинома $f(x)$, то той е $m-1$-кратен неразложим множител за $f'(x)$.

Действително имаме

$$f(x)=p^m(x)\,g(x),$$

гдето $g(x)$ е полином от $P[x]$, който не се дели на $p(x)$. За $f'(x)$ получаваме

$$f'(x)=mp^{m-1}(x)\,p'(x)\,g(x)+p^m(x)\,g'(x)=$$

$$=p^{m-1}(x)\,[mp'(x)\,g(x)+p(x)\,g'(x)\,].$$

Полиномът в скобите не се дели на $p(x)$, понеже полиномът $g(x)$ не се дели на $p(x)$ и производният полином $p'(x)$ не се дели на $p(x)$, а вторият член $p(x)\,g(x)$ се дели на $p(x)$. Прилагайки тогава това правило, получаваме, че $p(x)$ е $m-2$-кратен неразложим множител на $f''(x)$, $m-3$-кратен на $f'''(x)$ и т. н., прост на $f^{(m-1)}(x)$ и не е множител на $f^{(m)}(x)$.

От предната теорема следва, че общият най-голям делител на полинома (5) и на производната му е равен на

$$D(x)=\varphi_1^{\lambda_1-1}(x)\,\varphi_2^{\lambda_2-1}(x)\ldots\varphi_m^{\lambda_m-1}(x),$$

където естествено множителите $\varphi_s^{\lambda_s-1}(x)$ при $\lambda_s=1$ трябва да се заместят с 1.

На основание на предното свойство можем да отделим многократните множители, следвайки път, аналогичен на този при отделяне на многократните корени на уравненията. Нека в разлагането на полинома $f(x)$ на неразложими множители да означим с X_1 произведението на еднократните множители, с X_2^2 произведението на двукратните множители, с X_3^3 това на трикратните и т. н. Ако при някое k няма k-кратни множители, то полагаме $X_k=1$. Тогава $f(x)$ се представя така:

$$f(x)=X_1\,X_2^2\,X_3^3\ldots X_m^m,$$

където m е равно най-много на степента на полинома $f(x)$. Ако $D_1(x)$ е общият най-голям делител на $f(x)$ и $f'(x)$, то ще имаме

$$D_1(x)=X_2X_3^2\,X_4^3\ldots X_m^{m-1}.$$

Като означим по-нататък с $D_2(x)$ общия най-голям делител на $D_1(x)$ и $D_1'(x)$, с $D_3(x)$ този на полиномите $D_2(x)$ и $D_2'(x)$ и т. н., ще имаме

$$D_2(x)=X_3X_4^2\,X_5^3\ldots X_m^{m-2},$$

$$D_3(x)=X_4X_5^2\ldots X_m^{m-8},$$
$$\ldots\ldots\ldots\ldots$$
$$D_{m-1}(x)=X_m,$$
$$D_m(x)=1.$$

Оттук получаваме

$$f_1(x)=\frac{f(x)}{D_1(x)}=X_1X_2X_3\ldots X_m,$$
$$f_2(x)=\frac{D_1(x)}{D_2(x)}=X_2X_3\ldots X_m,$$
$$f_3(x)=\frac{D_2(x)}{D_3(x)}=X_3X_4\ldots X_m.$$
$$\ldots\ldots\ldots\ldots$$
$$f_m(x)=\frac{D_{m-1}(x)}{D_m(x)}=X_m.$$

Полиномите X_1, X_2, X_3,..., X_m се дават с

$$X_1=\frac{f_1(x)}{f_2(x)},\quad X_2=\frac{f_2(x)}{f_3(x)},\ldots,\ X_{m-1}=\frac{f_{m-1}(x)}{f_m(x)},\ X_m=f_m(x).$$

По изложения начин намираме така произведенията на еднаквократните неразложими множители; последните, броени като прости.

Ако α е произволен елемент от полето P, то остатъкът от делението на произволен полином $f(x)$ от пръстена $P[x]$ с $x-\alpha$ е равен на $f(\alpha)$.

Това свойство следва непосредствено от равенството

$$f(x)=(x-\alpha)f_1(x)+R$$

(R елемент от P), като поставим $x=\alpha$. Оттук следва, че $f(x)$ се дели само тогава на $x-\alpha$, когато α е корен на уравнението $f(x)=0$. Следователно намирането на корените на $f(x)=0$ е равносилно с намирането на линейните множители на полинома $f(x)$. Ако полиномът $f(x)$ се дели на $(x-\alpha)^m$, но не се дели на $(x-\alpha)^{m+1}$, то α се нарича, както ни е известно от по-рано, m-кратен корен на уравнението $f(x)=0$. Но числото на линейните множители в разлагането на полинома $f(x)$ на неразложими множители от $P[x]$ не може очевидно да надмине степента му, т. е. имаме следната теорема:

Броят на корените на уравнението не надминава степента му.

От тази теорема следва предложението: ако един полином се анулира за повече значения на неизвестното, отколкото степента му, то той е равен на нула. Непосредствено от това следва, че ако два полинома от степени, които не надминават n, са равни за повече от n различни стойности на неизвестното, то те са равни помежду си.

Видяхме, че всеки нормиран полином от пръстена $P[x]$ може да се разложи по единствен начин на произведение от неразложими нормирани полиноми, принадлежащи на $P[x]$. Ще разгледаме едно обоб-

щение на тази теорема. Ако C е пръстенът, съставен от всички цели числа, то той е очевидно комутативен и от равенството $ab=0$ за два негови елемента следва, че поне един от елементите a и b е равен на нула. Следователно C е област на цялостност с единичен елемент 1. Числата 1 и -1 са единиците в него. Абстрахирайки се от тези единици, всеки елемент от C се разлага, както е известно, по единствен начин на произведение от неразложими елементи, които са простите числа. Нека сега R е област на цялостност, притежаваща единичен елемент e. Обратният елемент e^{-1} на e като равен на e принадлежи също на областта R. Ще наричаме всеки елемент ε от R единица, ако обратният му елемент ε^{-1} принадлежи също на R. Така в областта C единичният елемент е 1, а единиците са 1 и -1. Съвкупността от комплексните числа $a+bi$, където a и b са произволни цели числа, образува област на цялостност с единичен елемент 1 и единици 1, -1, i и $-i$. Същото е в сила и за областта на цялостност, съставена от всички комплексни числа (която е и поле). Ако ε е единица от областта на цялостност R, то всеки елемент a от нея може да се разложи на множители

$$a=a\varepsilon\,.\,\varepsilon^{-1}.$$

Такова разлагане, в което единият множител е единица, ще наричаме тривиално разлагане. Всеки елемент, който не допуска друго разлагане освен тривиалното, се нарича *неразложим* или *прост елемент.* Нека R е област на цялостност с единичен елемент, в която важи еднозначното разлагане на елементите в произведение на прости елементи, т. е. всеки елемент от R се разлага на произведение от краен брой прости елементи и това разлагане, абстрахирайки се от единиците, е еднозначно. Такава е например областта на цялостност C и областта, съставена от полиномите на x, на които коефициентите принадлежат на дадено поле P. Ще установим следната теорема:

Областта на цялостност $R[x]$ важи теоремата за еднозначно разлагане на произведение от прости елементи.

Полиномът от $R[x]$

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$$

се нарича примитивен, ако коефициентите му a_0, a_1, a_2, ..., a_n нямат друг общ множител освен единици. Ще докажем следната теорема на Гаус:

Произведението на два примитивни полинома е примитивен полином.

Нека

$$\varphi(x)=b_0+b_1x+b_2x^2+\ldots+b_mx^n$$

е също примитивен полином и да означим с $F(x)$ произведението

$$F(x)=f(x)\varphi(x)=c_0+c_1x+c_2x^2+c_3x^3+\ldots$$

Да предположим, че коефициентите c_0, c_1, c_2, c_3, ... на полинома $F(x)$ имат общ делител d, различен от единица, и нека p е един прост де-

лител на d. Понеже коефициентите a_0, a_1, a_2, ..., a_n на полинома $f(x)$ не могат едновременно да се делят на p, то нека a_μ е първият поред коефициент, който не се дели на p. Също нека b_ν е първият поред коефициент от коефициентите b_0, b_1, b_2, ..., b_m на полинома $\varphi(x)$, който не се дели на p. Намираме лесно, че коефициентът пред $x^{\mu+\nu}$ в полинома $F(x)$ е равен на

$$(7) \quad c_{\mu+\nu} = a_\mu b_\nu + a_{\mu-1} b_{\nu+1} + a_{\mu-2} b_{\nu+2} + \ldots + a_{\mu+1} b_{\nu-1} + a_{\mu+2} b_{\nu-2} + \ldots$$

Но в дясната част на предното равенство първият член $a_\mu b_\nu$ не се дели на p, а останалите се делят на p и следователно цялата тази част не се дели на p. Но така идваме до противоречие, понеже лявата част на равенството (7) се дели на p. Значи коефициентите c_0, c_1, c_2, ... на полинома $F(x)$ нямат друг общ делител освен единица, т. е. полиномът $F(x)$ е примитивен.

Да означим сега с P полето от отношенията за областта R. Както знаем (§ 4), такова поле съществува. Нека $f(x)$ е произволен полином от пръстена $P[x]$, който в случая е и област на цялостност. Ако приведем всичките коефициенти на $f(x)$ в общ знаменател, то този полином приема формата $f(x) = \frac{\Phi(x)}{b}$, като $\Phi(x)$ е полином от $R[x]$ и b е елемент от R. Като означим с a общия най-голям делител на коефициентите на полинома $\Phi(x)$, то виждаме, че $f(x)$ се представя във вида

$$(8) \quad f(x) = \frac{a}{b} g(x),$$

където $\frac{a}{b}$ са елементи от P и $g(x)$ е примитивен полином в $R[x]$. Това представяне до множител единица е еднозначно. Да допуснем, че има друго представяне на $f(x)$, а именно

$$(9) \quad f(x) = \frac{c}{d} g_1(x),$$

където $\frac{c}{d}$ е от P и $g_1(x)$ е примитивен полином в $R[x]$. От (8) и (9) получаваме

$$(10) \quad ad\, g(x) = bc g_1(x).$$

Понеже полиномите $g(x)$ и $g_1(x)$ са примитивни от $R[x]$, следва, че

$$ad = bc\varepsilon,$$

където ε е единица. Като заместим в (10) ad с $bc\varepsilon$ и съкратим, получаваме

$$\varepsilon g(x) = g_1(x),$$

т. е. полиномите $g(x)$ и $g_1(x)$ до множител единица са идентични. Така установихме теоремата:

Всеки полином $f(x)$ от $P[x]$ се представя еднозначно до множител единица във формата

$$f(x)=\frac{a}{b}g(x),$$

където $\frac{a}{b}$ е елемент от P и $g(x)$ е примитивен полином в $R[x]$.

Ако полиномите от $f(x)$ и $\varphi(x)$ от $P[x]$ са представени по този начин, а именно

$$f(x)=\frac{a}{b}g(x),$$

$$\varphi(x)=\frac{c}{d}h(x),$$

то произведението им

$$f(x)\varphi(x)=\frac{ac}{bd}g(x)h(x)$$

има същото представяне, понеже $g(x)h(x)$ по теоремата на Гаус е примитивен полином. Следователно примитивният полином, който отговаря на произведението на два полинома от $P[x]$, е произведение на примитивните полиноми, отговарящи съответно на двата полинома.

Нека полиномът (8) е разложим в $P[x]$. Тогава той може да се представи в следната форма:

$$f(x)=\frac{c}{d}\,g_1(x)g_2(x), \tag{11}$$

където $\frac{c}{d}$ е елемент от полето P и $g_1(x)$ и $g_2(x)$ са примитивни полиноми от $R[x]$. Но тогава от (8) и (11) получаваме

$$g(x)=\frac{bc}{ad}g_1(x)g_2(x).$$

Понеже $g(x), g_1(x), g_2(x)$ са примитивни полиноми от $R[x]$, както преди заключаваме, че $\frac{bc}{ad}=\varepsilon$ и следователно

$$g(x)=\varepsilon g_1(x)g_2(x),$$

като ε означава една единица. Последното равенство показва, че полиномът $g(x)$ е разложим в $R[x]$. Обратно, ако полиномът $g(x)$ е разложим в $R[x]$, т. е. $g(x)=\alpha(x)\beta(x)$, $\alpha(x)$ и $\beta(x)$ полиноми от $R[x]$, то за $f(x)$ получаваме

$$f(x)=\frac{a}{b}\alpha(x)\beta(x)$$

и следователно $f(x)$ е разложим в $P[x]$. В заключение ще имаме предложението:

Полиномът (8) е само тогава разложим в $P[x]$, когато полиномът $g(x)$ е разложим в $R[x]$.

Това предложение показва, че при изследване на разложимостта на полиномите в $P[x]$ можем да се ограничим на разложимост само в $R[x]$. При това за примитивни полиноми се ограничаваме само на примитивни полиноми като възможни множители.

8. Идеали. Нека R е един комутативен пръстен. Една непразна подсъвкупност $\mathfrak{m}$ на R се нарича идеал, ако елементите ѝ притежават следните свойства:

I. Ако a и b са елементи от $\mathfrak{m}$ то и $a-b$ е елемент от $\mathfrak{m}$. 2. Ако a е елемент от $\mathfrak{m}$ и c — кой да е елемент от R, то ac принадлежи[1] на $\mathfrak{m}$.

Съгласно с дефиницията очевидно нулата е идеал, който наричаме нулев идеал, и пръстенът R е идеал, наречен единичен идеал или идеал единица. Понеже $a-a=0$, $0-a=-a$, $b+a=b-(-a)$, то всеки идеал съдържа нулата и сборът на два кои да е елемента от него е също така елемент, принадлежащ на идеала.

Нека a е произволен елемент от пръстена R. Ако $\mathfrak{m}$ е идеал, на който елементът a принадлежи, то съгласно със свойствата 1 и 2 в $\mathfrak{m}$ трябва да лежат и елементите

$$0, a, a+a=2a,\ 2a+a=3a, \ldots, -a,\ 2\ (-a)=-2a, \ldots,$$

както и елементите ac, където c са произволни елементи от R. Следователно в $\mathfrak{m}$ лежат и елементите

$$ac+na, \tag{1}$$

където n е произволно цяло число и c е произволен елемент от R. Но лесно се вижда, че елементите (1) образуват също така идеал. Действително сборът или разликата на два елемента от формата (1) имат същата форма и произведението на елемент от (1) с елемент d от пръстена R има същата форма, понеже

$$(ac+na)\,d=a\,(cd+nd)=ak+0\,.\,a,$$

и $k=cd+nd$ е елемент от R. Така полученият идеал се означава накъсо с (a) и се нарича главен идеал. От гореизложеното се вижда, че идеалът (a) е в известен смисъл най-малкият идеал, който съдържа елемента a. Ако пръстенът R има единичен елемент e, то за елементите (1) имаме $ac+na=a(c+ne)=ac'$, където c' е елемент от R. Тогава главният идеал (a) се състои от елементите ac', където c' е произволен елемент от R. Очевидно е тогава, че нулевият идеал е главен идеал (0) и единичният идеал $\mathfrak{R}$ е главен идеал (e), ако пръстенът R има елемент единица.

Като прост пример на главен идеал да разгледаме пръстена C от всички цели рационални числа. Ако m е кое да е цяло число, то очевидно

---

[1] Ще отбележим, че при некомутативен пръстен ще имаме десен и ляв идеал според това, дали вземаме произведенията ac или ca.

главният идеал (m) се състои от всички цели числа, които са кратни на m $(m \neq 0)$. Нулевият идеал (0) е числото нула и единичният идеал (1) в случая се състои от всички цели рационални числа.

От повече елементи $a_1, a_2, \ldots, a_k$ на пръстена R можем също да образуваме идеал $\mathfrak{g}$, представляващ съвкупността от елементите от вида

$$(8) \qquad a_1 r_1 + a_2 r_2 + \ldots + a_k r_k + n_1 a_1 + n_2 a_2 + \ldots + n_k a_k,$$

където $n_1, n_2, \ldots, n_k$ са цели числа и $r_1, r_2, \ldots, r_k$ са елементи от R. Този идеал се означава с $(a_1, a_2, \ldots, a_k)$ и елементите $a_1, a_2, \ldots, a_k$ се наричат база на идеала. Ако $\mathfrak{g}$ има единичен елемент e, то вместо (2) можем да се ограничим на елементите $\sum_{i=1}^{k} a_i r_i$. Също от една безкрайна съвкупност (M), принадлежаща на пръстена R, можем да образуваме идеал, означен с (M) и съставен от елементи от вида $\sum_{i=1}^{s} a_i r_i + \sum_{i=1}^{s} n_i a_i$, където a_i са елементи от M и n_i — цели числа $(s = 1, 2, 3, \ldots)$.

С използуване на хомоморфизъм на два пръстена можем да дадем на понятието идеал просто тълкуване, което показва освен това, че това понятие отговаря на понятието за инвариантна подгрупа, което разгледахме. Предварително ще разгледаме тъй наречените конгруенции (сравнения) по модул — идеал от полето R. Именно казваме, че два елемента a, b от R са сравними (конгруентни) по модул $\mathfrak{m}$ ($\mathfrak{m}$ е идеал от R), ако разликата им $a-b$ принадлежи на $\mathfrak{m}$. Ако C е пръстенът, съставен от всички цели числа, като $\mathfrak{m} = (m)$, където m е произволно цяло число, то предното понятие за конгруентност се свежда към класичното такова понятие в теорията на числата, въведено от Гаус. Конгруентността на a и b по модул $\mathfrak{m}$ означаваме с $a \equiv b\,(\mathfrak{m})$. Това означение се нарича конгруенция (сравнение). Ако идеалът $\mathfrak{m}$ е главен $\mathfrak{m} = (m)$, то конгруенцията се означава по-просто: $a \equiv b\,(m)$.

При равен модул конгруенциите имат някои от свойствата на равенствата. За късота модула $\mathfrak{m}$ не ще го пишем. Ще установим няколко основни свойства.

От $a \equiv b$ и $b \equiv c$ следва, че $a \equiv c$.

Съгласно с условията $a-b \in \mathfrak{m}$ и $b-c \in \mathfrak{m}$. Следователно $a-c = (a-b)+(b-c) \in \mathfrak{m}$.

От $a \equiv b$ следва, че $ak \equiv bk$, като k е произволен елемент от $\mathfrak{m}$.

Понеже $a-b$ по условие принадлежи на $\mathfrak{m}$, то $(a-b)\,k$ принадлежи на $\mathfrak{m}$.

От $a \equiv b$ и $c \in R$ следва $a+c \equiv b+c$.

Действително разликата $a+c-(b+c) = a-b$ принадлежи на $\mathfrak{m}$. От $a \equiv b$ и $c \equiv d$ следват конгруенциите

$$a \pm c \equiv b \pm d,$$

$$ac \equiv bd.$$

Първата следва от равенството $a \pm c - (b \pm d) = a - b \pm (c - d)$ или от предното свойство и втората се получава от конгруенциите $ac \equiv bc$, $bc \equiv bd$.

Всеки пръстен относно събирането представлява Абелова група, като всеки идеал $\mathfrak{m}$, принадлежащ на R, е негова подгрупа. Тогава подобно на теорията на групите ние можем да разделим пръстена R в класи, като във всяка класа да отнесем елементите от R, които са конгруентни помежду си спрямо модула $\mathfrak{m}$. С други думи, ако a е елемент от R, то под клас R_a ще разбираме съвкупността от всичките елементи на R, които са конгруентни с a по модул $\mathfrak{m}$. На основание на горните свойства на конгруенциите следва, че ако a_1 е елемент от класа R_a, то класът R_{a1} съвпада с R_a, т. е. класът е определен с кой да е негов елемент. Така дефинираните класове ще наричаме класове остатъци по модул $\mathfrak{m}$ по аналогия на случая, когато R е пръстенът от всички цели числа. В последния случай, ако $\mathfrak{m}$ е главен идеал, $\mathfrak{m} = (m)$, като всичките цели числа (елементите на R) се разделят на m класи $R_0, R_1, \ldots, R_{m-1}$, които разгледахме в § 2, гл. II.

Съществуват пръстени, в които всеки идеал е главен идеал. Например такъв е пръстенът C от всички цели числа. Действително нека $\mathfrak{m}$ е един идеал, принадлежащ на C. Ако $\mathfrak{m}$ е нулевият идеал $m = (0)$, то твърдението е очевидно. Ако $\mathfrak{m}$ не е нулевият идеал, то в него ще има цели числа, отлични от нула. Да означим тогава с δ най-малкото от положителните числа на идеала $\mathfrak{m}$. Такова има, понеже ако $\mathfrak{m}$ съдържа отрицателното число $-a$, то той съдържа и положителното число a. Всеки елемент b от идеала като цяло число може да се представи (с деление на δ) във вида

$$b = \delta k + r,$$

където r е цяло число, за което $0 \leqq r < \delta$. Лесно се вижда, че остатъкът r трябва да бъде равен на нула. Числата b и δk принадлежат на идеала $\mathfrak{m}$ и следователно числото $b - \delta k = r$ ще принадлежи също на идеала. Ако $r > 0$, то би следвало, че в идеала има положително число, по-малко от δ, което противоречи. Значи $r = 0$ и идеалът е съставен от всички цели числа, кратни на δ, т. е. той е главният идеал $[\delta]$. Друг аналогичен пръстен е пръстенът $P[x]$, съставен от всички полиноми на неизвестното x с коефициенти, принадлежащи на дадено поле P. Всеки идеал в пръстена $P[x]$ е главен идеал. Нека $\mathfrak{m}$ е ненулев идеал от $P[x]$. С $\varphi(x)$ да означим един полином от идеала, който е с най-ниска степен, и с $r(x)$ да означим остатъка от делението на произволен полином от идеала с полинома $\varphi(x)$, като $g(x)$ е частното от делението. Тогава ще имаме равенството $f(x) = \varphi(x) q(x) + r(x)$ и по напълно аналогичен начин на предидущия установяваме, че $r(x) = 0$.

В теорията на групите видяхме, че при хомоморфно изобразяване на една група $\mathbf{G}$ върху друга група $\bar{\mathbf{G}}$ образът на инвариантна подгрупа $\mathbf{A}$ на $\mathbf{G}$ е единица в $\bar{\mathbf{G}}$ и съседните комплекси на $\mathbf{A}$ се изобразяват в другите елементи на $\bar{\mathbf{G}}$. Подобно свойство съществува и при хомоморфното изобразяване на един пръстен върху друг пръстен. Ще установим теоремата:

При хомоморфното изобразяване на пръстена R върху пръстена $\overline{R}$ елементите на R се разпределят в класи, като класът $\mathfrak{m}$ от елементи, на които образът в $\overline{R}$ е нулевият елемент, образуват един идеал в R и другите класи са класите, остатъци спрямо този идеал.

Действително нека a и b са произволни елементи от $\mathfrak{m}$. При хомоморфното изобразяване те преминават в нулата на R. Следователно елементът $-b$ и впоследствие $a-b$ преминава също в нулата. Ако c е произволен елемент от R, то ac преминава в елемента 0. $\overline{c}=0$, т. е. също в нулата на пръстена $\overline{R}$. Значи съвкупността $\mathfrak{m}$ е идеал в R. Нека сега α е елемент от R, който не принадлежи на идеала $\mathfrak{m}$. Да означим с $\overline{\alpha}$ образа му в $\overline{R}$. Нека α_1 е кой да е друг елемент от R, на който образът е също елементът $\overline{\alpha}$. Тогава елементът $\alpha_1-\alpha$ преминава в нулата от $\overline{R}$ и следователно той принадлежи на идеала $\mathfrak{m}$. Обратно, всеки елемент $\beta-\alpha$, $\alpha\sim\overline{\alpha}$, $\beta\sim\overline{\alpha}$ преминава в нулата на $\overline{R}$. Следователно всичките елементи от R, на които образът е $\overline{\alpha}$, образуват един клас остатъци спрямо $\mathfrak{m}$. Подобно другите елементи на $\overline{R}$ са образи на останалите класове остатъци.

Изложеният начин за доказване на предната теорема ни позволява да намерим пръстен $\overline{R}$, хомоморфен на даден пръстен R, на който елементите са образи на класите остатъци спрямо даден идеал на R. Именно нека $\mathfrak{m}$ е идеал от пръстена R. Пръстенът R се разделя на класи остатъци спрямо $\mathfrak{m}$, които да означим с $R_a, R_b, R_c, \ldots$, между които фигурира естествено и идеалът $\mathfrak{m}$. Както знаем, всеки клас се състои от елементи, конгруентни с един кой да е елемент от него спрямо модул $\mathfrak{m}$. Под сума R_a+R_b на класовете R_a и R_b разбираме класа R_c, за който $c\equiv a+b\,(\mathfrak{m})$. Класът R_c е така напълно определен. Действително, ако a_1 и b_1 са произволни елементи от R_a и R_b, то на основание на свойствата на конгруенциите по модул $\mathfrak{m}$ имаме

$$a_1\equiv a,\quad b_1\equiv b,$$

$$a_1+b_1\equiv a+b\equiv c.$$

Под произведение на класите R_a и R_b разбираме класа R_d, определен с $d\equiv ab$. Ако a_1 и b_1 са произволни елементи от R_a и R_b, то от конгруенциите $a_1\equiv a$, $b_1\equiv b$ следва конгруенцията $a_1b_1\equiv ab\equiv d$, което показва, че така дефинираният клас R_d е еднозначно определен. Да означим тогава с $\overline{R}$ съвкупността, съставена от елементите $\overline{a}=R_a$, $\overline{b}=R_b$, $\overline{c}=R_c$, . . . Предните изводи показват, че сумата и произведението на два кои да е елемента от $\overline{R}$ са също елементи от $\overline{R}$ и следователно $\overline{R}$ представлява пръстен — хомоморфен образ на пръстена R. При това нулевият елемент на $\overline{R}$ отговаря на идеала $\mathfrak{m}$ и другите му елементи са образи на класовете остатъци по модул $\mathfrak{m}$. Този пръстен се нарича пръстен от класовете остатъци и се бележи с

$R/\mathfrak{m}$. В него на равенство на два елемента отговаря конгруентност по модул $\mathfrak{m}$ в пръстена R.

Ако C е пръстенът от целите числа, то, както видяхме, всеки идеал $\mathfrak{m}$ в него е главен идеал, т. е. $\mathfrak{m}=(m)$, като m е цяло число, което може да се приеме за положително, ако идеалът не е нулев. Класовете остатъци спрямо m да означим с $C_0, C_1, C_2, \ldots, C_{m-1}$. C_r означава съвкупността на всичките цели числа, които имат остатък r при делението им с m. В случая пръстенът $C/\mathfrak{m}$ се състои точно от елементите $C_0, C_1, C_2, \ldots, C_{m-1}$.

Понятието за делимост в теорията на целите числа и полиномите се разширява и за идеалите. Нека $\mathfrak{m}$ е един идеал от пръстена R. Ако a е елемент от $\mathfrak{m}$, то казваме, че a е делим на идеала $\mathfrak{m}$. Това означаваме на основание на въведените конгруенции спрямо един идеал с конгруенцията $a \equiv 0\,(\mathfrak{m})$. Ако всички елементи на един идеал $\mathfrak{a}$ са делими с $\mathfrak{b}$, т. е. принадлежат на $\mathfrak{b}$, то казваме, че идеалът $\mathfrak{a}$ е делим на идеала $\mathfrak{b}$ и това означаваме с

$$\mathfrak{a} \equiv 0\,(\mathfrak{b}).$$

Идеалът $\mathfrak{b}$ се нарича делител на $\mathfrak{a}$, а идеалът $\mathfrak{a}$ — кратен на $\mathfrak{b}$. В пръстена C от целите числа всеки идеал е главен идеал. Идеалът (m) е делител на идеала (m_1), ако цялото число m дели цялото число m_1, понеже съгласно с дефиницията m_1 трябва да е елемент от идеала (m), т. е. кратно число на m, и всеки елемент от (m_1) принадлежи на (m), като число, кратно на m_1 и следователно и на m.

Идеалът $\mathfrak{m}$ от пръстена R се нарича прост идеал, ако пръстенът от остатъци спрямо $\mathfrak{m}$ $R/\mathfrak{m}$ е област на цялостност. Следователно, ако $\bar{a}, \bar{b}, \ldots$ са елементите на $R/\mathfrak{m}$, то от равенството $\bar{a}\bar{b}=0$ и $\bar{a} \neq 0$ следва, че $\bar{b}=0$ за кои да е елементи от него. Ако a е кой да е елемент от класа, чийто образ е $\bar{a}$, то равенството $\bar{a}\bar{b}=0$ може да се замести с конгруенцията $ab \equiv 0\,(\mathfrak{m})$. Следователно дефиницията за прост идеал може да се замести със следното условие: от $ab \equiv 0\,(\mathfrak{m})$ (a и b елементи от различни класове) и $a \not\equiv 0\,(\mathfrak{m})$ следва, че $b \equiv 0\,(\mathfrak{m})$. Един идеал $\mathfrak{p}$ се нарича неделим, ако се дели само на себе си и на единичния идеал. Едно основно свойство е изразено със следната теорема:

Ако идеалът $\mathfrak{p}$ от пръстена R, притежаващ единица, е неделим и различен от R, то той е прост и пръстенът от класите остатъци $R/\mathfrak{p}$ е поле. Обратно, ако $R/\mathfrak{p}$ е поле, то $\mathfrak{p}$ е неделим.

За да установим, че $R/\mathfrak{p}$ е поле, трябва да докажем, че всяко уравнение

$$\bar{a}\bar{x}=\bar{b}$$

при $\bar{a} \neq 0$ има решение. Нека a не е елемент от $\mathfrak{p}$ и b — произволен елемент. Да разгледаме съвкупността F от елементи $k+al$, където k е произволен елемент от идеала $\mathfrak{p}$ и l е произволен елемент от пръстена R. Съвкупността F е идеал от R, понеже разликата от два елемента от F има същата форма, както и произведението на елемент от

F с елемент β от R. От друга страна, F съдържа $\mathfrak{p}$ и не съвпада с $\mathfrak{p}$. Значи F е делител на $\mathfrak{p}$ и понеже $\mathfrak{b}$ е неделим, то F трябва да съвпада с R. Следователно всеки елемент $\mathfrak{b}$ от R има формата $k+al$, където $k \in \mathfrak{p}$ и $l \in R$. За образите на предните елементи следва тогава, че $\bar{b}=\bar{a}\bar{e}$, т. е. уравнението $\bar{a}\bar{x}=\bar{b}$ има решение. Така установихме, че $R/\mathfrak{p}$ е поле и понеже в полето няма делители на нулата, то $\mathfrak{p}$ е прост идеал.

Обратно, да допуснем, че $R/\mathfrak{p}$ е поле. Ще установим, че идеалът $\mathfrak{p}$ е неделим. Нека $\mathfrak{a}$ е идеал, който дели $\mathfrak{p}$, но не съвпада с него. Ще има тогава поне един елемент a от $\mathfrak{a}$, който не принадлежи на $\mathfrak{p}$. Понеже $R/\mathfrak{p}$ е поле, то конгруенцията $ax \equiv b\,(\mathfrak{p})$ ще има винаги решение x за всеки елемент b и R. От предната конгруенция следва тогава конгруенцията $ax \equiv b\,(\mathfrak{a})$ или $b \equiv 0\,(\mathfrak{a})$, т. е. b принадлежи на идеала $\mathfrak{a}$. Последното свойство показва, че идеалът $\mathfrak{a}$ трябва да съвпада с пръстена R и следователно идеалът $\mathfrak{p}$ е прост, понеже се дели само на себе си и на единичния идеал R.

Ако $\mathfrak{m}$ и $\mathfrak{g}$ са два идеала, то лесно се вижда, че съвкупността, съставена от елементите $a+b$, където a е произволен елемент от $\mathfrak{m}$ и b произволен елемент от $\mathfrak{g}$, е също идеал, който бележим с $(\mathfrak{m}, \mathfrak{g})$. Действително, ако и a_1, и b_1 са също елементи съответно от идеалите $\mathfrak{m}$ и $\mathfrak{g}$, то разликата $a_1+b_1-(a+b)$ като равна на $(a_1-a)+(b_1-b)$ е елемент от същата форма и произведението от $a+b$ с елемент c от пръстена R има също така формата на сума от елемента ac от $\mathfrak{m}$ и елемент bc от $\mathfrak{g}$. Очевидно $(\mathfrak{m}, \mathfrak{g})$ е общ делител на $\mathfrak{m}$ и $\mathfrak{g}$ и той се дели от всеки общ делител на тези идеали. По тази причина той се нарича общ най-голям делител на идеалите $\mathfrak{m}$ и $\mathfrak{g}$. Подобно се вижда, че съвкупността $\mathfrak{A}$, съставена от общите елементи на идеалите $\mathfrak{m}$ и $\mathfrak{g}$, е също така идеал, наречен общо най-малко кратно, понеже този идеал се дели на двата идеала $\mathfrak{m}$ и $\mathfrak{g}$ и всеки идеал, който се дели на тези идеали, се дели и на него.

Ако P е поле, то освен нулевия идеал в него има само един идеал, който е главен идеал (1), понеже всеки идеал, който съдържа елемент, различен от нула, ще съдържа и елемента $a\,a^{-1}=1$. Както видяхме, в пръстена от целите числа и в този от полиномите на неизвестното x с коефициенти от дадено поле всичките идеали са главни идеали. Една област на цялостност с единичен елемент, в която всеки идеал е и главен идеал, се нарича пръстен с главни идеали. Такива са разгледаните по-горе пръстени. В тези пръстени известното разлагане на целите числа в произведение на прости числа, което е еднозначно до множители единици, остава в сила. На доказателството не ще се спираме.

9. Алгебрично и трансцендентно разширение. Нека R е комутативен пръстен без делители на нулата и притежаващ единица e, т. е. област на цялостност с единица. С R' да означим друг пръстен, съдържащ R, който е комутативен и притежава единица, която очевидно трябва да съвпада с e, понеже в противен случай би имало две единици в R'. Нека α е произволен елемент от R' и да извършим върху елемента α и елементите от R действията събиране, изваждане и умно-

жение. В резултат на това ще получим елементи от пръстена R', които ще имат вида

$$b_0\alpha^{\lambda_0}+b_1\alpha^{\lambda_1}+\ldots+b_p\alpha^{\lambda_p}, \tag{1}$$

гдето $b_0, b_1, \ldots, b_p$ са елементи от пръстена R. Изразът (1) представлява полином на α. Сборът, разликата и произведението на елементи от вида (1) са също елементи от същия вид, т. е. множеството от всичките елементи (1) образува пръстен, който очевидно е комутативен. Нека го означим с $R[\alpha]$. В частност при $\lambda_0=0$, $b_1=b_2=\ldots=b_p=0$ елементът (1) се свежда на елемента b_0 и следователно пръстенът $R[\alpha]$ съдържа пръстена R. Ако α не принадлежи на R, то пръстенът R ще бъде същински подпръстен на $R[\alpha]$, който от своя страна е подпръстен на R'. В случай, че два елемента

$$c_0+c_1\alpha+c_2\alpha^2+\ldots+c_m\alpha^m,$$

$$d_0+d_1\alpha+d_2\alpha^2+\ldots d_m\alpha^m$$

са равни, то ще имаме

$$(c_0-d_0)+(c_1-d_1)\alpha+\ldots+(c_m-d_m)\alpha^m=0.$$

Ако не всички коефициенти са равни на нула, то получаваме, че α е корен на едно уравнение с коефициенти, които принадлежат на R. В такъв случай казваме, че елементът α е **алгебрически** спрямо пръстена R. В противен случай, т. е. ако равенството от вида

$$p_0+p_1\alpha+p_2\alpha^2+\ldots+p_k\alpha^k=0$$

при елементи $p_0, p_1, \ldots, p_k$ от R е възможно само тогава, когато елементите $p_0, p_1, \ldots, p_k$ са равни на нула, елементът α се нарича **трансцендентен** спрямо R.

Когато елементът α е трансцендентен относно R, то той се нарича **неизвестно** и обикновено го бележим с x. Тогава два полинома на x са равни само тогава, когато коефициентите им съответно са равни и за пръстена $R[\alpha]$ се прилага теорията от предния параграф, понеже определянето на този пръстен е идентично с това, което вече разгледахме.

Можем да получим пръстен $R[y]$ с неизвестно y при друго разширение на пръстена R. Ще докажем сега теоремата:

Ако x, y са неизвестни относно пръстена R, които са избрани от едно и също или две различни разширения R' и R'' на пръстена R, то пръстенът $R[x]$ от полиномите на x над R е изоморфен на пръстена $R[y]$ от полиномите на y над R.

На всеки полином

$$f(x)=a_0+a_1x+\ldots+a_nx^n$$

от $R[x]$ правим да съответствува полиномът

$$f(y)=a_0+a_1y+\ldots+a_ny^n$$

от $R[y]$, който има същите коефициенти, $a_i \in R$, $0 \leqq i \leqq n$. Лесно се проверява, че това съответствие е изоморфно. Така нека

$$g(x)=b_0+b_1x+\ldots+b_mx^m$$

е друг полином от $R[x]$, на който ще съответствува полиномът

$$g(y)=b_0+b_1y+\ldots+b_my^m$$

от $R[y]$. Ако двата полинома $f(y)$ и $g(y)$ са равни, то трябва коефициентите им да са равни, т. е.

$$a_0=b_0, \quad a_1=b_1, \quad a_2=b_2, \ldots$$

(липсващите членове се считат с коефициенти, равни на нула). Но тогава и коефициентите на полиномите $f(x)$ и $g(x)$ са равни помежду си, т. е. $f(x)=g(x)$. Значи на разни полиноми от $R[y]$ отговарят разни полиноми от $R[x]$. Следователно съответствието $f(x) \longrightarrow f(y)$ е взаимно еднозначно.

За сумата на полиномите $f(x)$ и $g(x)$ имаме

$$f(x)+g(x)=\psi(x)=c_0+c_1x+c_2x^2+\ldots,$$

$$c_0=a_0+b_0, \quad c_1=a_1+b_1, \quad c_2=a_2+b_2, \ldots$$

На полинома $\psi(x)$ съответствува полиномът

$$\psi(y)=c_0+c_1y+c_2y^2+\ldots$$

от $R[y]$. Но очевидно имаме

$$\psi(y)=f(y)+g(y).$$

Следователно на полинома $f(x)+g(x)$ от $R[x]$ съответствува полиномът $f(y)+g(y)$ от $R[y]$. Подобно установяваме, че на полинома $f(x)g(x)$ съответствува полиномът $f(y)g(y)$, с което се привършва доказателството на теоремата.

Елементите на $R[x]$, които не принадлежат на R, са трансцендентни относно R.

Допускаме противното, т. е. съществува елемент

$$y=a_0+a_1x+\ldots+a_nx^n, \quad a_n \neq 0 \tag{2}$$

от $R[x]$, който удовлетворява уравнението

$$b_0y^m+b_1y^{m-1}+\ldots+b_m=0, \quad b_0 \neq 0, \tag{3}$$

на което коефициентите $b_0, b_1, \ldots, b_m$ са елементи от R. Като заместим в (3) y с равното му от (2), получаваме

$$b_0(a_nx^n+a^{n-1}x^{n-1}+\ldots+a_0)^m+b_1(a_nx^n+\ldots+a_0)^{m-1}+\ldots+b_m=0.$$

Като преработим лявата част на това равенство, получаваме

$$b_0a_n^m x^{nm}+c_1x^{nm-1}+\ldots+c_{nm}=0,$$

гдето $b_0 a_n^m$, c_1, $c_2, \ldots, c_{nm}$ ще бъдат елементи от R. Понеже x е трансцендентен елемент относно R, то коефициентите на това уравнение трябва да са равни на нула и специално $b_0 a_n^m = 0$, което противоречи на условието, че в пръстена R няма делители на нулата.

Ще разгледаме сега алгебричното разширение по-подробно. При това с оглед на приложенията ще излезем от поле вместо от пръстен. Нека P е произволно поле и α е алгебричен елемент относно него. Множеството от полиномите на α

$$f(\alpha) = p_0 + p_1\alpha + \ldots + p_m\alpha^m \tag{4}$$

с коефициенти от P образува пръстен, който ще означаваме с $P[\alpha]$. Това свойство е очевидно, понеже полето е същевременно комутативен пръстен без делители на нулата и притежава единица. Можем да предположим, че уравнението

$$g(x) = a_0 + a_1 x + \ldots + a_n x^n = 0, \tag{5}$$

което удовлетворява елемента α, е неразложимо в P. Действително, ако въпросното уравнение е разложимо, то представяме лявата му част като произведение на неразложими множители, които са полиноми от $P[x]$, и този от тях, който се анулира за $x = \alpha$, ще вземем за лявата част на уравнението (5). Нека отбележим, че степента на полинома $g(x)$ ще бъде най-малко равна на 2. В противен случай елементът α би принадлежал на полето P и не бихме имали никакво разширение на P.

Елементите (4) на пръстена $P[\alpha]$ могат да се представят във формата

$$q_0 + q_1\alpha + \ldots + q_{m-1}\alpha^{m-1}, \tag{6}$$

гдето $q_0, q_1, \ldots, q_{m-1}$ са елементи от P и $m \leqq n$. В това се убеждаваме по познат вече начин. Именно да разделим $f(x)$ с $g(x)$:

$$f(x) = g(x)\,Q(x) + R(x).$$

Тук полиномът $R(x)$ е от степен, по-малка от n, и принадлежи на $P[x]$. Но при $x = \alpha$ получаваме

$$f(\alpha) = R(\alpha)$$

и твърдението е установено.

Нека отбележим, че (6) е само тогава нула, когато всичките коефициенти $q_0, q_1, \ldots, q_{m-1}$ са равни на нула. Действително в противен случай бихме получили, че α е корен на уравнение от по-ниска степен от n, т. е. неразложимото уравнение (5) би имало общ корен с уравнение от по-ниска степен, което противоречи на известна нам теорема.

Да образуваме сега всички отношения от вида

$$\frac{\varphi(\alpha)}{\psi(\alpha)}, \tag{7}$$

$$\varphi(\alpha) = b_0 + b_1\alpha + \ldots + b_m\alpha^m, \quad \psi(\alpha) = c_0 + c_1\alpha + \ldots + c_k\alpha^k$$

в които коефициентите b_i, c_i са елементи от P и делителите не са равни на нула, като не правим никакво ограничение за степените m и k. Пон-же сумата, разликата и произведението на отношения от вида (7) имат същия вид, то множеството отношения (7) ще образуват поле, което ще означим с $P(\alpha)$. Полето $P(\alpha)$ съдържа полето P, понеже (7) при $\varphi(\alpha)=b_0$, $\psi(\alpha)=c_0$ се свежда на елемент от P. Ако $\psi(\alpha)=c_0$, то отношението (7) има формата (4). Нека $k \geqq 1$, $c_k \neq 0$. На основание на предните резултати можем полинома в знаменателя на (7) да представим във формата

$$\psi(\alpha)=d_0+d_1\alpha+\ldots+d_p\alpha^p,$$

гдето степента p е по-малка от n. Понеже знаменателят е отличен от нула, коефициентите на полинома $\psi(x)$ не могат всички да бъдат равни на нула. Полиномът $\psi(x)$ не е равен на нула и видяхме, че е взаимно прост с полинома $f(x)$. По известен резултат ще съществуват два полинома $F(x)$ и $\Phi(x)$ с коефициенти, от P, за които ще имаме

$$f(x)F(x)+\psi(x)\Phi(x)=e.$$

Като поставим тук $x=\alpha$, получаваме

$$\Phi(\alpha)=\frac{e}{\psi(\alpha)}.$$

С това се вижда, че отношението (7) е от формата (4).

В заключение на горното можем да изкажем следната теорема:

Пръстенът $P[\alpha]$ и полето $P(\alpha)$ са идентични.

Полето $P(\alpha)$ се нарича алгебрическо разширение на полето P с присъединяване (адюнгиране) към P на елемента α.

Например нека P е едно поле от числа и ω не е точен квадрат на число от P. Тогава полето $P(\sqrt{\omega})$ ще бъде съставено от всички числа от формата

$$a+b\sqrt{\omega},$$

гдето a и b са произволни числа от P. Независимо от изложената теория това се вижда непосредствено от равенствата

$$(a+b\sqrt{\omega})+(a_1+b_1\sqrt{\omega})=(a+a_1)+(b+b_1)\sqrt{\omega},$$

$$(a+b\sqrt{\omega})(a_1+b_1\sqrt{\omega})=(aa_1+bb_1\omega)+(ab_1+a_1b)\sqrt{\omega},$$

$$\frac{a+b\sqrt{\omega}}{a_1+b_1\sqrt{\omega}}=\frac{aa_1-bb_1\omega}{a_1^2-b_1^2\omega}+\frac{a_1b-ab_1}{a_1^2-b_1^2\omega}\sqrt{\omega}.$$

От важно значение е и следната теорема:

Всеки елемент от полето $P(\alpha)$ е алгебричен относно полето P.

Нека $\beta=f(\alpha)$ е произволен елемент от $P(\alpha)$. Елементът α е корен на уравнението (5). Очевидно произведенията $\beta\alpha, \beta\alpha^2, \ldots, \beta\alpha^{n-1}$ са също елементи от полето $P(\alpha)$ и следователно можем да пишем

$$\beta\alpha^s=a_{so}+a_{s1}\alpha+\ldots+a_{s,n-1}\alpha^{n-1}, \quad s=0, 1, 2, \ldots, n-1,$$

гдето $a_{s0}, a_{s1}, \ldots, a_{s,n-1}$ са елементи от P. Но предните равенства лесно се преобразуват във формата

$$a_{00}-\beta+a_{01}\alpha+a_{02}\alpha^2+\ \ldots\ +a_{0,n-1}\alpha^{n-1}=0,$$

$$a_{10}+(a_{11}-\beta)\alpha+\alpha_{12}\alpha^2+\ \ldots\ +a_{1,n-1}\alpha^{n-1}=0,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_{n-1,0}+a_{n-1,1}\alpha+a_{n-1,2}\alpha^2+\ \ldots\ +(a_{n-1,\,n-1}-\beta)\alpha^{n-1}=0.$$

Тези уравнения са линейни спрямо неизвестните $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ и по теоремата за хомогенни системи от линейни уравнения трябва детерминантата от коефициентите пред неизвестните $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ да бъде равна на нула:

$$\begin{vmatrix} a_{00}-\beta & a_{01} & a_{02} & \ldots & a_{0,n-1} \\ a_{10} & a_{11}-\beta & a_{12} & \ldots & a_{1,n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1,0} & a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,\,n-1}-\beta \end{vmatrix}=0.$$

Следователно β ще бъде корен на уравнението

$$\begin{vmatrix} a_{00}-x & a_{01} & a_{02} & \ldots & a_{0,n-1} \\ a_{10} & a_{11}-x & a_{12} & \ldots & a_{1,n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1,0} & a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n-1}-x \end{vmatrix}=0,$$

коефициентите на което се вижда лесно, че принадлежат на полето P, с което теоремата е установена.

10. **Съществуване на корен.** Нека P е едно произволно поле и $f(x)$ е произволен полином на неизвестното x с коефициенти, които принадлежат на P, т. е. полином от пръстена $P[x]$. Интересно е да се реши следният въпрос: съществува ли надполе P' на P, такова, че в него уравнението $f(x)=0$ да има корен. Отговорът на този въпрос е утвърдителен. Имаме следната теорема:

За всяко поле P и за всеки полином $f(x)$ от пръстена $P[x]$ съществува такова разширение на полето, в което уравнението $f(x)=0$ има корен.

Очевидно можем да предположим, че полиномът $f(x)$ е неразложим и е от степен, по-голяма от 1, понеже всяко уравнение от първа степен има корен, принадлежащ на P. Разделяме пръстена $P[x]$ от полиномите на класи, като всеки клас се състои от всички полиноми, които имат един и същ остатък спрямо полинома $f(x)$. Следователно два полинома $\varphi(x)$ и $\psi(x)$ ще принадлежат на един и същ клас, ако разликата им се дели на $f(x)$, което отбелязваме с

$$\varphi(x)\equiv\psi(x)\ (\mathrm{mod}\ f(x)).$$

Горната формула се нарича сравнение (конгруенция) и казваме, че полиномът $\varphi(x)$ е сравним (конгруентен) с полинома $\psi(x)$ по модул $f(x)$. Конгруенциите при един и същ модул можем да ги събираме, вадим и умножаваме почленно. Така нека имаме конгруенциите

$$g(\ )\equiv h(x)(\mathrm{mod}\, f(x));$$
$$u(x)\equiv v(x)(\mathrm{mod}\, f(x)).$$

Понеже разликите $g(x)-h(x)$ и $u(x)-v(x)$ се делят на $f(x)$, то и разликата

$$g(x)+u(x)-[h(x)+v(x)]$$

ще се дели на $f(x)$, т. е. имаме конгруенцията

(1) $$g(x)+u(x)\equiv h(x)+v(x)(\mathrm{mod}\ f(x)).$$

Подобно установяваме конгруенцията

(2) $$g(x)u(x)\equiv h(x)v(x)(\mathrm{mod}\, f(x)).$$

Като представим разликата $g(x)\ u\ (x)-h(x)v(x)$ във формата

$$g(x)u(x)-h(x)v(x)=[g(x)-h(x)]u(x)+[u(x)-v(x)]h(x),$$

виждаме, че тази разлика се дели на $f(x)$, т. е. ще имаме (2).

Също лесно се установяват следните свойства:

1. Ако $g(x)\equiv h(x)(\mathrm{mod}\, f(x))$, $t(x)\equiv h(x)(\mathrm{mod}\, f(x))$, то $g(x)\equiv t(x)$ $(\mathrm{mod}\, f(x))$.

2. От $g(x)\equiv h(x)(\mathrm{mod}\, f(x))$ следва за всеки полином $\varphi(x)$ от $P[x]$ конгруенцията $\varphi(x)g(x)\equiv\varphi(x)h(x)(\mathrm{mod}\, f(x))$.

Да означим класите, на които се разпада пръстенът $P[x]$, с буквите α, β, γ,... Ще докажем, че множеството P', съставено от елементите α, β, γ,... при подходящо определяне на сбор и произведение, е поле. Всеки клас може да бъде определен с кой да е негов елемент $\varphi(x)$. Именно елементите от този клас са полиномите от $P[x]$, които са сравними с $\varphi(x)$ по модул $f(x)$. Нека $\varphi_1(x)$ и $\psi_1(x)$ са полиноми от класа α и класа β. Да означим с γ класа, в който се намира полиномът $\varphi_1(x)+\psi_1(x)$, и с δ класа, съдържащ полинома $\varphi_1(x)\psi_1(x)$. Нека $\varphi_2(x)$ и $\psi_2(x)$ са други кои да е полиноми съответно от класа α и класа β. Тогава на основание на (1) и на (2) заключаваме, че полиномите $\varphi_2(x)+\psi_2(x)$ и $\varphi_2(x)\psi_2(x)$ принадлежат съответно на класа γ и δ. Така виждаме, че класовете γ и δ не зависят от избора на представителите на класовете α и β. Класа γ наричаме сума на класовете α и β и бележим $\gamma=\alpha+\beta$, а класа δ — произведение на α и β и означаваме това с $\delta=\alpha\beta$. Класът, съдържащ елемента 0, т. е. състоящ се от всички полиноми от $P[x]$, които се делят на $f(x)$, се нарича нулев и ще го означаваме с 0. Противоположен клас на класа α, съставен от полиномите, които по модул $f(x)$ имат остатък $\varphi(x)$, е класът от полиноми, които спрямо $f(x)$ имат остатък $-\varphi(x)$. Класът, съставен от всички полиноми, конгруентни с 1 по модул $f(x)$, ще наричаме единица и ще го бележим с E. Нека α е отличен от нула клас и $\varphi(x)$ е един произ-

волен полином от него. Понеже този полином не се дели на $f(x)$ и последният полином е неразложим, то $\varphi(x)$ и $f(x)$ са взаимно прости полиноми. Следователно ще има полиноми $F(x)$ и $\Phi(x)$ от $P[x]$, които удовлетворяват на равенството

$$\varphi(x)F(x)+f(x)\Phi(x)=e.$$

Оттук следва непосредствено, че $\varphi(x)F(x)\equiv e\ (\mathrm{mod}\, f(x))$, т. е. полиномът $\varphi(x)F(x)$ принадлежи на класа E. Ако означим с β класа, на който принадлежи полиномът $F(x)$, то следва, че

$$\alpha\beta=E,$$

отгдето имаме $\beta=\alpha^{-1}$. Така установихме, че всеки ненулев елемент има обратен, с което доказателството на предложението, че множеството от класове е поле, се завършва.

Да означим полученото така поле с P'. На всеки елемент a от полето P съответствува класът, съставен от всички полиноми от $P[x]$, които са конгруентни на a по модул $f(x)$. Очевидно и елементът a като конгруентен на себе си ще принадлежи на този клас. Множеството от всички такива класове, съответствуващи на елементите на P, ще образуват поле P_1, което е подполе на P' и е изоморфно с полето P. Действително взаимната еднозначност на съответствие на полетата P и P_1 е очевидна. Като избираме за представители на класовете елементите от P, които те съдържат, то ясно е, че сумата и произведението на елементите от полето P_1 съответствува на сумата и произведението на отговарящите им елементи от полето P. На това основание ние не ще различаваме елементите от P_1 от елементите на P.

Нека X е класът от полиноми, които са конгруентни с x по модул $f(x)$, и нека $g(x)$ е произволен полином от този клас. Ще имаме

$$g(x)\equiv x\ [\mathrm{mod}\, f(x)]. \tag{3}$$

Нека

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n.$$

На основание на (2) от (3) получаваме

$$g^k(x)\equiv x^k\ [\mathrm{mod}\, f(x)],\quad k=1,\ 2,\ldots,n$$

и понеже $a_n\equiv a_n[\mathrm{mod}\, f(x)]$, то ще имаме

$$T(x)=a_0g^n(x)+a_1g^{n-1}(x)+a_2g^{n-2}(x)+\ldots+a_n\equiv 0\ [\mathrm{mod}\, f(x)].$$

Предната конгруенция показва, че полиномът $T(x)$ принадлежи на нулевия клас, т. е. за класа X имаме уравнението

$$a_0X^n+a_1X^{n-1}+a_2X^{n-2}+\ldots+a_n=0,$$

с което изказаната теорема е установена.

Поле на разлагане на полинома $f(x)$ наричаме такова поле, в което $f(x)$ се разлага на линейни множители. Ще установим теоремата:

За всеки полином $f(x)$ от пръстена $P[x]$ съществува надполе на P, което е поле на разлагане за полинома $f(x)$.

Действително, ако $f(x)$ се разлага в P на линейни множители, то P е полето на разлагане. Ако нямаме този случай, нека в полето P полиномът $f(x)$ се разлага на неразложимите множители

$$f(x)=(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_p)\varphi_1(x)\varphi_2(x)\ldots\varphi_k(x).$$

Като присъединяваме корените $\alpha_1, \alpha_2, \ldots, \alpha_p$ към полето P, то това поле очевидно не се изменя, понеже въпросните корени принадлежат на него. Да разширим полето P до такова поле P', в което уравнението $\varphi_1(x)=0$ има корен. Тогава полиномът $\varphi_1(x)$ ще бъде разложим в полето P' и може евентуално и други от полиномите $\varphi_2(x), \ldots, \varphi_k(x)$ да бъдат разложими в P'. Следователно полиномът $f(x)$ може да се представи във формата

$$f(x)=(x-\beta_1)(x-\beta_2)\ldots(x-\beta_q)\psi_1(x)\ldots\psi_m(x),$$

където корените $\beta_1, \ldots, \beta_q$ принадлежат на P' и полиномите $\psi_1(x)$, $\psi_2(x), \ldots, \psi_m(x)$ са неразложими в P'. Разширяваме сега полето P' до надполе P'', в което уравнението $\psi_1(x)=0$ има поне един корен. Тогава този полином става разложим в P'' и полиномът $f(x)$ ще се разлага в P'' на произведение на линейни множители и на неразложими полиноми в полето P''. Като продължаваме така, очевидно получаваме последователно полета, в които броят на линейните множители в разлагането на полинома $f(x)$ расте. Следователно с краен брой подобни разширения на полетата ще получим поле Δ, в което полиномът $f(x)$ се разлага на произведение от линейни множители, броят на които е равен на степента на полинома. Следователно в полето Δ уравнението $f(x)=0$ има толкова корена, колкото степента му. Полето Δ се нарича **поле на разлагане** на полинома $f(x)$.

11. Крайно разширение и полета на Галоа. Нека P е поле и Γ е надполе на P, т. е. поле, което съдържа P. Ако $u_1, u_2, \ldots, u_k$ са елементи от Γ и $\lambda_1, \lambda_2, \ldots, \lambda_k$—елементи от P, то очевидно линейната комбинация $\lambda_1 u_1+\ldots+\lambda_k u_k$ представлява елемент от полето Γ. Казваме, че елементите $u_1, u_2, \ldots, u_p$ от полето Γ са **линейно зависими**, ако съществуват p елемента μ_l от полето P, поне единият от които е отличен от нула, за които имаме

$$\mu_1 u_1+\mu_2 u_2+\ldots+\mu_p u_p=0.$$

В противен случай елементите $u_1, u_2, \ldots, u_p$ се наричат **линейно независими**. Да предположим, че в полето Γ съществуват m линейно независими елементи $u_1, u_2, \ldots, u_m$ и че всеки $(m+1)$-елементи от него са линейно зависими. Тогава по напълно аналогичен начин, както в линейната алгебра, установяваме, че всеки елемент от полето Γ ще има формата

$$\tau_1 u_1+\tau_2 u_2+\ldots+\tau_m u_m,$$

където $\tau_1, \tau_2, \ldots, \tau_m$ са елементи от полето P. Следователно елементите $u_1, u_2, \ldots, u_m$ образуват базис на полето Γ. По известно свойство

от линейната алгебра всеки m-елементи от Γ, които са линейни форми на елементите $u_1, u_2, \ldots, u_m$ с коефициенти от полето P, детерминантата от които е отлична от нула, образуват също базис на полето. В разглеждания случай казваме, че полето Γ е получено от полето P с крайно разширение. Числото m на линейно независимите елементи се нарича степен на Γ относно P и се бележи обикновено с $(\Gamma:P)$. Очевидно тази степен не зависи от избрания базис.

Всеки елемент α от полето Γ е корен на алгебрично уравнение от m-та степен с коефициенти, принадлежащи на полето P.

Действително елементите $1, \alpha, \alpha^2, \ldots, \alpha^m$ принадлежат на Γ и броят им $m+1$ надминава m. Следователно между тях ще има линейна връзка

$$\lambda_0+\lambda_1\alpha+\lambda_2\alpha^2+\ldots+\lambda_m\alpha^m=0$$

с коефициенти, принадлежащи на P, от които единият поне е отличен от нула.

Ще установим няколко основни теореми за крайното разширение на полетата.

Ако Γ е крайно разширение на полето P и L е крайно разширение на Γ, то L е също крайно разширение на P и за степените имаме

$$(L:P)=(L:\Gamma)(\Gamma:P).$$

Нека $u_1, u_2, \ldots, u_m$ е един базис на L спрямо Γ и $v_1, v_2, \ldots, v_k$ е базис на Γ спрямо P. Всеки елемент α от L има формата

$$\alpha=s_1u_1+s_2u_2+\ldots+s_mu_m,$$

където $s_1, s_2, \ldots, s_m$ са елементи от Γ. Но последните елементи като принадлежащи на Γ ще имат формата

$$s_i=t_{i1}v_1+t_{i2}v_2+\ldots+t_{ik}v_k, \quad i=1, 2, \ldots, m,$$

където $t_{i1}, t_{i2}, \ldots, t_{ik}$ са елементи от полето P. Като заместим тези стойности на s_i в α, получаваме за α израза

$$\alpha=\sum_{i=1}^{m}u_i\sum_{j=1}^{k}t_{ij}v_j=\sum_{i=1}^{m}\sum_{j=1}^{k}t_{ij}u_iv_j.$$

Но произведенията u_iv_j са елементи от полето L. Следователно полето L представлява крайно разширение на P. За да установим, че степента на L спрямо P е равна точно на mk, ще трябва да докажем, че елементите u_iv_j са линейно независими. Да допуснем обратното, т. е. че между тях съществува връзката

$$\sum_{i=1}^{m}\sum_{j=1}^{k}\tau_{ij}u_iv_j=0,$$

където τ_{ij} са елементи от P. Предното равенство можем да пишем така:

$$\sum_{i=1}^{m} u_i \sum_{j=1}^{k} \tau_{ij} v_j = 0. \tag{1}$$

Ако положим

$$w_j = \sum_{j=1}^{k} \tau_{ij} v_j,$$

равенството (1) става

$$\sum_{i=1}^{m} u_i w_i = 0.$$

Но елементите $u_1, u_2, \ldots, u_m$ са линейно независими в L и понеже $w_1, w_2, \ldots, w_m$ са елементи от Γ, то следва, че трябва да бъдат равни на нула, т. е.

$$\sum_{i=1}^{k} \tau_{ij} v_j = 0 \qquad i = 1, 2, \ldots, m. \tag{2}$$

Но τ_{ij} са елементи от P и понеже $v_1, v_2, \ldots, v_k$ са линейно независими от (2), следва, че всичките елементи τ_{ij} са равни на нула. Така теоремата е установена.

Нека Γ е крайно разширение на полето P и Σ е някое разширение на P, което се съдържа в Γ. Тогава Σ е крайно разширение на P и степента $(\Sigma:P)$ е делител на $(\Gamma:P)$.

Да означим с n степента на Γ относно P. Понеже Σ е подполе на Γ, то в Σ ще има само краен брой линейно независими елементи. Нека $u_1, u_2, \ldots, u_m$ е една система от максимален брой линейно независими елементи в Σ. Очевидно ще имаме $m \leqq n$. Ако u е произволен елемент от Σ, то елементите $u, u_1, u_2, \ldots, u_m$ ще бъдат линейно зависими, т. е. ще съществува релацията

$$\lambda u + \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_m u_m = 0, \tag{3}$$

където $\lambda, \lambda_1, \lambda_2, \ldots, \lambda_m$ са елементи от P. В (3) елементът λ е сигурно отличен от нула, защото в противен случай бихме имали

$$\lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_m u_m = 0,$$

което равенство противоречи на факта, че елементите $u_1, u_2, \ldots, u_m$ са линейно независими. Но тогава от (3) получаваме

$$u = -\frac{\lambda_1}{\lambda} u_1 - \frac{\lambda_2}{\lambda} u_2 - \ldots - \frac{\lambda_m}{\lambda} u_m$$

и следователно елементите $u_1, u_2, \ldots, u_m$ образуват базис на полето Σ относно полето P и полето Σ е крайно разширение на P. Но полето Γ е също така крайно разширение на полето Σ. Действително, ако $v_1, v_2, \ldots, v_n$ е базис на Γ относно P, то същите елементи могат да се разглеждат като базис на Γ относително Σ. Тогава по предната теорема

имаме $(\Gamma:P)=(\Gamma:\Sigma)(\Sigma:P)$, откъдето непосредствено следва, че $(\Sigma:P$ дели $(\Gamma:P)$.

Нека P е едно поле и

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n$$

е полином с коефициенти, принадлежащи на P. Ако присъединим последователно към P нулите $x_1, x_2,\ldots,x_n$ на полинома $f(x)$, получаваме разширено поле Σ, което, както видяхме, се нарича поле на разлагане на полинома $f(x)$ относно P. От значение е следната теорема:

Нека Σ е полето на разлагане на полинома $f(x)$ относно полето P, на което принадлежат коефициентите на $f(x)$. Тогава всеки полином $\varphi(x)$ с коефициенти от P, който е неразложим в P, или няма нито един корен в Σ, или се разпада в Σ на линейни множители.

Понеже полето Σ се получава от P с присъединяване на корените $x_1, x_2,\ldots,x_n$, то елементите на Σ са полиноми на тези корени с коефициенти, принадлежащи на полето P. Нека полиномът $\varphi(x)$ има в Σ поне една нула y_1. Тогава y_1 ще бъде полином на $x_1, x_2,\ldots, x_n$:

$$y_1=\psi(x_1, x_2,\ldots,x_n).$$

Да означим с $y_1, y_2,\ldots,y_m$ стойностите на функцията ψ, когато разменяме $x_1, x_2,\ldots,x_n$ по всевъзможни начини. Известно ни е от трансформация на уравненията, че тези стойности са корени на уравнение

$$F(y)=y^m+A_1y^{m-1}+\ldots+A_m=0,$$

на което коефициентите са рационални функции от коефициентите на $f(x)$, т. е. принадлежат на полето P. Понеже $\varphi(x)$ е неразложим в P полином и $F(x)$ има поне една обща нула с него, то по известна теорема (§ 7) полиномът $F(x)$ трябва да се дели на $\varphi(x)$. Но $F(x)$ се разпада в Σ на произведение на линейни множители и следователно полиномът $\varphi(x)$ като делител на $F(x)$ ще се разпада на произведение на линейни множители в същото поле.

Ако полето P е числово, то всеки полином има толкова нули, колкото е степента му и полето на разлагане на полинома се получава с последователно присъединяване на нулите му.

Нека P е дадено поле. Всяко подполе на P, което не съвпада с P, се нарича собствено подполе на P. Едно поле се нарича *просто*, ако не притежава собствено подполе.

Всяко поле притежава едно просто поле и последното е единствено.

Действително сечението Q на всичките подполета на полето P е просто поле. В противен случай полето Q ще съдържа собствено подполе и сечението на подполетата на P не ще бъде полето Q, а част от него. Ако допуснем, че има две прости полета в полето P, то сечението им ще бъде също просто поле, което е тяхно собствено подполе, и въпросните полета не ще са прости.

Нека означим с P' простото подполе на даденото поле P. Очевидно P' ще съдържа нулата и единицата, която да означим с e.

Както видяхме и по-рано, P' ще съдържа и елементите ne, където n е цяло и произволно число, т. е. елементите

$$0,\ \pm e,\ \pm 2e,\ \pm 3e,\ldots \tag{4}$$

Да предположим, че елементите (4) са различни, т. е. ако $n \neq m$, то и $ne \neq me$ и да означим с Δ безкрайната съвкупност (4). Понеже за произволни цели числа m и n имаме

$$me+ne=(m+n)\,e,\quad me\,.\,ne=mne^2=mne,$$

то Δ е комутативен пръстен. Освен това от $me\,.\,ne=0$ следва, че или поне $m=0$, или $n=0$, т. е. $me=0$ или $ne=0$, което показва, че Δ е област на цялостност. Но тогава с въвеждане на елементите на отношения $\frac{pe}{qe}$ пръстенът Δ се разширява в поле Δ', което принадлежи на полето P'. Понеже от $ne=0$ следва, че $n=0$, то, както вече споменахме, полето P е с характеристика нула. Като направим да съответствува на всеки елемент $\frac{pe}{qe}$ от полето Δ' рационалното число $\frac{p}{q}$, лесно виждаме, че Δ' е изоморфно поле на полето от рационалните числа. Така достигаме до следния резултат:

Всяко поле с характеристика нула съдържа просто поле, което е изоморфно на полето от рационалните числа.

Да предположим сега, че елементите (4) не са всичките различни и следователно за някои цели числа p' и q' имаме $p'e=q'e$. Следвайки тогава начина на извеждане от § 2, получаваме, че съществува цяло положително число n, за което $ne=0$ и елементите

$$0,e,\ 2e;\ldots,\ (n-1)\,e \tag{5}$$

са различни помежду си. Освен това елементите (5) образуват само тогава поле, когато числото n е просто, $n=p$. Полето P' от тези елементи е изоморфно на полето от класите $A_0,\ A_1,\ldots,A_{p-1}$ на целите числа относно модул p т. е. на целите числа, които имат еднакви остатъци относно p. Полето P е с характеристика p. Така установихме, че всяко поле с характеристика p (p просто число) съдържа просто поле, изоморфно на полето от p-те класи на целите числа спрямо модул p.

Ако a и b са произволни елементи от поле P с характеристика p, то имаме

$$\begin{aligned}(a+b)^p&=a^p+b^p,\\(a-b)^p&=a^p-b^p.\end{aligned} \tag{6}$$

Действително по формулата за бинома имаме

$$(a+b)^p=a^p+\binom{p}{1}a^{p-1}b+\binom{p}{2}a^{p-2}b^2+\ldots+b^p. \tag{7}$$

Понеже p е просто число, то лесно се вижда, че биномните коефициенти

$$\binom{p}{\mu}=\frac{p(p-1)\ldots(p-\mu+1)}{\mu!},\quad 1\leqq\mu\leqq p-1,$$

се делят на p. Като вземем пред вид, че $pe=0$, то следва равенството

$$\binom{p}{\mu}a^{p-\mu}b^{\mu}=0,\quad 1\leqq\mu\leqq p-1$$

и от (7) получаваме първото равенство от (6). Второто равенство (6) се получава по подобен начин или като следствие от първото, като го приложим за елементите $a-b$ и b. Именно имаме

$$(a-b+b)^p=(a-b)^p+b^p,\quad \text{т. е. } (a-b)^p=a^p-b^p.$$

Ще отбележим, че по индуктивен път от (6) получаваме

$$(a+b)^{p^m}=a^{p^m}+b^{p^m}$$

за всяко естествено число m.

Ще разгледаме сега крайните полета, които се наричат полета на Галоа. Нека Γ е едно такова поле и n е броят на елементите му. Характеристиката на Γ е просто число, което да означим с p. Тази характеристика не може да бъде равна на нула, понеже броят на елементите на Γ е краен. Както видяхме, простото подполе Γ' на Γ се състои от p-те елемента $0,\ e,\ 2e,\ldots,(p-1)e$. Понеже полето Γ е крайно, то в него ще има краен брой линейно независими елементи спрямо полето Γ'. Нека m е максималният брой на линейно независимите елементи и $u_1,\ u_2,\ldots,u_m$ е една такава система от m елемента. Тогава Γ може да се разглежда като крайно разширение на Γ' и елементите от него ще имат формата

$$\alpha=\lambda_1u_1+\lambda_2u_2+\ldots+\lambda_mu_m,$$

където коефициентите λ_i принадлежат на Γ'. Но всеки коефициент λ_i може да взема p стойности. Следователно броят на елементите α на Γ ще бъде равен на p^m, т. е. $n=p^m$. Числото m е степента на Γ спрямо P'. Следователно съществува теоремата:

Броят на елементите на едно поле на Галоа е равен на степен на характеристиката му p. Степенният показател е равен на степента на Γ спрямо простото подполе Γ'.

Нека $\alpha\neq0$ е елемент от полето на Галоа Γ, броят на елементите на което е $n=p^m$. Да означим с $\alpha_1,\ \alpha_2,\ldots,\alpha_{n-1}$ елементите на Γ, като от Γ сме премахнали нулата. Очевидно елементът α ще бъде измежду тях. Произведенията

(8) $$\alpha\alpha_1,\ \alpha\alpha_2,\ldots,\alpha\alpha_{n-1}$$

са елементи от полето Γ. Ще установим, че те съвпадат, абстрахирайки се от реда им, с елементите $\alpha_1,\ \alpha_2,\ldots,\alpha_{n-1}$. Действително ни-

кой елемент от (8) не е равен на нула, понеже от $\alpha\alpha_i=0$ следва, че $\alpha_i=0$. Те са все различни, понеже от равенството $\alpha\alpha_i=\alpha\alpha_k$ следва $\alpha(\alpha_i-\alpha_k)=0$, от което имаме $\alpha_i-\alpha_k=0$, т. е. $\alpha_i=\alpha_k$. Следователно елементите (8) са всичките елементи на Γ, като сме изключили елемента нула. Но тогава произведението на елементите (8) е равно на произведението

$$\alpha_1\alpha_2 \ldots \alpha_{n-1}\alpha^{n-1}$$

или

$$\alpha_1\alpha_2 \ldots \alpha_{n-1}\alpha^{n-1}=\alpha_1\alpha_2 \ldots \alpha_{n-1}$$

и следователно

$$\alpha^{n-1}=1. \tag{9}$$

Като умножим предното равенство с α, получаваме

$$\alpha^n=\alpha,$$

което очевидно се удовлетворява и за елемента нула.

Така установихме теоремата:

Ако Γ е поле на Галоа и n е броят на елементите му, то за всеки елемент α от Γ имаме равенството $\alpha^n=\alpha$.

Нека отбележим, че равенството (9) може да се получи като следствие на теоремата на Лагранж за индекса на една подгрупа. Очевидно елементите $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ образуват група относно умножението, която е Абелева, понеже е комутативна. Ако μ е редът на елемента α от нея то имаме $\alpha^\mu=1$. Но μ трябва да дели реда $n-1$ на въпросната група. Тогава, като повдигнем равенството $\alpha^\mu=1$ в степен $\frac{n-1}{\mu}$, получаваме (9).

В частност, ако полето Γ съвпада с полето Γ', то за всеки негов отличен от нула елемент имаме $\alpha^{p-1}=1$. Като вземем пред вид, че $kp\epsilon=0$, k цяло, получаваме следната теорема:

Ако p е просто число и a е произволно число, което не се дели на p, то числото $a^{p-1}-1$ се дели на p.

Тази теорема от теорията на числата е открита от Ферма и за пръв път доказана от Ойлер.

Както видяхме, отличните от нула елементи $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ на полето на Галоа Γ удовлетворяват уравнението

$$x^{n-1}-1=0. \tag{10}$$

Като използуваме теорията на корените на единицата, заключаваме, че уравнението (10) има поне един примитивен корен ω и всичките му корени се дават с

$$\omega, \omega^2, \ldots, \omega^{n-1}.$$

Следователно отличните от нула елементи на едно поле на Галоа с n елементи образуват циклична група от ред $n-1$. Друго едно свойство на тези полета е следното:

Всяко поле на Галоа с характеристика p съдържа за всеки свой елемент α точно един p-ти корен $\alpha^{\frac{1}{p}}$.

Действително, ако елементите на полето са $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \alpha_n$, то и p-тите им степени

(11) $$\alpha_1^p, \alpha_2^p, \ldots, \alpha_{n-1}^p, \alpha_n^p$$

принадлежат също на полето. От равенството

$$(\alpha_i - \alpha_j)^p = \alpha_i^p - \alpha_j^p$$

следва, че елементите (11) са все различни, т. е. това са всичките елементи на полето. Следователно всеки елемент е p-та степен на някой елемент от полето.

12. Съвършени полета. Едно поле P, в което всяко уравнение с поне един многократен корен е разложимо, се нарича съвършено.

Всяко поле P с характеристика нула е съвършено.

Действително нека уравнението $f(x)=0$ с коефициенти от P има многократен корен. Тогава уравнението $f'(x)=0$, в което $f'(x)$ не е тъждествено равно на нула, ще има общ корен с даденото уравнение. Следователно полиномите $f(x)$ и $f'(x)$ ще имат общ делител $d(x)$ с коефициенти от P и полиномът $f(x)$ ще бъде разложим, понеже се дели на $d(x)$.

Ще изследваме сега кога едно поле P с характеристика $p>0$ е съвършено. Нека $f(x)$ е полином от полето P,

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n.$$

За производната $f'(x)$ на $f(x)$ имаме

$$f'(x)=a_1+2a_2x+3a_3x^2+\ldots+na_nx^{n-1}.$$

Ако полиномът $f'(x)$ не е тъждествено равен на нула, то, както преди установяваме, че всеки такъв полином $f(x)$, който има многократна нула, е разложим. Остава следователно да се разгледат полиноми, на които производната $f'(x)$ е равна тъждествено на нула, т. е. всичките коефициенти на $f'(x)$ са равни на нула. Понеже полето P е с характеристика p, то всичките коефициенти μ a_μ в $f'(x)$, на които индексът се дели на p, са равни на нула. Следователно полиномът $f(x)$ ще бъде тъждествено равен на нула, ако коефициентите a_μ, за които индексът μ не се дели на p, са равни на нула, т. е. полиномът $f(x)$ да има формата

(1) $$f(x)=a_0+a_px^p+a_{2p}x^{2p}+\ldots+a_{mp}x^{mp}.$$

Съвкупността от p-тите степени на елементите на полето P е също така поле. Действително, ако a и b са произволни елементи от P, то, както установихме, имаме $(a+b)^p=a^p+b^p$ и освен това $(ab)^p=a^pb^p$. Въпросното поле от p-тите степени на елементите на P да означим с P^p.

Необходимо и достатъчно условие полето P с характеристика p да бъде съвършено е полето P^p да съвпада с P, т. е. $P^p=P$.

Ще докажем отначало, че условието е достатъчно. Както видяхме, в горните предварителни бележки можем да се ограничим на полиноми

от вида (1). От предположението, че полето P^p съвпада с полето P, следва, че има елементи $c_0, c_1, \ldots, c_m$ от P, за които имаме

$$a_0 = c_0^p,\ a_p = c_1^p,\ \ldots,\ a_{mp} = c_m^p.$$

Но тогава, ако с $\psi(x)$ означим полинома

$$\psi(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_m x^m,$$

то лесно получаваме, че

$$\psi^p(x) = (c_0 + c_1 x + c_2 x^2 + \ldots + c_m x^m)^p = a_0 + a_p x^p + a_{2p} x^{2p} + \ldots + a_{mp} x^{mp} = f(x).$$

Последното равенство показва, че полиномът $f(x)$ е разложим.

Сега ще установим необходимостта на условието. Да предположим, че полетата P^p и P не съвпадат. Ще има тогава поне един елемент α от P, който не е p-та степен на елемент от същото поле. Да разгледаме специалния полином

$$f(x) = x^p - \alpha,$$

на който производната px^{p-1} е равна на нула. Да присъединим към P елемента $\beta = \alpha^{\frac{1}{p}}$. Ще имаме

$$f(x) = x^p - \alpha = (x - \beta)^p.$$

Да предположим, че $f(x)$ е разложим в полето P, т. е.

$$f(x) = \varphi(x)\,\psi(x).$$

Понеже в разширеното поле $P(\beta)$ $f(x)$ е равен на $(x-\beta)^p$, то следва, че

$$\varphi(x) = (x-\beta)^v,\ \psi(x) = (x-\beta)^{p-v}.$$

Понеже полиномът $\varphi(x)$ принадлежи на полето P, то $\varphi(0) = (-\beta)^v$ трябва да бъде елемент от P. Но числата v и p са взаимно прости и по едно елементарно свойство знаем, че съществуват цели числа l и k, отлични от нула, за които имаме

$$vl - pk = 1.$$

Но тогава ще имаме

$$\alpha = \alpha^{vl}\alpha^{-pk} = \beta^{vlp}\alpha^{-pk} = (-1)^{vlp}\gamma^p,\quad \gamma = (-\beta)^{vl}\alpha^{-k},$$

откъдето следва, че α е p-та степен на елемент от P, което противоречи. Следователно полиномът $f(x)$ е неразложим и изказаната теорема е така установена.

В предния параграф видяхме, че равенството $P^p = P$ е изпълнено за всяко крайно поле с характеристика p. Следователно всяко крайно поле е съвършено.

За съвършените полета ще установим още една основна теорема. Казваме, че едно алгебрично разширение на дадено поле е просто, ако

то се получава с присъединяване на един корен на дадено алгебрично уравнение.

Всяко крайно алгебрично разширение на дадено съвършено поле P е просто разширение на P.

Ако полето е с краен брой елементи, теоремата е очевидна, понеже елементите му са $1, \omega, \omega^2, \ldots, \omega^{n-1}$, където ω е примитивен корен на уравнението $x^{n-1}=1$. Можем следователно да предположим, че P съдържа безбройно много елементи. Трябва да установим, че полето $P(\alpha_1, \alpha_2, \ldots, \alpha_m)$, получено с присъединяване към P на корени на неразложимите уравнения

$$f_1(x)=0, \quad f_2(x)=0, \ldots, f_m(x)=0,$$

може да се получи от P с присъединяване само на корен на едно уравнение. Да се ограничим отначало на две уравнения. Нека следователно α и β са корени на неразложимите уравнения

$$f(x)=0, \quad \varphi(x)=0,$$

на които степените са съответно равни на m и n. Да означим с $\alpha_1=\alpha, \alpha_2, \ldots, \alpha_m$ корените на първото уравнение и с $\beta_1=\beta, \beta_2, \ldots, \beta_n$ тези на второто. Понеже полето P е съвършено, въпросните корени са прости. Да разгледаме числата

$$(2) \qquad \frac{\alpha_i-\alpha_1}{\beta_1-\beta_j}, \; i=1, 2, \ldots, m; \quad j=2, 3, \ldots, n,$$

които по причина на $\beta_1-\beta_j \neq 0$ са дефинирани. Понеже P съдържа безбройно много елементи, нека c е елемент от P, който е отличен от дробите (2). Следователно ще имаме $\alpha_i+c\beta_j \neq \alpha_1+c\beta_1$ за $i=1, 2, \ldots, m$; $j=1, 2, \ldots, n$ (освен $i=1, j=1$).
Елементът

$$\gamma=\alpha_1+c\,\beta_1=\alpha+c\,\beta$$

принадлежи на полето $P(\alpha, \beta)$. Но елементът β е корен на $\varphi(x)=0$, т. е.

$$\varphi(\beta)=0.$$

Като вземем пред вид, че $f(\alpha)=0$, получаваме

$$f(\gamma-c\beta)=0.$$

Значи уравненията

$$(3) \qquad f(\gamma-cx)=0, \quad \varphi(x)=0$$

имат общ корен $x=\beta$. Друг общ корен те не могат да имат. Действително, ако β_i, $i>1$, е друг общ корен на уравненията (3), то следва, че елементът $\gamma-c\beta_i$ е равен на някой на корените α_j, което е невъзможно, понеже елементите $\alpha_i+c\,\beta_j$ са все различни. Следователно общият най-голям делител $D(x)$ на полиномите $f(\gamma-cx)$ и $\varphi(x)$ ще бъде от първа степен, т. е.

$$D(x)=A+Bx.$$

Коефициентите A и B са очевидно елементи от полето $P(\gamma)$. Но тогава $\beta=-\frac{A}{B}$ и $\alpha=\gamma-c\beta$ ще са елементи от същото поле $P(\gamma)$, на което ще принадлежат следователно и елементите на полето $P(\alpha, \beta)$, т. е. $P(\alpha, \beta)=P(\gamma)$.

Така теоремата е установена за две уравнения. Предполагаме, че теоремата е установена за $k-1\,(k\geqq 3)$ уравнения, т. е. за полето $P(\alpha_1, \alpha_2, \ldots, \alpha_{k-1})$ имаме

$$P(\alpha_1, \alpha_2, \ldots, \alpha_{k-1})=P(\omega).$$

Но тогава за k уравнения (с прибавяне на корен α_k) имаме

$$P(\alpha_1, \alpha_2, \ldots, \alpha_{k-1}, \alpha_k)=P(\omega, \alpha_k)=P(\delta).$$

По този начин теоремата е доказана в общия случай.

13. Трансцендентно разширение. Нека P е едно поле и $P[x]$ означава пръстена, получен от P с присъединяване на неизвестното x. Както се вижда от разглежданията в § 6, този пръстен е комутативен без делители на нулата, т. е. образува област на цялостност. Да предположим, че пръстенът $P[x]$ принадлежи на едно поле П. Тогава за всеки два полинома $f(x)$ и $\varphi(x)\neq 0$ от него уравнението

$$\varphi(x)z=f(x)$$

ще има решение, което да означим с $\frac{f(x)}{\varphi(x)}$, което да наречем отношение на полиномите $f(x)$ и $\varphi(x)$. Понеже $P[x]$ е област на цялостност, разглежданията в § 4 се прилагат, като там разбираме под буквите $a, b, \ldots$ полиноми на x. Следователно върху отношенията $\frac{f(x)}{\varphi(x)}$ можем да извършваме същите действия, както с дробите, т. е. имаме

$$\frac{f(x)}{\varphi(x)}=\frac{f_1(x)}{\varphi_1(x)}$$

само тогава, когато $f(x)\varphi_1(x)=\varphi(x)f_1(x)$, $\varphi(x)\neq 0$, $\varphi_1(x)\neq 0$,

$$\frac{f(x)}{\varphi(x)}+\frac{f_1(x)}{\varphi_1(x)}=\frac{f(x)\varphi_1(x)+\varphi(x)f_1(x)}{\varphi(x)\varphi_1(x)}, \quad \varphi(x)\neq 0, \quad \varphi_1(x)\neq 0,$$

(1)
$$\frac{f(x)}{\varphi(x)}\cdot\frac{f_1(x)}{\varphi_1(x)}=\frac{f(x)f_1(x)}{\varphi(x)\varphi_1(x)}, \quad \varphi(x)\neq 0, \quad \varphi_1(x)\neq 0,$$

$$\frac{f(x)}{\varphi(x)}:\frac{f_1(x)}{\varphi_1(x)}=\frac{f(x)\varphi_1(x)}{\varphi(x)f_1(x)}, \quad \varphi(x)\neq 0, \quad \varphi_1(x)\neq 0, \quad f_1(x)\neq 0.$$

От свойствата (1) следва, че съвкупността на отношенията на всичките полиноми от $P[x]$ образува поле Σ, което е подполе на П. Полето Σ ще наречем поле на отношенията на пръстена $P[x]$ в П. Може да се случи, че пръстенът $P[x]$ е част от друго поле П′ и тогава ще получим ново

поле на отношенията Σ'. Като използуваме резултатите от поменатия параграф, виждаме, че полетата Σ и Σ' са изоморфни, т. е. съществува до изоморфизъм единствено поле на отношенията. Установихме, че за всяка област на цялостност съществува поле на отношения. Следователно имаме:

Поле на отношения съществува за всеки пръстен от полиноми $P[x]$. Да означим това поле на отношения с $P(x)$. Очевидно полето $P(x)$ ще се състои от всичките рационални функции на x с коефициенти от полето P на влизащите в тях полиноми. Полето $P(x)$ се нарича трансцендентно разширение на полето P. То е получено от P с присъединяване на трансцендентния спрямо P елемент x. Нека $\psi(x)$ е една рационална функция от полето $P(x)$. Да образуваме полето $P(\psi)$, което представлява съвкупността на всичките рационални функции на $\psi(x)$ с коефициенти, принадлежащи на полето P. Очевидно $P(\psi)$ е подполе на $P(x)$. Това поле представлява също така трансцендентно разширение на полето P, понеже $\psi(x)$ не може да удовлетворява алгебрично уравнение с коефициенти от P. Действително, ако $\psi(x)$ е корен на уравнение $g(x)=0$ с коефициенти от P, то очевидно уравнението

$$g[\psi(x)]=0$$

ще се сведе към уравнение $T(x)=0$ с коефициенти от P, което противоречи на условието, че x е трансцендентен елемент спрямо полето P. Лесно се вижда, че полето $P(\psi)$ е изоморфно на полето $P(x)$. Именно нека направим да отговаря на всеки елемент $\omega(x)$ от полето $P(x)$ елемента $\omega[\psi(x)]$ от полето $P(\psi)$. Тогава непосредствено се вижда, че на сумата на два елемента от $P(x)$ отговаря сумата на съответните елементи от $P(\psi)$ и на произведението на два елемента от първото поле отговаря произведението на съответните елементи от второто поле. Освен това на разни елементи от едното поле отговарят разни елементи от другото поле. Забележително е, че полето $P(x)$ не съдържа други подполета освен такива от вида на подполето $P(\psi)$. Именно имаме следната теорема на *Люрот*:

Всяко подполе P' на полето $P(x)$, което съдържа полето P, съдържа такъв елемент y, че $P(y)$ съвпада с полето P'.

Ще изложим доказателството на тази теорема, което е било дадено от Штайниц. Всеки елемент на полето $P(x)$ представлява дробна рационална функция от вида

$$\psi(x)=\frac{f(x)}{\varphi(x)},$$

където $f(x)$ и $\varphi(x)$ са полиноми с коефициенти от полето P. Ние винаги можем да предположим, че полиномите $f(x)$ и $\varphi(x)$ са взаимно прости. Под степен на функцията $\psi(x)$ разбираме най-голямата от степените на полиномите $f(x)$ и $\varphi(x)$. Една помощна теорема при доказателството на теоремата на Люрот е следната теорема на Штайниц:

Ако функциите $G(x)$ и $\psi(x)$ имат съответно степените n и m, то функцията $G[\psi(x)]$ има степен $n.\ m$.

Нека

$$G(x)=\frac{F(x)}{\Phi(x)}, \qquad \psi(x)=\frac{f(x)}{\varphi(x)},$$

като полиномите $F(x)$ и $\Phi(x)$ са взаимно прости и полиномите $f(x)$ и $\varphi(x)$ са взаимно прости. Тогава можем да пишем

$$F(x)=A_0x^n+A_1x^{n-1}+\ldots+A_n, \qquad \Phi(x)=B_0x^n+B_1x^{n-1}+\ldots+B_n,$$

$$f(x)=a_0x^m+a_1x^{m-1}+\ldots+a_m, \qquad \varphi(x)=b_0x^m+b_1x^{m-1}+\ldots+b_m,$$

като поне едно от числата A_0 и B_0 е отлично от нула и поне едно от числата a_0 и b_0 е също отлично от нула. Тогава за $G[\psi(x)]$ имаме

$$G[\psi(x)]=\frac{M(x)}{N(x)}, \tag{2}$$

където

$$M(x)=\varphi^n(x)\,F\left[\frac{f(x)}{\varphi(x)}\right]=A_0f^n+A_1f^{n-1}\varphi+\ldots+A_n\varphi^n,$$

$$N(x)=\varphi^n(x)\,\Phi\left[\frac{f(x)}{\varphi(x)}\right]=B_0f^n+B_1f^{n-1}\varphi+\ldots+B_n\varphi^n. \tag{3}$$

Ще покажем, че полиномите $M(x)$ и $N(x)$ са взаимно прости и че степента на (2) е точно равна на $n.\ m$. Съгласно с условието полиномите $F(x)$ и $\Phi(x)$ са взаимно прости и следователно има полиноми $L(x)$ и $K(x)$ от степен $\leqq n-1$, за които имаме

$$F(x)L(x)+\Phi(x)K(x)=1, \tag{4}$$

откъдето получаваме

$$F\left[\frac{f(x)}{\varphi(x)}\right]L\left[\frac{f(x)}{\varphi(x)}\right]+\Phi\left[\frac{f(x)}{\varphi(x)}\right]K\left[\frac{f(x)}{\varphi(x)}\right]=1.$$

Като умножим предното равенство с $\varphi^{2n-1}(x)$, получаваме

$$M(x)\varphi^{n-1}(x)L\left[\frac{f(x)}{\varphi(x)}\right]+N(x)\varphi^{n-1}(x)K\left[\frac{f(x)}{\varphi(x)}\right]=\varphi^{2n-1}(x), \tag{5}$$

където $\varphi^{n-1}(x)L\left[\frac{f(x)}{\varphi(x)}\right]$ и $\varphi^{n-1}(x)\ K\left[\frac{f(x)}{\varphi(x)}\right]$ са очевидно полиноми на x. От (5) следва, че всеки общ делител на полиномите $M(x)$ и $N(x)$ ще дели полинома $\varphi^{2n-1}(x)$ и следователно тези полиноми ще имат общ делител $d_1(x)$, който дели полинома $\varphi(x)$. Но тогава, понеже поне единият от коефициентите A_0 и B_0 е отличен от нула, би следвало, че $d_1(x)$ дели $f(x)$, т. е. полиномите $f(x)$ и $\varphi(x)$ ще имат общ делител, което противоречи. Следователно полиномите $M(x)$ и $N(x)$ са взаимно прости.

От (3) се вижда, че коефициентите пред най-високата степен x^{nm} в полиномите $M(x)$ и $N(x)$ са съответно равни на

$$C_0 = A_0 a_0^n + A_1 a_0^{n-1} b_0 + \ldots + A_n b_0^n = b_0^n F\left(\frac{a_0}{b_0}\right),$$

$$D_0 = B_0 a_0^n + B_1 a_0^{n-1} b_0 + \ldots + B_n b_0^n = b_0^n \Phi\left(\frac{a_0}{b_0}\right),$$

ако $b_0 \neq 0$, и равни на

$$C_0 = A_0 a_0^n, \quad D_0 = B_0 a_0^n,$$

ако $b_0 = 0$ (като тогава трябва $a_0 \neq 0$). Ако $b_0 = 0$, то очевидно е тогава, че поне един от коефициентите C_0 и D_0 е отличен от нула. Ако $b_0 \neq 0$, то като поставим в (4) $x = \frac{a_0}{b_0}$ и умножим с b_0^n, получаваме

$$b_0^n F\left(\frac{a_0}{b_0}\right) L\left(\frac{a_0}{b_0}\right) + b_0^n \Phi\left(\frac{a_0}{b_0}\right) K\left(\frac{a_0}{b_0}\right) = b_0^n$$

или

$$C_0 L\left(\frac{a_0}{b_0}\right) + D_0 K\left(\frac{a_0}{b_0}\right) = b_0^n.$$

Последното равенство очевидно показва, че двата коефициента C_0 и D_0 не могат едновременно да бъдат равни на нула. Така теоремата на Штайниц е установена.

Нека в полето P' да означим с

$$y = \frac{f(x)}{\varphi(x)}$$

елемент, на който степента ν е най-малка. Понеже полиномите $f(x)$ и $\varphi(x)$ имат коефициенти от полето P, то уравнението

$$f(t) - y\varphi(t) = 0,$$

на което x е корен, има коефициенти от полето $P(y) \subset P'$. Ако това уравнение е разложимо, нека $F(t) = 0$ е неразложимото уравнение в полето P', на което x е корен. Ако μ е степента му относно t, то очевидно $\mu \leqq \nu$. Коефициентите на полинома $F(t)$ пред степените на t са елементи от полето P' и следователно и степените им относно x не могат да бъдат по-малки от ν. Понеже уравнението $F(t) = 0$ е неразложимо, то полиномът $f(t) - y\varphi(t)$ трябва да се дели на $F(t)$. Като приведем в еднакъв знаменател коефициентите на степените на t в полинома $F(t)$ и умножим този полином с въпросния общ знаменател, получаваме полином $F(t, x)$, на който степента относно x не е по-малка от ν. Ще отбележим, че в така получения полином коефициентите пред степените на t не могат да имат общ множител, т. е. този

полином в полето, получено от P с присъединяване на неизвестното x, е примитивен полином. Полиномът

$$\varphi(x)[f(t)-y\varphi(t)]=\varphi(x)f(t)-f(x)\varphi(t)$$

е също примитивен полином. Понеже полиномът $f(t)-y\varphi(t)$ се дели на полинома $F(t)$, то ще имаме

$$\varphi(x)f(t)-f(x)\varphi(t)=F(t,\ x)\,Q(t,\ x), \tag{6}$$

където по лемата на Гаус $Q(t, x)$ е полином освен на t, но и на x. Но лявата част на (6) относно x е полином от степен γ, а в дясната част полиномът $F(t, x)$ е от степен $\geqq\gamma$ спрямо x. Следователно полиномът $Q(t, x)$ не може да съдържа x, т. е. $Q(t, x)$ е полином само на t, който да означим с $Q(t)$. От предните разглеждания следва, че полиномът $f(t)-y\varphi(t)$ се дели на полинома $Q(t)$. Нека $Q(t)$ не е константа. Да разделим тогава полиномите $f(t)$ и $\varphi(t)$ с $Q(t)$. Ще имаме

$$f(t)=Q(t)f_1(t)+R_1(t),\quad \varphi(t)=Q(t)\varphi_1(t)+R_2(t),$$

където остатъците са от степен, по-ниска от тази на полинома $Q(t)$. От предните равенства получаваме

$$f(t)-y\varphi(t)=[f_1(t)-y\varphi_1(t)]\,Q(t)+R_1(t)-yR_2(t),$$

от които следва, че полиномът $R_1(t)-yR_2(t)$ се дели на $Q(t)$, което е невъзможно, понеже първият полином има по-ниска степен от втория. Следователно остатъците $R_1(t)$ и $R_2(t)$ трябва да бъдат равни на нула, от което заключаваме, че полиномите $f(t)$ и $\varphi(t)$ се делят на $Q(t)$. Но това противоречи на факта, че полиномите $f(t)$ и $\varphi(t)$ са взаимно прости. Значи $Q(t)$ е константа и следователно полиномът $f(t)-y\varphi(t)$ е неразложим в P'.

Нека $z=\frac{\alpha(x)}{\beta(x)}$ е един кой да е елемент от полето P'. Уравнението

$$\alpha(t)-z\beta(t)=0$$

спрямо t, коефициентите на което принадлежат на полето P', ще има общ корен $t=x$ с неразложимото уравнение

$$f(t)-y\varphi(t)=0,$$

коефициентите на което също принадлежат на P'. Следователно полиномът $\alpha(t)-z\beta(t)$ трябва да се дели на полинома $f(t)-y\varphi(t)$. Остатъкът на делението на първия полином с втория е полином на t от степен $\gamma-1$, коефициентите на който са очевидно линейни функции на z. Но всички тези коефициенти трябва да бъдат равни на нула и така за z получаваме уравнение от първа степен $Az+B=0$, коефициентите A и B на което принадлежат на полето $P'=P(y)$. От последното уравнение получаваме $z=-\frac{B}{A}$ и следователно елементът z принадлежи на това поле. Така теоремата на Люрот е установена.

14. Алгебрични числа. Всяко реално или комплексно число, което е корен на някое алгебрично уравнение с рационални коефициенти, се нарича *алгебрично число*. Числата, които не са корени на такива уравнения, се наричат *трансцендентни*. Например всяко рационално число е алгебрично, понеже е корен на уравнение с рационални коефициенти. Съществуването на трансцендентни числа не е очевидно и ние ще покажем по-нататък, че действително има такива числа, и то безкрайно много.

Уравнението, което удовлетворява едно алгебрично число, може да се предполага за неразложимо в естествената област на рационалност. В противен случай разлагаме лявата му част на множители и вземаме този неразложим множител, който се анулира за даденото алгебрично число. Не е трудно да се види, че уравнението $f(x)=0$, на което алгебричното число α е корен, е тогава еднозначно определено до постоянен множител. Действително да предположим, че α е корен и на друго неразложимо уравнение $\varphi(x)=0$. Понеже неразложимите уравнения $f(x)=0$ и $\varphi(x)=0$ имат поне един общ корен α, то по известна теорема полиномът $f(x)$ трябва да се дели на $\varphi(x)$, т. е. двата полинома могат да се отличават един от друг само с един постоянен множител.

Очевидно можем да предполагаме, че уравнението $f(x)=0$, на което α е корен, е с цели рационални коефициенти. При това, ако коефициентите му са взаимно прости числа, то това уравнение е напълно определено. Предполагаме естествено, че то е неразложимо. Алгебричните числа, които са корени на едно и също неразложимо уравнение, се наричат спрегнати помежду си. Едно алгебрично число се нарича *цяло алгебрично число*, ако коефициентът пред най-високата степен на уравнението, което то удовлетворява, е равен на 1.

Ще *покажем сега, че множеството от алгебричните числа е поле*.

Затова ще установим именно, че сумата, разликата произведението и частното на две кои да е алгебрични числа са също алгебрични числа. Нека α и β са такива числа и съответните уравнения, на които те са корени, да бъдат

$$f(x)=0,\ \varphi(x)=0. \tag{1}$$

Да означим с $\alpha_1=\alpha, \alpha_2, \ldots, \alpha_n$ спрегнатите числа на α и с $\beta=\beta_1, \beta_2, \ldots, \beta_m$ спрегнатите числа на β. Уравнението, което има за корени числата

$$\alpha_i+\beta_j,\ i=1,\ 2,\ 3,\ldots,n,\ j=1,\ 2,\ 3,\ldots,\ m,$$

ще бъде

$$F(x)=\prod_{i=1,\ j=1}^{n,m}[x-(\alpha_i+\beta_j)]=0. \tag{2}$$

Коефициентите на това уравнение не се променят, като разместваме променливите $\alpha_1, \alpha_2, \ldots, \alpha_n$ помежду им, както и променливите $\beta_1, \beta_2, \ldots, \beta_m$. Следователно те са симетрични функции от корените на уравненията (1). Но като такива те ще бъдат рационални функции от коефициен-

тите на уравненията (1) с рационални коефициенти, т. е. ще са рационални числа. Сумата $\alpha+\beta$ като корен на уравнението (2), на което коефициентите са рационални числа, ще бъде алгебрично число. Понеже разликата $\alpha-\beta$ е корен на уравнението

$$\Phi(x)=\prod_{i=1,\ j=1}^{n,m}[x-(\alpha_i-\beta_j)]=0$$

и произведението $\alpha\beta$ е корен на уравнението

$$G(x)=\prod_{i=1,\ j=1}^{n,\ m}(x-\alpha_i\beta_j)=0,$$

то напълно аналогично се убеждаваме, че тези числа са алгебрични.

За да установим, че и частното $\frac{\alpha}{\beta}$ $(\beta\neq 0)$ е алгебрично число, достатъчно е да докажем, че числото $\frac{1}{\beta}$ е алгебрично. Но това веднага следва от факта, че числото $\frac{1}{\beta}$ е корен на уравнението

$$x^m\varphi\left(\frac{1}{x}\right)=0,$$

където m е степента на полинома $\varphi(x)$.

Ще отбележим, че от изложеното доказателство следва и следната теорема.

Сумата, разликата и произведението на две кои да е цели алгебрични числа са също цели алгебрични числа.

Ще докажем сега една теорема, която в известен смисъл показва, че полето на алгебричните числа е затворено.

Ако числото ω е корен на уравнение

$$f(x)=x^n+\alpha x^{n-1}+\beta x^{n-2}+\ldots+\varepsilon=0, \tag{3}$$

на което коефициентите са алгебрични числа, то числото ω е също алгебрично.

Нека $\alpha_i, \beta_j, \ldots, \varepsilon_\tau$ са спрегнатите числа на $\alpha, \beta, \ldots, \varepsilon$, като в тях броим и тези последни числа (съгласно с дефиницията). Да разгледаме всевъзможните полиноми от вида

$$\varphi_{i,j,\ldots,\tau}(x)=x^n+\alpha_i x^{n-1}+\beta_j x^{n-2}+\ldots+\varepsilon_\tau,$$

между които очевидно е и полиномът $f(x)$, и да образуваме произведението им

$$F(x)=\prod_{i,j,\ldots,\tau}\varphi_{i,j,\ldots,\tau}(x).$$

Коефициентите на това уравнение са симетрични функции на всяка от системите $\alpha_i, \beta_j, \ldots, \varepsilon_r$ и следователно ще се изразяват рационално посредством коефициентите на уравненията, на които тези системи от числа са корени. Това показва, че коефициентите на уравнението $F(x)=0$ са рационални числа и теоремата е установена.

Ще отбележим, че от приведеното доказателство следва и предложението:

Ако коефициентите на уравнението (3) са цели алгебрични числа, то и числото ω е цяло алгебрично число.

Като приложение на предните теореми лесно се убеждаваме, че числата

$$\sqrt{2}+3\sqrt{3}, \quad \sqrt{2-\sqrt{2}}+\frac{1}{1+\sqrt{3}}, \quad \sqrt[3]{1+2\sqrt{5}}, \ldots$$

ще са алгебрични числа и въобще всяко число, получено от рационални числа с прилагане на действията събиране, изваждане, умножение, деление и коренуване с цял показател, ще бъде алгебрично число. Но обратното не е в сила въобще.

За пръв път от Лиувил е установено, че съществуват трансцендентни числа. Ще изложим въпросното доказателство. Нека α е ирационално число, което е корен на едно неразложимо уравнение с цели коефициенти:

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0. \tag{4}$$

Нека $\frac{p}{q}$, $q>0$, е една приближена дроб на α, т. е. дроб, която се намира в интервала $(\alpha-\delta, \alpha+\delta)$, където δ е фиксирано положително число. По теоремата за крайните нараствания ще имаме

$$-f\left(\frac{p}{q}\right)=f(\alpha)-f\left(\frac{p}{q}\right)=\left(\alpha-\frac{p}{q}\right)f'(\zeta),$$

гдето ζ е число между α и $\frac{p}{q}$. Ако означим с M максимума на $|f'(x)|$ в интервала $(\alpha-\delta, \alpha+\delta)$, ще имаме тогава

$$\left|f\left(\frac{p}{q}\right)\right| \leqq M\left|\alpha-\frac{p}{q}\right|. \tag{5}$$

От друга страна, в равенството

$$q^n f\left(\frac{p}{q}\right)=a_0p^n+a_1p^{n-1}q+\ldots+a_nq^n$$

дясната част е цяло число, отлично от нула, понеже уравнението (4) не може да има рационален корен. Следователно $q^n\left|f\left(\frac{p}{q}\right)\right| \geqq 1$ и от (5) получаваме

$$\left|\alpha-\frac{p}{q}\right| \geqq \frac{1}{Mq^n}. \tag{6}$$

Сега ще покажем, че има ирационални числа, които не удовлетворяват неравенството (6), каквото и да е числото n, т. е. съществуват трансцендентни числа.

Предварително ще отбележим, че неравенството (6) е в сила и за рационални часла $\frac{p}{q}$, стига числото α да е отлично от $\frac{p}{q}$, както това се вижда от самото доказателство (или направо). Нека g е произволно цяло положително число, отлично от 1. Редът

$$\alpha = 1 + \frac{1}{g} + \frac{1}{g^{2!}} + \frac{1}{g^{3!}} + \dots \tag{7}$$

е сходящ и сумата му е означена с α. Да предположим, че числото α е алгебрично. Тогава то ще бъде корен на уравнение от вида (4). Нека S_m е m-тата парциална сума на реда (7). Ще имаме

$$\alpha - S_{m+1} = \frac{1}{g^{(m+1)!}} + \frac{1}{g^{(m+2)!}} + \dots < \frac{1}{g^{(m+1)!}} + \frac{1}{g^{2(m+1)!}} + \dots = \tag{8}$$

$$= \frac{1}{g^{(m+1)!}} \cdot \frac{1}{1 - \frac{1}{g^{(m+1)}}} < \frac{2}{g^{(m+1)!}}.$$

Като означим с q числото $g^{m!}$ и вземем под внимание, че S_{m+1} може да се пише във формата на дроб, $S_{m+1} = \frac{p}{q}$, гдето p е цяло число, неравенството (8) става

$$\alpha - \frac{p}{q} < \frac{2}{q^{m+1}}.$$

Лесно се вижда, че така достигаме до противоречие с неравенството (6). Именно каквото и да е числото M, можем винаги да вземем m така голямо, че числото

$$\frac{2}{q^{m+1}}$$

да е по-малко от $\frac{1}{Mq^n}$. Следователно числото α е трансцендентно.

Впрочем така установяваме съществуване на безбройно много трансцендентни числа.

По едно друго доказателство на Кантор, основаващо се на основни теореми от теорията на множествата, следва, че множеството на трансцендентните числа е много по-голямо (по-мощно) от множеството на алгебричните числа или по-точно първото множество е континиум, а второто е изброимо множество.

В 1873 г. Хермит установи, че числото e е трансцендентно, а десет години по-късно Линдеман доказа трансцендентността на числото π. В 1936 г. съв. математик Гелфонд с нови методи доказа трансцендент-

ността на всички числа от вида α^β, където $\alpha \neq 0,1$, α е алгебрично число и β е алгебрично число, което не е рационално.

15. Полиноми от няколко неизвестни. Нека R е комутативен пръстен с единица и без делители на нулата. Нека Q е някое комутативно разширение на R и $x_1, x_2, \ldots, x_n$ са n елемента от Q. Както при едно неизвестно върху елементите на пръстена R и системата елементи $x_1, x_2, \ldots, x_n$ прилагаме действията събиране и умножение. В резултат ще получим очевидно елементи от пръстена Q от формата

$$f(x_1, x_2, \ldots, x_n) =$$

$$(1) \quad = A_1 x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} + B x_1^{\beta_1} x_2^{\beta_2} \ldots x_n^{\beta_n} + \ldots + K x_1^{\delta_1} x_2^{\delta_2} \ldots x_n^{\delta_n},$$

гдето $\alpha, \beta, \ldots . \delta$ са цели неотрицателни числа, а $A, B, \ldots, K$ са елементи от пръстена R. С привеждане на подобните членове можем да считаме, че в горния израз нямаме такива членове. Елементите $x_1, x_2, \ldots, x_n$ се наричат **неизвестни**, ако всеки израз от формата (1) е равен на нула, когато всичките коефициенти $A, B, \ldots$ са равни на нула, и то само в такъв случай. Изразите от вида (1) се наричат **цели рационални функции или полиноми** от неизвестните $x_1, x_2, \ldots x_n$ над пръстена R. Ясно е, че нулата може да се разглежда като полином с коефициенти, равни на нула, и всеки елемент a от пръстена R като полином, съдържащ неизвестните в нулева степен. В бъдеще ще предполагаме, че членовете в (1), на които коефициентите са равни на нула, са изпуснати. Два полинома са само тогава равни, когато членовете на единия полином влизат в състава на другия полином и обратно. В противен случай за разликата, на двата полинома бихме получили полином с не всички коефициенти, равни на нула, и такъв полином не може да бъде равен на нула.

Непосредствено се вижда, че сумата и произведението на полиноми $f(x_1, x_2, \ldots, x_n)$, $g(x_1, x_2, \ldots, x_n), \ldots$ е пак полином на неизвестните $x_1, x_2, \ldots, x_n$. При това съдружителният, разместителният и разпределителният закон остават в сила. С това виждаме, че множеството полиноми (1) е пръстен, който ще означаваме с $R[x_1, x_2, \ldots, x_n]$. Очевидно този пръстен е комутативен. Ще покажем, че той не съдържа делители на нулата. Това твърдение следва от предложението:

Степента на произведението на два полинома е сума от степените на множителите.

Това предложение ни е известно за полиноми с едно неизвестно. Ние ще го докажем по индуктивен път, като предполагаме, че то е вече установено за полиноми с $n-1$ неизвестни. Нека отначало предположим, че са дадени два хомогенни полинома на неизвестните

$$f(x_1, x_2, \ldots, x_n), \quad g(x_1, x_2, \ldots, x_n)$$

от степени p и q. Да ги наредим по степените на едно на неизвестните, например по тези на x_n:

$$f(x_1, x_2, \ldots, x_n) = a_0 x_n^k + a_1 x_n^{k-1} + \ldots + a_k,$$

$$g(x_1, x_2, \ldots, x_n) = b_0 x_n^h + b_1 x_n^{h-1} + \ldots + b_h.$$

Тук $a_0, a_1, \ldots, a_k$, $b_0, b_1, \ldots, b_h$ са полиноми на $n-1$ неизвестни над пръстена R и $a_0(x_1, \ldots, x_{n-1})$ и $b_0(x_1, \ldots, x_{n-1})$ са отлични от нула. В произведението членът с най-висока степен ще бъде

$$a_0(x_1, \ldots, x_n)\, b_0(x_1, \ldots, x_n)\, x_n^{k+h}.$$

Понеже полиномите a_0 и b_0 не са равни на нула, то и полиномът $a_0\, b_0$ не е равен на нула. Освен това в произведението

$$fg = a_0\, b_0\, x_n^{k+h} + (a_0\, b_1 + a_1\, b_0) x_n^{k+h-1} + \ldots + a_k\, b_h$$

членовете след първия съдържат x_n в по-ниска степен от $k+h$ и следователно първият член не може да се унищожи от останалите членове. Но така установихме, че полиномът fg има поне един член от степен $p+q$, който е отличен от нула, и понеже степента на този полином не може да надвишава числото $p+q$, то тя е равна точно на $p+q$.

Нека сега дадените полиноми $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$ са от степени p и q. Ние можем да ги представим в сума от хомогенни полиноми, на които степените постепенно намаляват:

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1, x_2, \ldots, x_n) + f_2(x_1, x_2, \ldots, x_n) + f_i(x_1, x_2, \ldots, x_n),$$

$$g(x_1, x_2, \ldots, x_n) = g_1(x_1, x_2, \ldots, x_n) + g_2(x_1, x_2, \ldots, x_n) + \ldots + g_j(x_1, x_2, \ldots, x_n).$$

В произведението

$$fg = f_1 g_1 + f_1 g_2 + \ldots + f_1 g_j + f_2 g_1 + \ldots + f_i g_j$$

първият полином $f_1 g_1$ ще бъде от степен $p+q$ съгласно с току-що доказаното, а останалите полиноми $f_\alpha g_\beta$ ще бъдат от по-ниска степен. С това предложението е установено.

От предложението следва, че произведението на два полинома, отлични от нула, е полином, също отличен от нула, с което е установено, че в пръстена $R[x_1, x_2, \ldots, x_n]$ няма делители на нулата.

Нека отбележим, че предложението се обобщава непосредствено за произведение на повече от два полинома, т. е. при умножение на полиноми степените им се събират.

Както при едно неизвестно лесно се доказва, че пръстенът от полиноми на n неизвестни над пръстена R се явява единствен до изоморфизъм.

Нека $y_1, y_2, \ldots, y_n$ е друга система от неизвестни, взети от Q или от друго комутативно разширение Q', притежаващо единица. На всеки полином

$$(2) \qquad f(x_1, x_2, \ldots, x_n) = A_1 x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} + A_2 x_1^{\beta_1} x_2^{\beta_2} \ldots x_n^{\beta_n} + \ldots$$

от $R[x_1, x_2, \ldots, x_n]$ правим да съответствува полином

$$(3) \qquad f(y_1, y_2, \ldots, y_n) = A_1 y_1^{\alpha_1} y_2^{\alpha_2} \ldots y_n^{\alpha_n} + A_2 y_1^{\beta_1} y_2^{\beta_2} \ldots y_n^{\beta_n} + \ldots$$

от $R[y_1, y_2, \ldots, y_n]$. Полиномът (3) се получава от полинома (2) с простата смяна на неизвестните $x_1, x_2, \ldots, x_n$ с неизвестните $y_1, y_2, \ldots, y_n$.

Лесно се вижда, напълно аналогично на по-рано, че въведеното съответствие е изоморфизъм на въпросните два пръстена.

Пръстенът $R[x_1, x_2, \ldots, x_n]$ може да се получи, като към пръстена $R[x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n]$, образуван с прибавяне към R на неизвестните $x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n$ се прибави неизвестното x_p. Този начин на образуване се оправдава със следното предложение:

Всяко неизвестно x_p е трансцендентен елемент относно пръстена $R[x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n]$ от останалите неизвестни.

Да допуснем обратното. Тогава x_p ще удовлетворява уравнение

$$(4)\quad a_0(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n)x^m + \ldots + a_m(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n) = 0,$$

на което коефициентите a_i са елементи от пръстена R, като $a_0(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n)$ е отличен от нула елемент. Членовете в произведението $a_0 x_p^m$ не могат да се унищожат от членовете в произведенията $a_i x_p^k$, $i > 0$, понеже последните съдържат x_p в по-ниска от m-та степен. Нека

$$a_0(x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n)x_p^m = B_1 x_1^{\lambda_1} x_2^{\lambda_2} \ldots x_n^{\lambda_n} +$$

$$+ B_2 x_1^{\mu_1} x_2^{\mu_2} \ldots x_n^{\mu_n} + \ldots$$

Тогава ст (4) при $x = x_p$ ще имаме равенството

$$B_1 x^{\lambda_1} x_2^{\lambda_2} \ldots x_n^{\lambda_n} + B_2 x_1^{\mu_1} x_2^{\mu_2} \ldots x_n^{\mu_n} + \ldots +$$

$$+ a_m(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_p) = 0.$$

Но такова равенство, както постулирахме, е само тогава възможно, ако коефициентите на членовете в лявата му част са равни на нула и значи специално имаме $B_1 = 0$, $B_2 = 0, \ldots$ Но тогава ще имаме $a_0(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n) = 0$, което противоречи на допущането.

Значение на един произволен полином $f(x_1, x_2, \ldots, x_n)$ от пръстена $R[x_1, x_2, \ldots, x_n]$ за значения $x_1 = c_1$, $x_2 = c_2, \ldots, x_n = c_n$ от пръстена на неизвестните се разбира, както преди елементът d от R, дефиниран с $d = f(c_1, c_2, \ldots, c_n)$.

Ще отбележим, че по аналогичен начин на случая при едно неизвестно може да се установи, че пръстенът $R[x_1, x_2, \ldots, x_n]$, получен с присъединяване на неизвестните $x_1, x_2, \ldots, x_n$ към дадено поле P, може да се разшири в поле от отношения.

Също така относно въпроса за съществуване на неизвестни имаме аналогична теорема на случая на едно неизвестно.

За всеки комутативен пръстен R с единица и без делител на нулата съществува такъв комутативен пръстен R' с единица, който представлява разширение на R и в който има система независими (относно R) неизвестни $x_1, x_2, \ldots, x_n$ с произволен брой n.

Тук под независими неизвестни относно пръстена R разбираме, че един полином от неизвестните $x_1, x_2, \ldots, x_n$ с коефициенти от R е само тогава равен на нула, ако всичките му коефициенти са равни на

нула. При това предполага се естествено, че подобните членове в него са сведени към един член с подходящ коефициент.

16. Хиперкомплексни числа. В началото разгледахме теорията на комплексните числа, които са по-широка система от числа, съдържаща реалните числа. Сега ще изучим по-общи системи от числа, наречени хиперкомплексни системи или алгебри.

Нека A_n е едно n-мерно векторно пространство, в което координатите на векторите са произволни числа от дадено поле P. Можем да предположим, че полето P е произволно не само числово. Над векторите от пространството A_n извършваме събирането и умножение с елемент от полето P, както по-рано, т. е. за произволен вектор

$$\alpha=(a_1,\ a_2,\ldots,a_n)$$

и произволен елемент λ от P се дефинира произведението

$$\lambda\alpha=\alpha\lambda=(\lambda a_1,\ \lambda a_2,\ldots,\lambda a_n)$$

и за вектора α и вектора

$$\beta=(b_1,\ b_2,\ldots,b_n)$$

дефинираме сумата им $\alpha+\beta$ с

$$\alpha+\beta=(a_1+b_1,\ a_2+b_2,\ldots,a_n+b_n).$$

Под нулев вектор се разбира векторът с координати нула:

$$0^*=(0,\ 0,\ldots,\ 0).$$

За всеки вектор α имаме $\alpha+0^*=\alpha$. Противоположният вектор на α е векторът

$$-\alpha=(-a_1,\ -a_2,\ldots,-a_n).$$

Разликата $\beta-\alpha$ може да се дефинира като сумата от β и $-\alpha$:

$$\beta-\alpha=\beta+(-\alpha)=(b_1-a_1,\quad b_2-a_2,\ldots,b_n-a_n).$$

Умножението с елементи от полето P притежава свойствата:

$$a(b\alpha)=(ab)\alpha,$$

$$a(\alpha+\beta)=a\alpha+a\beta,\quad (a+b)\alpha=a\alpha+b\alpha,$$

$$e\alpha=\alpha,\ (-e)\alpha=-\alpha,$$

$$0.\alpha=0^*,\ a.\ 0^*=0^*.$$

Тук e означава единицата на полето P и 0 — нулата. Нека отбележим, че от $a\alpha=0^*$, гдето a е елемент от P и α е вектор, следва, че или $\alpha=0^*$, или поне $a=0$.

Нека $e_1, e_2,\ldots,e_n$ е един произволен базис на пространството A_n. Тогава знаем, че всеки вектор α от него се представя във формата

$$\alpha=a_1e_1+a_2e_2+\ldots+a_ne_n,$$

гдето $a_1,\ a_2,\ldots,a_n$ са елементи от полето P. В пространството A_n ние дефинирахме равенството на два вектора с равенството на съответните

им координати. Също видяхме, че остава в сила и при представянето на векторите с произволен базис, а именно два вектора

$$\alpha = a_1 e_1 + a_2 e_2 + \ldots + a_n e_n,$$
$$\beta = b_1 e_1 + b_2 e_2 + \ldots + b_n e_n$$

са равни само тогава, когато $a_1 = b_1$, $a_2 = b_2, \ldots, a_n = b_n$. Сборът на двата вектора α и β ще бъде равен на

$$\alpha + \beta = (a_1 + b_1) e_1 + (a_2 + b_2) e_2 + \ldots + (a_n + b_n) e_n$$

и произведението αc при произволен елемент c от полето P е равно на

$$\alpha c = (a_1 c) e_1 + (a_2 c) e_2 + \ldots + (a_n c) e_n.$$

Ще въведем сега умножение на вектори от пространството A_n. Именно умножението на векторите α и β извършваме, както при умножение на суми от числа (при използуване на разпределителния закон). Така получаваме

$$\alpha\beta = (a_1 e_1 + a_2 e_2 + \ldots + a_n e_n)(b_1 e_1 + b_2 e_2 + \ldots + b_n e_n) =$$
$$= a_1 b_1 e_1^2 + a_1 b_2 e_1 e_2 + \ldots + a_1 b_n e_1 e_n + a_2 b_1 e_2 e_1 + \ldots + a_n b_n e_n^2.$$

Тук получените символи $e_i e_j$, $i, j = 1, 2, \ldots, n$ не са дефинирани. Ние полагаме по-нататък, че това са вектори от пространството A_n. С други думи полагаме, че произведенията $e_i e_j$ са линейни комбинации на базисните вектори $e_1, e_2, \ldots, e_n$

$$e_i e_j = d_{ij}^{(1)} e_1 + d_{ij}^{(2)} e_2 + \ldots + d_{ij}^{(n)} e_n. \tag{1}$$

$i, j = 1, 2, \ldots, n$, $d_{ij}^{(k)}$ елементи от полето P. Тогава произведението $e_i e_j$ ще бъде също вектор от пространството A_n. За да има по-голямо приложение дефинираното така умножение, постулираме и валидността на съдружителния закон. Достатъчно е затова да бъде той в сила за базичните вектори, т. е. да имаме

$$e_i(e_j e_s) = (e_i e_j) e_s, \; i, j, s = 1, 2, 3, \ldots, n. \tag{2}$$

С въведеното умножение виждаме, че множеството от n-мерните вектори над полето P образува пръстен. Този пръстен се нарича **асоциативна алгебра или хиперкомплексна система** над полето P от краен ранг. Броят n на базисните елементи $e_1, e_2, \ldots, e_n$ се нарича **ранг** на алгебрата.

Системата равенства (1) се нарича **таблица за умножение** в алгебрата при базис $e_1, e_2, \ldots, e_n$. Ясно е, че тази таблица зависи от избрания базис и при друг базис тя може да се промени.

При умножението ние не изискваxме, щото разместителният закон да бъде в сила, т. е. алгебрата A_n може да бъде и некомутативна.

Ще разгледаме някои примери. При алгебра от втори ранг системата равенства (1) става

$$e_1^2 = d_{11}^{(1)} e_1 + d_{11}^{(2)} e_2, \; e_1 e_2 = d_{12}^{(1)} e_1 + d_{12}^{(2)} e_2,$$
$$e_2 e_1 = d_{21}^{(1)} e_1 + d_{21}^{(2)} e_2, \; e_2^2 = d_{22}^{(1)} e_1 + d_{22}^{(2)} e_2.$$

Тук коефициентите $d_{ij}^{(s)}$ трябва така да бъдат подбрани, че да са изпълнени условията (2). Като пръв пример на такава алгебра да разгледаме случая, когато равенствата (1) се свеждат на следните:

$$e_1^2 = e_1, \quad e_1 e_2 = e_2, \tag{3}$$

$$e_2 e_1 = e_2, \quad e_2^2 = -e_1,$$

и полето P е полето D на реалните числа. Не е трудно да се провери, че при така избраната таблица за умножение условията (2) са изпълнени. Така например имаме

$$e_1(e_1 e_1) = e_1 e_1^2 = e_1^2 = e_1,$$

$$e_1(e_2 e_2) = e_1 e_2^2 = -e_1 e_1 = -e_{12} = -e_1, \ldots$$

Да разгледаме действията във въведената така алгебра. За кои да е елементи от нея

$$\alpha = ae_1 + be_2, \quad \beta = ce_1 + de_2$$

имаме

$$\alpha + \beta = (a+c)e_1 + (b+d)e_2,$$

$$\alpha c = (ac)e_1 + (bc)e_2.$$

Като използуваме таблицата (3) за умножение, получаваме

$$\alpha\beta = (ac - bd)e_1 + (ad + bc)e_2.$$

От предните равенства виждаме, че правилата за събиране и умножение са еднакви с тези за комплексните числа, в които вместо елементите e_1 и e_2 имаме единиците 1 и i. По-точно нека на всеки елемент от алгебрата A_2 направим да съответствува комплексното число $z = a + bi$ от полето K на комплексните числа. Това съответствие е взаимно еднозначно $z = a + bi \leftrightarrow ae_1 + be_2 = \alpha$. Ако $\beta = ce_1 + de_2$ е също елемент от A_2, то съответното му комплексно число ще бъде $u = c + di$. Понеже

$$z + u = a + c + (b+d)i \leftrightarrow (a+c)e_1 + (b+d)e_2 = \alpha + \beta,$$

$$zu = ac - bd + (ad + bc)i \leftrightarrow (ac - bd)e_1 + (ad + bc)e_2 = \alpha\beta,$$

то виждаме, че на елемента $\alpha + \beta$ съответствува комплексното число $z + u$ и на елемента $\alpha\beta$ съответствува числото zu. С това установихме, че алгебрата A_2 и полето на комплексните числа са изоморфни. На това основание между тях няма разлика и можем също алгебрата A_2 да наречем поле на комплексните числа, като в случая e_1 играе ролята на 1 и e_2 тази на i.

Да разгледаме друга алгебра от ранг 2 при таблица за умножение:

$$e_1^2 = e_1, e_1 e_2 = e_2 e_1 = e_2 e_2^2 = 0.$$

Тук 0 е нулата от полето P. Получаваме така комутативна алгебра, в която има обаче делители на нулата. Например елементът e_2 не е нула, но $e_2 . e_2 = 0$.

Една забележителна система от хиперкомплексни числа с 4 единици са тъй наречените к в а т е р н и о н и. Именно кватернионите са системи с 4 базисни вектора, които да означим с e, i, j, k. Таблицата за умножение се дава с

$$e^2=e,\ ei=ie=i,\ ej=je=j,\ ek=ke=k,$$

$$i^2=j^2=k^2=-1,\ ij=-ji=k,\ jk=-kj=i,\ ki=-ik=j.$$

Оттук се вижда, че произведението на кой да е кватернион с e е равно на самия кватернион, т. е. e е единицата на пръстена от кватерниони. Ние просто я означаваме с 1. Следователно кватернионите ще имат формата

$$\alpha=a+bi+cj+dk,$$

гдето a, b, c, d са произволни реални числа и символите i, j, k са подчинени на следните правила за умножение:

$$i^2=j^2=k^2=-1,$$

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

Оттук се вижда, че умножението не е комутативно. Проверява се лесно, че то е асоциативно, т. е. за кои да е кватерниони α, β, γ имаме

$$(\alpha\beta)\gamma=\alpha(\beta\gamma).$$

Кватернионът

$$\alpha=a-bi-cj-dk$$

се нарича конюгован на α. За произведението $\alpha\bar{\alpha}$ получаваме

$$\alpha\bar{\alpha}=a^2+b^2+c^2+d^2.$$

Това произведение се нарича норма на α и се бележи с $N(\alpha)$. Очевидно нормата $N(\alpha)$ е само тогава равна на нула, ако $\alpha=0$. Числото $\sqrt{N(\alpha)}$ се нарича модул на кватерниона α. Ако

$$\beta=a_1+b_1 i+c_1 j+d_1 k$$

е също кватернион, то с лесно пресмятане получаваме

$$\overline{\alpha\beta}=\bar{\beta}\bar{\alpha}.$$

За квадрата на модула на $\alpha\beta$ ще имаме тогава

$$\alpha\beta\,.\,\overline{\alpha\beta}=\alpha\beta\,\bar{\beta}\,\bar{\alpha}=\alpha\bar{\alpha}\,.\,\beta\bar{\beta}. \tag{4}$$

Следователно модулът на произведение на кватерниони е равен на произведението от модулите им. Оттук следва непосредствено, че едно произведение от кватерниони е само тогава равно на нула, ако поне единият от множителите е равен на нула, т. е. системата кватерниони не съдържа делители на нулата. В разкрита форма равенството (4) е еквивалентно с тъждеството на *Лагранж*. Да означим с $(\alpha\beta)$ израза

$$(\alpha\beta)=aa_1+bb_1+cc_1+dd_1.$$

Тогава имаме

$$\alpha\bar{\beta}+\beta\bar{\alpha}=2(\alpha\beta)$$

и специално

$$\bar{\alpha}\alpha=\alpha\bar{\alpha}=(\alpha\alpha).$$

Наричаме $a=S\alpha$ скаларна част и $bi+cj+dk=V\alpha$ векторна част на кватерниона α. Ако в кватерниона α скаларната му част $S\alpha$ е равна на нула, то той се нарича „вектор“. За два вектора α и β ($a=a_1=0$) числото

$$(\alpha\beta)=bb_1+cc_1+dd_1$$

се нарича скаларно произведение. Лесно се вижда, че $(\alpha\beta)=S\bar{\alpha}\beta$ и векторът

$$\frac{1}{2}(\alpha\beta-\beta\alpha)=i(cd_1-c_1d)+j(db_1-d_1b)+k(bc_1-b_1c)=V\alpha\beta$$

се нарича векторно произведение на α и β. Ако α е един вектор, то $\frac{\alpha}{\sqrt{\alpha\bar{\alpha}}}$ се нарича орта на α.

Да разгледаме делението. Нека α и β са два вектора, като $\alpha\neq 0$. Да решим уравненията

$$x\alpha=\beta, \quad \alpha y=\beta.$$

Като умножим първото уравнение надясно с $\frac{\bar{\alpha}}{N(x)}$ и второто наляво с $\frac{\bar{\alpha}}{N(\alpha)}$, получаваме

$$x\frac{\alpha\bar{\alpha}}{N(\alpha)}=\beta\frac{\bar{\alpha}}{N(\alpha)}, \quad \frac{\bar{\alpha}}{N(\alpha)}\alpha y=\frac{\bar{\alpha}}{N(\alpha)}\beta.$$

Понеже $\alpha\bar{\alpha}=N(\alpha)$, то горните равенства стават

$$x=\beta\frac{\bar{\alpha}}{N(\alpha)}, \quad y=\frac{\bar{\alpha}}{N(\alpha)}\beta.$$

Кватернионът x можем да наречем ляво частно на кватернионите α и β, а y — дясно частно.

Първоначалните изследвания върху кватернионите датират още от Ойлер в 1747 г. и Гаус в 1819 г. Теорията им е развита от Хамилтон в 1840 г. Приложението на кватернионите в другите области на математиката е голямо, но то се покрива в същността си с известното по-късно развито и векторно смятане.

Ще разгледаме още един пример на хиперкомплексни системи. Нека a, b, c, d са реални числа и в системата с 4 единици

$$ae_1+be_2+ce_3+de_4$$

да положим $e_1=1$, $e_2=j$, $e_3=j^2$, $e_4=j^3$, $e_4^4=2e_2^2-1$.

По този начин таблицата за умножение на единиците е напълно определена. Следователно векторите от така дефинираното 4-мерно пространство ще имат формата

$$\alpha = a + bj + cj^2 + dj^3,$$

като за j имаме $j^4 = 2j^2 - 1$. Лесно се проверява, че асоциативното свойство остава в сила и че умножението е комутативно. Под модул на α разбираме числото

$$\sqrt{(c-a)^2 + (d-b)^2}$$

и следователно има вектори, които не са нулеви, но модулът им е равен на нула. Не е трудно да се види, че делението е винаги възможно, стига модулът на делителя да е отличен от нула.

С непосредствена проверка намираме, че квадратът на векторите

$$\pm\frac{1}{2}(3j + j^3)$$

е равен на -1. Полагаме тогава

$$i = \frac{1}{2}(3j + j^3), \tag{1}$$

като i е имагинерната единица. Но векторът α може да се пише в следната форма:

$$\alpha = \left(\alpha' + \frac{3j + j^3}{2}\alpha''\right) + \left(\beta' + \frac{3j + j^3}{2}\beta''\right)(1 + j^2),$$

гдето

$$\alpha' = a - c, \qquad \alpha'' = b - d,$$

$$\beta' = c, \qquad \beta'' = \frac{3d - b}{2}.$$

Като вземем пред вид равенството (1), изразите

$$\alpha' + \frac{3j + i^3}{2}\alpha'', \qquad \beta' + \frac{3j + j^2}{2}\beta''$$

представляват две комплексни числа z и u и α приема вида

$$\alpha = z + u\omega,$$

гдето $\omega = 1 + j^2$. За ω^2 имаме

$$\omega^2 = 1 + 2j^2 + j^4 = 0.$$

Така получаваме, че общата форма на векторите от разгледаната хиперкомплексна система е следната:

$$z + u\omega,$$

гдето z и u са произволни комплексни числа и $\omega^2=0$. За сумата и произведението на два кои да е вектора $\alpha=z+u\omega$, $\beta=z_1+u_1\omega$

получаваме

$$\alpha+\beta=z+z_1+(u+u_1)\,\omega,$$

$$\alpha\beta=zz_1+(zu_1+z_1u)\,\omega.$$

Понеже $\omega.\omega=0$, то в системата има делители на нулата. Разгледаната система е въведена от Собреро за решение на въпроси от теорията на еластичността.

Накрая ще установим една теорема, която оправдава сменяването на една от единиците с числото 1, което извършихме в разгледаните примери.

Ако алгебрата A_n има единица e, то полето P ще се съдържа в A_n и умножението на елементи от алгебрата с елементи от полето P се обръща в умножение на елементи от самата алгебра.

Нека M е множеството от елементите ae на алгебрата A, гдето a е произволен елемент от полето P. На всеки елемент a от полето P можем да съпоставим елемента ae от M. Това съответствие е взаимно еднозначно. Действително, ако на равни елементи a и b от полето P съответстват равни елементи ae и be от M, то бихме имали $ae-be=(a-b)\,e=0$*. Понеже $e\neq 0$, заключаваме, че $a-b=0$, което противоречи на условието $a\neq b$.

Следователно по този начин получаваме взаимно еднозначно съответствие между елементите a на полето P и елементите ae на множеството M, т. е. $a\longleftrightarrow ae$. Ако b е произволен елемент от P, то от $a\longleftrightarrow ae$ и $b\longleftrightarrow be$ получаваме

$$a+b\longleftrightarrow(a+b)\,e=ae+be,\quad ab\longleftrightarrow abe=abe^2=$$
$$=a\,(be^2)=a\,(ebe)=(ae)\,(be).$$

С това е доказано, че полето P и множеството M са изоморфни. Благодарение на изоморфизма ние можем да не отличаваме елемента ae от елемента a на полето P и по този начин полето P ще се съдържа в алгебрата A_n. На основание на асоциативния закон имаме $(ae)\alpha=a\,(e\alpha)=a\alpha$, което равенство показва, че произведението на произволен елемент a от полето P с елемент α от алгебрата A_n може да се разглежда като произведение на елементите ae и α от алгебрата A_n.

Видяхме, че полето на реалните числа може да се разшири в полето на комплексните числа, които притежават при действията всичките свойства на реалните числа. При кватернионите умножението не е комутативно въобще. Явява се въпросът, дали не съществуват други алгебри с подобни свойства. Нека отбележим, че кватернионите образуват некомутативно тяло. Множеството реални числа и множеството от комплексните числа са комутативни тела, т. е. полета. Ако някоя алгебра A_n от ранг n над полето P е тяло, то тя се нарича алгебра с деление над полето P. Разбира се, предполага се, че алгебрата е асоциативна. Отговорът на поставения по-горе въпрос е отрицателен. По-точно имаме следната теорема на Фробениус:

Над полето на реалните числа съществуват само три алгебри с деление от краен ранг, а именно самото поле на реалните числа, полето на комплексните числа и алгебрата на кватернионите.

17. Неразложимост на полиномите. Ще разгледаме по-подробно въпроса за неразложимостта на полиномите в естествената област на рационалност. Един полином $f(x)$ с рационални коефициенти е разложим в тази област R, ако представлява произведение на полиноми, на които степените са най-малко равни на 1, с коефициенти — рационални числа. Ако приведем в еднакъв знаменател коефициентите на полинома $f(x)$, то очевидно въпросът за разложимостта на полинома $f(x)$ се свежда към този за разложимост на полином с цели рационални коефициенти. Но имаме следната лема на Гаус (доказана също в § 7):

Ако един полином с цели рационални коефициенти се разлага на произведение на два полинома с дробни рационални коефициенти, то той се разлага на произведение от полиноми с цели рационални коефициенти.

Нека полиномът $f(x)$ с цели рационални коефициенти се разлага на произведение от полиномите $\varphi(x)$ и $\psi(x)$,

$$f(x)=\varphi(x)\,\psi(x),$$

на които коефициентите са дробни числа. Привеждаме коефициентите на полинома $\varphi(x)$ и тези на полинома $\psi(x)$ към еднакъв знаменател. Ще имаме

$$(1)\qquad f(x)=\frac{b_0+b_1x+b_2x^2+\dots}{d_1}\cdot\frac{c_0+c_1x+c_2x^2+\dots}{d_2},$$

гдето числата $b_0, b_1, b_2, \dots$ са цели рационални и нямат общ делител с цялото число d_1 и числата $c_0, c_1, c_2, \dots$ са също цели рационални и нямат общ делител с числото d_2 (освен единицата). Нека p е просто число, което е делител на d_1. Понеже p не може да дели всичките коефициенти $b_0, b_1, b_2, \dots$, то нека b_r е първият коефициент поред, който не се дели на p. Ще покажем, че p трябва да дели всичките коефициенти $c_0, c_1, c_2, \dots$ Да предположим, че това не е така и нека c_s е първият поред коефициент, който не се дели на p. Но съгласно с равенството (1) коефициентът на x^{r+s} в $f(x)$ ще е равен на

$$\frac{1}{d_1d_2}(b_rc_s+b_{r-1}c_{s+1}+b_{r-2}c_{s+2}+\dots+b_{r+1}c_{s-1}+b_{r+2}c_{s-2}+\dots)$$

и по условие това число трябва да бъде цяло. Понеже p е делител на d_1, то това число трябва да дели и израза в скобите. Но по предположение числата

$$b_{r-1},\ b_{r-2},\dots,\ b_0,\ c_{s-1},\ c_{s-2},\dots,\ c_0$$

се делят на p и следователно сумата в скобите от втория член нататък са числа, делими на p, а първият член не се дели на p., т. е. въпросната сума представлява число, което не се дели на p. Следователно p трябва да дели всичките числа $c_0, c_1, c_2, \dots$ Като съкратим

тогава на p и продължаваме така за делителите на d_1 и d_2, ще получим полином с цели рационални коефициенти и лемата е така установена.

На основание на тази лема при въпроса за разложимост в естествената област на рационалност можем да се ограничим само на полиноми с цели рационални коефициенти.

Въпросът за разложимостта на полиномите е свързан с преодоляване на редица трудности. Ние ще разгледаме някои прости и удобни критерии, които ни позволяват да установим неразложимостта на някои категории от полиноми.

Теорема на Айзенщайн и Шонеман. Ако един полином $f(x)$ има цели рационални коефициенти, първият от които не се дели на едно просто число p, останалите коефициенти се делят на това число и последният не се дели на p^2, то този полином е неразложим.

По условие полиномът $f(x)$ е от формата

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n,$$

дето $a_1, a_2, \ldots, a_n$ са цели рационални числа, които се делят на p, a_0 не се дели на p и a_n не се дели на p^2. Да допуснем, че $f(x)$ е разложим на произведение от два полинома, които съгласно с лемата на Гаус могат да бъдат предположени с цели рационални коефициенти, т. е.

$$f(x)=(b_0x^r+b_1x^{r-1}+\ldots+b_r)(c_0x^s+c_1x^{s-1}+\ldots+c_s).$$

Като приравним свободните членове, получаваме

$$a_n=b_rc_s.$$

Понеже a_n се дели на p, то трябва поне едно от числата b_r и c_s да се дели на p. От друга страна, и двете числа b_r и c_s не могат да се делят на p, защото в противен случай тяхното произведение би се делило на p^2, което противоречи на условието на теоремата. Нека тогава c_s се дели на p, а b_r не се дели на p. Но тогава, като умножим вдясно полинома $(b_0x^r+b_1x^{r-1}+\ldots+b_r)$ с c_s и пренесем получените членове вдясно, ще получим равенството

$$a_0x^n+p\varphi_1(x)=(b_0x^r+b_1x^{r-1}+\ldots+b_r)\ (c_0x^s+c_1x^{s-1}+\ldots+c_{s-1}x),$$

гдето полиномът $\varphi_1(x)$ е от степен $<n$ и има цели рационални коефициенти. Съгласно с това равенство коефициентът b_rc_{s-1} на x в дясната му част трябва да се дели на p, т. е. c_{s-1} трябва да се дели на p. Като умножим с $c_{s-1}x$ и пренесем получените членове от дясната част в лявата, получаваме

$$a_0x^n+p\varphi_2(x)=(b_0x^r+b_1x^{r-1}+\ldots+b_r)\ (c_0x^s+c_1x^{s-1}+\ldots+c_{s-2}x^2),$$

гдето $\varphi_2(x)$ е полином с цели рационални коефициенти от степен $<n$. Но тогава като преди виждаме, че c_{s-2} трябва да се дели на p. Продължавайки така, достигаме до равенството

$$a_0x^n+p\varphi_s(x)=(b_0x^r+b_1x^{r-1}+\ldots+b_r)c_0x^s,$$

като полиномът $\varphi_s(x)$ е от степен $< n$ и коефициентите му са цели рационални числа. Но оттук следва, че коефициентът $b_r c_0$ на x^s трябва да се дели на p, което е невъзможно, понеже $a_0 = b_0 c_0$ не се дели на p.

Между разните други критерии ще разгледаме и един друг по принцип критерий за неразложимост, даден от Перон.

Нека полиномът

$$(a) \qquad f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n, \; a_n \neq 0,$$

с цели рационални коефициенти има в кръга $|x| < 1$ от комплексната равнина точно $n-1$ нули или $n-2$ нули, като останалите 2 нули са комплексни. Тогава този полином е неразложим.

Да предположим, че $f(x)$ е разложим, т. е.

$$f(x) = \varphi(x)\,\psi(x),$$

$$f(x) = x^p + b_1 x^{p-1} + \ldots + b_p, \quad \psi(x) = x^q + c_1 x^{q-1} + \ldots + c_q,$$

гдето $b_1, b_2, \ldots, b_p, c_1, c_2, \ldots, c_q$ са цели рационални числа и $b_p \neq 0$, $c_q \neq 0$. Никой от двата полинома $\varphi(x)$ и $\psi(x)$ не може да има само нули в кръга $|x| < 1$, понеже в противен случай би трябвало свободният му член да бъде по абсолютна стойност по-малък от 1. Но така идваме до противоречие. Именно в първия случай следва, че предположената нула извън кръга $|x| < 1$ не анулира полинома $f(x)$ и във втория случай имаме същото заключение за предположените две нули (които не са реални числа).

От самото доказателство следва и следното предложение:

Ако полиномът (a) с цели рационални коефициенти има $n-2$ нули в кръга $|x| < 1$ и двете нули вън от него, то той е неразложим при предположение, че няма нула, която е цяло рационално число.

Сега ще изложим един метод на Кронекер, по който винаги можем с краен брой действия да познаем дали един полином е разложим или е неразложим. Нека $f(x)$ е полином с цели рационални коефициенти. Допускаме, че е разложим, т. е. имаме

$$f(x) = \varphi(x)\;\psi(x),$$

гдето полиномите $\varphi(x)$ и $\psi(x)$ съгласно с лемата на Гаус могат да бъдат предположени с цели рационални коефициенти. Ако n е степента на $f(x)$, то очевидно степента на един от полиномите $\varphi(x)$ и $\psi(x)$ не ще надминава числото $\frac{n}{2}$. Да означим с $\varphi(x)$ този от полиномите $\varphi(x)$ и $\psi(x)$, който има степен $\leqq \frac{n}{2}$ (винаги има поне един такъв полином).

Ако α е произволно цяло число, за което $f(\alpha) \neq 0$, от равенството

$$f(\alpha) = \varphi(\alpha)\;\psi(\alpha)$$

виждаме, че цялото число $\varphi(\alpha)$ трябва да дели цялото число $f(\alpha)$. Нека тогава m е цяло число, което не е по-малко от $\frac{n}{2}+1$, и да взе-

мем най-малкото такова число. Вземаме си произволни различни m цели числа $\alpha_1, \alpha_2, \ldots, \alpha_m$. Ако за някое от тия α_i полиномът $f(x)$ се анулира, то α_i ще е корен на уравнението $f(x)=0$ и следователно това уравнение ще бъде разложимо. С отстраняване на този корен свеждаме въпроса за неразложимост към такъв за полином от по-ниска степен. Можем следователно да приемем, че числата $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_m)$ са отлични от нула. Съгласно с формулата на Лагранж за полинома $\varphi(x)$ ще имаме

$$(2)\quad \varphi(x)=\varphi(\alpha_1)\frac{(x-\alpha_2)\ldots(x-\alpha_m)}{(\alpha_1-\alpha_2)\ldots(\alpha_1-\alpha_m)}+\varphi(\alpha_2)\frac{(x-\alpha_1)\ldots(x-\alpha_m)}{(\alpha_2-\alpha_1)\ldots(\alpha_2-\alpha_m)}+\ldots+$$
$$+\varphi(\alpha_m)\frac{(x-\alpha_1)\ldots(x-\alpha_{m-1})}{(\alpha_m-\alpha_1)\ldots(\alpha_m-\alpha_{m-1})}.$$

Тук числата $\varphi(\alpha_1), \varphi(\alpha_2), \ldots, \varphi(\alpha_m)$ трябва да бъдат делители съответно на числата $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_m)$. Тогава образуваме всички възможни полиноми (2), в които за $\varphi(\alpha_k)$, $k=1, 2, \ldots, m$, вземаме за стойности възможни делители на числата $f(\alpha_k)$, $k=1, 2, \ldots, m$. Ако някой така получен полином (2) има и дробни рационални коефициенти, то естествено трябва да го отстраним. За всеки такъв полином с цели рационални коефициенти трябва да изпитаме дали той дели полинома $f(x)$. В случай, че го дели, то полиномът $f(x)$ ще е разложим и по-нататък нашето внимание може да бъде пренесено върху новите полиноми, които са по-прости за изследване. Ако никой полином $\varphi(x)$ не дели $f(x)$, то този полином е неразложим.

Ще разгледаме въпроса за неразложимостта на уравнението на примитивните корени на биномните уравнения. Като пример за приложение на теоремата на Айзенщайн — Шонеман ще докажем неразложимостта на уравнението

$$(3)\qquad \frac{x^p-1}{x-1}=x^{p-1}+x^{p-2}+\ldots+x+1=0$$

на примитивните корени на

$$x^p-1=0,$$

гдето p е просто число. Ако в (3) поставим $x=z+1$, получаваме

$$(4)\qquad \frac{(z+1)^p-1}{z}=z^{p-1}+\binom{p}{1}z^{p-2}+\ldots+p=0.$$

В това уравнение всички коефициенти

$$\binom{p}{1}, \binom{p}{2}, \ldots, \binom{p}{p-1}=p$$

се делят на p и последният се дели само на p в първа степен. Действително, понеже числото

$$\binom{p}{m}=\frac{p(p-1)\ldots(p-m+1)}{m!}, \quad m<p,$$

е цяло и p е просто, то трябва

$$\frac{(p-1)(p-2)\ldots(p-m+1)}{m!}$$

да е цяло число, т. е. $\binom{p}{m}$ се дели на n при $m<p$. Следователно уравнението (4) е неразложимо и от това непосредствено следва, че и уравнението (3) е неразложимо.

По други начини се установява общо, че уравнението

$$\varphi_n(x)=0$$

на примитивните корени на уравнението

$$x^n-1=0$$

е неразложимо.

За случая $n=p^\lambda$, където p е просто число, ще изложим един метод на Кронекер. Да допуснем, че уравнението

$$\varphi_n(x)=\frac{x^{p^\lambda}-1}{x^{p^{\lambda-1}}-1}=0 \tag{5}$$

е разложимо и нека това разлагане да е следното:

$$\varphi_n(x)=\psi_1(x)\,\psi_2(x)\ldots\psi_k(x),$$

като полиномите с цели рационални коефициенти $\psi_s(x)$ са предположени неразложими. Но лесно се вижда, че

$$\varphi_n(1)=p$$

и от това

$$\psi_1(1)\,\psi_2(1)\ldots\psi_k(1)=p,$$

което равенство показва, че едно от числата $\psi_1(1), \psi_2(1), \ldots, \psi_k(1)$ трябва да е равно на $\pm p$, а останалите са ± 1. Нека $\psi_1(1)=\pm p$ и ω да е корен на уравнението

$$\psi_1(x)=0. \tag{6}$$

Понеже корените на (5) са степени на ω, то уравнението

$$\psi_2(x)=0$$

ще има корен ω^g, където g е цяло положително число. Но тогава уравнението

$$\psi_2(x^g)=0$$

ще има общ корен ω с неразложимото уравнение (6) и следователно полиномът $\psi_2(x^g)$ трябва да се дели на полинома $\psi_1(x)$, т. е.

$$\psi_2(x^g)=\psi_1(x)\,h(x),$$

гдето полиномите $\psi_1(x)$ и $h(x)$ са с цели рационални коефициенти. Но при $x=1$ получаваме оттук, че 1 трябва да се дели на p, което е невъзможно. Това доказателство на Кронекер може да се разшири за произволна степен n, но изисква използуването на по-сложни теореми от теорията на числата.

ЧАСТ VIII

АБЕЛЕВИ И БИНОМНИ УРАВНЕНИЯ

Глава I

Абелеви уравнения

1. Дефиниция и групиране на корените. Както ще видим, уравненията от пета степен нагоре в общия си вид не са решими алгебрически. Съществуват обаче специални класи уравнения, които са решими. Една такава важна класа са уравненията на Абел. Едно уравнение

(1) $$f(x)=0$$

се нарича абелево, ако в една дадена област на рационалност, в която принадлежат коефициентите му, е неразложимо и между два негови корена x_1 и x' има релацията

$$x'=\theta(x_1),$$

гдето $\theta(x)$ е рационална функция на x, на която коефициентите принадлежат на същата област на рационалност.

От условието $f(x')=0$, т. е. $f[\theta(x_1)]=0$, следва, че уравненията

(2) $$f(x)=0,\quad f[\theta(x)]=0$$

има общ корен x_1. Понеже първото е неразложимо, то второто ще допуска за корени всичките корени на (1). Следователно ще имаме

$$f[\theta(x')]=0.$$

За простота да означим с

$$\theta[\theta(x)]=\theta^2(x)=\theta^2 x,$$
$$\theta[\theta^2(x)]=\theta^3(x)=\theta^3 x,$$
$$\cdots\cdots\cdots$$

тогава последното равенство може да се пише

(3) $$f[\theta^2(x_1)]=0.$$

Но (3) показва, че уравнението

$$f[\theta^2(x)]=0$$

има общ корен x_1 с (1). Следователно, понеже (1) е неразложимо, ще имаме

$$f[\theta^2(x')]=0$$

или

(4) $$f[\theta^3(x_1)]=0.$$

От (2), (3), (4) виждаме, че $\theta x_1, \theta^2 x_1, \theta^3 x_1$ са пак корени на (1). Продължавайки така, се убеждаваме, че редицата числа

(5) $$x_1, \theta x_1, \theta^2 x_1, \theta^3 x_1, \ldots$$

са все корени на (1). Понеже (1) има краен брой корени, трябва числата (5) да се повтарят. Следователно ще имаме за две от тях поне

$$\theta^{\mu+\nu} x_1 = \theta^{\nu} x_1.$$

Но понеже $\theta^{\mu+\nu} x_1 = \theta^{\mu}(\theta^{q\nu} x_1) = \theta^{\nu} x_1$, то уравнението

(6) $$\theta^{\mu}(x_1) - x = 0$$

има общ корен $\theta^{\nu} x_1$ с (1). Следователно за всеки корен на (1) и специално за x_1 ще имаме

$$\theta^{\mu}(x_1) = x_1.$$

Значи в редицата (5) ще има след първото число непременно едно, което е равно на x_1. Нека първото такова поред да бъде $\theta^r x_1 = x_1$ т. е. нека r е най-малкото число с това свойство. Тогава в редицата (5) различни са само

(7) $$x_1, \theta x_1, \theta^2 x_1, \ldots, \theta^{r-1} x_1; \ \theta^r x_1 = x_1,$$

като всички други са повторения на тези. Ако допуснем, че две числа от (7) са равни помежду си, то, както по-горе бихме дошли до заключение, че има число $s < r$ такова, че

$$\theta^s x_1 = x_1,$$

което е невъзможно.

Ако степента n на $f(x)=0$ е равна на r, то така с редицата (7) са изчерпани всички негови корени. Ако $r < n$, ще има поне един корен x_2, който не се съдържа в (7). Но видяхме, че уравнението $f[\theta(x)]=0$ допуска всички корени на (1). Следователно $f(x_2)=0$, което показва, че θx_2 е пак корен на (1). С това е установено, че числата

$$x_2, \theta x_2, \theta^2 x_2, \ldots, \theta^{r-1} x_2 \ldots$$

са все корени на (1). Освен това, понеже $\theta^r x - x = 0$ има общ корен с (1), то ще имаме $\theta^r x_2 = x_2$ и r е най-малкото число, за което това равенство съществува. Следователно x_2 дава редицата от r корена на (1)

(8) $$x_2, \theta x_2, \theta^2 x_2, \ldots, \theta^{r-1} x_2; \ \theta^r x_2 = x_2.$$

Числата (8) са отлични от (7). Действително, ако допуснем противното, например

$$\theta^k x_2 = \theta^q x_1,$$

то получаваме

$$\theta^{r-k}\theta^k x_2 = \theta^{r-k+q} x_1,$$
$$\theta^{r-k+q} x_1 = \theta^r x_2 = x_2,$$

т. е. x_2 би принадлежало на (7), което не е вярно. Следователно или $n=2r$, или $n>2r$. Във втория случай ще има корен x_3, който не се съдържа нито в (7), нито в (8) и за който ще получим една подобна редица от r корена. Така продължавайки, достигаме до общия резултат: Степента n на (1) е кратна на r, $n=m.r$ и всичките корени се групират на m реда по r по следния начин:

$$
\begin{array}{l}
x_1,\ \theta x_1,\ \theta^2 x_1,\ \ldots,\ \theta^{r-1} x_1,\\
x_2,\ \theta x_2,\ \theta^2 x_2,\ \ldots,\ \theta^{r-1} x_2,\\
\cdots\cdots\cdots\cdots\\
x_m,\ \theta x_m,\ \theta^2 x_m,\ \ldots,\ \theta^{r-1} x_m
\end{array}
\tag{9}
$$

и $\theta^r(x_i)=x_i,\ i=1,\ 2,\ldots,m$.

2. Редукция към помощни уравнения. Нека

$$f(x)=0 \tag{10}$$

е абелево уравнение, на което корените са представени с таблицата (9). Нека с $y=\varphi(\alpha_1,\ \alpha_2,\ldots,\alpha_r)$ означим една рационална симетрична функция на

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_r$$

и поставим

$$y_1=\varphi(x_1,\ \theta x_1,\ \theta^2 x_1,\ \ldots,\ \theta^{r-1} x_1)=\varphi(x_1),$$

гдето $\varphi(x_1)$ ще бъде рационална функция на x_1. Ако сега заместим x_1 с θx_1 в израза φ, то ще получим

$$y_1=\varphi(\theta x_1,\ \theta^2 x_1,\ \ldots,\ \theta^{r-1} x_1,\ x_1)=\varphi(\theta x_1),$$

понеже φ е симетрична функция. Изобщо ще имаме

$$y_1=\varphi(x_1)=\varphi(\theta x_1)=\varphi(\theta^2 x_1)=\ldots=\varphi(\theta^{r-1} x_1).$$

Ако сега с $y_2,\ y_3,\ldots,y_m$ означим стойностите, които получава y_1, като вместо x_1 поставим съответно $x_2,\ x_3,\ldots,x_m$, ще имаме

$$
\begin{array}{l}
\cdots\cdots\\
y_2=\varphi(x_2)=\varphi(\theta x_2)=\varphi(\theta^2 x_2)=\ldots=\varphi(\theta^{r-1} x_2),\\
\cdots\cdots\cdots\cdots\\
y_m=\varphi(x_m)=\varphi(\theta x_m)=\varphi(\theta^2 x_m)=\ldots=\varphi(\theta^{r-1} x_m).
\end{array}
$$

Числата $y_1,\ y_2,\ldots,y_m$ ще бъдат корени на едно уравнение от m-та степен:

$$y^m+A_1 y^{m-1}+A_2 y^{m-2}+\ldots+A_m=0. \tag{11}$$

За пресмятане на коефициентите му най-напред ще пресметнем степенните сборово. Имаме

$$y_i^k=\frac{1}{r}[\varphi^k(x_i)+\varphi^k(\theta x_i)+\ldots+\varphi^k(\theta^{r-1} x_i)],$$

следователно

$$y_1{}^k+y_2{}^k+\ldots+y_m{}^k=\frac{1}{r}\sum\varphi^k(x),$$

гдето сумирането е разпространяно върху всички корени на уравнението (10). Но тази сума е симетрична функция на корените на това уравнение и следователно е известна. Намирайки така степенните сборове за (11) по формулите на Нютон, получаваме и коефициентите A_i.

След като решим уравнението (11), т. е. намерим корените му $y_1, y_2, \ldots, y_m$, лесно се пресмятат и корените на даденото уравнение, като намерим m уравнения от r-та степен, корените на които са числата от различните редици на (9). Действително нека допуснем, че φ е така избрана, че стойностите $y_1, y_2, \ldots, y_m$ са все различни помежду си. В това по-нататък лесно ще се убедим. Тогава уравнението

$$\varphi(x)-y_1=0$$

има общи корени с $f(x)=0$ само числата

(12) $$x_1,\ \theta x_1,\ \theta^2 x_1, \ldots, \theta^{r-1} x_1,$$

т. е. първата редица числа от таблицата (9). Че (12) са общи корени на поменатите уравнения, следва веднага, понеже

$$\varphi(\theta^k x_1)=y_1.$$

От друга страна, уравнението

$$\varphi(x)-y_1=0$$

при $x=\theta^s x_i$, $i>1$, не се удовлетворява, понеже лявата му част става равна на

$$y_i-y_1 \neq 0.$$

Следователно корените (12) ще са дадени с уравнението

(13) $$D(x, y_1)=x^r+\alpha_1(y_1)x^{r-1}+\ldots+\alpha_r(y_1)=0,$$

гдето $D(x, y_1)$ е общият най-голям делител на

$$\varphi(x)-y_1=0 \text{ и } (10).$$

Корените

$$x_2,\ \theta x_2, \ldots, \theta^{r-1} x_2$$

ще бъдат дадени с уравнението

(14) $$D(x, y_2)=x^r+\alpha_1(y_2)x^{r-1}+\ldots+\alpha_r(y_2)=0$$

и т. н. Най-сетне корените

$$x_m,\ \theta x_m, \theta^2 x_m, \ldots, \theta^{r-1} x_m$$

ще бъдат дадени с уравнението

(15) $$D(x, y_m)=x^r+\alpha_1(y_m)x^{r-1}+\ldots+\alpha_r(y_m)=0.$$

Сега ще установим, че действително функцията φ може да се подбере така, че стойностите ѝ $y_1, y_2, \ldots, y_m$ да бъдат все различни. Така нека поставим

(16) $$y_i=(\alpha-x_i)\ (\alpha-\theta x_i)\ldots(\alpha-\theta^{r-1} x_i),\ i=1,\ 2, \ldots, m.$$

Ако положим $y_i = y_k$, $i \neq k$, то за α получаваме уравнение най-много от $r-1$-ва степен, следователно такова равенство най-много може да бъде изпълнено за $r-1$ стойности на α. Ако поставим $i=1, 2, \ldots, m$; $k=1, 2, \ldots, m$, то виждаме, че най-много за

$$q = \frac{m(m-1)}{2}(r-1)$$

стойности на α може да има две равни стойности на y. Ако следователно изберем α отлично от тези q стойности, то функцията (16) ще приема само различни помежду си значения:

$$y_1, y_2, \ldots, y_m.$$

Така решението на уравнението (10) сведохме посредством уравнението от m-та степен (11) към решение на m уравнения от r-та степен. Уравнението (11) въобще не е абелево, но уравненията (13), (14),...,(15) са абелеви, на които корените образуват един период. Такива абелеви уравнения се наричат ц и к л и ч н и.

Един прост пример на абелевите уравнения са реципрочните. Ако

$$f(x) = 0$$

е реципрочно уравнение и x_1 е един кой да е негов корен, то ще има и корена

$$\frac{1}{x_1} = \theta(x_1),\ \theta^2(x_1) = x_1.$$

Тогава, ако степента е $2m$, то ще имаме следната редица от корени:

$$x_1,\ \theta x_1,$$
$$x_2,\ \theta x_2,$$
$$\ldots\ldots$$
$$x_m, \theta x_m.$$

С въвеждане на

$$y = x + \frac{1}{x} = x + \theta x$$

получихме уравнение от m-та степен спрямо y и на всеки корен y_i на това уравнение ще съответствуват два на даденото, получени от

$$x^2 - y_i x + 1 = 0.$$

3. Циклични уравнения. Да разгледаме случая, когато $n=r$, т. е. уравнението е ц и к л и ч н о. Следователно корените на уравнението от n-та степен

(16′) $$f(x) = 0$$

ще бъдат

$$x_1,\ \theta x_1,\ \theta^2 x_1, . , . , \theta^{n-1} x_1,\ \theta^n x_1 = x.$$

Нека α е произволен корен на $x^n=1$ и нека поставим

$$y=\varphi(x)=(x_1+\alpha\theta x_1+\alpha^2\theta^2 x_1+\ \ldots\ +\alpha^{n-1}\theta^{n-1}x_1)^n.$$

Тази функция от корените, наречена Лагранжова, е симетрична функция на корените на (16). Действително, ако сменим x_1 с $\theta^i x_1$, получаваме

$$\varphi(\theta^i x_1)=(\theta^i x_1+\alpha\theta^{i+1}x_1+\ldots+\alpha^{n-1}\theta^{n-1+i}x_1)^n=$$

$$=(\alpha^{-i})^n(\alpha^i\theta^i x_1+\alpha^{i+1}\theta^{i+1}x_1+\ \ldots\ +\alpha^{n-1+i}\theta^{n-1+i}x_1)^n.$$

Понеже

$$\alpha^{n-1+i}=\alpha^{i-1},\ \theta^{n-1+i}x_1=\theta^{i-1}x,\ \ldots\ \text{и}\ (\alpha^{-i})^n=\alpha^{-in}=1,$$

то ще имаме

$$\varphi(\theta^i x_1)=\varphi(x_1),$$

отгдето при $i=1, 2, 3, \ldots, n-1$,

$$\varphi(x_1)=\varphi(\theta x_1)=\varphi(\theta^2 x_1)=\ \ldots\ =\varphi(\theta^{n-1}x_1),$$

$$\varphi(x_1)=\frac{1}{n}[\varphi(x_1)+\varphi(\theta x_1)+\ldots+\varphi(\theta^{n-1}x_1)].$$

Следователно $\varphi(x_1)$ е симетрична функция от корените на (16) и ще може да се изрази рационално посредством коефициентите му.

Нека сега $y_0, y_1, y_2, \ldots, y_{n-1}$ са стойностите на y, когато за α вземем n-те корена на $x^n=1$:

$$1,\ \alpha_1, \alpha_2,\ \ldots\ , \alpha_{n-1}.$$

$\sqrt[n]{y_0}$ представлява сумата от корените на (16), така че, ако A е коефициентът пред x^{n-1} в това уравнение, а коефициентът пред x^n е 1, ще имаме

$$\sqrt[n]{y_0}=-A.$$

Така ще получим уравненията:

$$\left\{\begin{array}{llllll}
x_1+ & \theta x_1+ & \theta^2 x_1+\ldots+ & & & \theta^{n-1}x_1=-A,\\
x_1+ & \alpha_1\theta x_1+ & \alpha_1{}^2\theta^2 x_1+\ \ldots\ + & \alpha_1{}^{n-1}\theta^{n-1}x_1=\sqrt[n]{y_1},\\
x_1+ & \alpha_2\theta x_1+ & \alpha_2{}^2\theta^2 x_2+\ \ldots\ + & \alpha_2{}^{n-1}\theta^{n-1}x_1=\sqrt[n]{y_2},\\
\ldots & \ldots & \ldots & \ldots\\
x_1+\alpha_{n-1}\theta x_1+ & \alpha^2{}_{n-1}\theta^2 x_1+\ \ldots\ + & & \alpha_{n-1}^{n-1}\theta^{n-1}x_1=\sqrt[n]{y_{-1}}.
\end{array}\right.$$

Ако съберем тия равенства, ще имаме

$$x_1=\frac{1}{n}(-A+\sqrt[n]{y_1}+\sqrt[n]{y_2}+\ldots+\sqrt[n]{y_{n-1}}). \tag{17}$$

Като ги умножим с α^{i-m} и съберем, ще имаме

$$\theta^m x_1 = \frac{1}{n}\left(-A+\alpha_1^{-m}\sqrt[n]{y_1}+\alpha_2^{-m}\sqrt[n]{y_2}+\ldots+\alpha_{n-1}^{-m}\sqrt[n]{y_{n-1}}\right). \tag{18}$$

Ако в (17) и (18) вземем за $\sqrt[n]{\ }$ всички възможни стойности, то за x получаваме n^{n-1} стойности, а уравнението (10) има само n корена. Това усложняване обаче се премахва, както непосредствено ще видим, то обстоятелството, че стойността на единия радикал, бидейки избрана, стойностите на останалите радикали се напълно определят.

Действително нека α е примитивен корен на

$$x^n = 1.$$

Тогава корените $1, \alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ могат да се представят така:

$$1, \alpha, \alpha^2, \ldots, \alpha^{n-1}.$$

Следователно ще имаме

$$\sqrt[n]{y_1} = x_1+\alpha\theta x_1+\alpha^2\theta^2 x_1+\ldots+\alpha^{n-1}\theta^{n-1}x_1,$$
$$\sqrt[n]{y_k} = x_1+\alpha^k\theta x_1+\alpha^{2k}\theta^2 x_1+\ldots+\alpha^{(n-1)k}\theta^{n-1}x_1.$$

Ако в тези формули сменим x_1 с $\theta^n x_1$, то $\sqrt[n]{y_1}$ се заменя с $\alpha^{n-m}\sqrt[n]{y_1}$, $\sqrt[n]{y_k}$ с $\alpha^{k(n-m)}\sqrt[n]{y_k}$. Следователно изразът

$$\psi(x_1) = (\sqrt[n]{y_1})^{n-k}\sqrt[n]{y_k}$$

при заместване на x_1 с $\theta^m x_1$ става

$$\psi(\theta^m x_1) = (\alpha^{n-m}\sqrt[n]{y_1})^{n-k}[\alpha^{k(n-m)}\sqrt[n]{y_k}] =$$
$$= \alpha^{n(n-m)}\,(\sqrt[n]{y_1})^{n-k}\sqrt[n]{y_k} = \psi(x_1).$$

Оттук имаме

$$\psi(x_1) = \psi(\theta x_1) = \ldots = \psi(\theta^{n-1}x_1) =$$
$$= \frac{1}{n}[\psi(x_1)+\psi(\theta x_1)+\ldots+\psi(\theta^{n-1}x_1)];$$

$\psi(x_1)$ е симетрична функция от корените на (16) и следователно се изразява рационално от коефициентите на $f(x)$. Ако с a_k означим стойността на $\psi(x_1)$, то

$$\sqrt[n]{y_k} = \frac{a_k}{y_1}(\sqrt[n]{y_1})^k. \tag{19}$$

Формулата (17) става

$$x = \frac{1}{n}\left[\left(-A+\sqrt[n]{y_1}+\frac{a_2}{y_1}(\sqrt[n]{y_1})^2+\frac{a_3}{y_1}(\sqrt[n]{y_1})^3+\ldots+\frac{a_{n-1}}{y_1}(\sqrt[n]{y_1})^{n-1}\right]. \tag{20}$$

Този израз дава n-те корена на уравнението $f(x)=0$, като даваме n-те стойности на радикала $\sqrt[n]{y_1}$.

Както ще видим по-нататък, биномните уравнения са решими алгебрически. Следователно уравненията на Абел, в които корените образуват само един период

$$x_1,\ \theta x_1,\ \theta^2 x_1,\ldots,\theta^{n-1}x_1,\ \theta^n x_1=x_1,$$

са решими алгебрически.

4. Уравнения с реални коефициенти. В случая, когато коефициентите на $f(x)$ и $\theta(x)$ са реални, y_1 съдържа само комплексното число α. Ако α сменим в y_1 с $\alpha^{n-1}=\frac{1}{\alpha}$, то ще получим y_{n-1}. Понеже α и $\frac{1}{\alpha}$ са конюговани, то следва оттук, че y_1 и y_{n-1} са конюговани. Следователно можем да поставим

$$y_1=\rho(\cos\varphi+i\sin\varphi),\quad y_{n-1}=\rho(\cos\varphi-i\sin\varphi), \tag{21}$$

гдето $\rho>0$ и φ е аргументът на y_1.

Тогава от (19) при $k=n-1$ имаме

$$\sqrt[n]{y_1}\cdot\sqrt[n]{y_{n-1}}=a_{n-1}. \tag{22}$$

Понеже при промяна на α с $\frac{1}{\alpha}$, a_{n-1} не се изменя, то следва, че α_{n-1} е реално число. Но тогава от (21) и (22) следва

$$\rho^2=\alpha_{n-1}^n,\ \sqrt[n]{\rho}=\sqrt{a_{n-1}},$$

$$\sqrt[n]{y_1}=\sqrt{a_{n-1}}\left(\cos\frac{\varphi+2r\pi}{n}+i\sin\frac{\varphi+2r\pi}{n}\right),$$

гдето $r=0,\ 1,\ 2,\ldots,\ n-1$, и от (19)

$$\sqrt[n]{y_k}=\frac{a_k}{y_1}\sqrt{a_{n-1}^k}\left(\cos\frac{\varphi+2r\pi}{n}+i\sin k\,\frac{\varphi+2r\pi}{n}\right).$$

Тук a_k, y_1 са рационални функции от коефициентите на $f(x)$, $\theta(x)$ и от $\alpha=\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n}$. Ако заместим в (20), получаваме (пишейки $a_{n-1}=a$ за простота)

$$x=\frac{1}{n}\left\{-A+\sqrt{a}\left(\cos\frac{\varphi+2r\pi}{n}+i\sin\frac{\varphi+2r\pi}{n}\right)+\right.$$

$$+(g_2+h_2 i)\,a\left(\cos 2\frac{\varphi+2r\pi}{n}+i\sin 2\,\frac{\varphi+2r\pi}{n}\right)+$$

$$+ (g_2+h_3i)\sqrt[3]{a^3}\left(\cos 3\,\frac{\varphi+2r\pi}{n}+\mathrm{i}\sin 3\,\frac{\varphi+2r\pi}{n}\right)+\ldots\Big\},$$

гдето g, h се изразяват рационално от същите количества, както a_k и y_1. Поставяйки $r=0,\ 1,\ 2,\ldots,n-1$, получаваме всичките корени на $f(x)=0$.

От последната формула се вижда, че решението зависи от α, т. е. от $\cos\frac{2\pi}{n}$ и $\sin\frac{2\pi}{n}$, от $\frac{\varphi}{n}$, гдето φ е ъгъл, даден от (21), чийто тангенс се изразява рационално чрез α, и от извличане на квадратен корен на едно реално число a.

Понеже $\theta(x)$ е реална функция, то очевидно, ако един корен x_1 е реален, всички са реални. Корените на уравнението са всички реални или всички имагинерни.

5. Циклично уравнение със съставна степен. Нека уравнението на Абел.

$$f(x)=0 \tag{23}$$

е циклично и корените му са

$$x_1,\ \theta x_1,\ \theta^2x_1,\ldots,\ \theta^{n-1}x_1,$$

гдето $\theta^n x_1=x_1$, и нека предположим, че степента $n=mr$.

Ще видим, че решението на уравнението може значително да се опрости. Да разпределим корените му в една правоъгълна матрица така:

$$\begin{array}{llll}
x_1, & \theta^m x_1, & \theta^{2m}x_1 & ,\ldots,\ \theta^{(r-1)m}x_1,\\
\theta x_1, & \theta^{m+1}x_1, & \theta^{2m+1}x_1 & ,\ldots.,\ \theta^{(r-1)m+1}x_1,\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
\theta^{m-1}x_1, & \theta^{2m-1}x_1, & \theta^{3m-1}x_1 & ,\ldots,\ \theta^{rm-1}x_1.
\end{array}$$

Ако поставим

$$x_2=\theta x_1,\ x_3=\theta^2x_1,\ldots,x_m=\theta^{m-1}x_1,\ \theta^m x=\theta_1 x,$$

предната матрица става

$$\begin{array}{l}
x_1,\ \theta_1x_1,\ \theta_1^2x_1,\ldots,\theta_1{}^{r-1}x_1,\\
x_2,\ \theta_1x_2,\ \theta_1^2x_2,\ldots,\theta_1{}^{r-1}x_2,\\
\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\
x_m,\ \theta_1x_m,\ \theta_1^2x_m,\ldots,\theta_1{}^{r-1}x_m
\end{array} \tag{24}$$

и освен това

$$\theta_1{}^r x_1=\theta^{rm}x_1=\theta^n x_1=x_1,\ \theta_1{}^r x_2=\theta^{rm}\theta x_1=\theta^n x_2=x_2,\ldots,\theta_1{}^r x_m=x_m.$$

Така имаме същата система от корени, както в пар. 1 за абелево уравнение, на което корените могат да се разпределят в m реда по r,

само че вместо θ имаме θ_1. Следователно може да се приложи методът, изложен в пар. 2. Именно, ако поставим

$$y_1 = F(x_1,\ \theta_1 x_1, \ldots, \theta_1{}^{r-1} x_1) = \varphi(x_1),$$
$$y_2 = F(x_2,\ \theta_1 x_2, \ldots, \theta_1{}^{r-1} x_2) = \varphi(x_2),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$y_m = F(x_m, \theta_1 x_m, \ldots, \theta_1{}^{r-1} x_m) = \varphi(x_m),$$

гдето $F(\alpha_1, \alpha_2, \ldots, \alpha_r)$ е симетрична функция на $\alpha_1, \alpha_2, \ldots, \alpha_r$, то числата $y_1, y_2, \ldots, y_m$ са корени на едно уравнение

$$y^m + A_1 y^{m-1} + \ldots + A_m = 0, \tag{25}$$

на което коефициентите се изразяват рационално посредством коефициентите на даденото уравнение. Редиците по r корена в (24) са корени на следните m абелеви уравнения:

$$\begin{aligned} &x^r + \alpha_1(y_1) x^{r-1} + \ldots + \alpha_r(y_1) = 0,\\ &x^r + \alpha_1(y_2) x^{r-1} + \ldots + \alpha_r(y_2) = 0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &x^r + \alpha_1(y_m) x^{r-1} + \ldots + \alpha_r(y_m) = 0. \end{aligned} \tag{26}$$

Тези уравнения са циклични и както видяхме, са решими алгебрически. Уравнението (15) в общия случай, когато уравнението (23) не е циклично, е въобще нерешимо алгебрически, но в дадения случай, понеже (23) е решимо, то следва, че и (25) е решимо алгебрически. Ние ще покажем, че това последно уравнение е пак едно абелево циклично уравнение.

От

$$y_2 = F(x_2, \theta_1 x_2, \ldots, \theta_1{}^{r-1} x_2) = F(\theta x_1,\ \theta\theta_1 x_1, \theta\theta_1^2 x_1, \ldots, \theta\theta_1{}^{r-1} x_1)$$

се вижда, че y_2 е симетрична функция на корените

$$x_1,\ \theta_1 x_1,\ \theta_1 x_1^2, \ldots, \theta_1{}^{r-1} x_1. \tag{27}$$

Но тези корени са дадени с първото уравнение (26), коефициентите на което са рационални функции на y_1. Следователно y_2 ще бъде рационална функция на y_1, т. е.

$$y_2 = \varphi(x_2) = \varphi(\theta x_1) = \omega(y_1) = \omega\varphi(x_1).$$

Подобно имаме вследствие пълната симетричност

$$y_3 = \varphi(\theta^2 x_1) = \omega\varphi(\theta x_1) = \omega^2\varphi(x_1),$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$y_m = \varphi(\theta^{m-1} x_1) = \omega\varphi(\theta^{m-2} x_1) = \ldots = \omega^{m-1}\varphi(x_1).$$

Корените на (25) образуват една редица

$$y_1,\ \omega y_1,\ \omega^2 y_1, \ldots,\ \omega^{m-1} y,\ \omega^m y_1 = \varphi(\theta^m x_1) = y_1.$$

Ако m е пак съставно число, $m = m_1 r_1$, то върху него може да се приложи същият анализ, като се сведе решението на абелеви циклични от степен m_1 и от r_1-та степен. Продължавайки така, достигаме до следното заключение: ако

$$n = p^\lambda q^\mu r^\nu \ldots,$$

то решението на едно абелево уравнение от n-та степен, на което корените образуват само един период, се свежда към решението на подобни уравнения от степен p, q, r, ...

Оттук се вижда, че всяко абелево циклично уравнение от степен 2^λ е решимо с квадратни радикали.

Глава II

Алгебрическо решение на биномните уравнения

1. Биномните уравнения, разглеждани като абелеви. Едни от най-простите абелеви уравнения са биномните. По-рано видяхме, че решението на биномните уравнения

$$x^n = 1$$

се свежда към решение на биномни уравнения, степените на които са прости числа. Следователно можем да се ограничим с биномното уравнение

$$x^p = 1,$$

в което числото p е просто. Като отстраним корена 1, получаваме уравнението

$$\frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \ldots + x + 1 = 0, \tag{1}$$

на което, ако α е един примитивен корен, всички корени ще бъдат

$$\alpha,\ \alpha^2,\ \alpha^3, \ldots, \alpha^{p-1}. \tag{2}$$

Оттук очевидно веднага следва, че уравнението (1) е абелево, понеже е неразложимо. За корените на уравнението, дадени с (2), ще докажем, че образуват един период, т. е. че това уравнение е циклично. Нека r е един примитивен корен на

$$x^{p-1} \equiv 1 \pmod{p};$$

тогава числата

$$1,\ r.r^2, \ldots, r^{p-2},$$

разделени с p, имат за остатъци числата

$$1,\ 2, \ldots, p-1,$$

само че изобщо в друг ред. Понеже

$$\alpha^p = 1,$$

то корените (2) могат да се представят така:

$$\alpha,\ \alpha^r,\ \alpha^{r^2},\ \ldots,\ \alpha^{r^{p-2}}\ (\alpha^{r^{p-1}} = \alpha).$$

Ако поставим $\theta(x) = x^r$, то очевидно тази редица може да се напише така:

$$\alpha,\ \theta\alpha,\ \theta^2\alpha,\ \ldots,\ \theta^{p-2}\alpha\ \ (\theta^{p-1}\ \alpha = \alpha).$$

Следователно решението на уравнението (1) може да стане по изложения вече метод. Нека β е корен на

$$x^{p-1} = 1 \tag{3}$$

и да образуваме функцията

$$y = \varphi(\alpha) = (\alpha + \beta\alpha^r + \beta^2\alpha^{r^2} + \ldots + \beta^{p-2}\alpha^{r^{p-2}})^{p-1},$$

на която стойността ще зависи от β и е симетрична функция на корените на (1) и следователно ще бъде известна, ако е известно β. Но следвайки индуктивния път, ние можем да предположим, че биномните уравнения от степен, по-ниска от p, са решими алгебрически. Ако тогава

$$y_0,\ y_1,\ \ldots,\ y_{p-2}$$

са стойности на y, когато на β даваме $p-1$ на брой стойности, равни на корените на (3), то по предидущата глава от (17) получаваме за корените x на (1)

$$x = \frac{1}{p-1}\left\{-1 + \sqrt[p-1]{y_1} + \sqrt[p-1]{y_2} + \ldots + \sqrt[p-1]{y_{p-2}}\ \right\}.$$

Останалите радикали освен $\sqrt[p-1]{y_1}$ можем да определим, както в глава VIII, пар. 3. Така получихме, че и биномните уравнения от степен p са решими алгебрически, понеже са решими уравненията $x^2 = 1$, $x^3 = 1$, а оттам следва по индуктивен път, че всички биномни уравнения са решими алгебрически.

Понеже $p-1$ не е просто число, то можем уравнението (1) да го сведем на циклични уравнения от по-ниски степени, докато достигнем до такива със степени, равни на прости числа.

Така, ако $p = 2^k + 1$, то уравнението (1) се решава с квадратни радикали, понеже води до решение на квадратни уравнения.

2. Друго решение. Уравнението (1) е реципрочно и следователно със субституцията

$$z = x + \frac{1}{x}$$

ще можем да го сведем към уравнение от два пъти по-ниска степен. Ако $p=2v+1$, новополученото уравнение с означенията в глава II ще бъде от степен v:

$$P_v+P_{v-1}+\ldots+P_1+1=0, \tag{4}$$

гдето

$$P_k=x^k+\frac{1}{x^k}.$$

Уравнението (4) е пак абелево. Понеже корените на (1) са

$$x_k=\cos k\frac{2\pi}{p}+i\sin k\frac{2\pi}{p},$$

$$k=1,\ 2,\ldots,\ p-1,$$

то корените на (4) ще бъдат

$$z=2\cos\frac{2k\pi}{p},\ k=1,\ 2,\ 3,\ldots,\ v=\frac{p-1}{2},$$

тъй като двата корена x_k и x_{p-k} са, както знаем, конюговани, т. е. реципрочни.

За да наредим корените на (4) в една редица с един период, трябва да изследваме кои корени от

$$\alpha,\ \alpha^r,\ \alpha^{r^2},\ldots,\alpha^{r^{p-2}} \tag{5}$$

на уравнението

$$\frac{x^p-1}{x-1}=0$$

са реципрочни помежду си. Ако $\alpha^{r^i}, \alpha^{r^k}$ са два такива, то от условието

$$\alpha^{r^i}\ \alpha^{r^k}=\alpha^{r^i+r^k}=1$$

следва

$$r^i+r^k\equiv 0\ (\mathrm{mod}\,p),$$

$$r^i\equiv -r^k,$$

отгдето, ако $i>k$,

$$r^{i-k}\equiv -1,\quad r^{2(i-k)}\equiv 1.$$

Но понеже $i,\ k\leqq p-2,\ i-k<p-1,\ 2(i-k)<2(p-1)$ и r е примитивен корен на $x^{p-1}\equiv 1\ (\mathrm{mod}\ p)$, то следва или

$$2(i-k)=p-1,\ i-k=v,\ \text{или}\ i-k=0,$$

което е невъзможно, понеже първата от двете горни формули би се обърнала в $1\equiv -1$, което при $p>2$ е невъзможно.

Следователно в (5) на корена с r в нулева степен е реципрочен този с r в степен v. На корена с r в първа е реципрочен този с r във $v+1$ и т. н., на $v-1$ отговаря този с

$$2v-1=p-2.$$

Следователно корените на уравнението (4) могат да се пишат, като поставим $\frac{2\pi}{p}=\frac{2\pi}{2v+1}=a,$

(6) $$2\cos a,\quad 2\cos ra,\quad 2\cos r^2a,\ldots,2\cos r^{v-1}a.$$

Обаче известно е, че $2\cos ra$ е полином на $2\cos a$, т. е. ако поставим $2\cos a=x$, то $2\cos ra=P(x)$. Тогава редицата (6) може да се пише

$$2\cos a=x,\ P(x),\ P^2(x),\ldots,P^{v-1}(x),\ P^v(x)=2\cos r^va=$$

$$=2\cos a=x,\text{ понеже } r^v\equiv -1,$$

гдето с $P^2(x)$ е означен полиномът $P[P(x)]$, и т. н.

3. Примери. 1. Дадено е уравнението

(6) $$x^5-1=0,$$

от което с отстраняване на корена 1 добиваме

(7) $$\frac{x^5-1}{x-1}=x^4+x^3+x^2+x+1=0.$$

Това реципрочно уравнение решаваме, като поставим

$$x+\frac{1}{x}=z.$$

За z получаваме уравнението

(8) $$z^2+z-1=0.$$

Ако z_1, z_2 са корените на това уравнение, то корените на (7) ще бъдат дадени с

$$x^2-z_1x+1=0,\ x^2-z_2x+1=0.$$

Ако обаче върху (8) приложим за упражнение общия метод, то трябва да намерим примитивен корен на

$$x^4\equiv 1\ (\mathrm{mod}\ 5).$$

Най-малкият примитивен корен е 2. Трябва да образуваме функцията

$$y=(2\cos a+\alpha\,.\,2\cos 2a)^2,$$

гдето α е корен на $x^2-1=0$, т. е. $\alpha=\pm 1$. За $\alpha=1$ ще имаме

$$\sqrt{y_0}=2\cos a+2\cos 2a=-1,$$

а за $\alpha=-1$:

$$y_1=4\cos^2 a+4\cos^2 2a-8\cos a\cos 2a=(z_1+z_2)^2-4z_1z_2=5.$$

Следователно корените на (8) ще бъдат

$$z_{1,2}=\frac{1}{2}(-1\pm\sqrt{5}).$$

Ако върху (7) искаме да приложим метода, изложен в § 5 на предната глава, трябва да намерим един примитивен корен r на конгруенцията

$$x^4\equiv 1\pmod 5,$$

например корена $r=2$, и да образуваме редицата

$$\alpha,\ \alpha^r,\ \alpha^{r^2},\ \alpha^{r^3},$$

която е редицата

$$\alpha,\ \alpha^2,\ \alpha^4,\ \alpha^3,$$

гдето α е примитивен корен на (7). Тези корени разделяме на две групи по два:

$$\alpha,\ \alpha^4,$$

$$\alpha^2,\ \alpha^3$$

и образуваме

$$y_1=\alpha+\alpha^4,$$

$$y_2=\alpha^2+\alpha^3.$$

Тогава от

$$y_1+y_2=\alpha+\alpha^2+\alpha^3+\alpha^4=-1,$$

$$y_1y_2=\alpha^3+\alpha^4+\alpha^6+\alpha^7=\alpha^2+\alpha^4+\alpha+\alpha^2=-1$$

следва, че y_1, y_2 са корени на квадратното уравнение

$$y^2+y-1=0.$$

Понеже $\alpha.\alpha^4=1$, $\alpha^2.\alpha^3=1$, то α, α^4 са корени на уравнението

$$x^2-y_1x+1=0,$$

а α^2, α^3 — на уравнението

$$x^2-y_2x+1=0.$$

2. Нека е дадено уравнението

$$x^7-1=0,$$

което, като отстраним корена 1, става

(9) $$x^6+x^5+x^4+x^3+x^2+x+1=0.$$

Един примитивен корен е

$$\alpha = \cos\frac{2\pi}{7} + i\sin\frac{2\pi}{7}$$

и всички корени са

(10) $$\alpha,\ \alpha^2,\ \alpha^3,\ \alpha^4,\ \alpha^5,\ \alpha^6.$$

За да приложим методите за решението на абелевите уравнения, търсим примитивен корен на

$$x^6 \equiv 1 \pmod 7.$$

Такива са 3 и 5. Ако използуваме по-малкия 3, то (10) се нареждат в реда

$$\alpha,\ \alpha^3,\ \alpha^{3^2},\ \alpha^{3^3},\ \alpha^{3^4},\ \alpha^{3^5} \quad (\alpha^{3^6}=\alpha)$$

или

$$\alpha,\ \alpha^3,\ \alpha^2,\ \alpha^6,\alpha^4,\ \alpha^5.$$

Понеже $6=2.3$, то можем да ги разделим в три групи по два:

(11)
$$\begin{array}{ll} \alpha\ ,\ \alpha^{3^3} \text{ или } \alpha,\ \alpha^6, \\ \alpha^3,\ \alpha^{3^1} \qquad \alpha^3,\ \alpha^4, \\ \alpha^{3^2},\ \alpha^{3^5} \qquad \alpha^2,\ \alpha^5. \end{array}$$

Ако положим

$$z_1=\alpha+\alpha^6,\ z_2=\alpha^3+\alpha^4,\ z_3=\alpha^2+\alpha^5,$$

то z_1, z_2, z_3 са корени на кубичното уравнение

$$z^3+pz^2+qz+r=0,$$

гдето

$$-p=z_1+z_2+z_3=\Sigma\alpha^i=-1$$

и $\sum$ е разпростряна върху всички корени на (9).

$$q=(\alpha+\alpha^6)(\alpha^3+\alpha^4)+(\alpha+\alpha^6)(\alpha^2+\alpha^5)+(\alpha^3+\alpha^4)(\alpha^2+\alpha^5)=$$

$$=2\sum\alpha=-2,\ -r=(\alpha+\alpha^6)(\alpha^3+\alpha^4)(\alpha^2+\alpha^5)=2+\sum\alpha=1.$$

Така че уравнението за z ще бъде

$$z^3+z^2-2z-1=0.$$

Понеже $\alpha.\alpha^6=1$, $\alpha^3.\alpha^4=1$, $\alpha^2.\alpha^5=1$, то корените от (11) ще бъдат дадени с

$$x^2-z_1x+1=0,$$
$$x^2-z_2x+1=0,$$
$$x^2-z_3x+1=0.$$

Вместо да разделим шестте корена в три групи по два, можем да ги разделим в две групи по три:

$$\alpha\,,\ \alpha^{3^2},\ \alpha^{3^4} \qquad\qquad \alpha\,,\ \alpha^2,\ \alpha^4$$

(12) или

$$\alpha^3,\ \alpha^{3^3},\ \alpha^{3^5} \qquad\qquad \alpha^3,\ \alpha^6,\ \alpha^5.$$

Тогава поставяме

$$y_1=\alpha+\alpha^2+\alpha^4,$$

$$y_2=\alpha^3+\alpha^6+\alpha^5,$$

при което y_1, y_2 са корени на уравнението

$$y^2+Ay+B=0,$$

където

$$-A=\sum\alpha=-1,\quad A=1.$$

$$B=(\alpha+\alpha^2+\alpha^4)\,(\alpha^3+\alpha^6+\alpha^5)=$$

$$=3+\sum\alpha=2.$$

Следователно уравнението за y е

$$y^2+y+2=0.$$

Корените от първата група на (12) са дадени с

$$x^3+px^2+qx+r=0,$$

където

$$-p=\alpha+\alpha^2+\alpha^4=y_1,$$

$$q=\alpha\alpha^2+\alpha\alpha^4+\alpha^2\alpha^4=y_2=-1-y_1,$$

$$-r=\alpha\,.\,\alpha^2\,.\,\alpha^4=\alpha^7=1,$$

така че уравнението е

$$x^3-y_1x^2-(1+y_1)\,x-1=0.$$

Корените от втората група на (12) ще бъдат дадени с

$$x^3-y_2x^2-(1+y_2)\,x-1=0.$$

Тези две уравнения са пак абелеви с един период, които лесно решаваме.

Можем върху (9) да приложим другия метод. Поставяме отначало

$$x+\frac{1}{x}=z,$$

така че получаваме уравнението

(13) $$z^3+z^2-2z-1=0,$$

корените на което ще бъдат

$$2\cos a,\ 2\cos 3a,\ 2\cos 3^2a\ (2\cos 3^3a = 2\cos 2a),\quad a=\frac{2\pi}{7},$$

или

$$2\cos a,\ 2\cos 3a,\ 2\cos 2a.$$

Образуваме функцията

$$y=(2\cos a+\beta\,.\,2\cos 3a+\beta^2\,.\,2\cos 2a)^3,$$

гдето β **е корен на**

$$x^3=1.$$

При $\beta=1$

$$y_0=\left(\sum z\right)^3=-1.$$

Ако β е имагинерен корен, то понеже $\beta^3=1$, $\beta^4=\beta$, $\beta^5=\beta^2$, **ще имаме**

$$y=8\cos^3 a+8\cos^2 3a+8\cos^3 2a+$$
$$+24\,(\cos^2 a\cos 3a+\cos^2 2a\cos a+\cos^2 3a\cos 2a)\beta+$$
$$+24\,(\cos a\cos^2 3a+\cos 2a\cos^2 a+\cos 3a\cos^2 2a)\,\beta^2+$$
$$+48\cos a\cos 2a\cos 3a.$$

Ако сега преработим всеки член, като използуваме релацията

$$2\cos ma\,.\cos na=\cos(m+n)\,a+\cos(m-n)\,a$$

и като забележим, че $\cos 4a=\cos 3a$, $\cos 5a=\cos 2a$, **то**

$$\cos^2 a\cos 3a=\frac{1}{2}\cos a\,(\cos 3a+\cos 2a),$$

$$\cos a\cos^2 3a=\frac{1}{2}\cos 3a\,(\cos 3a+\cos 2a),$$

. . . . . .

отгдето получаваме, че коефициентът на β **е** $6\sum z_1z_2$, **а коефициентът на** β^2 **е**

$$3\left(\sum z_1^2+\sum z_1z_2\right).$$

Следователно ще имаме

$$y=\sum z_1^3+6z_1z_2z_3+6\beta\sum z_1z_2+3\beta^3\left(\sum z_1^2+\sum z_1z_2\right)$$

или понеже $\beta^2+\beta+1=0$, $\beta^2=-\beta-1$,

$$y=\sum z_1^3+6z_1z_2z_3-3\left(\sum z_1^2+\sum z_1z_2\right)+$$
$$+3\beta\left(\sum z_1z_2-\sum z_1^2\right).$$

Но от уравнението (13) имаме

$$\sum z_1^3 = -a_1^3 + 3a_1a_2 - 3a_3 = 4,$$

$$\sum z_1^2 = a_1^2 - 2a_2 = 5,$$

$$\sum z_1z_2 = -2,$$

$$z_1z_2z_3 = 1,$$

така че

$$y_1 = -7 - 21\beta_1, \quad y_2 = -7 - 21\beta_2,$$

$$z = \frac{1}{3}\left(-1 + \sqrt[3]{-7-21\beta_1} + \sqrt[3]{-7-21\beta_2}\right),$$

или понеже

$$\beta_{1,2} = \frac{-1 \pm \sqrt{-3}}{2},$$

отгдето

$$z = \frac{1}{3}\left(-1 + \sqrt[3]{\frac{7}{2} - \frac{7}{2}\,3\sqrt{-3}} + \sqrt[3]{\frac{7}{2} + \frac{7}{2}\,3\sqrt{-3}}\right).$$

Същия израз можем да получим с формулата на Кардано.

3. Нека е дадено уравнението

$$x^{13} = 1.$$

Като отстраним корена 1 и положим

$$z = x + \frac{1}{x},$$

получаваме уравнението

(19) $$z^6 + z^5 - 5z^4 - 4z^3 + 6z^2 + 3z - 1 = 0,$$

на което корените са

$$z = 2\cos\frac{2\pi k}{13}, \quad k = 1, 2, \ldots, 6.$$

Ако r е примитивен корен на конгруенцията

$$x^{12} \equiv 1 \pmod{13},$$

то корените на (19) са

(20) $$2\cos a,\ 2\cos ra, \ldots,\ 2\cos r^5a,$$

гдетo $a = \frac{2\pi}{13}$. Конгруенцията има примитивни корени 2, 6, 7, 11. Да вземем най-малкия $r = 2$. Тогава редицата (20) е

$$2\cos a,\ 2\cos 2a,\ 2\cos 4a,\ 2\cos 5a,\ 2\cos 3a,\ 2\cos 6a.$$

Можем вече върху абелевото уравнение (19) да приложим общите методи.

Понеже $6=2\,.\,3$, разделяме тези корени в две групи по три:

$$\begin{matrix} 2\cos a,\ 2\cos r^2q,\ 2\cos r^4a \\ \text{или} \\ 2\cos ra,\ 2\cos r^3a,\ 2\cos r^5a \end{matrix} \qquad \begin{matrix} 2\cos a,\ 2\cos 4a,\ 2\cos 3a, \\ \\ 2\cos 2a,\ 2\cos 5a,\ 2\cos 6a. \end{matrix} \tag{21}$$

Образуваме, следвайки общия метод, една симетрична функция от редовете, например

$$y_1=2\cos a+2\cos 4a+2\cos 3a,$$

$$y_2=2\cos 2a+2\cos 5a+2\cos 6a.$$

Тогава y_1+y_2 е сумата на всички корени на (19), т. е.

$$y_1+y_2=-1.$$

След това, като вземем под внимание формулата

$$2\cos\alpha\cos\beta=\cos(\alpha+\beta)+\cos(\alpha-\beta),$$

получаваме

$$y_1y_2=3\times\text{сумата на всички корени}=-3.$$

Следователно y_1, y_2 са корени на уравнението

$$y^2+y-3=0.$$

Лесно е да намерим уравнение, на което корените са първият ред в (21). Нека означим

$$x_1=2\cos a,\quad x_2=\cos 4a,\quad x_3=2\cos 3a;$$

тогава получаваме

$$\sum x_1=x_1+x_2+x_3=y_1,$$

$$\sum x_1x_2=\sum z_1=-1,$$

гдето $\sum z_1$ е сумата от корените на уравнението (19).

$$x_1x_2x_3=2\cos 5a+2\cos 6a+2\cos 2a+2=y_2+2=1-y_1.$$

Уравнението за x_1, x_2, x_3 следователно ще бъде

$$x^3-y_1x^2-x-1+y_1=0.$$

Тогава уравнението, на което корените са от втория ред на (21), ще бъде

$$x^3-y_2x^2-x-1+y_2=0.$$

Вместо да разделяме корените на две групи по три, можем да ги разделим на три групи по два, именно

$$\begin{matrix} 2\cos a, & 2\cos 5a, \\ 2\cos 2a, & 2\cos 3a, \\ 2\cos 4a, & 2\cos 6a. \end{matrix} \tag{22}$$

Нека тогава

$$y_1 = 2\cos a + 2\cos 5a,$$

$$y_2 = 2\cos 2a + 2\cos 3a,$$

$$y_3 = 2\cos 4a + 2\cos 6a;$$

получаваме $$y_1 + y_2 + y_3 = \sum z_1 = -1,$$

$$y_1 y_2 = \sum z_1 + 2\cos 2a + 2\cos 3a = -1 + y_2,$$

$$y_2 y_3 = -1 + y_3, \quad y_3 y_1 = -1 + y_1.$$

Следователно

$$y_1 y_2 + y_2 y_3 + y_1 y_3 = -3 + y_1 + y_2 + y_3 = -4.$$

По-нататък имаме

$$y_1 y_2 y_3 = (-1 + y_2) y_3 = y_2 y_3 - y_3 = -1.$$

Следователно y_1, y_2, y_3 са корени на уравнението

$$y^3 + y^2 - 4y + 1 = 0.$$

След това лесно се получават трите квадратни уравнения, които дават корените от трите групи в (22). Така имаме

$$2\cos a + 2\cos 5a = y_1,$$

$$2\cos a \,.\, 2\cos 5a = 2\cos 4a + 2\cos 6a = y_3 = \frac{-1 + y_1}{y_1}.$$

Значи корените в първия ред на (22) се дават с уравнението

$$x^2 - y_1 x + \frac{y_1 - 1}{y_1} = 0.$$

Ако заместим тук y_1 с y_2 и y_3, получаваме корените от втория и третия ред.

4. Да разгледаме сега биномното уравнение

$$x^{17} = 1.$$

Ако α е негов примитивен корен, а r е примитивен корен на конгруенцията

$$x^{16} \equiv 1 \pmod{17},$$

то корените на уравнението

$$\frac{x^{17}-1}{x-1}=0 \tag{23}$$

ще бъдат

$$\alpha,\ \alpha^r,\ \alpha^{r^2},\ \ldots,\ \alpha^{r^{15}}. \tag{24}$$

Най-малкият примитивен корен r е равен на 3. Като редуцираме степените r^k на техните положителни и отрицателни остатъци по модула 17, получаваме за редицата (24)

$$\alpha,\ \alpha^3,\ \alpha^{-8},\ \alpha^{-7},\ \alpha^{-4},\ \alpha^5,\ \alpha^{-2},\ \alpha^{-6},\ \alpha^{-1},\ \alpha^{-3},\ \alpha^8,\ \alpha^7,\ \alpha^4,\ \alpha^{-5},\ \alpha^2,\ \alpha^6.$$

Понеже $16=2.8$, разделяме тези корени на две групи по 8, като в първата влизат четни степени на r, а във втората — нечетни, т. е.

$$\alpha,\ \alpha^{-8},\ \alpha^{-4},\ \alpha^{-2},\ \alpha^{-1},\ \alpha^8,\ \alpha^4,\ \alpha^2,$$

$$\alpha^3,\ \alpha^{-7},\ \alpha^5,\ \alpha^{-6},\ \alpha^{-3},\ \alpha^7,\ \alpha^{-5},\ \alpha^6.$$

Като поставим

$$y_1=\sum_{k=0}^{7}\alpha^{r^{2k}},$$

$$y_2=\sum_{k=0}^{7}\alpha^{r^{2k+1}},$$

имаме

$$y_1+y_2=-1.$$

Произведението $y_1 y_2$ се състои от 64 члена, които са степени на α. Ако в това произведение заместим α^r с α^{r^g}, то при g четно y_1 и y_2 не се променят, а при g нечетно y_1 става y_2, а последното y_1, така че $y_1\ y_2$ не се изменя. Тъй лесно се убеждаваме, че всеки корен (24) влиза еднакво число пъти, в което впрочем можем да се убедим директно, като умножим y_1 и y_2. Така получаваме

$$y_1 y_2=4\sum\alpha^{r^s}=-4,$$

отгдето следва, че y_1 и y_2 са корени на уравнението

$$y^2+y-4=0. \tag{26}$$

След това полагаме

$$z_1=\alpha+\alpha^{-4}+\alpha^{-1}+\alpha^4, \qquad t_1=\alpha^3+\alpha^5+\alpha^{-3}+\alpha^{-5},$$

$$z_2=\alpha^{-8}+\alpha^{-2}+\alpha^8+\alpha^2, \qquad t^2=\alpha^{-7}+\alpha^{-6}+\alpha^7+\alpha^6.$$

Лесно се вижда, че

$$z_1+z_2=y, \qquad t_1+t_2=y_2,$$

$$z_1 z_2=y_1+y_2=-1, \qquad t_1t_2=-1,$$

т. е. z_1, z_2 са корени на уравнението

(26′) $$z^2-y_1 z-1=0,$$

а t_1, t_2 са корени на уравнението

$$t^2-y_2 t-1=0.$$

Ако поставим

$$u_1=\alpha+\alpha^{-1}, \quad u_2=\alpha^{-4}+\alpha^4,$$

получаваме

$$u_1+u_2=z_1,$$

$$u_1 u_2=t_1,$$

т. е. u_1, u_2 са корени на уравнението

(27) $$u^2-z_1 u+t_1=0,$$

отгдето, понеже $\alpha.\alpha^{-1}=1$, $\alpha^{-4}.\alpha^4=1$, корените α, α^{-1} се дават с

(28) $$x^2-u_1 x+1=0,$$

а α^{-4}, α^4 с уравнението

$$x^2-u_2 x+1=0.$$

Аналогично имаме за останалите корени α^5, α^{-5} и т. н.

За уравнението (23) можем да приложим и другия начин на решение. За

$$z=x+\frac{1}{x}$$

получаваме уравнението

(29) $$z^8+z^7-7z^6-6z^5+15z^4+10z^3-10z^2-4z+1=0,$$

на което корените са

$$z_k=2\cos\frac{2k\pi}{17}, \; k=1, 2, \ldots, 8.$$

Ако r е примитивен корен на конгруенцията

$$x^{16}\equiv 1 \pmod{17},$$

то корените на (29) са

$$\left(a=\frac{2\pi}{7}\right),$$

$$2\cos a, \; 2\cos ra, \; 2\cos r^2 a; \ldots, \; 2\cos r^7 a;$$

$$2\cos r^8 a=2\cos(-a)=2\cos a$$

или ако вземем $r=3$, който е най-малкият примитивен корен на конгруенцията, като заместим степените на r с остатъците им спрямо модула 17, ще имаме

$$2\cos a, \quad 2\cos 3a, \; 2\cos 9a, \quad 2\cos 10a,$$

$$2\cos 13a,\ 2\cos 5a,\quad 2\cos 15a,\ 2\cos 11\,a$$

или

$$2\cos a,\ 2\cos 3a,\ 2\cos 8a,\ 2\cos 7a,\ 2\cos 4a,\ 2\cos 5a,\ 2\cos 2a,\quad 2\cos 6a.$$

Разделяме тези корени в две групи по четири:

$$\begin{matrix} 2\cos a,\quad 2\cos 8a,\ 2\cos 4a,\ 2\cos 2a, \\ 2\cos 3a,\ 2\cos 7a,\ 2\cos 5a,\ 2\cos 6a. \end{matrix} \tag{30}$$

С y_1 означаваме сумата на първите, а с y_2 сумата на вторите:

$$y_1 = 2\cos a + 2\cos 8a + 2\cos 4a + 2\cos 2a,$$

$$y_2 = 2\cos 3a + 2\cos 7a + 2\cos 5a + 5\cos 6a.$$

Тогава y_1+y_2 е равна на сумата на корените на (20), следователно равна на -1. След това лесно се получава, че $y_1 y_2$ е равна на 4 пъти сумата на корените, т. е. равно на -4. Числата y_1, y_2 са корени на уравнението

$$y^2+y-4=0. \tag{30'}$$

Лесно се получават уравненията

$$x^4+\alpha_1(y_1)\,x^3+\alpha_2(y_1)\,x^2+\alpha_3(y_1)\,x+\alpha_4(y_1)=0,$$

$$x^4+\alpha_1(y_2)\,x^3+\alpha_2(y_2)\,x^2+\alpha_3(y_2)\,x+\alpha_4(y_2)=0,$$

които дават корените на първата и втората група от (30). Това са пак абелеви уравнения, на които корените образуват един период. Тяхното решение обаче можем да извършим, без да сме ги получили.

Именно първата редица от (30) разделяме на две:

$$2\cos a,\quad 2\cos 4a,$$

$$2\cos 8a,\quad 2\cos 2a,$$

и полагаме $u_1 = 2\cos a + 2\cos 4a$, $u_2 = 2\cos 8a + 2\cos 2a$; имаме

$$u_1+u_2=y_1,$$

$u_1 u_2 =$ сумата на всички корени $= -1$. Следователно u_1, u_2 са корени на

$$u^2-y_1u-1=0. \tag{31}$$

Също втората редица (30) разлагаме на две:

$$2\cos 3a,\ 2\cos 5a,$$

$$2\cos 7a,\ 2\cos 6a.$$

Ако поставим

$$v_1 = 2\cos 3a + 2\cos 5a,\quad v_2 = 2\cos 7a + 2\cos 6a,$$

получаваме

$$v_1+v_2=y_2, \quad v_1v_2=-1,$$

т. е. v_1, v_2 са корени на уравнението

$$v^2-y_2v-1=0, \tag{32}$$

или понеже $y_1+y_2=-1$, на

$$v^2+(1+y_1)v-1=0.$$

Ако сега означим с

$$x_1=2\cos a, \quad x_2=2\cos 4a,$$

то имаме

$$x_1+x_2=u_1, \quad x_1x_2=2\cos 3a+2\cos 5a=v_1,$$

т. е. x_1, x_2 са корени на уравнението

$$x^2-u_1x+v_1=0. \tag{33}$$

Подобно се получава, че

$$x_3=2\cos 8a, \quad x_4=2\cos 2a$$

са корени на

$$x^2-u_2x+v_2=0,$$

по-нататък

$$x_5=2\cos 3a, \; x_6=2\cos 5a$$

са корени на

$$x^2-v_1x+u_2=0$$

и най-сетне

$$x_7=2\cos 7a, \; x_8=2\cos 6a$$

са корени на уравнението

$$x^2-v_2x+u_1=0.$$

Обаче за решението очевидно стига само да определим единия корен $2\cos a$, което става с решението на квадратните уравнения (30′), (31), (32), (33).

4. Решими с линийка и пергел конструктивни задачи. При геометрическите построения, за които ще става дума, ще се използуват линийката и пергелът, за да се получават от дадени точки нови търсени фигури. Линийката се използува да се съединяват дадени или намерени точки с прави. Пергелът се използува, за да начертаем окръжност с център, който е дадена или намерена точка, и радиус, равен на разстоянието между две дадени или намерени точки. Нови точки се получават като пресечни точки на така прекараните линии и окръжности. Тези операции трябва да бъдат извършени в краен брой.

Нека отнесем всичките точки към една правоъгълна координатна система. Тогава всички геометрични задачи за построения могат да се

формулират така: По даден краен брой точки $P_1, P_2, \ldots, P_m$ да се намерят краен брой точки $X_1, X_2, \ldots, X_n$, които във връзка с дадените да удовлетворяват на дадени геометрични условия с използуване на линийката и пергела. Координатите на дадените точки ще принадлежат на една област R на рационалност (най-малка такава), като координатната система е така избрана, че точката $(1,0)$ фигурира между точките P_i. Тогава, понеже правата линия, която съединява две точки P_1 и P_2 или построени две такива, се представя с уравнение от първа степен с коефициенти, които са рационални функции на координатите им, координатите на пресечната точка на две такива прави ще бъдат рационални функции от координатите на точките, които ги дефинират. Понеже уравнението на окръжността е от втора степен, то координатите на пресечните ѝ точки с права се дават с квадратно уравнение, както, в което лесно се убеждаваме, и координатите на пресечните точки с друга окръжност. Оттук става ясно, че координатите на точките, построени с пергел и линийки, се получават с решение на редица квадратни уравнения.

Решението на квадратно уравнение става с квадратни радикали, така че от това става ясно, че координатите на получените точки принадлежат на една реална област на рационалност, наречена квадратна област на рационалност, получена чрез адюнгиране на квадратни радикали към областта на рационалност, в която принадлежат координатите на дадените точки.

Обратно, ако елементите на една фигура или координатите на точки принадлежат на една реална квадратна област на рационалност, то задачата е ре́шима с пергел и линийка. Достатъчно е да докажем, че ако може да построим точките, на които координатите принадлежат на една област на рационалност R, ще можем да построим и точките, на които координатите са от друга област на рационалност R_1, получена от R с адюнгиране на радикал $\sqrt{A}$, $A>0$. Видяхме по-рано, че всяко число от R_1 има форма

$$M+N\sqrt{A},$$

гдето M, N, A са от R. Нека вземем една отсечка за единица. Но отсечката $\alpha=\sqrt{A}=\sqrt{A\,.\,1}$ като средна геометрична на A и 1 се построява лесно. Също $\alpha_1=N\alpha$, понеже

$$\frac{\alpha_1}{N}=\frac{\alpha}{1}$$

е построена пак с пергел и линийка, което читателят лесно ще изпълни. Ако означим отсечката 1 с буква, то изразите за търсените отсечки ще бъдат хомогенни функции от първа степен спрямо дадените отсечки.

Нека отбележим, че всяко тяло от числа или функции съдържа тялото P, състоящо се от всички рационални числа. Понеже даденото тяло ще има елемент $a\neq 0$ и оттам ще съдържа $\frac{a}{a}=1$, $1+1=2$,

$2+1=3, \ldots, 0-1=-1, \ -2, \ldots,$ и $\frac{p}{q}$, p и q цели, $q\neq 0$.

Сега ще установим една теорема, която ни дава необходимо условие, за да бъде едно уравнение решимо с квадратни радикали.

Ако едно неразложимо уравнение в една област на рационалност R има корен, принадлежащ на една квадратна област на рационалност, получена от R с адюнгиране на квадратни радикали, то степента му трябва да е равна на 2^{μ}.

Действително нека

$$f(x)=0 \tag{34}$$

е неразложимо уравнение в областта на рационалност R и нека x_1 е корен, който принадлежи на една квадратна област на рационалност R_k, получена от R с адюнгиране на квадратни радикали, като така последователно сме получили областите $R, R_1, R_2, \ldots, R_k$.

Полиномът $f(x)$ е разложим в R_k, понеже имаме

$$f(x)=(x-x_1)f_1(x),$$

гдето полиномите $x-x_1$, $f_1(x)$ имат коефициенти в R_k. Следователно ще има област R_p, $1\leqq p\leqq k$, в която $f(x)$ е разложим, а в R_{p-1} е неразложим ($R_0=R$). Ако $\varphi(x)$ е неразложимият в R_p множител, който съдържа x_1, то

$$f(x)=\varphi(x)\psi(x), \tag{35}$$

гдето очевидно $\psi(x)$ има коефициенти от R_p. Нека R_p се получава от R_{p-1} с адюнгиране на радикала $r=\sqrt{\alpha}$, гдето α принадлежи на R_{p-1}, но r не принадлежи на тази област. Но приравнявайки коефициентите в двете части на (35), ще получим линейни уравнения спрямо r, коефициентите на които трябва да бъдат равни на нула, защото иначе r би принадлежало на областта R_{p-1}. Следователно, ако поставим $-r$ вместо r, тъждеството (35) не ще се наруши, т. е. ще имаме

$$f(x)=\varphi_0(x)\psi_0(x), \tag{36}$$

гдето $\varphi_0(x)$ и $\psi_0(x)$ се получават съответно от $\varphi(x)$ и $\psi(x)$, като сменим r с $-r$. Но тогава, като умножим (35) с (36), получаваме

$$f^2(x)=\varphi(x)\varphi_0(x)\psi(x)\psi_0(x).$$

Полиномите $\varphi(x)\varphi_0(x)$ и $\psi(x)\psi_0(x)$ имат вече коефициенти, принадлежащи на R_{p-1}, в която област $f(x)$ е неразложим. Следователно горното тъждество е възможно само тогава, когато

$$f(x)=A\varphi(x)\varphi_0(x), \quad f(x)=B\psi(x)\psi_0(x), \quad AB=1,$$

гдето A и B са константи. Ако n е степента на $f(x)$, а m тази на $\varphi(x)$, то $n=2m$. Ако $m>1$, то като приложим абсолютно същите разсъждения за $\varphi(x)$, получаваме, че

$$m=2m_1, \text{ т. е. } m=2^2m_1,$$

и т. н., докато достигнем до $n=2^\mu$.

5. Деление на ъгъла на три равни части. От теоремата в предния параграф веднага следва предложението: за да бъде едно кубично уравнение с коефициенти, принадлежащи на една област на рационалност R, решимо с квадратни радикали, необходимо и достатъчно е да има корен, принадлежащ на R. Понеже, ако има такъв корен, то като го отстраним, получаваме квадратно уравнение. Обратно, ако е решимо с квадратни радикали, понеже степента му не е 2^μ, то трябва да бъде разложимо, т. е. да има един линеен множител, който дава рационален корен, принадлежащ на R.

Едно интересно приложение на това е в проблемата за трисекцията на ъгъла. Нека φ е даден ъгъл; тогава директно или по формулата на Моавър имаме

$$2\cos\varphi = \left(2\cos\frac{\varphi}{3}\right)^3 - 3\left(2\cos\frac{\varphi}{3}\right).$$

Ако поставим

$$2\cos\varphi=a,\ 2\cos\frac{\varphi}{3}=x,$$

за x ще имаме кубичното уравнение

$$x^3-3x-a=0.$$

Трисекцията на ъгъла φ е възможна с пергел и линийка, ако това уравнение има рационален корен в областта на рационалност на a. Ако например $\varphi=\frac{\pi}{3}$, то $a=1$ и уравнението е

$$x^2-3x-1=0.$$

Единствените рационални корени могат да бъдат 1 и —1, но със заместване се убеждаваме, че те не го удовлетворяват. Следователно ъгълът от 60^0 не може да се раздели на три равни части с пергел и линийка. При $\varphi=45^0$, $a=\sqrt{2}$, уравнението

$$x^3-2x-\sqrt{2}=0$$

има рационалния корен $x=-\sqrt{2}$ в областта $R(\sqrt{2})$.

Трисекцията е възможна с пергел и линийка. Подобно се вижда лесно, че не може да се построи с пергел и линийка страната на куба, на който обемът е два пъти по-голям от обема на даден куб.

6. Построяване на правилни многоъгълници. Нека в равнината на x опишем една окръжност с радиус 1 около началото. Ако нанесем корените

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

на уравнението

$$x^n = 1, \tag{37}$$

ще получим върховете на един правилен многоъгълник: именно тези точки делят прекараната окръжност на n равни части. Единият връх е точката 1, следващият е точката

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

Другите ще се получат с нанасяне на тази дъга $\overset{\frown}{1\alpha}$ с пергела. От предния параграф се вижда, че въпросът за построението на правилния n-ъгълник с пергел и линийка се свежда към въпроса за решимост с квадратни радикали на уравнението (37).

Отначало да разгледаме случая, когато $n=p$ е просто число. Тогава видяхме, че уравнението

$$\frac{x^p-1}{x-1} = x^{p-1}+\ldots+x+1=0$$

е неразложимо и ако $p=2^\lambda+1$, то е решимо с квадратни радикали. Обратно, ако то е решимо с квадратни радикали, трябва p да е равно на $2^\lambda+1$. Следователно правилният p-ъгълник, където p е просто, е само тогава построим с пергел и линийка, когато p има форма $2^\lambda+1$.

За да бъде числото $p=2^\lambda+1$ просто, трябва λ да няма делител, който е нечетно число. Защото, ако

$$\lambda = h(2m+1),$$

то $p=(2^h)^{2m+1}+1$ се дели на 2^h+1, т. е. няма да бъде просто. Трябва λ да има форма 2^k, т. е.

$$p = 2^{2^k}+1.$$

За $k=0$, $p=3$; за $k=1$, $p=5$, $k=2$, $p=17$. След това за $k=3$, $p=257$, но при $k=5$, $p=4294967297$, едно число, което се дели на 641, следователно не е просто. Не е известно колко числа има от тази форма.

Ако сега $n=p^k$, то видяхме, че уравнението

$$\frac{x^{p^k}-1}{x^{p^{k-1}}-1} = 0,$$

което дава примитивните корени, е неразложимо. Степента му е

$$p^k - p^{k-1} = p^{k-1}(p-1)$$

и ако $k \geq 1$, $p \neq 2$, то тя няма форма 2^{μ}, следователно уравнението е неразрешимо с квадратни радикали, в което се убеждаваме непосредствено. Значи в този случай ($n = p^k$, $k > 1$) само тогава правилният n-ъгълник е построим с линийка и пергел, когато $p = 2$.

Ако сме построили правилния a-ъгълник и правилния b-ъгълник и ако a и b са взаимно прости, то лесно построяваме и правилния ab-ъгълник. Действително корените γ на уравнението

$$x^{ab} = 1$$

се получават, като умножим корените α на

$$x^a = 1$$

с корените β на $x^b = 1$. А построяването на точката $\gamma = \alpha\beta$ става по точките α и β лесно с пергел и линийка.

Обратно, ако сме построили правилния n-ъгълник и $n = ab$, то, като съединим върховете през b, ще получим върховете на правилния a-ъгълник. По същия начин със съединяване през a получаваме правилния b-ъгълник. Ако

$$n = 2^{\lambda} p_1 p_2 \ldots p_{\varrho},$$

където $p_r, \ldots, p_{\varrho}$ са различни прости числа от вида $2^k + 1$, то на основание на горното следва, че правилният n-ъгълник е построим с пергел и линийка. Ако едно от простите числа p няма форма $2^k + 1$ или n съдържа множител p^{μ}, $\mu > 1$, $p > 2$, то n-ъгълникът не е построим с пергел и линийка. Така достигаме до общата теорема на Гаус: Правилният многоъгълник е само тогава построим с пергел и линийка, когато броят на страните n има формата

$$n = 2^{\lambda} p_1 p_2 \ldots p_{\varrho},$$

където

$$p_1, p_2, \ldots, p_{\varrho}$$

са различни прости числа от форма $2^d + 1$.

7. Квадратура на кръга. Тази класична задача се състои в следното: Да се построи квадрат, лицето на който е равно на това на даден кръг. Ако означим с x страната на квадрата и вземем за простота радиуса на кръга, равен на 1, то $x = \sqrt{\pi}$. Ако задачата е възможна с пергел и линийка, трябва π да е число от една квадратна област на рационалност R_2, получена от натуралната област P с адюнгиране на квадратни радикали. Но тогава по глава IV, § 3, трябва π да е корен на едно уравнение с цели рационални коефициенти. Обаче в 1882 г. Линдеман доказа, че π не удовлетворява на никое такова уравнение (доказателството на този факт е изложено в сборника ми по висша алгебра). Следователно квадратурата на кръга в горния смисъл е невъзможна.

Глава III

Алгебрична нерешимост на уравненията от степен, по-висока от четири

1. Увод. Както вече видяхме, уравненията от трета и четвърта степен са решими алгебрически, т. е. можем да намерим един алгебрически израз на коефициентите, който, заместен вместо неизвестното, обръща уравнението в тъждество. Явява се въпросът за алгебрическо решение на уравненията от по-високи степени. Някои специални уравнения от степен, по-голяма от четири, са решими алгебрически, каквито са например биномните и абелеви циклични уравнения. Обаче, както ще видим, уравненията от общ вид, т. е. коефициентите на които са произволни параметри от степен, по-висока от четири, са нерешими алгебрически. Тази теорема е била доказана за пръв път от италианския математик Паоло Руфини[1] в съчинението Teoria generale delle Equazioni, in cui si dimonstra impossibile la soluzione algebraica delle equazioni generali di grado superiore al quarto, Bologna, 1799 г. Доказателството на Руфини не било оценено добре от съвременниците му и останало в неизвестност.

След това норвежкият математик N. Abel, един от основателите на модерната математика, дава отново в 1826 г. строго доказателство (Beweis der Unmöglichkeit, algebraische Gleichungen von höheren Graden als dem vierten allgemein aufzulösen. Crelles Journ. Bd. I, S. 65).

Ще изложим доказателството на Абел, опростено от Ванцел.

2. Обща форма на алгебрична функция. Под алгебрична функция в тесен смисъл на думата ще разбираме такава, в която променливите са подложени само на алгебричните действия, в краен брой, събиране, изваждане, умножение, деление, степенуване и коренуване с цял показател.

Нека $x_1, x_2, \ldots, x_n$ да са променливите. Както вече видяхме, съвкупността на рационалните функции на тези променливи с коефициенти от дадена област на рационалност образува една област на рационалност. Към тази област, ако прибавим един радикал от една функция $\sqrt[p]{f}$, получаваме една разширена област на рационалност, като предполагаме, че f не е p-та степен на рационална функция. Очевидно можем да считаме показателя p на корена за просто число, защото иначе адюнгирането на радикала би станало последователно с адюнгиране на радикали с прости показатели.

Нека областта на рационалните функции да означим с R_0. С адюнгиране на радикали от вида

$$\sqrt[p]{v_0},$$

гдето v_0 са функции от R_0, получаваме една нова област на рационалност R_1. С адюнгиране на радикали от същия вид

$$\sqrt[p]{v_1},$$

---

[1] Роден в 1765 г., умрял в 1822 г.; Руфини по професия бил лекар.

гдето v_1 е функция от R_1, получаваме нова област на рационалност R_2 и т. н.; въобще областта R_μ се получава от $R_{\mu-1}$ с адюнгиране на радикали

$$\sqrt[s]{v_{\mu-1}}, \tag{1}$$

гдето $v_{\mu-1}$ принадлежи на $R_{\mu-1}$. Всяка функция от R_μ ще наричаме алгебрична функция от ред μ и ако броят на радикалите от вида (1) в нея е m, то тя ще бъде от степен m. Следователно общата форма на алгебрична функция от ред μ и степен m е

$$u=f(\sqrt[p]{v}, v_1, v_2, \ldots),$$

гдето f е рационална функция, p — просто число, v — функция от ред $\mu-1$, $v_1, v_2, \ldots$ са функции от ред μ и степен $m-1$.

Теорема. Всяка алгебрична функция от ред μ и степен m може да се представи във формата

$$u=u_0+u_1v^{\frac{1}{p}}+u_2v^{\frac{2}{p}}+\ldots+u_{p-1}v^{\frac{p-1}{p}},$$

гдето p е просто число, $u_0, u_1, \ldots, u_{p-1}$ са от ред μ и степен $m-1$, v е от ред $\mu-1$, $\sqrt[p]{v}$ не е рационална функция на $u_0, u_1, \ldots, u_{p-1}$; може да се приеме $u_1=1$.

Понеже всяка рационална функция на няколко променливи може да се представи като отношение на две цели рационални функции на същите променливи, то u може да се пише

$$\frac{\varphi_0+\varphi_1v^{\frac{1}{p}}+\varphi_2v^{\frac{2}{p}}+\ldots}{\psi_0+\psi_1v^{\frac{1}{p}}+\psi_2v^{\frac{2}{p}}+\ldots},$$

гдето φ_i, ψ_i са цели рационални функции на $v_1, v_2, v_3, \ldots$

Като означим

$$\varphi(x)=\varphi_0+\varphi_1x+\varphi_2x^2+\ldots,$$
$$\psi(x)=\psi_0+\psi_1x+\psi_2x^2+\ldots,$$

ще имаме

$$u=\frac{\varphi(v^{\frac{1}{p}})}{\psi(v^{\frac{1}{p}})}=\frac{\varphi(v^{\frac{1}{p}})\,\psi(\omega v^{\frac{1}{p}})\ldots\psi(\omega^{p-1}v^{\frac{1}{p}})}{\psi(v^{\frac{1}{p}})\,\psi(\omega v^{\frac{1}{p}})\ldots\psi(\omega^{p-1}v^{\frac{1}{p}})},$$

гдето ω е примитивен корен на $x^p=1$. Очевидно

$$v^{\frac{1}{p}},\ \omega v^{\frac{1}{p}},\ \ldots,\ \omega^{p-1}v^{\frac{1}{p}}$$

са всичките корени на уравнението

$$x^p=v. \tag{2}$$

Изразът

$$T=\psi(v^{\frac{1}{p}})\,\psi(\omega v^{\frac{1}{p}})\ldots\psi(\omega^{p-1}v^{\frac{1}{p}})$$

е симетрична функция на корените на уравнението (2) и следователно е рационална функция на коефициентите му, т. е. на v. Така виждаме, че в T няма да фигурира $v^{\frac{1}{p}}$.

Изразът

$$T'=\psi(\omega v^{\frac{1}{p}})\,\psi(\omega^2 v^{\frac{1}{p}})\ldots\psi(\omega^{p-1}v^{\frac{1}{p}})$$

представя цяла рационална симетрична функция на корените на уравнението

$$\frac{x^p-v}{x-v^{\frac{1}{p}}}=0,$$

коефициентите на което са цели рационални функции на $v^{\frac{1}{p}}$. Следователно T' ще бъде цяла рационална функция на $v^{\frac{1}{p}}$:

$$\varphi(v^{\frac{1}{p}})\,T'=w_0+w_1v^{\frac{1}{p}}+w_2v^{\frac{2}{p}}+\ldots+w_kv^{\frac{k}{p}},$$

гдето $w_0, w_1, \ldots, w_k$ са цели рационални функции на

$$\varphi_0, \varphi_1, \varphi_2, \ldots, v,$$

т. е. на $v_0, v_1, v_2, v_3, \ldots$ Понеже

$$v^{\frac{p+1}{p}}=vv^{\frac{1}{p}},\ v^{\frac{p+2}{p}}=vv^{\frac{2}{p}},\ldots,$$

то можем да приемем, че $k\leqq p-1$. Ако тогава поставим

$$\frac{w_0}{T}=u_0,\ \frac{w_1}{T}=u_1,\ \frac{w_2}{T}=u_2,\ldots,$$

то ще имаме

$$u=u_0+u_1v^{\frac{1}{p}}+u_2v^{\frac{2}{p}}+\ldots+u_{p-1}v^{\frac{p-1}{p}},$$

гдето $u_0, u_1, \ldots u_{p-1}$ като рационални функции на $v, v_1, v_2, \ldots$ са алгебрични функции от ред μ и степен $m-1$, а v е от ред $\mu-1$.

Сега ще установим, че на u може да се даде такава форма, че u_1 да бъде равно на 1. Действително, ако $u_1\neq 0$, полагаме

$$u_1v^{\frac{1}{p}}=g^{\frac{1}{p}},\ g=u_1^pv$$

и u става

$$u=g_0+g^{\frac{1}{p}}+g_2g^{\frac{2}{p}}+\ldots+g_{p-1}g^{\frac{p-1}{p}},$$

гдето $g_0, g_2, \ldots, g_{p-1}$ като равни съответно на $\frac{u_2}{u_1^2}, \ldots, \frac{u_{p-1}}{u_1^{p-1}}$ са функции от ред μ и степен $m-1$, а също g е от ред μ и степен $m-1$. Ако $u_1=0$, нека за общност

$$u_1=0,\ u_2=0,\ \ldots,\ u_{q-1}=0,\ u_q\neq 0.$$

Полагаме

$$u_q\sqrt[q]{v^p}=\sqrt[1]{w^p},\ w=u_q^p v^q;$$

$\underline{s}$

тогава всяка степен v^p, $s>q$, ще може да се изрази посредством w. Понеже q и p са взаимно прости, неопределеното уравнение

$$s=qy-px$$

има цяло положително решение x, y. Но тогава

$$\sqrt[s]{v^p}=\sqrt[q]{v^p}^y \qquad v^{-x}=\frac{v^{-x}}{u_q^y}\sqrt[y]{w^p}$$

и понеже, както лесно се вижда, $y>1$, то на u може да се даде формата

$$u=w_0+\sqrt[1]{w^p}+w_2\sqrt[2]{w^p}+\ldots+w_{p-1}\sqrt[p-1]{w^p},$$

гдето $w, w_0, w_1, \ldots, w_{p-1}$ са от ред μ и степен $m-1$.

3. **Доказване на теоремата за алгебрична нерешимост.** Отначало ще установим една помощна теорема.

Нека p е просто число и $v, v_0, v_1, \ldots$ да принадлежат на една област на рационалност, но $\sqrt[p]{v}$ да не принадлежи на тази област. Тогава, ако имаме уравнението

$$v_0+v_1\sqrt[1]{v^p}+v_2\sqrt[2]{v^p}+\ldots+v_{p-1}\sqrt[p-1]{v^p}=0,$$

то трябва да имаме

$$v_0=0,\ v_1=0,\ \ldots,\ v_{p-1}=0.$$

От условието на теоремата следва, че $\sqrt[p]{v}$ е общ корен на уравненията

$$x^p-v=0,\ v_0+v_1x+v_2x^2+\ldots+v_{p-1}x^{p-1}=0.$$

Ако не всички коефициенти $v_0, v_1, \ldots$ са нули, то левите им части ще имат един общ най-голям делител, на който коефициентите като рационални функции на v_i, v ще бъдат в същата област на рационалност. Нека лявата част на уравнението

$$r_0+r_1x+\ldots+r_kx^k=0 \tag{3}$$

да бъде един неразложим делител на общия най-голям делител и което уравнение съдържа корена $\sqrt[p]{v}$. Степента на (3) е най-малко две, понеже ако $k=1$, то $\sqrt[p]{v}$ би принадлежало в областта на рационалност. Но тогава (3) има поне още един общ корен $\omega\sqrt[p]{v}$ с уравнението

$$x^p - v = 0, \tag{4}$$

гдето $\omega^p = 1$. Следователно уравнението

$$r_0 + r_1\omega x + \ldots + r_k\omega^k x^k = 0$$

има корен $x = v^{\frac{1}{p}}$. Или като го извадим от (3), следва, че уравнението от степен $k-1$

$$r_1(1-\omega) + \ldots + r_k(1-\omega^k)x^{k-1} = 0$$

има общ корен $\sqrt[p]{v}$ с неразложимото уравнение (3), което е невъзможно. При това считаме, че към всяка област на рационалност са адюнгирани корените на биномните уравнения, които, както знаем, са решими алгебрически.

Доказателството може да се извърши също по друг начин, като се установи, че (4) е неразложимо уравнение.

Теорема. Алгебрическата функция от коефициентите, която дава общия израз на корена на едно реши мо с радикали уравнение, се изразява рационално посредством корените му.

На основание на по-раншните теореми на алгебрическата функция, която удовлетворява едно дадено уравнение

$$a_0x^n - a_1x^{n-1} + \ldots + a_n = 0, \tag{5}$$

може да се даде формата

$$x = u_0 + v^{\frac{1}{p}} + u_2v^{\frac{2}{p}} + \ldots + u_{p-1}v^{\frac{p-1}{p}}, \tag{6}$$

гдето u_i, v принадлежат на една област на рационалност R, получена от областта на коефициентите $a_0, \ldots, a_n$ с адюнгиране на радикали. $\sqrt[p]{v}$ не се изразява рационално посредством u_i. Ако заместим израза (6) за x в уравнението, то получаваме

$$v_0 + v_1v^{\frac{1}{p}} + v_2v^{\frac{2}{p}} + \ldots + v_{p-1}v^{\frac{p-1}{p}} = 0,$$

гдето $v_0, v_1, \ldots, v_{p-1}$ принадлежат на R. Но тогава според помощната теорема трябва

$$v_0 = 0,\ v_1 = 0, \ldots,\ v_{p-1} = 0.$$

Оттук обаче следва, че уравнението (5) се удовлетворява от числата x, получени от (6), като на $\sqrt[p]{v}$ дадем всички негови p значения

$$v^{\frac{1}{p}}, \omega v^{\frac{1}{p}}, \omega^2 v^{\frac{1}{p}}, \ldots, \omega^{p-1} v^{\frac{1}{p}},$$

гдето ω е примитивен корен на $x^p = 1$.

Следователно имаме

$$\begin{aligned}
x_1 &= u_0 + v^{\frac{1}{p}} + u_2 v^{\frac{2}{p}} + \ldots + u_{p-1} v^{\frac{p-1}{p}}, \\
x_2 &= u_0 + \omega v^{\frac{1}{p}} + u_2 \omega^2 v^{\frac{2}{p}} + \ldots + u_{p-1} \omega^{p-1} v^{\frac{p-1}{p}}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_p &= u_0 + \omega^{p-1} v^{\frac{1}{p}} + u_2 \omega^{2p-2} v^{\frac{2}{p}} + \ldots + u_{p-1} \omega^{(p-1)2} v^{\frac{p-1}{p}},
\end{aligned} \tag{7}$$

гдето $x_1, x_2, \ldots, x_p$ са корени на даденото уравнение (5) и ω е примитивен корен на $x^p = 1$.

Ако умножим последователно тези равенства с

$$1, \omega^{-1}, \omega^{-2}, \ldots, \omega^{-p+1}$$

и съберем, ще получим

$$\sqrt[p]{v} = \frac{1}{p}(x_1 + \omega^{-1} x_2 + \omega^{-2} x_3 + \ldots + \omega^{-p+1} x_p). \tag{8}$$

От (7) по аналогичен начин получаваме

$$u_r = p^{r-1} \frac{x_1 + \omega^{p-r} x_2 + \omega^{2p-2r} x_3 + \ldots}{(x_1 + \omega^{-1} x_2 + \omega^{-2} x_3 + \ldots)^r} \quad (r = 0, 2, \ldots, p-1). \tag{9}$$

От (8) и (9) се вижда, че $\sqrt[p]{v}$, u_0, $u_2, \ldots, u_{p-1}$ са рационални функции на корените на (5). Следователно, ако y е коя да е от тези функции може да се пише

$$y = f(x_1, x_2, x_3 \ldots),$$

гдето f е една рационална функция. Нека $y_1, y_2, \ldots, y_m$ са значенията на f, когато разместваме корените $x_1, x_2, \ldots, x_n$ по всевъзможни начини. Тогава, както видяхме при трансформацията на уравненията, $y_1, y_2, \ldots, y_m$ са корени на едно уравнение $\varphi(\nu) = 0$, коефициентите на което се изразяват рационално посредством коефициентите

$$a_0, a_1, \ldots, a_n$$

на (5). Но имаме

$$y = v_0 + w^{\frac{1}{q}} + v_2 w^{\frac{2}{q}} + \ldots + v_{q-1} w^{\frac{q-1}{q}},$$

гдето $v_0, v_2, \ldots, v_{q-1}, w$ са алгебрични функции от степен с единица по-малко. Понеже y е корен на

$$\varphi(y) = 0,$$

то по горното следва, че v_i, $\sqrt[q]{v}$ са рационални функции на корените му y_1, y_2, . . . и следователно ще бъдат рационални функции на x_1, x_2, . . .

Така продължавайки, получаваме очевидно, че всеки радикал със заместване на коефициентите с изразите им посредством корените се обръща в рационална функция на последната, с което теоремата е установена.

Сега лесно ще установим теоремата на Абел, че всяко общо уравнение от степен, по-висока от четири, е неразрешимо алгебрически.

Да допуснем, че уравнението е решимо алгебрически и нека $\sqrt[p]{u}$ е един от първите радикали, които се срещат в израза на корените, гдето u е рационална функция от коефициентите. По доказаната теорема $\sqrt[p]{u}$ е рационална функция на корените на уравнението (5) т. е. имаме

$$\sqrt[p]{u} = \varphi(x_1, x_2, x_3, \ldots),$$
$$u = \varphi^p(x_1, x_2, x_3 \ldots).$$

Понеже u е симетрична функция на x_1, x_2, . . ., то като направим една произволна транспозиция, например $(x_1\ x_2)$, ще имаме

$$u = \varphi^p(x_2, x_1, x_3, \ldots),$$

отгдето

(10) $$\varphi(x_2, x_1, x_3, \ldots) = \omega\varphi(x_1, x_2, x_3 \ldots),$$

гдето ω е корен на $x^p = 1$. Очевидно $\omega \neq 1$, понеже иначе φ ще бъде симетрична функция на корените x_1, x_2, . . ., т. е. ще се изразява рационално посредством коефициентите му и $\sqrt[p]{u}$ ще бъде привидно ирационалност. Ако тогава в (10) направим транспозицията $(x_1\ x_2)$, ще получим

$$\varphi(x_1, x_2, x_3, \ldots) = \omega\varphi(x_2, x_1, x_3, \ldots)$$

и като умножим (10) и (11), получаваме

$$\omega^2 = 1, \quad \omega = -1,$$

което е възможно само тогава, когато $p = 2$. Първият радикал, който се среща, е квадратен. Функцията φ е алтернативна. Комбинирайки сега φ с рационални функции от коефициентите и алтернативни, получаваме област на рационалност от първи ред, в която елементите са двузначни рационални функции на корените на уравнението, т. е. такива, които се изменят при една транспозиция и имат само две стойности. Но, както видяхме по-рано, те не се изменят, ако приложим една циклична субституция от три елемента.

Да допуснем, че степента n на (5) е по-голяма от 2. Нека $\sqrt[q]{v}$ да е нов радикал, който се среща в израза за корените, гдето v ще бъде

една двузначна функция на корените. По доказаната теорема $\sqrt[q]{v}$ е рационална функция на корените $x_1, x_2, \ldots$, т. е.

$$\sqrt[q]{v} = \psi(x_1, x_2, x_3, x_4, \ldots),$$
$$v = \psi^q(x_1, x_2, x_3, x_4, \ldots).$$

Ако приложим субституцията $(x_1\ x_2\ x_3)$, то според казаното v не се променя:

$$v = \psi^q(x_2, x_3, x_1, x_4, \ldots),$$

отгдето

$$\psi(x_2, x_3, x_1, x_4, \ldots) = \omega\psi(x_1, x_2, x_3, x_4, \ldots), \tag{12}$$

гдето $\omega \neq 1$ е корен на $x^q = 1$. Ако в (12) направим два пъти кръговата субституция $(x_1\ x_2\ x_3)$, получаваме

$$\begin{aligned}\psi(x_3, x_1, x_2, x_4, \ldots) &= \omega\psi(x_2, x_3, x_1, x_4, \ldots),\\ \psi(x_1, x_2, x_3, x_4, \ldots) &= \omega\psi(x_3, x_1, x_2, x_4, \ldots).\end{aligned} \tag{13}$$

Като умножим (12) и (13), получаваме

$$\omega^3 = 1,$$

отгдето, понеже $\omega^q = 1$ и q е просто число, $q = 3$. Вторият радикал трябва да бъде кубичен.

Да допуснем сега, че $n \geqq 5$. Лесно е да се убедим, че ако в двузначната функция v извършим една циклична субституция $(x_1x_2x_3x_4x_5)$ от пет елемента, то v не се променя. Действително субституцията $(x_1x_2x_3x_4x_5)$ е еквивалентна на четири транспозиции:

$$(x_1\ x_2)\,(x_1\ x_3)\,(x_1\ x_4)\,(x_1\ x_5),$$

с които v се изменя четири пъти, т. е. връща се към първоначалната си стойност. Но тогава, извършвайки във

$$v = \psi^q(x_1, x_2, x_3, x_4, x_5, \ldots)$$

субституцията $(x_1\ x_2\ x_3\ x_4\ x_5)$, получаваме

$$\psi(x_2, x_3, x_4, x_5, x_1, \ldots) = \alpha\psi(x_1, x_2, x_3, x_4, x_5, \ldots),$$

гдето α е корен на $x^q = 1$. Ако в това уравнение извършим четири пъти същата субституция и умножим, ще получим $\alpha^5 = 1$, отгдето $q = 5$, което противоречи с $q = 3$. С това теоремата на Абел е установена.

Глава IV

Неразложимият случай при кубичното уравнение

Както видяхме при алгебричното решение на уравнението от трета степен

$$x^3+px+q=0,$$

в случая, когато уравнението има само реални корени, формулата на Кардано не ни дава възможност да намерим тези корени чрез реални радикали от коефициентите му. Това не е една слабост на формулата на Кардано, но едно съществено свойство на алгебричната решимост на уравненията. Ще установим именно следната теорема:

Нека уравнението от n-та степен $f(x)=0$ има само реални корени и е неразложимо в едно поле P от реални числа, на което принадлежат коефициентите му. Нека освен това степента му n да се дели на някое нечетно просто число p. Тогава корените на това уравнение не могат да се изразят чрез реални радикали.

Предварително ще установим едно предложение, което ще използуваме по-нататък.

Ако биномното уравнение $x^p-a=0$ е разложимо в някое поле P, като a принадлежи на P и p е просто число, то a е точна p-та степен на някой елемент от полето P.

Действително нека

$$x^p-a=\varphi(x)\,\psi(x),$$

като $\varphi(x)$ и $\psi(x)$ са полиноми от полето P, т. е. коефициентите им принадлежат на P. Да означим с $\alpha_1, \alpha_2, \ldots, \alpha_p$ корените на $x^p=1$ и с θ един кой да е от корените на уравнението $x^p-a=0$. Корените на последното уравнение, както знаем, ще бъдат

$$\alpha_1\theta,\ \alpha_2\theta,\ \ldots,\ \alpha_p\theta. \tag{1}$$

Корените на уравнението $\varphi(x)=0$ ще бъдат измежду числата (1). Ако означим с $(-1)^s b_s$ свободния член в полинома $\varphi(x)$, като коефициентът пред x^s в този полином е равен на 1, ще имаме

$$b_s=\alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_s}\,\theta^s.$$

Оттук получаваме

$$b_s^p=\alpha_{i_1}^p\,\alpha_{i_2}^p\ldots\alpha_{i_s}^p\,\theta^{sp}=a^s. \tag{2}$$

Но числото s е взаимно просто с p. Известно е тогава, че съществуват цели числа μ и ν, за които имаме

$$\mu s+\nu p=1.$$

Но тогава от (2) получаваме

$$b_s^{\mu p} = a^{s\mu} = a^{1-\nu p},$$

$$a = \left(\frac{b_s^{\mu}}{a^{\nu}}\right)^p.$$

Понеже a^{ν} и b_s^{μ} принадлежат на полето P, то и $\alpha = \frac{b_s^{\mu}}{a^{\nu}}$ принадлежи на същото поле. Равенството $a = \alpha^p$ показва, че a е точна p-та степен на елемента α, принадлежащ на полето P, с което предложението е установено.

Ще пристъпим сега към доказване на теоремата. Предполагаме, че уравнението $f(x) = 0$ е решимо с реални радикали, т. е. корените му са елементи на едно поле Σ, получено от P с последователно присъединяване на реални радикали, на които показателите, както в предната глава, можем да предполагаме, че са прости числа. Именно разширяваме полето P с присъединяване на радикал $\sqrt[p_1]{a_0} = g_1$, където a_0 е елемент от P и не е точна p_1-та степен на елемент от същото поле и от възможните p_1-ти корени на a_0 сме взели реалната стойност на този радикал. Към така полученото поле $P(g_1)$ присъединяваме радикал $\sqrt[p_2]{a_1} = g_2$, като a_1 е елемент от това поле, който не е точна p_2-та степен, и получаваме разширеното поле $P(g_1, g_2)$ и т. н. и най-после с присъединяване на радикала $\sqrt[p_m]{a_{m-1}} = g_m$, където a_{m-1} е елемент от полето $P(g_1, g_2, \ldots, g_{m-1})$, получаваме разширеното ново поле $\Sigma = P(g_1, g_2, \ldots, g_{m-1}, g_m)$, в което лежат корените $x_1, x_2, \ldots, x_n$ на даденото уравнение $f(x) = 0$. Уравнението

$$x^{p_k} - a_{k-1} = 0$$

е неразложимо в полето $P(g_1, g_2, \ldots, g_{n-1})$. Действително, ако предположим, че това уравнение е разложимо от доказаното по-горе предложение, следва, че a_{k-1} ще бъде точна p_k-та степен на елемент от полето $P(g_1, g_2, \ldots, g_{k-1})$, което противоречи. Съгласно с теоремите за крайното алгебрично разширение (глава II) за степента $(\Sigma : P)$ на полето Σ спрямо P ще имаме

$$(\Sigma : P) = p_1 p_2 \ldots p_m.$$

Ако с l означим броя на числата от $p_1, p_2, \ldots p_m$, които са равни на 2, и с $q_1, q_2, \ldots, q_r$ останалите числа p_i, то ще имаме

$$(\Sigma : P) = 2^l q_1 q_2 \ldots q_r. \tag{3}$$

Да означим с Ω полето на разлагане на полинома $f(x)$, т. е. полето, получено от P с присъединяване на всичките корени на уравнението $f(x) = 0$. На същото основание, както по-горе, заключаваме, че степента $(\Omega : P) = h$ се дели на степента n на уравнението $f(x) = 0$. Съгласно с предположението полето Ω ще бъде подполе на полето Σ.

Да присъединим сега радикалите към полето Ω, като от полето $\Omega_{k-1}=\Omega\ (g_1, g_2, \ldots, g_{k-1})$ с присъединяване на радикала $g_k=\sqrt[p_k]{a_{k-1}}$ се получава полето $\Omega_k=\Omega(g_1, g_2, \ldots, g_k)$. Ако $p_k=2$ и уравнението $x^2-a_{k-1}=0$ е неразложимо, то така получаваме разширение от втора степен. Ако предното уравнение е разложимо по предложението, a_{k-1} е квадрат на елемент от полето Ω_{k-1} и фактически не получаваме никакво разширение. Нека сега показателят p_k е отличен от 2. Ако уравнението $x^{p_k}-a_{k-1}=0$ е разложимо, то в полето Ω_{k-1} по предложението a_{k-1} е точна p_k-та степен на елемент от това поле и следователно радикалът $g_k=\sqrt[p_k]{a_{k-1}}$ трябва да лежи в същото поле. Ако с $P(g_1, g_2, \ldots, g_k)=P_k$ означим полето, получено от разширение на P с присъединяване на радикалите $g_1, g_2, \ldots, g_k$, то очевидно полето $\Omega(g_1, g_2, \ldots, g_k)=\Omega_k$ се получава от полето P_k с присъединяване на корените $x_1, x_2, \ldots, x_n$ на уравнението $f(x)=0$, т. е. Ω_k представлява полето на разлагане на полинома $f(x)$ спрямо полето P_k. Но тогава съгласно с теоремата от глава II, част VII в полето Ω_{k-1} трябва да лежат всичките корени на уравнението $x^{p_k}-a_{k-1}=0$. Но при $p_k>2$ това уравнение има само един реален корен и следователно полето Ω_{k-1} трябва да съдържа имагинерни числа, което противоречи. Така доказахме, че уравнението $x^{p_k}-a_{k-1}=0$ е неразложимо в полето Ω_{k-1} и следователно полето Ω_k представлява действително крайно разширение на полето Ω_{k-1} от степен p_k. От предните разглеждания следва, че за степента $(\Sigma:\Omega)$ на полето $\Sigma=\Omega_m$ относно Ω ще имаме

$$(\Sigma:\Omega)=2^{l'}\ q_1\ q_2 \ldots q_\nu,\ l'\leqq l.$$

Оттук, като вземем пред вид, че $(\Sigma:P)=(\Sigma:\Omega)\ (\Omega:P)$, получаваме

$$(\Sigma:P)=2^{l'}\,.\,q_1\ q_2 \ldots q_\nu\, h. \tag{4}$$

От (3) и (4) следва, че $h=2^{l-l'}$. Но това равенство е невъзможно, понеже числото h се дели на n и следователно и на простото нечетно число p. Така теоремата е установена напълно. Първо доказателство за невъзможността на алгебрическата решимост с реални радикали на уравнението от трета степен, притежаващо само реални корени, е дадено от О. Холдер.

ЧАСТ IX

ПРИЛОЖЕНИЕ НА ТЕОРИЯТА НА ГРУПИТЕ ЗА АЛГЕБРИЧНОТО РЕШАВАНЕ НА УРАВНЕНИЯТА

Глава I

Субституции

1. Основни свойства. Както видяхме при метода на Лагранж, алгебричното решение на уравненията от трета и четвърта степен е възможно благодарение на съществуването на рационална функция от корените на уравнението, която при всевъзможните разместявания на променливите взема на брой по-малко различни стойности от степента му. Обаче същият метод (на Лагранж) е неприложим за уравнения от пета степен, понеже не съществува такава функция от пет променливи, която взема стойности на брой, по-малък от пет, освен едно-значните (симетрични) функции и двузначните функции. Въпросът за възможността на алгебричните уравнения от обща форма за степен, по-голяма от 4, бе решен отрицателно с теоремата на Абел. Съществуват обаче специални уравнения от произволно високи степени като биномните и цикличните абелеви, които са решими алгебрически. Следователно явява се за разрешение проблемата за намиране на условията, при които едно дадено алгебрично уравнение е решимо алгебрически или нерешимо. Решението на този въпрос е дадено от Галоа и ние ще видим, че алгебричната решимост на едно уравнение зависи изключително от структурата на група от субституции, която му съответствува.

Нека имаме n елемента, които ще номерираме с $1, 2, 3, \ldots, n$. Ако всеки елемент заместим с елемент от дадените n така, че всеки два различни елемента се заместват с два различни, то това действие се нарича с у б с т и т у ц и я. С други думи, ако $a_1, a_2, \ldots, a_n$ означава една произволна пермутация на елементите $1, 2, 3, \ldots, n$, то преминаването от основната пермутация $1\ 2\ 3 \ldots n$ към пермутацията $a_1\, a_2\, a_3 \ldots a_n$ е субституция. Както видяхме по-рано, въпросната субституция означаваме с

$$S=\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}.$$

Оттук веднага следва, че от n елемента имаме $n!$ субституции. Субституцията, която не променя основната пермутация, я наричаме единица и я означаваме с E:

$$E=\begin{pmatrix} 1 & 2 & 3 \ldots n \\ 1 & 2 & 3 \ldots n \end{pmatrix}.$$

Под произведение ST на две субституции разбираме субституцията, която се получава, като извършим отначало субституцията S и след това субституцията T. Умножението не притежава комутативното свойство. Ако $ST=TS$, то такива субституции се наричат ***комутативни.*** Лесно се вижда, че асоциативният (съдружителният) закон при умножението е в сила. Действително за субституциите

$$S=\begin{pmatrix} a & \dots \\ b & \dots \end{pmatrix}, \qquad T=\begin{pmatrix} b & \dots \\ c & \dots \end{pmatrix}, \qquad U=\begin{pmatrix} c & \dots \\ d & \dots \end{pmatrix}$$

получаваме

$$TU=\begin{pmatrix} b & \dots \\ c & \dots \end{pmatrix}\begin{pmatrix} c & \dots \\ d & \dots \end{pmatrix}=\begin{pmatrix} b & \dots \\ d & \dots \end{pmatrix},$$

$$ST=\begin{pmatrix} a & \dots \\ b & \dots \end{pmatrix}\begin{pmatrix} b & \dots \\ c & \dots \end{pmatrix}=\begin{pmatrix} a & \dots \\ c & \dots \end{pmatrix}$$

и следователно

$$S(TU)=\begin{pmatrix} a & \dots \\ d & \dots \end{pmatrix}, \quad (ST)U=\begin{pmatrix} a & \dots \\ d & \dots \end{pmatrix},$$

т. е. двете субституции $S(TU)$ и $(ST)U$ са една и съща субституция. По употребен начин следва от това, че асоциативният закон остава верен и за произведение от произволен брой субституции.

Под S^{-1} разбираме обратната субституция на S, т. е.

$$S^{-1}=\begin{pmatrix} a_1, & a_2, & a_3, & \dots, & a_n \\ 1, & 2, & 3, & \dots, & n \end{pmatrix}.$$

Очевидно ще имаме $SS^{-1}=S^{-1}S=E$. Под S^m при m цяло положително число разбираме субституцията, която се получава, като умножим S m пъти и под S^{-m}, m цяло положително число, се разбира субституцията $(S^{-1})^m$. Полагаме $S^0=E$ и вместо S^1 пишем само S. Така виждаме, че правилата за умножение на степени

$$S^kS^l=S^{k+l}, \quad (S^k)^l=S^{kl}$$

(k и l цели числа) са в сила. Лесно се вижда, че

$$(ST)^{-1}=T^{-1}S^{-1}.$$

Действително, ако поставим $U=T^{-1}S^{-1}$, то получаваме

$$U(ST)=T^{-1}S^{-1}ST=T^{-1}(S^{-1}S)\,T=T^{-1}ET=E.$$

Изобщо ще имаме

$$(STV\dots)^{-1}=\dots V^{-1}T^{-1}S^{-1}.$$

Нека S е произволна субституция и да разгледаме редицата от степените на S:

$$S, \; S^2, \; S^3, \dots$$

Понеже броят на субституциите от n елемента е краен, то ще има поне две равни степени $S^k=S^l, k>l$. Като умножим предното равенство

с S^{-1}, получаваме $S^{k-1}=E$, т. е. има степен на S, която е равна на E. Нека g е най-малкото положително цяло число, за което $S^g=E$. Тогава g се нарича ред (период) на субституцията S. В този случай субституциите

$$E,\ S,\ S^2,\ \ldots,\ S^{g-1} \tag{1}$$

са различни помежду си, понеже ако $S^\mu=S^\nu$, $0\leqq\mu<\nu\leqq g-1$, то би следвало, че $S^{\nu-\mu}=E$, което е невъзможно, тъй като $\nu-\mu<g$. Нека m е произволно цяло число и S е субституцията в (1). Нека q и r са частното и остатъкът от делението на m с g, т. е. $m=qg+r$, $0\leqq r\leqq g-1$. Тогава от равенствата

$$S^m=S^{qg}S^r=(S^g)^qS^r=E^qS^r=S^r$$

следва, че S^m е само тогава равна на E, когато $S^r=E$. Но последното е възможно само ако $r=0$. Следователно S^m е само тогава равна на E, когато m се дели без остатък на g. Да намерим реда на субституцията S^k. Трябва да се намери най-малкото цяло положително число μ, за което $(S^k)^\mu=S^{k\mu}=E$. Последното равенство показва, че $k\mu$ трябва да се дели на g. Ако d е общият най-голям делител на числата k, g, то $k=k_1d$ и $g=g_1d$, като k_1 и g_1 са взаимно прости. Но частното $\frac{k\mu}{g}$ е равно на $\frac{k_1\mu}{g_1}$ и понеже трябва да е цяло число, то следва, че μ се дели на g_1, т. е. $\mu=g_1$. Така установихме, че редът на S^k е равен на $\frac{g}{d}$ и следователно само тогава е равен на g, когато $d=1$, т. е. k и g са взаимно прости числа. В последния случай степените на S^k ще са еднакви с тези на S. Достатъчно е да докажем, че всяка положителна степен S^p на S е степен на S^k. Понеже g и k са взаимно прости числа, известно е, че съществуват цели числа x и y, които удовлетворяват уравнението

$$kx+gy=p.$$

Но тогава

$$S^p=S^{kx}S^{gy}=(S^k)^x.$$

2. Кръгови субституции. Една субституция се нарича кръгова или циклична, ако замества всеки елемент със следващия го и последния с първия, т. е. субституция от вида

$$S=\begin{pmatrix} a_1 & a_2 & \ldots & a_{p-1} & a_p \\ a_2 & a_3 & \ldots & a_p & a_1 \end{pmatrix}.$$

Отбелязваме и накратко с $(a_1\ a_2\ \ldots\ a_{p-1}\ a_p)$. На първо място може да се постави кой да е елемент: така например

$$(bcda)=(cdab)=(dabc).$$

1. Всяка субституция може да се представи като произведение на кръгови субституции, наречени цикли. За да покажем това, достатъчно е да си послужим с един пример. Да разгледаме субституцията S от 9 елемента

$$S=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\\ 4 & 5 & 6 & 7 & 2 & 3 & 9 & 8 & 1\end{pmatrix}.$$

Тази субституция замества елемента 1 с 4, 4 със 7, 7 с 9 и 9 с 1. Така получаваме един цикъл (1 4 7 9). Елементът 2 се замества от S с 5, 5 с 2 и така получаваме нов цикъл (2 5), след това трети цикъл (3 6) и четвърти цикъл (8), състоящ се само от един елемент. Следователно можем да пишем

$$S=(1479)\ (25)(36)(8).$$

Обикновено, ако са известни елементите на субституцията, циклите, съставени само от един елемент, не се пишат. От начина на извеждането е ясно, че отделните цикли нямат общи елементи и като такива са очевидно комутативни при умножение.

Редът на всяка циклична субституция е равен на броя на влизащите в нея елементи. Нека именно $S=(a_1\ a_2\ \ldots a_{p-1}\ a_p)$ S^2 ще замества a_1 с a_3, a_2 с a_4, . . ., a_{p-2} с a_p, a_{p-1} с a_1, т. е. всеки елемент a_i с a_{i+2}, като, ако $i+2>p$, индексът се замества с остатъка му по модул p; S^3 замества a_i с a_{i+3} и т. н., а S^p замества a_i с a_i, т. е. $S^p=E$.

2. *Редът на една субституция е равен на най-малкото общо кратно на редовете на циклите, на които тя се разпада, в случай, че последните нямат общи елементи.* Действително нека S се разпада на циклите C_1, C_2, . . ., C_k, т. е. $S=C_1\ C_2\ \ldots\ C_k$ и циклите да нямат общи елементи. Понеже C_i са комутативни помежду си, то $S^m=C_1^m C_2^m\ \ldots\ C_k^m$ и за да имаме $S^m=E$, трябва

$$C_1^m=C_2^m=\ \ldots\ =C_k^m=E,$$

което е само тогава възможно, когато m е кратно на редовете на циклите, от което следва и предложението.

Всяка кръгова субституция от два елемента се нарича транспозиция.

3. *Всяка субституция е или транспозиция, или произведение на транспозиции.* Действително всяка субституция е произведение на кръгови субституции, а всяка кръгова субституция е произведение от транспозиции, както се вижда от равенството

$$(1\ 2\ 3\ \ldots m)=(1\ 2)\ (1\ 3)\ \ldots\ (1m)$$

Обаче представянето на една субституция с транспозиции не е еднозначно. Така например имаме

$$(1\ 3)(1\ 2)(1\ 3)(1\ 2)=(1\ 2)(1\ 3)=(1\ 2\ 3).$$

Но в сила е предложението.

4. *Ако една субституция се представя като произведение на транспозиции по няколко начина, то броят на транспозициите за всичките представяния е или четен, или нечетен.*

Нека субституцията

$$S=\begin{pmatrix}1 & 2 & 3 & \ldots & n\\ a_1 & a_2 & a_3 & \ldots & a_n\end{pmatrix}$$

е представена като произведение на транспозиции

$$S=(b_1c_1)\ (b_2c_2)\ \ldots\ (b_pc_p).$$

Значи пермутацията $a_1a_2a_3 \ldots a_n$ се получава от пермутацията 1 2 3 . . . n, като отначало сменим елементите b_1 и c_1 помежду им, сетне елементите b_2 и c_2 и т. н. и най-после елементите b_p и c_p. По лемата на стр. 17 пермутацията $a_1a_2a_3 \ldots a_n$ ще бъде четна или нечетна, ако съответно p е четно или нечетно число, т. е. четността или нечетността на p ще бъде напълно определена.

Една субституция се нарича *регулярна*, ако циклите, на които се разпада, са с еднакъв брой елементи, като се предполага естествено, че два кои да са цикъла нямат общи елементи.

5. Ако S е кръгова субституция от ред p, то S^q е регулярна субституция, състояща се от d цикъла, като d е общият най-голям делител на p и q.

Действително нека

$$S=(a_1a_2 \ldots a_p).$$

Тогава в S^q получаваме отначало един цикъл $(a_1\, a_{1+q}\, a_{1+2q} \ldots)$, гдето всеки индекс, по-голям от p, трябва да бъде заместен с остатъка си по модул p. Ако i е броят на елементите в предния цикъл, то за да се върнем към първия елемент a_1, трябва да имаме

$$1+iq=1+jp,$$

откъдето $i=\frac{jp}{q}$. Ако p и q са взаимно прости, то следва, че $j=q$, $i=p$, т. е. S^q е кръгова. Ако $p=p_1d$, $q=q_1d$, като p_1 и q_1 са взаимно прости, то ще имаме $j=q_1$, $i=p_1$, $d=\frac{p}{i}$, т. е. S^q се състои от d цикъла. Така например за

$$S=(1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6)$$

ще имаме

$$S^2=(1\ \ 3\ \ 5)\ (2\ \ 4\ \ 6),\ S^3=(1\ \ 4)\ (2\ \ 5)\ (3\ \ 6),\ S^4=(1\ \ 5\ \ 3)\ (2\ \ 6\ \ 4),$$

$$S^5=(1\ \ 6\ \ 5\ \ 4\ \ 3\ \ 2),\ S^6=E.$$

6. Всяка регулярна субституция е степен на една кръгова. Нека регулярната субституция е

$$S=(a_1b_1c_1 \ldots l_1)(a_2b_2c_2 \ldots l_2) \ldots (a_mb_mc_m \ldots l_m).$$

Ако поставим

$$U=(a_1a_2 \ldots a_mb_1b_2 \ldots b_m \ldots l_1l_2 \ldots l_m),$$

очевидно

$$S=U^m.$$

3. **Подобни комутативни субституции.** Две субституции се наричат подобни, ако са съставени от еднакъв брой цикли с еднакъв брой елементи в тях. Както преди споменахме, субституцията $S_1=T^{-1}ST$ се нарича трансформирана на S посредством T, или S_1 е спрегната (конюгована) на S. Нека $(abc \ldots)$ е един цикъл на S и a_1, b_1, c_1, . . . са елементите, с които се заместват при прилагане на T. Тогава T^{-1} замества a_1 с a, S замества a с b и T замества b с b_1 така, че S_1 замества a_1 с b_1. Също S_1 замества b_1 с c_1 и т. н., т. е. на цикъла $(abc \ldots)$ в S съответствува цикъла $(a_1b_1c_1 \ldots)$ в S_1. Оттук се вижда, че S_1 се получава от S, като в циклите на последната субституция

извършим субституцията T. Следователно S и S_1 са подобни субституции. Така, ако

$$S=(1\ 2\ 3)(4\ 5), \qquad T=(1\ 3\ 4)(2\ 5),$$

то $T^{-1}ST=(3\ 5\ 4)(1\ 2)$. Обратно, нека S и S' са две подобни субституции. Тогава лесно може да се намери субституцията T, която трансформира S в S'. Тази субституция именно замества всеки елемент от S с елемент от S', който заема същото място. Така, ако $S=(1\ 2\ 3\ 4)$, $S'=(3\ 4\ 1\ 2)$,

то

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

Ще установим няколко лесно доказуеми свойства. От равенството

$$T^{-1}(SU)T=T^{-1}ST\,.\,T^{-1}UT$$

се вижда, *че трансформираната субституция на едно произведение е произведението от трансформираните субституции на множителите*. От

$$ST=T^{-1}(TS)T$$

следва, че ST е трансформирана субституция на TS и следователно ST и TS са подобни субституции. От предположението $ST=TS$ следва, че

$$U^{-1}SU\,.\,U^{-1}TU=U^{-1}TU\,.\,U^{-1}SU,$$

т. е. *комутативните субституции се трансформират с една субституция пак в комутативни субституции.*

Ще установим сега предложението:

7. Ако m е цяло положително число, взаимно просто с реда на една субституция S, то има субституция T, за която

$$T^{-1}ST=S^m. \tag{1}$$

Понеже от предложението 2 следва, че редът на S се дели на реда на всеки от циклите, на които тази субституция се разпада, то m ще бъде взаимно просто с редовете на въпросните цикли. Следователно S^m е подобна субституция на S, понеже всеки цикъл, повдигнат в m-та степен, дава цикъл от същия ред. По предните изводи ще има субституции T, които удовлетворяват (1). Можем да намерим всички тези субституции. Именно нека T_1 е една от тези субституции. Тогава (1) може да се напише така:

$$T_1^{-1}ST_1=T^{-1}ST,$$

откъдето получаваме

$$TT_1^{-1}S=STT_1^{-1},$$

т. е. субституциите S и TT_1^{-1} са комутативни. Нека тогава U е комутативна субституция на S и нека $U=TT_1^{-1}$, откъдето $T=UT_1$. Тази субституция удовлетворява (1), понеже имаме

$$(UT_1)^{-1}SUT_1=T_1^{-1}U^{-1}SUT_1=T_1^{-1}ST_1=S^m.$$

Следователно всичките решения на (1) получаваме, като умножим едно специално решение с всички комутативни субституции на S.

4. Групи от субституции. Един комплекс **G** от субституции на дадени n елемента е група, ако като произведението на кои да са две равни или не субституции от **G** принадлежи също на G. Ако S е една субституция от **G**, то видяхме, че има степен на S^k, която е равна на E. Следователно субституцията единица принадлежи на комплекса **G**. Понеже $S^{-1}=S^{k-1}$, то всеки елемент от **G** има обратен S^{-1}, който принадлежи също на **G**. Следователно от основните постулати за група, достатъчно е да бъде изпълнен първият постулат.

Очевидно съвкупността на всичките n субституции от n-те елемента 1, 2, 3, . . ., n образуват група, която обикновено се бележи с **S** и се нарича симетричната група. Лесно се вижда, че съвкупността на всички субституции от n-те елемента, които се разлагат на четен брой транспозиции, образува група **A** от ред $\frac{n}{2}$, наречена а л т е р н а т и в н а група. Степените на всяка субституция образуват група, наречена к р ъ г о в а или ц и к л и ч н а, на която редът е равен на реда на субституцията.

По теоремата на Лагранж редът на всяка група от субституции трябва да дели реда $n!$ на симетричната група **S**.

Една група **G** от елементи 1, 2, 3, . . ., n се нарича т р а н з и т и в н а, ако субституциите ѝ сменят всеки елемент j с всеки друг. В противен случай тя се нарича и н т р а н з и т и в н а. Лесно следва предложението: *Една група е само тогава транзитивна, когато сменя елемента 1 с всичките останали елементи.*

Действително нека a и b са два кои да са елемента от 1, 2, 3, . . ., n. В групата ще има субституция S, която сменя 1 с a, и субституция T, сменяща 1 с b. Тогава субституцията $S^{-1}T$ принадлежи на групата и сменя a с b, с което предложението е установено. Нека $S_1=E, S_2, \ldots, S_k$ да бъдат субституциите от групата **G**, които не сменят елемента 1. Очевидно те образуват една подгрупа **A**. Нека T_j е субституция, която сменя 1 с j $(1<j\leqq n)$. Тогава субституциите от **G**, които сменят 1 с j, ще са субституциите на комплекса $\mathbf{G}T_j$. Понеже ако U е субституция, сменяща 1 с j, то UT^{-1} не ще променя 1, т. е. ще бъде равна на някоя субституция S_p от **A**. Но тогава от $UT_j^{-1}=S_p$ получаваме $U=S_pT_j$, с което се установява предното твърдение. Разлагането на групата **G** по **A** ще бъде следователно

$$\mathbf{G}=\mathbf{A}+\mathbf{A}T_2+\ldots+\mathbf{A}T_n.$$

Оттук следва непосредствено теоремата:

8. *Редът на всяка транзитивна група от n-та степен се дели на n.*

Значи редът на всяка транзитивна група от степен n е най-малко равен на n. Лесно е да дадем пример за транзитивна група от ред, равен на степента ѝ. Така цикличната група от ред n:

$$E+S+S^2+\ldots+S^{n-1},\ S=(1\ 2\ 3\ \ldots\ n)$$

е транзитивна, понеже при всяко j, $1<j\leqq n$, има субституция, която сменя 1 с j. Това е именно субституцията S^{j-1}.

Една транзитивна група **G** от субституции се нарича и м п р и м и т и в н а, ако можем така да разпределим елементите ѝ на повече от

една система, че със субституциите на G елементите на една коя да е система или да се разместват помежду им, или да преминават в елементите на друга система. Очевидно във всяка система трябва да има еднакъв брой елементи. Ако G е от степен n, то можем да вземем за елементи числата 1, 2, . . ., n. Тогава, ако m е броят на елементите във всяка система, то всичките елементи могат да се представят в следната таблица:

$$\begin{array}{ll} (A') & a_1'\ a_2', \ldots, a_m', \\ (A'') & a_1''\ a_2'', \ldots, a_m'', \\ (1) & \ldots\ldots\ldots\ldots \\ (A^{\nu}) & a_1^{(\nu)}, a_2^{(\nu)}, \ldots, a_m^{(\nu)}, \end{array}$$

като $\nu = \frac{n}{m}$, $\nu > 1$.

Нека $T_1 = E$, T_2, T_3, . . ., T_μ са субституциите от G, които разместват само елементите на системата (A'). Очевидно те ще образуват една подгрупа A на G.

В групата на G ще има субституция V_i, която сменява системата (A') с една произволна друга система $(A^{(i)})$ от (1), което да означим символично с равенството $V_i(A') = A^{(i)}$, $V_1 = E$. Тогава очевидно ще имаме и

$$(2) \qquad T_k . V_i(A') = A^{(i)},\ k = 1, 2, 3, \ldots, \mu.$$

Ако за някоя субституция S от G имаме също $S(A') = A^{(i)}$, то бяхме получили

$$S . V_i^{-1}(A') = V_i^{-1}(A^{(i)}) = A',$$

т. е. SV_i^{-1} е една субституция T_j измежду разглежданите и следователно $S = T_j V_i$. Следователно съседният комплекс $\mathbf{A}V_i$ се състои от всичките субституции на G, които трансформират (A') в $(A^{(i)})$. Това можем да означим с равенство

$$\mathbf{A}V_i(A') = (A^{(i)}).$$

Но всяка субституция S на G трансформира системата (A') в една от системите (1). Следователно S се съдържа в един от комплексите

$$(3) \qquad \mathbf{A}V_1,\ \mathbf{A}V_2,\ \mathbf{A}V_3, \ldots \mathbf{A}V_\nu$$

т. е. групата G се разпада относно подгрупата A на комплексите (3).

Ако редът на G е равен на g, то ще имаме оттук $g = \mu\nu$ и понеже $n = m\nu$, ще имаме $\mu = \frac{mg}{n}$. Виждаме, че числото mg е кратно на n.

Всички конюговани с A групи ще се дават с

$$(4) \qquad V_1^{-1}\mathbf{A}V_1,\ V_2^{-1}\mathbf{A}V_2,\ V_3^{-1}\mathbf{A}V_3, \ldots, V_\nu^{-1}\mathbf{A}V_\nu,$$

които очевидно не всички са различни.

Лесно се убеждаваме, че групата

$$\mathbf{A}_i = V_i^{-1}\mathbf{A}V_i,\ \mathbf{A}_1 = \mathbf{A},$$

се състои от всичките субституции от **G**, които само разместват елементите на системата $(A^{(i)})$ помежду им. Тогава разрезът

$$\mathbf{D}=(\mathbf{A}_1,\ \mathbf{A}_2,\ \mathbf{A}_3,\ \ldots)$$

ще се състои от всички субституции на **G**, които всяка от системите (1) трансформират в себе си. Може да се случи, че $\mathbf{D}=\mathbf{E}$, т. е. групите (4) имат само E за общ елемент. Ако редът на **D** е равен на μ, то групите (4) съвпадат с **A** и **A** е и. пг. на **G**. Тогава **A** е една интранзитивна група, сменяваща само елементите на отделните системи помежду им, като системата (1) е нейната система на интранзитивност.

Ако една група **G** не е импримитивна, то тя се нарича примитивна.

5. Представяне на една група с група от субституции. Имаме следното предложение:

9. Всяка крайна група **G** от ред n е изоморфна на една група от субституции от степен n.

Нека групата **G** се състои от елементите

$$A_1,\ A_2,\ \ldots,\ A_n. \tag{1}$$

Ако A_i е произволен елемент от редицата (1), то редицата

$$A_1A_i,\ A_2A_i,\ \ldots,\ A_nA_i \tag{2}$$

е еднаква с (1), понеже **G** е група, т. е. (2) представлява една пермутация на (1), която при $A_i\neq E$ е отлична от (1).

Ако означим с

$$S_i=\begin{pmatrix} A_1 & A_2 & \ldots & A_n \\ A_1A_i & A_2A_i & \ldots & A_nA_i \end{pmatrix},$$

то получаваме така n субституции $S_1, S_2, \ldots, S_n$. Ако A_1 е единицата на **G**, то $S_1=E$. Субституциите S_i са различни, защото от $A_kA_i=A_kA_j$ следва, че $A_i=A_j$. Лесно е да видим, че те образуват група, изоморфна на **G**. Действително, ако

$$A_iA_j=A_r,$$

то S_j сменя пермутацията (2) с

$$A_1A_r,\ A_2A_r,\ \ldots,\ A_nA_r, \tag{3}$$

която се получава от (1) с извършване на субституцията S_r. Следователно ще имаме

$$S_iS_j=S_r,$$

с което твърдението е установено.

Да напишем всичките редици (2) за елементите A_i, $i=1, 2, \ldots, n$

Така получаваме една таблица (Кейли),

$$\begin{array}{c} A_1A_1,\ A_2A_1,\ \ldots,\ A_nA_1, \\ A_1A_2,\ A_2A_2,\ \ldots,\ A_nA_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ A_1A_n,\ A_2A_n,\ \ldots,\ A_nA_n, \end{array}$$

6. Ред на разлагане на симетричната група. Задачата да се намери редът на разлагането на една група е доста сложна. Ние ще изследваме разлагането на симетричната група, което ще има важно значение нататък. Симетричната група **S** от n елемента съдържа $n!$ субституции, а алтернативната **A** има $\frac{n!}{2}$ субституции. **A** е подгрупа на **S** с индекс 2. Ако T е коя да е субституция от **S**, то понеже комплексът $T^{-1}\mathbf{A}T$ се състои пак от четни субституции, то той е равен на **A**. Следователно **A** е и. пг. на **S** и понеже индексът е прост, тя е максимална.

При $n=2$ имаме $\mathbf{A}=\mathbf{E}$ и имаме реда на разлагане

$$\mathbf{S},\ \mathbf{E}$$

с индекса 2. Групата **S** е метациклична.

При $n=3$ групата **A** е от трети ред, т. е. от ред, който е просто число. **A** има само истинска подгрупа **E** и ще имаме за **S** реда на разлагане

$$\mathbf{S},\ \mathbf{A},\ \mathbf{E}$$

с индексен ред: 2, 3. Групата **S** е метациклична.

При $n=3$ да разгледаме комплекса

$$\mathbf{B}=E+(12)(34)+(13)(24)+(14)(23),$$

за който лесно се проверява, че представлява група. Понеже **B** се състои от субституции, съдържащи четен брой транспозиции, то **B** е подгрупа на алтернативната **A**, дадена с

$$\mathbf{A}=E+(12)(34)+(13)(24)+(14)(23)+(123)+(134)+$$
$$+(243)+(142)+(132)+(234)+(124)+(143).$$

Разлагането на **A** по **B** е следното:

$$\mathbf{A}=\mathbf{B}+\mathbf{B}(123)+\mathbf{B}(132).$$

Лесно се вижда, че **B** е и. пг. на **A** и на S, понеже всяка конюгована на **B** подгрупа **S** ще има три субституции по два цикъла, следователно ще е идентична с **B**, понеже други субституции от този вид няма в **S**. **B** е м. и. пг. на **A** с индекс 3, просто число.

Групата

$$\mathbf{C}=E+(12)(34)$$

е една м. и. пг. на **B** и така получаваме реда на разлагане

$$\mathbf{S},\ \mathbf{A},\ \mathbf{B},\ \mathbf{C},\ \mathbf{E}$$

със съответния индексен ред 2, 3, 2, 2. **S** и **A** са метациклични. Вместо **C** може да се вземат групите

$$\mathbf{C}_1=E+(13)(24),\quad \mathbf{C}_2=E+(14)(23).$$

При $n>4$ се получава съвсем различен резултат. Именно имаме теоремата:

10. Алтернативната субституционна група **A** от степен $n>4$ е проста. Следователно за симетричната група **S** имаме реда на разлагане

S, A, E

с индексен ред: 2, $\frac{1}{2}n!$. **S** и **A** не са метациклични групи.

Ще забележим, че цикличните субституции от нечетен ред се съдържат в **A**, понеже, както вече видяхме отначало, те са произведение на четен брой транспозиции. Нека **U** да бъде и. пг. на **A**. Последователно ще докажем отделни предложения, от които последното дава изказаната теорема.

A. **U** не съдържа цикли (кръгови субституции) от трети ред.

B. **U** не съдържа субституции, в представянето на които по цикли да фигурира цикъл с повече от три елемента.

C. **U** не съдържа субституция, в която, разложена на цикли, да влиза цикъл от три елемента.

D. **U** съдържа само субституцията E.

Доказване на A. Ако **U** съдържа един цикъл (123) и ако $(m\ p\ q)$ е друг произволен цикъл от трети ред, то **A** ще съдържа поне една от субституциите

$$S=\begin{pmatrix}1&2&3\\m&p&q\end{pmatrix},\qquad T=\begin{pmatrix}1&2&3\\m&q&p\end{pmatrix}.$$

Понеж **U** е и. пг. на **A**, то **U** ще съдържа поне една от субституциите

$$S^{-1}(123)\,S=(mpq),$$
$$T^{-1}(123)\,T=(mqp)$$

и във втория случай също $(mqp)^2=(mpq)$. Това показва, че **U** съдържа всички кръгови субституции (т. е. цикли) от трети ред. Но тогава **U** ще съдържа произведенията на циклите от трети ред и по формулите

$$(ab)\,(ac)=(abc),$$
$$(ab)\,(cd)=(abc)\,(adc)$$

ще съдържа произведенията по четен брой транспозиции. Следователно **U** е равно на **A**, а не истинска подгрупа.

B. Да допуснем обратното на тази точка. Тогава **U** ще съдържа субституцията S, дадена с

$$S=(1\ 2\ 3\ \ldots\ p)\,L,$$

гдето $p>3$, L е произведение на другите цикли или $L=E$, ако няма такива. Тогава, понеже **U** е и. пг. на **A**, то **U** ще съдържа субституцията

$$(123)^{-1}\,S\,(123),$$

следователно и субституцията

$$S^{-1}(123)^{-1}\,S\,(123)=$$
$$=(p\ \ldots\ 3\ \ 2\ \ 1)\ L^{-1}(1\ \ 3\ \ 2)\ (1\ \ 2\ \ 3\ \ldots\ p)\ \ L\,(1\ \ 2\ \ 3).$$

Понеже елементите 1, 2, . . . , p не влизат в L, то може да се разместят местата на L и L^{-1}, на които произведението е равно на E, така че горното произведение става

$$(p \ldots 3\ 2\ 1)\ (1\ 3\ 2)\ (1\ 2\ 3 \ldots p)\ \ (1\ 2\ 3) = (1\ 2\ 4).$$

Така достигаме до заключение, че **U** съдържа цикъла (1 2 4) от трети ред, което е невъзможно по A.

C. Ако в **U** имаме субституция S, в представянето на която по цикли влиза един от трети ред, то по **B** другите цикли на **S** могат да бъдат само от втори или трети ред. В S^2 влизат само цикли от трети ред и по *A.* трябва да има поне два такива. Нека следователно

$$S^2 = (1\ 2\ 3)\ (4\ 5\ 6)\, L,$$

гдето L е произведение на другите цикли. Но тогава **U** като и. пг. на **A** съдържа субституцията

$$(1\ 4\ 2)^{-1}\ S^2 (1\ 4\ 2)$$

и следователно и субституцията

$$S^{-2}(1\ 4\ 2)^{-1} S^2 (1\ 4\ 2) =$$
$$= (132)\ (465)\, L^{-1}\ (124)\,(123)\,(456)\ L\,(142) = (14235).$$

U ще съдържа значи един цикъл от пети ред, което противоречи на B.

D. Ако **U** съдържа субституция S, отлична от E, то в S могат да влизат по B. и C. само транспозиции, и то естествено в четен брой.

Нека S съдържа само два такива, например

$$S = (12)\,(34).$$

Понеже $n > 4$, ще има поне един още елемент 5. Тогава, понеже **U** е и. пг. на **A**, тя ще съдържа и субституцията

$$(125)^{-1}\ S\,(125),$$

следователно и субституцията

$$S(1\ 2\ 5)^{-1}\ S(1\ 2\ 5) = (12)(34)(152)\,(12)\,(34)\,(125) = (152),$$

т. е. един цикъл от трети ред, което по A. е невъзможно.

Нека обаче S съдържа повече от два такива цикъла, т. е. нека

$$S = (1\ 2)(3\ 4)(5\ 6)\, L,$$

гдето L е произведението на другите цикли. Понеже **U** е и. пг. на **A**, то тя ще съдържа субституцията $(1\ 3\ 5)^{-1} S(1\ 3\ 5)$, също и субституцията

$$S^{-1}(1\ 3\ 5)^{-1} S(1\ 3\ 5) =$$
$$= (1\ 2)(3\ 4)(5\ 6)\, L^{-1}(1\ 5\ 3)(1\ 2)(3\ 4)(5\ 6)\, L(1\ 3\ 5) = (1\ 3\ 5)(2\ 6\ 4),$$

което по **C.** е невъзможно.

7. Линейна група от субституции. В много въпроси от теорията на групите е полезно да се въведе аналитично представяне на субституциите. Нека

$$S = \begin{pmatrix} 0 & 1 & 2 & \ldots & n-1 \\ a & b & c & \ldots & l \end{pmatrix}$$

е произволна субституция на елементите 0, 1, 2, . . . , $n-1$. По формулата на Лагранж можем да построим полином $f(z)$ от степен, най-много равна на $n-1$, който за $z=0, 1, 2, \ldots, n-1$ взема съответно стойностите $a, b, c, \ldots, l$. Тогава S може да се пише накратко така:

$$S=\begin{pmatrix} z \\ f(z) \end{pmatrix}.$$

По-удобно е да се вземе друг полином, на който остатъците на стойностите му по модул n при $z=0, 1, 2, \ldots, n-1$ са числата $a, b, c, \ldots, l$. Тогава S може да се пише така:

$$S=\begin{pmatrix} z \\ \varphi(z) \end{pmatrix},$$

или по-кратко $|z\,\varphi(z)|$.

Не ще разглеждаме общия случай, а ще се спрем на специална класа субституции, при които полиномът $\varphi(z)$ е от първа степен, т. е.

$$\varphi(z)=az+b,\ a \text{ и } b \text{ цели числа, } a>0.$$

Нека a е взаимно просто с n. Тогава, когато z взема стойностите $0, 1, 2, \ldots, n-1$, изразът $az+b$, както знаем, взема стойностите на една неконгруентна система числа по модул n. Следователно остатъците им ще бъдат същите цели числа от 0 до $n-1$, но въобще в друг ред. Същото е в сила, когато z взема стойностите на коя да е неконгруентна по модул n система числа. Значи при a, взаимно просто с n, ще разгледаме субституции от вида

$$|z\ az+b|,$$

които ще бележим с $S_{a,b}$. Очевидно можем да се ограничим на $0 \leqq b<n$, понеже ако $b \geqq n$, ще заместим b с остатъка му r по n и тогава имаме

$$S_{a,b}=S_{a,r}.$$

Лесно се вижда, че всички субституции $S_{a,b}$, гдето a взема за стойности всички прости с n числа и по-малки от n, а b — стойностите 0, 1, 2, . . . , $n-1$, образуват група от ред $n\varphi(n)$, гдето $\varphi(n)$ е индикаторът на n, наречена главна линейна група. Действително първо субституциите $S_{a,b}$ са все различни. Понеже от

$$S_{a,b}=S_{c,d}$$

следва при $z=0, 1, 2, \ldots, n-1$,

$$az+b \equiv cz+d \pmod{n}.$$

Специално при $z=0$ получаваме $b \equiv d$, т. е. $b=d$, и при $z=1$ получаваме $a \equiv c$, отгдето $a=c$.

За произведението на две субституции ще имаме

$$S_{a,b}.S_{c,d}=|z\ az+b|.|y\ cy+d|=$$
$$=|z\ az+b\ .|\ az+bc(az+b)+d|=|z\ acz+bc+d|.$$

Следователно

$$S_{a,b}.S_{c,d}=S_{ac,\ bc+d}, \tag{15}$$

гдето вдясно индексите трябва да се заместят с остатъците им по модул n. Понеже a и c са прости с n, то и ac е просто с n, така че в дясната част субституцията принадлежи на комплекса, т. е. последният е група. Единицата е $S_{1,0}=E$.

Ако $a=1$, $b=0, 1, 2, \ldots, n-1$, то

$$S_{1,b}=|z\ z+b|=S^b, \text{ гдето } S=|z\ z+1|.$$

Следователно цикличната група

$$E, S, S^2, \ldots, S^{n-1}$$

е подгрупа на линейната.

Също при $b=0$, $d=0$ от (11) имаме

$$S_{a,0}\ S_{c,0}=S_{ac,0},$$

т. е. субституциите $S_{a,0}$, гдето a взема за стойности всички прости с n числа, образуват една подгрупа от ред $\varphi(n)$.

Пример: $n=4$. a взема стойностите 1 и 3, а b — стойностите 0, 1, 2, 3. За линейната група имаме

$$S_{1,b}=|z\ z+b|=(1234)^b,$$

$$S_{3,b}=|z\ 3z+b|=|z\ 3z|\ .\ |3z\ 3z+b|=(13)\ (1234)^b.$$

Следователно тя ще бъде

$$E+(1234)+(13)(24)+(1\ 4\ 3\ 2)+(13)+(14)(23)+$$
$$+(24)+(12)(34).$$

Глава II

Полиноми, принадлежащи на една група от субституции

1. Връзка между реда на групата и броя на стойностите на функцията. Нека $f(x_1, x_2, \ldots, x_n)$ е един полином на независимите променливи $x_1, x_2, \ldots, x_n$. Върху променливите x_i да извършим субституцията

$$S=\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \alpha & \beta & \gamma & \ldots & \varepsilon \end{pmatrix}.$$

Тогава полиномът $f(x_1, x_2, \ldots, x_n)$ ще стане равен на

$$\varphi(x_1, x_2, \ldots, x_n)=f(x_\alpha, x_\beta, x_\gamma, \ldots, x_\varepsilon).$$

Ако алгебрически полиномът φ е равен на f, т. е. имаме идентично

$$f(x_\alpha, x_\beta, \ldots, x_\varepsilon)=f(x_1, x_2, \ldots, x_n),$$

то казваме, че S не *променя* полинома.

Нека $S_1=E, S_2, \ldots, S_m$ са всички субституции, които не променят полинома f. Тогава лесно се вижда, че те образуват една група $\mathbf{G}$, която казваме, че принадлежи на f. Действително, ако извършим една субституция S_p, то f не се променя, т. е. f взема стойност $f_1=f$. Ако сега приложим субституцията S_g, то f_1 се обръща в $f_2=f$. Следователно $S_p S_g$ изменя f в $f_2=f$, т. е. ще бъде някой S_r, $1 \leqq r \leqq m$.

Така на функцията

(1) $$f(x_1, x_2, x_3, x_4)=x_1x_2+x_3x_4$$

принадлежащата ѝ група е

(2) $$\mathbf{A}=E+(12)+(34)+(12)(34)+(13)(24)+(14)(23)+$$
$$+(1324)+(1423).$$

На симетричните функции групата е от всички субституции, ето защо последната се нарича симетрична група. На алтернативните функции е алтернативна група.

Нека

(3) $$\mathbf{G}=S_1+S_2+\ \ldots\ +S_m,\ S_1=E$$

е групата на една функция $f(x_1, x_2, \ldots, x_n)$. Ако $m<n!$, то ще има субституция T_2, която променя стойността на $f=f_1$ в друга f_2.
Тогава субституциите от комплекса

(4) $$\mathbf{G}\,T_2$$

ще изменят f_1 в f_2 и само те ще бъдат с това свойство. Действително, ако U е такава субституция, то субституцията UT_2^{-1} не ще променя f_1, т. е. ще принадлежи на $\mathbf{G}$. Следователно ще имаме

$$UT_2^{-1}=S_g,$$

отгдето $U=S_g\,T_2$. Така ще имаме нови m-субституции (4). Ако $2m=n!$, то функцията f ще бъде двузначна. В противен случай ще има субституция T_3, която изменя f_1 в нова стойност f_3 и същото ще правят всички субституции от комплекса $\mathbf{G}T_3$. Това са също единствените субституции със свойство да променят f_1 в f_3. Продължавайки така, стигаме до общото заключение: Нека $f_1, f_2, \ldots, f_r$ са всички възможни стойности на функцията $f(x_1, x_2, \ldots, x_n)$. Ако $T_2, T_3, \ldots, T_r$ са субституции, които изменят f_1 съответно в $f_2, f_3, \ldots, f_r$, то всички $n!$ субституции могат да се представят в следната таблица:

(5)
$$\begin{array}{l}
\mathbf{G}T_1; S_1, \quad S_2, \quad \ldots, S_m,\\
\mathbf{G}T_2; S_1T_2, S_2T_2, \ldots, S_mT_2,\\
\cdots\cdots\cdots\cdots\\
\mathbf{G}T_r; S_1T_r, S_2T_r, \ldots, S_mT_r,
\end{array}$$

където $T_1=S_1=E$. Очевидно два комплекса $\mathbf{G}T_i$, $\mathbf{G}T_j$, $i\neq j$ нямат общ елемент, така че ще имаме

$$rm=n!,\ r=\frac{n!}{m}.$$

11. *Следователно броят на различните стойности на полинома f е р·вен на индекса на групата му.* Да намерим групата на f_i. Нека S е субституция, която принадлежи на групата $\mathbf{G}_i$ на тази функция. T_i изменя f_1 и f_i; S не променя f_i и $T_i^{-1}f_i$ с f_1, т. е. субституцията

$$T_iS\,T_i^{-1}$$

не променя f_1, следователно ще бъде равна на някоя субституция S_k от $\mathbf{G}$, т. е.

$$T_i S\, T_i^{-1} = S_k.$$

Оттук лесно получаваме $S = T_i^{-1} S_k\, T_i$.

Следователно групата $\mathbf{G}_i$ на f_i *е равна на*

$$T_i^{-1}\, \mathbf{G}\, T_i.$$

Така например за функцията (1) имаме: с транспозицията $\sigma_2 = (23)$ функцията f взема нова стойност

$$f_2 = x_1x_3 + x_2x_4,$$

на която групата ще бъде

$$\sigma_2^{-1}\mathbf{A}\sigma_2 = E + (13) + (24) + (13)(24) + (12)(34) + \\ + (14)(23) + (1234) + (1432).$$

Със $\sigma_3 = (24)$ тя приема нова стойност

$$f_3 = x_1x_4 + x_2x_3,$$

на която групата ще бъде

$$\sigma_3^{-1}\mathbf{A}\sigma_3 = E + (14) + (23) + (14)(23) + \\ + (13)(24) + (12)(34) + (1342) + (1243).$$

Нека полиномът $f(x_1, x_2, \ldots, x_n)$ има r стойности:

(6) $$f_1, f_2, \ldots, f_r.$$

Да разгледаме една цяла рационална симетрична функция на стойностите (6):

$$S(f) = S(f_1, f_2, \ldots, f_r).$$

Ако заместим функциите f_i с изразите им спрямо x_k, така ще получим една рационална функция на променливите x_k, която ще бъде и симетрична, понеже разместването на тези променливи помежду им е еквивалентно с разместване на полиномите f_i, при което $S(f)$ не се изменя. Следователно $S(f)$ ще се изразява като цяла рационална функция на елементарните симетрични функции c_i на променливите x_k. Така получаваме теоремата:

12. *Всички стойности* $f_1, f_2, \ldots, f_r$ *на един полином*

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

ще бъдат корени на уравнението от r-та степен

$$f^r + a_1 f^{r-1} + \ldots + a_r = 0,$$

в което коефициентите a_i *са цели рационални функции на елементарните симетрични функции* $c_1, c_2, \ldots, c_n$ *на променливите* $x_1, x_2, \ldots, x_n$.

Интересно е да се знае какъв може да бъде броят r на стойностите на една рационална функция от n променливи. Или което, както видяхме, е все едно, какъв е индексът r на грпата от n-та степен. При

$n>4$ пръв Коши установи, че ако $r>2$, то и $r \geqq p$, гдето p е най-голямото просто число $\leqq n$. По-сетне Бертран (1845), основавайки се на един аритметичен постулат, доказа при $n>4$, че при $r>2$ имаме $r \geqq n$. Опростено доказателство, неизползуващо този постулат, даде Сере (1850).

2. Полиноми, принадлежащи на една група, и теорема на Лагранж. Видяхме, че на всеки полином принадлежи една група. Ще установим, обратно, че за всяка група **G** съществуват полиноми, които не изменят стойностите си само за субституции, принадлежащи на **G**. Да разгледаме линейната функция

$$u=\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_n x_n,$$

където числата $\alpha_1, \alpha_2, \ldots, \alpha_n$ са различни помежду си, а променливите $x_1, x_2, \ldots, x_n$ предполагаме, че са линейно независими, т. е. релация от вида

$$\beta_1 x_1+\beta_2 x_2+\ldots+\beta_n x_n=0 \tag{1}$$

е само тогава възможна, ако всичките числа $\beta_1, \beta_2, \ldots, \beta_n$ са равни на нула. Като разместваме променливите $x_1, x_2, \ldots, x_n$ по всички възможни начини, u ще приема все различни стойности, понеже предположението, че две стойности на u са равни, би довело до равенство от вида (1), в което поне едно от числата $\beta_1, \beta_2, \ldots, \beta_n$ е отлично от нула, което е невъзможно. Нека

$$\mathbf{G}=S_1+S_2+\ldots+S_m, \quad S_1=E$$

е групата и $u_1, u_2, \ldots, u_m$ са стойностите на u, като прилагаме субституциите на **G**. Тогава полиномът

$$\varphi(x_1, x_2, \ldots, x_n)=u_1 u_2 \ldots u_m$$

ще принадлежи на **G** и всяка субституция, която не принадлежи на **G**, ще променя алгебрически полинома φ. Ясно е, че можем да построим така безбройно много функции с това свойство. Съществува важна теорема на Лагранж, която дава връзката между отделните функции, която теорема гласи:

13. *Ако групата на една рационална функция ψ на променливите $x_1, x_2, \ldots, x_n$ съдържа групата на друга рационална функция φ, то ψ се изразява като полином на φ с коефициенти, които са симетрични рационални функции на променливите $x_1, x_2, \ldots, x_n$.*

Нека **G** е групата на функцията φ. Тогава, както видяхме, всички $n!$ субституции могат да се представят с

$$\mathbf{G}+\mathbf{G}T_2+\mathbf{G}T_3+\ldots+\mathbf{G}T_p,$$

гдето $T_2, T_3, \ldots, T_p$ изменят $\varphi_1=\varphi$ във $\varphi_2, \varphi_3, \ldots, \varphi_p$. Всички тези стойности на φ са корени на уравнението

$$F(z)=(z-\varphi_1)(z-\varphi_2)\ldots(z-\varphi_p)=0$$

с коефициенти, които са симетрични функции на променливите. Понеже групата на ψ съдържа G, то субституциите на тази група не ще я променят. Нека $\psi_2, \psi_3, \ldots, \psi_p$ са стойностите (не винаги всички различни), които взема $\psi_1=\psi$, като ѝ приложим субституциите $T_2, T_3, \ldots, T_p$. Ако $f_i(\varphi_i, \psi_i)$ е рационална функция на φ_i и ψ_i, то всяка симетрична функция на

$$f_1, f_2, \ldots, f_p$$

ще бъде симетрична и на променливите x, понеже на разместването на последните ще отговаря разместване на f_i. Но тогава сумата

$$\Phi(z)=\sum_{i=1}^{p}\psi_i\frac{F(z)}{z-\varphi_i}\frac{1}{F'(\varphi_i)}$$

ще бъде симетрична функция на променливите x, отгдето следва, че ще имаме

$$\Phi(z)=A_0z^{p-1}+A_1z^{p-1}+A_2z^{p-2}+\ldots+A_{p-1},$$

гдето A_k са симетрични функции на $x_1, x_2, \ldots, x_n$. По формулата на Лагранж при $z=\varphi_k$ имаме

$$\Phi(\varphi_k)=\psi_k,$$

с което теоремата е доказана.

От тази теорема следва, че ако две рационални функции имат една и съща група, то те се изразяват рационално една чрез друга, като коефициентите са симетрични функции. Обратно, ако две функции се изразяват рационално една друга т. е. $\psi=R(\varphi)$ и $\varphi=R_1(\psi)$, то групата им е една и съща, понеже субституциите, които не изменят едната, не изменят и другата.

Нека на независимите променливи $x_1, x_2, \ldots, x_n$ в полинома $f(x_1, x_2, \ldots, x_n)$ дадем n специални стойности $\zeta_1, \zeta_2, \ldots, \zeta_n$ (зависими променливи, числа и пр.); така полученият израз $f(\zeta_1, \zeta_2, \ldots, \zeta_n)$ ще наричаме полином на специалните стойности. Нека $f(x_1, x_2, \ldots, x_n)$ посредством субституцията S преминава в полинома

$$\varphi(x_1, x_2, \ldots, x_n).$$

Ако между специалните стойности ζ имаме връзката

$$f(\zeta_1, \ldots, \zeta_n)=\varphi(\zeta_1, \ldots, \zeta_n),$$

то полиномът на n-те специални стойности $f(\zeta_1, \ldots, \zeta_n)$ не ще изменя числено своята стойност. Може формално $f(x_1, \ldots, x_n)$ като полином на независимите променливи $x_1, \ldots, x_n$ да се измени.

Така например, ако ζ_1, ζ_2 са независими променливи, но

$$\zeta_3=\frac{1}{2}(\zeta_1+\zeta_2),$$

то изразът

$$\zeta_1-\zeta_3$$

е числено равен на изразите

$$\zeta_3-\zeta_2,\ 2\zeta_1+\zeta_2-3\zeta_3,\ \frac{1}{2}(\zeta_1-\zeta_2)$$

и други, които обаче формално (т. е. като считаме ζ_1, ζ_2, ζ_3 като независими променливи) са все различни. Лесно е да се провери, че само субституциите

$$E,\ (132)$$

не изменят числено ζ_1—ζ_3. Видяхме, че субституциите, които не изменят формално един полином, образуват група. Обаче същото не е въобще валидно за всички субституции, които числено не изменят един полином. Така субституциите (7) не образуват група. Имаме обаче теоремата

14. Нека **G** е една група от субституции и нека $\zeta_1, \zeta_2, \ldots, \zeta_n$ са n различни помежду си количества (независими променливи или специални стойности); тогава има полином $f(\zeta_1, \ldots, \zeta_n)$, който принадлежи на **G**, т. е. който не си променя числената стойност само за субституции, принадлежащи на **G**. Полиномът f може да се избере с коефициенти, които са цели числа.

Да разгледаме линейната форма на ζ:

$$\varphi = \zeta_1 + t\,\zeta_2 + t^2\zeta_3 + \ldots + t^{n-1}\zeta_n,$$

гдето t отначало не е определено. При всички възможни размествания на $\zeta\varphi$ ще приема $m = n!$ стойности:

$$\varphi_1, \varphi_2, \ldots, \varphi_m.$$

Условието две стойности φ_l и φ_k да бъдат равни се изразява с едно уравнение спрямо t най-много от $n-1$-ва степен:

$$\varphi_i - \varphi_k = 0, \tag{8}$$

в което не всички коефициенти са равни на нула. Всички такива уравнения (8) имат най-много

$$\frac{m(m-1)}{2}(n-1)$$

корени t_p, т. е. краен брой. Ако на t дадем стойност, равна на кое да е цяло число, отлично от корените t_p, то очевидно функцията φ ще приема все различни стойности.

Нека от субституциите на **G** функцията φ приема стойностите

$$\varphi_1, \varphi_2, \ldots, \varphi_g, \tag{9}$$

гдето g е редът на **G**. Ако върху (9) приложим субституции от **G**, то φ_i променят само реда си. Приложим ли друга субституция, непринадлежаща на **G**, то те се изменят. Ако следователно приложим върху полинома на x

$$F(x) = (x-\varphi_1)(x-\varphi_2)\ldots(x-\varphi_g)$$

субституции от **G**, той не се изменя. От други субституции той се изменя в други полиноми. Нека всички тези получени полиноми са

$$F_1(x), F_2(x), \ldots, F_q(x).$$

Тогава полиномите (броят на които е краен)

$$F(x)-F_1(x), F(x)-F_2(x), \ldots, F(x)-F_q(x)$$

не са идентично равни на нула и следователно можем на x да дадем както по-горе, стойност x_0, равна на цяло число, така че числата

$$F(x_0)-F_1(x_0),\ F(x_0)-F_2(x_0), \ldots, F(x_0)-F_0(x_0)$$

да бъдат всичките отлични от нула. Но тогава полиномът

$$F(x_0)=F(x_0;\ \zeta_1,\ldots,\zeta_n)$$

не се изменя числено само от субституцията на G и има цели коефициенти

3. Полиноми, принадлежащи на една линейна група. Да разгледаме случая, когато n е просто число. Нека r е примитивен корен на конгруенцията

$$x^{n-1}\equiv 1 \pmod n$$

и да разгледаме субституциите

$$T_i=|z\ r^i z|,\ i=1,\ 2,\ldots,\ n-1.$$

Тъй като остатъците от деленията на

$$r^i,\ i=1,\ 2,\ldots,\ n-1,$$

с n са числата от 1 до $n-1$, то очевидно T_i са субституциите

$$|z\,az|=S_{a,0},$$

гдето на a даваме стойностите 1, 2,..., $n-1$. Произведението на субституциите T_i със субституциите

$$S_{1,b}=|z\,z+b|,\ b=0,\ 1,\ 2,\ldots,\ n-1$$

дава линейната група от ред $n(n-1)$.

Сега ще построим полином на променливите

$$x_0,\ x_1,\ldots,x_{n-1},$$

който принадлежи на линейната група. Нека α е кой да е корен на уравнението $\frac{x^n-1}{x-1}=0$ и да образуваме полиномите

$$X_1=(x_0+\alpha x_1+\ldots+\alpha^{n-1}\ x_{n-1})^n,$$
$$X_2=(x_0+\alpha x_r+\ldots+\alpha^{n-1}\ x_{(n-1)r})^n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$X_{n-1}=(x_0+\alpha x_r{}^{n-2}+\ldots+\alpha^{n-1}\ x_{(n-1)r^{n-2}})^n.$$

Явно е, че T_i ги пермутират кръгово. За да видим влиянието на $S_{1,b}$, ще им дадем подходяща форма. Така X_{p+1}, даден с

$$X_{p+1}=(x_0+\alpha x_r{}^p+\ldots+\alpha^{n-1}x_{(n-1)r^p})^n,$$

преобразуваме по следния начин: да намерим такова число h, че да имаме $hr^p\equiv 1 \pmod n$. Оттук получаваме

$$h\equiv r^{n-1-p},$$

отгдето $\alpha^h=\alpha^{r^{n-1-p}}=\beta$, β е пак корен на $x^n=1$. Но тогава членът $\alpha^h x_{hr^p}$ става βx_1, а общо членът $\alpha^k x_{kr^p}$ става равен на $\beta^m x_n$, ако по-

ставим $kr^p \equiv m$, $k \equiv mr^{n-1-p}$, $\alpha^k = \beta^m$. Следователно X_{p+1} приема формата

$$X_{p+1} = (x_0 + \beta x_1 + \beta^2 x_2 + \ldots + \beta^{n-1} x_{n-1})^n,$$

гдето е ясно, че полиномите X_i не се променят от субституциите на $S_{1,b}$.

Тогава полиномът

$$M_i = (X_1 + \omega_i X_2 + \ldots + \omega_i^{n-2} X_{n-1})^{n-1},$$

гдето ω_i е корен на $x^{n-1} - 1 = 0$, ще принадлежи на линейната група, понеже субституциите от нея или не ще изменят полиномите X, или ще ги пермутират кръгово, при което M_i не ще се пак променя. Ако са известни стойностите на функциите M_i, $i = 1, 2, \ldots, n-1$, то от уравненията

$$X_1 + \omega_i X_2 + \ldots + \omega_i^{n-2} X_{n-1} = \sqrt[n-1]{M_i}$$

лесно намираме стойностите на X_k. Така ще имаме следната теорема:

15. Ако $x_0, x_1, \ldots, x_{n-1}$ са корени на алгебрично уравнение и полиномите

$$M_i, \; i = 1, 2, \ldots, n-1,$$

са алгебрично известни, то уравнението е решимо алгебрически.

Действително от уравненията

$$x_0 + \beta x_1 + \ldots + \beta^{n-1} x_{n-1} = \sqrt[n]{X_{p+1}}, p = 0, 1, 2, \ldots, n-2,$$

$$x_0 + x_1 + \ldots + x_{n-1} = -A,$$

гдето A е коефициент на x^{n-1} в даденото уравнение, по познат вече път намираме корените x_k.

4. Исторически бележки. След дълги опити на математиците да решат алгебрично уравненията от степен, по-висока от четири, в 17·8 г. италианският учен Паоло Руфини установява, че общите такива уравнения са нерешими алгебрично. Доказателството на Руфини не е било пълно. Строго доказателство на въпросната нерешимост дава Абел (1802—1829), прочут норвежки математик, който е дал основни работи в разни области от математиката. Отговор на открития въпрос за условията, на които трябва да удовлетворява едно уравнение, за да бъде решимо алгебрично, е дал Еварист Галоа (1811—1832), един от най-крупните математици на човечеството. Теорията на Галоа играе ръководна роля в съвременната математика.

По-нататъшното развитие на теорията на групите, основите на която са в работите на Галоа, е направено от Жордан, Силов, Фробениус и в най-ново време от съветските алгебристи начело с Шмит. Теорията на Галоа получава по-нататъшно развитие с фундаменталните резултати на съветските математици Н. Г. Чеботарев, Б. Н. Делоне и др.

Глава III

Теория на Галоа

1. Група на едно уравнение. Както ще видим по-нататък, на всяко уравнение в една дадена област на рационалност отговаря една група от субституции, която го характеризира в известно отношение. Именно от свойствата на тази група, откритието на която се дължи на Галоа, зависи алгебричното решение на уравнението. Нека

$$f(x)=0 \tag{1}$$

е уравнение от n-та степен с корени $x_1, x_2, \ldots, x_n$, които са все различни помежду си, и с коефициенти, принадлежащи на една област на рационалност R. Както видяхме (в теорема 21), съществува цяла рационална функция V от корените, например

$$V=a_1x_1+a_2x_2+\ldots+a_nx_n,$$

гдето a_i са рационални числа, която взема все различни стойности:

$$V_1, V_2, \ldots, V_m, m=n!, \tag{2}$$

когато разместваме корените помежду им по всевъзможни начини. Ще установим отначало една основна теорема, дадена от Галоа.

16. Корените $x_1, x_2, \ldots, x_n$ са рационални функции на V с коефициенти, принадлежащи на R.

Нека в една стойност V_1 на V да оставим x_1 на постоянно място, а да разместим $x_2, x_3, \ldots, x_n$ помежду им по всевъзможни начини. Така получаваме g на брой стойности на V_1, именно

$$V_1, V_2, \ldots, V_g, g=(n-1)!. \tag{3}$$

Тези стойности ще бъдат корени на уравнение, коефициентите на което ще бъдат симетрични функции на корените на

$$\frac{f(x)}{x-x_1}=0$$

и следователно ще бъдат рационални функции на x_1. Нека означим това уравнение с

$$\varphi(V, x_1)=0, \tag{4}$$

гдето φ е полином на V и x_1. Следва, че уравнението

$$\varphi(V_1, x)=0 \tag{5}$$

има общ корен x_1 с (1). Уравненията (5) и (1) не могат да имат други общи корени. Понеже ако x_2 удовлетворява (5), то бихме имали

$$\varphi(V_1, x_2)=0. \tag{6}$$

Но уравнението

$$\varphi(V, x_2)=0$$

има за корени стойностите на V, които получаваме, като заместим x_1 с x_2 и всички ще бъдат отлични от V_1, т. е. (6) е невъзможно. Общият най-голям делител на $f(x)$ и $\varphi(V_1, x)$ ще бъде от първа степен, отгдето следва, че x_1 ще бъде рационална функция на V_1. Подобно и $x_2, \ldots, x_n$ ще бъдат рационални функции на V_1, с което не само се доказва теоремата, но се дава и начин за намиране на тези рационални функции.

Нека сега

$$\Psi(V)=0 \tag{7}$$

е уравнението, на което корените са стойностите (2).

По теорема 19 коефициентите на това уравнение принадлежат на областта R. Понеже $V_2, \ldots, V_m$ са рационални функции на $x_1, \ldots, x_n$, които от своя страна са рационални функции на V_1, то корените на (7) са рационални функции на единия от тях.

Ако уравнението $\Psi(V)=0$ е разложимо в областта R, нека $\psi(V)$ е един неразложим множител на $\Psi(V)$ и $V_1, V_2, \ldots, V_\nu$ да бъдат корените на уравнението

$$\psi(V)=0. \tag{8}$$

Видяхме, че корените на (1) могат да се пишат

$$x_1=R_1(V_1),\ x_2=R_2(V_1), \ldots, x_n=R_n(V_1),$$

гдето R_i са рационални функции. Лесно се вижда също, че числата

$$R_1(V_k),\ R_2(V_k), \ldots,\ R_n(V_k), \tag{9}$$

гдето V_k е кой да е корен $V_1, V_2, \ldots, V_\nu$, също са корените $x_1, x_2, \ldots, x_n$, само че в друг ред. Действително, понеже $R_i(V_1)$ е корен на (1), т. е. $f[R_i(V_1)]=0$, то уравнението

$$f\ R_i(V)]=0$$

има общ корен V_1 с неразложимото уравнение (8), отгдето следва, че то ще се удовлетворява за всички негови корени. Редицата числа (9) са корени на (1). Остава да се докаже, че са различни помежду си. Нека допуснем обратното, например $R_i(V_k)=R_j(V_k)$. Уравнението

$$R_i(V)-R_j(V)=0$$

ще има общ корен V_k с (8) и следователно ще допуска за корени всички негови такива, отгдето специално ще имаме

$$R_i(V_1)-R_j(V_1)=0,$$

което е невъзможно, понеже $x_i \neq x_j$.

Така получихме, че корените $x_1, x_2, \ldots, x_n$ могат да се представят с кой да е хоризонтален ред на таблицата

$$
\begin{array}{l}
R_1(V_1),\ R_2(V_1),\ldots,R_n(V_1),\\
R_1(V_2),\ R_2(V_2),\ldots,R_n(V_2),\\
\cdots\cdots\cdots\cdots\cdots\cdots\\
R_1(V_\nu),\ R_2(V_\nu),\ldots,R_n(V_\nu).
\end{array}
\tag{10}
$$

Така получаваме ν пермутации на корените, които нека означим с A_1, $A_2,\ldots,A_\nu$. Две пермутации са виниги различни помежду си, понеже иначе би трябвало изразът

$$a_1x_1+\ldots+a_nx_n$$

да представлява едновременно две стойности на V. Нека с

$$S_1=E,\ S_2,\ldots,\ S_\nu \tag{11}$$

да означим субституциите, които сменят A_1 съответно с A_1, $A_2,\ldots,A_\nu$. Ще докажем, че субституциите (11) образуват група. Видяхме, че

$$V_k=\theta_k(V_1),$$

гдето θ_k е рационална функция. Заместването на V_1 в първата редица на (1) с $\theta_k(V_1)$ е еквивалентно на извършване на субституцията S_k. Значи S_k дава пермутацията

$$R_1(\theta_k V_1),\ R_2(\theta_k V_1),\ldots,\ R_n(\theta_k V_1),$$

гдето за простота е писано вместо $\theta_k(V_1)$ само $\theta_k V_1$. Ако сега извършим субституцията S_l, то трябва да заместим V_1 с $\theta_l V_1$. Така, прилагайки субституцията S_kS_l, от пермутацията A_1 ще получим пермутацията

$$R_1(\theta_k\theta_l V_1),\ R_2(\theta_k\theta_l V_1),\ldots,R_n(\theta_k\theta_l V_1).$$

Понеже $\psi(\theta_k V_1)=0$, то уравнението

$$\psi[\theta(V)]=0$$

има общ корен с неразложимото уравнение (8), т. е. ще допуска всички негови корени. Следователно

$$\psi[\theta_k(\theta_l V_1)]=0,$$

което показва, че $\theta_k\theta_l V_1$ е някой корен $\theta_s V_1$ на (8), отгдето заключаваме, че $S_kS_l=S_s$, т. е. субституциите (11) образуват група от ред ν. Тази група Galois нарича група на уравнението в дадената област на рационалност. Ние ще я бележим с **Г**.

2 Свойства на групата на едно уравнение. Групата на уравнението притежава свойство, което я характеризира напълно. Именно имаме теоремите:

17. Всяка рационална функция U от корените на уравнението (1), която не се променя числено от субституциите на групата **Г**, принадлежи на областта на рационалност R.

Функцията U може да се изрази рационално посредством един корен, например V_1 на (8), т. е.

$$U=\varphi(V_1),$$

гдето φ е рационална функция. Понеже U не си променя стойността от субституциите на Γ, които са еквивалентни на сменяване V_1 с V_1, $V_2,\ldots,\ V_\nu$, то ще имаме

$$U=\varphi(V_1)=\varphi(V_2)=\ldots=\varphi(V_\nu),$$

т. е.

$$U=\frac{1}{\nu}[\varphi(V_1)+\ldots+\varphi(V_\nu)],$$

което показва, че U е симетрична функция от корените на (8) и следователно ще бъде рационално известна, т. е. ще принадлежи на областта P.

18. Една рационална функция от корените на (1), която се изразява като число от областта R, остава числено непроменена от субституциите на групата Γ.

Нека изразим във функцията корените посредством V_1. Тогава, понеже функцията има числена стойност, принадлежаща на R, то ще имаме

$$\varphi(V_1)-a=0,$$

гдето φ е рационална функция, а a е число от R. Но тогава уравнението

$$\varphi(V)-a=0$$

има общ корен V_1 с уравнението (8) и последното е неразложимо; ще имаме

$$\varphi(V_i)-a=0,\ i=2,\ 3,\ldots,\ \nu,$$

т. е. $\varphi(V_1)=\varphi(V_2)=\ldots=\varphi(V_\nu)$, което показва, че φ не се изменя от субституциите на Γ.

19. Няма други субституции освен тези, които принадлежат на Γ, които не изменят числено всяка рационално известна функция от корените. Да разгледаме действително функцията

$$(\alpha-V_1)\ (\alpha-V_2)\ldots(\alpha-V_\nu),$$

гдето α е рационално число. Тя не се изменя за субституциите от Γ, но, както по-рано, виждаме, че може α така да се подбере, че тя да се изменя за всяка субституция, непринадлежаща на Γ.

20. Групата Γ е транзитивна или интранзитивна според това, дали уравнението $f(x)=0$ е неразложимо или разложимо.

Да допуснем, че групата Γ на неразложимото уравнение (1) е интранзитивна. Тогава Γ замества елемента x_1 с $x_1,\ x_2,\ldots,\ x_p$, гдето $p<n$. Тя ще пермутира тези корени помежду им. Действително нека

T замества x_q с x_m, $q \leqq p$, $m > p$, и нека U замества x_1 с x_q. Тогава Γ съдържа субституцията UT, която замества x_1 с x_m, т. е. x_m ще се намира между горните елементи. Всяка симетрична функция на

$$x_1, x_2, \ldots, x_p$$

не се изменя от Γ и следователно е рационално известна, отгдето следва, че $f(x)$ допуска за делител

$$(x - x_1)(x - x_2)\ldots(x - x_p)$$

с коефициенти, принадлежащи на областта R, което е невъзможно.

За да установим обратната част, нека допуснем, че $f(x)$ има делител

$$g(x) = (x - x_1)\ldots(x - x_r), \quad r < n$$

с рационално известни коефициенти. Ако Γ е транзитивна група, то има субституция S от нея, която сменява x с x_i, $i > r$. При прилагане на S множителят $x - x_1$ се изменя в $x - x_i$ и при подбрано рационално x $g(x)$ си променя числено рационалната стойност в R.

Ще направим една важна забележка относно рационалните функции от корените. Нека

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

е една такава функция. Да предположим, че тя запазва числено стойността, като ѝ приложим субституцията

$$S = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ p & q & r & \ldots & u \end{pmatrix},$$

т. е. ще имаме равенството

$$\varphi(x_1, x_2, x_3, \ldots, x_n) = \varphi(x_p, x_q, x_r, \ldots, x_u),$$

което не е необходимо да бъде изпълнено, ако x са независими променливи. Използувайки релациите между корените, можем на φ да дадем друга форма:

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

и тогава въобще няма да имаме

$$\psi(x_1, x_2, x_3, \ldots, x_n) = \psi(x_p, x_q, x_r, \ldots, x_u).$$

Значи, когато се говори, че една функция не променя числено стойността си при прилагане на една субституция S, трябва да се показва и формата, в която е взета. Това неудобство се избягва, ако S принадлежи на групата Γ. Действително, ако

$$h(x_1, x_2, \ldots, x_n) = \varphi(x_1, x_2, \ldots, x_n) - \psi(x_1, x_2, \ldots, x_n) = 0,$$

то h е рационална функция, която принадлежи на областта R, понеже има стойност нула. По теорема 18 тя не ще си измени стойността при прилагане на S, т. е. ще имаме

$$\varphi(x_p, x_q, \ldots, x_u) - \psi(x_p, x_q, \ldots, x_u) = 0.$$

Значи, ако φ не се изменя числено на S, то и ψ не се изменя. Следователно може да се абстрахираме от различните изрази, които приема една рационална функция от корените при прилагане на субституции, принадлежащи на групата Г.

Може да се изследва как се изменя групата, ако вместо $\psi(V)$ вземем друг неразложим множител на $\Psi(V)$. Доказва се, че така се получават конюговани на Г групи.

3. Примери за групи на Галоа. Да разгледаме общото уравнение от n-та степен:

$$x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n=0, \tag{12}$$

гдето на $a_1, a_2, \ldots, a_n$ гледаме като на независими променливи. Областта на рационалността R_0 се състои от всички рационални функции на $a_1, a_2, \ldots, a_n$ с коефициенти, които са рационални числа. Ще докажем, че групата на Галоа Г е симетричната група.

Разглеждаме една коя да е рационална функция от корените x_i, която е равна на число нт R_0, т. е. на една рационална функция на $a_1, a_2, \ldots, a_n$. Така ще имаме релация от вида

$$F(x_1, x_2, \ldots, x_n, a_1, a_2, \ldots, a_n)=0,$$

гдето F е рационална функция спрямо x и a. Ако заместим коефициентите a посредством корените x, то така получаваме едно тъждество спрямо независимите променливи

$$x_1, x_2, \ldots, x_n,$$

което естествено ще остава в сила при произволното им пермутиране. С това се вижда, че горната релация остава в сила при прилагане на всяка субституция, т. е. групата Г е симетрична група.

Ако групата на едно уравнение не е симетричната група, то между корените на уравнението ще има рационална връзка в областта на рационалност, която не остава в сила за всички субституции, които извършваме върху корените. Действително имаме

$$\psi(a_1x_1+a_2x_2+\ldots+a_nx_n)=0,$$

понеже $\psi(V_1)=0$. Функцията $\psi(a_1x_1+\ldots+a_nx_n)$, на която числената стойност е нула, не се променя от субституциите на Г, но се променя за всяка друга субституция, понеже на такава отговаря стойност на V, която не е корен на (8). Следователно горната релация не е едно тъждество спрямо корените x, което доказва нашето твърдение.

Обратно, ако имаме релацията между корените

$$\varphi(x_1, x_2, \ldots, x_n)=0,$$

която не остава в сила за всички субституции, то лесно се вижда, че групата на Галоа Г не е симетричната група. Действително, ако заместим корените x посредством V, то горната релация, като използуваме уравнението (7), дава уравнение

$$\psi_1(V)=0,$$

степента на което е по-малка от $n!$, понеже иначе φ не би се изменила за всички субституции. Резолвентното уравнение $\psi(V)=0$ ще има неразложим делител от степен, по-малка от $n!$.

Нека да се върнем към общото уравнение (12) с произволни коефициенти, които са независими променливи. Нека към областта на рационалност R_0, съставена от всички рационални функции на параметрите a, адюнгираме $\sqrt{D}$, гдето D е дискриминантата на уравнението. Ще докажем, че в така разширената област R_1 на рационалност групата на Галоа е алтернативната група A. Действително да разгледаме една коя да е рационална функция от корените на (12), която е равна на число от R_1. Така ще получим релация от вида

$$F(x_1, x_2, \ldots, x_n, a_1, a_2, \ldots, a_n, \sqrt{D})=0,$$

гдето F е рационална функция спрямо x, a_k и $\sqrt{D}$ с рационални коефициенти. Ако изразим a_k, $\sqrt{D}$ посредством корените x_i, то се получава тъждество, което остава естествено валидно за всички субституции на $x_1, x_2, \ldots, x_n$. Субституциите от A не изменят $\sqrt{D}$, a_k, т. е. F остава непроменено от тях. Следователно групата на Галоа съдържа групата A. Всяка субституция, която не принадлежи на A, не може да се съдържа в групата на Галоа, понеже функцията $\sqrt{D}$ принадлежи на областта на рационалност R_1 и като полином на x_i променя знака си при прилагане на S.

4. Друго определение на група на едно уравнение. Нека

$$f(x)=0$$

е едно уравнение и P е числово поле (област на рационалност) към което принадлежат коефициентите му. Да означим с Ω полето на разлагане на полинома $f(x)$, т. е. това поле, което се получава от P с присъединяване корените $x_1, x_2, \ldots, x_n$ на уравнението $f(x)=0$. Предполагаме естествено, че въпросното уравнение няма многократни корени, което не е ограничение, понеже с рационални действия всяко уравнение с многократни корени се свежда към такова със само прости корени. Видяхме, че всяка рационална функция от корените на уравнението $f(x)=0$, която има стойност, равна на число от полето P, не се изменя при прилагане субституциите, които принадлежат на групата на уравнението $f(x)=0$. Това характерно свойство на групата на уравнението позволява да се даде едно видоизменено определение на тази група. Под автоморфизъм s на полето Ω относно P разбираме такъв изоморфизъм на Ω в себе си, при който всеки елемент a от P остава неизменен, т. е. елементът a отговаря на себе си. Нека при автоморфизма s на елемента α от Ω отговаря елементът α'. Съгласно с изискването това съответствие е еднозначно и в обратна посока, т. е. елементът α' е съответен само на елемента α. Нека сега на α' при автоморфизма t на полето Ω отговаря елементът α'' от Ω. Тогава съответствието $\alpha \to \alpha''$ е също така взаимно еднозначно. Лесно се вижда, че това съответствие е също автоморфизъм на полето Ω, който ще

наричаме произведение на автоморфизмите s и t и ще бележим с st. Действително автоморфизмите s и t не изменят елементите на P и следователно, приложени последователно, те също не изменят тези елементи. Освен това, ако на елементите α и β съответствуват елементите α'' и β'', то на $\alpha+\beta$ съответствува $\alpha''+\beta''$ и на $\alpha\beta - \alpha''\beta''$. Ще установим следната теорема:

21. Съвкупността A от автоморфизмите на полето Ω образува относно въведеното умножение група, която съвпада до изоморфизъм с групата на уравнението $f(x)=0$.

Именно ще установнм, че на всеки автоморфизъм съответствува еднозначно и обратимо една субституция върху корените на уравнението и на произведението на два автоморфизма съответствува произведението на съответните суоституции. Ще означаваме автоморфизмите за разлика от субституциите с неглавни букви $s, t, \ldots$ Нека автоморфизмът s преобразува корена x_k на уравнението $f(x)=0$ в елемента γ. Тогава γ принадлежи на полето Ω и понеже $f(x_k)$ като равно на нула принадлежи на полето P, то трябва да имаме $f(\gamma)=0$. Следователно γ е също корен на уравнението $f(x)=0$. При това два различни корена се трансформират пак в различни корени, защото в противен случай два елемента биха се трансформирали в един и същ елемент, което противоречи на условието на еднозначност в двете посоки. Следователно корените $x_1, x_2, \ldots, x_n$ ще се трансформират в автоморфизма s в същите корени, но въобще в друг ред $x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n}$. Ясно е тогава, че всяка релация в полето

$$\varphi(x_1, x_2, \ldots, x_n)=0$$

между корените на уравнението $f(x)=0$ автоморфизмът s превежда в релацията

$$\varphi(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n})=0.$$

Да означим с S субституцията

$$S=\begin{pmatrix}1 & 2 & 3 & \ldots & n\\ \mu_1 & \mu_2 & \mu_3 & \ldots & \mu_n\end{pmatrix}.$$

Поставяме тогава в съответствие на автоморфизма субституцията

$$s \to S.$$

Нека t е друг автоморфизъм, на който отговаря също така субституцията S. Тогава t трябва да преобразува корените $x_1, x_2, \ldots, x_n$ съответно в корените $x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n}$. Но всеки елемент α от полето Ω представлява цяла рационална функция от корените $x_1, x_2, \ldots, x_n$ с коефициенти P, $\alpha=\psi(x_1, x_2, \ldots, x_n)$. С автоморфизмите s и t елементът α преминава в един и същ елемент, а именно $\psi(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n})$ и всеки елемент от полето P остава непроменен. Следователно $s=t$, т. е. на различни автоморфизми отговарят различни субституции. Може

лесно да се види, че за всяка субституция S от групата Г има автоморфизъм s от полето Ω такъв, че $s \to S$. Ако

$$S = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \mu_1 & \mu_2 & \mu_3 & \ldots & \mu_n \end{pmatrix},$$

то кой да е елемент $\alpha = g(x_1, x_2, \ldots, x_n)$ от Ω преминава в елемента $\alpha' = g(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n})$ при прилагане на S. При това, ако изразим този елемент по някой друг начин от корените $x_1, x_2, \ldots, x_n$, т. е. $\alpha = h(x_1, x_2, \ldots, x_n)$, където $h(x_1, x_2, \ldots, x_n)$ означава също така цяла рационална функция с коефициенти от P, то субституцията S привежда елемента α пак в елемента α'. Това се дължи на факта, че субституцията S принадлежи на групата на уравнението Г и при прилагането ѝ връзката

$$g(x_1, x_2, \ldots, x_n) = h(x_1, x_2, \ldots, x_n)$$

между корените на $f(x) = 0$ не се променя. Така виждаме, че независимо от начина на изразяването на елемента α чрез корените $x_1, x_2, \ldots, x_n$ субституцията S привежда α в α', т. е. съответствието $\alpha \to \alpha'$ не зависи от въпросния начин. Елементите от полето P остават непроменени при прилагане на S. Друг елемент β от Ω, $\beta \to \alpha'$ не съществува. Действително в противен случай субституцията S^{-1}, която също принадлежи на Г, би привеждала елемента α' в два различни елемента α и β, което противоречи. Лесно се вижда, че за всеки елемент α' от Ω отговаря елемент α от същото поле, който с прилагане на субституцията S преминава в α'. Именно α е елементът, в който преминава α' при прилагане на субституцията S^{-1}. От всичко гореказано следва, че съотношението $\alpha \to \alpha'$ представлява взаимно еднозначно изобразяване на полето Ω в себе си, като при това изобразяване елементите на полето P остават непроменени. Това съответствие е и автоморфизъм s на Ω, на който отговаря субституцията S. Наистина, ако

$$\alpha = g(x_1, x_2, \ldots, x_n) \to \alpha' = g(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n}),$$

$$\beta = h(x_1, x_2, \ldots, x_n) \to \beta' = h(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n}),$$

то субституцията привежда $\alpha + \beta = g(x_1, x_2, \ldots, x_n) + h(x_1, x_2, \ldots, x_n)$ в сумата $g(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n}) + h(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n}) = \alpha' + \beta'$ и произведението $\alpha\beta$ в $\alpha'\beta'$, т. е. имаме $\alpha + \beta \to \alpha' + \beta'$ и $\alpha\beta \to \alpha'\beta'$. С това се убеждаваме, че съответствието $s \to S$ е взаимно еднозначно съответствие между съвкупността А и групата Г.

Нека сега s и t са произволни автоморфизми, на които отговарят субституциите S и T:

$$S = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \mu_1 & \mu_2 & \mu_3 & \ldots & \mu_n \end{pmatrix}, \quad T = \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 & \ldots & \mu_n \\ \nu_1 & \nu_2 & \nu_3 & \ldots & \nu_n \end{pmatrix}.$$

Ако върху произволен елемент $\alpha = g(x_1, x_2, \ldots, x_n)$ от Ω приложим субституцията S, то той преминава в елемента $\alpha' = g(x_{\mu_1}, x_{\mu_2}, \ldots, x_{\mu_n})$ и с прилагане на субституцията T последният елемент преминава в елемента $\alpha'' = g(x_{\nu_1}, x_{\nu_2}, \ldots, x_{\nu_n})$. Но очевидно елементът α'' може да се получи от α с прилагане на субституцията

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \nu_1 & \nu_2 & \nu_3 & \ldots & n \end{pmatrix} = ST.$$

Следователно на произведението st на автоморфизмите s и t отговаря произведението ST на съответните им субституции. Така достигаме до заключението, че групата автоморфизми **A** на полето Ω и групата **Г** на уравнението $f(x)=0$ са изоморфни. Това ни дава основание да считаме тези групи за неразлични и при построяване на теорията на Галоа в голяма част от нея да си служим с групата **A** от автоморфизми.

5. Понижаване на групата на едно уравнение с адюнгиране на рационална функция от корените. Видяхме, че редът ν на групата **Г** на Галоа се даваше със степента на неразложимото уравнение за V. Ако към областта на рационалност R адюнгираме нови ирационалности, то може да се случи това уравнение да стане разложимо в разширената област R_1. Понеже корените на новото уравнение са между числата $V_1, \ldots, V_\nu$, то новата група на Галоа, отговаряща на областта R_1, ще бъде подгрупа на **Г**.

22. **Нека Г е групата на уравнението $f(x)=0$ и φ е една произволна рационална функция на корените. Субституциите от Г, които не изменят числено φ, образуват група Н, която ще бъде групата на уравнението, когато към областта на рационалност се адюнгира φ.**

Явно е, че субституциите от **Г**, които не изменят числено φ, образуват група. При присъединяването на φ групата на уравнението не може да съдържа субституции освен такива, които са от **Г**, и понеже те не трябва да изменят φ, то очевидно, че те ще бъдат субституции от **Н**. Ще покажем, че групата на уравнението ще съдържа всички субституции **Н**, т. е. ще се слива с нея. Действително нека рационалната функция $\varphi(x_1, x_2, \ldots, x_n)$ се изразява рационално в разширената област на рационалност

$$\psi = R(\varphi).$$

Понеже функцията $\psi - R(\varphi)$ като равна числено на нула принадлежи на първоначалната си област на рационалност, то като приложим кой да е субституция от **Н**, ще имаме

$$\psi_1 - R(\varphi) = 0,$$

отгдето $\psi_1 = \psi$. Значи всяка рационална функция от корените на уравнението, която е в разширена област на рационалност и има числената си стойност от нея, не се изменя от субституциите на **Н**. Групата на уравнението **Г** значи се редуцира на **Н**.

По-общо присъединяване към областта на рационалност на рационалните функции

$$\varphi_1, \varphi_2, \ldots, \varphi_p$$

от корените на уравнението редуцира групата Г на подгрупата Н, която не изменя числената стойност в R на всичките функции φ_i.

По-рано доказахме теорема на Лагранж, която се отнасяше до групата от субституции, които не изменят алгебрически една функция. Съществува аналогична теорема за числената стойност на функциите.

23. **Нека една рационална функция φ от корените да запазва числената си стойност за групата субституции Н, която е подгрупа на групата Г на уравнението. Тогава всяка рационална функция ψ от корените, която запазва числовата си стойност за субституциите на Н, се изразява рационално от φ.**

Действително, ако присъединим към първоначалната област на рационалност R функцията φ, то по теорема 22 групата на уравнението става Н. Тогава по теорема 13 функцията ψ ще има стойност от разширената област, т. е. ще бъде рационална функция на φ с коефициенти, принадлежащи на първоначалната област.

Сега ще установим някои теореми, които ще имат приложение нататък.

24. **Редът на групата на едно неразложимо уравнение, на което корените са рационални функции на един от тях, е равен на степента му.**

Нека даденото уравнение да бъде $f(x)=0$ с корени $x_1, x_2, \ldots, x_n$ и

$$V=a_1x_1+\ldots+a_nx_n$$

е резолвентната функция на Галоа. Понеже $x_2, x_3, \ldots, x_n$ са рационални функции на x_1, то $V=R(x_1)$, гдето $R(x)$ е рационална функция. Уравнението

$$[V-R(x_1)]\,[V-R(x_2)]\ldots[V-R(x_n)]=0$$

е с коефициенти, принадлежащи на дадената област на рационалност, и е от n-та степен. Понеже редът на групата Г на $f(x)=0$ се дава от степента на неразложимото уравнение, на което V е корен, ясно е, че той не е по-голям от n и понеже групата по теорема 20 е транзитивна, то редът е най-малко равен на n, т. е. ще бъде точно n.

25. **Обратно, ако групата на уравнението е транзитивна и ако редът ѝ е равен на степента, то всички корени се изразяват рационално посредством единия от тях.**

Понеже групата на уравнението Г е транзитивна и редът ѝ е равен на степента, то ще има една субституция, именно E, която не изменя елемента x_1 например. Ако към областта на рационалност присъединим x_1, то групата на уравнението ще се редуцира на Е. Ре-

золвентното уравнение ще бъде от първа степен, т. е. V ще бъде рационална функция на x_1, откъдето следва, че и $x_2, \ldots, x_n$ ще бъдат рационални функции на x_1.

Да разгледаме случая, когато n е просто число. Тогава уравнението като абелево от проста степен ще е решимо алгебрически. В това се убеждаваме лесно и така. Групата на уравнението Г ще бъде циклична. Действително нека S е субституция от Г, отлична от E. Редът на S като делител на n ще бъде равен на n, понеже това число е просто. Да разложим S на цикли без общи елементи:

$$S=C_1C_2\ldots C_k,$$

гдето C_i е от ред n_i и $n_1+n_2+\ldots+n_k=n$. Понеже редът на S е най-малкото общо кратно на числата n_i, то заключаваме, че едно от числата n_i е равно на n, а другите са равни на нула, т. е. S е циклична субституция.

Групата Г ще бъде

$$E,\ S,\ S^2,\ S^3,\ \ldots,S^{n-1}.$$

Всяка циклична функция

$$(x_1+\alpha x_2+\alpha^2x_3+\ \ldots+\alpha^{n-1}x_n)^n,\ \alpha^n=1,$$

ще бъде рационално известна, откъдето и корените x_k ще се намерят като алгебрически функции от познати величини, т. е. уравнението е решимо алгебрически.

6. Адюнгиране на корени на помощни уравнения. Нека към областта на рационалност R присъединим корен z_1 на неразложимото уравнение

$$F(z)=0 \tag{13}$$

и да допуснем, че в така разширената област на рационалност резолвентното уравнение на Галоа

$$\psi(V)=0 \tag{14}$$

става разложимо. Тогава, ако $z_1, z_2, \ldots, z_k$ са всички корени на (13) и $\varphi(V, z_1)$ е един неразложим делител на $\psi(V)$, то ще установим, че $\psi(V)$ се дели на полиномите

$$\varphi(V, z_2), \ldots, \varphi(V, z_k).$$

Затова нека разделим $\psi(V)$ на $\varphi(V, z)$ и да означим частното и остатъка от делението съответно с Q и $R(V, z)$, които са полиноми на V с коефициенти рационални функции на z. Значи ще имаме

$$\varphi(V)=\varphi(V, z)Q+R(V, z),$$

гдето

$$R(V, z)=a_0(z)\,V^m+a_1(z)\,V^{m-1}+\ldots+a_m(z),$$

$a_i(z)$ са рационални функции на z. При $z=z_1$ по условие имаме $R(V, z_1)=0$, т. е. $a_g(z_1)=0$, $g=0, 1, 2, \ldots, m$. Понеже (13) е неразложимо уравнение, то следва, че и

$$a_g(z_i)=0,\ i=1,\ 2, \ldots, k,$$

т. е. $R(V, z_i)=0$, с което се доказва, че $\psi(V)$ се дели и на $\varphi(V, z_i)$.

Произведението

$$H(V)=\varphi(V, z_1)\varphi(V,z_2)\ldots\varphi(V, z_k)$$

е с коефициенти, които принадлежат на първоначалната област на рационалност и корените на уравнението

$$H(V)=0$$

принадлежат на (14). Понеже последното уравнение е неразложимо, то лесно следва, че ще имаме

(15) $$H(V)=[\psi\ (V)]^q,$$

като и в двата полинома сме взели коефициентите пред нйй-високата степен равни на единица. Във верността на (15) се убеждаваме лесно. Именно нека най-високата степен на $\psi(V)$, която дели $H(V)$, да бъде m-тата, т. е. да имаме

$$H(V)=h(V)[\psi(V)]^m.$$

Понеже корените на $h(V)=0$ са такива и на (14), то трябва $h(V)$, ако не е константа, да се дели на $\psi(V)$, т. е. m няма да бъде най-високата степен. Значи $h(V)$ е константа, която може да се приеме равна на 1.

Ако две уравнения

(16) $$\varphi(V, z_1)=0,\quad \varphi(V, z_2)=0$$

имат общ корен, то и всичките им корени ще бъдат общи. Нека общият им корен бъде V_1. Понеже всички корени на (16) са рационални функции на единия, то първото уравнение (16) ще има корен $V_2=\theta(V_1)$, гдето $\theta(x)$ е рационална функция, която може да се счита и полином.

Уравнението

$$\varphi[\theta(V), z_1]=0$$

има общ корен V_1 с неразложимото първо уравнение (16), т. е. полиномът

$$\varphi[\theta(V), z]$$

за $z=z_1$ се дели точно на $\varphi(V, z)$ и следователно аналогично на по-рано заключаваме, че ще се дели и за $z=z_2$. С това получаваме, че

$$(\varphi[\theta(V_1), z_2]=\varphi(V_2; z_2)=0,$$

т. е. V_2 е също корен на второто уравнение (16). Значи две уравнения като (16) са или идентични, или нямат нито един общ корен. Понеже

всеки корен на (14) трябва по релацията (15) да бъде корен на q полинома φ, то заключаваме, че полиномите φ по q се повтарят, т. е. има от тях само $r=\frac{k}{q}$ на брой различни.

С $\varphi(V, z_1)$ означихме един неразложим делител на $\psi(V)$ в областта на z_1. Лесно е да установим, че и $\varphi(V, z_i)$ е неразложим в областта на z_i. Понеже, ако допуснем, че $\varphi(V, z_i)$ се дели на $\varphi_1(V, z_i)$, то лесно, както по-горе, установяваме, че $\varphi(V, z_1)$ се дели на $\varphi_1(V, z_1)$, което е невъзможно. Така получаваме, че корените z_i се разпределят на r системи по q корена, така че групата Г се свежда на различни редуцирани групи, ако към областта на рационалност адюнгираме корени от разни системи, докато при адюнгиране на корени от една и съща система се получават еднакви редуцирани групи. В случай, че степента k на помощното уравнение е просто число, то $q=1$ и редът на редуцираната група на Галоа ще бъде равен на реда на Г, делен с k.

Ако адюнгираме към областта на рационалност на една помощна рационална функция ω от корените, то така адюнгираме корен на уравнението, което удовлетворява ω. Така че такова адюнгиране е в същност еквивалентно на горното.

Нека сега $\mathbf{H}_1$ и $\mathbf{H}_2$ са групите на Галоа, на които се редуцира Г, когато към областта на рационалност адюнгираме корени z_1 и z_2 от различна система. Нека V_1 и V_2 са два корена съответно на

$$\varphi(V, z_1)=0, \quad \varphi(V, z_2)=0. \tag{17}$$

Видяхме, че ако корените на първото уравнение са

$$V_1, \theta_1(V_1), \theta_2(V_1)\ldots,$$

то корените на второто ще бъдат

$$V_2, \theta_1(V_2)\ \theta_2(V_2), \ldots,$$

гдето θ са рационални функции. Нека S_m е субституция от $\mathbf{H}_1$, която замества V_1 с $\theta_m(V_1)$, а S_m' е субституция от $\mathbf{H}_2$, заместваща V_2 с $\theta_m(V_2)$. Тогава, ако T е субституция от Г, която сменява V_1 с V_2, очевидно ще имаме

$$TS_m'=S_mT,$$

понеже и субституциите в двете части сменяват едновременно V_1 с $\theta_m(V_2)$. Оттук получаваме

$$T^{-1}S_mT=S_m', \text{ т. е. } T^{-1}\mathbf{H}_1T=\mathbf{H}_2.$$

Значи субституцията T трансформира групата $\mathbf{H}_1$ в групата $\mathbf{H}_2$. С това се доказва, че групите Н, на които се редуцира групата Г на уравнението с адюнгиране на корени на уравнението (13), са конюговани на една от тях.

Нека допуснем, че корените на (13) са рационални функции на единия от тях z_1, коефициентите на които функции, разбира се, принадлежат на първоначалната област на рационалност. Ако адюнгираме

корена z_1, то с това са адюнгирани и всички други корени. Следователно групите $\mathbf{H}_1$, $\mathbf{H}_2$,... съвпадат с една група $\mathbf{H}$, която по горното предложение ще се трансформира със субституциите на Γ сама в себе си, т. е. $\mathbf{H}$ е и. пг. на Γ. Така получаваме теоремата:

2·. Ако корените на помощното уравнение са рационални функции на единия от тях и ако адюнгирането им редуцира групата Γ на $\mathbf{H}$, то $\mathbf{H}$ е една и. пг. на Γ. Редът на $\mathbf{H}$ е равен на реда на Γ, делен със степента на помощното уравнение, в случай, че то е просто число.

Сега ще установим една теорема, която решава обратната задача.

27. Ако групата на уравнението Γ от ред $g=kq$ съдържа и. пг. $\mathbf{H}$ от ред q, то Γ може да се редуцира на $\mathbf{H}$ с адюнгиране на корените на едно абелево уравнение от степен k.

Нека

$$\mathbf{H}=S_1+S_2+\ldots+S_q,\ S_1=E$$

и

$$\Gamma=\mathbf{H}T_1+\mathbf{H}T_2+\ldots+\mathbf{H}T_k,\quad T_1=E.$$

Нека $V_1, V_2, \ldots, V_q$ са стойностите на V, които се получават от V с прилагане на субституциите на $\mathbf{H}$, и нека t е цяло число, така подбрано, че функцията

$$\theta=(t-V_1)(t-V_2)\ldots(t-V_q),$$

която остава непроменена за субституциите на $\mathbf{H}$, се изменя за всички други субституции от Γ. Тогава, ако адюнгираме към областта на рационалност θ, групата Γ на уравнението ще се редуцира на тая своя подгрупа. Лесно ще видим, че θ се дава като корен на едно абелево уравнение. Нека $\theta_1, \theta_2, \ldots, \theta_k$ са стойностите на θ, които се получават, като в θ_1 се извършват субституциите

$$T_1, T_2, \ldots, T_k.$$

Нека сега F е една симетрична функция на $\theta_1, \theta_2, \ldots, \theta_k$. Понеже $\mathbf{H}$ е и. пг. на Γ, то функциите θ ще остават непроменени от субституцията S. Функцията F не ще се изменя за всички субституции $S_\alpha T_\beta$, понеже S_α не изменя θ, а T само сменява функциите θ една с друга. Действително имаме например

$$T_\beta\theta_i=T_\beta T_{i-1}\theta_1=T_\gamma S_\delta\theta_1=T_\gamma\theta_1=\theta_{\gamma+1},$$

понеже $T_\beta T_{i-1}$ е субституция от Γ и може да се представи от формата $S_a T_b$, която има форма $T_\gamma S_\delta$, понеже $\mathbf{H}$ е и. пг. на Γ.

Следователно функциите θ са корени на уравнението

$$(x-\theta_1)(x-\theta_2)\ldots(x-\theta_k)=0, \tag{18}$$

на което коефициентите принадлежат на първоначалната област на рационалност, като функции, които не се променят за субституциите

на групата Г. Както по-рано, лесно се убеждаваме, че уравнението (18) е неразложимо. Понеже $\theta_1, \theta_2, \ldots, \theta_k$ не се изменят от субституциите S на Н, то всички те са рационални функции на единия от тях, т. е. (18) е абелево уравнение. Редът на групата му по теорема 26 е равен на k. Ако k е просто число, абелевото уравнение (18) е решимо алгебрически. Така получаваме, че ако индексът на подгрупата Н на Г е просто число, то с присъединяване на радикали можем да сведем групата на уравнението Г на и. пг. Н.

Ще изследваме сега въпроса за адюнгиране на корени на няколко помощни уравнения. Това можем да сведем към адюнгиране на корени на едно уравнение. Именно имаме предложението: ако са дадени няколко различни корена на уравнения с коефициенти, принадлежащи на една област на рационалност R, можем да ги изразим рационално посредством един само корен V на едно уравнение с коефициенти в R.

Действително нека да имаме например две уравнения с коефициенти, принадлежащи на областта R:

$$h(x)=0, \quad g(x)=0, \tag{19}$$

съответно с корени $x_1, x_2, \ldots, x_n$; $y_1, y_2, \ldots, y_m$. Тези числа ще бъдат прости корени на уравнение

$$F(x)=0$$

с коефициенти в R, което се получава лесно. По теорема 16 числата x_i и y_i ще се изразяват рационално посредством V, което е корен на уравнение с коефициенти в R.

7. Обща теорема на Галоа за алгебрична разрешимост. На основание на предните теореми лесно ще намерим условието едно дадено уравнение да бъде решимо алгебрически. Като излизаме от първоначалната област на рационалност, то ако уравнението е решимо алгебрично, с присъединяване на радикали ще достигнем до област на рационалност, в която ще принадлежат корените му. Тогава групата му ще се сведе на групата единица Е. Очевидно адюнгирането на радикала е еквивалентно с адюнгиране на корени на биномни уравнения от прости степени.

Нека $f(x)=0$ е уравнението и Г е групата му от ред m. Да предположим, че p е най-малкото просто число, така че адюнгирането на корени на биномното уравнение

$$r^p = A_i \tag{20}$$

редуцира Г на своя подгрупа Γ_1. A_i принадлежи на област на рационалност, получена от първоначалната с адюнгиране на корени на биномни уравнения от по-ниска степен:

$$r^\alpha = A_1, \; r^\beta = A_2, \; \ldots, \; r^\tau = A_{t-1},$$

като A_g принадлежи на област на рационалност, получена от първоначалната евентуално с адюнгиране на корени на предшествуващите в тази редица биномни уравнения.

Може да считаме, че в разширената така област сме адюнгирали корените на $x^p=1$, тъй като, както ни е известно, те се получават от биномни уравнения, на които степените са делители на $p-1$. Но понеже корените на (20) са рационални функции на единия от тях r_1 като равни на

$$r_1,\ \omega r_1,\ \ldots,\ \omega^{p-1}r_1,$$

гдето ω е примитивен корен на $x^p=1$, то по теорема 26 групата се редуцира на една своя и. пг. Γ_1 с индекс p, което е просто число, т. е. Γ_1 е м. и. пг. По-нататък с присъединяване на нови корени на биномни уравнения групата Γ_1 ще се сведе до своя м. и. пг. Γ_2 с индекс, равен на просто число, и т. н., докато стигнем до групата единица E. От това следва, че за да бъде едно уравнение решимо алгебрически, трябва за групата му Γ да съществува един ред на разлагане

$$\Gamma,\ \Gamma_1,\ \Gamma_2,\ \ldots,\ \Gamma_{q-1},\ \Gamma_q=E, \tag{21}$$

така че всяка група Γ_i е м. и. пг. на Γ_{i-1} с индекс просто число, т. е. Γ да е метациклична група. Редът от индексите на разлагане, както ни е известно по теоремата на Жордан, е един и същ за различните разлагания на Γ.

Нека сега, обратно, да допуснем, че Γ е метациклична група и (21) е един неин ред на разлагане. Тогава, ако m и m_1 са редовете на Γ и Γ_1 по теорема 26 с адюнгиране на корен на абелевото уравнение (18) от степен $\frac{m}{m_1}$, равна на просто число, групата Γ се свежда на Γ_1. Уравнението (18) като абелево от проста степен е решимо алгебрически. Значи с адюнгиране на радикали Γ може да се свежда на Γ_1. Продължавайки така, очевидно ще достигнем до групата единици и уравнението ще се реши алгебрически. Така получихме основната теорема на Galois:

28. Необходимото и достатъчно условие едно уравнение да се реши алгебрически се състои в това групата му да е метациклична.

От тази теорема веднага получаваме теоремата на Абел за алгебрична нерешимост на уравненията от степен, по-висока от четири. Видяхме в глава I, § 6, че симетричната група от n елемента при $n=3,4$ е метациклична, отгдето следва алгебрическата решимост на уравненията от трета и четвърта степен. При $n>4$ индексите на разлагането са числата 2 и $\frac{n!}{2}$, което не е просто число, отгдето следва поменатата теорема на Абел.

От теоремата на Галоа следва, че необходимо и достатъчно условие едно уравнение да бъде решимо с квадратни радикали се състои в това, индексите на разлагането на групата му да са равни на две.

Оттук веднага се получава, че степента на едно неразложимо такова уравнение трябва да има формата 2^λ. Понеже групата му Γ трябва

да бъде от ред 2^n и от неразложимостта следва, че Г е транзитивна, а по теорема 17 трябва степента на уравнението да дели 2^m, т. е. да има форма 2^λ.

Глава IV

Приложение на теорията на Галоа

1. Уравнения на Галоа. Когато уравнението е от степен просто число, условията за разрешимост се трансформират, както показва това Galois, в условия, отнасящи се до връзки между корените му. Отначало ще установим една помощна теорема:

29. Ако едно уравнение $f(x)=0$ от степен n, която е просто число, става разложимо с присъединяване на корени на неразложимото уравнение $F(r)=0$ от степен m, на което корените са рационални функции на едните от тях, то m се дели на n.

Действително нека $r_1, r_2, \ldots, r_m$ са корените на

$$F(r)=0,\ r_k=\theta_k(r_1)$$

и нека $\psi(x, r_1)$ е един неразложим делител на $f(x)$ в разширената рационална област. Както по-рано (глава VI, § 5), се убеждаваме, че $f(x)$ се дели на $\psi(x, r_2), \ldots, \psi(x, r_m)$ и от последните два кои да е полинома, включително $\psi(x, r_1)$, или нямат нито един общ корен, или имат всичките си корени общи. Произведението

$$\prod(x)=\psi(x, r_1)\,\psi(x, r_2)\ldots\psi(x, r_m),$$

което е с коефициенти от първоначалната област на рационалност, ще е точна степен на $f(x)$, т. е.

$$\prod(x)=[f(x)]^q,$$

отгдето, ако степента на $\psi(x, r_1)$ е g, ще имаме $nq=mg$. Понеже n е просто число, то трябва m да се дели на n, тъй като $g<n$.

Интересен е за приложенията нататък и случаят, когато m е просто число. Тогава трябва m да е равно на n и $\psi(x, r_1)$ ще бъде от първа степен и следователно даденото уравнение ще се реши с адюнгиране на r_1. Значи с присъединяване на корените на уравнението $F(r)=0$ групата Г на $f(x)=0$ трябва да се редуцира на Е и по теорема 29 редът на групата Г ще бъде равен на n. Следователно Г е съставена от степените на една кръгова субституция от n-ти ред, т. е. е циклична група. Сега ще установим следната теорема на Галоа:

30. За да бъде едно неразложимо уравнение от степен просто число, решимо алгебрично, необходимо и достатъчно условие е групата му да бъде линейната група или подгрупа на нея.

Отначало ще докажем, че условието е достатъчно. Ако групата е линейната или подгрупа на нея, то всяка рационална функция от корените, която не се изменя за субституциите ѝ, ще има стойност, принадлежаща на областта на рационалност, т. е. ще бъде рационално известна. По теорема 15 корените на уравнението ще се изразят алгебрично посредством коефициентите му.

Обратно, да допуснем, че уравнението е решимо алгебрично и нека Γ е групата му. Тогава по теорема 28 следва, че Γ се разлага в ред

$$\Gamma, \Gamma_1, \Gamma_2, \ldots, \Gamma_r = E$$

с индекси съответно $i_1, i_2, \ldots, i_r$, които са прости числа, като последователно групата Γ се свежда на $\Gamma_1, \Gamma_2, \ldots, \Gamma_r$ с адюнгиране на корени на абелеви уравнения от прости степени $i_1, i_2, \ldots, i_r$. По теорема 29 и следствието ѝ даденото уравнение

$$f(x)=0 \tag{2'}$$

остава неразложимо, докато се сведе групата му на Γ_{r-1} и при последното адюнгиране на корени тя се свежда от Γ_{r-1} на E, при което то се разлага на линейни множители. Групата Γ_{r-1} е циклична и следователно ще се състои от субституциите

$$S^\alpha = |z\ z+\alpha|,\ \alpha=0,\ 1,\ 2,\ \ldots,\ p-1,$$

където p е степен на (22). Нека $T=|z\ \varphi(z)|$ е субституция от Γ_{r-2}. Понеже Γ_{r-1} е и. пг. на Γ_{r-2}, то ще имаме

$$T^{-1} S^a T = T^b.$$

отгдето, като изразим, че $S^aT = TS^b$, получаваме

$$\varphi(z+a) \equiv \varphi(z) + b \pmod p,$$

което трябва да е вярно за $z=0,\ 1,\ 2,\ \ldots,\ p-1$. Ако вместо z поставим $z,\ z+a,\ z+2a,\ \ldots,\ z+qa$ и съберем, ще получим

$$\varphi(z+qa) \equiv \varphi(z)+qb$$

и при $z=0$, като положим $\varphi(0)=c,\ a=1$,

$$\varphi(q) \equiv qb+c.$$

Това показва, че Γ_{r-2} се състои от субституции от вида $|z$

т. е. Γ_{r-2} е или линейната група, или подгрупа на нея.

Групата Γ_{r-2} не съдържа други субституции от ред p освен тия на Γ_{r-1}. Действително, ако

$$U=|z\ bz+c'\ ,$$

то

$$U^m=|\ z\ b^mz+c(b^{m-1}+b^{m-2}+\ldots+1)\ |$$

и $U^m=E$, ако $b^m \equiv 1,\ c(b^{m-1}+b^{m-2}+\ldots+1) \equiv 0 \pmod p$.

Второто условие е следствие от първото при $b \neq 1$, което по теоремата на Ферма се удовлетворява при

$$m = p - 1 < p.$$

Нека сега T_1 е субституция от Γ_{r-3}, S_1 — субституция от Γ_{r-1}, която е също така и субституция от Γ_{r-2}. Понеже Γ_{r-2} и и. пг. на Γ_{r-3}, то трансформираната субституция

$$T_1^{-1} S_1 T_1$$

принадлежи на Γ_{r-2} и тъй като е от ред p, то тя ще принадлежи и на Γ_{r-1}, т. е. тази последна група ще бъде и. пг. на Γ_{r-3}. По горния начин се убеждаваме веднага, че групата Γ_{r-3} ще е съставена само от линейни субституции и т. н., докато достигнем до Γ, с което теоремата е доказана.

Галоа намира още друго условие, дадено с теоремата:

31. Необходимо и достатъчно условие едно неразложимо уравнение от степен просто число да бъде алгебрично решимо се състои в това, че всички негови корени да бъдат рационални функции на два от тях.

Първо, условието е необходимо. Нека следователно уравнението е решимо и да присъединим към областта на рационалност два негови корена, x_α и x_β. Групата на уравнението Γ ще се сведе на тази своя подгрупа Γ_1, която не изменя нито x_α, нито x_β. Ако $S = |z\ az+b$ бъде субституция от Γ с това свойство, то трябва да имаме

$$a\alpha + b \equiv \alpha,\ a\beta + b \equiv \beta,$$

отгдето следва, че $a = 1$, $b = 0$. Групата $\Gamma_1 = E$. Уравнението става решимо и значи всички негови корени са рационални функции на x_α и x_β.

Ще докажем сега обратното. Допускаме значи, че дадено уравнение е неразложимото от p-та степен и всички корени са рационални функции на два от тях, x_α и x_β. Нека Γ е групата на уравнението и редът ѝ е равен на m. С присъединяване на x_α към областта на рационалност групата Γ се свежда на Γ_α от ред $\frac{m}{p}$, понеже x_α е рационална функция, която за субституциите на Γ приема p стойности x_0, $x_1, \ldots, x_{p-1}$, които са корените на уравнението. Ако с $T_0 = E, T_1, T_2, \ldots, T_{p-1}$ означим субституциите от Γ, които сменят x_0 съответно с $x_0, x_1, x_2, \ldots, x_{p-1}$, а с Γ_0 групата, която съответствува на x_0, то групата за x_i е

$$T_i^{-1} \Gamma_0 T_i.$$

Субституциите T съществуват, понеже Γ е транзитивна група. Две групи Γ_α и Γ_β не могат да имат общи субституции освен E, защото би следвало, че Γ има субституции, които не изменят два корена, а следователно не изменят всички останали корени, което е невъзможно.

Да адюнгираме сега x_β, който е корен на

$$\frac{f(x)}{x - x_\alpha} = 0,$$

или ако това уравнение е разложимо, на негов неразложим делител. Групата ще се сведе на група G_β от ред

$$\frac{m}{pv}, v \leqq p-1,$$

и тъй като уравнението се решава, то $G_\beta = E$, откъдето следва, че $m = pv$. Следователно редът на всяка група Γ_i е равен на v. Броят на субституциите във всичките групи Γ_i е равен на $p(v-1)$, като сме изключили единичната E. От това заключаваме, че броят на субституциите на Γ, които едновременно не изменят всички елементи, е равен на $(v-1)p+1$. Но тогава броят на субституциите на Γ, които изменят всички елементи x, ще бъде равен на

$$pv-(v-1)p-1 = p-1.$$

Тези $p-1$ субституции са степените на една кръгова субституция от ред p. Действително нека T е субституция от Γ, която изменя всички елементи. Разлагаме я на цикли. Понеже p е просто число, то всички нейни цикли не могат да имат по еднакъв брой елементи. Ще има един цикъл от минимален ред $g > 1$. Тогава субституцията $T^g \neq E$ не би изменяла g елементи и с присъединяването на тях групата Γ не би се редуцирала на E, което е невъзможно. Степените на T образуват група G от p-ти ред. G е и. пг. на Γ, понеже, ако U е субституция от Γ, а $S_k = T^k$ е от G, то $U^{-1} S_k U$ е кръгова субституция от p-ти ред и следователно принадлежи на G, тъй като Γ не съдържа други субституции, които изменят p елемента. Но тогава по начина, следван по-горе, доказваме, че Γ се състои само от линейни субституции $|z\ az+b|$, което показва съгласно с теорема 15, че уравнението е решимо алгебрически.

2. Числени уравнения, на които групата е симетричната група. Както видяхме, въпросът, дали едно уравнение е разрешимо алгебрически или не, се свежда до намиране на групата му и до установяване дали тази група е метациклична или не. Ще разгледаме примери на уравнения, на които групата е симетричната група или, както се казва, са без афект. Предварително ще установим едно елементарно предложение:

32. *Ако едно неразложимо уравнение с реални коефициенти от степен просто число е решимо алгебрически, то или всичките му корени са реални, или само един от корените му е реален.*

Действително, ако уравнението има поне два реални корена x_1, x_2, то понеже останалите корени са рационални функции на тях с коефициенти, принадлежащи на областта на рационалност, т. е. реални числа, всичките корени ще бъдат реални.

Достатъчно е значи едно неразложимо уравнение от проста степен да има само два имагинерни корена, за да не бъде решимо алгебрически. Такива уравнения съставяме лесно. Така да разгледаме уравнението

$$(1) \qquad f(x) = x^3(x-2)(x-4)\ldots(x-2p+6)-2 = 0$$

от проста степен $p>4$. Понеже всичките му коефициенти освен първия се делят на 2, а последният на 2 в първа степен, то по теоремата на Айзенцайн уравнението е неразложимо. От неравенствата

$$f(0)<0,\ f(1)>0,\ f(3)<0,\ldots,f(2p-7)<0,\ f(2p-5)>0$$

се вижда, че това уравнение има поне $p-2$ реални корена и понеже липсват членовете с x и x^2, то ще има поне два имагинерни корена, т. е. точно два такива. По въпросното предложение уравнението (1) е нерешимо алгебрически.

Ще разгледаме сега примери на уравнения без афект. Предварително ще установим следната теорема:

33. *Ако една транзитивна група от субституции от степен просто число съдържа една транспозиция, то тя е симетричната група.*

Нека групата е **G** и елементите са 1, 2, 3,..., n и нека в нея да влиза транспозицията (12). Може евентуално в нея да влизат и други транспозиции, единият елемент на които е 1, и нека всичките такива транспозиции са

$$(12),\ (13),\ldots,\ (1m).$$

Понеже $(a\,b)=(1a)\,(1b)(1a)$, то групата **G** не съдържа всичките транспозиции на елементите 1, 2, 3,..., m и следователно и симетричната група **M** от тези елементи. Ако $m=n$, то **G** се слива със симетричната група **S** от елементите 1, 2, 3,..., n. Нека $n>m$. Групата **G** не съдържа транспозиции от вида $(i,\ k)$, гдето $1\leqq i\leqq m,\ k>m$. Действително инак тя би съдържала произведението $(1i)(ik)(1k)=(1k)$, което противоречи на допущането. Но **G** е транзитивна група и следователно ще има субституция T от нея, която замества 1 с p, $p>m$. Тогава T не може да сменя елементи от системата P (1, 2, 3,..., m) пак с елементи от нея. Така, ако T замества i с j, $i,\ j\leqq m$, то в групата **G** ще има субституция $T^{-1}(1\,i)\,T=(p\,j)$, която замества j с $p>m$, което е невъзможно. Следователно субституциите на **G** ще сменят или елементите на P помежду им, или ще ги сменят с една различна система **P′** и в **G** ще влизат всички субституции от елементите на P'. Ако така не се изчерпят всичките n елемента, то ще има субституции от **G**, които ще сменят елементите от P с нови елементи на една друга система P'', която няма общи елементи с P'. Действително нека с a означим кое да е число $0<a\leqq n$ и с a_s, a_t означим числата, в които a се изменя съответно със субституциите S и T. Тогава a_{st} е елементът, в който преминава a, при прилагане на субституцията $S\,T$. Ако $a,\ b,\ c,\ldots$ са елементи от P, то ще имаме

$$T^{-1}(ab)\,T=(a_t\,b_t),$$
$$ST^{-1}(ab)\,TS^{-1}=(a_{ts-1},\ b_{ts-1})=(cd).$$

Ако тогава $a_t=c_s$ се намира в P', като T води от P в P'', то $a_{ts-1}=c$ се съдържа в P и следователно $b_{ts-1}=d$ ще се съдържа в P, понеже транспозицията $(c,\ d)$ се съдържа в **G**. Следователно имаме $a_t=c_s$, $b_t=d_s$, т. е. ако един елемент от P'' се намира в P', то трябва P' и P'' да са идентични.

С това е установено, че елементите 1, 2, 3,..., n се разпределят в системи P, P', P'',... по m във всяка. Понеже n е просто число, то трябва m да е равно на n, като по-голямо от 1, с което предложението е доказано, тъй като групата M, която е подгрупа на G, трябва да се слива със симетричната група от n-та степен.

Въз основа на предното предложение ще установим следното предложение :

34. *Ако едно неразложимо уравнение с реални коефициенти от степен просто число има само два имагинерни корена, то групата му е симетричната група.*

Нека предположим, че групата на въпросното уравнение от проста степен p не е симетричната група. Ако към областта на рационалност причислим реалните му корени, броят на които е $p-2$, групата му ще се редуцира на единица. Именно друга субституция не може да има в редуцираната група, понеже възможна е само такава, която сменя двата имагинерни корена. Но тогава въпросната субституция е транспозиция и по предното предложение групата на уравнението трябва да е симетричната. Понеже с присъединяването на реалните корени групата на уравнението се сведе до единичната група, то дадата имагинерни корена трябва да принадлежат на разширената област на рационалност и следователно трябва да бъдат реални, което противоречи. Впрочем това може да се установи малко по-другояче. Нека x_1, x_2 са имагинерните корени и x_3, x_4,..., x_p са реалните корени. От условието, че всяка реална рационална функция от корените на уравнението е известна, т. е. е реална

$$R(x_1,\ x_2,\ x_3,\ldots,\ x_p)=A+iB,\ B=0,$$

следва със смяната на i с $-i$

$$R(x_2,\ x_1,\ x_3,\ldots,x_p)=A-iB=A,$$

т. е. транспозицията (12) трябва да принадлежи на групата на уравнението и по предното предложение следва, че тази група е симетричната.

По подобен начин можем да образуваме произволен брой числени уравнения без афект. Стига именно да вземем уравнения от проста степен с реални коефициенти

$$x^n+a,\ x^{n-1}+\ldots+a_n=0,$$

които имат само прости корени, два от които са имагинерни, а другите — реални. В това уравнение да изменим малко коефициентите и да ги заместим с числа

$$\frac{b,}{c},\ \frac{b_2}{c},\ldots,\frac{b_n}{c},$$

в които b_1, b_2,..., b_n са цели числа, делящи се на просто число q, като b_n се дели само на q в първа степен и c е цяло число, което не се дели на q. При това, понеже корените са непрекъснати функции на коефициентите, може това изменение на коефициентите да бъде така малко, че новото уравнение да има пак $n-2$ реални и два имагинерни корена. Понеже по теоремата на Айзенщайн въпросното уравнение е

неразложимо, то по доказаната теорема следва, че групата му ще бъде симетричната група.

Ше се спрем още на уравненията без афект. Ако $a=1$, a_1, $a_2,\ldots,a_n$ са коефициентите на едно уравнение от n-та степен, то резолвентната функция на Галоа

$$V=p_1x_1+p_2x_2+\ldots+p_nx_n$$

удовлетворява едно уравнение

$$\psi(V)=0$$

от $n!$ степен, което при a_k независими променливи е неразложимо. Полиномът $\psi(V)$ е цяла рационална функция от V, a_1, $a_2,\ldots,a_n$. Като поставим вместо коефициентите a_k стойности от една област на рационалност R, може да се случи $\psi(V)$ да се разложи в R. Тогава групата на уравнението ще бъде истинска подгрупа на симетричната или, както се казва, уравнението има афект. Въпрос е дали може да се поставят вместо a_k рационални стойности така, че $\psi(V)$ да остане неразложим в областта на рационалност, т. е. групата Γ на уравнението да бъде симетричната. Този въпрост е разрешен от Хилберт, който е установил следната теорема: Във всеки неразложим полином на няколко променливи с рационални коефициенти можем винаги да дадем на колкото искаме от тях рационални стойности, така че полиномът остава неразложим спрямо останалите променливи. От теоремата веднага следва, че има безбройно много уравнения с коефициенти рационални числа, които са без афект от произволна степен. Други начини за получаване на уравнения без афект читателят може да намери в цитирана алгебра на О. Хаупт, т. II, стр. 567—573. Въпросът за намиране на всичките уравнения от една степен, имащи дадена група, е решен само в някои частни случаи.

3. Уравнения на Абел. Ще разгледаме отначало някои помощни въпроси. Очевидно всяка транзитивна група от степен просто число е примитивна. Циклична група от съставна степен е импримитивна. Така групата от степените на (123456)

$$\mathbf{G}=E+(123456)+(135)(246)+(14)(25)(36)+(153)(264)+(165432)$$

е импримитивна. Елементите могат де се разделят в две системи

$$(1,\ 3,\ 5),\ (2,\ 4,\ 6)$$

или в три системи

$$(1,\ 4),\ (2,\ 5),\ (3,\ 6).$$

Субституциите на една импримитивна група могат да се наредят в една таблица, подобна на по-раншните. Нека

$$S_1,\ S_2,\ldots,S_r$$

да бъдат субституциите, които пермутират елементите във всяка редица, като не ги сменяват с елементи от друга редица. Очевидно те образуват група $\mathbf{G}_1$, която ще бъде и. пг. на $\mathbf{G}$. Нека T_2 е субституи-

ция, която сменява поне един елемент от една редица с един от друга. Тогава произведенията

$$S_1 T_2, \ldots, S_r T_2$$

са различни помежду си и от първите субституции S; така получаваме втора редица от субституции. Ако не се изчерпва с това групата G, продължаваме така, докато получим цялата група. По този начин се убеждаваме, че G може да се разложи по подгрупи

$$G = G_1 + G_1 T_2 + \ldots + G_1 T_{p-1},$$

отгдето следва, че $pg_1 = g$, гдето g е редът на G, а g_1 — този на G_1. Подобно на групите $H_1, H_2, \ldots$, които са съставени от субституции, пермутиращи само елементите последователно във всяка отделна редица, без да сменяват елементите в другите редици, са и. пг. на G и G_1 и са подобни. Имаме следната теорема:

35. Уравнението, което се получава с елиминиране на y между двете неразложими уравнения

$$f(y) = y^m + a_1 y^{m-1} + \ldots + a_m = 0,$$
$$\varphi(x,y) = x^n + b_1(y) x^{n-1} + \ldots + b_n(y) = 0,$$

гдето $b_i(y)$ са рационални функции на y, има импримитивна група и, обратно, всяко уравнение с импримитивна група може да се получи по този начин.

Нека $y_1, y_2, \ldots, y_m$ са корените на първото уравнение. Уравнението за x ще бъде

$$F(x) = \varphi(x, y_1)\varphi(x, y_2)\ldots\varphi(x, y_m) = 0.$$

Нека с $x_{i1}, x_{i2}, \ldots, x_{in}$ да означим корените на уравнението

$$\varphi(x, y_i) = 0.$$

Всяка симетрична функция u_i на корените x_{ik} се изразява като рационална функция на y_i и всяка симетрична функция на $u_1, u_2, \ldots, u_m$ се изразява като рационална функция на коефициентите на $f(y) = 0$, т. е. ще има стойност, принадлежаща на областта на рационалност. Групата на уравнението G ще се състои само от субституции, които не променят такива функции, т. е. оставащи или u_i непроменени, или ги сменяват помежду си. Тези субституции сменяват значи корените от системите

(39) $$x_{i1}, x_{i2}, \ldots, x_{in}$$

или само помежду им, или една система с друга система, т. е. G е импримитивна група. Обратно, нека G е импримитивна група и елементите x се разделят на m редици по n корена (39). Да разгледаме симетрични функции от елементите във всяка редица:

$$y_i = F_i(x_{i1}, x_{i2}, \ldots, x_{in});$$

те или остават непроменени, или се пермутират от субституциите на G. Следователно произведението

$$f(y) = (y - y_1)(y - y_2)\ldots(y - y_m)$$

не се променя от субституциите на G и ще има коефициенти, принадлежащи в областта на рационалност. Ако познаваме y_i, то всяка

симетрична функция на $x_{i1}, x_{i2}, \ldots, x_{in}$ ще бъде рационална функция на y_i, т. е. ще можем да образуваме уравнение

$$\varphi(x, y_i)=0,$$

на което тези числа ще бъдат корени. Но тогава даденото уравнение като неразложимо ще бъде

$$F(x)=\varphi(x, y_1)\ \varphi\ (x, y_2)\ldots\varphi(x, y_m)=0$$

и ще се получава с елиминиране на y между две уравнения от вида (39).

Нека разгледаме сега едно неразложимо абелево уравнение

(40) $$f(x)=0,$$

в което предполагаме, че има рационална връзка между два негови корена x_1 и x_2:

(41) $$x_2=\theta(x_1).$$

Групата на уравнението трябва да съдържа субституция S, която замества x_1 с x_2; нека

$$(x_1\ x_2\ldots x_r)$$

е един от нейните цикли. Понеже функцията $x_2-\theta(x_1)$ има стойност нула, т. е. принадлежи на областта на рационалност, то при прилагане на S и степените ѝ тя не променя стойността си. Следователно ще имаме

(42) $$x_2=\theta(x_1),\ x_3=\theta(x_2),\ldots,x_1=\theta(x_r).$$

Ако приложим една субституция, която замества един от кореният $x_1, x_2, \ldots, x_r$ с друг, който не фигурира в тази редица, то, както лесно се вижда, тази субституция ще замества всички корени от тази редица с корени от нова една редица, понеже иначе уравнението би имало равни корени. Ако сега една субституция между кореният в редица (42) трансформираме посредством субституция, която ги сменява в други r, то така получаваме една нова редица от r корена, свързани с релации, подобни на (42). Така, като продължаваме, достигаме до известния ни вече резултат, че кореният се разпределят в m системи по r, гдето $mr=n$ е степента му. Групата на уравнението ще бъде импримитивна и ще сменява или само елементи от всяка система, или всички елементи на една коя да е система на всички такива на друга. По теорема 35 решението на даденото уравнение се свежда към решение на едно уравнение от m-та степен и на m абелеви уравнения от r-та степен, резултат, който получихме по-рано.

Под общо абелево уравнение ще разбираме такова, в което всички корени са рационални функции на единия от тях x_1 и ако $\theta_\alpha(x_1)$ и $\theta_\beta(x_1)$ са два кои да са корена, то

(43) $$\theta_\alpha\,\theta_\beta(x_1)=\theta_\beta\,\theta_\alpha(x_1).$$

Едно такова неразложимо уравнение по теорема 29 има група Г от ред, равен на степента му n, която ще се състои от субституции

S_α, $\alpha=1, 2, \ldots, n$, $S_1=E$, които сменяват x_1 съответно с $x_1, x_2, \ldots, x_n$. Тогава от релацията (43) е ясно, че за кои да са две субституции ще имаме

$$S_\alpha S_\beta = S_\beta S_\alpha,$$

т. е. **Г** е абелева група. Обратно, ще установим теоремата

36. Ако групата на едно уравнение е абелева, то и уравнението е абелево.

Ако уравнението е разложимо, то групите на отделните неразложими делители са подгрупи на общата група и са следователно абелеви. Остава значи да се ограничим само на транзитивни групи. В този случай има n субституции $S_1, S_2, \ldots, S_n$ в групата **Г**, които сменяват корена x_1 с всички корени $x_1, x_2, \ldots, x_n$ на уравнението. Лесно се вижда, че това са всички субституции на **Г**.

Всички субституции от **Г**, които не изменят x_1, образуват група Γ_1, която е подгрупа на **Г**. Тогава конюгованата подгрупа

$$S_k^{-1} \Gamma_1 S_k$$

се състои от всички субституции, които не изменят x_k и понеже **Г** е абелева група, то Γ_1 е и. пг., т. е. тази група се слива с Γ_1. Следователно Γ_1 не изменя никой корен, т. е. се състои само от субституцията E. По теорема 23 всички корени ще бъдат рационални функции на единия от тях, например на x_1. Значи ще имаме

$$x_k = \theta_k(x_1).$$

Субституцията $S_l S_k$ ще трансформира x_1 с

$$x_i = \theta_k[\theta_l(x_1)]$$

и понеже $S_k S_l = S_l S_k$, ще получим лесно, че

$$\theta_k[\theta_l(x_1)] = \theta_l[\theta_k(x_1)],$$

т. е. уравнението е абелево.

Върху алгебричната решимост на уравнения от степен p^m, гдето p е просто число, са работили главно Жордан и Силов. Много геометрични въпроси водят до разрешими уравнения. Такова например е уравнението от 9-та степен на Хес, до което се идва при търсене инфлексните точки на кривите от трета степен, но на такива въпроси няма да се спрем, понеже биха ни завели твърде далеч в подробности. Тях любознателният читател ще може да намери в цитираните книги

ИЗПОЛЗУВАНА ЛИТЕРАТУРА

А. Я. Окунев, Высшая алгебра, Москва, 1949.

А. Г. Курош, Курс высшей алгебры, Москва, 1949.

А. К. Сушкевич, Основы высшей алгебры, Москва, 1941.

А. И. Мальцев, Основы линейной алгебры, Москва, 1948.

Г. Шапиро, Высшая алгебра, Москва, 1938.

И. М. Гельфанд, Лекции по линейной алгебре, Москва, 1951.

Вандер Варден, Современная алгебра, т. I, II, Москва, 1947.

А. Г. Курош, Теория групп, Москва, 1953.

Л. Е. Диксон, Линейные алгебры, Москва, 1935.

H. Veber, Lehrbuch des Algebra, Braunschweig, m. I, II (1896), m. III) 1908).

C. Runge, Praxis des Gleichungen, Berlin, 1921.

R. Fricke, Lehrbuch des Algebra, Braunschweig, m. I (1924), m. II (1926), m. III (1928).

O. Haupt, Einführung in die Algebra, Leipzig (1929).

E. Netto, Vorlesungen über Algebra, Leipzig, m. I, II (1900).

A. Speiser, Theorie der Gruppen von endlicher Ordnung, Berlin (1923).

Энциклопедия элементарной математики, II, Алгебра, Москва (1951).

Н. Г. Чеботарев, Теория алгебрических функций, Москва (1948).

Н. Г. Чеботарев, Основы теории Галуа, Москва, т. I и II (1937).

O. Perron, Algebra, Leipzig und Berlin (1927).

Н. Обрéшков, Quelques classes de fonctions entières limites de polynomes et de fonctions méromorphes limites de fonctions rationnelles, Actualites scientifiques et industrielles, Paris, 1941.

J. Dieudonné, La théorie analytiqye des polynomes d'une variable (à coefficients quelconques), Mémorial des sciences mathématiques, Paris, 1938.

M. Marden, The geometry of the zeros of a polynomial in a complex variable, New York, 1949.

ПЕЧАТНИ ГРЕШКИ

Стр.	Ред	Напечатано	Да се чете	По вина на
21	6 отдолу	$-1)^{[\beta\delta\ldots\gamma]}$	$(-1)^{[1\beta\delta\ldots\gamma]}$	коректора
80	14 отгоре	$a_{in}x_m$	$a_{im}x_m$	автора
163	20 „	$=Ae'p$	$-Ae'_p$	печатницата
163	21 „	$=\nu_r h_r$	$+\nu_r h_r$	коректора
291	7 отдолу	$a_n-a_{n-1}<$	$\lvert a_n-a_{n-1}\rvert <$	печатницата
292	7 „	$a_1+a_1-a\dfrac{1}{1-q}$	$a_1+\lvert a_1-a\rvert \dfrac{1}{1-q}$	„
294	1 „	$-\dfrac{f(a_1)}{f(a)}$	$-\dfrac{f(a_1)}{f'(a)}$	„
305	6 „	не е	е	автора
306	15 „	$=a_p x^p$	$=\lvert a_p x^p\rvert$	печатницата
314	14 отгоре	$+a_{n-1}x^{n-1}$	$+\bar{a}_{n-1}x^{n-1}$	автора
322	4 отдолу	$\lvert <1$	$\lvert y\rvert <1$	печатницата
360	9 отгоре	$=z_k$	$=\bar{z}_k$	автора
413	9 „	$e^{-\lambda}$	e^{-1}	„
468	6 „	$\alpha^{i+1}\theta^{i+1}x_1$	$\alpha^{i+1}\theta^{i+1}x_1$	печатницата
494	5 отдолу	$\psi(\omega^{p-1}v^p)$	$\psi(\omega^{p-1}v^{\frac{1}{p}})$	печатницата
495	3 отдолу	$=g^p$	$=g^{\frac{1}{p}}$	печатницата