Лекция N 18 (В.Т. Марков)
(11 ноября 2011 г.)

Задача 3.19.11. Если m — нечётное число, $m > 1$, то $\Phi_{2m}(x) = \Phi_m(-x)$.

Задача 3.19.12. $\Phi_n(x) = x^{\varphi(n)}\Phi_n(x^{-1})$.

Задача 3.19.13. Для любого числа $n \in \mathbb{N}$ многочлен $\Phi_n(x)$ неприводим над полем рациональных чисел $\mathbb{Q}$. Однако многочлены $\Phi_n(x)$ могут быть приводимы над конечными полями. Например, в кольце $\mathbb{Z}_2[x]$ над полем $\mathbb{Z}_2$

$$\Phi_4(x) = (x+1)^2; \quad \Phi_{15}(x) = (x^4+x^3+1)(x^4+x+1).$$

В кольце $\mathbb{Z}_3[x]$ над полем $\mathbb{Z}_3$

$$\Phi_8(x) = (x^2+x-1)(x^2-x-1).$$

Упражнение 3.19.14. Покажите, что $x^{15}-1 = \Phi_{15}(x)\Phi_5(x)\Phi_3(x)\Phi_1(x)$ в $\mathbb{Z}_2[x]$.

3.20. Рациональные дроби и функции

Пусть K — поле, $K[x]$ — кольцо многочленов над полем K. *Рациональной дробью* называется пара многочленов

$$\frac{f(x)}{g(x)}, \quad f(x), g(x) \in K[x], \quad g(x) \neq 0.$$

На рациональных дробях рассматривается следующее отношение эквивалентности: две рациональные дроби $\dfrac{f(x)}{g(x)}$ и $\dfrac{f_1(x)}{g_1(x)}$ *равны*, если $f(x)g_1(x) = g(x)f_1(x)$.

Упражнение 3.20.1. Показать, что классы равных между собой рациональных дробей (называемые *рациональными функциями*) с естественно определёнными операциями сложения и умножения образуют поле $K(x)$ рациональных функций (поле частных кольца $K[x]$). Так как $K \subset K[x] \subset K(x)$, то $\operatorname{char} K = \operatorname{char} K(x)$.

Рациональная дробь $\dfrac{f(x)}{g(x)}$ называется *несократимой*, если многочлены $f(x)$ и $g(x)$ взаимно просты, и *правильной*, если $\deg f(x) < \deg g(x)$ (или если $f(x) = 0$).

Лемма 3.20.2. *Всякая рациональная дробь* $\dfrac{f(x)}{g(x)}$ *равна некоторой несократимой дроби* $\dfrac{f_1(x)}{g_1(x)}$, *определяемой однозначно, с точностью до множителя* $0 \neq c \in K$ *(т. е.* $\dfrac{cf_1(x)}{cg_1(x)}$*).*

Доказательство. 1) Если $d(x) = \text{НОД}\big(f(x), g(x)\big)$, $f(x) = d(x)f_1(x)$, $g(x) = d(x)g_1(x)$, $\big(f_1(x), g_1(x)\big) = 1$, то $\dfrac{f(x)}{g(x)} = \dfrac{f_1(x)}{g_1(x)}$.

2) Если $\dfrac{f(x)}{g(x)} = \dfrac{\varphi(x)}{\psi(x)}$ — две равные несократимые дроби, то $f(x)\psi(x) = g(x)\varphi(x)$. Так как правая часть делится на $g(x)$ и $\big(f(x), g(x)\big) = 1$, то $\psi(x)$ делится на $g(x)$. Аналогично, левая часть делится на $\psi(x)$, $\big(\varphi(x), \psi(x)\big) = 1$, и поэтому $g(x)$ делится на $\psi(x)$. Таким образом, $cg(x) = \psi(x)$, $0 \neq c \in K$. Поэтому $cf(x)g(x) = g(x)\varphi(x)$. Сокращая на $g(x)$, получаем $\varphi(x) = cf(x)$. □

Лемма 3.20.3. *Всякая рациональная дробь* $\dfrac{f(x)}{g(x)}$ *представима (и единственным образом) в виде суммы многочлена и правильной дроби.*

Доказательство. 1) Пусть $f(x) = g(x)q(x) + r(x)$, где или $r(x) = 0$, или $\deg r(x) < \deg g(x)$. Тогда $\dfrac{f(x)}{g(x)} = q(x) + \dfrac{r(x)}{g(x)}$.

2) Если $\dfrac{f(x)}{g(x)} = \bar{q}(x) + \dfrac{\varphi(x)}{\psi(x)}$, где $\bar{q}(x) \in K[x]$, или $\varphi(x) = 0$, или $\deg \varphi(x) < \deg \psi(x)$, то

$$q(x) - \bar{q}(x) = \frac{\varphi(x)g(x) - \psi(x)r(x)}{\psi(x)g(x)}.$$

Из сравнения степеней видим, что $q(x) = \bar{q}(x)$, $\varphi(x)g(x) - \psi(x)r(x) = 0$, т. е. $\dfrac{\varphi(x)}{\psi(x)} = \dfrac{r(x)}{g(x)}$. □

Правильная рациональная дробь $\dfrac{f(x)}{g(x)}$ называется *простейшей*, если $g(x) = p^k(x)$, где $p(x)$ — неприводимый многочлен, $k \geqslant 1$ и $\deg f(x) < \deg p(x)$.

Теорема 3.20.4. *Всякая правильная рациональная дробь разлагается (и единственным образом) в сумму простейших.*

Доказательство существования разложения

Сначала докажем следующее утверждение.

Лемма 3.20.5. *Если* $\dfrac{f(x)}{g(x)h(x)}$ — *правильная дробь и* $\big(g(x), h(x)\big) = 1$, *то*

$$\frac{f(x)}{g(x)h(x)} = \frac{v(x)}{g(x)} + \frac{u(x)}{h(x)},$$

где $\dfrac{u(x)}{h(x)}, \dfrac{v(x)}{g(x)}$ — *правильные дроби.*

Доказательство. Пусть

$$g(x)\bar{u}(x) + h(x)\bar{v}(x) = 1.$$

Умножая на $f(x)$, получаем

$$g(x)\bar{u}(x)f(x) + h(x)\bar{v}(x)f(x) = f(x).$$

Пусть $\bar{u}(x)f(x) = h(x)q(x) + u(x)$, где или $u(x) = 0$, или $\deg u(x) < \deg h(x)$. Тогда

$$g(x)u(x) + h(x)v(x) = f(x),$$

где $v(x) = \bar{v}(x)f(x) + g(x)q(x)$. Из сравнения степеней получаем, что $\deg v(x) < \deg g(x)$. Тогда

$$\frac{f(x)}{g(x)h(x)} = \frac{u(x)}{h(x)} + \frac{v(x)}{g(x)},$$

$\dfrac{u(x)}{h(x)}$ и $\dfrac{v(x)}{g(x)}$ — правильные дроби. □

Если $g(x) = p_1^{k_1}(x) \dots p_r^{k_r}(x)$ — разложение в произведение неприводимых многочленов, то

$$\frac{f(x)}{g(x)} = \frac{u_1(x)}{p_1^{k_1}(x)} + \dots + \frac{u_r(x)}{p_r^{k_r}(x)},$$

где слагаемые в правой части — правильные дроби.

Далее, если $\dfrac{u(x)}{p^k(x)}$ — правильная дробь и $p(x)$ — неприводимый многочлен, то, проводя последовательно деления с остатком

$$\begin{aligned} u &= p^{k-1}s_1 + u_1, \\ u_1 &= p^{k-2}s_2 + u_2, \\ &\dots \\ u_{k-2} &= ps_{k-1} + u_{k-1}, \end{aligned}$$

получаем

$$u = p^{k-1}s_1 + p^{k-2}s_2 + \dots + ps_{k-1} + u_{k-1},$$

и поэтому

$$\frac{u}{p^k} = \frac{u_{k-1}}{p^k} + \frac{s_{k-1}}{p^{k-1}} + \dots + \frac{s_1}{p} -$$

разложение в сумму простейших дробей. □

Доказательство единственности представления правильной рациональной дроби в виде суммы простейших дробей

Допустим противное, т. е. что некоторая правильная рациональная дробь допускает два различных представления в виде суммы простейших дробей. Вычитая из одного представления другое и приводя подобные члены, приходим к нетривиальной сумме простейших дробей, равной нулю. Пусть $p_1(x), \dots, p_s(x)$ — все различные неприводимые многочлены, входящие своими степенями в знаменатели. Пусть k_i — наивысшая среди них степень многочлена $p_i(x)$, $i = 1, \dots, s$. Умножим наши равенства на $p_1^{k_1-1}(x)p_2^{k_2}(x) \cdots p_s^{k_s}(x)$, тогда все слагаемые в нашей сумме окажутся многочленами, кроме одного, который из $\dfrac{u(x)}{p_1^{k_1}(x)}$ превратится в дробь $\dfrac{u(x)p_2^{k_2}(x) \cdots p_s^{k_s}(x)}{p_1(x)}$. Так как многочлен $p_1(x)$ неприводим, а все множители числителя взаимно просты с $p_1(x)$, то числитель не делится нацело на знаменатель. Разделив числитель на $p_1(x)$ с остатком, мы получим, что сумма многочлена и отличной от нуля правильной дроби равна нулю, что приводит нас к противоречию. □

Пример 3.20.6. $\dfrac{f(x)}{g(x)} \in \mathbb{R}(x)$,

$$\begin{aligned} f(x) &= 2x^4 - 10x^3 + 7x^2 + 4x + 3, \\ g(x) &= x^5 - 2x^3 + 2x^2 - 3x + 2. \end{aligned}$$

Так как $g(x) = (x+2)(x-1)^2(x^2+1)$, то разложение в сумму простейших дробей ищем в виде

$$\frac{f(x)}{g(x)} = \frac{A}{x+2} + \frac{B}{(x-1)^2} + \frac{C}{x-1} + \frac{Dx+E}{x^2+1}.$$

Приводя дроби к общему знаменателю и сравнивая коэффициенты при степенях, вычисляем, что

$$A = 3, \quad B = 1, \quad C = -2, \quad D = 1, \quad E = -3,$$

т. е.

$$\frac{f(x)}{g(x)} = \frac{3}{x+2} + \frac{1}{(x-1)^2} - \frac{2}{x-1} + \frac{x-3}{x^2+1}.$$

Замечание 3.20.7. В частном случае, когда знаменатель $g(x)$ рациональной дроби $\frac{f(x)}{g(x)}$, $\deg f < \deg g$, имеет различные корни $x_1, x_2, \ldots, x_n$,

$$g(x) = (x - x_1)(x - x_2) \cdots (x - x_n),$$

рациональная дробь имеет разложение в простейшие:

$$\frac{f(x)}{g(x)} = \sum_{i=1}^{n} \frac{A_i}{x - x_i}.$$

Умножая обе части равенства на $x - x_i$ и подставляя $x = x_i$, получаем

$$A_i = \frac{f(x_i)}{\prod\limits_{\substack{j=1 \\ i \neq j}}^{n} (x_i - x_j)} = \frac{f(x_i)}{g'(x_i)},$$

следовательно,

$$\frac{f(x)}{g(x)} = \sum_{i=1}^{n} \frac{f(x_i)}{(x - x_i)g'(x_i)}.$$

Отсюда (умножая на $g(x)$ это равенство) получаем другой вывод интерполяционной формулы Лагранжа:

$$f(x) = \sum_{i=1}^{n} \frac{g(x)f(x_i)}{(x - x_i)g'(x_i)}.$$

Упражнение 3.20.8. Покажите, что

$$\frac{n!}{(x+1)(x+2)\cdots(x+n)} = \frac{C_n^1}{x+1} - \frac{2C_n^2}{x+2} + \frac{3C_n^3}{x+3} + \ldots + (-1)^{n+1}\frac{n \cdot C_n^n}{x+n}.$$

3.21. Границы корней многочлена с действительными коэффициентами

Лемма 3.21.1. *Пусть*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 \in \mathbb{R}[x],$$

$a_n \neq 0$, $A = \max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}$, $K = 1 + \frac{A}{|a_n|}$. *Тогда если* $x_0 \in \mathbb{R}$ *и* $|x_0| \geqslant K$, *то* x_0 *не является корнем многочлена* $f(x)$. *Таким образом, если многочлен* $f(x)$ *имеет действительные корни, то все они принадлежат интервалу* $(-K, K)$.

Доказательство. Ясно, что $K \geqslant 1$. Если $K = 1$, то $A = 0$ и $a_0 = a_1 = \ldots = a_{n-1} = 0$, то есть $f(x) = a_n x^n$, все его корни равны нулю. Ясно, что число x_0 такое, что $|x_0| \geqslant 1$, не является его корнем.

Если же $K > 1$, то $|x_0| \geqslant K > 1$. Тогда

$$\begin{aligned} D &= |a_{n-1}x_0^{n-1} + \ldots + a_1 x_0 + a_0| \leqslant \\ &\leqslant |a_{n-1}|\,|x_0|^{n-1} + \ldots + |a_1|\,|x_0| + |a_0| \leqslant \\ &\leqslant A(|x_0|^{n-1} + \ldots + |x_0| + 1) = A\frac{|x_0|^n - 1}{|x_0| - 1} < A\frac{|x_0|^n}{|x_0| - 1}. \end{aligned}$$

Так как $|x_0| \geqslant 1 + \dfrac{A}{|a_n|}$, то $|x_0| - 1 \geqslant \dfrac{A}{|a_0|}$, следовательно, $|a_0| \geqslant \dfrac{A}{|x_0| - 1}$, и поэтому $D < |a_0 x_0^n|$. Ясно, что $|f(x_0)| \geqslant |a_0 x_0^n| - D > 0$, поэтому $f(x) \neq 0$. □

Замечание 3.21.2. Аналог утверждения леммы 3.21.1 справедлив для комплексных многочленов (неравенство Коши): если $f(z) = a_n z^n + \ldots + a_0 \in \mathbb{C}[z]$, $a_n \neq 0$, $f(x) = 0$ $(x \in \mathbb{C})$, то

$$|x| < 1 + \frac{\max\{|a_0|, |a_1|, \ldots, |a_n|\}}{|a_n|}.$$

См. также следствие 3.29.4.

Для положительных корней более точная (верхняя) граница даётся следующей оценкой Маклорена.

Теорема 3.21.3. *Пусть*

$$f(x) = a_n x^n + \ldots + a_1 x + a_0 \in \mathbb{R}[x],$$

$a_n > 0$, $K_1 = 1 + \sqrt[m]{\dfrac{M}{a_n}}$, *где M — максимум модуля отрицательных коэффициентов многочлена $f(x)$, $(n - m)$ — номер первого отрицательного коэффициента, считая с левой стороны. Если $x_0 \in \mathbb{R}$ и $x_0 \geqslant K_1$, то $f(x_0) \neq 0$.*

Доказательство. Если многочлен $f(x)$ не имеет отрицательных коэффициентов, то он не имеет положительных корней.

Если же $M > 0$ и $x_0 \geqslant K_1$, то $x_0 > 1$ и

$$\begin{aligned} f(x_0) &= a_n x_0^n + \ldots + a_{n-m+1}x_0^{n-m+1} + a_{n-m}x_0^{n-m} + \ldots + a_1 x_0 + a_0 \geqslant \\ &\geqslant a_n x_0^n - M(x_0^{n-m} + \ldots + x_0 + 1) = a_n x_0^n - M\frac{x_0^{n-m+1} - 1}{x_0 - 1} = \\ &= \frac{x_0^{n-m+1}(a_n x_0^{m-1}(x_0 - 1) - M)}{x_0 - 1} + \frac{M}{x_0 - 1} > \\ &> \frac{x_0^{n-m+1}(a_n x_0^{m-1}(x_0 - 1) - M)}{x_0 - 1} > \frac{x_0^{n-m+1}(a_n(x_0 - 1)^m - M)}{x_0 - 1} \geqslant \\ &\geqslant \frac{x_0^{n-m+1}\left(a_n\left(\sqrt[m]{\frac{M}{a_n}}\right)^m - M\right)}{x_0 - 1} = \frac{x_0^{n-m+1}(M - M)}{x_0 - 1} = 0. \end{aligned}$$

Итак, мы показали, что $f(x_0) > 0$. □

Пусть $f(x) \in \mathbb{R}[x]$, $\deg f(x) = n > 0$, N_0 — верхняя граница его положительных корней, $\varphi_1(x) = x^n f\left(\frac{1}{x}\right)$, $\varphi_2(x) = f(-x)$, $\varphi_3(x) = x^n f\left(-\frac{1}{x}\right)$, N_1, N_2, N_3 — соответственно верхние границы положительных корней многочленов $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$. Ясно, что $\frac{1}{N_1}$ — нижняя граница положительных корней многочлена $f(x)$, $-N_2$ и $-\frac{1}{N_3}$ — соответственно нижняя и верхняя границы отрицательных корней многочлена $f(x)$.

Замечание 3.21.4 (Ньютон). Из представления многочлена $f(x)$ в виде многочлена Тейлора в точке $c \in \mathbb{R}$,

$$f(x) = f(c) + f'(c)(x-c) + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n,$$

следует, что если

$$f(c) > 0,\ f'(c) > 0, \ldots,\ f^{(n)}(c) > 0,$$

то c — верхняя граница положительных корней многочлена $f(x)$. Как и выше, нижняя граница отрицательных корней находится с помощью верхней границы положительных корней многочлена $f(-x)$.

Пример 3.21.5. $f(x) = x^4 - 3x^3 + 8x^2 - 5 = 0$.

а) Тогда $K = 1 + \frac{8}{1} = 9$, то есть все действительные корни многочлена $f(x)$ (если они есть) принадлежат интервалу $(-9, 9)$.

б) Оценка Маклорена даёт $N_0 = K_1 = 1 + \frac{5}{1} = 6$ (верхняя граница положительных корней). Для получения нижней оценки отрицательных корней рассматриваем

$$\varphi_2(x) = f(-x) = x^4 + 3x^3 + 8x^2 - 5,$$

$N_2 = K_1' = 1 + \sqrt[4]{\frac{5}{1}}$, $N_2 < 2{,}5$, $-N_2 > -2{,}5$. Таким образом, для корней получаем интервал $(-2{,}5, 6)$ $(\subset (-9, 9))$.

в) Рассмотрим теперь производные многочлена $f(x)$:

$$\begin{aligned} f(x) &= x^4 - 3x^3 + 8x^2 - 5, \\ f'(x) &= 4x^3 - 9x^2 + 16x, \\ f''(x) &= 12x^2 - 18x + 16, \\ f^{(3)}(x) &= 24x - 18, \\ f^{(4)}(x) &= 24. \end{aligned}$$

Нетрудно видеть, что

$$f(1) > 0,\ f'(1) > 0,\ f''(1) > 0,\ f^{(3)}(1) > 0,\ f^{(4)}(1) > 0.$$

Таким образом, 1 — верхняя граница положительных корней многочлена $f(x)$. Рассматривая многочлен $\varphi_2(x) = f(-x)$, с помощью метода Ньютона убеждаемся в том, что -1 — нижняя граница отрицательных корней многочлена $f(x)$. Итак, все действительные корни многочлена $f(x)$ (если они есть) принадлежат интервалу $(-1, 1)$.

Теорема 3.21.6 (Коши). *Пусть*

$$f(x) = x^n - b_{n-1}x^{n-1} - \ldots - b_1 x - b_0 \in \mathbb{R}[x],$$

$b_i \in \mathbb{R}$, $b_i \geqslant 0$, *при этом хотя бы один из этих коэффициентов* b_i *отличен от нуля. Тогда:*

1) *многочлен* $f(x)$ *имеет единственный положительный корень с кратности* 1,

2) *модули* $|\alpha|$ *остальных корней* α *не превосходят числа* c, *т. е.* $|\alpha| \leqslant c$.

Доказательство. 1) Функция

$$F(x) = -\frac{f(x)}{x^n} = -1 + \frac{b_{n-1}}{x} + \ldots + \frac{b_0}{x^n}$$

при возрастании x от 0 к $+\infty$ строго убывает от $+\infty$ до -1, следовательно, она обращается в нуль ровно в одной точке $c > 0$. Для $x \neq 0$ условия $f(x) = 0$ и $F(x) = 0$ равносильны, поэтому $f(c) = 0$.

Так как

$$F'(c) = -\frac{f'(c)}{c^n} + \frac{nf(c)}{c^{n+1}} = -\frac{f'(c)}{c^n} = -\frac{b_{n-1}}{c^2} - \ldots - \frac{nb_0}{c^{n+1}} < 0,$$

то $f'(c) \neq 0$, поэтому кратность корня многочлена $f(x)$ равна 1.

2) Пусть $f(\alpha) = 0$, $\alpha \in \mathbb{C}$. Допустим противное, т. е. что $c < |\alpha|$. Тогда

$$0 = F(c) > F(|\alpha|) = -\frac{f(|\alpha|)}{|\alpha|^n}.$$

Следовательно, $f(|\alpha|) > 0$. С другой стороны,

$$\alpha^n = b_{n-1}\alpha^{n-1} + \ldots + b_1\alpha + b_0,$$

поэтому

$$|\alpha|^n \leqslant b_{n-1}|\alpha|^{n-1} + \ldots + b_1|\alpha| + b_0,$$

т. е. $f(|\alpha|) \leqslant 0$, что приводит к противоречию.

Итак, мы показали, что $|\alpha| \leqslant c$. □

Замечание 3.21.7. В условиях теоремы Коши возможно, что $|\alpha| = c$ для $\alpha \neq c$: так как

$$x^2 - x - 1 = \left(x - \left(\frac{1}{2} + \frac{\sqrt{5}}{2}\right)\right)\left(x - \left(\frac{1}{2} - \frac{\sqrt{5}}{2}\right)\right),$$

то многочлен $x^{2n} - x^n - 1$ имеет ровно n корней, модули которых равны единственному положительному корню этого многочлена.

Замечание 3.21.8. Если в теореме Коши дополнительно предположить, что наибольший общий делитель номеров положительных коэффициентов a_i равен 1, то многочлен $f(x)$ имеет единственный положительный корень $c > 0$, а модули остальных корней α строго меньше, чем c (т. е. $|\alpha| < c$).

Упражнение 3.21.9. Пусть $z_1, \ldots, z_n \in \mathbb{C}$ — точки, в которых расположены единичные массы,

$$m = \frac{z_1 + \ldots + z_n}{n} -$$

центр масс точек $z_1, \ldots, z_n$. Если $z_0 \in \mathbb{C}$, то

$$m(z_0) = z_0 + n\left(\frac{1}{z - z_0} + \ldots + \frac{1}{z_n - z_0}\right)^{-1} -$$

центр масс точек $z_1, \ldots, z_n$ относительно точки z_0. Пусть

$$f(z) = (z - z_1)\ldots(z - z_n) \in \mathbb{C}[z],$$

тогда центр масс корней многочлена $f(z)$ относительно произвольной точки $z_0 \in \mathbb{C}$ вычисляется по формуле

$$m(z_0) = z_0 - n\frac{f(z_0)}{f'(z_0)}.$$

Действительно, следует учесть, что

$$\frac{f'(z)}{f(z)} = (z - z_1)^{-1} + \ldots + (z - z_n)^{-1}.$$

Упражнение 3.21.10 (теорема Лагерра). Пусть $f(z) \in \mathbb{C}[z]$, $\deg f(z) = n$, c — его корень кратности 1, тогда центр масс всех $n - 1$ остальных корней относительно точки $c \in \mathbb{C}$ равен $c - 2(n-1)\dfrac{f'(c)}{f''(c)}$.

3.22. Число действительных корней многочлена с действительными коэффициентами на отрезке

Пусть $f(x) \in \mathbb{R}[x]$ — многочлен с действительными коэффициентами. Одно из достижений алгоритмической (компьютерной) алгебры — *теорема Штурма* (1829 г.), дающая алгоритм для вычисления числа действительных корней многочлена $f(x) \in \mathbb{R}[x]$ на отрезке $[a, b]$, $a, b \in \mathbb{R}$, $a < b$ (случай $a = -\infty$, $b = +\infty$ для расширенной прямой $\overline{\mathbb{R}}$ даёт число всех вещественных корней многочлена $f(x)$).

Ясно, что достаточно эту задачу решить для многочлена без кратных корней (общий случай сводится к этому переходом от многочлена $f(x)$ к многочлену $f(x)/\big(f(x), f'(x)\big)$, имеющему в точности те же корни, что и многочлен $f(x)$, но кратности, равной 1). В этом случае НОД$\big(f(x), f'(x)\big) = c \neq 0$, $c \in \mathbb{R}$.

3.23. Алгоритм Евклида и система многочленов Штурма

На основе алгоритма Евклида нахождения наибольшего общего делителя построим следующую *каноническую систему многочленов Штурма* для многочлена $f(x) \in \mathbb{R}[x]$.

Пусть $f_0(x) = f(x)$, $f_1(x) = f'(x)$. Далее используем следующую модификацию в алгоритме Евклида (остатки от последующих делений будем брать с противоположным знаком):

$$\begin{aligned} f_0(x) &= f_1(x)q_1(x) - f_2(x), \\ &\vdots \\ f_{k-1}(x) &= f_k(x)q_k(x) - f_{k+1}(x), \\ &\vdots \\ f_{s-2}(x) &= f_{s-1}(x)q_{s-1}(x) - f_s(x), \\ f_{s-1}(x) &= f_s(x)q_s(x) - 0, \end{aligned}$$

здесь $f_s(x) = \text{НОД}\big(f(x), f'(x)\big) = c \neq 0$, $c \in \mathbb{R}$ (т. е. ненулевая константа).

Под *канонической системой Штурма* для многочлена $f(x) \in \mathbb{R}[x]$ без кратных корней понимаем следующую систему многочленов:

$$f_0(x) = f(x),\ f_1(x) = f'(x),\ f_2(x), \ldots,\ f_s(x) = c \neq 0.$$

Свойства системы многочленов Штурма $f_0(x), f_1(x), \ldots, f_s(x)$

1) Соседние многочлены $f_k(x)$ и $f_{k+1}(x)$ не имеют общих корней.

Доказательство. Пусть

$$f_k(\alpha) = 0 = f_{k+1}(\alpha), \quad \text{где} \quad \alpha \in \mathbb{R}.$$

Тогда

$$f_{k-1}(\alpha) = f_k(\alpha)q_k(\alpha) - f_{k+1}(\alpha) = 0,$$

и поэтому, поднимаясь вверх по схеме алгоритма Евклида, имеем $f'(\alpha) = 0$, $f(\alpha) = 0$, что противоречит отсутствию кратных корней для $f(x)$. □

2) Ясно, что последний многочлен $f_s(x) = c \neq 0$, $c \in \mathbb{R}$, не имеет действительных корней.

3) Если $1 \leqslant k \leqslant s-1$ и $f_k(\alpha) = 0$, $\alpha \in \mathbb{R}$, то

$$f_{k-1}(\alpha)f_{k+1}(\alpha) < 0$$

(т. е. действительные ненулевые числа $f_{k-1}(\alpha)$ и $f_{k+1}(\alpha)$ имеют противоположные знаки).

Доказательство. Так как

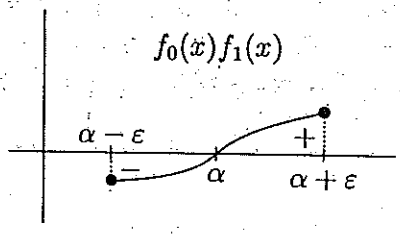
$$f_{k-1}(x) = f_k(x)q_k(x) - f_{k+1}(x),$$

то

$$f_{k-1}(\alpha) = -f_{k+1}(\alpha),$$

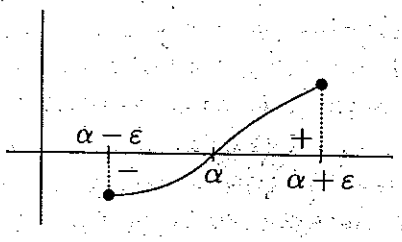
при этом в силу 1) $f_{k-1}(\alpha) \neq 0$. □

4) Если $f(\alpha) = 0$ для $\alpha \in \mathbb{R}$, то многочлен $f_0(x)f_1(x)$ при переходе через $x = \alpha$ меняет знак $-$ на $+$, т. е. график многочлена $f_0(x)f_1(x)$ в достаточно малой окрестности $O(\alpha) = (\alpha - \varepsilon, \alpha + \varepsilon)$ точки α имеет следующий вид:

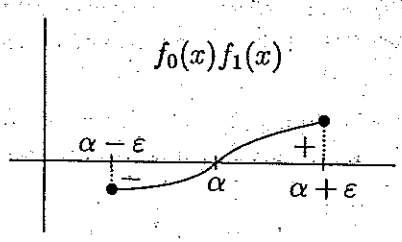


Доказательство. Так как $f_0(\alpha) = f(\alpha) = 0$, то в силу 1) $f_1(\alpha) = f'(\alpha) \neq 0$.

Случай а). Пусть $f_1(\alpha) = f'(\alpha) > 0$. Тогда $f_1(x) = f'(x) > 0$ во всех точках x достаточно малой окрестности $O(\alpha) = (\alpha - \varepsilon, \alpha + \varepsilon)$ точки α. Таким образом, в этой окрестности $O(\alpha)$ функция $f(x)$ строго возрастающая, т. е.

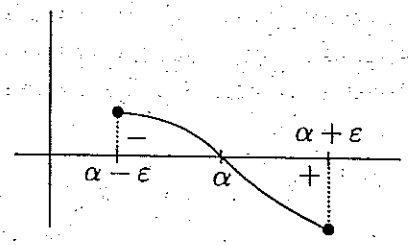


Умножая в этой окрестности на непрерывную функцию $f_1(x)$, где $f_1(x) > 0$, получаем, что функция $f_0(x)f_1(x)$

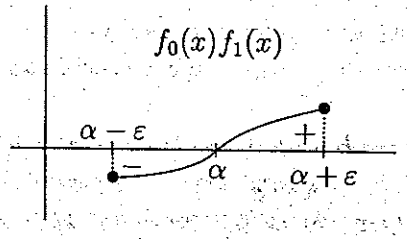


переходит со знака $-$ на $+$ в достаточно малой окрестности $O(\alpha)$ точки α.

Случай б). Пусть $f_1(\alpha) = f'(\alpha) < 0$. Тогда $f_1(x) = f'(x) < 0$ во всех точках x достаточно малой окрестности $O(\alpha) = (\alpha - \varepsilon, \alpha + \varepsilon)$ точки α. Таким образом, в этой окрестности $O(\alpha)$ функция $f(x)$ строго убывающая, т. е.



Умножая в этой окрестности на непрерывную функцию $f_1(x)$, где $f_1(x) < 0$, получаем, что функция $f_0(x)f_1(x)$



переходит со знака $-$ на $+$ в достаточно малой окрестности $O(\alpha)$ точки α. □

3.24. Число перемен знаков в системе значений многочленов системы Штурма

Если $f(x) \in \mathbb{R}[x]$, $\big(f(x), f'(x)\big) = 1$,

$$f_0(x) = f(x),\ f_1(x) = f'(x), \ldots,\ f_s(x) -$$

система многочленов Штурма для многочлена $f(x)$, $c \in \mathbb{R}$, то в ряду действительных чисел

$$f_0(c),\ f_1(c), \ldots,\ f_s(c)$$

выбросим нулевые значения и подсчитаем число перемен знаков $W(c)$ в оставшемся ряду ненулевых действительных чисел.

Теорема Штурма. *Пусть $f(x) \in \mathbb{R}[x]$ — многочлен с действительными коэффициентами без кратных корней* (т. е. НОД$\big(f(x), f'(x)\big) = 1$),

$$f_0(x) = f(x),\ f_1(x) = f'(x), \ldots,\ f_s(x) -$$

его система Штурма, $a, b \in \mathbb{R}$, $a < b$ (возможно, $a = -\infty$, $b = +\infty$ для расширенной действительной оси), $f(a) \neq 0$, $f(b) \neq 0$. Тогда:

1) $W(a) \geqslant W(b)$;

2) *разность $W(a) - W(b)$ равна числу действительных корней между a и b* (т. е. *в интервале (a, b)*).

Доказательство. Проанализируем поведение знаков значений многочленов системы Штурма

$$f_0(\alpha),\ f_1(\alpha), \ldots,\ f_s(\alpha)$$

при движении $\alpha \in \mathbb{R}$ от $-\infty$ к $+\infty$.

Случай 1 (переход через корень α многочлена $f_k(\alpha)$, $1 \leqslant k \leqslant s-1$). Пусть $f_k(\alpha) = 0$ для $1 \leqslant k \leqslant s-1$. Тогда в силу свойства 1) $f_{k-1}(\alpha) \neq 0$, $f_{k+1}(\alpha) \neq 0$. Тогда в достаточно малой окрестности $O(\alpha) = (\alpha - \varepsilon, \alpha + \varepsilon)$ многочлены $f_{k-1}(x)$ и $f_{k+1}(x)$ не имеют корней, и поэтому сохраняют знаки своих значений, которые, в силу 3), противоположны.

Рассмотрим таблицу значений многочленов $f_{k-1}(x)$, $f_k(x)$, $f_{k+1}(x)$ в точках $\alpha - \varepsilon$, α, $\alpha + \varepsilon$:

$$\begin{array}{c|ccc} \alpha - \varepsilon & f_{k-1}(\alpha - \varepsilon) & f_k(\alpha - \varepsilon) & f_{k+1}(\alpha - \varepsilon) \\ \alpha & f_{k-1}(\alpha) & f_k(\alpha) = 0 & f_{k+1}(\alpha) \\ \alpha + \varepsilon & f_{k-1}(\alpha + \varepsilon) & f_k(\alpha + \varepsilon) & f_{k+1}(\alpha + \varepsilon) \end{array},$$

в которой для среднего столбца возможны столбцы

$$\begin{pmatrix} - \\ 0 \\ + \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} + \\ 0 \\ - \end{pmatrix},$$

крайние столбцы (первый и третий) имеют постоянные противоположные знаки, т. е.

$$\begin{pmatrix} - \\ - \\ - \end{pmatrix} \text{ и } \begin{pmatrix} + \\ + \\ + \end{pmatrix}, \quad \text{или} \quad \begin{pmatrix} + \\ + \\ + \end{pmatrix} \text{ и } \begin{pmatrix} - \\ - \\ - \end{pmatrix}.$$

Таким образом, во всех возможных четырёх случаях при переходе через α от $\alpha-\varepsilon$ к $\alpha+\varepsilon$ число перемен знаков $W(x)$ не меняется (оно равно 1 для нашей таблицы):

$$\begin{pmatrix} - & - & + \\ - & 0 & + \\ - & + & + \end{pmatrix}, \quad \begin{pmatrix} + & - & - \\ + & 0 & - \\ + & + & - \end{pmatrix};$$

$$\begin{pmatrix} - & + & + \\ - & 0 & + \\ - & - & + \end{pmatrix}, \quad \begin{pmatrix} + & + & - \\ + & 0 & - \\ + & - & - \end{pmatrix}.$$

Случай 2 (переход через корень α многочлена $f_0(x) = f(x)$). Пусть $f_0(\alpha) = f(\alpha) = 0$ для $\alpha \in \mathbb{R}$. Тогда в силу 1) $f_1(\alpha) \neq 0$. Поэтому в достаточно малой окрестности $O(\alpha) = (\alpha-\varepsilon, \alpha+\varepsilon)$ точки α многочлен $f_1(x)$ сохраняет знак (не обращаясь в нуль).

Вариант а): $f_1(x) > 0$ для всех $x \in O(\alpha)$. Так как в силу 4) при переходе через α произведение $f(x)f_1(x)$ меняет знак $-$ на знак $+$, то многочлен $f(x)$ при переходе через α также меняет знак $-$ на знак $+$. Таким образом, возможна следующая таблица для

$$\begin{array}{c|cc} \alpha-\varepsilon & f_0(\alpha-\varepsilon) & f_1(\alpha-\varepsilon) \\ \alpha & f_0(\alpha) & f_1(\alpha) \\ \alpha+\varepsilon & f_0(\alpha+\varepsilon) & f_1(\alpha+\varepsilon) \end{array}:$$

$$\begin{pmatrix} - & + \\ 0 & + \\ + & + \end{pmatrix},$$

т. е. наш счётчик $W(x)$ в этой таблице при переходе через α уменьшил своё значение на 1 (от 1 к 0).

Вариант б): $f_1(x) < 0$ для всех $x \in O(\alpha)$. Так как в силу 4) при переходе через α произведение $f(x)f_1(x)$ меняет знак $-$ на знак $+$, то многочлен $f(x)$ при переходе через α меняет знак $+$ на знак $-$. Таким образом, возможна лишь следующая таблица:

$$\begin{pmatrix} + & - \\ 0 & - \\ - & - \end{pmatrix},$$

т. е. наш счётчик $W(x)$ в этой таблице при переходе через α уменьшил своё значение на 1 (от 1 к 0). □

Пример 3.24.1. $f(x) = x^3 + 3x^2 - 1 \in \mathbb{R}[x]$. Тогда $f_1(x) = f'(x) = 3x^2 + 6x$. Ясно, что НОД$(f(x), f'(x)) = 1$. Далее по алгоритму Евклида $f_2(x) = 2x + 1$, $f_3(x) = 1$. Следовательно,

	f_0	f_1	f_2	f_3	$W(x)$
$x = -\infty$	$-$	$+$	$-$	$+$	3
$x = +\infty$	$+$	$+$	$+$	$+$	0

,

т. е. $f(x)$ имеет три действительных корня.

Более того, теорема Штурма является эффективным средством (в комбинации с определением границ корней) для решения *проблемы локализации* (указания интервалов, содержащих ровно один действительный корень; это позволяет к этому интервалу применять алгоритмы нахождения корня). Например, в нашем случае, так как $x^3+3x^2 = x^2(x+3) > 1$ при $x \geqslant 1$ и для $x = -z$ многочлен $-f(z) = z^3 - 3z^2 - 1$ при $z \geqslant 4$ не имеет корней ($z^2(z-3) > 1$ при $z \geqslant 4$), то все действительные корни многочлена $f(x)$ принадлежат интервалу $(-4, 1)$. Более точно,

	f_0	f_1	f_2	f_3	$W(x)$
-3	$-$	$+$	$-$	$+$	3
-1	$+$	$-$	$-$	$+$	2
0	$-$	0	$+$	$+$	1
1	$+$	$+$	$+$	$+$	0

,

т. е. интервалы $(-3, -1)$, $(-1, 0)$, $(0, 1)$ содержат в точности по одному действительному корню.

Замечание 3.24.2. Теорема Штурма справедлива для любой *системы многочленов Штурма*

$$\varphi_0(x),\ \varphi_1(x), \ldots,\ \varphi_s(x), \quad \text{где} \quad \varphi_0(x) = f(x),$$

удовлетворяющих рассмотренным свойствам 1)—4) последовательности многочленов в теореме Штурма.

Пример 3.24.3. Система

$$\lambda_0 f_0(x),\ \lambda_1 f_1(x), \ldots,\ \lambda_s f_s(x),$$

где $\lambda_0, \ldots, \lambda_s > 0$, $f_0, \ldots, f_s$ — каноническая система Штурма, является системой Штурма.

Замечание 3.24.4 (система Штурма для отрезка $[a, b]$). Последовательность ненулевых многочленов

$$f_0(x) = f(x),\ f_1(x), \ldots,\ f_s(x)$$

называется системой Штурма для многочлена $f(x)$ на отрезке $[a, b]$, если выполнены следующие условия:

1) если $f_k(c) = 0$ для $c \in [a, b]$ и $1 \leqslant k \leqslant s-1$, то $f_{k-1}(c) f_{k+1}(c) < 0$;

2) последний многочлен $f_s(x)$ не имеет корней на $[a, b]$;

3) $f_0(a) f_0(b) \neq 0$;

4) если $f(c) = 0$ для $c \in [a, b]$, то функция $f_0(x) f_1(x)$ меняет знак с $-$ на $+$, если x, возрастая, проходит через точку c.

Теорема Штурма для интервала. *Число действительных корней многочлена $f(x)$ степени $n \geqslant 1$ на интервале (a, b) равно $W(a) - W(b)$ (для любой фиксированной системы Штурма для многочлена $f(x)$ на отрезке $[a, b]$).*

Доказательство аналогично приведённому доказательству теоремы Штурма. □

Замечание 3.24.5 (теорема Декарта). Так как для $0 < c \in \mathbb{R}$ число перемен знаков в системе коэффициентов многочлена $f(x) \in \mathbb{R}[x]$ меньше числа перемен знаков в системе коэффициентов произведения $(x-c)f(x)$ на нечётное число, то из этого выводится теорема Декарта (более слабая оценка, чем в теореме Штурма): число положительных корней многочлена $f(x) \in \mathbb{R}[x]$, засчитываемых каждый столько раз, какова его кратность, равно числу перемен знаков в системе коэффициентов этого многочлена или меньше этого числа на чётное число (если все корни многочлена $f(x)$ действительны, то число положительных корней равно числу перемен знаков в системе коэффициентов многочлена $f(x)$).

Замечание 3.24.6 (теорема Бюдана—Фурье). Оценку для числа корней на интервале (a, b) (а не только на интервале $(0, +\infty)$, как в теореме Декарта) даёт следующая теорема Бюдана—Фурье (Бюдан, 1822 г.; Фурье, 1820 г.).

Пусть $f(x) \in \mathbb{R}[x]$, $\deg f(x) = n$, $c \in \mathbb{R}$, $f(c) \neq 0$, $S_+(c)$ — число перемен знаков в системе чисел

$$f(c),\ f'(c), \ldots,\ f^{(n)}(c). \tag{3.2}$$

Если $f^{(k)}(c) = f^{(k+1)}(c) = \ldots = f^{k+l-1}(c) = 0$, но $f^{(k-1)}(c) \neq 0$, $f^{(k+l)}(c) \neq 0$, то считаем, что $f^{(k+i)}(c)$, $0 \leqslant i \leqslant l-1$, имеет такой же знак, как $f^{(k+l)}(c)$, если разность $l - i$ чётная, и противоположный знак, если это число нечётно, а через $S_-(c)$ обозначим число перемен знаков в системе (3.2) с учётом знаков, приписанных этим способом числам $f^{(k)}(c), \ldots, f^{(k+l-1)}(c)$.

Теорема 3.24.7 (Бюдан—Фурье). *Пусть $f(x) \in \mathbb{R}[x]$, $a, b \in \mathbb{R}$, $a < b$, $f(a) \neq 0$, $f(b) \neq 0$. Тогда число корней многочлена $f(x)$ в интервале (a, b), подсчитываемых каждый столько раз, какова его кратность, равно $S_+(a) - S_-(b)$ или меньше этого числа на чётное число.*

Следствие 3.24.8. *Пусть все корни многочлена $f(x) \in \mathbb{R}[x]$ вещественны, $a, b \in \mathbb{R}$, $f(a) \neq 0$, $f(b) \neq 0$. Тогда число корней многочлена $f(x)$, лежащих в интервале (a, b), равно $W\big(f_a(x)\big) - W\big(f_b(x)\big)$, где*

$$f_a(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} x^k = f(x+a),$$

$$f_b(x) = \sum_{k=0}^{n} \frac{f^{(k)}(b)}{k!} x^k = f(x+b),$$

$W\big(f_a(x)\big)$ и $W\big(f_b(x)\big)$ — число перемен знаков в системе коэффициентов многочленов $f_a(x)$ и $f_b(x)$ соответственно.

Упражнение 3.24.9. Число перемен знаков в системе коэффициентов многочлена $f(x) = x^3 - 3x + 1$ равно 2. Поэтому число положительных корней многочлена $f(x)$ равно либо 2, либо 0. Число перемен знаков в системе коэффициентов многочлена $f(-x) = -x^3 + 3x + 1$ равно 1, следовательно, число отрицательных корней многочлена $f(x)$ равно 1. Дополнительно, вычисляя дискриминант кубического многочлена $f(x)$, получаем, что все корни многочлена $f(x)$ действительны.

Упражнение 3.24.10. Для характеристического многочлена $f(x) = -x^3 + 14x + 20$ симметрической матрицы

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 3 & 1 \end{pmatrix}$$

известно, что все корни действительны. Поэтому, применяя теорему Декарта, получаем, что число положительных корней многочлена $f(x)$ равно 1. Так как число перемен знаков в системе коэффициентов многочлена $f(-x) = x^3 - 14x + 20$ равно 2, то число отрицательных корней многочлена $f(x)$ равно 2.

Упражнение 3.24.11. Пусть

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

(урезанная экспонента), $a, b \in \mathbb{R}$, $a < b < 0$. Положим $\varphi_0(x) = f(x)$, $\varphi_1(x) = f'(x)$, $\varphi_2(x) = -f(x) + f'(x) = -\dfrac{x^n}{n!}$.

Нетрудно видеть, что система $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$ является системой Штурма для многочлена $f(x)$ на отрезке $[a, b]$. Полагая, что число a достаточно большое по модулю, а число b достаточно малое, применяя теорему Штурма для отрезка, получаем, что при чётном n многочлен $f(x)$ не имеет вещественных корней, а при нечётном n имеет один отрицательный корень. При $n = 3$ каноническая система Штурма содержит четыре многочлена

$$f_0(x) = f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6},$$
$$f_1(x) = f'(x) = 1 + x + \frac{x^2}{2},$$
$$f_2(x) = -\left(\frac{1}{3}x + \frac{2}{3}\right),$$
$$f_3(x) = -1,$$

а построенная выше система Штурма для отрезка $[a, b]$ содержит три многочлена $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$.

Задача 3.24.12. Пусть $f \in \mathbb{Z}[x]$, $\deg f = n$, старший коэффициент многочлена f равен 1. Докажите, что невозможна ситуация, когда многочлен f имеет n корней (считая кратность) на интервале $(m, m+1)$, где $n \in \mathbb{Z}$.

Упражнение 3.24.13. *Многочлены Лежандра* (сферические функции) определяются рекуррентной формулой

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0,$$

где $P_0(x) = 1$, $P_1(x) = x$, или формулами

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n((x^2-1)^n)}{dx^n}.$$

Нетрудно показать, что при всех n $P_n(x)$ — многочлен степени n, а также что $P_n(x)$ — интеграл дифференциального уравнения

$$(x^2-1)y'' + 2xy' - n(n+1)y = 0;$$

$$\frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} P_n(x) z^n.$$

Покажите, что система многочленов

$$P_m(x),\ P_{m-1}(x), \ldots,\ P_1(x),\ P_0(x) -$$

система Штурма для многочлена $f(x) = P_m(x)$.

Упражнение 3.24.14. Покажите с использованием теоремы Бюдана—Фурье, что многочлен

$$f(x) = 3x^8 - 2x^5 + x^4 + 4x^2 - x - 1$$

имеет ровно один корень на интервале $(-2, 0)$.

Задача 3.24.15. Каждый действительный многочлен $f(x)$ можно представить в виде рациональной дроби, числитель и знаменатель которой — действительные многочлены, в знаменателе нет перемены знаков, а в числителе число перемен знаков равно числу положительных корней многочлена $f(x)$.

Указание. $f(x) = \dfrac{f(x)(1+x)^k}{(1+x)^k}$ для достаточного большого k.

Задача 3.24.16 (теорема Шура). Пусть f — многочлен степени n, имеющий лишь действительные корни, λ_n — наибольший корень многочлена f, λ_{n-k} — наибольший корень производной $f^{(k)}$, $k = 1, 2, \ldots, n-1$. Тогда

$$\lambda_n - \lambda_{n-1} \leqslant \lambda_{n-1} - \lambda_{n-2} \leqslant \ldots \leqslant \lambda_2 - \lambda_1.$$

Задача 3.24.17 (А. Г. Хованский). Пусть $f \in \mathbb{R}[x]$, $\deg f \geqslant 1$, $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, $a_n \neq 0$,

$$j(f) = \begin{cases} 0, & \text{если } n \text{ — чётное число,} \\ 1, & \text{если } n \text{ нечётное и } a_n > 0, \\ -1, & \text{если } n \text{ нечётное и } a_n < 0. \end{cases}$$

Рациональная дробь $\dfrac{f'}{f}$ однозначно может быть представлена в виде непрерывной дроби

$$\frac{f'}{f} = \left[0; \frac{1}{f_1}, \frac{1}{f_2}, \ldots, \frac{1}{f_k}\right].$$

Покажите, что число вещественных корней многочлена $f(x)$ (без учёта кратностей) равно

$$j(f_1) - j(f_2) + \ldots + (-1)^{k-1} j(f_k).$$

Например, пусть $f(x) = -x^3 + 1$. Тогда

$$f'(x) = -3x^2,$$

$$\frac{f'(x)}{f(x)} = \frac{-3x^2}{-x^3+1} = \frac{1}{\left(\frac{x^3-1}{3x^2}\right)} = \frac{1}{\frac{1}{3}x + \frac{1}{-3x^2}},$$

$$f_1(x) = \frac{1}{3}x, \quad f_2(x) = -3x^2.$$

$$j(f_1) = 1, \quad j(f_2) = 0, \quad j(f_1) - j(f_2) = 1,$$

и многочлен $f(x)$ имеет один действительный корень.

3.25. Приближённое вычисление корней многочленов с действительными коэффициентами

Метод деления отрезка (метод дихотомии, или метод вилки)

Пусть многочлен $f(x) \in \mathbb{R}[x]$ имеет единственный корень на отрезке $[a, b]$ (пусть $f(x)$ имеет разные знаки в точках a и b; этого можно достичь, отделяя кратные корни многочлена $f(x)$). Деля на части отрезок $[a, b]$ и определяя знак многочлена $f(x)$ в точках деления, мы можем сужать отрезок, содержащий корень, и осуществлять приближённое вычисление корня. Однако скорость сходимости этого метода мала (по сравнению с другими методами).

Метод непрерывных дробей

Пусть действительный многочлен $f(x) \in \mathbb{R}[x]$ степени n имеет простой (то есть кратности 1) корень c в интервале $(a, a+1)$. Тогда $c - a \in (0, 1)$. Положим $y = \dfrac{1}{x-a}$, то есть $x = \dfrac{1}{y} + a$. Многочлен $g(y) = y^n f\left(\dfrac{1}{y} + a\right)$ имеет корень $b = \dfrac{1}{c-a}$ в интервале $(1, +\infty)$. Пусть $a_1 < b < a_1 + 1$, $z = \dfrac{1}{y - a_1}$, $d = \dfrac{1}{b - a_1}$, $a_2 < d < a_2 + 1$, $w = \dfrac{1}{z - a_2}$, $e = \dfrac{1}{d - a_2}$, $a_3 < e < a_3 + 1$ и так далее. Тогда:

$$a + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + 1}}} < c < a + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3}}}.$$

Продолжая этот процесс, получаем всё более точные приближения корня c непрерывными дробями.

Метод хорд и метод Ньютона

Пусть функция $f(x)$ и её производные $f'(x)$, $f''(x)$ непрерывны на отрезке $[a, b]$, $f(a)f(b) < 0$, производные $f'(x)$ и $f''(x)$ сохраняют знак на $[a, b]$. В частности, из этого следует, что существует такая точка c, $a < c < b$, что $f(c) = 0$; функция $f(x)$ возрастает (или убывает) на $[a, b]$, поэтому корень c единственен; график функции $f(x)$ выпуклый вниз (вверх) на $[a, b]$.

Эти предположения относительно функции $f(x)$ легко могут быть реализованы для многочленов: достаточно отделить кратные корни многочлена $f(x)$, затем локализовать корни многочлена $f(x)$. Так как $f'(x)$ и $f''(x)$ — многочлены, то, отделяя корни многочлена $f(x)$ и применяя теорему Штурма к многочленам $f(x)$ и $f'(x)$, мы рассматриваем такой отрезок $[a, b]$, что он содержит лишь один корень c многочлена $f(x)$ (кратности один), $f(a)f(b) < 0$ и многочлен $f'(x)$ не имеет корней на отрезке $[a, b]$. Если при этом $f''(c) = 0$, то задача о приближённом вычислении корня c многочлена $f(x)$ сводится к вычислению корня c многочлена $f''(x)$, имеющего меньшую степень, чем многочлен $f(x)$. Если же $f''(c) \neq 0$, то, вновь применяя теорему Штурма и сужая отрезок $[a, b]$, мы приходим к ситуации, когда $f(a)f(b) < 0$, многочлен $f(x)$ имеет лишь один корень внутри отрезка $[a, b]$, многочлены $f'(x)$ и $f''(x)$ сохраняют знак на $[a, b]$.

Метод хорд. За приближённое значение корня принимается число

$$x_1 = a - \frac{(b-a)f(a)}{f(b)-f(a)} = b - \frac{(b-a)f(b)}{f(b)-f(a)}.$$

Геометрически это означает, что вместо корня c, где точка $(c, 0)$ является точкой пересечения графика функции $f(x)$ с осью абсцисс, берётся пересечение с осью абсцисс хорды, соединяющей точки $\big(a, f(a)\big)$ и $\big(b, f(b)\big)$. Продолжая этот процесс, можно построить приближение корня с любой степенью точности.

Так, например, если $f' > 0$, $f'' > 0$ на $[a, b]$, $f(a) < 0$, $f(b) > 0$, то $f(x_1) < 0$, и последовательные приближения вычисляем по формуле

$$x_{n+1} = x_n - \frac{(b-x_n)f(x_n)}{f(b)-f(x_n)}.$$

При этом для любого n: $a < x_n < x_{n+1} < c$, где $f(c) = 0$, $\lim\limits_{n\to\infty} x_n = c$, $|x_n - c| \leqslant \dfrac{|f(x_n)|}{m}$, где m — наименьшее значение функции $|f'(x)|$ на отрезке $[a, b]$.

Метод хорд часто называют методом пропорциональных частей, методом линейной интерполяции или методом ложного положения.

Более эффективным методом является метод Ньютона.

Метод Ньютона (метод касательных, или метод линеаризации). Для первого приближения к корню c положим

$$x_1 = b - \frac{f(b)}{f'(b)}.$$

Геометрически x_1 — это абсцисса точки пересечения с осью x касательной к графику функции $f(x)$ в точке $\big(b, f(b)\big)$. Если $f(b)$ одного знака с $f''(x)$ на $[a, b]$, то x_1 лежит между c и b ($f(c) = 0$). Если $f(a)$ одного знака с $f''(x)$, то полагаем

$$x_1' = a - \frac{f(a)}{f'(a)}$$

(и тогда x_1' — абсцисса точки пересечения с осью x касательной к графику функции $f(x)$ в точке $\big(a, f(a)\big)$, x_1' лежит между a и c).

Повторяя этот процесс, мы получаем последовательность убывающих чисел x_n, $b > x_n > x_{n+1} > c$ (или последовательность возрастающих чисел x_n', $a < x_n' < x_{n+1}' < c$), где

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

$$x_{n+1}' = x_n' - \frac{f(x_n')}{f'(x_n')},$$

при этом $\lim\limits_{n\to\infty} x_n = c$ ($\lim\limits_{n\to\infty} x_n' = c$). Если, как и раньше, m — наименьшее значением функции $|f'(x)|$ на $[a, b]$, M — наибольшее значение функции $|f''(x)|$ на $[a, b]$, то

$$|x_{n+1} - c| \leqslant \frac{M}{2m}|x_n - c|^2$$

(аналогичная формула справедлива для x_n'). Этим обеспечивается достаточно близкое приближение x_n к c (x_n' к c).

При определённых условиях метод Ньютона можно использовать для приближённого вычисления комплексных корней комплексных многочленов.

Задача о глобальной сходимости метода Ньютона для рациональных эндоморфизмов римановой сферы является открытой проблемой.

Упражнение 3.25.1. Покажите, что метод Ньютона для нахождения корней многочлена $x^2 - x - 1$ с начальным условием $x_0 = 1$ сходится к корню $\frac{1+\sqrt{5}}{2}$, а с начальным условием $x_0 = 0$ — к корню $\frac{1-\sqrt{5}}{2}$.

Иногда целесообразно применять **модифицированный метод Ньютона**, в котором последовательные приближения определяются формулами

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_1)},$$

то есть $(x_{n+1}, 0)$ — точка пересечения прямой, проходящей через точку $(x_n, f(x_n))$ и имеющей угловой коэффициент $f'(x_1)$, с осью x. Эта прямая лишь на первом шаге совпадает с касательной к графику функции $f(x)$. В случае применения модифицированного метода Ньютона вычислительная схема упрощается, однако скорость сходимости модифицированного метода Ньютона меньше, чем обычного метода Ньютона.

Часть при практических вычислениях удобно одновременно использовать метод Ньютона и метод хорд.

Метод секущих использует итерационный процесс

$$x_{n+1} = x_n - \frac{f(x_n)(x_{n-1} - x_n)}{f(x_{n-1}) - f(x_n)},$$

при этом приближения x_0 и x_1 задаются заранее. Приближение x_0 можно выбирать, как и в методе Ньютона, а приближение x_1 брать вблизи точки x_0, между x_0 и предполагаемым корнем. Достоинством метода секущих является очень простая вычислительная схема, однако в плане численной устойчивости он проигрывает методу Ньютона.

Пример 3.25.2. Многочлен $f(x) = x^3 - 2x^2 - 4x - 7$ имеет единственный корень внутри отрезка $[3, 4]$, на этом отрезке его производные $f'(x)$ и $f''(x)$ сохраняют знак, наименьшее значение $f'(x)$ на этом отрезке $m = 11$. Для первого приближения по методу хорд полагаем

$$x_1 = 3 - \frac{f(3)}{f(4) - f(3)} = 3 + \frac{10}{19}.$$

Для вычисления корня с точностью до 0,01, используя оценку

$$|x_n - c| \leqslant \frac{|f(x_n)|}{m},$$

где $f(c) = 0$, необходимо совершить три шага метода:

$$x_3 \approx 3{,}63.$$

При использовании метода Ньютона полагаем

$$x_1 = 4 - \frac{f(4)}{f'(4)} = 4 - \frac{9}{28}.$$

Уже $x_2 \approx 3{,}63$ даёт приближение корня с точностью до 0,01.

Конец лекции N 18.

Лекция № 19-20-21 (22.XI.2011) (25.XI.2011)

Метод итерации

Пусть $f(x) \in \mathbb{C}[x]$. Преобразуем уравнение $f(x) = 0$ к виду $x = \varphi(x)$ (либо перенося a_1x в правую часть и деля на $-a_1$ при $a_1 \neq 0$, либо полагая $x = x + df(x)$, $d \neq 0$).

Пусть x_0 — начальное приближение к корню уравнения. Определим последовательность x_n, полагая $x_{n+1} = \varphi(x_n)$, $n = 0, 1, 2, \ldots$. Можно показать, что если для любых x', x'' из круга $\{x \in \mathbb{C} \mid |x - x_0| \leqslant \delta\}$ функция $\varphi(x)$ удовлетворяет неравенству

$$|\varphi(x') - \varphi(x'')| \leqslant q|x' - x''|,$$

где $0 < q < 1$, и выполнено неравенство

$$\frac{m}{1-q} \leqslant \delta,$$

где $m = |x_0 - \varphi(x_0)|$, то уравнение $x = \varphi(x)$ имеет в круге $\{x \in \mathbb{C} \mid |x - x_0| \leqslant \delta\}$ единственный корень c, к которому сходится последовательность $\{x_n\}$, при этом скорость сходимости определяется неравенством

$$|x_n - c| \leqslant \frac{m}{1-q} q^n.$$

Замечание 3.25.3. Существуют и другие методы приближённого вычисления корней многочленов. Например, метод Лобачевского (1834), иногда называемый методом Лобачевского—Греффе—Данделена, имеющий довольно сложную вычислительную схему, не требует знания начального приближения к корню многочлена и применим к нахождению начального приближения. Метод Лобачевского позволяет одновременно приближённо вычислять все корни многочлена.

3.26. Проблема Рауса—Гурвица, устойчивые многочлены

Многочлен $f(z) \in \mathbb{C}[z]$ называется *устойчивым*, если все его комплексные корни лежат в левой полуплоскости (т. е. вещественные части a всех корней $c = a + bi$ многочлена $f(z)$ отрицательны, $a < 0$). Этот класс многочленов возник в теории устойчивости движения (в частности, для применения теоремы А. А. Ляпунова). Ясно, что делитель устойчивого многочлена устойчив, произведение устойчивых многочленов также является устойчивым.

Замечание 3.26.1. Пусть дана линейная однородная система дифференциальных уравнений

$$\frac{d\hat{X}}{dt} = A \cdot \hat{X}$$

с постоянной матрицей $A \in \mathrm{M}_n(\mathbb{R})$,

$$\hat{X} = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t). \end{pmatrix}$$

Решение этой системы даётся формулой

$$X_{\hat{C}}(t) = e^A \cdot \hat{C} \quad (t \geqslant 0),$$

см. ?? (начальное условие $\hat{X}(0) = \hat{C} \in \hat{\mathbb{R}}_n$).

Система $\dfrac{d\hat{X}}{dt} = A \cdot \hat{X}$ асимптотически устойчива, если все решения $\hat{X} = \hat{X}(t)$ обладают свойством $\lim\limits_{t\to\infty} \hat{X}(t) = (0)$.

Критерий А. М. Ляпунова асимптотической устойчивости системы: система $\dfrac{d\hat{X}}{dt} = A \cdot \hat{X}$ асимптотически устойчива тогда и только тогда, когда все собственные числа матрицы A имеют отрицательные вещественные части (иными словами, когда характеристический многочлен матрицы A устойчив).

Теорема А. М. Ляпунова (1892). *Пусть $A \in \mathrm{M}_n(\mathbb{R})$. Тогда все собственные числа матрицы A имеют отрицательные действительные части в том и только в том случае, когда существует (и единственна) такая положительно определённая симметрическая матрица $B \in \mathrm{M}_n(\mathbb{R})$ ($B^t = B$, $X \cdot V \cdot X^t > 0$ для любой строки $X \in \mathbb{R}^n$), что*

$$A^t \cdot B + B \cdot A = -E.$$

Эта теорема позволяет свести задачу об определении устойчивости системы к решению системы линейных уравнений.

Проблема Рауса—Гурвица заключается в следующем: по коэффициентам $a_0, a_1, \ldots, a_n$ многочлена $f(z) = \sum\limits_{i=0}^{n} a_i z^i \in \mathbb{C}[z]$ определить, устойчив ли он.

Бо́льшая часть известных критериев устойчивости относится к многочленам с действительными коэффициентами. Это объясняется возможностью редукции общей проблемы к случаю действительных коэффициентов.

Лемма 3.26.2 (о редукции проблемы устойчивости для многочленов с комплексными коэффициентами к случаю многочленов с вещественными коэффициентами). *Пусть*

$$f(z) = a_n z^n + \ldots + a_1 z + a_0 \in \mathbb{C}[z],$$
$$\bar{f}(z) = \bar{a}_n z^n + \ldots + \bar{a}_1 z + \bar{a}_0 \in \mathbb{C}[z],$$
$$g(z) = f(z)\bar{f}(z).$$

Тогда:

1) *многочлен $f(z)$ устойчив тогда и только тогда, когда устойчив многочлен $g(z) = f(z)\bar{f}(z)$;*

2) *$g(z) = f(z)\bar{f}(z) \in \mathbb{R}[z]$, т. е. $g(z)$ — многочлен с действительными коэффициентами.*

Доказательство. Так как отображение $c = a + bi \mapsto \bar{c} = a - bi$ является автоморфизмом поля комплексных чисел $\mathbb{C}$, то отображение $f \mapsto \bar{f}$ является автоморфизмом кольца многочленов $\mathbb{C}[z]$.

Если

$$f(z) = a_n z^n + \ldots + a_0 = a_n(z - c_1)\ldots(z - c_n),$$

где $c_1, \ldots, c_n \in \mathbb{C}$ — корни многочлена $f(z)$, то

$$\bar{f}(z) = \bar{a}_n z^n + \ldots + \bar{a}_0 = \bar{a}_n(z - \bar{c}_1)\ldots(z - \bar{c}_n),$$

где корень $c = a + bi$ многочлена $f(z)$ соответствует корню $\bar{c} = a - bi$ многочлена $\bar{f}(z)$, при этом вещественная часть a у них одна и та же. Таким образом, многочлен $f(z)$ устойчив тогда и только тогда, когда устойчив многочлен $g(z) = f(z)\bar{f}(z)$.

Так как

$$g(z) = f(z)\bar{f}(z) = a_n\bar{a}_n \prod_{i=1}^{n}(z - c_i)(z - \bar{c}_i), \quad a_n\bar{a}_n \in \mathbb{R},$$

$$(z - c)(z - \bar{c}) = \big(z - (a + bi)\big)\big(z - (a - bi)\big) = z^2 - 2a + (a^2 + b^2) \in \mathbb{R}[z],$$

то $g(z) \in \mathbb{R}[z]$, т. е. $g(z)$ — многочлен с действительными коэффициентами. □

Лемма 3.26.3. *Если многочлен*

$$f(z) = a_n z^n + \ldots + a_1 z + a_0 = a_n(z - c_1)\ldots(z - c_n) \in \mathbb{C}[z]$$

с комплексными коэффициентами ($a_n \neq 0$, $c_1, \ldots, c_n$ — *его корни*, $c_j = a_j + b_j i$) *является устойчивым* (*т. е.* $a_j < 0$, $j = 1, \ldots, n$), *то:*

1) $a_0 \neq 0$;

2) $\operatorname{Re}\left(\dfrac{a_1}{a_0}\right) > 0$.

Доказательство. 1) Так как $f(0) = a_0$, то $a_0 = 0$ тогда и только тогда, когда 0 является корнем многочлена $f(z)$. Так как многочлен $f(z)$ устойчивый, то $a_0 \neq 0$.

2) Так как $c_j \neq 0$, $j = 1, \ldots, n$, и

$$\frac{1}{c_j} = \frac{a_j}{a_j^2 + b_j^2} - \frac{b_j}{a_j^2 + b_j^2} i,$$

то

$$\operatorname{Re}\left(\frac{1}{c_j}\right) = \frac{a_j}{a_j^2 + b_j^2} < 0.$$

Поэтому

$$\operatorname{Re}\left(\sum_{j=1}^{n} \frac{1}{c_j}\right) < 0.$$

В силу следствия к теореме Виета:

$$-\frac{a_1}{a_0} = \sum_{j=1}^{n} \frac{1}{c_j},$$

следовательно,

$$\operatorname{Re}\left(-\frac{a_1}{a_0}\right) = \operatorname{Re}\left(\sum_{j=1}^{n} \frac{1}{c_j}\right) < 0.$$ □

Лемма 3.26.4 (свойство левой полуплоскости в $\mathbb{C}$). *Если* $z = a + bi, w = u + vi \in \mathbb{C}$, $a = \operatorname{Re}(z) < 0$, $u = \operatorname{Re}(w) < 0$, *то*

$$|z + \bar{w}| > |z - w|.$$

Доказательство.

$$|z+\bar{w}|^2-|z-w|^2=|(a+u)+(b-v)i|^2-|(a-u)+(b-v)i|^2=$$
$$=(a+u)^2+(b-v)^2-(a-u)^2-(b-v)^2=4au>0,$$

следовательно,

$$|z+\bar{w}|>|z-w|.$$

□

Замечание 3.26.5. Если

$$f(z)=a_nz^n+\ldots+a_1z+a_0\in\mathbb{C}[z],$$

то определим

$$\tilde{f}(z)=z^nf\left(\frac{1}{z}\right)=a_0z^n+\ldots+a_{n-1}z+a_n$$

(«обращённый» многочлен). Так как знаки $\mathrm{Re}(c)$ и $\mathrm{Re}\left(\frac{1}{c}\right)$ для $0\neq c\in\mathbb{C}$ совпадают, то многочлен $f(z)$ устойчив тогда и только тогда, когда устойчив многочлен $\tilde{f}(z)$.

Теорема 3.26.6 (теорема А. Стодолы, 1894 г.). *Если многочлен*

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0\in\mathbb{R}[x]$$

с действительными коэффициентами степени n, $a_n>0$, является устойчивым, то все его коэффициенты $a_n,a_{n-1},\ldots,a_1,a_0$ положительны.

Доказательство. Пусть

$$f(x)=a_nx^n+\ldots+a_1x+a_0=a_n(x-c_1)\ldots(x-c_n),$$

где $c_j=a_j+b_ji\in\mathbb{C}$, $j=1,\ldots,n$, — корни многочлена $f(x)$. Так как многочлен $f(x)$ устойчивый, то $a_j<0$ для всех $j=1,\ldots,n$.

Отрицательному действительному корню $c\in\mathbb{R}$, $c<0$, соответствует множитель $x-c$ с положительными коэффициентами.

Паре ненулевых сопряжённых корней $a+bi$, $a-bi$, где $a<0$, в каноническом разложении соответствует множитель

$$\big(x-(a+bi)\big)\big(x-(a-bi)\big)=x^2-2ax+a^2+b^2$$

с положительными коэффициентами.

Так как $a_n>0$ и произведение многочленов с положительными коэффициентами является многочленом с положительными коэффициентами, то $a_n>0$, $a_{n-1}>0,\ldots,a_1>0$, $a_0>0$.

□

Замечание 3.26.7. Многочлен $f(x)\in\mathbb{R}[x]$ степени 1 или 2 с вещественными коэффициентами и с положительным старшим членом устойчив тогда и только тогда, когда все его коэффициенты положительны.

Действительно, для многочлена первой степени a_1x+a_0, $a_1>0$, его единственный корень $-\frac{a_0}{a_1}$ отрицателен тогда и только тогда, когда $a>0$.

Многочлен второй степени $a_2x^2 + a_1x + a_0$, $a_2 > 0$, имеет корни

$$\frac{-a_1 \pm \sqrt{a_1^2 - 4a_0a_2}}{2a_2}.$$

Если дискриминант $D = a_1^2 - 4a_0a_2 < 0$, то $a_0 > 0$ и вещественная часть обоих корней равна $-\dfrac{a_1}{2a_2}$, т. е. устойчивость многочлена в этом случае равносильна тому, что $a_1 > 0$.

Если же $D = a_1^2 - 4a_0a_2 > 0$, то оба корня вещественны. Если $a_1 > 0$ и $a_0 > 0$, то оба корня отрицательны.

Пример 3.26.8 (показывающий, что при степени $\geqslant 3$ положительность коэффициентов недостаточна для его устойчивости). Многочлен

$$x^3 + x^2 + 4x + 30 \in \mathbb{R}[x]$$

имеет *все положительные коэффициенты, но не является устойчивым* (его корни -3, $1 \pm 3i$; среди них два корня имеют положительную вещественную часть).

Упражнение 3.26.9 (критерий Вышнеградского устойчивости многочлена третьей степени). Многочлен с вещественными коэффициентами

$$f = a_3x^3 + a_2x^2 + a_1x + a_0 \in \mathbb{R}[x],$$

$a_3 > 0$, устойчив тогда и только тогда, когда все его коэффициенты положительны (т. е. $a_3 > 0$, $a_2 > 0$, $a_1 > 0$, $a_0 > 0$) и имеет место неравенство $a_1a_2 > a_0a_3$.

Упражнение 3.26.10. Многочлен

$$x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 \in \mathbb{R}[x]$$

устойчив тогда и только тогда, когда

$$a_3 > 0, \quad a_0 > 0, \quad a_3a_2 > a_1, \quad a_1(a_3a_2 - a_1) > a_3^2a_0.$$

Следующая теорема даёт прозрачный критерий устойчивости многочленов с действительными коэффициентами (продолжая линию: что надо добавить к условию положительности всех коэффициентов).

Теорема 3.26.11. *Пусть*

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \in \mathbb{R}[x] -$$

многочлен с действительными коэффициентами,

$$g(x) = x^m + b_{m-1}x^{m-1} + \ldots + b_1x + b_0 -$$

многочлен степени $m = \dfrac{n(n-1)}{2}$, *корнями которого служат все суммы пар корней многочлена* $f(x)$. *Тогда:*

1) $g(x) \in \mathbb{R}[x]$;

2) *многочлен $f(x)$ устойчив тогда и только тогда, когда все коэффициенты многочленов $f(x)$ и $g(x)$ положительны.*

Доказательство. 1) Многочлен $g(x)$ имеет действительные коэффициенты, поскольку его комплексные, не являющиеся действительными, корни разбиваются на пары сопряжённых корней. Действительно, если $r \in \mathbb{R}$, $c, d \in \mathbb{C} \setminus \mathbb{R}$, $f(r) = f(c) = f(d) = 0$, то $f(\bar{c}) = f(\bar{d}) = 0$, при этом $\dfrac{3(3-1)}{2} = 3$ суммы пар из $\{r, c, \bar{c}\}$, $c \neq \bar{c}$, имеют вид $c + \bar{c} \in \mathbb{R}$, $r + c, r + \bar{c} \in \mathbb{C} \setminus \mathbb{R}$, где $\overline{r+c} = r + \bar{c}$; $\dfrac{4(4-1)}{2} = 6$ сумм пар из $\{c, \bar{c}, d, \bar{d}\}$, $c \neq \bar{c}$, $d \neq \bar{d}$, имеют вид $c + \bar{c}, d + \bar{d} \in \mathbb{R}$, $c + d$, $\bar{c} + \bar{d} = \overline{c+d}$, $c + \bar{d}$, $\bar{c} + d = \overline{c + \bar{d}}$ (в случае кратных корней $c = d$: $c + \bar{d} = c + \bar{c}, \bar{c} + d = \bar{c} + c \in \mathbb{R}$).

2а) Если многочлен $f(x)$ устойчив, то по определению вещественные части всех его корней отрицательны, следовательно, вещественные части всех сумм пар его корней также отрицательны, т. е. многочлен $g(x)$ также устойчив.

Применяя к устойчивым многочленам $f(x), g(x) \in \mathbb{R}[x]$ теорему Стодолы, убеждаемся в том, что все коэффициенты многочленов $f(x)$ и $g(x)$ положительны.

2б) Допустим, что все коэффициенты многочленов $f(x)$ и $g(x)$ положительны. Тогда все действительные корни многочленов $f(x), g(x) \in \mathbb{R}[x]$ отрицательны.

Таким образом, если $c \in \mathbb{R}$ и $f(c) = 0$, то $c < 0$, если $c = a + bi, \bar{c} = a - bi \in \mathbb{C} \setminus \mathbb{R}$, $f(c) = f(\bar{c}) = 0$, — пара сопряжённых корней многочлена $f(x)$, то действительное число $2a = (a + bi) + (a - bi) \in \mathbb{R}$ является действительным корнем многочлена $g(x)$, и поэтому $2a < 0$, следовательно, $a < 0$.

Итак, мы показали, что действительные части всех корней многочлена $f(x)$ отрицательны, т. е. что $f(x)$ — устойчивый многочлен. □

3.27. Преобразования И. Шура

Пусть

$$f(z) = a_n z^n + \ldots + a_1 z + a_0 \in \mathbb{C}[z], \quad a_n \neq 0;$$
$$\bar{f}(z) = \bar{a}_n z^n + \ldots + \bar{a}_1 z + \bar{a}_0$$

(замена всех коэффициентов a_k на их сопряжённые $\bar{a}_k$);

$$f^*(z) = \bar{f}(-z),$$

т. е.

$$f^*(z) = (-1)^n \bar{a}_n z^n + (-1)^{n-1} \bar{a}_{n-1} z^{n-1} + \ldots + \bar{a}_0$$

(перемена знаков в многочлене $\bar{f}(z)$ при нечётных степенях переменной z).

Лемма 3.27.1. *Если $f, f_1, f_2 \in \mathbb{C}[z]$, то*

$$(f_1 + f_2)^* = f_1^* + f_2^*;$$
$$(f_1 f_2)^* = f_1^* f_2^*;$$
$$(cf)^* = \bar{c} f^* \quad \text{для} \quad c \in \mathbb{C};$$
$$f^{**} = f$$

(в частности, отображение $f(z) \mapsto f^*(z)$ является автоморфизмом кольца $\mathbb{C}[z]$).

Доказательство. Так как отображение $c \mapsto \bar{c}$ является автоморфизмом поля комплексных чисел $\mathbb{C}$, то (как мы видели) отображение $f \mapsto \bar{f}$ является автоморфизмом кольца многочленов $\mathbb{C}[z]$, при этом $\overline{(cf)}(z) = \bar{c}\bar{f}(z)$ для $c \in \mathbb{C}$. Далее:

$$(f_1 + f_2)^*(z) = \overline{(f_1 + f_2)}(-z) = (\bar{f}_1 + \bar{f}_2)(-z) = \bar{f}_1(-z) + \bar{f}_2(-z) = f_1^*(z) + f_2^*(z);$$
$$(f_1 f_2)^*(z) = \overline{(f_1 f_2)}(-z) = (\bar{f}_1 \bar{f}_2)(-z) = \bar{f}_1(-z)\bar{f}_2(-z) = f_1^*(z) f_2^*(z);$$
$$(cf)^* = \overline{(cf)}(-z) = (\bar{c}\bar{f})(-z) = \bar{c}\big(\bar{f}(-z)\big) = \bar{c} f^*(z);$$
$$f^{**}(z) = \big(\bar{f}(-z)\big)^* = f(z).$$

□

Следствие 3.27.2. *Если $c \in \mathbb{C}$, то $f(c) = 0$ тогда и только тогда, когда $f^*(-\bar{c}) = 0$, при этом кратности корней c и $-\bar{c}$ одинаковы.*

Доказательство. Пусть $c_1, \ldots, c_n \in \mathbb{C}$ — корни многочлена $f(z)$, т. е.

$$f(z) = a_n(z - c_1)\ldots(z - c_n).$$

Тогда (в силу леммы 3.27.1)

$$f^*(z) = \bar{a}_n(-z - \bar{c}_1)\ldots(z - \bar{c}_n) = (-1)^n \bar{a}_n\big(z - (-\bar{c}_1)\big)\ldots\big(z - (-\bar{c}_n)\big),$$

т. е. $-\bar{c}_1, \ldots, -\bar{c}_n$ — корни многочлена $f^*(z)$. □

Лемма 3.27.3. *Если $f(z) \in \mathbb{C}[z]$ — устойчивый многочлен, то для $c = a + bi \in \mathbb{C}$:*

1) $0 \leqslant |f(c)| < |f^*(c)|$, *если* $\mathrm{Re}(c) = a < 0$;

2) $0 \leqslant |f^*(c)| < |f(c)|$, *если* $\mathrm{Re}(c) = a > 0$;

3) $0 < |f(c)| = |f^*(c)|$, *если* $\mathrm{Re}(c) = a = 0$.

Доказательство. Пусть $c_1 = a_1 + b_1 i, \ldots, c_n = a_n + b_n i$ — все корни устойчивого многочлена $f(z)$, тогда $a_j < 0$ для всех $j = 1, \ldots, n$. Для $c \in \mathbb{C}$

$$\begin{aligned}
&|c + \bar{c}_j|^2 - |c - c_j|^2 = \\
&\quad = |(a + a_j) + (b - b_j)i|^2 - |(a - a_j) + (b - b_j)|^2 = \\
&\quad = (a + a_j)^2 + (b - b_j)^2 - (a - a_j)^2 - (b - b_j)^2 = 4aa_j.
\end{aligned}$$

Таким образом,

$$\begin{aligned}
&|c - c_j| < |c + \bar{c}_j|, \quad \text{если } a < 0;\\
&|c + \bar{c}_j| < |c - c_j|, \quad \text{если } a > 0;\\
&|c - c_j| < |c + \bar{c}_j|, \quad \text{если } a = 0.
\end{aligned}$$

Так как

$$\begin{aligned}
f(c) &= a_n(c - c_1)\ldots(c - c_n),\\
f^*(c) &= (-1)^n \bar{a}_n\big(c - (-\bar{c}_1)\big)\ldots\big(c - (-\bar{c}_n)\big),
\end{aligned}$$

то из полученных неравенств следуют утверждения 1)—3) леммы. □

Определение 3.27.4. Если $\alpha, \beta \in \mathbb{C}$, $|\alpha| > |\beta|$, то многочлен $g(z) = \alpha f(z) - \beta f^*(z)$ называется преобразованием Шура многочлена $f(z)$.

Замечание 3.27.5. Если многочлен $g(z)$ является преобразованием Шура многочлена $f(z)$, то многочлен $f(z)$ также будет преобразованием Шура многочлена $g(z)$.

Действительно, так как $g = \alpha f - \beta f^*$, то по лемме 3.27.1 $g^* = \bar{\alpha} f^* - \bar{\beta} f$. Поэтому если

$$\alpha_1 = \frac{\bar{\alpha}}{|\alpha|^2 - |\beta|^2}, \quad \beta_1 = \frac{-\beta}{|\alpha|^2 - |\beta|^2},$$

то

$$\alpha\alpha_1 + \bar{\beta}\beta_1 = \frac{\alpha\bar{\alpha} - \beta\bar{\beta}}{|\alpha|^2 - |\beta|^2} = \frac{|\alpha|^2 - |\beta|^2}{|\alpha|^2 - |\beta|^2} = 1;$$

$$\alpha_1\beta + \bar{\alpha}\beta_1 = \frac{\bar{\alpha}\beta - \bar{\alpha}\beta}{|\alpha|^2 - |\beta|^2} = 0.$$

Следовательно,

$$\alpha_1 g - \beta_1 g^* = \alpha_1(\alpha f - \beta f^*) - \beta_1(\bar{\alpha} f^* - \bar{\beta} f) = (\alpha\alpha_1 + \bar{\beta}\beta_1) f - (\alpha_1\beta + \bar{\alpha}\beta_1) f^* = f.$$

Ясно, что

$$|\alpha_1| = \frac{|\alpha|}{|\alpha|^2 - |\beta|^2} > \frac{|\beta|}{|\alpha|^2 - |\beta|^2} = |\beta_1|. \qquad \square$$

Предложение 3.27.6. *Многочлен $f(z) \in \mathbb{C}[z]$ устойчив тогда и только тогда, когда устойчивым является любое (некоторое) его преобразование Шура $g(z) = \alpha f(z) - \beta f^*(z)$, где $\alpha, \beta \in \mathbb{C}$, $|\alpha| > |\beta|$.*

Доказательство. 1) Пусть многочлен $f(z)$ устойчив. Если $c = a + bi \in \mathbb{C}$, $\mathrm{Re}(c) = a \geqslant 0$, то в силу леммы 3.27.3 (п. 2) и 3))

$$|f(c)| > |f^*(c)| \quad \text{или} \quad |f(c)| = |f^*(c)| > 0.$$

Так как $|\alpha| > |\beta|$, то

$$|\alpha f(c)| = |\alpha|\,|f(c)| > |\beta|\,|f^*(c)| = |\beta|\,|f^*(c)|.$$

Таким образом, если $g(c) = 0$, т. е. $g(c) = \alpha f(c) - \beta f^*(c) = 0$, следовательно $\alpha f(c) = \beta f^*(c)$, и поэтому $a < 0$. Итак, мы показали, что многочлен $f(z)$ устойчивый.

2) Если многочлен $g(z)$ устойчивый, то в силу замечания 3.27.5 многочлен $f(z)$ является преобразованием Шура многочлена $g(z)$, поэтому в силу 1) $f(z)$ — устойчивый многочлен. $\square$

Теорема 3.27.7 (редукция проблемы к многочлену меньшей степени). *Пусть $c = a + bi \in \mathbb{C}$, $\mathrm{Re}(c) = a < 0$, $f(z) \in \mathbb{C}[z]$, $\deg f(z) = n$. Многочлен $f(z)$ устойчив тогда и только тогда, когда*

1) $|f(c)| < |f^*(c)|$;

2) *многочлен*

$$F(z, c) = \frac{f^*(c) f(z) - f(c) f^*(z)}{z - c} \in \mathbb{C}[z]$$

степени $n - 1$ является устойчивым.

Доказательство. а) Пусть многочлен $f(z)$ устойчив. Так как $\mathrm{Re}(c) = a < 0$, то в силу п. 1 леммы 3.27.3

$$|f(c)| < |f^*(c)|.$$

Поэтому многочлен

$$g(z) = f^*(c)f(z) - f(c)f^*(z)$$

является преобразованием Шура устойчивого многочлена $f(z)$; по предложению 3.27.6 $g(z)$ — устойчивый многочлен.

Многочлен $g(z)$ имеет степень n (ясно, что $\deg g(z) \leqslant n$; его коэффициент при z^n равен $f^*(c)a_n - f(c)(-1)^n\bar{a}_n$; так как $|a_n| > 0$ и $|f(c)| < |f^*(c)|$, то

$$|f^*(c)a_n| = |f^*(c)|\,|a_n| < |f(c)|\,|a_n| = |f(c)|\,|\bar{a}_n| = |f(c)(-1)^n\bar{a}_n|,$$

и поэтому коэффициент при z^n многочлена $g(z)$ ненулевой).

Так как

$$g(c) = f^*(c)f(c) - f(c)f^*(c) = 0,$$

то c — корень многочлена $g(z)$, и поэтому многочлен $g(z)$ делится на $z - c$, $g(z) = (z - c)F(z, c)$,

$$F(z, c) = \frac{g(z)}{z - c} \in \mathbb{C}[z]$$

является многочленом степени $n - 1$. Так как $g(z)$ — устойчивый многочлен, то его делитель $F(z, c)$ также является устойчивым многочленом.

б) Пусть выполнены условия 1) и 2).

Так как

$$g(z) = f^*(c)f(z) - f(c)f^*(z) = (z - c)F(z, c)$$

и многочлен $F(z, c)$ устойчив (условие 2)), то устойчив и многочлен $g(z)$.

Так как $|f(c)| < |f^*(c)|$ (условие 1)), то многочлен

$$g(z) = f^*(c)f(z) - f(c)f^*(z)$$

является преобразованием Шура многочлена $f(z)$, при этом $g(z)$ — устойчивый многочлен. Согласно предложению 3.27.6 многочлен $f(z)$ также устойчив. □

Рассмотрим многочлен от двух переменных (z и c):

$$F_f(z, c) = F(z, c) = \frac{f^*(c)f(z) - f(c)f^*(z)}{z - c} \in \mathbb{C}[z, c].$$

Так как $F(z, c) = F(c, z)$, то многочлен $F(z, c)$ симметричен, и поэтому степень многочлена $F(z, c)$ по c равна его степени по z, т. е. равна $n - 1$, где $n = \deg f(z)$.

Разлагая многочлен $F(z, c)$ по степеням c^k переменной c, получаем

$$F(z, c) = F_{n-1}(z)c^{n-1} + \ldots + F_1(z)c + F_0(z),$$

где $F_k(z) \in \mathbb{C}[z]$, $\deg F_k(z) \leqslant n - 1$, $0 \leqslant k \leqslant n - 1$.

Многочлен

$$T_f(z, c) = T(z, c) = F_1(z)c + F_0(z)$$

назовём *c-хвостом* многочлена $F_f(z, c)$.

Лемма 3.27.8. *Если $T(z,c) = F_1(z)c + F_0(z)$ — c-хвост многочлена $F(z,c) \in \mathbb{C}[z,c]$,*

$$\varphi(z,c) = \bar{a}_0(z+c) - \bar{a}_1 zc \in \mathbb{C}[z,c],$$
$$\psi(z,c) = a_0(z+c) + a_1 zc \in \mathbb{C}[z,c],$$

то

$$z^2T(z,c) = f(z)\varphi(z,c) - f^*(z)\psi(z,c).$$

Доказательство. Так как

$$(z-c)F(z,c) = f^*(c)f(z) - f(c)f^*(z),$$

то

$$(z-c)\sum_{k=0}^{n-1} F_k(z)c^k = f^*(c)f(z) - f(c)f^*(z),$$

поэтому

$$-F_{n-1}(z)c^n + \sum_{k=1}^{n-1}[zF_k(z) - F_{k-1}(z)]c^k + zF_0(z) = \left(\sum_{k=0}^{n}(-1)^k\bar{a}_k c^k\right)f(z) - \left(\sum_{k=0}^{n} a_k c^k\right)f^*(z).$$

Приравнивая соответственно коэффициенты при нулевой (свободный член) и первой степенях по c, получаем

$$\begin{aligned} zF_0(z) &= \bar{a}_0 f(z) - a_0 f^*(z); \\ zF_1(z) - F_0(z) &= -\bar{a}_1 f(z) - a_1 f^*(z). \end{aligned} \qquad (*)$$

Отсюда

$$\begin{aligned} z^2T(z,c) &= z^2\big(F_1(z)c + F_0(z)\big) = \\ &= zc[zF_1(z) - F_0(z)] + (z+c)zF_0(z) = \\ &= zc[-\bar{a}_1 f(z) - a_1 f^*(z)] + (z+c)[\bar{a}_0 f(z) - a_0 f^*(z)] = \\ &= f(z)[\bar{a}_0(z+c) - \bar{a}_1 zc] - f^*(z)[a_0(z+c) + a_1 zc], \end{aligned}$$

то есть

$$z^2T(z,c) = f(z)\varphi(z,c) - f^*(z)\psi(z,c)$$

(утверждение леммы). □

Следствие 3.27.9.

$$z^2[T(z,c)\varphi^*(z,c) + T^*(z,c)\psi(z,c)] = f(z)[\varphi(z,c)\varphi^*(z,c) - \psi(z,c)\psi^*(z,c)].$$

Доказательство. При фиксированном $c \in \mathbb{C}$ применим к установленному в лемме 3.27.8 равенству автоморфизм $f \mapsto f^*$ кольца $\mathbb{C}[z]$ (см. лемму 3.27.1):

$$z^2T^*(z,c) = [z^2T(z,c)]^* = [f(z)\varphi(z,c) - f^*(z)\psi(z,c)] = f^*(z)\varphi^*(z,c) - f(z)\psi^*(z,c).$$

Следовательно,

$$\begin{aligned} z^2[T(z,c)\varphi^*(z,c) + T^*(z,c)\psi(z,c)] &= [f(z)\varphi(z,c) - f^*(z)\psi(z,c)]\varphi^*(z,c) + \\ &+ [f^*(z)\varphi^*(z,c) - f(z)\psi^*(z,c)]\psi(z,c) = f(z)[\varphi(z,c)\varphi^*(z,c) - \psi(z,c)\psi^*(z,c)]. \end{aligned}$$ □

Теорема 3.27.10 (критерий устойчивости Шура). *Пусть*

$$f(z) = a_n z^n + \ldots + a_1 z + a_0 \in \mathbb{C}[z] -$$

многочлен с комплексными коэффициентами степени n, $a_n \neq 0$. Тогда:

1) *если многочлен $f(z)$ устойчив, то $a_0 \neq 0$, $\operatorname{Re}\left(\frac{a_1}{a_0}\right) > 0$ и для всех $c \in \mathbb{C}$ таких, что $\operatorname{Re}(c) < 0$, c-хвост $T(z,c) = F_1(z)c + F_0(z)$ многочлена*

$$F(z,c) = \frac{f^*(c)f(z) - f(c)f^*(z)}{z - c}$$

устойчив (как многочлен степени $n-1$ из $\mathbb{C}[z]$);

2) *если $a_0 \neq 0$, $\operatorname{Re}\left(\frac{a_1}{a_0}\right) > 0$ и многочлен $T(z,c) \in \mathbb{C}[z]$ устойчив хотя бы для одного $c \in \mathbb{C}$ такого, что $\operatorname{Re}(c) < 0$, то многочлен $f(z) \in \mathbb{C}[z]$ также устойчив.*

Доказательство. 1а) Если $f(z)$ — устойчивый многочлен, то в силу леммы 3.26.3 $a_0 \neq 0$; $\operatorname{Re}\left(\frac{a_1}{a_0}\right) > 0$.

Для доказательства устойчивости многочлена $T(z,c)$ (как многочлена из $\mathbb{C}[z]$) для всех $0 \neq c \in \mathbb{C}$ таких, что $\operatorname{Re}(c) < 0$, надо показать, что $T(d,c) \neq 0$ для любого $d \in \mathbb{C}$ такого, что $\operatorname{Re}(d) \geqslant 0$.

Случай а). Пусть $F_0(d) \neq 0$. Рассмотрим многочлен

$$\begin{aligned}\Phi(t) &= t^{n-1} F\left(d, \frac{1}{t}\right) = \\ &= t^{n-1}\left(F_{n-1}(d)\frac{1}{t^{n-1}} + \ldots + F_0(d)\right) = \\ &= F_{n-1}(d) + F_{n-2}(d)t + \ldots + F_1(d)t^{n-2} + F_0(d)t^{n-1} \in \mathbb{C}[t]\end{aligned}$$

(так как $F_0(d) \neq 0$, то $\deg \Phi(t) = n-1$).

Заметим, что действительные части всех корней t_0 многочлена $\Phi(t)$ неотрицательны. В самом деле, если $0 \neq t_0 \in \mathbb{C}$, $\operatorname{Re}(t_0) < 0$, то $\operatorname{Re}\left(\frac{1}{t_0}\right) < 0$. В силу теоремы 3.27.7 (см. п. 2)) многочлен

$$F\left(z, \frac{1}{t_0}\right) = f_{\frac{1}{t_0}}(z) \in \mathbb{C}[z]$$

устойчив. Так как $\operatorname{Re}(d) \geqslant 0$, то $F\left(d, \frac{1}{t_0}\right) \neq 0$, следовательно, $\Phi(t_0) = t^{n-1} F\left(d, \frac{1}{t_0}\right) \neq 0$.

По теореме Виета сумма всех корней многочлена $\Phi(t) \in \mathbb{C}[t]$, $\Phi(t) = F_0(d)t^{n-1} + F_1(d)t^{n-2} + \ldots + F_{n-1}(d)$, $F_0(d) \neq 0$, равна $-\frac{F_1(d)}{F_0(d)}$, и поэтому

$$\operatorname{Re}\left(-\frac{F_1(d)}{F_0(d)}\right) \geqslant 0.$$

Допустим противное: $T(d,c)=0$, то есть $F_1(d)c+F_0(d)=0$, где $0\neq c\in\mathbb{C}$, $\mathrm{Re}(c)<0$. Тогда $-\dfrac{F_1(d)}{F_0(d)}=\dfrac{1}{c}$, и поэтому $\mathrm{Re}\left(\dfrac{1}{c}\right)\geqslant 0$, следовательно, $\mathrm{Re}(c)\geqslant 0$, что противоречит нашему выбору числа $c\in\mathbb{C}$.

Таким образом, мы показали, что $T(d,c)\neq 0$ для любого $d\in\mathbb{C}$ такого, что $\mathrm{Re}(d)\geqslant 0$ и $F_0(d)\neq 0$.

Случай б). Пусть $F_0(d)=0$.

Допустим противное: $T(d,c)=0$, то есть $F_1(d)c+F_0(d)=0$, где $0\neq c\in\mathbb{C}$, $\mathrm{Re}(c)<0$. В силу равенств $(*)$ из доказательства леммы 3.27.8 имеем

$$0=\bar{a}_0f(d)-a_0f^*(d);$$
$$0=\bar{a}_1f(d)+a_1f^*(d).$$

Поэтому

$$(\bar{a}_0a_1+a_0\bar{a}_1)f(d)=0.$$

Так как многочлен $f(z)\in\mathbb{C}[z]$ устойчив и $\mathrm{Re}(d)\geqslant 0$, то $f(d)\neq 0$. Поэтому $\bar{a}_0a_1+a_0\bar{a}_1=0$, при этом $0\neq a_0\in\mathbb{C}$. Следовательно, $\dfrac{a_1}{a_0}=-\overline{\left(\dfrac{a_1}{a_0}\right)}$, то есть $\mathrm{Re}\left(\dfrac{a_1}{a_0}\right)=0$, что противоречит уже установленному неравенству $\mathrm{Re}\left(\dfrac{a_1}{a_0}\right)>0$. Итак, мы пришли к противоречию, тем самым доказано, что и в случае б) имеем $T(d,c)\neq 0$ для всех $0\neq c\in\mathbb{C}$, $d\in\mathbb{C}$, $\mathrm{Re}(c)<0$, $\mathrm{Re}(d)\geqslant 0$.

1б) Покажем, что степень нашего многочлена

$$F(z,c)=\frac{f^*(c)f(z)-f(c)f^*(z)}{z-c}\in\mathbb{C}[z]$$

равна в точности $n-1$, если $a_0\neq 0$ и $\mathrm{Re}\left(\dfrac{a_1}{a_0}\right)>0$.

Ясно, что $\deg F(z,c)\leqslant n-1$. Допустим, что $\deg F(z,c)<n-1$. Тогда коэффициент при z^{n+1} в многочлене из леммы 3.27.8

$$z^2T(z,c)=f(z)\varphi(z,c)-f^*(z)\psi(z,c)$$

равен нулю, то есть

$$a_n(\bar{a}_0-\bar{a}_1c)-(-1)^n\bar{a}_n(a_0+a_1c)=0.$$

Поэтому ($|a_n|=|\bar{a}_n|\neq 0$)

$$|\bar{a}_0-\bar{a}_1c|=|a_0+a_1c|.$$

Однако если $z=\dfrac{1}{c}$, $w=-\dfrac{a_1}{a_0}$, поскольку $\mathrm{Re}(z)<0$, $\mathrm{Re}(w)<0$, то $|z+\bar{w}|>|z-w|$ (см. лемму 3.26.4), т. е.

$$\left|\frac{1}{c}-\frac{\bar{a}_1}{\bar{a}_0}\right|>\left|\frac{1}{c}+\frac{a_1}{a_0}\right|.$$

Умножая это неравенство на $|\bar{a}_0c|=|a_0c|$, получаем

$$|\bar{a}_0-\bar{a}_1c|>|a_0+a_1c|,$$

что противоречит установленному ранее равенству.

2) Допустим теперь, что $a_0 \neq 0$, $\operatorname{Re}\left(\dfrac{a_1}{a_0}\right) > 0$ и для некоторого $c \in \mathbb{C}$ такого, что $\operatorname{Re}(c) < 0$, многочлен

$$T(z,c) = F_1(z)c + F_0(z) \in \mathbb{C}[z]$$

устойчив. Покажем, что многочлен $f(z) \in \mathbb{C}[z]$ устойчив. Допустим противное, то есть предположим, что $f(d) = 0$ для некоторого $0 \neq d \in \mathbb{C}$ такого, что $\operatorname{Re}(d) \geqslant 0$. В силу следствия 3.27.9

$$d^2[T(d,c)\varphi^*(d,c) + T^*(d,c)\psi(d,c)] = 0.$$

Следовательно,

$$T(d,c)\varphi^*(d,c) + T^*(d,c)\psi(d,c) = 0.$$

Поэтому

$$|T(d,c)| \cdot |\varphi^*(d,c)| = |T^*(d,c)| \cdot |\psi(d,c)|.$$

Так как многочлен $T(z,c) \in \mathbb{C}[z]$ устойчив и $\operatorname{Re}(d) \geqslant 0$, то по лемме 3.27.3 (п. 2) и 3))

$$0 \leqslant |T^*(d,c)| < |T(d,c)|$$

или

$$0 < |T(d,c)| = |T^*(d,c)|.$$

Поэтому

$$|\varphi^*(d,c)| \leqslant |\psi(d,c)|. \tag{3.3}$$

Однако если $z, w \in \mathbb{C}$ и $\operatorname{Re}(z) < 0$, $\operatorname{Re}(w) < 0$, то

$$|z + \bar{w}| > |z - w|$$

(см. лемму 3.26.4).

Если $z = -\dfrac{a_1}{a_0} - \dfrac{1}{d}$, $w = \dfrac{1}{c}$, то: так как $\operatorname{Re}(c) < 0$, то $\operatorname{Re}(w) < 0$; так как $\operatorname{Re}(d) \geqslant 0$ и $\operatorname{Re}\left(\dfrac{a_1}{a_0}\right) > 0$, то $\operatorname{Re}(z) < 0$. Поэтому

$$\left|\frac{a_1}{a_0} + \frac{1}{d} - \frac{1}{\bar{c}}\right| = \left|-\frac{a_1}{a_0} - \frac{1}{d} + \frac{1}{\bar{c}}\right| > \left|-\frac{a_1}{a_0} - \frac{1}{d} - \frac{1}{c}\right| = \left|\frac{a_1}{a_0} + \frac{1}{d} + \frac{1}{c}\right|.$$

Умножая это неравенство на $|a_0 d\bar{c}| = |a_0 dc|$, получаем

$$|a_1 d\bar{c} + a_0\bar{c} - a_0 d| > |a_1 dc + a_0 c + a_0 d|.$$

Так как (см. лемму 3.27.8)

$$\varphi(z,c) = \bar{a}_0(z+c) - \bar{a}_1 zc, \quad \psi(z,c) = a_0(z+c) + a_1 zc,$$

то

$$\varphi^*(z,c) = a_0(-z + \bar{c}) + a_1 z\bar{c},$$

и поэтому

$$\begin{aligned} |\varphi^*(d,c)| &= |a_0(-d + \bar{c}) + a_1 d\bar{c}| = \\ &= |a_1 d\bar{c} + a_0\bar{c} - a_0 d| > |a_1 dc + a_0 c + a_0 d| = \\ &= |a_0(d+c) + a_1 dc| = |\psi(d,c)|. \end{aligned}$$

Это приводит нас к противоречию с $|\varphi^*(d,c)| \leqslant |\psi(d,c)|$ (см. (3.3)), что завершает доказательство теоремы. □

Следствие 3.27.11. *Пусть*

$$f(x) = a_n x^n + \ldots + a_1 x + a_0 \in \mathbb{R}[x], \quad a_n > 0,$$
$$\tilde{f}(x) = x^n f\left(\frac{1}{x}\right) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n,$$

$c = -1$, $f^{(1)}_{(x)} = \widetilde{\frac{1}{2} T_{\tilde{f}}(x, -1)}$. *Тогда многочлен $f(x)$ устойчив тогда и только тогда, когда $a_{n-1} > 0$ и многочлен $f^{(1)}_{(x)}$ устойчив.*

Доказательство. В силу замечания 3.26.5 многочлен $f(x)$ устойчив тогда и только тогда, когда устойчив многочлен $\tilde{f}(x)$.

Применяя теорему Шура к $\tilde{f}(x)$ (при $c = -1$), убеждаемся в том, что устойчивость многочлена $\tilde{f}(x)$ равносильна условиям

$$a_n \neq 0, \quad \operatorname{Re} \frac{a_{n-1}}{a_n} > 0, \quad T_{\tilde{f}}(x, -1) \text{ — устойчивый многочлен,}$$

то есть, поскольку $a_n > 0$, равносильна условиям

$$a_{n-1} > 0, \quad f^{(1)}(x) = \widetilde{\frac{1}{2} T_{\tilde{f}}(x, -1)} \text{ — устойчивый многочлен.} \qquad \square$$

3.28. Теорема Гурвица об устойчивых многочленах (с действительными коэффициентами)

Из критерия устойчивости Шура мы выведем сейчас один из наиболее ярких результатов теории устойчивых многочленов — теорему Гурвица. Мы ограничимся рассмотрением случая многочлена с действительными коэффициентами.

Пусть

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x],$$

$a_n > 0$, $\deg f(x) = n$.

Матрицей Гурвица многочлена $f(x)$ называется квадратная $(n \times n)$-матрица

$$H(f) = \begin{pmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \ldots & a_{-n+1} \\ a_n & a_{n-2} & a_{n-4} & \ldots & a_{-n+2} \\ 0 & a_{n-1} & a_{n-3} & \ldots & a_{-n+3} \\ 0 & a_n & a_{n-2} & \ldots & a_{-n+4} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_0 \end{pmatrix} \in \mathrm{M}_n(\mathbb{R}),$$

при чётном $n = 2k$ последняя строка имеет вид

$$\begin{pmatrix} 0 & \ldots & 0 & a_n & a_{n-2} & \ldots & a_0 \end{pmatrix},$$

при нечётном $n = 2k + 1$ последняя строка имеет вид

$$\begin{pmatrix} 0 & \ldots & 0 & a_{n-1} & a_{n-3} & \ldots & a_0 \end{pmatrix},$$

при этом предполагается, что $a_k = 0$ при $k < 0$ и при $k > n$, таким образом: $H(f) = (h_{ij}) \in \mathrm{M}_n(\mathbb{R})$, где $h_{ij} = a_{n+i-2j}$, а i-я строка матрицы $H(f)$ имеет вид

$$H_i = (a_{n+i-2}, a_{n+i-4}, a_{n+i-6}, \ldots, a_{n+i-2j}, \ldots, a_{-n+i}).$$

Определителями Гурвица многочлена $f(x) \in \mathbb{R}[x]$ называются главные миноры матрицы Гурвица $H(f)$ многочлена $f(x)$:

$$D_1 = a_{n-1}, \quad D_2 = \begin{vmatrix} a_{n-1} & a_{n-3} \\ a_n & a_{n-2} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{n-1} & a_{n-3} & a_{n-5} \\ a_n & a_{n-2} & a_{n-4} \\ 0 & a_{n-1} & a_{n-3} \end{vmatrix}, \ldots, \quad D_n = |H(f)|.$$

Предложение 3.28.1 (связь определителей Гурвица многочленов $f(x), f^{(1)}(x) \in \mathbb{R}[x]$). *Пусть*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x], \quad a_n > 0, \quad \deg f(x) = n;$$
$$\tilde{f}(x) = x^n f\left(\frac{1}{x}\right) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{R}[x];$$
$$f^{(1)}(x) = \frac{1}{2} T_{\tilde{f}}(x, -1) = b_{n-1}x^{n-1} + \ldots + b_1 x + b_0 \in \mathbb{R}[x]$$
$$(T_{\tilde{f}}(x, -1) = 2b_{n-1} + 2b_{n-2}x + \ldots + 2b_1 x^{n-2} + 2b_0 x^n);$$

$D_1, D_2, \ldots, D_n$ — главные миноры матрицы Гурвица $H(f) \in \mathrm{M}_n(\mathbb{R})$; $D_1^{(1)}, D_2^{(1)}, \ldots, D_{n-1}^{(1)}$ — главные миноры матрицы Гурвица $H(f^{(1)}) \in \mathrm{M}_{n-1}(\mathbb{R})$. Тогда для любого индекса $j = 1, \ldots, n-1$ имеем

$$D_j^{(1)} = \begin{cases} (a_n a_{n-1})^{\frac{j-1}{2}} D_{j+1}, & \text{если } j \text{ нечётно}, \\ a_n (a_n a_{n-1})^{\frac{j-2}{2}} D_{j+1}, & \text{если } j \text{ чётно}. \end{cases}$$

Доказательство. Пусть

$$\tilde{g}(x) = a_n + a_{n-2}x^2 + \ldots;$$
$$\tilde{h}(x) = a_{n-1}x + a_{n-3}x^3 + \ldots.$$

Тогда

$$\tilde{f}(x) = \tilde{g}(x) + \tilde{h}(x), \quad \tilde{f}^*(x) = \tilde{g}(x) - \tilde{h}(x).$$

В силу леммы 3.27.8 (применённой к $\tilde{f}(x)$, $c = -1$)

$$\begin{aligned} x^2 T_{\tilde{f}}(x, -1) &= \tilde{f}(x)\varphi(x, -1) - \tilde{f}^*\psi(x, -1) = \\ &= \big(\tilde{g}(x) + \tilde{h}(x)\big)\big(a_n(x-1) + a_{n-1}x\big) - \\ &- \big(\tilde{g}(x) - \tilde{h}(x)\big)\big(a_n(x-1) - a_{n-1}x\big) = \\ &= 2a_{n-1}x\tilde{g}(x) + 2a_n(x-1)\tilde{h}(x). \end{aligned}$$

Так как

$$T_{\tilde{f}}(x, -1) = 2b_{n-1} + \ldots + 2b_1 x^{n-2} + 2b_0 x^{n-1},$$

то, сокращая на 2, получим

$$x^2(b_{n-1} + \ldots + b_0 x^n) = a_{n-1}x(a_n + a_{n-2}x^2 + \ldots) + a_n(x-1)(a_{n-1}x + a_{n-3}x^3 + \ldots).$$

Приравнивая коэффициенты при степенях x^k получаем

$$\begin{aligned}
&b_{n-1} = a_n a_{n-1} \quad (\text{при } x^2),\\
&b_{n-2} = a_{n-1}a_{n-2} - a_n a_{n-3} \quad (\text{при } x^3),\\
&\dots\\
&\left.\begin{aligned}
&b_{n-(2k-1)} = a_n a_{n-(2k-1)} \quad (\text{при } x^{2k}),\\
&b_{n-2k} = a_{n-1}a_{n-2k} - a_n a_{n-(2k+1)} \quad (\text{при } x^{2k+1})
\end{aligned}\right\} \; k = 1, 2, \dots,\\
&\dots
\end{aligned}$$

Это означает, что матрица Гурвица многочлена $f^{(1)}(x)$ имеет вид

$$H(f^{(1)}) = \begin{pmatrix} b_{n-2} & b_{n-4} & b_{n-6} & \dots \\ b_{n-1} & b_{n-3} & b_{n-5} & \dots \\ 0 & b_{n-2} & b_{n-4} & \dots \\ \dots & b_{n-1} & b_{n-3} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} =$$

$$= \begin{pmatrix} a_{n-1}a_{n-2} - a_n a_{n-3} & a_{n-1}a_{n-4} - a_n a_{n-5} & a_{n-1}a_{n-6} - a_n a_{n-7} & \dots \\ a_n a_{n-1} & a_n a_{n-3} & a_n a_{n-5} & \dots \\ 0 & a_{n-1}a_{n-2} - a_n a_{n-3} & a_{n-1}a_{n-4} - a_n a_{n-5} & \dots \\ 0 & a_n a_{n-1} & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \in \mathrm{M}_{n-1}(\mathbb{R}),$$

при этом, в наших обозначениях, её главные миноры имеют вид

$$\begin{aligned}
&D_1^{(1)} = b_{n-2} = a_{n-1}a_{n-2} - a_n a_{n-3};\\
&D_2^{(1)} = \begin{vmatrix} b_{n-2} & b_{n-4} \\ b_{n-1} & b_{n-3} \end{vmatrix} = \begin{vmatrix} a_{n-1}a_{n-2} - a_n a_{n-3} & a_{n-1}a_{n-4} - a_n a_{n-5} \\ a_n a_{n-1} & a_n a_{n-3} \end{vmatrix};\\
&\dots\\
&D_{n-1}^{(1)} = |H(f^{(1)})|.
\end{aligned}$$

Рассмотрим следующее окаймление матрицы $H(f^{(1)})$ до $(n \times n)$-матрицы:

$$C = \begin{pmatrix} a_n a_{n-1} & a_n a_{n-3} & a_n a_{n-5} & \dots \\ 0 & a_{n-1}a_{n-2} - a_n a_{n-3} & a_{n-1}a_{n-4} - a_n a_{n-5} & \dots \\ 0 & a_n a_{n-1} & a_n a_{n-3} & \dots \\ 0 & 0 & a_{n-1}a_{n-2} - a_n a_{n-3} & \dots \\ 0 & 0 & a_n a_{n-1} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \in \mathrm{M}_n(\mathbb{R}),$$

при этом главные миноры матрицы C равны

$$a_n a_{n-1}, \quad a_n a_{n-1} D_1^{(1)}, \quad a_n a_{n-1} D_2^{(1)}, \dots, \quad a_n a_{n-1} D_{n-1}^{(1)}.$$

Прибавляя в матрице C первую строку ко второй ($C_2' = C_2 + C_1$), третью строку к четвёртой ($C_4' = C_4 + C_3$) и т. д., мы приходим к матрице

$$C' = \begin{pmatrix} a_n a_{n-1} & a_n a_{n-3} & a_n a_{n-5} & \dots \\ a_n a_{n-1} & a_{n-1}a_{n-2} & a_{n-1}a_{n-4} & \dots \\ 0 & a_n a_{n-1} & a_n a_{n-3} & \dots \\ 0 & a_n a_{n-1} & a_{n-1}a_{n-2} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

с теми же главными минорами, что и для матрицы C.

Вынося из нечётных строк матрицы C' общий множитель a_n, а из чётных строк общий множитель a_{n-1}, получаем матрицу Гурвица $H(f)$ многочлена $f(x)$. Поэтому главные миноры матрицы C' (а следовательно, и матрицы C) имеют вид

$$a_n a_{n-1}, \quad a_n a_{n-1} D_2, \ldots, \quad a_n^k a_{n-1}^k D_{2k}, \quad a_n^{k+1} a_{n-1}^k D_{2k+1}, \ldots.$$

Сравнивая выражения главных миноров матрицы через $D_k^{(1)}$ и через D_k, получаем утверждение предложения. □

Теорема 3.28.2 (теорема Гурвица (1895 г.) для многочленов с действительными коэффициентами). *Пусть*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x],$$

$a_n > 0$, $D_1, \ldots, D_n$ — *главные миноры матрицы Гурвица* $H(f)$ *многочлена* $f(x)$. *Многочлен* $f(x)$ *устойчив тогда и только тогда, когда*

$$D_1 > 0, \quad D_2 > 0, \ldots, \quad D_n > 0.$$

Доказательство. Проведём индукцию по степени $n = \deg f(x)$. Для $n = 1$ утверждение очевидно. Пусть утверждение теоремы Гурвица верно для всех многочленов, степень которых $\leqslant n-1$.

Пусть теперь $f(x) = a_n x^n + \ldots + a_0 \in \mathbb{R}[x]$, $\deg f(x) = n$, $a_n > 0$. В силу следствия к теореме Шура, $f(x)$ — устойчивый многочлен тогда и только тогда, когда $a_{n-1} > 0$ и устойчив многочлен $f^{(1)}(x)$ степени $\leqslant n-1$. Используя индуктивное предположение, получаем, что многочлен $f(x)$ устойчив тогда и только тогда, когда $a_{n-1} > 0$ и все определители Гурвица $D_1^{(1)}, \ldots, D_{n-1}^{(1)}$ матрицы Гурвица $H(f^{(1)})$ положительны (старший коэффициент b_{n-1} многочлена $f^{(1)}(x)$ равен $a_n a_{n-1}$ (см. доказательство замечания), и, поскольку $a_n > 0$, условие $a_{n-1} > 0$ равносильно тому, что $b_{n-1} > 0$). Согласно предложению 3.28.1 неравенства

$$a_{n-1} > 0, \quad D_1^{(1)} > 0, \ldots, \quad D_{n-1}^{(1)} > 0$$

(с учётом $a_n > 0$) равносильны неравенствам

$$D_1 = a_{n-1} > 0, \quad D_2 > 0, \ldots, \quad D_n > 0.$$

Таким образом, многочлен $f(x)$ устойчив тогда и только тогда, когда

$$D_1 > 0, \quad D_2 > 0, \ldots, \quad D_n > 0.$$

□

Упражнение 3.28.3. Сформулируйте критерий Гурвица для многочленов $f(z) \in \mathbb{C}[z]$ с комплексными коэффициентами.

Упражнение 3.28.4 (теорема Льенара—Шипара, 1914 г.). Пусть

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x],$$

$a_n > 0$, тогда

1) если $n = 2m+1$ — нечётное число, то многочлен $f(x)$ устойчив тогда и только тогда, когда

$$D_2 > 0, \quad D_4 > 0, \ldots, \quad D_{n-1} > 0;$$

2) если $n = 2m$ — чётное число, то многочлен $f(x)$ устойчив тогда и только тогда, когда

$$D_1 > 0, \quad D_3 > 0, \ldots, \quad D_{n-1} > 0.$$

Следствие 3.28.5. *Положительность всех определителей Гурвица чётного порядка равносильна положительности всех определителей Гурвица нечётного порядка.*

Пример 3.28.6. $f(x) = x^3 + 2x^2 + 3x + 1 \in \mathbb{R}[x]$, $a_3 = 1 > 0$,

$$H(f) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix},$$

$$D_1 = 2 > 0, \ D_2 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5 > 0, \ D_3 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix} = 5 > 0,$$

следовательно, $f(x)$ — устойчивый многочлен.

Пример 3.28.7. $f(x) = x^3 + 2x^2 + x + 3 \in \mathbb{R}[x]$, $a_3 = 1 > 0$,

$$H(f) = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 1 & 0 \\ 0 & 2 & 3 \end{pmatrix},$$

$$D_2 = \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = -1 < 0,$$

следовательно, многочлен $f(x)$ не является устойчивым (хотя все его коэффициенты положительны), т. е. хотя бы один его корень лежит в правой полуплоскости (поскольку у него нет чисто мнимых корней).

Пример 3.28.8 (схема Рауса, 1875 г.). На основе теории Штурма и теории индексов английский механик Раус предложил алгоритм для определения числа k корней многочлена $f(x) \in \mathbb{R}[x]$ с действительными коэффициентами, расположенных в правой полуплоскости $\{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$; при $k = 0$ этот алгоритм даёт критерий устойчивости.

Пусть

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x],$$

$a_n > 0$. Составим таблицу Рауса

$$R(f) = \begin{pmatrix} r_0^{(1)} & r_1^{(1)} & r_2^{(1)} & \ldots \\ r_0^{(2)} & r_1^{(2)} & r_2^{(2)} & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ r_0^{(p)} & r_1^{(p)} & r_2^{(p)} & \ldots \end{pmatrix},$$

где первая строка R_1 равна $H_2 = (a_n a_{n-2}, a_{n-4}, \ldots)$ (второй строке матрицы Гурвица $H(f)$); вторая строка R_2 равна $H_1 = (a_{n-1}, a_{n-3}, a_{n-5}, \ldots)$ (первой строке матрицы Гурвица $H(f)$); при $i \geqslant 3$ строка R_i определяется по формуле

$$r_j^{(i)} = r_{j+1}^{(i-2)} - \frac{r_0^{(i-2)}}{r_0^{(i-1)}} r_{j+1}^{(i-1)}$$

(т. е. $R_{i-2} - \frac{r_0^{(i-2)}}{r_0^{(i-1)}} R_{i-1} = (0, R_i)$, другими словами: из $(i-2)$-й строки вычитается $(i-1)$-я строка, умноженная на такое число, чтобы начальный элемент строки обратился в нуль; этот нулевой член отбрасывается, получившаяся строка сдвигается на одну позицию влево); построение останавливается на p-й строке, если $r_0^{(p+1)} = 0$.

Критерий Рауса. Многочлен

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x]$$

с действительными коэффициентами, $a_n > 0$, устойчив тогда и только тогда, когда процесс построения таблицы Рауса $R(f)$ не останавливается до $(n+1)$-й строки (регулярный случай), при этом все элементы начального столбца этой таблицы Рауса положительны.

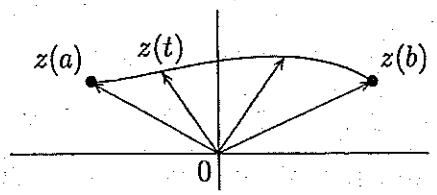
Замечание. В регулярном случае число k корней многочлена $f(x)$, лежащих в правой полуплоскости, равно числу перемен знаков в первом столбце таблицы Рауса $R(f)$.

Замечание 3.28.9. Используя начальные топологические понятия и результаты, связанные с индексами особых точек векторных полей, покажите, что многочлен $f(z)$ степени n устойчив тогда и только тогда, когда точка $f(it)$ при изменении t от $-\infty$ до $+\infty$ обходит начало координат $\frac{n}{2}$ раз (в сторону от 1 к i).

3.29. Комментарии к вопросу о распределении корней многочлена с комплексными коэффициентами на комплексной плоскости

1. Согласованный выбор аргумента точки комплексной плоскости, движущейся по непрерывной линии.

Пусть $z\colon [a,b] \to \mathbb{C}$ — непрерывная функция (т. е. $z(t) = x(t) + y(t)i$ — непрерывная функция от t, где $a \leqslant t \leqslant b$, $a, b \in \mathbb{R}$, принимающая комплексные значения), при этом $z(t) \neq 0$ для всех $t \in [a,b]$ (т. е. линия $z(t)$, $t \in [a,b]$, не проходит через точку 0 в комплексной плоскости $\mathbb{C}$)



и поэтому определена многозначная функция аргумент $\operatorname{Arg} z(t) = \arg z(t) + 2\pi k$.

Из непрерывности функции $z(t)$ следует непрерывность функций $x(t)$, $y(t)$ и $|z(t)| = \sqrt{x(t)^2 + y(t)^2}$ на компакте $[a,b]$, а поэтому и их равномерная непрерывность. Поэтому:

а) функция $|z(t)|$ на отрезке $[a,b]$ достигает своего минимума и, следовательно,

$$r = \inf\{|z(t)| \mid t \in [a,b]\} > 0;$$

б) можно накрыть отрезок $[a,b]$ конечным числом интервалов прямой $\mathbb{R}$, на пересечении каждого из которых с $[a,b]$ колебание функции не превосходит $\frac{r}{2}$, и следовательно, на каждом таком интервале можно считать, что $\arg z(t)$ меняется непрерывно (при фиксации значения в начале интервала);

в) таким образом, выбрав значение аргумента $\arg z(a)$ в начале пути $z(a)$, можно выбрать значение аргумента $\arg z(t)$ при всех t так, что функция $\arg z(t)$ является непрерывной функцией от t.

Замечание 3.29.1. Условие $z(t) \neq 0$ для $t \in [a,b]$ существенно. Действительно, если $z(t) = t$ для $t \in [-1,1]$, то:

$$\text{при } t < 0 \text{ имеем } \operatorname{Arg} z(t) = (2k+1)\pi;$$
$$\text{при } t > 0 \text{ имеем } \operatorname{Arg} z(t) = 2k\pi.$$

Итак, нельзя согласовать выбор $\arg z(t)$ для $t \in [-1,1]$ (доопределив значение при $t = 0$) так, чтобы получить непрерывность при $t = 0$.

2. Согласование выбора аргумента для произведения непрерывных комплексных функций $z(t) = z_1(t)\dots z_k(t)$, $t \in [a,b]$ ($z_i(t) \neq 0$ для всех $i = 1,\dots,k$, $t \in [a,b]$).

Выберем при $t = a$ значения аргументов $\operatorname{Arg} z_1(t),\dots,\operatorname{Arg} z_k(t)$ и $\operatorname{Arg} z(t)$ так, что

$$\arg z(a) = \sum_{j=1}^{k} \arg z_j(a).$$

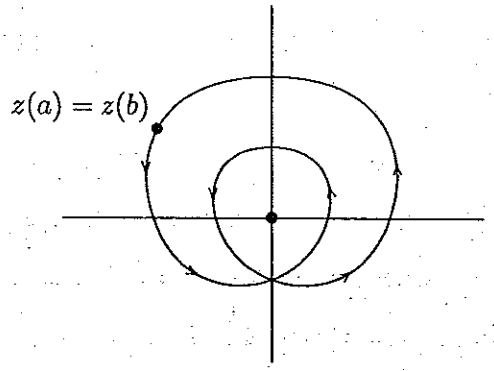
Тогда при непрерывном изменении аргументов (возможном в силу 1) равенство

$$\arg z(t) = \sum_{j=1}^{k} \arg z_j(t)$$

сохранится при всех $t \in [a,b]$ (как непрерывная функция разность $\arg z(t) - \sum\limits_{j=1}^{k} \arg z_k(t)$, принимающая значения $2k\pi$ и равная 0 при $t = a$, равна 0 при всех $t \in [a,b]$).

3. Замкнутая непрерывная линия: $z(t)$ — непрерывная комплекснозначная функция действительного переменного $t \in [a,b] \subset \mathbb{R}$, $z(a) = z(b)$ ($z(t) \neq 0$ для всех $t \in [a,b]$). В этом случае при непрерывном изменении аргумента $\arg z(t)$ разность значений при $t = a$ и $t = b$ может отличаться на $k \cdot 2\pi$, $k \in \mathbb{Z}$. Геометрический смысл целого числа k: число полных оборотов вокруг начала координат $z = 0$ точки $z(t)$ при обходе этой точки в направлении возрастания параметра t от a к b (с учётом знака в соответствии с направлением обхода, т. е. + для обхода против часовой стрелки, − для обхода по часовой стрелке). Например,

для указанного направления обхода $k = 2$ (при противоположном обходе $k = -2$).



4. Принцип аргумента. Под *простым замкнутым контуром* Γ будем понимать непрерывную замкнутую линию в комплексной плоскости $\mathbb{C}$ без самопересечений (т. е. $z(t)$ — непрерывная функция, $t \in [a, b]$, $z(a) = z(b)$, $z(t_1) \neq z(t_2)$ при $a < t_1 < t_2 < b$). *Теорема Жордана* утверждает, что простой замкнутый контур разбивает $\mathbb{C} \setminus \Gamma$ на две связные части: внутренность контура и внешняя часть контура. Пример простого замкнутого контура:

$$\Gamma = \Gamma(z_0, r) = \{z \in \mathbb{C} \mid |z - z_0| = r\} -$$

окружность радиуса r с центром в $z_0 \in \mathbb{C}$.

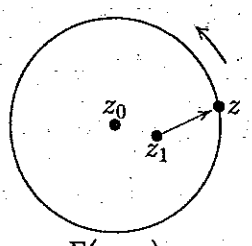
Лемма 3.29.2. *Пусть $z = z(t)$ обходит простой замкнутый контур в положительном направлении, $z_1 \in \mathbb{C} \setminus \Gamma$. Тогда*

$$\Delta_\Gamma \arg(z - z_1) = \begin{cases} 2\pi, & \text{если } z_1 \text{ внутри } \Gamma; \\ 0, & \text{если } z_1 \text{ вне } \Gamma \end{cases}$$

(*здесь* $\Delta_\Gamma \arg(z - z_1)$ — приращение $\arg(z - z_1)$ *при обходе по* Γ *вокруг* z_1 *в* $\mathbb{C}$).

Доказательство для произвольного простого замкнутого контура достаточно сложно и требует техники, выходящей за рамки данного курса. Мы проведём его для частного случая окружности $\Gamma(z_0, r)$.

Случай а): z_1 находится внутри контура $\Gamma(z_0, r)$.

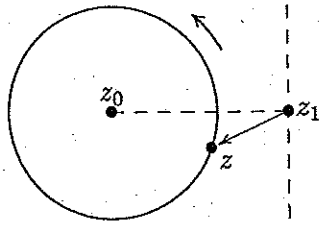


$\Gamma(z_0, r)$

Ясно, что при обходе один раз в положительном направлении вектор $z - z_1$ обернётся вокруг своего начала z_1 тоже один раз в положительном направлении, то есть

$$\Delta_\Gamma \arg(z - z_1) = 2\pi.$$

Случай б): z_1 находится вне контура $\Gamma(z_0, r)$



$\Gamma(z_0, r)$

Так как при обходе точкой z контура $\Gamma(z_0, r)$ в этом случае колебания аргумента $\arg(z - z_1)$ не превосходят π,

$$\Delta_\Gamma \arg(z - z_1) = 0. \qquad \square$$

Теорема 3.29.3 (принцип аргумента). *Пусть многочлен $f(z) \in \mathbb{C}[z]$ не имеет корней на простом замкнутом контуре $\Gamma \subset \mathbb{C}$. Тогда число комплексных корней многочлена $f(z)$ (с учётом их кратностей) внутри контура Γ равно*

$$\frac{\Delta_\Gamma \arg f(z)}{2\pi}$$

(здесь $\Delta_\Gamma \arg f(z)$ — приращение аргумента $\arg f(z)$ при обходе точкой $z = z(t)$ контура Γ один раз в положительном направлении).

Доказательство. Пусть $\deg f(z) = n > 0$,

$$f(z) = a_n(z - z_1)\dots(z - z_n),$$

где $z_1, \dots, z_n$ — корни многочлена (с учётом кратностей). При обходе точкой $z = z(t)$ простого контура сомножители и их произведение $f\big(z(t)\big)$ меняются непрерывно. В силу п. 2 можно считать, что

$$\arg f(z) = \arg a_n + \arg(z - z_1) + \dots + \arg(z - z_n).$$

Следовательно,

$$\Delta_\Gamma \arg f(z) = \sum_{j=1}^{n} \Delta_\Gamma \arg(z - z_j)$$

(при однократном обходе $z = z(t)$ по контуру Γ).

В силу леммы 3.29.2

$$\Delta_\Gamma \arg(z - z_j) = \begin{cases} 2\pi, & \text{если } z_j \text{ внутри } \Gamma, \\ 0, & \text{если } z_j \text{ вне } \Gamma. \end{cases}$$

Поэтому

$$\Delta_\Gamma \arg f(z) = m \cdot 2\pi,$$

где m — число корней z_j, расположенных внутри контура Γ. Следовательно,

$$m = \frac{\Delta_\Gamma \arg f(z)}{2\pi}. \qquad \square$$

5. Теорема Руше. *Пусть $f(z), g(z) \in \mathbb{C}[z]$, Γ — простой замкнутый контур в комплексной плоскости $\mathbb{C}$. Если*

$$|f(z) - g(z)| < |g(z)|$$

для всех $z \in \Gamma$, то внутри контура Γ располагается одинаковое число корней многочленов $f(z)$ и $g(z)$ (с учётом кратностей).

Доказательство. 1) Проверим, что к многочленам $g(z)$ и $f(z)$ можно применить принцип аргумента. Так как

$$|g(z)| > |f(z) - g(z)| \geqslant 0,$$

то

$$|g(z)| > 0$$

для всех $z \in \Gamma$; кроме того,

$$|f(z)| = \big|g(z) + \big(f(z) - g(z)\big)\big| \geqslant |g(z)| - |f(z) - g(z)| > 0$$

для всех $z \in \Gamma$. Таким образом, $g(z)$ и $f(z)$ не обращаются в 0 на контуре Γ.

Так как

$$f(z) = g(z)\left(1 + \frac{f(z) - g(z)}{g(z)}\right),$$

то

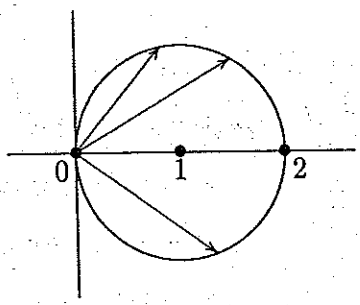
$$\Delta_\Gamma \arg f(z) = \Delta_\Gamma \arg g(z) + \Delta_\Gamma \arg\left(1 + \frac{f(z) - g(z)}{g(z)}\right)$$

(точка $z = z(t)$ пробегает контур Γ).

Поскольку

$$\left|\frac{f(z) - g(z)}{g(z)}\right| < 1,$$

комплексное число $1 + \dfrac{f(z) - g(z)}{g(z)}$ лежит в правой полуплоскости,



поэтому

$$\Delta_\Gamma \arg\left(1 + \frac{f(z) - g(z)}{g(z)}\right) = 0.$$

Итак,

$$\Delta_\Gamma \arg f(z) = \Delta_\Gamma \arg g(z).$$

В силу принципа аргумента число корней многочлена $f(z)$ и число корней многочлена $g(z)$ внутри контура Γ совпадают:

$$\frac{\Delta_\Gamma \arg f(z)}{2\pi} = \frac{\Delta_\Gamma \arg g(z)}{2\pi}. \qquad \square$$

Следствие 3.29.4 (другое доказательство теоремы Гаусса с оценкой для модуля корней). *Пусть*

$$f(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 \in \mathbb{C}[z]$$

и $R = 1 + a$, *где* $a = \max\{|a_i|, 0 \leqslant i \leqslant n-1\}$. *Тогда внутри круга*

$$\{z \in \mathbb{C} \mid |z| \leqslant R\}$$

расположены все n *корней многочлена* $f(z)$ *(с учётом их кратностей).*

Доказательство. Внутри рассматриваемого круга

$$\{z \in \mathbb{C} \mid |z| \leqslant R\}$$

многочлен $g(z) = z^n$ имеет корень 0 кратности n. Проверим условия теоремы Руше в этом случае, то есть что для всех $\{z \in \mathbb{C} \mid |z| = R = 1 + a\}$ имеем

$$\begin{aligned} |f(z) - g(z)| &= |a_{n-1}z^{n-1} + \ldots + a_1 z + a_0| \leqslant \\ &\leqslant a(|z|^{n-1} + |z|^{n-2} + \ldots + 1) = a\frac{|z|^n - 1}{|z| - 1} = \\ &= \frac{a(|z|^n - 1)}{a} = |z|^n - 1 < |z|^n = |g(z)|. \end{aligned}$$

□

Пример 3.29.5. Определим число корней многочлена

$$f(z) = z^8 - 4z^5 + z^2 - 1,$$

модуль которых меньше единицы. Положим $g(z) = -4z^5$ и $\Gamma = \{z \in \mathbb{C} \mid |z| = 1\}$. Тогда $h(z) = f(z) - g(z) = z^8 + z^2 - 1$. Для $z \in \Gamma$ (т. е. $|z| = 1$) имеем:

$$\begin{aligned} &|g(z)| = |-4z^5| = 4; \\ &|h(z)| = |z^8 + z^2 + 1| \leqslant |z^8| + |z^2| + 1 = 3 < 4 = |g(z)|. \end{aligned}$$

Следовательно, по теореме Руше многочлен $f(z)$ имеет внутри окружности $\Gamma = \{z \in \mathbb{C} \mid |z| = 1\}$ столько же корней, сколько и многочлен $g(z) = -4z^5$, то есть пять корней. Итак, многочлен $f(z)$ имеет пять корней, по модулю меньших единицы.

6. Непрерывная зависимость комплексных корней многочлена от его коэффициентов.

Теорема 3.29.6. *Пусть* $g(z) \in \mathbb{C}[z]$, $\deg g(z) = n \geqslant 1$. *Тогда корни многочлена* $g(z)$ *при достаточно малом изменении его коэффициентов меняются сколь угодно мало (при этом кратные корни могут распадаться в совокупность корней, количество которых совпадает с кратностью исходного корня).*

Доказательство. Пусть $c \in \mathbb{C}$ — корень кратности k многочлена $g(z)$.

Выберем окружность $\Gamma = \Gamma(c, r)$ достаточно малого радиуса r такого, что внутри контура Γ нет корней многочлена $g(z)$, отличных от c. Пусть

$$M = \inf\{|g(z)|, z \in \Gamma\}.$$

Так как функция $|g(z)|\colon \Gamma \to \mathbb{R}$ непрерывна на компакте Γ и $|g(z)| \neq 0$ для всех $z \in \Gamma$, то $M > 0$.

На линейном пространстве $\mathbb{C}_n[z]$ многочленов $f(z) \in \mathbb{C}[z]$, $\deg f(z) \leqslant n$,

$$\mathbb{C}_n[z] = \{f(z) = a_n z^n + \ldots + a_1 z + a_0 \mid a_i \in \mathbb{C}\} \cong \mathbb{C}^n,$$

рассматриваем топологию, в которой базис окрестностей для $f(z)$ состоит из совокупностей многочленов вида

$$V(a_n)z^n + \ldots + V(a_1)z + V(a_0),$$

где $V(a_i)$ — окрестность для $a_i \in \mathbb{C}$, $i = 0, 1, \ldots, n$.

Выберем коэффициенты многочлена $h(z) \in \mathbb{C}_n[z]$ настолько малыми, что

$$|h(z)| < M$$

для всех $z \in \Gamma$.

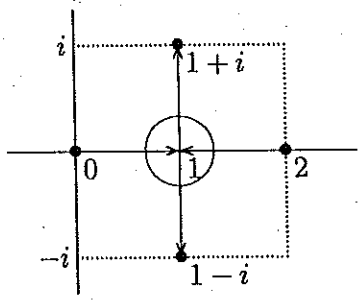
Применим к многочленам $f(z) = g(z) + h(z), g(z) \in \mathbb{C}[z]$ теорему Руше на контуре Γ, поскольку

$$|f(z) - g(z)| = |h(z)| < |g(z)|$$

для всех $z \in \Gamma$. Следовательно, внутри контура $\Gamma = \Gamma(c, r)$ «деформированный» многочлен $f(z) = g(z) + h(z)$ имеет столько же корней (с учётом кратностей), сколько их имел многочлен $g(z)$, то есть k корней. □

Замечание 3.29.7. Простой корень (то есть $k = 1$) при малом изменении коэффициентов многочлена $g(z)$ остаётся простым корнем многочлена $f(z) = g(z) + h(z)$.

Замечание 3.29.8. Если $f(z) = z^2 - 2z + t \in \mathbb{C}[z]$, $t \in \mathbb{R}$, $0 \leqslant t \leqslant 2$, то корни имеют вид $z_{1,2} = 1 \pm \sqrt{1-t}$. При изменении t от 0 до 1 корни $z_1 = 0$ и $z_2 = 2$ (при $t = 0$) сближаются по вещественной оси $\mathbb{R} \subset \mathbb{C}$, превращаясь в корень $z_{1,2} = 1$ кратности 2 (при $t = 1$). При изменении t от 1 до 2 корень $z_{1,2}$ кратности 2 (при $t = 1$) расходится по прямой $\operatorname{Re} z = 1$ до корней $1 + i$ и $1 - i$ (при $t = 2$):



Таким образом, для многочлена

$$g(z) = z^2 - 2z + 1 = (z-1)^2$$

и его корня $c = 1$ кратности 2 в любой его малой окрестности $\{z \in \mathbb{C} \mid |z - 1| < \varepsilon\}$, $0 < \varepsilon \in \mathbb{R}$, малое шевеление $f(z) = g(z) + \delta$, $0 < \delta = \delta(\varepsilon) \in \mathbb{R}$, многочлена гарантирует, что в окрестности $\{z \in \mathbb{C} \mid |z - 1| < \varepsilon\}$ многочлен $f(z)$ будет иметь также два корня (но уже различные).

3.30.

Если A, B и C — положительно определённые матрицы, то корни многочлена $|\lambda^2 A + \lambda B + C|$ имеют отрицательные вещественные части.

Задача 3.30.1.

1) Пусть $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$, $a_{ij} \geqslant 0$ для всех i, j, $i \neq j$, и существуют такие положительные числа $\alpha_1, \ldots, \alpha_n$, что

$$\sum_{j=1}^{n} \alpha_j a_{ij} < 0, \quad i = 1, \ldots, n.$$

Тогда A — устойчивая матрица.

2) Пусть $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{C})$ и

$$\mathrm{Re}(a_{ii}) < -\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad i = 1, \ldots, n.$$

Тогда A — устойчивая матрица.

Задача 3.30.2. Пусть $A \in \mathrm{M}_n(\mathbb{C})$,

$$A = \begin{pmatrix} a_1 + b_1 & a_2 & 0 & 0 & \ldots\ldots\ldots\ldots \\ -1 & b_2 & a_3 & 0 & \ldots\ldots\ldots\ldots \\ 0 & -1 & b_3 & a_4 & \ldots\ldots\ldots\ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots\ldots\ldots\ldots \\ \ldots\ldots & \ldots & \ldots & -1 & b_{n-1} \quad a_n \\ \ldots\ldots & \ldots & \ldots & 0 & -1 \quad b_n \end{pmatrix}$$

(все элементы вне главной диагонали и двух соседних с ней равны нулю), $a_j \in \mathbb{R}$, $b_k = 0$ или $b_k \in i \cdot \mathbb{R}$. Докажите, что число положительных членов в последовательности $a_1, a_1 a_2, \ldots, a_1 a_2 \cdots a_n$ равно числу собственных значений матрицы A, имеющих положительные действительные части.

Замечание 3.30.3 (критерий устойчивости в терминах цепных дробей, теорема Уолла—Франка). Пусть

$$f(z) = z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 \in \mathbb{C}[z], \quad a_k = d_k + i b_k, \quad k = 0, 1, \ldots, n-1,$$
$$g(z) = \alpha_{n-1} z^{n-1} + i\beta_{n-2} z^{n-2} + \alpha_{n-3} z^{n-3} + i\beta_{n-4} z^{n-4} + \ldots.$$

Тогда $f(z)$ является устойчивым многочленом в том и только в том случае, когда

$$\frac{g(z)}{f(z)} = \left[0; \frac{1}{1 + c_1 + d_1 z}, \frac{1}{c_2 + d_2 z}, \ldots, \frac{1}{c_n + d_n z}\right]$$

(обыкновенная конечная цепная дробь), где $\mathrm{Re}(c_j) = 0$ и $d_j > 0$, $j = 1, 2, \ldots, n$.

Следствие 3.30.4. *Пусть*

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \in \mathbb{R}[x],$$
$$g(x) = a_{n-1}x^{n-1} + a_{n-3}x^{n-3} + \ldots .$$

Тогда $f(x)$ является устойчивым многочленом в том и только в том случае, когда

$$\frac{g(x)}{f(x)} = \left[0; \frac{1}{1+b_1x}, \frac{1}{b_2x}, \frac{1}{b_3x}, \ldots, \frac{1}{b_nx}\right],$$

где $b_i > 0$, $i = 1, 2, \ldots, n$.

Отметим, что обыкновенная цепная дробь, описанная в теореме и следствии, может быть получена с помощью алгоритма деления.

Упражнение 3.30.5. Используя следствие 3.30.4 и алгоритм деления, покажите, что многочлен

$$f(x) = x^4 + 5x^3 + 10x^2 + 10x + 4$$

является устойчивым.

Указание. $g(x) = 5x^3 + 10x$,

$$\frac{g(x)}{f(x)} = \frac{1}{\left(\frac{f(x)}{g(x)}\right)} = \frac{1}{1+h_1(x)},$$

где

$$h_1(x) = \frac{1}{5}x + \frac{8x^2+4}{5x^3+10x} = \left(\frac{1}{5}x\right) + \frac{1}{h_2(x)},$$

где

$$h_2(x) = \frac{5}{8}x + \frac{\frac{15}{2}x}{8x^2+4} = \left(\frac{5}{8}x\right) + \frac{1}{h_3(x)}, \quad h_3(x) = \frac{16}{15}x + \frac{1}{\frac{15}{8}x}.$$

$$\frac{g(x)}{f(x)} = \left[0; \frac{1}{1+\frac{1}{5}x}, \frac{1}{\frac{5}{8}x}, \frac{1}{\frac{16}{15}x}, \frac{1}{\frac{15}{8}x}\right].$$

Задача 3.30.6. При каких вещественных значениях параметра c многочлен

$$f(x) = x^4 + 5x^3 + 10x^2 + 10x + c$$

устойчив.

Ответ. $0 < c < 16$.

Замечание 3.30.7. Теорема Гаусса—Люка утверждает, что если выпуклое подмножество комплексной плоскости содержит все корни многочлена $f(z) \in \mathbb{Z}[z]$, то это подмножество содержит также все корни производной $f'(z)$.

Действительно, пусть

$$f(z) = c(z - z_1)\cdots(z - z_n),$$

где $c, z_i \in \mathbb{C}$, $1 \leqslant i \leqslant n$. Тогда

$$\frac{f'(z)}{f(z)} = \frac{1}{z - z_1} + \ldots + \frac{1}{z - z_n}.$$

Если $w \in \mathbb{C}$, $f'(w) = 0$, $f(w)$ и w не принадлежат выпуклой оболочке точек $z_1, \ldots, z_n$ на комплексной плоскости, то через точку w можно провести прямую, не пересекающую выпуклую оболочку точек $z_1, \ldots, z_n$. При этом векторы $w - z_1, \ldots, w - z_n$ лежат в одной полуплоскости, определяемой указанной прямой. Поэтому векторы $\dfrac{1}{w - z_1}, \ldots, \dfrac{1}{w - z_n}$ также лежат в одной полуплоскости, и следовательно,

$$\frac{f'(w)}{f(w)} = \frac{1}{w - z_1} + \ldots + \frac{1}{w - z_n} \neq 0.$$

Полученное противоречие доказывает теорему Гаусса—Люка. □

Задача 3.30.8. Пусть $n \geqslant 3$,

$$q_n(z) = \frac{(1+z)^n - 1 - z^n}{z} \in \mathbb{C}[z].$$

Для каких n все корни многочлена $q_n(z)$ лежат на единичной окружности.

Указание. Все корни многочленов

$$q_3(z) = 3(1+z), \quad q_4(z) = 4(2 + 3z + 2z^2),$$
$$q_5(z) = 5(1+z)(1+z+z^2), \quad q_7(z) = 7(1+z)(1+z+z^2)^2$$

расположены на единичной окружности. Многочлен $q_6(z)$ может быть записан в виде $(a + bz + az^2)(c + dz + cz^2)$, где ad и bc — корни многочлена $t^2 - 15t + 48$. Отсюда вытекает, что все корни многочлена $q_6(z)$ лежат на единичной окружности.

Для $n \geqslant 8$ многочлен $q_n'(z)$ имеет корень, по модулю превосходящий 1. По теореме Гаусса—Люка это справедливо и для $q_n(z)$.

Задача 3.30.9. Пусть

$$f(z) = a_n z^n + \ldots + a_1 z + a_0 \in \mathbb{R}[z], \quad a_0 > a_1 > \ldots > a_n > 0.$$

Докажите, что единичной круг на комплексной плоскости $|z| \leqslant 1$ не содержит ни одного корня многочлена $f(z)$.

Указание. При $z = 1$

$$f(z) = a_n + \ldots + a_0 > 0;$$

при $z \in \mathbb{R}$, $0 \leqslant z < 1$ имеем

$$|(a_0 - a_1)z + (a_1 - a_2)z + \ldots + (a_{n-1} - a_n)z^n + a_n z^{n-1}| < a_0 - a_1 + a_1 - a_2 + \ldots + a_{n-1} - a_n + a_n = a_0,$$

и поэтому

$$|(1-z)f(z)| \geqslant a_0 - |(a_0 - a_1)z + \ldots + a_n z^{n+1}| > 0;$$

при $|z| \geqslant 1$, $z \notin \{x \in \mathbb{R} \mod 0 \leqslant x \leqslant 1\}$ получаем

$$|(1-z)f(z)| \geqslant a_0 - |(a_0 - a_1)z + (a_1 - a_2)z^2 + \ldots + (a_{n-1} - a_n)z^n + a_n z^n| > a_0 - (a_0 - a_1 + \ldots + a_n) = 0$$

(числа $(a_0 - a_1)z, (a_1 - a_2)z^2, \ldots, a_n z^{n+1}$ не могут одновременно иметь один аргумент: если φ — аргумент числа z и $\varphi = 2\varphi$ (аргумент числа z^2), то $\varphi = 2\pi k$, $k \in \mathbb{Z}$, и в нашем случае $z \in \mathbb{R}$, $0 \leqslant z \leqslant 1$).

Задача 3.30.10 (критерий Шура—Кона).

1) Оба корня $w = t_1, t_2$ действительного многочлена $t^2 + bt + c$ удовлетворяют условию $|w| < 1$ тогда и только тогда, когда $|b| < 1 + c < 2$.

2) Все три корня w действительного многочлена $t^3 + bt^2 + ct + d$ удовлетворяют условию $|w| < 1$ тогда и только тогда, когда $|bd - c| < 1 - d^2$, $|b + d| < |1 + c|$.

Упражнение 3.30.11 (обобщение теоремы Ролля). Пусть $f \in \mathbb{C}[x]$, $\deg f = n \geqslant 2$. Если $z_1, z_2 \in \mathbb{C}$, $z_1 \neq z_2$ и $f(z_1) = f(z_2)$, то диск

$$\left\{ z \in \mathbb{C} \;\middle|\; \left| z - \frac{z_1 + z_2}{2} \right| \leqslant \left| \frac{z_1 - z_2}{2} \right| \cdot \operatorname{ctg} \frac{\pi}{n} \right\}$$

содержит по крайней мере один корень многочлена $f'(x)$.

Задача 3.30.12 (неравенство Ландау). Пусть

$$f(z) = a_n z^n + \ldots + a_1 z + a_0 \in \mathbb{C}[z], \quad a_n \neq 0, \quad f(z) = a_n \prod_{i=1}^{n} (z - z_i),$$

$$\|f\|_2 = \sqrt{\sum_{i=0}^{n} |a_i|^2}, \quad M(f) = |a_n| \cdot \prod_{i=1}^{n} \max\{1, |z_i|\}.$$

Докажите, что $M(f) \leqslant \|f\|_2$.

Указание. Пусть $z_1, \ldots, z_k$ — корни многочлена $f(z)$, находящиеся вне единичного круга. Тогда

$$M(f) = |a_n| \cdot |z_1| \cdots |z_k|.$$

Положим

$$h(x) = a_n \cdot \prod_{j=1}^{k} (\bar{z}_j x - 1) \prod_{j=k+1}^{n} (x - z_j) = b_n x^n + \ldots + b_0 \in \mathbb{C}[x].$$

Если $g \in \mathbb{C}[x]$ и $z \in \mathbb{C}$, то

$$\|(x + z)g(x)\|_2 = \|(\bar{z}x + 1)g(x)\|_2.$$

Применяя k раз это соображение, получаем $\|f\|_2 = \|h\|_2$. Но $\|h\|_2 \geqslant |b_n| = M(f)$.

Задача 3.30.13. Пусть $f(z) \in \mathbb{C}[z]$, $\deg f = n$, f имеет не менее двух различных корней. Тогда многочлен $F = f \cdot f' \cdots f^{(n-1)}$ степени $\dfrac{n(n+1)}{2}$ имеет не менее $n+1$ различных корней.

Задача 3.30.14 (Коши, Кнут). Пусть $n \geqslant 1$,

$$f(z) = a_n z^n + \ldots + a_0 \in \mathbb{C}[z], \quad a_n \neq 0, \quad f(x) = 0 \quad (x \in \mathbb{C}).$$

Тогда:

$$|x| \leqslant \max\left\{ \left| \frac{n \cdot a_{n-1}}{a_n} \right|, \left| \frac{n \cdot a_{n-2}}{a_n} \right|^{1/2}, \ldots, \left| \frac{n \cdot a_0}{a_n} \right|^{1/n} \right\};$$

$$|x| \leqslant 2 \max\left\{ \left| \frac{a_{n-1}}{a_n} \right|, \left| \frac{a_{n-2}}{a_n} \right|^{1/2}, \ldots, \left| \frac{a_0}{a_n} \right|^{1/n} \right\}.$$

Если $a_0 \neq 0$, то с помощью этих оценок легко получить границы минимального модуля корня многочлена $f(z)$ (заменяя в исходном многочлене z на $\frac{1}{z}$ и находя границу R максимального модуля корня многочлена $a_0z^n + \ldots + a_n$; $\frac{1}{R}$ будет искомой границей минимального модуля корня).

Задача 3.30.15 (Энестрём, Какея). Все корни многочлена

$$f(z) = \sum_{j=0}^{n} a_j z^j \in \mathbb{R}[z]$$

с положительными коэффициентами принадлежат множеству

$$\left\{ z \in \mathbb{C} \;\middle|\; \min_{1 \leqslant i \leqslant n} \left(\frac{a_{i-1}}{a_i} \right) \leqslant |z| \leqslant \max_{1 \leqslant i \leqslant n} \left(\frac{a_{i-1}}{a_i} \right) \right\}.$$

Замечание 3.30.16. Пусть $f \in \mathbb{C}[t]$. Через $n_0(f)$ обозначим количество различных корней многочлена f.

Теорема Мейсона—Стотерса. *Пусть $f, g, h \in \mathbb{C}[t]$ — не равные константе взаимно простые многочлены, $f + g = h$. Тогда*

$$\max\{\deg(f), \deg(g), \deg(h)\} \leqslant n_0(fgh) - 1.$$

С использованием теоремы Мейсона—Стотерса покажите, что если $n \in \mathbb{N}$, $n \geqslant 3$, то не существует решений уравнения

$$\big(x(t)\big)^n + \big(y(t)\big)^n = \big(z(t)\big)^n$$

с не равными константе взаимно простыми многочленами $x(t), y(t), z(t) \in \mathbb{C}[t]$.

Гипотеза Смейла (1981 г.). Пусть $f(z) \in \mathbb{C}[z]$ — такой многочлен степени n, что $f(0) = 0$, $f'(0) \neq 0$. Тогда

$$\min \left\{ \left| \frac{f(z)}{zf'(0)} \right| \;\middle|\; f'(z) = 0 \right\} \leqslant M,$$

где $M = 1$ или, возможно, $M = \frac{n-1}{n}$.

Смейл доказал гипотезу для $n = 4$. Рассматривая $f(z) = a_1 z + a_n z^n$, легко убедиться, что M не может быть меньше, чем $\frac{n-1}{n}$.

Задача 3.30.17. Пусть

$$f(z) = z^n + b_{n-1}z^{n-1} + \ldots + b_1 z + b_0 \in \mathbb{C}[z].$$

Покажите, что

$$\max_{|z|=1}\{|f(z)|\} \geqslant 1$$

(равенство достигается лишь для $f(z) = z^n$).

Задача 3.30.18.

1) *Неравенство Маркова.* Пусть $\deg f \leqslant n$, $|f(x)| \leqslant M$ при $-1 \leqslant x \leqslant 1$. Тогда $|f'(x)| \leqslant M \cdot n^2$ при $-1 \leqslant x \leqslant 1$.

2) *Неравенство Бернштейна.* Пусть $f(z) \in \mathbb{C}[z]$, $\deg f \leqslant n$ и $|f(z)| \leqslant M$ при $|z| = 1$. Тогда $|f'(z)| \leqslant M \cdot n$ при $|z| = 1$.

Гипотеза Сендова. Пусть

$$f(z) = \prod_{i=1}^{n}(z - z_i), \quad n \geqslant 2,$$

все корни лежат в замкнутом единичном диске. Тогда каждый из замкнутых дисков $\bar{D}(z_1, 1), \ldots, \bar{D}(z_n, 1)$ содержит корень производной $f'(z)$.

Глава 4

Многочлены от многих переменных

4.1. Многочлены от многих переменных

Пусть K — поле (например, $K = \mathbb{R}$ или $K = \mathbb{C}$). Под *многочленом* $f(x_1, x_2, \ldots, x_n)$ *от* n *переменных* $x_1, x_2, \ldots, x_n$, $n \geqslant 1$, понимаем сумму конечного числа одночленов вида

$$a_{k_1,\ldots,k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n},$$

где $a_{k_1,\ldots,k_n} \in K$, $\mathbb{N}_0 \ni k_i$, $i = 1, \ldots, n$, по различным строчкам $(k_1, k_2, \ldots, k_n) \in \mathbb{N}_0^n$. При этом считаем, что

$$a \equiv a x_1^0 x_2^0 \cdots x_n^0 \quad \text{для} \quad a \in K.$$

$K \to K[x_1, \ldots, x_n]$

Два многочлена *равны*, если равны их коэффициенты при одинаковых *мономах* $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$. Таким образом, формально мы определили многочлен от n переменных как отображение $\mathbb{N}_0^n \to K$, $(k_1, k_2, \ldots, k_n) \mapsto a_{k_1,k_2,\ldots,k_n}$, для которого *почти все* (то есть все, кроме конечного числа) элементы в его образе равны нулю.

Каждый многочлен

$$f(x_1, x_2, \ldots, x_n) = \sum_{(k_1,k_2,\ldots,k_n) \in \mathbb{N}_0^n} a_{k_1,k_2,\ldots,k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

от n переменных $x_1, x_2, \ldots, x_n$ определяет функцию

$$f \colon K^n \to K,$$

$$(c_1, c_2, \ldots, c_n) \in K^n \mapsto f(c_1, c_2, \ldots, c_n) = \sum_{(k_1,k_2,\ldots,k_n) \in \mathbb{N}_0^n} a_{k_1,k_2,\ldots,k_n} c_1^{k_1} c_2^{k_2} \cdots c_n^{k_n}$$

(иногда называемую *полиномиальной функцией* от переменных $x_1, x_2, \ldots, x_n$; связь многочленов от n переменных и соответствующих им полиномиальных функций будет рассмотрена несколько позже).

Через $K[x_1, x_2, \ldots, x_n]$ обозначим множество всех многочленов от n переменных $x_1, x_2, \ldots, x_n$ с коэффициентами из поля K. На множестве $K[x_1, x_2, \ldots, x_n]$ рассмотрим две операции:

1) суммой двух многочленов $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$ называется многочлен, коэффициенты которого при мономе $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ являются суммами соответствующих коэффициентов при мономе $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ в многочленах $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$;

2) произведение одночленов определяется равенством

$$(ax_1^{k_1}\cdots x_n^{k_n})(bx_1^{l_1}\cdots x_n^{l_n}) = abx_1^{k_1+l_1}\cdots x_n^{k_n+l_n},$$

а произведение многочленов $f(x_1,\ldots,x_n)$ и $g(x_1,\ldots,x_n)$ определяется как результат умножения одночленов с последующим приведением подобных членов.

Замечания 4.1.1.

1) $K[x_1,x_2,\ldots,x_n] = R[x_n]$, где $R = K[x_1,x_2,\ldots,x_{n-1}]$, то есть каждый многочлен от переменных $x_1,x_2,\ldots,x_n$ с коэффициентами из поля K можно рассматривать как многочлен от одной переменной x_n, коэффициенты которого — многочлены из $R = K[x_1,x_2,\ldots,x_{n-1}]$.

Пример 4.1.2. $x_1^2x_2^2+x_1^2x_2+x_1x_2^2+x_2^2+x_1+x_2+3 = (x_1^2+x_1+1)x_2^2+(x_1^2+1)x_2+(x_1+3)$.

2) Ясно, что $K \subseteq K[x_1,x_2,\ldots,x_n]$, $a \equiv ax_1^0x_2^0\cdots x_n^0$ для $a \in K$; таким образом, $0 \in K$ является нейтральным элементом относительно сложения, а $1 \in K$ — нейтральным элементом относительно умножения в $K[x_1,x_2,\ldots,x_n]$.

Теорема 4.1.3. *Многочлены $K[x_1,x_2,\ldots,x_n]$ от n переменных $x_1,x_2,\ldots,x_n$ с коэффициентами из поля K с операциями сложения и умножения образуют коммутативное кольцо с единицей.*

Доказательство. Мы убедились ранее, что если R — коммутативное кольцо с единицей, то $R[x]$ — коммутативное кольцо с единицей. Принимая во внимание, что $K[x_1,x_2,\ldots,x_n] = (K[x_1,\ldots,x_{n-1}])[x_n]$, и проводя индукцию по n (начало индукции $n=1$), получаем, что $K[x_1,x_2,\ldots,x_n]$ — коммутативное кольцо с единицей. □

Замечания 4.1.4.

1) Так как многочлены $K[x_1,x_2,\ldots,x_n]$ являются линейным пространством над полем K и

$$k(fg) = (kf)g = f(kg)$$

для всех $k \in K$, $f,g \in K[x_1,x_2,\ldots,x_n]$, то $K[x_1,x_2,\ldots,x_n]$ не только кольцо, но и алгебра над полем K (кратко: K-алгебра).

2) Если поле K лежит в поле или в коммутативном кольце R, $K \subseteq R$, $\alpha_1,\alpha_2,\ldots,\alpha_n \in R$, то для

$$f(x_1,x_2,\ldots,x_n) = \sum a_{l_1,l_2,\ldots,l_n}x_1^{l_1}x_2^{l_2}\cdots x_n^{l_n} \in K[x_1,x_2,\ldots,x_n]$$

определено значение

$$f(\alpha_1,\alpha_2,\ldots,\alpha_n) = \sum a_{l_1,l_2,\ldots,l_n}\alpha_1^{l_1}\alpha_2^{l_2}\cdots\alpha_n^{l_n} \in R,$$

при этом отображение

$$K[x_1,x_2,\ldots,x_n] \to R,$$
$$f(x_1,x_2,\ldots,x_n) \mapsto f(\alpha_1,\alpha_2,\ldots,\alpha_n)$$

является гомоморфизмом алгебр. Образ этого гомоморфизма

$$K[\alpha_1, \alpha_2, \ldots, \alpha_n] = \{f(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]\}$$

является наименьшей K-подалгеброй, содержащей элементы $\alpha_1, \alpha_2, \ldots, \alpha_n \in R$. Ядро этого гомоморфизма — идеал кольца многочленов $K[x_1, x_2, \ldots, x_n]$,

$$\{f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n] \mid f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0\} \lhd K[x_1, x_2, \ldots, x_n].$$

Если система элементов $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset R$ *алгебраически независима* (это означает, что $f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0$, $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n] \implies f(x_1, x_2, \ldots, x_n) = 0$), то ядро нашего гомоморфизма нулевое, и поэтому имеем изоморфизм

$$K[\alpha_1, \alpha_2, \ldots, \alpha_n] \cong K[x_1, x_2, \ldots, x_n].$$

При $n = 1$ алгебраически независимая система $\alpha \in R$ называется *трансцендентным элементом*, $K[\alpha] \cong K[x]$. В противном случае, если $\alpha \in R$ и существует ненулевой многочлен $f(x) \in K[x]$ такой, что $f(x) = 0$, то элемент $\alpha \in R$ называется *алгебраическим* элементом над полем K.

Лемма 4.1.5. *Если R — кольцо без делителей нуля, то кольцо многочленов $R[x]$ также не имеет делителей нуля.*

Доказательство. Если

$$f(x) = a_m x^m + \ldots + a_1 x + a_0 \neq 0, \quad a_m \neq 0,$$
$$g(x) = b_n x^n + \ldots + b_1 x + b_0 \neq 0, \quad b_n \neq 0,$$

то

$$f(x)g(x) = a_m b_n x^{m+n} + \ldots + \sum_{0 \leqslant i \leqslant k < m+n} a_i b_{k-i} x^k + \ldots + a_0 b_0,$$

при этом $a_m b_n \neq 0$, поскольку R — кольцо без делителей нуля.

Итак, $f(x)g(x) \neq 0$ (фактически доказано, что $\deg(f(x)g(x)) = m + n = \deg f(x) + \deg g(x)$). □

Следствие 4.1.6. *Кольцо многочленов $K[x_1, x_2, \ldots, x_n]$ от переменных $x_1, x_2, \ldots, x_n$ с коэффициентами из поля K не имеет делителей нуля.*

Доказательство. Так как

$$K[x_1, x_2, \ldots, x_n] = (K[x_1, x_2, \ldots, x_{n-1}])[x_n],$$

то наше утверждение получаем, проводя индукцию по n (начало индукции $n = 1$). □

4.2. Степень $\deg f(x_1, x_2, \ldots, x_n)$ многочлена от переменных $x_1, x_2, \ldots, x_n$ по совокупности переменных

Степенью монома $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$, $k_i \geqslant 0$, назовём сумму

$$k_1 + k_2 + \ldots + k_n;$$

под *степенью* $\deg f(x_1, x_2, \ldots, x_n)$ ненулевого многочлена $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ будем понимать наивысшую степень мономов его ненулевых одночленов (которых может быть несколько: $f(x_1 x_2) = \underline{x_1^2 x_2} + \underline{x_1 x_2^2} + x_1 + x_2$, $\deg f(x_1, x_2) = 3$).

4.3. Степень $\deg_{x_i} f(x_1, x_2, \ldots, x_n)$ многочлена $f(x_1, x_2, \ldots, x_n)$ из $K[x_1, x_2, \ldots, x_n]$ по переменной x_i

Так как

$$K[x_1, x_2, \ldots, x_n] = (K[x_1, \ldots, \widehat{x_i}, \ldots, x_n])[x_i],$$

то для $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ определим его степень $\deg_{x_i} f(x_1, x_2, \ldots, x_n)$ по переменной x_i как степень многочлена $f(x_1, x_2, \ldots, x_n)$ от x_i, рассматриваемого как многочлен с коэффициентами из $K[x_1, \ldots, \widehat{x_i}, \ldots, x_n]$.

Пример 4.3.1. $f(x_1, x_2) = x_1^3 x_2 + x_1^3 + x_1 x_2^2 + x_2^2 + x_1 + x_2 + 1 = (x_2 + 1)x_1^3 + (x_2^2 + 1)x_1 + (x_2^2 + x_2 + 1) = (x_1 + 1)x_2^2 + (x_1^3 + 1)x_2 + (x_1^3 + x_1 + 1)$. Поэтому: $\deg f(x_1, x_2) = 4$; $\deg_{x_1} f(x_1, x_2) = 3$; $\deg_{x_2} f(x_1, x_2) = 2$.

4.4. Однородные многочлены

Если все члены многочлена $f(x_1, x_2, \ldots, x_n)$ имеют по совокупности переменных одну и ту же степень s, то многочлен $f(x_1, x_2, \ldots, x_n)$ называется *однородным многочленом* (или формой) *s-й степени*. Это равносильно тому, что коэффициенты $a_{k_1, k_2, \ldots, k_n}$ многочлена $f(x_1, x_2, \ldots, x_n)$ равны нулю, если $k_1 + k_2 + \ldots + k_n \neq s$.

Примеры 4.4.1.

1) *линейные формы* $\sum\limits_{i=1}^{n} a_i x_i$ при $s = 1$;

2) *квадратичные формы* $\sum\limits_{i=1}^{n} a_i x_i^2 + \sum\limits_{1 \leqslant i < j \leqslant n} a_{ij} x_i x_j$ при $s = 2$.

Замечание 4.4.2. Однородные многочлены степени s от n переменных $x_1, x_2, \ldots, x_n$ образуют конечномерное подпространство размерности

$$\frac{n(n+1)\cdots(n+s-1)}{s!}$$

(число сочетаний с повторениями из n по s) в бесконечномерном линейном пространстве ${}_K V = K[x_1, x_2, \ldots, x_n]$.

Любой многочлен $f = f(x_1, x_2, \ldots, x_n)$ *однозначно представляется* в виде суммы

$$f = f_0 + f_1 + \ldots + f_d,$$

где f_s — однородный многочлен степени s (называемый *s-й однородной компонентой*), $0 \leqslant s \leqslant d$, $d = \deg f(x_1, x_2, \ldots, x_n)$. Ясно, что s-я однородная компонента f_s многочлена f является суммой всех одночленов степени s многочлена $f = f(x_1, x_2, \ldots, x_n)$.

Лемма 4.4.3. *Если*

$$f(x_1, x_2, \ldots, x_n), g(x_1, x_2, \ldots, x_n) \in K[x_1, \ldots, x_n] -$$

ненулевые многочлены от n переменных $x_1, x_2, \ldots, x_n$, то:

1) $\deg(f + g) \leqslant \max(\deg f, \deg g)$ (*если* $f + g \neq 0$);

2) $\deg fg = \deg f + \deg g$.

Доказательство. 1) Если $s > \max(\deg f, \deg g)$, $h = f + g$, то $h_s = f_s + g_s = 0 + 0 = 0$. Итак, $\deg(f + g) \leqslant \max(\deg f, \deg g)$.

2) Если

$$f = f_0 + f_1 + \ldots + f_m, \quad f_m \neq 0, \quad m = \deg f,$$
$$g = g_0 + g_1 + \ldots + g_n, \quad g_n \neq 0, \quad n = \deg g,$$

то

$$h = fg = f_0 g_0 + \ldots + f_m g_n, \quad f_m g_n \neq 0, \quad \deg(f_m g_n) = m + n$$

(поскольку $K[x_1, x_2, \ldots, x_n]$ — кольцо без делителей нуля), то есть

$$h = fg = h_0 + h_1 + \ldots + h_{m+n}, \quad h_{m+n} = f_m g_n \neq 0.$$

Итак, $\deg fg = m + n = \deg f + \deg g$. □

Лемма 4.4.4 (критерий однородности многочлена). *Пусть $x_1, \ldots, x_n$ — переменные над коммутативным кольцом R. Многочлен $f(x_1, \ldots, x_n) \in R[x_1, \ldots, x_n]$ является однородным степени d тогда и только тогда, когда*

$$f(yx_1, \ldots, yx_n) = y^d f(x_1, \ldots, x_n)$$

в кольце многочленов $R[x_1, \ldots, x_n, y]$.

Доказательство. 1) Пусть $f(x_1, \ldots, x_n)$ — однородный многочлен степени d. Тогда $f = \sum_{i=1}^{t} r_i F_i$, где $0 \neq r_i \in R$, F_i — моном от переменных $x_1, \ldots, x_n$ степени d.

Если $F = x_1^{\alpha(1)} \cdots x_n^{\alpha(n)}$, $\alpha(1) + \ldots + \alpha(n) = d$, то

$$F(yx_1, \ldots, yx_n) = (yx_1)^{\alpha(1)} \cdots (yx_n)^{\alpha(n)} = y^d x_1^{\alpha(1)} \cdots x_n^{\alpha(n)} = y^d F(x_1, \ldots, x_n).$$

Поэтому

$$f(yx_1, \ldots, yx_n) = y^d f(x_1, \ldots, x_n).$$

2) Пусть в кольце $R[x_1, x_2, \ldots, x_n, y]$:

$$f(yx_1, \ldots, yx_n) = y^d f(x_1, \ldots, x_n);$$

пусть

$$f(x_1, \ldots, x_n) = \sum r_{\alpha(1)\alpha(2)\cdots\alpha(n)} x_1^{\alpha(1)} x_2^{\alpha(2)} \cdots x_n^{\alpha(n)},$$

$f = f_0 + f_1 + \ldots + f_m$, где f_i — либо 0, либо однородный многочлен степени i, $f_m \neq 0$ и $m = d(f)$. Тогда в кольце $R[x_1, x_2, \ldots, x_n, y] = (R[x_1, x_2, \ldots, x_n])[y]$:

$$y^d f(x_1, \ldots, x_n) = f(yx_1, \ldots, yx_n) = \sum_{i=0}^{m} f_i(yx_1, \ldots, yx_n) = \sum_{i=0}^{m} y^i f_i(x_1, \ldots, x_n).$$

Итак, приравнивая коэффициенты при степенях y^i, получаем: $m = d$, $f_1 = \ldots = f_{d-1} = 0$, $f_d = f$. Следовательно, $f = f_d$ — однородный многочлен степени d. □

4.5. Формальное и функциональное равенство многочленов из $K[x_1, x_2, \ldots, x_n]$ над бесконечным полем K

Как мы уже видели ранее, в случае многочленов от одной переменной из функционального равенства двух многочленов $f(x), g(x) \in K[x]$ (то есть $f(c) = g(c)$ для всех $c \in K$) может не следовать равенство многочленов $f(x) = g(x)$ в $K[x]$; если поле K бесконечно, то формальное и функциональное равенство двух многочленов оказываются равносильны.

Аналогичное утверждение имеет место и для многочленов от n переменных над бесконечным полем K.

Теорема 4.5.1. *Пусть K — бесконечное поле,*

$$f(x_1, x_2, \ldots, x_n), g(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n].$$

Тогда

$$f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n) \quad \text{в} \quad K[x_1, x_2, \ldots, x_n]$$

в том и только в том случае, если многочлены $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$ определяют одну и ту же функцию $\hat{f} = \hat{g}$ из K^n в K (то есть в этом случае формальное равенство и функциональное равенство двух многочленов равносильны).

Доказательство. Достаточно доказать, что если многочлен $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ определяет нулевую функцию $\hat{f}\colon K^n \to K$, $\hat{f}(c_1, c_2, \ldots, c_n) = f(c_1, c_2, \ldots, c_n) = 0$ для всех $(c_1, c_2, \ldots, c_n) \in K^n$, то $f(x_1, x_2, \ldots, x_n) = 0$ в $K[x_1, x_2, \ldots, x_n]$.

Доказательство проведём индукцией по n. Начало индукции: $n = 1$ (см. ??).

Допустим, что наше утверждение верно для $n - 1$. Пусть

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{d} g_i(x_1, \ldots, x_{n-1}) x_n^i \in K[x_1, x_2, \ldots, x_n] = (K[x_1, x_2, \ldots, x_{n-1}])[x_n],$$

где $d = \deg_{x_n} f(x_1, x_2, \ldots, x_n)$, $g_i(x_1, \ldots, x_{n-1}) \in K[x_1, \ldots, x_{n-1}]$.

Зафиксируем произвольную строчку $(c_1, c_2, \ldots, c_{n-1}) \in K^{n-1}$. Тогда $f(c_1, c_2, \ldots, c_{n-1}, x_n) \in K[x_n]$, и этот многочлен обращается в нуль в каждой точке $c_n \in K$, поскольку $f(c_1, c_2, \ldots, c_{n-1}, c_n) = 0$ для всех $(c_1, c_2, \ldots, c_n) \in K^n$. Следовательно,

$$f(c_1, c_2, \ldots, c_{n-1}, x_n) = 0 \text{ в } K[x_n],$$

поэтому все его коэффициенты при степенях x_n равны нулю, то есть $g_i(c_1, c_2, \ldots, c_{n-1}) = 0$ для всех $0 \leqslant i \leqslant d$. Так как точка $(c_1, c_2, \ldots, c_{n-1}) \in K^{n-1}$ была произвольной, то функция

$$K^{n-1} \to K, \quad (c_1, c_2, \ldots, c_{n-1}) \mapsto g_i(c_1, c_2, \ldots, c_{n-1}),$$

является нулевой для каждого $0 \leqslant i \leqslant d$.

В силу индуктивного предположения

$$g_i(x_1, x_2, \ldots, x_{n-1}) = 0 \text{ в } K[x_1, x_2, \ldots, x_{n-1}].$$

Следовательно,

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{d} g_i(x_1, x_2, \ldots, x_n) x_n^i = 0 \text{ в } K[x_1, x_2, \ldots, x_n]. \qquad \square$$

Следствие 4.5.2 (принцип несущественности алгебраических неравенств). *Пусть K — бесконечное поле, $f, g, h \in K[x_1, \ldots, x_n]$, $h \neq 0$, $f(\alpha_1, \ldots, \alpha_n) = g(\alpha_1, \ldots, \alpha_n)$ для всех $\alpha = (\alpha_1, \ldots, \alpha_n) \in K^n$, для которых $h(\alpha_1, \ldots, \alpha_n) \neq 0$. Тогда $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_n)$.*

Доказательство. Из условия $((f-g)h)(\alpha_1, \ldots, \alpha_n) = 0$ для всех $\alpha = (\alpha_1, \ldots, \alpha_n) \in K^n$. Так как поле K бесконечно, по нашей теореме $(f-g)h = 0$. Поскольку $h \neq 0$ и в кольце $K[x_1, \ldots, x_n]$ нет делителей нуля, $f = g$. $\square$

Замечания 4.5.3.

1) Если поле K конечно, $|K| = q < \infty$, то конечно и линейное пространство K^n, $|K^n| = q^n < \infty$, а поэтому конечно и множество K^{K^n} всех функций $K^n \to K$, $|K^{K^n}| = (q^n)^q = q^{nq} < \infty$. Так как множество многочленов $K[x_1, x_2, \ldots, x_n]$ всегда бесконечно (даже при $n = 1$ оно содержит бесконечное подмножество $\{1, x, x^2, \ldots, x^k, \ldots\} = \{x^k \mid k \in \mathbb{N}_0\}$), то утверждение об эквивалентности формального и функционального равенства не имеет места ни для какого конечного поля K.

2) Если $|K| = q$, то полиномиальные функции, задаваемые одночленами $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$, где $k_1 < q, \ldots, k_n < q$, образуют базис в линейном пространстве K^{K^n} всех функций из K^n в K.

3) Если $|K| = q$, $f = f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ — многочлен от n переменных над K такой, что $\deg_{x_i} f < n$ по каждой переменной x_i, $1 \leqslant i \leqslant n$ (такой многочлен f называется *редуцированным*) и $f = f(x_1, x_2, \ldots, x_n)$ индуцирует нулевую функцию $\hat{f}\colon K^n \to K$, $\hat{f}(c_1, c_2, \ldots, c_n) = f(c_1, c_2, \ldots, c_n) = 0$ для всех $(c_1, c_2, \ldots, c_n) \in K^n$, то $f = f(x_1, x_2, \ldots, x_n) = 0$ в $K[x_1, x_2, \ldots, x_n]$.

   Таким образом, для каждого многочлена $f \in K[x_1, x_2, \ldots, x_n]$ существует и единствен редуцированный многочлен f^*, задающий ту же самую полиномиальную функцию $\hat{f}\colon K^n \to K$, то есть $\hat{f} = \widehat{f^*}$.

4) Если точка $(c_1, c_2, \ldots, c_n) \in K^n$ зафиксирована, то отображение

$$K[x_1, x_2, \ldots, x_n] \to K,$$
$$f(x_1, x_2, \ldots, x_n) \mapsto f(c_1, c_2, \ldots, c_n),$$

является гомоморфизмом колец (K-алгебр).

5) Каноническое отображение, сопоставляющее каждому многочлену $f = f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ полиномиальную функцию

$$\hat{f}\colon K^n \to K,$$
$$(c_1, c_2, \ldots, c_n) \mapsto \hat{f}(c_1, c_2, \ldots, c_n) = f(c_1, c_2, \ldots, c_n),$$

является гомоморфизмом колец (K-алгебр)

$$K[x_1, x_2, \ldots, x_n] \to K^{K^n}, \quad f \mapsto \hat{f}.$$

Полиномиальные функции $\hat{f}$, где $f \in K[x_1, \ldots, x_n]$ (образы этого канонического гомоморфизма), образуют подкольцо (K-подалгебру) в кольце (K-алгебре) всех функций K^{K^n} от n переменных. В тех случаях, когда имеет место равносильность формального и функционального равенства многочленов, многочлены $f = f(x_1, x_2, \ldots, x_n)$ можно отождествлять с определяемыми ими полиномиальными функциями $\hat{f}\colon K^n \to K$, что позволяет в этом случае рассматривать K-алгебру многочленов $K[x_1, x_2, \ldots, x_n]$ как подалгебру всех полиномиальных функций от n переменных со значениями в поле K.

Так как любое коммутативное кольцо R является коммутативной $\mathbb{Z}$-алгеброй, то для любого набора элементов $r_1, \ldots, r_n \in R$ имеем гомоморфизм $\mathbb{Z}$-алгебр

$$\mathbb{Z}[x_1, \ldots, x_n] \to R,$$

образ которого совпадает с подкольцом $\mathbb{Z}[r_1, \ldots, r_n]$ в кольце R, порождённым элементами $r_1, \ldots, r_n$.

Лемма 4.5.4 (о специализации). *Пусть $f(x_1, \ldots, x_n), g(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ и $f(j_1, \ldots, j_n) = g(j_1, \ldots, j_n)$ для всех $j_1, \ldots, j_n \in \mathbb{N}$. Тогда для любого коммутативного кольца R и любых $r_1, \ldots, r_n \in R$:*

$$f(r_1, \ldots, r_n) = g(r_1, \ldots, r_n).$$

Доказательство. Положим

$$F(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) - g(x_1, \ldots, x_n).$$

Тогда $F(j_1, \ldots, j_n) = 0$ для всех $j_1, \ldots, j_n \in \mathbb{N}$. Покажем, что $F(x_1, \ldots, x_n) = 0$ (тогда для любого коммутативного кольца R и любого набора элементов $r_1, \ldots, r_n \in R$

$$\psi(r_1, \ldots, r_n)(F) = \psi(r_1, \ldots, r_n)(0) = 0,$$

в частности, $f(r_1, \ldots, r_n) = g(r_1, \ldots, r_n)$).

Проведём индукцию по n.

Пусть $n = 1$. Тогда многочлен $F(x_1) \in \mathbb{Z}[x_1]$ имеет бесконечное число корней в $\mathbb{Z}$, в силу ??, $F(x_1) = 0$.

Пусть $n > 1$. Если $d(F) \leqslant 0$, то $F = 0$, поскольку F имеет бесконечное число нулей. Пусть $d(F) \geqslant 1$, тогда F содержит хотя бы один ненулевой одночлен, при этом можно считать, что x_n входит в этот одночлен в положительной степени (в противном случае можно перенумеровать переменные $x_1, \ldots, x_n$). Тогда

$$F(x_1, \ldots, x_n) = \sum_{i=0}^{p} G_i(x_1, \ldots, x_{n-1}) x_n^i,$$

$$G_i \in \mathbb{Z}[x_1, \ldots, x_{n-1}], \quad G_p(x_1, \ldots, x_{n-1}) \neq 0.$$

Так как $F(j_1, \ldots, j_n) = 0$ для всех $j_1, \ldots, j_n \in \mathbb{N}$, то $G_i(j_1, \ldots, j_n) = 0$ для всех $j_1, \ldots, j_{n-1} \in \mathbb{N}$ и всех $0 \leqslant i \leqslant p$. Действительно, если $G_i(j_1, \ldots, j_{n-1}) \neq 0$, то

$$F(j_1, \ldots, j_{n-1}, x_n) = \sum_{i=0}^{p} G_i(j_1, \ldots, j_{n-1}) x_n^i -$$

ненулевой многочлен в $\mathbb{Z}[x_n]$ с бесконечным числом нулей, что невозможно (см. ??).

В силу индуктивного предположения $G_i(x_1, \ldots, x_{n-1}) = 0$ для всех $0 \leqslant i \leqslant p$, в частности для $i = p$, что приводит к противоречию с $G_p \neq 0$. □

Следствие 4.5.5. *Пусть*

$$f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n].$$

Тогда $f(r_1, \ldots, r_n) = 0$ для всех коммутативных колец R и всех наборов элементов $r_1, \ldots, r_n \in R$ тогда и только тогда, когда $f(j_1, \ldots, j_n) = 0$ для всех $(j_1, \ldots, j_n) \in \Delta^n$, где Δ — любое бесконечное подмножество в $\mathbb{N} \cup \{0\}$.

4.6. Делимость в кольцах многочленов $K[x_1, x_2, \ldots, x_n]$ от n переменных над полем K

Многочлен $\varphi = \varphi(x_1, x_2, \ldots, x_n)$ называется *делителем* многочлена $f = f(x_1, x_2, \ldots, x_n)$ (или будем в этом случае также говорить, что f *делится* на φ), если в $K[x_1, x_2, \ldots, x_n]$ найдётся многочлен $\psi = \psi(x_1, x_2, \ldots, x_n)$ такой, что $f = \varphi\psi$.

Общие свойства делимости 1)—9) для многочленов в $K[x]$ от одной переменной справедливы также и в $K[x_1, x_2, \ldots, x_n]$, при этом с теми же доказательствами.

Мы приведём их здесь для полноты изложения.

1) Если f делится на g, g делится на h, то f делится на h.

2) Если f и g делятся на φ, то $f \pm g$ делятся на φ.

3) Если f делится на φ, то fg также делится на φ для всех g.

4) Если f_i, $1 \leqslant i \leqslant k$, делятся на φ, то $\sum_{i=1}^{k} f_i g_i$ делится на φ для любых $g_1, g_2, \ldots, g_k$.

5) Если $0 \neq c \in K$, то любой многочлен f делится на c.

6) Если f делится на φ, то f делится на cf, где $0 \neq c \in K$.

7) Многочлены cf, где $0 \neq c \in K$, и только они являются делителями многочлена f, имеющими ту же степень, что и f.

8) Многочлены f и g одновременно делятся друг на друга тогда и только тогда, когда $g = cf$ для $0 \neq c \in K$.

9) Многочлены f и cf, $0 \neq c \in K$, имеют одно и то же множество делителей.

Многочлен $f = f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$, $\deg f = k \geqslant 1$ называется *неприводимым* над полем K, если он не разлагается в произведение многочленов $f = gh$, $g, h \in K[x_1, x_2, \ldots, x_n]$, таких, что $\deg g < \deg f$, $\deg h < \deg f$. В противном случае, если $f = gh$, где $\deg g < \deg f$, $\deg h < \deg f$, многочлен f называется *приводимым*.

Следующая теорема является естественным обобщением результата о факториальности кольца многочленов $K[x]$ от одной переменной над полем K.

Теорема 4.6.1 (о факториальности кольца многочленов от n переменных над полем). *Всякий многочлен $f = f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$, K — поле, $\deg f \geqslant 1$, разлагается в произведение неприводимых многочленов, при этом это разложение однозначно (с точностью до порядка сомножителей и множителей нулевой степени).*

Доказательство. 1) *Существование разложения.* Если многочлен f неприводим, то всё доказано. Если f — приводимый многочлен, то $f = gh$, $\deg g \geqslant 1$, $\deg h \geqslant 1$, $\deg f = \deg g + \deg h$. Продолжая этот процесс, после конечного числа шагов мы придём к его остановке, то есть к разложению многочлена $f(x_1, x_2, \ldots, x_n)$ в произведение неприводимых сомножителей.

2) *Единственность разложения.* Доказательство проведём индукцией по n. Начало индукции $n = 1$. Пусть теорема доказана для многочленов от n переменных $x_1, x_2, \ldots, x_n$. Покажем, что наше утверждение верно для многочленов от $n+1$ переменной $x_1, x_2, \ldots, x_n, x = x_{n+1}$. Так как $K[x_1, x_2, \ldots, x_n, x] = (K[x_1, x_2, \ldots, x_n])[x]$, то $f(x_1, x_2, \ldots, x_n, x) = \varphi(x)$, где коэффициенты многочлена $\varphi(x)$ лежат в $K[x_1, x_2, \ldots, x_n]$, для них, в силу индуктивного предположения, теорема верна, и поэтому каждый из них однозначно разлагается в произведение неприводимых сомножителей из $K[x_1, x_2, \ldots, x_n]$.

2а) Если коэффициенты многочлена $\varphi(x)$ не содержат ни одного общего неприводимого множителя (то есть в совокупности взаимно просты), то назовём многочлен $\varphi(x)$ *примитивным* (над кольцом $K[x_1, x_2, \ldots, x_n]$).

Нам понадобится в доказательстве следующее полезное утверждение.

Лемма 4.6.2 (лемма Гаусса). *Произведение двух примитивных многочленов само является примитивным многочленом.*

Доказательство. Допустим противное: пусть

$$f(x) = \sum_{i=0}^{k} a_i x^i, \; g(x) = \sum_{j=0}^{l} b_j x^j, \quad a_i, b_j \in K[x_1, x_2, \ldots, x_n], \; -$$

два примитивных многочлена с коэффициентами из $K[x_1, x_2, \ldots, x_n]$,

$$f(x)g(x) = \sum_{t=0}^{k+l} c_t x^t -$$

их произведение, которое не является примитивным, скажем, $p = p(x_1, x_2, \ldots, x_n)$ — общий неприводимый множитель коэффициентов $c_0, c_1, \ldots, c_{k+l}$. Так как $f(x)$ и $g(x)$ — примитивные многочлены, то все их коэффициенты $a_0, a_1, \ldots, a_k$ и $b_0, b_1, \ldots, b_l$ не могут делиться на p. Пусть коэффициенты в $f(x)$ $a_k, a_{k-1}, \ldots, a_{i+1}$ делятся на p, а коэффициент a_i (при x^i) — первый, считая слева, который не делится на p; аналогично, пусть коэффициенты в $g(x)$ $b_l, b_{l-1}, \ldots, b_{j+1}$ делятся на p, а коэффициент b_j (при x^j) не делится на p. Тогда рассмотрение коэффициента при x^{i+j} в $f(x)g(x)$

$$c_{i+j} = a_i b_j + (a_{i-1}b_{j+1} + \ldots + a_{i+j-l}b_l) + (a_{i+1}b_{j-1} + \ldots + a_k b_{i+j-k})$$

показывает, что c_{i+j} и две суммы в скобках делятся на p, поэтому $a_i b_j$ делится на p, что противоречит тому, что a_i и b_j не делятся на p. □

Продолжение доказательства теоремы. Кольцо многочленов $K[x_1, x_2, \ldots, x_n]$ содержится в своём поле частных — поле рациональных дробей $Q = K(x_1, x_2, \ldots, x_n)$, поэтому

$$(K[x_1, x_2, \ldots, x_n])[x] \subset Q[x].$$

Если $\varphi(x) \in Q[x]$, то каждый его коэффициент является дробью, в числителе и знаменателе которой стоят многочлены из $K[x_1, x_2, \ldots, x_n]$. Вынося общий знаменатель этих частных и общие простые множители из числителей, можно записать многочлен $\varphi(x)$ в следующем виде: $\varphi(x) = \dfrac{a}{b} f(x)$, где $a, b \in K[x_1, x_2, \ldots, x_n]$, $f(x) \in (K[x_1, x_2, \ldots, x_n])[x]$, при этом $f(x)$ — примитивный многочлен.

Построенный примитивный многочлен определён однозначно (с точностью до ненулевого элемента поля K). Действительно, если

$$\varphi(x) = \frac{a}{b} f(x) = \frac{c}{d} g(x),$$

где $g(x)$ — примитивный многочлен, то

$$adf(x) = bcg(x).$$

Из предполагаемого (по индуктивному предположению) свойства факториальности для $K[x_1, x_2, \ldots, x_n]$ следует, что ad и bc отличаются лишь на ненулевой элемент поля K. Отсюда следует, что на этот же множитель отличаются и примитивные многочлены $f(x)$ и $g(x)$.

2б) Отметим, что соответствие $\varphi(x) \mapsto f(x)$ сохраняет произведение. Действительно, если

$$\varphi(x) = \frac{a}{b} f(x), \quad \psi(x) = \frac{c}{d} g(x),$$

где $f(x)$, $g(x)$ — примитивные многочлены, $a, b, c, d \in K[x_1, x_2, \ldots, x_n]$, то

$$\varphi(x)\psi(x) = \frac{ac}{bd} f(x)g(x),$$

при этом $f(x)g(x)$ — примитивный многочлен, $ac, bd \in K[x_1, x_2, \ldots, x_n]$.

2в) Отметим, что многочлен $\varphi(x) \in Q[x]$ неприводим над полем Q тогда и только тогда, когда соответствующий ему примитивный многочлен $f(x) \in (K[x_1, x_2, \ldots, x_n])[x]$ неприводим над полем $Q = K(x_1, x_2, \ldots, x_n)$.

Действительно, если f приводим, $f = f_1 f_2$, $\deg f_1 \geqslant 1$, $\deg f_2 \geqslant 1$, поэтому f_1 и f_2 содержат ненулевые члены со степенями x, и следовательно,

$$\varphi(x) = \frac{a}{b} f(x) = \left(\frac{a}{b} f_1\right) f_2 -$$

приводимый многочлен над Q.

2г) Разложим примитивный многочлен f на неприводимые множители над полем Q:

$$f = f_1 f_2 \cdots f_k.$$

Тогда все f_i, $1 \leqslant i \leqslant k$, содержат переменную x и являются примитивными многочленами. Это разложение примитивного многочлена f однозначно (с точностью до ненулевых множителей из поля K). Действительно, в силу 2в) это разложение соответствует разложению $\varphi(x)$ на неприводимые множители в $Q[x]$ над полем Q, где однозначность разложения нам известна — с точностью до ненулевых множителей из поля $Q = K(x_1, x_2, \ldots, x_n)$. Учитывая, что все рассматриваемые нами сомножители f_i примитивные, получаем, что они определены однозначно с точностью до ненулевых элементов поля K.

3) *Завершение доказательства теоремы.* Как мы показали, всякий неприводимый многочлен кольца $K[x_1, x_2, \ldots, x_n, x]$ является или неприводимым многочленом из кольца $K[x_1, x_2, \ldots, x_n]$ (1-й тип, x не входит в многочлен), или неприводимым примитивным многочленом (2-й тип, x входит в ненулевой степени).

Если нам задано некоторое разложение многочлена $\varphi(x_1, x_2, \ldots, x_n, x)$ в произведение неприводимых множителей, то, объединяя сомножители 1-го и 2-го типа, получаем:

$$\varphi(x_1, x_2, \ldots, x_n, x) = a(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n, x),$$

где $a(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ не содержит $x = x_{n+1}$, $f(x_1, x_2, \ldots, x_n, x) \in Q[x]$ — примитивный многочлен. Это разложение уже однозначно для φ (с точностью до ненулевых элементов из поля K).

По предположению индукции имеет место однозначность разложения на неприводимые множители для многочлена $a(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$; однозначность разложения примитивного многочлена $f(x_1, x_2, \ldots, x_n, x)$ на неприводимые примитивные сомножители установлена в 2г) сведением задачи к случаю кольца $Q[x]$. Таким образом, единственность разложения в произведение неприводимых сомножителей установлена для случая $n+1$ неизвестных $x_1, x_2, \ldots, x_n, x = x_{n+1}$. □

Следствие 4.6.3 (из единственности разложения на неприводимые множители). *Если произведение fg двух многочленов*

$$f = f(x_1, x_2, \ldots, x_n), g = g(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$$

делится на неприводимый многочлен

$$p = p(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n],$$

то хотя бы один из многочленов f или g делится на p.

Доказательство. Допустив противное, получаем, что для произведения fg имеем два разложения на неприводимые множители, в одно из которых p входит, а другое разложение не содержит p. □

Упражнение 4.6.4. Пусть $X = (x_{ij})$, x_{ij} — переменные в кольце многочленов $K[x_{ij} \mid 1 \leqslant i, j \leqslant n]$. Тогда определитель $|X|$ — неприводимый многочлен от n^2 переменных x_{ij}.

Указание. $|X|$ — однородный многочлен степени 1 от элементов каждой строки и каждого столбца.

4.7. Поле рациональных дробей $K(x_1, x_2, \ldots, x_n)$ от n переменных $x_1, x_2, \ldots, x_n$ над полем K

Как мы видели (см. ??), кольцо многочленов над полем $R = K[x_1, x_2, \ldots, x_n]$ является коммутативным кольцом без делителей нуля. Тогда (см. ??) кольцо R вкладывается в своё поле частных $Q = Q(K) = K(x_1, x_2, \ldots, x_n)$, называемое *полем рациональных дробей от n переменных* $x_1, x_2, \ldots, x_n$ (при $n = 1$ имеем рассмотренное ранее поле рациональных дробей $K(x)$ от одной переменной x над полем K).

Как и в случае $n = 1$, всякая упорядоченная пара $\big(f(x_1, x_2, \ldots, x_n), g(x_1, x_2, \ldots, x_n)\big)$, где $g(x_1, x_2, \ldots, x_n) \neq 0$, называется рациональной дробью с числителем $f(x)$ и знаменателем $g(x)$: запись $\dfrac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)}$.

Две рациональные дроби $\dfrac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)}$ и $\dfrac{\varphi(x_1, x_2, \ldots, x_n)}{\psi(x_1, x_2, \ldots, x_n)}$ традиционно называются *равными* (на самом деле это отношение эквивалентности), если в кольце $K[x_1, x_2, \ldots, x_n]$ имеет место равенство многочленов

$$f(x_1, x_2, \ldots, x_n)\psi(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n)\varphi(x_1, x_2, \ldots, x_n).$$

Классы равных между собой рациональных дробей с операциями сложения и умножения на их представителях

$$\frac{f}{g} + \frac{\varphi}{\psi} = \frac{f\psi + g\varphi}{g\psi}; \quad \frac{f}{g} \cdot \frac{\varphi}{\psi} = \frac{f \cdot \varphi}{g \cdot \psi},$$

которые являются корректно определёнными (то есть не зависят от выбора представителей), образуют поле $K(x_1, x_2, \ldots, x_n)$, называемое *полем рациональных дробей*.

Сопоставление каждому многочлену $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ класса дробей, равных дроби $\dfrac{f(x_1, x_2, \ldots, x_n)}{1}$, осуществляет вложение кольца многочленов $K[x_1, x_2, \ldots, x_n]$ в поле $P(x_1, x_2, \ldots, x_n)$,

$$K[x_1, x_2, \ldots, x_n] \subset K(x_1, x_2, \ldots, x_n).$$

Так как

$$\frac{f(x_1, x_2, \ldots, x_n)}{1} \cdot \frac{1}{g(x_1, x_2, \ldots, x_n)} = \frac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)},$$

то все элементы нашего поля $K(x_1, x_2, \ldots, x_n)$ можно считать частными многочленов из кольца $K[x_1, x_2, \ldots, x_n]$.

4.8. Лексикографический порядок на множестве мономов в кольце многочленов $K[x_1, x_2, \ldots, x_n]$ от n переменных $x_1, x_2, \ldots, x_n$

Как мы уже видели, степень $d = \deg f(x_1, x_2, \ldots, x_n)$ многочлена $f(x_1, x_2, \ldots, x_n)$ по совокупности переменных $x_1, x_2, \ldots, x_n$ не позволяет нам ввести отношение линейного порядка на множестве мономов (эта степень лишь определяет старшую однородную компоненту f_d многочлена f, которая может содержать несколько разных одночленов совокупной степени d).

Необходимость рассмотрения линейных порядков на множестве всех мономов в кольце $K[x_1, x_2, \ldots, x_n]$, согласованных с умножением, вызвана нашим желанием иметь аналог *старшего члена* для многочлена от n переменных, обладающего рядом хороших свойств. Одним из таких отношений порядка на множестве всех мономов является *лексикографический порядок*, к рассмотрению которого мы приступаем (прилагательное «лексикографический» связано с некоторой аналогией с упорядочиванием слов в словаре).

Пусть $x_1 > x_2 > \ldots > x_n$ и мономы $u = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$, $v = x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n}$ различны, $u \neq v$ (это равносильно тому, что строки их показателей $(k_1, \ldots, k_n), (l_1, \ldots, l_n) \in \mathbb{N}_0^n$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, различны). Тогда хотя бы одна из разностей $k_i - l_i$, $i = 1, \ldots, n$, отлична от нуля. Будем говорить, что первый моном *старше* второго, $u > v$, если первая ненулевая из этих разностей (считая слева) положительна: $k_i - l_i > 0$ $(k_1 - l_1 = 0, \ldots, k_{i-1} - l_{i-1} = 0)$, то есть $k_1 = l_1, \ldots, k_{i-1} = l_{i-1}$, но $k_i > l_i$.

Предложение 4.8.1. *Отношение лексикографического отношения порядка $\geqslant$ на множестве всех мономов в $K[x_1, x_2, \ldots, x_n]$ обладает следующими свойствами:*

1) *отношение $u \geqslant v$ ($u = v$, или $u > v$ для $u \neq v$) является линейным порядком, то есть:*

   1.1) $u \geqslant u$;

   1.2) *если $u \geqslant v$, $v \geqslant u$, то $u = v$;*

   1.3) *если $u \geqslant v$, $v \geqslant w$, то $u \geqslant w$;*

   1.4) *если $u \neq v$, то либо $u > v$, либо $v > u$ (линейность порядка);*

2) *отношение порядка $u \geqslant v$ согласовано с умножением мономов: если $u \geqslant v$ и w — моном, то $uw \geqslant vw$ (и как следствие:*

3) *если $u \geqslant v$, $w \geqslant z$, то $uw \geqslant vz$).*

Доказательство. 1) 1.1. Ясно, что $u \geqslant u$.

1.2. Если $u \geqslant v$ и $v \geqslant u$, то $u > v$ и $v > u$ одновременно невозможно, поэтому $u = v$.

1.3. Пусть мономам u, v, w отвечают строки показателей $(k_1, k_2, \ldots, k_n)$, $(l_1, l_2, \ldots, l_n)$, $(m_1, m_2, \ldots, m_n)$ соответственно. Если $u > v$, то $k_1 = l_1, \ldots, k_{i-1} = l_{i-1}$, $k_i > l_i$ для некоторого $1 \leqslant i \leqslant n$; если $v > w$, то $l_1 = m_1, \ldots, l_{j-1} = m_{j-1}$, $l_j > m_j$ для некоторого $1 \leqslant j \leqslant n$. Для $t = \min\{i, j\}$: $k_1 = l_1 = m_1, \ldots, k_{t-1} = l_{t-1} = m_{t-1}$, $k_t \geqslant l_t \geqslant m_t$, при этом либо $k_t > l_t$ (если $t = i$), либо $l_t > m_t$ (если $t = j$), то есть всегда $k_t > m_t$, и поэтому $u > w$.

1.4 следует из определения нашего отношения в случае $u \neq v$: либо $u > v$, либо $u < v$.

2) Пусть мономам u, v, w отвечают строки показателей $(k_1, k_2, \ldots, k_n)$, $(l_1, l_2, \ldots, l_n)$ и $(m_1, m_2, \ldots, m_n)$ соответственно. Если $u = v$, то ясно, что $uw = vw$. Если $u > v$, то $k_1 = l_1, \ldots, k_{i-1} = l_{i-1}$, $k_i > l_i$ для некоторого $1 \leqslant i \leqslant n$. Тогда в строках показателей $(k_1 + m_1, k_2 + m_2, \ldots, k_n + m_n)$, $(l_1 + m_1, l_2 + m_2, \ldots, l_n + m_n)$, отвечающих мономам uw и vw соответственно, имеем $k_1 + m_1 = l_1 + m_1, \ldots, k_{i-1} + m_{i-1} = l_{i-1} + m_{i-1}$, $k_i + m_i > l_i + m_i$. Это означает, что $uw > vw$.

Свойство 3) является следствием свойства 2): $uw \geqslant vw \geqslant vz$. □

Замечания 4.8.2.

1) Если число переменных $n \geqslant 2$, то $x_1 > x_2^k$ для всех $k \in \mathbb{N}$. Таким образом, в конусе $\{u \mid u \leqslant x_1\}$ всех мономов u таких, что $u \leqslant x_1$, находится бесконечное число элементов. Это является существенным отличием лексикографического порядка на мономах от линейного порядка $\{\mathbb{N}, \leqslant\}$ на множестве $\mathbb{N}$ натуральных чисел.

2) Может быть, например, что при $n \geqslant 2$: $x_1 > x_2^2$, но $\deg x_1 = 1 < 2 = \deg x_2^2$.

3) Если рассмотреть упорядочение переменных $x_1, x_2, \ldots, x_n$, отличное от $x_1 > x_2 > \ldots > x_n$, то получим другой лексикографический порядок. Если специально не оговорено упорядочение на переменных $x_1, x_2, \ldots, x_n$, то считаем, что $x_1 > x_2 > \ldots > x_n$.

4) При любом лексикографическом порядке 1 меньше любого неединичного монома.

4.9. Старший член ненулевого многочлена $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ (относительно лексикографического порядка на мономах)

Среди конечного числа ненулевых одночленов ненулевого многочлена

$$f = f(x_1, x_2, \ldots, x_n) = \sum a_{l_1, l_2, \ldots, l_n} x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n} \in K[x_1, x_2, \ldots, x_n], \quad a_{l_1, l_2, \ldots, l_n} \neq 0,$$

имеется *единственный* одночлен

$$\mathrm{Lt}(f) = a_{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}, \quad a_{k_1, k_2, \ldots, k_n} \neq 0,$$

называемый *старшим членом* многочлена $f(x_1, x_2, \ldots, x_n)$, моном

$$x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

которого *старше* всех остальных мономов $x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n}$ ненулевых членов.

Пример 4.9.1. Многочлен

$$f(x_1, x_2) = 5x_1^3 + 2x_1^2 x_2^2 + 3x_1^2 x_2 + 4x_2^3 + 1$$

записан *лексикографически* (т. е. в лексикографическом порядке его мономы с ненулевыми коэффициентами образуют убывающую цепь $x_1^3 > x_1^2 x_2^2 > x_1^2 x_2 > x_2^3 > 1$), его старший член равен $5x_1^3$.

Следующее утверждение распространяет хорошо нами освоенное утверждение о том, что старший член произведения двух многочленов от одной переменной является произведением старших членов, на случай многочленов от нескольких переменных.

Лемма 4.9.2. *Старший член произведения*

$$h(x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n)g(x_1, x_2, \ldots, x_n)$$

двух ненулевых многочленов

$$f = f(x_1, x_2, \ldots, x_n), g = g(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$$

равен произведению

$$abx_1^{k_1+l_1}x_2^{k_2+l_2}\cdots x_n^{k_n+l_n}, \quad ab \neq 0,$$

старших членов $ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$, $a \neq 0$, *и* $bx_1^{l_1}x_2^{l_2}\cdots x_n^{l_n}$, $b \neq 0$, *сомножителей* $f = f(x_1, x_2, \ldots, x_n)$ *и* $g = g(x_1, x_2, \ldots, x_n)$ *соответственно.*

Доказательство. Пусть $u_1 = x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ — моном старшего члена многочлена $f(x_1, x_2, \ldots, x_n)$, v_1 — моном любого другого ненулевого одночлена в $f(x_1, x_2, \ldots, x_n)$. Тогда $u_1 > v_1$.

Аналогично, пусть $u_2 = x_1^{l_1}x_2^{l_2}\cdots x_n^{l_n}$ — моном старшего члена многочлена $g(x_1, x_2, \ldots, x_n)$, v_2 — моном любого другого ненулевого одночлена в $g(x_1, x_2, \ldots, x_n)$. Тогда $u_2 > v_2$.

В силу свойств лексикографического порядка на мономах из $u_1 > v_1$, $u_2 > v_2$ следует, что $u_1u_2 > v_1v_2$. Таким образом, при приведении подобных членов в произведении fg одночлен abu_1u_2, $ab \neq 0$, не сокращается и является старшим членом. □

4.10. Симметрические многочлены

Определение 4.10.1. Многочлен $f(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$, K — поле, называется *симметрическим*, если он не меняется ни при какой подстановке неизвестных (это означает, что $f(x_{\tau(1)}, \ldots, x_{\tau(n)}) = f(x_1, \ldots, x_n)$ для любой подстановки $\tau \in \mathrm{S}_n$).

Замечание 4.10.2.

1) При указанном действии группы S_n на множестве мономов от переменных $x_1, \ldots, x_n$ для $\sigma \in \mathrm{S}_n$ соответствующее отображение мономов

$$x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n} \mapsto x_{\sigma(1)}^{k_1}x_{\sigma(2)}^{k_2}\cdots x_{\sigma(n)}^{k_n}$$

является биекцией на множестве мономов.

2) Если многочлен $f(x_1, x_2, \ldots, x_n) \in K[x_1, \ldots, x_n]$ является симметрическим, то для любой подстановки $\sigma \in \mathrm{S}_n$ соответствующая ей биекция мономов переставляет одночлены многочлена $f(x_1, x_2, \ldots, x_n)$.

3) Поскольку любая подстановка $\tau \in \mathrm{S}_n$ является произведением транспозиций (циклов длины 2) (i, j), $i \neq j$, то для доказательства симметричности многочлена $f(x_1, \ldots, x_n)$ достаточно убедиться в том, что он не меняется ни при какой транспозиции двух неизвестных $x_i \leftrightarrow x_j$, $i \neq j$.

4) Однородные компоненты симметрического многочлена $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ также являются симметрическими многочленами, поскольку для $\sigma \in S_n$ степени мономов $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ и $x_{\sigma(1)}^{k_1} x_{\sigma(2)}^{k_2} \cdots x_{\sigma(n)}^{k_n}$ совпадают и равны $k_1 + k_2 + \ldots + k_n$.

Симметрический многочлен от переменных $x_1, x_2, \ldots, x_n$, все члены которого могут быть получены из одного одночлена применением подстановок переменных $x_1, x_2, \ldots, x_n$, называется *моногенным*. Если $ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$, $a \in K$, $k_i \geqslant 0$, — старший член моногенного многочлена, то этот многочлен обозначается через $S(ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n})$. Итак,

$$S(ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}) = \sum_{\sigma \in S_n} a x_{\sigma(1)}^{k_1} x_{\sigma(2)}^{k_2} \cdots x_{\sigma(n)}^{k_n}.$$

Например, от переменных x_1, x_2, x_3:

$$S(x_1^2 x_2) = x_1^2 x_2 + x_1^2 x_3 + x_2^2 x_1 + x_2^2 x_3 + x_3^2 x_1 + x_3^2 x_2.$$

Примеры 4.10.3.

1) Многочлен $f(x_1, x_2) = x_1^2 + x_2 \in \mathbb{R}[x_1, x_2]$ не является симметрическим, поскольку при транспозиции $x_1 \leftrightarrow x_2$:

$$f(x_2, x_1) = x_2^2 + x_1 \neq x_1^2 + x_2 = f(x_1, x_2).$$

2) *Элементарные симметрические многочлены*:

$$\sigma_1(x_1, \ldots, x_n) = x_1 + x_2 + \ldots + x_n$$

(сумма всех неизвестных);

$$\sigma_2(x_1, \ldots, x_n) = x_1 x_2 + x_1 x_3 + \ldots + x_{n-1} x_n = \sum_{i<j} x_i x_j$$

(сумма всех произведений $x_i x_j$, $i < j$, двух различных переменных x_i, x_j);

$$\ldots$$

$$\sigma_k(x_1, \ldots, x_n) = x_1 x_2 \cdots x_k + \ldots + x_{k+1} x_{k+2} \cdots x_n = \sum_{i_1 < \ldots < i_k} x_{i_1} \ldots x_{i_k}$$

(сумма всех C_n^k произведений из k различных переменных $x_{i_1} x_{i_2} \cdots x_{i_k}$, $i_1 < i_2 < \ldots < i_k$);

$$\ldots$$

$$\sigma_n(x_1, \ldots, x_n) = x_1 x_2 \cdots x_n.$$

Иногда удобно считать, что $\sigma_0 = 1$, $\sigma_k(x_1, \ldots, x_n) = 0$ при $k > n$.

Напомним, что эти многочлены уже возникали у нас в *формулах Виета*, связывающих коэффициенты многочлена $f(x) \in \mathbb{C}[x]$ с его корнями $x_1, \ldots, x_n \in \mathbb{C}$:

$$(x-x_1)\cdots(x-x_n) = x^n - \sigma_1(x_1,\ldots,x_n)x^{n-1} + \\ + \sigma_2(x_1,\ldots,x_n)x^{n-2} + \ldots + (-1)^n\sigma_n(x_1,\ldots,x_n),$$

или, иначе, если

$$(x-x_1)\cdots(x-x_n) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0,$$

то

$$\sigma_k(x_1,\ldots,x_n) = (-1)^k a_{n-k}.$$

Элементарные симметрические многочлены задают (и определяются) *производящей функцией*:

$$\sigma(t) = \sum_{k=0}^{\infty}\sigma_k t^k = \prod_{i=1}^{n}(1+tx_i).$$

3) *Полные однородные симметрические многочлены:*

$$p_k(x_1,\ldots,x_n) = \sum_{i_1+\ldots+i_n=k} x_1^{i_1}\ldots x_n^{i_n}.$$

Им соответствует производящая функция

$$p(t) = \sum_{k=0}^{\infty}p_k t^k = \prod_{i=1}^{n}(1-tx_i)^{-1}.$$

4) *Степенные суммы:*

$$s_1 = x_1 + x_2 + \ldots + x_n;$$
$$s_2 = x_1^2 + x_2^2 + \ldots + x_n^2$$

(сумма всех квадратов x_i^2);

$$\ldots$$
$$s_k = x_1^k + x_2^k + \ldots + x_n^k$$

(сумма всех k-х степеней x_i^k).

Степенным суммам соответствует производящая функция

$$s(t) = \sum_{k=0}^{\infty}s_k t^{k-1} = \sum_{i=1}^{n}x_i(1-tx_i)^{-1}.$$

5) *Мономиальные симметрические многочлены:*

$$m_{i_1\ldots i_n}(x_1,\ldots,x_n) = \sum_{\sigma\in S_n} x_{\sigma(1)}^{i_1}\cdots x_{\sigma(n)}^{i_n}.$$

6) Определитель Вандермонда

$$V(x_1, \ldots, x_n) = \prod_{i>j}(x_i - x_j)$$

не является симметрическим многочленом (если $\operatorname{char} K \neq 2$).

Определитель Вандермонда является кососимметрическим многочленом:

$$V(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma)V(x_1, \ldots, x_n)$$

(это равносильно тому, что при любой транспозиции переменных $x_i \leftrightarrow x_j$, $1 \leqslant i < j \leqslant n$, многочлен $V(x_1, \ldots, x_n)$ меняет знак).

Более того, любой кососимметрический многочлен $f(x_1, \ldots, x_n)$ представим в виде

$$f(x_1, \ldots, x_n) = V(x_1, \ldots, x_n)g(x_1, \ldots, x_n),$$

где $g(x_1, \ldots, x_n)$ — симметрический многочлен.

Однако симметрическим многочленом является квадрат определителя Вандермонда:

$$V(x_1, x_2, \ldots, x_n)^2 = \prod_{i>j}(x_i - x_j)^2.$$

Замечание 4.10.4 (о связях некоторых производящих функций симметрических многочленов).

1)

$$\sigma(t)p(-t) = \prod_{i=1}^{n}(1 + tx_i)\prod_{i=1}^{n}(1 + tx_i)^{-1} = 1,$$

и следовательно (приравнивая коэффициенты при t^n, $n \geqslant 1$, в левой и правой частях):

$$\sum_{r=0}^{n}(-1)^r\sigma_r p_{n-r} = 0. \tag{4.1}$$

Зафиксировав $\sigma_1, \ldots, \sigma_k$, получаем систему линейных уравнений для $p_1, \ldots, p_k$, откуда:

$$\sigma_k = \begin{vmatrix} p_1 & 1 & 0 & \ldots & 0 \\ p_2 & p_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & p_{k-3} & \ldots & 1 \\ p_k & p_{k-1} & p_{k-2} & \ldots & p_1 \end{vmatrix}.$$

Зафиксировав $p_1, \ldots, p_k$, получаем систему линейных уравнений для $\sigma_1, \ldots, \sigma_k$, откуда:

$$p_k = \begin{vmatrix} \sigma_1 & 1 & 0 & \ldots & 0 \\ \sigma_2 & \sigma_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{k-1} & \sigma_{k-2} & \sigma_{k-3} & \ldots & 1 \\ \sigma_k & \sigma_{k-1} & \sigma_{k-2} & \ldots & \sigma_1 \end{vmatrix}.$$

2)

$$s(t) = \sum_{i=1}^{n} x_i(1-tx_i)^{-1} = \left(\prod_{i=1}^{n}(1-tx_i)^{-1}\right)' \prod_{i=1}^{n}(1-tx_i) = \frac{d}{dt}\ln p(t) = \frac{p'(t)}{p(t)},$$

и следовательно,

$$s(t)p(t) = p'(t).$$

Приравнивая коэффициенты при t^{n+1}, получаем

$$\sum_{r=1}^{n} s_r p_{n-r} = np_n. \tag{4.2}$$

Поэтому:

$$s_k = (-1)^{k-1}\begin{vmatrix} p_1 & 1 & 0 & \dots & 0 \\ 2p_2 & p_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ kp_k & p_{k-1} & p_{k-2} & \dots & p_1 \end{vmatrix}; \quad p_k = \frac{1}{k!}\begin{vmatrix} s_1 & -1 & 0 & \dots & 0 \\ s_2 & s_1 & -2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_{k-2} & s_{k-3} & \dots & -(k-1) \\ s_k & s_{k-1} & s_{k-2} & \dots & s_1 \end{vmatrix}.$$

3)

$$s(-t) = \sum_{i=1}^{n} x_i(1+tx_i)^{-1} = -\frac{d}{dt}\left(\ln\prod_{i=1}^{n}(1+tx_i)\right) = -\frac{d}{dt}\ln\sigma(t) = -\frac{\sigma'(t)}{\sigma(t)},$$

и следовательно,

$$s(-t)\sigma(t) = -\sigma'(t).$$

Приравнивая коэффициенты при t^{n+1}, получаем

$$\sum_{r=1}^{n}(-1)^{r-1} s_r \sigma_{n-r} = n\sigma_n \tag{4.3}$$

(формулы Ньютона). Поэтому:

$$s_k = \begin{vmatrix} \sigma_1 & 1 & 0 & \dots & 0 \\ 2\sigma_2 & \sigma_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k\sigma_k & \sigma_{k-1} & \sigma_{k-2} & \dots & \sigma_1 \end{vmatrix}; \quad \sigma_k = \frac{1}{k!}\begin{vmatrix} s_1 & 1 & 0 & \dots & 0 \\ s_2 & s_1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_{k-2} & s_{k-3} & \dots & k-1 \\ s_k & s_{k-1} & s_{k-2} & \dots & s_1 \end{vmatrix}.$$

Лемма 4.10.5. *Симметрические многочлены от n переменных $x_1, \dots, x_n$ с коэффициентами из поля K образуют подкольцо в кольце $K[x_1, \dots, x_n]$, называемое* кольцом симметрических многочленов от n переменных над полем K.

Доказательство. Так как сумма, разность и произведение двух симметрических многочленов являются симметрическими многочленами, то симметрические многочлены образуют подкольцо в кольце $K[x_1, \dots, x_n]$. □

Следствие 4.10.6. *Если $F(X_1,\dots,X_m) \in K[X_1,\dots,X_m]$ — произвольный многочлен от m переменных $X_1,\dots,X_m$, $f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n) \in K[x_1,\dots,x_n]$ — m произвольных симметрических многочленов от переменных $x_1,\dots,x_n$, то $F(f_1,f_2,\dots,f_m) \in K[x_1,\dots,x_n]$ — симметрический многочлен от переменных $x_1,\dots,x_n$.*

В частности, любой многочлен $F(\sigma_1,\dots,\sigma_n)$, $F(X_1,\dots,X_n) \in K[X_1,\dots,X_n]$, от элементарных симметрических многочленов $\sigma_1,\dots,\sigma_n \in K[x_1,\dots,x_n]$ является симметрическим многочленом от переменных $x_1,\dots,x_n$.

Пример 4.10.7. Многочлен

$$f(x_1,x_2,x_3) = \sigma_1\sigma_2 + 2\sigma_3 = x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_2^2x_3 + x_1x_3^2 + x_2x_3^2 + 5x_1x_2x_3$$

является симметрическим многочленом ($F(X_1,X_2,X_3) = X_1X_2 + 2X_3$).

Глубоким обращением этого простого замечания о том, что любой многочлен от элементарных симметрических многочленов оказывается симметрическим многочленом, явилась следующая теорема.

Теорема 4.10.8 (основная теорема о симметрических многочленах). *Всякий симметрический многочлен $f(x_1,x_2,\dots,x_n) \in K[x_1,x_2,\dots,x_n]$ над полем K единственным образом представляется в виде многочлена $F(\sigma_1,\sigma_2,\dots,\sigma_n)$, $F(X_1,X_2,\dots,X_n) \in K[X_1,X_2,\dots,X_n]$, от элементарных симметрических многочленов $\sigma_1(x_1,x_2,\dots,x_n)$, $\sigma_2(x_1,x_2,\dots,x_n),\dots,\sigma_n(x_1,x_2,\dots,x_n)$ (с коэффициентами из исходного поля K).*

Эта теорема явилась одним из краеугольных камней *теории инвариантов* (подкольцо R^G неподвижных элементов при действии на кольце многочленов $R = K[x_1,\dots,x_n]$ группы $G = \mathrm{S}_n$ оказалось совпадающим с подкольцом $K[\sigma_1,\dots,\sigma_n]$, а потому изоморфным кольцу многочленов от n переменных над полем K).

Для доказательства этой теоремы нам понадобятся следующие два ключевых утверждения о старших членах симметрических многочленов.

Лемма 4.10.9 (о старшем члене симметрического многочлена). *Пусть*

$$a_{k_1,\dots,k_n}x_1^{k_1}x_2^{k_2}\dots x_n^{k_n}, \quad 0 \neq a = a_{k_1,\dots,k_n} \in K, \ —$$

старший одночлен симметрического многочлена $f(x_1,\dots,x_n) \in K[x_1,\dots,x_n]$. Тогда

$$k_1 \geqslant k_2 \geqslant \dots \geqslant k_n.$$

Доказательство. Допустим противное: пусть $k_i < k_{i+1}$ для некоторого i, $1 \leqslant i \leqslant n-1$. Так как $f(x_1,\dots,x_n) = f(x_{\sigma(1)},\dots,x_{\sigma(n)})$ для $\sigma = (i,i+1)$, то (см. замечание ??) $f(x_1,\dots,x_n)$ содержит одночлен

$$ax_1^{k_1}\dots x_i^{k_{i+1}}x_{i+1}^{k_i}\dots x_n^{k_n},$$

получающийся из старшего члена подстановкой $x_i \leftrightarrow x_{i+1}$. В наших предположениях

$$(k_1,\dots,k_{i+1},k_i,\dots,k_n) > (k_1,\dots,k_i,k_{i+1},\dots,k_n),$$

что противоречит определению старшего члена многочлена $f(x_1,\dots,x_n)$. □

Следствие 4.10.10. *В мономе $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ старшего члена симметрического многочлена $f(x_1,x_2,\ldots,x_n)\in K[x_1,\ldots,x_n]$ имеет место неравенство $k_1\geqslant k_i$ для всех $i=1,2,\ldots,n$.*

Лемма 4.10.11 (о старшем члене произведения элементарных симметрических многочленов).

1) *Старший член симметрического многочлена $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}\in K[x_1,\ldots,x_n]$, где $l_i\geqslant 0$, $i=1,\ldots,n$, равен*

$$x_1^{k_1}\cdots x_n^{k_n},\quad \text{где}\quad \begin{cases} k_1 = \quad l_1+l_2+\ldots+l_n,\\ k_2 = \qquad l_2+\ldots+l_n,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ k_n = \qquad\qquad\qquad l_n. \end{cases}$$

2) *Для любого монома $x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$, где $k_1\geqslant k_2\geqslant\ldots\geqslant k_n$, существуют неотрицательные числа $l_1,l_2,\ldots,l_n\in\mathbb{N}_0$ такие, что старший член симметрического многочлена $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}$ совпадает с $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ (этим условием строка $(l_1,l_2,\ldots,l_n)$ определена однозначно).*

Доказательство. 1) Так как старший член многочлена $\sigma_k(x_1,\ldots,x_n)$ равен $x_1x_2\cdots x_k$, то в силу леммы ?? (о старшем члене произведения), старший член многочлена $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}$ равен

$$x_1^{l_1}(x_1x_2)^{l_2}\cdots(x_1x_2\cdots x_n)^{l_n}=x_1^{l_1+\ldots+l_n}x_2^{l_2+\ldots+l_n}\ldots x_n^{l_n}.$$

Таким образом:

$$\begin{cases} l_1+l_2+\ldots+l_n=k_1,\\ \qquad l_2+\ldots+l_n=k_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ \qquad\qquad\qquad l_n=k_n. \end{cases}$$

2) Полученная система линейных уравнений имеет единственное решение

$$l_i=k_i-k_{i+1},\quad i=1,2,\ldots,n-1,\quad l_n=k_n.$$

Если $k_1\geqslant k_2\geqslant\ldots\geqslant k_n$, то $l_i\geqslant 0$, $i=1,\ldots,n$. □

Следствия 4.10.12.

1) *Равенство $l_1+l_2+\ldots+l_n=k_1$ означает, что степень монома $X_1^{l_1}X_2^{l_2}\cdots X_n^{l_n}$ по совокупности переменных равна степени старшего члена $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ многочлена $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}\in K[x_1,\ldots,x_n]$ по переменной x_1.*

2) *$\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}\in K[x_1,x_2,\ldots,x_n]$ — однородный симметрический многочлен от переменных $x_1,x_2,\ldots,x_n$ степени*

$$s=l_1+2l_2+\ldots+(n-1)l_{n-1}+nl_n$$

(число $l_1+2l_2+\ldots+(n-1)l_{n-1}+nl_n$ называется весом $\omega(X^{l_1}X_2^{l_2}\ldots X_{n-1}^{l_{n-1}}X_n^{l_n})$ монома $X^{l_1}X_2^{l_2}\ldots X_{n-1}^{l_{n-1}}X_n^{l_n}$, если заданы веса $\omega(X_1)=1,\omega(X_2)=2,\ldots,\omega(X_{n-1})=n-1$, $\omega(X_n)=n$).

Доказательство-алгоритм представления симметрического многочлена $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ в виде многочлена $F(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $F(X_1, X_2, \ldots, X_n) \in K[X_1, X_2, \ldots, X_n]$, от элементарных симметрических многочленов $\sigma_1, \sigma_2, \ldots, \sigma_n \in K[x_1, x_2, \ldots, x_n]$

1) Если $f = 0$, то возьмём $F = 0$.

2) Если $f_1 = f(x_1, x_2, \ldots, x_n) \neq 0$, то рассмотрим его старший член

$$u_1 = a_{k_1,k_2,\ldots,k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}, \quad a_{k_1,k_2,\ldots,k_n} \neq 0,$$

где в силу леммы 4.10.9

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_n.$$

В силу леммы 4.10.11 найдём такой одночлен $F_1 = a_{k_1,k_2,\ldots,k_n} X_1^{l_1} X_2^{l_2} \ldots X_n^{l_n}$, что старший член симметрического многочлена

$$F_1(\sigma_1, \sigma_2, \ldots, \sigma_n) = a_{k_1,k_2,\ldots,k_n} \sigma_1^{l_1} \sigma_2^{l_2} \cdots \sigma_n^{l_n} \in K[x_1, \ldots, x_n]$$

совпадает со старшим членом u_1 исходного многочлена $f_1 = f(x_1, x_2, \ldots, x_n)$.

Разность

$$f_2(x_1, x_2, \ldots, x_n) = f_1 - F_1(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n]$$

является симметрическим многочленом от $x_1, x_2, \ldots, x_n$ (как разность симметрических многочленов).

Если $f_2 = 0$, то положим $F = F_1$. Если $f_2(x_1, x_2, \ldots, x_n) \neq 0$, то повторяем нашу процедуру: пусть u_2 — старший член многочлена $f_2(x_1, x_2, \ldots, x_n)$, $F_2 \in K[X_1, X_2, \ldots, X_n]$ — такой одночлен, что старший член многочлена $F_2[\sigma_1, \sigma_2, \ldots, \sigma_n] \in K[x_1, x_2, \ldots, x_n]$ равен одночлену u_2,

$$f_3(x_1, x_2, \ldots, x_n) = f_2 - F_2(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n] -$$

симметрический многочлен.

Если $f_3 = 0$, то положим $F = F_1 + F_2$. Если $f_3(x_1, x_2, \ldots, x_n) \neq 0$, то продолжаем процесс и получаем последовательность симметрических многочленов

$$f_1(x_1, \ldots, x_n) = f(x_1, \ldots, x_n), \ f_2(x_1, \ldots, x_n), \ f_3(x_1, \ldots, x_n), \ldots,$$

старшие члены которых в лексикографическом порядке образуют цепь

$$u_1 = u > u_2 > u_3 > \ldots.$$

В силу следствия к лемме 4.10.9 степень любой переменной в u_i, $i \geqslant 1$, не превосходит степени переменной x_1, которая в этих мономах не превосходит k_1. Поэтому число различных возможных таких мономов не превосходит числа k_1^n. Это доказывает, что описанный процесс обрывается. Таким образом, после конечного числа m шагов имеем $f_{m+1} = 0$, и поэтому можно положить $F = F_1 + F_2 + \ldots + F_m$. □

Доказательство единственности представления симметрического многочлена $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ в виде многочлена от элементарных симметрических многочленов

Пусть $F(X_1, \ldots, X_n), G(X_1, \ldots, X_n) \in K[X_1, \ldots, X_n]$ и $f(x_1, x_2, \ldots, x_n) = F(\sigma_1, \sigma_2, \ldots, \sigma_n) = G(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $H = F - G \in K[X_1, \ldots, X_n]$. Тогда $H(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0$ (как многочлен от $x_1, x_2, \ldots, x_n$). Наша цель — доказать, что $F = G$ в $K[X_1, \ldots, X_n]$, то есть что $H = 0$ как многочлен от $X_1, X_2, \ldots, X_n$.

Допустим противное: пусть $H \neq 0$, $H = \sum\limits_{i=1}^{s} H_i$, где $H_1, H_2, \ldots, H_s$ — все ненулевые одночлены многочлена $H(X_1, X_2, \ldots, X_n)$.

Рассмотрим $u_i(x_1, x_2, \ldots, x_n)$ $(i = 1, 2, \ldots, s)$ — старший член многочлена $H_i(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n]$. В силу леммы 4.10.9 мономы одночленов $u_1, u_2, \ldots, u_s$ различны. Выберем из них самый старший, пусть это будет одночлен u_k. Тогда в силу нашего построения:

а) одночлен u_k старше всех остальных одночленов многочлена $H_k(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n]$;

б) одночлен u_k старше старшего члена u_i многочлена $H_i(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, \ldots, x_n]$, и поэтому старше всех членов многочленов $H_i(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $i \neq k$.

Поэтому после приведения подобных членов в сумме

$$H(\sigma_1, \sigma_2, \ldots, \sigma_n) = \sum_{i=1}^{s} H_i(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n]$$

одночлен $u_k(x_1, \ldots, x_n)$ не сократится, и поэтому $H(\sigma_1, \sigma_2, \ldots, \sigma_n)$ — ненулевой многочлен от переменных $x_1, x_2, \ldots, x_n$, что противоречит нашему предположению. □

Следствие 4.10.13. *Пусть $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ — симметрический многочлен, $a_{k_1,k_2,\ldots,k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ — его старший член, $f(x_1, x_2, \ldots, x_n) = \Phi(\sigma_1, \sigma_2, \ldots, \sigma_n)$, где $\Phi(X_1, X_2, \ldots, X_n) \in K[X_1, X_2, \ldots, X_n]$. Тогда:*

$$\deg \Phi(X_1, X_2, \ldots, X_n) = k_1 = \deg_{x_1} f(x_1, x_2, \ldots, x_n)$$
$$(= \deg_{x_i} f(x_1, x_2, \ldots, x_n) \text{ для всех } i,\ 1 \leqslant i \leqslant n).$$

Доказательство. 1) Из определения старшего члена следует, что $k_1 \geqslant l_1$ для любого одночлена $a_{l_1,l_2,\ldots,l_n} x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n}$ многочлена $f(x_1, x_2, \ldots, x_n)$. Таким образом,

$$k_1 = \deg_{x_1} f(x_1, x_2, \ldots, x_n).$$

2) Из симметричности многочлена $f(x_1, x_2, \ldots, x_n)$ (в частности, его инвариантности при подстановке $x_1 \leftrightarrow x_i$, $2 \leqslant i \leqslant n$) следует, что

$$\deg_{x_1} f(x_1, x_2, \ldots, x_n) = \deg_{x_i} f(x_1, x_2, \ldots, x_n), \quad 1 \leqslant i \leqslant n.$$

3) Для $\Phi_1(X_1, X_2, \ldots, X_n) = X_1^{k_1-k_2} X_2^{k_2-k_3} \cdots X_{n-1}^{k_{n-1}-k_n} X_n^{k_n}$ имеем:

$$\deg \Phi_1(X_1, X_2, \ldots, X_n) = (k_1 - k_2) + (k_2 - k_3) + \ldots + (k_{n-1} - k_n) + k_n =$$
$$= k_1 = \deg_{x_1} f(x_1, x_2, \ldots, x_n).$$

Далее,

$$\deg \Phi_2(X_1, X_2, \ldots, X_n) = \deg_{x_1} f_2(x_1, x_2, \ldots, x_n) \leqslant \deg_{x_1} f(x_1, x_2, \ldots, x_n) = k_1.$$

Таким образом, для всех i, $1 \leqslant i \leqslant m$,

$$\deg \Phi_i(X_1, X_2, \ldots, X_n) \leqslant \deg_{x_1} f(x_1, x_2, \ldots, x_n) = k_1,$$

и поэтому для $\Phi = \sum\limits_{i=1}^{m} \Phi_i$:

$$\deg \Phi(X_1, X_2, \ldots, X_n) \leqslant \deg_{x_1} f(x_1, x_2, \ldots, x_n) = k_1.$$

Так как мономы одночленов $\Phi_i(X_1, X_2, \ldots, X_n)$ отвечают различным старшим членам $u_i(x_1, x_2, \ldots, x_n)$ симметрических многочленов $f_i(x_1, x_2, \ldots, x_n)$, $1 \leqslant i \leqslant m$, то $\Phi_i(X_1, X_2, \ldots, X_n)$ не сокращаются в $\Phi = \sum\limits_{i=1}^{m} \Phi_i$. Поэтому наличие в $\Phi = \sum\limits_{i=1}^{m} \Phi_i$ одного одночлена Φ_1 с $\deg \Phi_1 = k_1$ показывает, что $\deg \Phi(X_1, X_2, \ldots, X_n) = k_1$. □

Следствие 4.10.14. *Если $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ — однородный симметрический многочлен, $\deg f(x_1, x_2, \ldots, x_n) = s$ (то есть $f = f_s$), $f(x_1, x_2, \ldots, x_n) = \Phi(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $\Phi(X_1, X_2, \ldots, X_n) = \sum\limits_{i=1}^{m} \Phi_i \in K[X_1, X_2, \ldots, X_n]$ (здесь $\Phi_1, \ldots, \Phi_m$ — одночлены от $X_1, X_2, \ldots, X_n$ в нашем алгоритме), то веса всех мономов одночленов Φ_i, $1 \leqslant i \leqslant m$, совпадают и равны s (напомним, что под весом монома $X_1^{l_1} X_2^{l_2} \ldots X_n^{l_n}$ понимается число $\omega(X_1^{l_1} X_2^{l_2} \ldots X_n^{l_n}) = l_1 + 2l_2 + \ldots + nl_n$, определяемое следующим заданием весов переменных: $\omega(X_1) = 1, \omega(X_2) = 2, \ldots, \omega(X_n) = n$).*

Доказательство. Если $a_{k_1,k_2,\ldots,k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ — старший член однородного многочлена $f_1 = f(x_1, x_2, \ldots, x_n)$, то $s = k_1 + k_2 + \ldots + k_n$. Так как моном одночлена $\Phi_1(X_1, X_2, \ldots, X_n)$ равен $X_1^{k_1-k_2} X_2^{k_2-k_3} \ldots X_{n-1}^{k_{n-1}-k_n} X_n^{k_n}$, то

$$\begin{aligned}\omega(X_1^{k_1-k_2} X_2^{k_2-k_3} \ldots X_{n-1}^{k_{n-1}-k_n} X_n^{k_n}) &= \\ = (k_1 - k_2) + 2(k_2 - k_3) + \ldots + (n-1)(k_{n-1} - k_n) + nk_n &= \\ = k_1 + k_2 + k_3 + \ldots + k_n = s.&\end{aligned}$$

Далее, $f_2 = f_1 - \Phi_1(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n]$ также однородный многочлен от $x_1, x_2, \ldots, x_n$ степени s (как разность двух однородных многочленов степени s), поэтому вес монома одночлена $\Phi_2(X_1, X_2, \ldots, X_n)$ будет равен s. Продолжая это рассуждение, получаем наше утверждение для всех i, $1 \leqslant i \leqslant m$. □

Замечания 4.10.15. Из доказательства основной теоремы о симметрических многочленах следует ряд следствий.

1) Коэффициенты найденного для симметрического многочлена $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ не только принадлежат основному полю K, но выражаются через коэффициенты многочлена $f(x_1, x_2, \ldots, x_n)$ с помощью операций сложения и вычитания, и поэтому принадлежат подкольцу поля K, порождаемому коэффициентами многочлена $f(x_1, x_2, \ldots, x_n)$.

2) Утверждение о единственности представления симметрического многочлена в виде многочлена от элементарных симметрических многочленов можно переформулировать в следующей форме: элементарные симметрические многочлены $\sigma_1, \sigma_2, \ldots, \sigma_n$ алгебраически независимы в коммутативной алгебре $K[x_1, x_2, \ldots, x_n]$ над полем K (см. ??).

Примеры применения основной теоремы о симметрических многочленах

1) $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 = \sigma_1^2 - 3\sigma_2 + 3$ $(F(X_1, X_2) = X_1^2 - 3X_2 + 3)$.

2) *Степенные суммы.* Ясно, что:

$$s_1 = x_1 + \ldots + x_n = \sigma_1 \quad (F = X_1);$$
$$s_2 = x_1^2 + \ldots + x_n^2 = \sigma_1^2 - 2\sigma_2 \quad (F = X_1^2 - 2X_2).$$

Для представления многочлена

$$s_3 = x_1^3 + \ldots + x_n^3$$

в виде многочлена от $\sigma_1, \ldots, \sigma_n$ запишем *применение алгоритма из доказательства основной теоремы* в виде таблицы:

m	$f_m = f_{m-1} - F_{m-1}(\sigma_1, \ldots, \sigma_n)$	u_m	$F_m(\sigma_1, \sigma_2, \ldots, \sigma_n)$
1	$s_3 = x_1^3 + x_2^3 + \ldots + x_n^3$	x_1^3	$\sigma_1^3 = \sum\limits_{i=1}^{n} x_i^3 + 3\sum\limits_{i \neq j} x_i^2 x_j + 6\sum\limits_{i<j<k} x_i x_j x_k$
2	$-3\sum\limits_{i \neq j} x_i^2 x_j - 6\sum\limits_{i<j<k} x_i x_j x_k$	$-3x_1^2 x_2$	$-3\sigma_1\sigma_2 = -3\sum\limits_{i \neq j} x_i^2 x_j - 9\sum\limits_{i<j<k} x_i x_j x_k$
3	$3\sum\limits_{i<j<k} x_i x_j x_k$	$3x_1x_2x_3$	$3\sigma_3 = 3\sum\limits_{i<j<k} x_i x_j x_k$
4	0		

Таким образом:

$$s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 \quad (F = X_1^3 - 3X_1X_2 + 3X_3).$$

Замечание 4.10.16. Следующие формулы Ньютона и Варинга (см. с. 135) дают в общем случае выражения s_k через $\sigma_1, \ldots, \sigma_n$, а также выражения σ_k через $s_1, \ldots, s_k, \ldots$.

Степенные суммы $s_k = x_1^k + x_2^k + \ldots + x_n^k$ с элементарными симметрическими многочленами связаны *формулами Ньютона:*

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \ldots + (-1)^{k-1}s_1\sigma_{k-1} + (-1)^k k\sigma_k = 0 \quad (k \leqslant n);$$
$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \ldots + (-1)^n s_{k-n}\sigma_n = 0 \quad (k > n).$$

Эти формулы позволяют последовательно находить выражения $s_1, s_2, s_3, \ldots$ через $\sigma_1, \sigma_2, \ldots, \sigma_n$ (или наоборот).

Пример 4.10.17 ($n \geqslant 3$). $s_1 = \sigma_1$; $s_2 - s_1\sigma_1 + 2\sigma_2 = 0$, откуда $s_2 = \sigma_1^2 - 2\sigma_2$; $s_3 - s_2\sigma_1 + s_1\sigma_2 - 3\sigma_3 = 0$, откуда $s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$.

Общая формула, выражающая s_k $(k = 1, 2, 3, \ldots)$ через $\sigma_1, \sigma_2, \ldots, \sigma_n$ для поля K, $\operatorname{char} K = 0$, то есть $\mathbb{Q} \subseteq K$ (первая формула Варинга):

$$s_k = k\sum(-1)^{2\lambda_1 + 3\lambda_2 + \ldots + (n+1)\lambda_n}\frac{(\lambda_1 + \ldots + \lambda_n - 1)!\sigma_1^{\lambda_1}\cdots\sigma_n^{\lambda_n}}{\lambda_1!\cdots\lambda_n!},$$

где суммирование ведётся по всем наборам $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ неотрицательных чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ таких, что $\lambda_1 + 2\lambda_2 + \ldots + n\lambda_n = k$.

Пример 4.10.18. Выражение для s_4. При $k = 4$ имеем следующие наборы:

$$\begin{aligned}
&\lambda_1 = 4, \quad \lambda_2 = \ldots = \lambda_n = 0;\\
&\lambda_1 = 2, \quad \lambda_2 = 1, \quad \lambda_3 = \ldots = \lambda_n = 0;\\
&\lambda_1 = 1, \quad \lambda_2 = 0, \quad \lambda_3 = 1, \quad \lambda_4 = \ldots = \lambda_n = 0;\\
&\lambda_1 = 0, \quad \lambda_2 = 2, \quad \lambda_3 = \ldots = \lambda_n = 0;\\
&\lambda_1 = \lambda_2 = \lambda_3 = 0, \quad \lambda_4 = 1, \quad \lambda_5 = \ldots = \lambda_n = 0;
\end{aligned}$$

откуда

$$s_4 = 4\left((-1)^8\frac{3!\sigma_1^4}{4!} + (-1)^7\frac{2!\sigma_1^2\sigma_2}{2!1!} + (-1)^6\frac{1!\sigma_1\sigma_3}{1!1!} + \right.$$
$$\left. + (-1)^6\frac{1!\sigma_2^2}{2!} + (-1)^5\frac{\sigma_4}{1!}\right) = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_1\sigma_3 + 2\sigma_2^2 - 4\sigma_4.$$

Элементарные симметрические многочлены $\sigma_1, \sigma_2, \ldots, \sigma_n$ от переменных $x_1, x_2, \ldots, x_n$ выражаются через степенные суммы над полем K, $\operatorname{char} K = 0$, то есть $\mathbb{Q} \subseteq K$ (вторая формула Варинга):

$$\sigma_k = \sum \frac{(-1)^{\lambda_1+\ldots+\lambda_k+k}}{1^{\lambda_1}2^{\lambda_2}\ldots k^{\lambda_k}\lambda_1!\cdots\lambda_k!}s_1^{\lambda_1}\cdots s_k^{\lambda_k},$$

суммирование ведётся по тем же наборам $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, что и в первой формуле Варинга.

Пример 4.10.19.

$$\sigma_3 = \frac{(-1)^{3+3}}{1^3\cdot 3!}s_1^3 + \frac{(-1)^{1+1+3}}{1\cdot 2\cdot 1!\cdot 1!} + \frac{(-1)^{1+3}}{3\cdot 1!}s_3 = \frac{1}{6}s_1^3 - \frac{1}{2}s_1s_2 + \frac{1}{3}s_3.$$

3) Для нахождения выражения симметрического многочлена через элементарные удобно сначала разбить исходный многочлен на однородные компоненты, собирая вместе все члены многочлена, имеющие одну и ту же степень $s = k_1 + k_2 + \ldots + k_n$ по совокупности переменных, а затем к однородным компонентам, также являющимся симметрическими многочленами, применять наш алгоритм.

В случае однородного симметрического многочлена со старшим членом $u_1 = a_1x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$, $a_1 \in K$, $(k_1, k_2, \ldots, k_n)$ — строчка показателей степеней, можно *сразу выписать всевозможные строчки*

$$(l_1, l_2, \ldots, l_n), \quad l_i \geqslant 0,$$

показателей степеней такие, что:

1) $l_1 + l_2 + \ldots + l_n = k_1 + k_2 + \ldots + k_n$;

2) $l_1 \geqslant l_2 \geqslant \ldots \geqslant l_n$;

3) моном $x_1^{l_1}x_2^{l_2}\cdots x_n^{l_n}$ не выше монома $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$.

Далее, для каждого такого набора $l = (l_1, l_2, \ldots, l_n)$ составляем произведение $F_l = \sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\cdots\sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n}$. Так как мономы одночленов $u_1, u_2, \ldots, u_m$ в нашем алгоритме удовлетворяют свойствам 1)—3), то многочлен $f(x_1, x_2, \ldots, x_n)$ можно искать в виде суммы построенных произведений F_l *с неопределёнными коэффициентами* a_l:

$$f(x_1, x_2, \ldots, x_n) = \sum_l a_lF_l$$

(коэффициент a_1 при $\sigma_1^{k_1-k_2}\cdots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n}$ уже задан).

В силу нашей теоремы (утверждение о единственности представления) коэффициенты $a_l \in K$ определены однозначно. Их можно найти, выбирая различные наборы значений из поля K для переменных $x_1, x_2, \ldots, x_n$ в левой и правой частях равенства.

Пример 4.10.20. Однородный симметрический многочлен

$$f(x_1,x_2,x_3) = -x_1^3x_2 - x_1x_2^3 - x_1^3x_3 - x_1x_3^3 - x_2^3x_3 - x_2x_3^3 + x_1^2x_2x_3 + x_1x_2^2x_3 + x_1x_2x_3^2,$$

$K = \mathbb{Q}$, старший член: $-x_1^3x_2$, $(3,1,0)$.

Составляем таблицу всех строк l, удовлетворяющих условиям 1)—3), начиная со строки $(3,1,0)$, и соответствующих одночленов $F_l(\sigma_1,\sigma_2,\sigma_3)$:

3	1	0	$\sigma_1^2\sigma_2$
2	2	0	σ_2^2
2	1	1	$\sigma_1\sigma_3$

.

Будем искать представление многочлена в следующем виде:

$$f(x_1,x_2,x_3) = -\sigma_1^2\sigma_2 + a\sigma_2^2 + b\sigma_1\sigma_3, \quad a,b \in \mathbb{Q}.$$

Далее:

x_1	x_2	x_3	$f(x_1,x_2,x_3)$	σ_1	σ_2	σ_3
1	−1	0	2	0	−1	0
1	1	1	−3	3	3	1

;

$$\begin{cases} 2 = a, \\ -3 = -27 + 18 + 3b. \end{cases}$$

Итак, $a = 2$, $b = 2$, то есть

$$f(x_1,x_2,x_3) = -\sigma_1^2\sigma_2 + 2\sigma_2^2 + 2\sigma_1\sigma_3.$$

Пример 4.10.21. Симметрический многочлен

$$f(x_1,x_2,\ldots,x_n) = S(x_1^2x_2) + S(2x_1^2)$$

является суммой двух однородных компонент $g_1 = S(x_1^2x_2)$ и $g_2 = S(2x_1^2)$.

Случай 1, $g_1 = S(x_1^2x_2)$:

2	1	0	0	...	0	$\sigma_1\sigma_2$
1	1	1	0	...	0	σ_3

;

$g_1(x_1,x_2,\ldots,x_n) = \sigma_1\sigma_2 + a\sigma_3$;

x_1	x_2	x_3	x_4	...	x_n	g_1	σ_1	σ_2	σ_3
1	1	1	0	...	0		3	3	1

;

$6 = 3\cdot 3 + a$, $a = -3$; $g_1 = S(x_1^2x_2) = \sigma_1\sigma_2 - 3\sigma_3$.

Случай 2, $g_2 = S(2x_1^2) = 2\sigma_1^2 - 4\sigma_2$. Итак,

$$f(x_1,x_2,\ldots,x_n) = \sigma_1\sigma_2 - 3\sigma_3 + 2\sigma_1^2 - 4\sigma_2.$$

4) *Вычисление значений любого симметрического многочлена* $f(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n]$ *от* n *переменных от корней* $c_1, c_2, \ldots, c_n$ *многочлена* $g(x) = a_n x^n + \ldots + a_1 x + a_0 = a_n(x - c_1) \cdots (x - c_n) \in K[x]$ n*-й степени от одной переменной.*

Пусть $F(X_1, X_2, \ldots, X_n) \in K[X_1, X_2, \ldots, X_n]$ — такой многочлен, что

$$f(x_1, x_2, \ldots, x_n) = F(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

В силу теоремы Виета

$$\sigma_1(c_1, \ldots, c_n) = -\frac{a_{n-1}}{a_n}, \ \sigma_2(c_1, \ldots, c_n) = \frac{a_{n-2}}{a_n}, \ldots, \ \sigma_n(c_1, \ldots, c_n) = (-1)^n \frac{a_0}{a_n},$$

поэтому

$$f(c_1, c_2, \ldots, c_n) = F\left(-\frac{a_{n-1}}{a_n}, \frac{a_{n-2}}{a_n}, \ldots, (-1)\frac{a_0}{a_n}\right).$$

Пример 4.10.22. Найти сумму кубов $c_1^3 + c_2^3 + c_3^3 + c_4^3$ корней c_1, c_2, c_3, c_4 многочлена $g(x) = 7x^4 - 14x^2 - 7x + 2$.

Так как для

$$g(x) = 7x^4 - 14x^2 - 7x + 2 = 7(x - c_1)(x - c_2)(x - c_3)(x - c_4)$$

имеем $\sigma_1(c_1, c_2, c_3, c_4) = 2$, $\sigma_2(c_1, c_2, c_3, c_4) = 0$, $\sigma_3(c_1, c_2, c_3, c_4) = 1$ (формулы Виета) и

$$f(x_1, x_2, x_3, x_4) = x_1^3 + x_2^3 + x_3^3 + x_4^3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3,$$

то

$$f(c_1, c_2, c_3, c_4) = c_1^3 + c_2^3 + c_3^3 + c_4^3 = 8 - 0 + 3 = 11.$$

Пример 4.10.23. *Кубическая резольвента уравнения* $x^4 + px^2 + qx + r = 0$: надо найти коэффициенты уравнения третьей степени

$$y^3 + a_1 y^2 + a_2 y + a_3 = 0,$$

корнями которого являются $d_1 = c_1c_2 + c_3c_4$, $d_2 = c_1c_3 + c_2c_4$, $d_3 = c_1c_4 + c_2c_3$, где c_1, c_2, c_3, c_4 — корни исходного уравнения.

По формулам Виета:

$$a_1 = -(d_1 + d_2 + d_3), \quad a_2 = d_1d_2 + d_1d_3 + d_2d_3, \quad a_3 = -d_1d_2d_3;$$
$$d_1 + d_2 + d_3 = \sigma_2, \quad d_1d_2 + d_1d_3 + d_2d_3 = \sigma_1\sigma_3 - 4\sigma_4, \quad d_1d_2d_3 = \sigma_1^2\sigma_4$$

здесь $\sigma_i = \sigma_i(c_1, c_2, c_3, c_4)$;

$$\sigma_1 = 0, \quad \sigma_2 = p, \quad \sigma_3 = -q, \quad \sigma_4 = r.$$

Поэтому

$$a_1 = -p, \quad a_2 = -4r, \quad a_3 = 4pr - q^2.$$

Итак, кубическая резольвента имеет вид:

$$y^3 - py^2 - 4ry + (4pr - q^2) = 0.$$

Упражнение 4.10.24.

$$(c_1+c_2-c_3-c_4)^2=4(d_1-p);$$
$$(c_1-c_2+c_3-c_4)^2=4(d_2-p);$$
$$(c_1-c_2-c_3+c_4)^2=4(d_3-p);$$
$$(c_1+c_2-c_3-c_4)(c_1-c_2+c_3-c_4)(c_1-c_2-c_3+c_4)=-8q.$$

Таким образом, если $\operatorname{char} K \neq 2$, то

$$c_{1,2,3,4}=\frac{1}{2}\left(\pm\sqrt{d_1-p}\pm\sqrt{d_2-p}\pm\sqrt{d_3-p}\right),$$

где число минусов должны быть чётно, а исходные значения квадратных корней надо выбирать так, чтобы их произведение равнялось $-q$.

Итак, решение уравнения четвёртой степени

$$x^4+px^2+qx+r=0$$

сведено к решению уравнения третьей степени (его кубической резольвенты)

$$y^3-py^2-4ry+4(pr-q^2)=0.$$

5) Если

$$f(x)=a_nx^n+\ldots+a_1x+a_0=a_n(x-c_1)\ldots(x-c_n)\in K[x],\quad a_n\neq 0,\quad a_0\neq 0$$

(то есть $f(0)=a_0\neq 0$, что равносильно условию $c_j\neq 0$ для всех $j=1,\ldots,n$), то

$$\sum_{j=1}^{n}\frac{1}{c_j}=-\frac{a_1}{a_0}.$$

Доказательство. Так как

$$\sigma_n(c_1,\ldots,c_n)=c_1\ldots c_n=(-1)^n\frac{a_0}{a_n},$$
$$\sigma_{n-1}(c_1,\ldots,c_n)=\sum_{j=1}^{n}c_1\cdots\widehat{c_j}\cdots c_n=(-1)^{n-1}\frac{a_1}{a_n},$$

то

$$\sum_{j=1}^{n}\frac{1}{c_j}=\sum_{j=1}^{n}\frac{c_1\cdots\widehat{c_j}\cdots c_n}{c_1\cdots c_n}=\frac{\sigma_{n-1}(c_1,\ldots,c_n)}{\sigma_n(c_1,\ldots,c_n)}=\frac{(-1)^{n-1}\frac{a_1}{a_n}}{(-1)^n\frac{a_0}{a_n}}=-\frac{a_1}{a_0}.\qquad\square$$

4.11. Симметрические рациональные дроби

Назовём ненулевую рациональную дробь

$$\frac{f(x_1,x_2,\ldots,x_n)}{g(x_1,x_2,\ldots,x_n)}\in K(x_1,x_2,\ldots,x_n),$$

Конец лекции №21

Лекция N 22 (29.XI.2011)

$$f = f(x_1, x_2, \ldots, x_n), g = g(x_1, x_2, \ldots, x_n) \in K[x_1, x_2, \ldots, x_n], \quad g \neq 0,$$

симметрической, если для любой подстановки $\sigma \in \mathrm{S}_n$

$$\frac{f^\sigma}{g^\sigma} = \frac{f}{g},$$

где $f^\sigma = f^\sigma(x_1, x_2, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$, то есть

$$f^\sigma g = f g^\sigma.$$

Это определение корректно: если $\dfrac{f}{g} = \dfrac{\varphi}{\psi}$, $\psi \neq 0$, то $f\psi = g\varphi$, и поэтому $f^\sigma \psi^\sigma = g^\sigma \varphi^\sigma$. Если $f = 0$, то $f^\sigma = 0$, и поэтому, так как $g^\sigma \neq 0$, имеем $\varphi^\sigma = 0$, то есть $\varphi = 0$. Если $f \neq 0$, то, умножая равенство на f, получаем $f f^\sigma \psi^\sigma = f g^\sigma \varphi^\sigma = f^\sigma g \varphi^\sigma$. Сокращая на $0 \neq f^\sigma \in K[x_1, x_2, \ldots, x_n]$ (в кольце без делителей нуля), получаем: $\dfrac{\varphi^\sigma}{\psi^\sigma} = \dfrac{f}{g} = \dfrac{\varphi}{\psi}$.

Теорема 4.11.1. *Всякая симметрическая рациональная дробь* $\dfrac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)} \in$ $\in K(x_1, x_2, \ldots, x_n)$ *от* n *переменных над полем* K *равна рациональной дроби вида* $\dfrac{\Phi(\sigma_1, \sigma_2, \ldots, \sigma_n)}{\Psi(\sigma_1, \sigma_2, \ldots, \sigma_n)}$, *где* $\Phi(X_1, X_2, \ldots, X_n), \Psi(X_1, X_2, \ldots, X_n) \in K[X_1, X_2, \ldots, X_n]$, $\Psi \neq 0$.

Доказательство. Многочлен

$$G = G(x_1, x_2, \ldots, x_n) = \prod_{\sigma \in \mathrm{S}_n} g^\sigma, \quad g^\sigma(x_1, x_2, \ldots, x_n) = g(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}),$$

является симметрическим, и поэтому многочлен $\dfrac{Gf}{g} \in K[x_1, x_2, \ldots, x_n]$ также является симметрическим. Поэтому $\dfrac{f}{g} = \left(\dfrac{Gf}{g}\right) \Big/ G$. Остаётся применить к симметрическим многочленам $\dfrac{Gf}{g}$ и G основную теорему. □

Упражнение 4.11.2. Если несократимая рациональная дробь $\dfrac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)}$ симметрическая, то многочлены $f(x_1, x_2, \ldots, x_n)$, $g(x_1, x_2, \ldots, x_n)$ симметрические. Это позволяет дать другое доказательство теоремы о симметрических рациональных дробях.

4.12. Дискриминант

Напоминание из школьной программы

Для квадратного уравнения

$$ax^2 + bx + c = 0, \quad a \neq 0, \quad a, b, c \in \mathbb{R},$$

выделяя полный квадрат

$$a\left(x^2 + 2\frac{b}{2a}x + \frac{b^2}{4a^2}\right) = \frac{b^2}{4a} - c,$$

$$\left(x+\frac{b}{2a}\right)^2=\frac{b^2-4ac}{4a^2},$$

получаем формулу для корней

$$x_{1,2}=-\frac{b}{2a}\pm\frac{\sqrt{b^2-4ac}}{2a}.$$

Выражение $D=b^2-4ac$ называется *дискриминантом квадратного трёхчлена* ax^2+bx+c.

Замечание 4.12.1.

$$D=b^2-4ac=a^2\left(\left(\frac{b}{a}\right)^2-4\frac{c}{a}\right)=a^2\big((x_1+x_2)^2-4x_1x_2\big)=$$
$$=a^2(x_2-x_1)^2=a^2\big(\sigma_1^2(x_1,x_2)-4\sigma_2(x_1,x_2)\big),$$

поскольку

$$\frac{b}{a}=-(x_1+x_2),\quad \frac{c}{a}=x_1x_2.$$

Если $a=1$, то $D=b^2-4ac=(x_2-x_1)^2=\sigma_1^2-4\sigma_2$.

Определитель Вандермонда и симметрические многочлены

Пусть K — поле, $K[x_1,x_2,\dots,x_n]$ — кольцо многочленов от переменных $x_1,\dots,x_n$. Рассмотрим определитель Вандермонда

$$\Delta_n=\begin{vmatrix}1 & 1 & \dots & 1\\ x_1 & x_2 & \dots & x_n\\ \dots & \dots & \dots & \dots\\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1}\end{vmatrix}=\prod_{1\leqslant j<i<n}(x_i-x_j)\in\mathbb{Z}[x_1,\dots,x_n].$$

Этот многочлен $\Delta_n(x_1,\dots,x_n)$ не является симметрическим (при подстановке $x_i\leftrightarrow x_j$, $i\neq j$, он меняет знак). Однако его квадрат $\Delta_n^2=\prod\limits_{i<j}(x_i-x_j)^2\in\mathbb{Z}[x_1,\dots,x_n]$ уже является симметрическим многочленом. Поэтому *существует и единственный* многочлен, называемый *дискриминантом*,

$$\mathrm{Dis}(X_1,\dots,X_n)\in\mathbb{Z}[X_1,\dots,X_n],$$

такой, что

$$\Delta_n^2(x_1,\dots,x_n)=\mathrm{Dis}(\sigma_1,\dots,\sigma_n).$$

Замечание 4.12.2. Так как

$$\begin{vmatrix}1 & 1 & \dots & 1\\ x_1 & x_2 & \dots & x_n\\ \dots & \dots & \dots & \dots\\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1}\end{vmatrix}=\begin{vmatrix}1 & x_1 & \dots & x_1^{n-1}\\ 1 & x_2 & \dots & x_2^{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_n & \dots & x_n^{n-1}\end{vmatrix},$$

то

$$\Delta_n^2(x_1,\ldots,x_n)=\begin{vmatrix}1 & 1 & \ldots & 1\\ x_1 & x_2 & \ldots & x_n\\ \ldots & \ldots & \ldots & \ldots\\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1}\end{vmatrix}\begin{vmatrix}1 & x_1 & \ldots & x_1^{n-1}\\ 1 & x_2 & \ldots & x_2^{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_n & \ldots & x_n^{n-1}\end{vmatrix}=\begin{vmatrix}n & s_1 & s_2 & \ldots & s_{n-1}\\ s_1 & s_2 & s_3 & \ldots & s_n\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ s_{n-1} & s_n & s_{n+1} & \ldots & s_{2n-2}\end{vmatrix},$$

где $s_k(x_1,\ldots,x_n)=x_1^k+\ldots+x_n^k$.

Следствие 4.12.3 ($n=2$).

$$\Delta_2^2(x_1,x_2)=\begin{vmatrix}2 & s_1\\ s_1 & s_2\end{vmatrix}=\begin{vmatrix}2 & \sigma_1\\ \sigma_1 & \sigma_1^2-2\sigma_2\end{vmatrix}=2(\sigma_1^2-2\sigma_2)-\sigma_1^2=\sigma_1^2-4\sigma_2.$$

Итак, $\mathrm{Dis}(\sigma_1,\sigma_2)=\sigma_1^2-4\sigma_2$ *(как и в случае дискриминанта унитарного квадратного трёхчлена).*

Упражнение 4.12.4 ($n=3$).

а) $s_1=\sigma_1\sigma_2$, $s_2=\sigma_1^2-2\sigma_2$, $s_3=\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3$, $s_4=\sigma_1^4-4\sigma_1^2\sigma_2+4\sigma_1\sigma_3+2\sigma_2^2-4\sigma_4$,

$$\Delta_3^2(x_1,x_2,x_3)=\begin{vmatrix}3 & s_1 & s_2\\ s_1 & s_2 & s_3\\ s_2 & s_3 & s_4\end{vmatrix}=$$
$$=3s_2s_4+2s_1s_2s_3-s_2^3-s_1^2s_4-3s_3^2=$$
$$=\sigma_1^2\sigma_2^2-4\sigma_2^3-4\sigma_1^3\sigma_3-27\sigma_3^2+18\sigma_1\sigma_2\sigma_3.$$

б)

$$\Delta_3^2(x_1,x_2,x_3)=(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2=$$
$$=\sigma_1^2\sigma_2^2-4\sigma_2^3-4\sigma_1^3\sigma_3-27\sigma_3^2+18\sigma_1\sigma_2\sigma_3=\mathrm{Dis}_3(\sigma_1,\sigma_2,\sigma_3).$$

в) Попробуйте вычислить $\Delta_4^2(x_1,x_2,x_3,x_4)=\mathrm{Dis}(\sigma_1,\sigma_2,\sigma_3,\sigma_4)$.

Дискриминант многочлена и кратные корни

Для унитарного многочлена $f(x)\in K[x]$,

$$f(x)=x^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0=(x-x_1)\cdots(x-x_n)=$$
$$=x^n-\sigma_1(x_1,\ldots,x_n)x^{n-1}+\ldots+(-1)^n\sigma_n(x_1,\ldots,x_n),$$

где $x_1,x_2,\ldots,x_n$ — корни многочлена $f(x)$, $n\geqslant 2$, *дискриминант многочлена* $f(x)$ определяется следующим образом:

$$D(f(x))=\prod_{j<i}(x_i-x_j)^2\left(=\pm\prod_{i\neq j}(x_i-x_j)\right)=\mathrm{Dis}(\sigma_1,\ldots,\sigma_n)$$

(последнее равенство позволяет вычислять $D(f)$ через коэффициенты многочлена $f(x)$, не зная его корней).

Теорема 4.12.5. *$D(f(x)) = 0$ тогда и только тогда, когда $f(x)$ имеет кратные корни.*

Доказательство. $D(f(x)) = \prod_{j<i}(x_i - x_j)^2 = 0$ тогда и только тогда, когда $x_i = x_j$ при $i \neq j$, т. е. когда $f(x)$ имеет кратные корни. □

Замечание 4.12.6. Ранее другой критерий наличия кратных корней для поля K, $\operatorname{char} K = 0$, был связан с НОД$(f(x), f'(x))$.

Замечание 4.12.7. Для многочлена

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = a_n(x - x_1)\cdots(x - x_n) \in K[x]$$

с $a_n \neq 1$ *дискриминант* определяется как

$$D(f) = a_n^{2n-2}\prod_{i>j}(x_i - x_j)^2 = a_n^{2n-2}\operatorname{Dis}(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

(появление множителя a_n^{2n-2} в определении $D(f)$ будет ясно из примеров).

Таким образом: если

$$f(x) = a_2x^2 + a_1x + a_0, \quad a_2 \neq 0,$$

то

$$D(f) = a_2^2(x_1 - x_2)^2 = a_2^2(\sigma_1^2 - 4\sigma_2) = \frac{a_2^2(a_1^2 - 4a_2a_0)}{a_2^2} = a_1^2 - 4a_2a_0;$$

если

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0, \quad a_3 \neq 0,$$

то

$$\begin{aligned} D(f) &= a_3^4 \prod_{1\leqslant j<i\leqslant 3}(x_i - x_j)^2 = a_3^4\operatorname{Dis}_3(\sigma_1, \sigma_2, \sigma_3) = \\ &= a_3^4(\sigma_1^2\sigma_2^2 - 4\sigma_2^3 - 4\sigma_1^3\sigma_3 - 27\sigma_3^2 + 18\sigma_1\sigma_2\sigma_3) = \\ &= a_2^2a_1^2 - 4a_2^3a_0 - 4a_3a_1^3 - 27a_3^2a_0^2 + 18a_3a_2a_1a_0, \end{aligned}$$

поскольку

$$\sigma_1 = -\frac{a_2}{a_3}, \quad \sigma_2 = \frac{a_1}{a_3}, \quad \sigma_3 = -\frac{a_0}{a_3}.$$

Для неполного кубического многочлена $f(x) = x^3 + px + q$ имеем $D(f) = -4p^3 - 27q^2$.

Упражнение 4.12.8. Пусть $p, q \in \mathbb{R}$, $f(x) = x^3 + px + q \in \mathbb{R}[x]$. Тогда:

1) если $D(f) > 0$, то все три корня действительны и различны;

2) если $D(f) = 0$, то все три корня действительны и хотя бы два из них совпадают;

3) если $D(f) < 0$, то все три корня различны, один из них действительный, два других не являются действительными и сопряжены.

Формулы решения уравнений третьей степени (формулы Кардано) см. упражнение ??.

Задача 4.12.9. Если $f(x) = x^4 + px^2 + qx + r$ (неполное уравнение четвёртой степени), то

$$D\big(f(x)\big) = 16p^4r - 4p^3q^2 - 128p^2r^2 + 144pq^2r - 27q^4 + 256r^3.$$

Указание. $D\big(f(x)\big) = D(x^3 - 2px^2 + (p^2 - 4r)x + q^2)$.

Задача 4.12.10. Если $f(x) \in \mathbb{Z}[x]$, то $D(f)$ либо делится на 4, либо имеет остаток 1 при делении на 4.

Задача 4.12.11. Пусть $f(x) \in \mathbb{R}[x]$, $n = \deg f(x) \geqslant 2$, $D(f) \neq 0$, r — число пар взаимно сопряжённых мнимых корней. Тогда знаки чисел $D(f)$ и $(-1)^r$ совпадают. В частности, если $f(x)$ не имеет действительных корней, то знак числа $D(f)$ равен $(-1)^{n/2}$.

Упражнения 4.12.12.

1) Пусть $f(x) = x^n - 1 = \prod\limits_{i=1}^{n}(x - \varepsilon_i)$. Тогда $\sigma_1 = \sigma_2 = \ldots = \sigma_{n-1} = 0$, $\sigma_n = (-1)^{n-1}$, и в силу формул Ньютона $s_1 = s_2 = \ldots = s_{n-1} = 0$, $s_n = n$. Тогда

$$D\big(f(x)\big) = \mathrm{Dis}(\sigma_1, \ldots, \sigma_n) = \Delta_n^2(x_1, \ldots, x_n) =$$

$$= \begin{vmatrix} n & s_1 & s_2 & \ldots & s_{n-1} \\ s_1 & s_2 & s_3 & \ldots & s_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{n-1} & s_n & s_{n+1} & \ldots & s_{2n-2} \end{vmatrix} = n \begin{vmatrix} 0 & \ldots & n \\ \vdots & \iddots & \\ n & & * \end{vmatrix} =$$

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

2) $D(x^n + a) = (-1)^{\frac{n(n-1)}{2}} n^n a^{n-1}$.

3) $D(x^{n-1} + x^{n-2} + \ldots + 1) = (-1)^{\frac{(n-1)(n-2)}{2}} n^{n-2}$.

4) $D\left(1 + x + \dfrac{x^2}{2!} + \ldots + \dfrac{x^n}{n!}\right) = (-1)^{\frac{n(n-1)}{2}} (n!)^{-n+2}$.

5) Трудная задача: вычислить $D(\Phi_n)$, где Φ_n — n-й многочлен деления круга.

Упражнения 4.12.13.

1) Пусть $f = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$. Покажите, что

$$na_n + (n-1)a_{n-1}\frac{\partial D}{\partial a_{n-2}} + \ldots + a_1\frac{\partial D}{\partial a_0} = 0,$$

где $D = D(a_0, \ldots, a_n) = D(f)$ — дискриминант многочлена f.

2) Пусть K — поле, $\mathrm{char}\, K = 0$ и $f(x) = \prod\limits_{i=1}^{n}(x - \alpha_i)$. Тогда

$$D(f) = (-1)^{\frac{n(n-1)}{2}} \prod_{i=1}^{n} f'(\alpha_i).$$

3) Для многочлена $f = a_m x^m + \ldots + a_0 \in \mathbb{Z}[x]$, $a_m \neq 0$, положим

$$\|f\|_2 = \sqrt{\sum_{i=0}^{m} a_i^2}.$$

Докажите, что если $f, g \in \mathbb{Z}[x]$, $m = \deg(f)$, $n = \deg(g)$, то

$$|R(f,g)| \leqslant \|f\|_2^m \cdot \|g\|_2^n.$$

Задача 4.12.14. Вычислить дискриминанты:

а) многочленов Эрмита

$$f_n(x) = (-1)^n e^{x^2/2} \frac{d^n(e^{-x^2/2})}{dx^n};$$

б) многочленов Лагерра

$$g_n(x) = (-1)^n e^x \frac{d^n(x^n e^{-x})}{dx^n};$$

в) многочленов Чебышёва

$$h_n(x) = 2\cos\left(n \cdot \arccos\left(\frac{x}{2}\right)\right).$$

Ответы. а) $1 \cdot 2^2 \cdot 3^3 \cdot \ldots \cdot (n-1)^{n-1} \cdot n^n$; б) $1 \cdot 2^2 \cdot 3^5 \cdot \ldots \cdot n^{2n-1}$; в) $2^{n-1} \cdot n^n$.

4.13. Результант двух многочленов от одной переменной над полем

Пусть K — поле, $f, g \in K[x]$. Алгоритм Евклида позволяет конструктивно найти наибольший общий делитель $d(x)$ многочленов $f(x)$ и $g(x)$, что позволяет ответить на вопрос: имеют ли многочлены $f(x)$ и $g(x)$ общий множитель, отличный от константы ($\deg d(x) \geqslant 1$). Нам хотелось бы получить критерий в терминах коэффициентов многочленов $f(x)$ и $g(x)$, с этой целью будет введён результант $R(f,g)$ многочленов f и g.

Лемма 4.13.1. *Пусть K — поле, $f, g \in K[x]$, $f \neq 0$, $g \neq 0$,*

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \ldots + a_0,$$
$$g(x) = b_n x^n + b_{n-1}x^{n-1} + \ldots + b_0,$$

(возможно, $a_m = 0$ или $b_n = 0$). Тогда многочлены $f(x)$ и $g(x)$ имеют общий множитель, отличный от константы, или $a_m = 0 = b_n$ тогда и только тогда, когда существуют многочлены $f_1, g_1 \in K[x]$ такие, что

$$fg_1 = gf_1, \quad f_1 \neq 0, \quad g_1 \neq 0, \quad \deg f_1 < m, \quad \deg g_1 < n. \tag{4.4}$$

Доказательство. 1а) Пусть $d = \text{НОД}(f, g)$, $\deg d > 0$, тогда $f = df_1$, $g = dg_1$, $f_1 \neq 0$, $g_1 \neq 0$, $\deg f_1 < m$, $\deg g_1 < n$, т. е. выполнено условие (4.4).

1б) Если $a_m = 0 = b_n$, то (4.4) имеет место (даже если $\deg d = 0$): $fg_1 = gf_1$, $\deg f_1 \leqslant \deg f < m$, $\deg g_1 \leqslant \deg g < n$.

2) Пусть выполнено условие (4.4) и $d = (f, g) = 1$. Тогда f делит f_1, g делит g_1, и поэтому $\deg f \leqslant \deg f_1 < m$, $\deg g \leqslant \deg g_1 < n$, следовательно, $a_m = 0 = b_n$. □

Замечание 4.13.2. В условии (4.4)

$$fg_1 = gf_1, \quad \deg f_1 < m, \quad \deg g_1 < n.$$

Поэтому $\deg g_1 \leqslant n-1$, и следовательно, $\deg fg_1 \leqslant m+n-1$; $\deg f_1 \leqslant m-1$, и следовательно, $\deg gf_1 \leqslant m+n-1$. Считая коэффициенты $u_{m-1}, \ldots, u_1, u_0, v_{n-1}, \ldots, v_1, v_0$ многочленов

$$f_1 = u_{m-1}x^{m-1} + \ldots + u_1x + u_0,$$
$$g_1 = v_{n-1}x^{n-1} + \ldots + v_1x + v_0$$

неизвестными и приравнивая коэффициенты при различных степенях x^k, $0 \leqslant k \leqslant m+n-1$, в равенстве

$$(a_mx^m + \ldots + a_0)(v_{n-1}x^{n-1} + \ldots + v_0) = fg_1 =$$
$$= gf_1 = (b_nx^n + \ldots + b_0)(u_{m-1}x^{m-1} + \ldots + u_0),$$

получаем квадратную однородную систему $m+n$ линейных уравнений от $m+n$ неизвестных $u_{m-1}, \ldots, u_0, v_{n-1}, \ldots, v_0$:

$$\begin{aligned} a_mv_{n-1} &= b_nu_{m-1}, \\ a_{m-1}v_{n-1} + a_mv_{n-2} &= b_{n-1}u_{m-1} + b_nu_{m-2}, \\ &\ldots \\ a_1v_0 + a_0v_1 &= b_1u_0 + b_0u_1, \\ a_0v_0 &= b_0u_0. \end{aligned}$$

Перенося правую часть равенств налево и рассматривая транспонированный определитель системы (от переменных $v_{n-1}, \ldots, v_0, -u_{m-1}, \ldots, -u_0$, чтобы не иметь знаков $-$ в определителе)

$$R(f,g) = |\mathcal{S}(a_m, \ldots, a_0, b_n, \ldots, b_0)| = \left|\begin{matrix} a_m & a_{m-1} & \ldots & a_0 & & & \\ & a_m & a_{m-1} & \ldots & a_0 & & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ & & & a_m & a_{m-1} & \ldots & a_0 \\ b_n & b_{n-1} & \ldots & b_0 & & & \\ & b_n & b_{n-1} & \ldots & b_0 & & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ & & & b_n & b_{n-1} & \ldots & b_0 \end{matrix}\right| \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \end{matrix} \tag{4.5}$$

(называемый *результантом* многочленов $f(x)$ и $g(x)$; всюду, где ничего не написано, подразумеваются нули), получаем, что наша система линейных уравнений имеет ненулевое решение

$$(v_{n-1}, \ldots, v_0, u_{m-1}, \ldots, u_0)$$

тогда и только тогда, когда $R(f,g)=0$.

Приведённый метод исключения восходит к Эйлеру, вид результанта $R(f,g)$ в (4.5) восходит к Сильвестру.

Выведем теперь основные свойства результанта $R(f,g)$.

Теорема 4.13.3. *Пусть K — поле, $f(x), g(x) \in K[x]$, $f \neq 0$, $g \neq 0$,*

$$f(x) = a_m x^m + \ldots + a_0,$$
$$g(x) = b_n x^n + \ldots + b_0.$$

Тогда результант $R(f,g)$ многочленов $f(x)$ и $g(x)$ является многочленом с целыми коэффициентами от переменных $a_m, \ldots, a_0, b_n, \ldots, b_0$,

$$R(f,g) \in \mathbb{Z}[a_m, \ldots, a_0, b_n, \ldots, b_0],$$

обладающим следующими свойствами:

1) $R(c,g) = c^n$ *для* $c \in K$; $R(f,f) = 0$; $R(f,g) = (-1)^{mn} R(g,f)$; $R(cf,g) = c^n R(f,g)$, $R(x^k f, g) = b_0^k R(f,g)$ *при* $k > 0$;

2) $R(f,g) = 0$ *тогда и только тогда, когда* $f(x)$ *и* $g(x)$ *имеют общий множитель, отличный от константы, или* $a_m = 0 = b_n$;

3) *многочлен* $R(f,g)$ *однороден степени* n *от переменных* $a_m, \ldots, a_0$ *и однороден степени* m *от переменных* $b_n, \ldots, b_0$ *(таким образом, многочлен* $R(f,g)$ *является однородным многочленом степени* $m+n$ *от переменных* $a_m, \ldots, a_0, b_n, \ldots, b_0$, *при этом изобаричен веса* mn*);*

4) *существуют многочлены* $F(x), G(x) \in K[x]$, $\deg F < m$, $\deg G < n$, *такие что*

$$R(f,g) = Gf + Fg;$$

5) *если многочлены* $f(x)$ *и* $g(x)$ *разлагаются на линейные множители,*

$$f(x) = a_m \prod_{i=1}^{m} (x - \alpha_i), \quad g(x) = b_n \prod_{j=1}^{n} (x - \beta_j),$$

то

$$R(f,g) = a_m^n \prod_{i=1}^{m} g(\alpha_i) = (-1)^{mn} b_n^m \prod_{j=1}^{n} f(\beta_j) = a_m^n b_n^m \prod_{i,j} (\alpha_i - \beta_j).$$

Доказательство. 1) Приведённые свойства вытекают непосредственно из свойств определителя.

2) следует из леммы 4.13.1 и замечания 4.13.2.

3) Если в (4.5) в определителе $R(f,g)$ многочлен f заменить на tf (каждый коэффициент a_i заменяется на ta_i, поэтому первые n строк определителя $R(f,g)$ умножаются каждая на t), то $R(tf,g) = t^n R(f,g)$, это показывает, что $R(f,g)$ является однородным многочленом степени n от переменных $a_m, \ldots, a_0$. Аналогично, $R(f,tg) = t^m R(f,g)$, и поэтому $R(f,g)$ является однородным многочленом степени m от переменных $b_n, \ldots, b_0$.

Вес члена определителя

$$M = \prod_{i=1}^{m+n} s_{i\sigma(i)}, \quad \mathcal{S}(a_m, \ldots, a_0, b_n, \ldots, b_0) = (s_{ij}),$$

равен

$$\omega(M) = \sum_{i=1}^{n}(\sigma(i) - i) + \sum_{i=n+1}^{m+n}(n + \sigma(i) - i) = \sum_{i=1}^{m+n}\sigma(i) - \sum_{i=1}^{m+n} i + \sum_{i=n+1}^{m+n} n = \sum_{i=n+1}^{m+n} n = mn,$$

поскольку $\sum\limits_{i=1}^{m+n}\sigma(i) = \sum\limits_{i=1}^{n} i$ для любой подстановки $\sigma \in \mathrm{S}_{m+n}$. Итак, $R(f,g)$ — изобарический многочлен веса mn от переменных $a_m, \ldots, a_0, b_n, \ldots, b_0$.

4) Так как

$$\mathcal{S}(a_m, \ldots, a_0, b_n, \ldots, b_0)\begin{pmatrix} x^{n+m-1} \\ x^{n+m-2} \\ \vdots \\ x^m \\ x^{m-1} \\ \vdots \\ x \\ 1 \end{pmatrix} = \begin{pmatrix} x^{n-1}f(x) \\ x^{n-2}f(x) \\ \vdots \\ f(x) \\ x^{m-1}g(x) \\ \vdots \\ xg(x) \\ g(x) \end{pmatrix},$$

то, умножая в этом равенстве каждую i-ю строку на алгебраическое дополнение c_i к элементу матрицы $\mathcal{S}$, стоящему на пересечении i-й строки и $(m+n)$-го столбца, $i = 1, \ldots, m+n$, и складывая, получаем

$$\begin{aligned} &c_1x^{n-1}f(x) + c_2x^{n-2}f(x) + \ldots + c_nf(x) + c_{n+1}x^{m-1}g(x) + c_{n+2}x^{m-2}g(x) + \ldots + c_{m+n}g(x) = \\ &= \sum_{j=1}^{m+n} c_1s_{1j}x^{m+n-j} + \sum_{j=1}^{m+n} c_2s_{2j}x^{m+n-j} + \ldots + \sum_{j=1}^{m+n} c_{m+n}s_{m+n,j}x^{m+n-j} = \\ &= \sum_{i=1}^{m+n}\sum_{j=1}^{m+n} c_is_{ij}x^{m+n-j} = \sum_{j=1}^{m+n}\left(\sum_{i=1}^{m+n} c_is_{ij}\right)x^{m+n-j} = R(f,g). \end{aligned}$$

Итак,

$$G(x)f(x) + F(x)g(x) = R(f,g),$$

где

$$\begin{aligned} G(x) &= c_1x^{n-1} + c_2x^{n-2} + \ldots + c_n, \\ F(x) &= c_{n+1}x^{m-1} + c_{n+2}x^{m-2} + \ldots + c_{m+n}. \end{aligned}$$

5) Так как

$$f(x) = a_m\prod_{i=1}^{m}(x - \alpha_i) = \sum_{i=1}^{m} a_ix^i, \quad g(x) = b_n\prod_{j=1}^{n}(x - \beta_j) = \sum_{j=1}^{n} b_jx^j,$$

то a_i и b_j являются симметрическими функциями от независимых переменных α_i и от β_j соответственно, умноженными на переменные a_m и b_n. Как мы показали, $R(f,g)$ является

однородным многочленом степени n от a_i и однородным многочленом степени m от b_j, и поэтому $R(f,g)$ равен элементу $a_m^n b_n^m$, умноженному на симметрическую функцию от переменных α_i и β_j. Если $\alpha_k = \beta_l$, то многочлены $f(x)$ и $g(x)$ имеют общий множитель $x - \alpha_k = x - \beta_l$, и поэтому $R(f,g) = 0$, следовательно, $R(f,g)$ делится на одночлен $\alpha_i - \beta_j$ для всех $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, и, таким образом, делится на произведение одночленов $\prod\limits_{i,j}(\alpha_i - \beta_j)$, поскольку сомножители взаимно просты. Так как $a_m^n b_n^m \prod\limits_{i,j}(\alpha_i - \beta_j)$ и $R(f,g)$ имеют как однородные многочлены одинаковую степень, они отличаются на константу. На самом деле они равны, поскольку они содержат член

$$a_m^n b_n^m (-1)^{mn} \left(\prod_{j=1}^{n} \beta_j\right)^m = a_m^n b_0^m$$

(здесь мы использовали равенство $b_0 = (-1)^n b_n \prod\limits_{j=1}^{n} \beta_j$). Итак,

$$R(f,g) = a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (\alpha_i - \beta_j).$$

Поскольку $g(\alpha_i) = b_n \prod\limits_j (\alpha_i - \beta_j)$,

$$R(f,g) = a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (\alpha_i - \beta_j) = a_m^n \prod_{i=1}^{m} \left(b_n \prod_{j=1}^{n} (\alpha_i - \beta_j)\right) = a_m^n \prod_{i=1}^{n} g(\alpha_i).$$

Так как $R(f,g) = (-1)^{mn} R(g,f)$, то

$$R(f,g) = (-1)^{mn} b_n^m \prod_{j=1}^{n} f(\beta_j). \qquad \square$$

Следствие 4.13.4. *Для любых многочленов $f(x), g(x), h(x) \in K[x]$ над полем K*

$$R(fg,h) = R(f,h)R(g,h);$$
$$R(f,gh) = R(f,g)R(f,h).$$

Доказательство. а) В алгебраически замкнутом расширении $\bar{K}$ поля K (см. ??)

$$f = a_m \prod_i (x - \alpha_i), \quad g = b_n \prod_j (x - \beta_j),$$

следовательно,

$$fg = a_m b_n \prod_i (x - \alpha_i) \prod_j (x - \beta_j),$$

поэтому в силу п. 4 нашей теоремы

$$R(fg,h) = (a_m b_n)^r \prod_i h(\alpha_i) \prod_j h(\beta_j) =$$
$$= \left(a_m^r \prod_i h(\alpha_i)\right)\left(b_n^r \prod_j h(\beta_j)\right) = R(f,h)R(g,h), \quad r = \deg(h).$$

б) $R(f,gh) = (-1)^{m(n+r)} R(gh,f) = (-1)^{m(n+r)} R(g,f)R(h,f) = (-1)^{mn} R(g,f)(-1)^{mr} R(h,f) = R(f,g)R(f,h)$. $\square$

Дискриминант многочлена как результант многочлена и его производной

Пусть K — поле, $\operatorname{char} K = 0$, $f(x) \in K[x]$ и $g(x) = f'(x)$ — производная многочлена

$$f(x) = a_m \prod_{i=1}^{m} (x - \alpha_i).$$

Тогда

$$f'(\alpha_i) = a_m \prod_{j, j \neq i} (\alpha_i - \alpha_j),$$

и поэтому

$$R(f, f') = a_m^{m-1} \prod_{i=1}^{m} a_m f'(\alpha_i) = a_m^{2m-1} \prod_{i=1}^{m} \Big(\prod_{j, j \neq i} (\alpha_i - \alpha_j) \Big) = a_m (-1)^{\frac{m(m-1)}{2}} D(f),$$

где $D(f) = a_m^{2m-2} \prod_{i>j} (\alpha_i - \alpha_j)^2$ — дискриминант многочлена $f(x)$. Эквивалентная форма:

$$D(f) = (-1)^{\frac{m(m-1)}{2}} a_m^{-1} R(f, f').$$

Из формулы умножения для результанта следует, что

$$D(fg) = D(f) D(g) R^2(f, g);$$
$$D(fgh) = D(f) D(g) D(h) R(f, g)^2 R(g, h)^2 R(h, f)^2.$$

Действительно,

$$\begin{aligned} &D(f) D(g) R(f, g)^2 = \\ &= \Big(a_m^{2m-2} \prod_{i>j} (\alpha_i - \alpha_j)^2 \Big) \Big(b_n^{2n-2} \prod_{k>l} (\beta_k - \beta_l)^2 \Big) \Big(a_m^n b_n^m \prod_{s,t} (\alpha_s - \beta_t) \Big)^2 = \\ &= a_m^{2m-2} b_n^{2n-2} a_m^{2n} b_n^{2m} \prod_{i>j} (\alpha_i - \alpha_j)^2 \prod_{k>l} (\beta_k - \beta_l)^2 \prod_{s>t} (\alpha_s - \beta_t)^2 = \\ &= (a_m b_n)^{2(m+n)-2} \prod_{u>v} (\gamma_u - \gamma_v)^2 = D(fg). \end{aligned}$$

Вторая формула следует из первой:

$$\begin{aligned} D(fgh) &= D\big((fg)h\big) = D(fg) D(h) R^2(fg, h) = \\ &= D(f) D(g) R^2(f, g) \cdot D(h) R^2(f, h) R^2(g, h) = \\ &= D(f) D(g) D(h) R^2(f, g) R^2(g, h) R^2(f, h). \end{aligned}$$

□

Упражнения 4.13.5.

1) $R\big(x - \alpha, f(x)\big) = f(\alpha)$.

2) $R(f, g)$ — неприводимый многочлен от коэффициентов многочленов f и g.

3) Пусть $r, s \in \mathbb{N}$, $d = \text{НОД}(r, s)$, $r_1 = \dfrac{r}{d}$, $s_1 = \dfrac{s}{d}$. Тогда:

$$R(x^r - a, x^s - b) = (-1)^s (a^{s_1} - b^{r_1})^d.$$

4) Пусть $n, k \in \mathbb{N}$, $n > k$, $d = \text{НОД}(n, k)$, $n_1 = \dfrac{n}{d}$, $k_1 = \dfrac{k}{d}$. Тогда:

$$D(x^n + ax^k + b) = (-1)^{\frac{n(n-1)}{2}} b^{k-1} (n^{n_1} b^{n_1 - k_1} + (-1)^{n_1+1} (n-k)^{n_1-k_1} k^{k_1} a^{n_1})^d.$$

Указание. $D(f) = \dfrac{(-1)^{\frac{n(n-1)}{2}}}{a^n} R(f, f')$, $R(f, x^m g) = \big(f(0)\big)^m R(f, g)$, и воспользуйтесь результатом задачи 3).

5) Пусть $f(x) = a_m x^m + \ldots + a_0$, $g(x) = a_{m-1} x^{m-1} + a_{m-2} x^{m-2} + \ldots + a_1$. Покажите, что $R\big(f(x), g(x)\big) \vdots\!= a_m^m \cdot a_0^{m-1}$.

Задача 4.13.6. Пусть $m, n \in \mathbb{N}$, $m \geqslant n$, Φ_m и Φ_n — многочлены деления круга (см. ??). Докажите, что

а) $R(\Phi_m, \Phi_n) = \begin{cases} 0 & \text{при } m = n, \\ p^{\varphi(n)} & \text{при } m = np^k,\ p \text{ — простое число, } k \in \mathbb{N}, \\ 1 & \text{в остальных случаях;} \end{cases}$

б) $D(\Phi_n) = (-1)^{\varphi(n)/2} \dfrac{n^{\varphi(n)}}{\prod\limits_{\substack{p \mid n \\ p \text{ простое}}} p^{\varphi(n)/(p-1)}}$.

Задача 4.13.7 (способ Безу нахождения результанта). Докажите, что результант многочленов

$$f(x) = a_m x^m + \ldots + a_0$$

и

$$g(x) = b_m x^m + \ldots + b_0$$

равен определителю, составленному из коэффициентов многочленов $\varphi_k(x)$ степени не выше $(m-1)$:

$$\varphi_k(x) = (a_m x^{k-1} + a_{m-1} x^{k-2} + \ldots + a_{m-k+1}) g(x) - \\ - (b_m x^{k-1} + b_{m-1} x^{k-2} + \ldots + b_{m-k+1}) f(x), \quad k = 1, \ldots, m.$$

При этом

$$\varphi_1(x) = a_m g(x) - b_m f(x),$$
$$\varphi_k(x) = x\varphi_{k-1}(x) + a_{m-k+1} g(x) - b_{m-k+1} f(x).$$

Если

$$\varphi_k(x) = c_{k1} + c_{k2} x + \ldots + c_{km} x^{m-1}, \quad 1 \leqslant k \leqslant m,$$

то

$$\begin{vmatrix} c_{11} & \ldots & c_{1m} \\ \vdots & \ldots & \vdots \\ c_{n1} & \ldots & c_{nm} \end{vmatrix} = R(f, g).$$

Задача 4.13.8 (способ Эрмита нахождения результанта). Докажите, что результант многочленов

$$f(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_0$$

и

$$g(x) = b_n x^n + \ldots + b_0$$

равен определителю, составленному из коэффициентов остатков при делении многочленов $g(x), xg(x), \ldots, x^{m-1}g(x)$ на $f(x)$. При этом остатки расположены в порядке возрастания степеней переменной x:

$$x^{k-1}g(x) = f(x)q_k(x) + r_k(x), \quad r_k(x) = c_{k1} + c_{k2}x + \ldots + c_{kn}x^{n-1};$$

$$\begin{vmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & \ldots & \vdots \\ c_{n1} & \ldots & c_{nn} \end{vmatrix} = R(f, g).$$

Например, пусть $f(x) = x^2 - x + 1$, $g(x) = x^4 - 2x^2 + 3$. Тогда: $g(x) = x^4 - 2x^2 + 3 = f(x) \cdot q_1(x) + r_1(x)$, где $q_1(x) = x^2 + x - 2$, $r_1(x) = 5 - 3x$; $x \cdot g(x) = x^5 - 2x^3 + 3x = f(x) \cdot q_2(x) + r_2(x)$, где $q_2(x) = x^3 + x^2 - 2x - 3$, $r_2(x) = 3 + 2x$. Поэтому

$$R(f, g) = \begin{vmatrix} 5 & -3 \\ 3 & 2 \end{vmatrix} = 19.$$

Замечание 4.13.9. Результанты играют важную роль в теории символьного интегрирования. Например, пусть K — поле, $f, g \in K[x]$, $\deg(g) > 0$, $\deg(f) < \deg(g)$, многочлен g не имеет кратных корней (в алгебраическом замыкании $\bar{K}$ поля K), НОД$(f, g) = 1$, $r(t) = R_x(g, f - tg') \in K[t]$. Тогда

$$\int \frac{f}{g}\,dx = \sum_{a, r(a)=0} a \ln\bigl(\text{НОД}(g, f - ag')\bigr)$$

(алгоритм Ротстейна—Трегера).

Пример 4.13.10. $f = x^4 - 3x^2 + 6$, $g = x^6 - 5x^4 + 5x^2 + 4 \in \mathbb{Q}[x]$, $r(t) = R_x\bigl(x^6 - 5x^4 + 5x^2 + 4, x^4 - 3x^2 + 6 - t(6x^5 - 20x^3 + 10x)\bigr) = 45796(4t^2 + 1)^3$. Пусть $a \in \bar{\mathbb{Q}} \subseteq \mathbb{C}$ и $4a^2 + 1 = 0$. Тогда

$$\text{НОД}\bigl(x^6 - 5x^4 + 5x^2 + 4, x^4 - 3x^2 + 6 - a(6x^5 - 20x^3 + 10x)\bigr) = x^3 + 2ax^2 - 3x - 4a,$$

$$\int \frac{x^4 - 3x^2 + 6}{x^6 - 5x^4 + 5x^2 + 4}\,dx = \sum_{a, 4a^2+1=0} a \ln(x^3 + 2ax^2 - 3x - 4a) =$$

$$= \frac{i}{2}\ln(x^3 + ix^2 - 3x - 2i) - \frac{i}{2}\ln(x^3 - ix^2 - 3x + 2i).$$

4.14. Результант над коммутативным кольцом

Матрица и определитель Сильвестра

Пусть $u_m, \ldots, u_1, u_0, v_n, \ldots, v_1, v_0$ — множество $m+n+2$ переменных над кольцом целых чисел $\mathbb{Z}$, при этом $m+n \geqslant 1$. Рассмотрим следующую $((m+n), (m+n))$-матрицу Сильвестра $\mathcal{S} = \mathcal{S}(u_m, \ldots, u_0; v_n, \ldots, v_0) \in \mathrm{M}_{m+n}(\mathbb{Z}[u_m, \ldots, u_0, v_n, \ldots, v_0])$:

$$\begin{pmatrix} u_m & u_{m-1} & u_{m-2} & \ldots & u_0 & 0 & \ldots & 0 \\ 0 & u_m & u_{m-1} & \ldots & u_1 & u_0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots & \ldots & \ldots & u_0 \\ v_n & v_{n-1} & v_{n-2} & \ldots & v_0 & 0 & \ldots & 0 \\ 0 & v_n & v_{n-1} & \ldots & v_1 & v_0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & v_n & v_{n-1} & \ldots & v_0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \end{matrix}$$

На этой картинке в первых n строках расположены последовательные сдвиги на один шаг вправо отрезка $[u_m, u_{m-1}, \ldots, u_1, u_0]$, а в следующих m строках — соответственно последовательные сдвиги на один шаг вправо отрезка $[v_n, v_{n-1}, \ldots, v_1, v_0]$. Таким образом, $\mathcal{S} = (s_{ij})$, где

$$s_{ij} = \begin{cases} u_{m-j+i} & \text{для } i = 1, \ldots, n, \\ v_{-j+i} & \text{для } i = n+1, \ldots, m+n \end{cases}$$

(здесь: $u_k = 0$ при $k < 0$ и при $k > m$; $v_k = 0$ при $k < 0$ и при $k > n$). В частности, на диагонали матрицы $\mathcal{S}$ стоят элементы

$$s_{ii} = \begin{cases} u_m, & 1 \leqslant i \leqslant n, \\ v_0, & n+1 \leqslant i \leqslant m+n. \end{cases}$$

Примеры 4.14.1.

1) Если $m = 0$, $n \geqslant 1$, то

$$\mathcal{S} = \begin{pmatrix} u_0 & & 0 \\ & \ddots & \\ 0 & & u_0 \end{pmatrix} = d(u_0, \ldots, u_0) \in \mathrm{M}_n(R[u_0, v_n, \ldots, v_0]);$$

если $m \geqslant 1$, $n = 0$, то

$$\mathcal{S} = \begin{pmatrix} v_0 & & 0 \\ & \ddots & \\ 0 & & v_0 \end{pmatrix} = d(v_0, \ldots, v_0) \in \mathrm{M}_m(R[u_m, \ldots, u_0, v_0]).$$

2) $\mathcal{S}(u_0, v_1, v_0) = (u_0)$; $\mathcal{S}(u_1, u_0, v_0) = (v_0)$.

3)

$$\mathcal{S}(u_2, u_1, u_0, v_2, v_1, v_0) = \begin{pmatrix} u_2 & u_1 & u_0 & 0 \\ 0 & u_2 & u_1 & u_0 \\ v_2 & v_1 & v_0 & 0 \\ 0 & v_2 & v_1 & v_0 \end{pmatrix}.$$

4)

$$\mathcal{S}(u_2,u_1,u_0;v_3,v_2,v_1,v_0)=\begin{pmatrix} u_2 & u_1 & u_0 & 0 & 0\\ 0 & u_2 & u_1 & u_0 & 0\\ 0 & 0 & u_2 & u_1 & u_0\\ v_3 & v_2 & v_1 & v_0 & 0\\ 0 & v_3 & v_2 & v_1 & v_0 \end{pmatrix}.$$

Определитель

$$S(u_m,\dots,u_0;v_n,\dots,v_0)=|\mathcal{S}(u_m,\dots,u_0;v_n,\dots,v_0)|\in\mathbb{Z}[u_m,\dots,u_0,v_n,\dots,v_0]$$

матрицы Сильвестра $\mathcal{S}(u_m,\dots,u_0;v_n,\dots,v_0)$ называется *определителем Сильвестра* (а его специализации в коммутативных кольцах R называются *результантами*).

Замечание 4.14.2. Определитель Сильвестра $S(u_m,\dots,u_0;v_n,\dots,v_0)$ — всегда ненулевой многочлен в кольце $\mathbb{Z}[u_m,\dots,u_0;v_n,\dots,v_0]$, поскольку диагональный член определителя $u_m^n v_0^m$ не сокращается ни с каким другим членом определителя $|\mathcal{S}((u_m,\dots,u_0;v_n,\dots,v_0))|$ в кольце $\mathbb{Z}[(u_m,\dots,u_0;v_n,\dots,v_0)]$.

Результант двух многочленов над коммутативным кольцом

Пусть R — коммутативное кольцо, $f(x)=\sum\limits_{i=0}^{m}a_ix^i, g(x)=\sum\limits_{j=0}^{n}b_jx^j\in R[x]$, $a_m\neq 0$, $b_n\neq 0$, $m+n\geqslant 1$. Под *результантом двух многочленов* $f(x)$ и $g(x)$ будем понимать

$$R(f,g)=S(a_m,\dots,a_0;b_n,\dots,b_0)\in R$$

(специализацию $u_m\to a_m,\dots,u_0\to a_0;v_n\to b_n,\dots,v_0\to b_0$). Таким образом,

$$R(f,g)=|\mathcal{S}(a_m,\dots,a_0;b_n,\dots,b_0)|\in R.$$

Замечание 4.14.3. В нашем изложении $R(f,g)$ не определяется, если:

1) либо $f=0$, либо $g=0$;

2) f и g — константы в R.

Пример 4.14.4. Пусть $R=\mathbb{Z}$, $f(x)=2x^3-5x^2+x+2$, $g(x)=x^2-3x+2$, $m=3$, $n=2$, $m+n=5$. Тогда

$$R(f,g)=S(2,-5,1,2;1,-3,2)=\begin{vmatrix} 2 & -5 & 1 & 2 & 0\\ 0 & 2 & -5 & 1 & 2\\ 1 & -3 & 2 & 0 & 0\\ 0 & 1 & -3 & 2 & 0\\ 0 & 0 & 1 & -3 & 2 \end{vmatrix}=0.$$

Упражнение 4.14.5. Пусть R — коммутативное кольцо без делителей нуля, $Q=Q(R)$ — его поле частных. Если $f(x)=\sum\limits_{i=0}^{m}a_ix^i, g(x)=\sum\limits_{j=0}^{n}b_jx^j\in R[x]$ и $\mathcal{S}=\mathcal{S}(a_m,\dots,a_0;b_n,\dots,b_0)$ — матрица Сильвестра многочленов $f(x)$ и $g(x)$, то, приводя матрицу $\mathcal{S}$ над полем Q элементарными преобразованиями строк к ступенчатому виду, в последней ненулевой строке получаем коэффициенты многочлена НОД$(f(x),g(x))$.

Теорема 4.14.6 (элементарные свойства результанта $R(f,g)$ двух многочленов над коммутативным кольцом R). *Пусть* $f = \sum\limits_{i=0}^{m} a_i x^i, g(x) = \sum\limits_{j=0}^{n} b_j x^j \in R[x]$, $m > 0$, $n > 0$, $0 \neq c \in R$. *Тогда:*

1) $R(c,g) = c^n$;

2) $R(f,f) = 0$;

3) $R(f,g) = (-1)^{mn} R(g,f)$;

4) $R(cf,g) = c^n R(f,g)$;

5) $R(x^k f, g) = b_0^k R(f,g)$, $k > 0$;

6) *если* $\varphi\colon R \to R'$ — *гомоморфизм коммутативных колец с* 1, $\Phi\colon R[x] \to R'[x]$, $\Phi\Big(\sum\limits_{i=0}^{m} a_i x^i\Big) = \sum\limits_{i=0}^{m} \varphi(a_i) x^i$, $\deg\big(\Phi(f)\big) = m$, $\deg\big(\Phi(g)\big) = k$, $0 \leqslant k \leqslant n$, *то*

$$\Phi\big(R(f,g)\big) = \varphi(a_m)^{n-k} R\big(\Phi(f), \Phi(g)\big).$$

Доказательство. Утверждения 1)—4) следуют непосредственно из определения $R(f,g)$.

Так как $R(xf,g) = b_0 R(f,g)$, то имеем 5).

Если $\deg\big(\Phi(g)\big) = n$, то $\varphi\big(R(f,g)\big) = R\big(\Phi(f), \Phi(g)\big)$, поскольку $\big|(\varphi(s_{ij}))\big| = \varphi|(s_{ij})|$ для матрицы Сильвестра $\mathcal{S} = (s_{ij})$. Если же $\deg\big(\Phi(g)\big) = k < n$, то в первых $n-k$ столбцах матрица Сильвестра является верхнетреугольной с элементом $\varphi(a_m)$ на диагонали. Применяя теорему Лапласа (по первым $n-k$ столбцам), получаем наше утверждение 6). □

Теорема 4.14.7. *Пусть* R — *коммутативное кольцо с* 1 *без делителей нуля,* $f(x) = a_m \prod\limits_{i=1}^{n}(x - \alpha_i), g(x) = b_n \prod\limits_{j=1}^{n}(x - \beta_j) \in R[\alpha_i, \beta_j, x]$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$. *Тогда:*

1) $R(f,g) = a_m^n b_n^m \prod\limits_{i=1}^{m} \prod\limits_{j=1}^{n} (\alpha_i - \beta_j)$;

2) $R(f,g) = a_m^n \prod\limits_{i=1}^{m} g(\alpha_i)$;

3) $R(f,g) = (-1)^{mn} b_n^m \prod\limits_{j=1}^{n} f(\beta_j)$.

Доказательство. Поскольку результант является однородным многочленом степени m по переменным α_i и степени n по переменным β_j, $f(x) = \sum\limits_{i=0}^{m} a_i x^i$, $g(x) = \sum\limits_{j=0}^{n} b_j x^j$, a_i и b_j — симметрические многочлены от переменных α_i и β_j соответственно, то результант является произведением $a_m^n b_n^m$ на симметрический многочлен от α_i, β_j.

В силу критерия Сильвестра $R(f,g)$ равен нулю, если $\alpha_k = \beta_l$ (многочлены f и g имеют общий множитель). Это означает, что $R(f,g)$ делится на $(\alpha_i - \beta_j)$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, и поэтому делится на произведение $\prod\limits_{i=1}^{m} \prod\limits_{j=1}^{n} (\alpha_i - \beta_j)$ (сомножители взаимно просты). Так как

$R(f,g)$ и правая часть равенства в 1) имеют одну и ту же степень, то они отличаются на константу. Они равны, поскольку они содержат член $(-1)^{mn}a_m^n b_n^m\Big(\prod\limits_{j=1}^{n}\beta_j\Big)^m = a_m^n b_0^m$ (мы используем равенство $b_0 = (-1)^n b_n \prod\limits_{j=1}^{n}\beta_j$).

Утверждения 2) и 3) следуют из 1). □

Замечание 4.14.8. Если α_i и β_j — соответственно корни многочленов $f(x)$ и $g(x)$, то равенство 1) в теореме (специализация переменных α_i, β_j) во многих математических текстах и нами в предшествующем параграфе используется как определение $R(f,g)$.

Теорема 4.14.9. *Пусть R — коммутативное кольцо с 1, $f(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_1x + a_0 \in R[x]$, $m = \deg f > 0$, $0 \neq g(x) = b_nx^n + b_{n-1}x^{n-1} + \ldots + b_1x + b_0 \in R[x]$, $n = \deg g \geqslant 0$, $(f) = (R[x])f$ — главный идеал в кольце $R[x]$, порождённый элементом f.*

Тогда следующие условия эквивалентны:

1) *$g + (f)$ — делитель нуля в фактор-кольце $R[x]/(f)$;*

2) *$R(f,g)$ — делитель нуля в кольце R.*

Доказательство. 1) $\Longrightarrow$ 2). Пусть $g \notin (f)$, $g + (f)$ — делитель нуля в $R[x]/(f)$. Это означает, что существует такой многочлен $h(x) \in R[x] \setminus (f)$, что $h(x)g(x) \in (f)$ (тогда $\big(h + (f)\big)\big(g + (f)\big) = hg + (f) = (f)$). Так как $1 \in U(R)$, то (см. ??)

$$h(x) = f(x)q(x) + r(x),$$

где $q(x), r(x) \in R[x]$, $r = 0$ или $\deg\big(r(x)\big) < m$. Поскольку $h(x) \notin (f)$, $r(x) \neq 0$. Кроме того,

$$r(x)g(x) = \big(h(x) - f(x)q(x)\big)g(x) = h(x)g(x) - f(x)q(x)g(x) \in (f),$$

следовательно, $r(x)g(x) = k(x)f(x)$ для некоторого $k(x) \neq 0$. Если $k(x) \neq 0$, то

$$\deg k + m = \deg(kf) = \deg(rg) \leqslant \deg r + \deg g < m + n,$$

и поэтому $\deg k \leqslant n - 1$. Итак, либо $k(x) = 0$, либо $\deg\big(k(x)\big) \leqslant n - 1$.

Пусть

$$r(x) = c_{m-1}x^{m-1} + \ldots + c_1x + c_0 \in R[x],$$
$$k(x) = -(d_{n-1}x^{n-1} + \ldots + d_1x + d_0) \in R[x]$$

(так как $\deg r \leqslant m - 1$, то возможно, что $c_{m-1} = 0$, но есть коэффициент $c_i \neq 0$, $0 \leqslant i \leqslant m - 1$; многочлен $k(x)$ может быть нулевым, поэтому мы допускаем, что все d_i могут равняться нулю). Итак:

$$(b_nx^n + \ldots + b_0)(c_{m-1}x^{m-1} + \ldots + c_0) + (d_{n-1}x^{n-1} + \ldots + d_0)(x^m + a_{m-1}x^{m-1} + \ldots + a_0) = 0.$$

Приравнивая коэффициенты, получаем:

$$b_nc_{m-1} + d_{n-1} = 0 \qquad (\text{при } x^{m+n-1});$$
$$b_nc_{m-2} + b_{n-1}c_{m-1} + d_{n-1}a_{m-1} + d_{n-2} = 0 \qquad (\text{при } x^{m+n-2});$$
$$b_nc_{m-3} + b_{n-1}c_{m-2} + b_{n-2}c_{m-1} + d_{n-1}a_{m-2} + d_{n-2}a_{m-1} + d_{n-3} = 0 \qquad (\text{при } x^{m+n-3});$$

$$\dots$$
$$b_1c_0 + b_0c_1 + a_0d_1 + a_1d_0 = 0 \qquad \text{(при } x\text{)};$$
$$b_0c_0 + a_0d_0 = 0 \qquad \text{(при } x^0 = 1\text{)}.$$

Таким образом, мы имеем $m+n$ линейных уравнений от $m+n$ переменных $c_{m-1}, \dots, c_0, d_{n-1}, \dots, d_0$:

$$(d_{n-1}, \dots, d_0, c_{m-1}, \dots, c_0)\mathcal{S}(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0) = (0, \dots, 0),$$

$\mathcal{S}(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0) \in \mathrm{M}_{m+n}(R)$, $m+n = \deg(f(x)) + \deg(g(x)) \geqslant 1$. Транспонируя, получаем, что квадратная однородная система

$$\mathcal{S}(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0)^{\mathrm{t}} \begin{pmatrix} d_{n-1} \\ \vdots \\ d_0 \\ c_{m-1} \\ \vdots \\ c_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathrm{M}_{m+n,1}(R)$$

имеет нетривиальное решение. В силу теоремы ?? $|\mathcal{S}(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0)^{\mathrm{t}}|$ — делитель нуля в кольце R. Поэтому

$$R(f,g) = S(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0) = \\ = |\mathcal{S}(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0)| = |\mathcal{S}(1, a_{m-1}, \dots, a_0; b_n, \dots, b_0)^{\mathrm{t}}| -$$

делитель нуля в кольце R.

2) $\Longrightarrow$ 1). Доказательство следует полным обращением первой части доказательства. Если $R(f,g)$ — делитель нуля в кольце R, то рассмотренная система линейных уравнений имеет нетривиальное решение $(d_{n-1}, \dots, d_0; c_{m-1}, \dots, c_0)$. Таким образом,

$$g(x)r(x) = k(x)f(x).$$

Если $c_{m-1} = \dots = c_0 = 0$, то $r(x) = 0$, и поэтому $k(x)f(x) = 0$. Так как старший коэффициент многочлена $f(x)$ равен 1, то $k(x) = 0$, но тогда

$$(d_{n-1}, \dots, d_0, c_{m-1}, \dots, c_0) = (0, \dots, 0),$$

что противоречит нетривиальности решения. Следовательно, $c_i \neq 0$ при некотором $0 \leqslant i \leqslant n-1$, и, таким образом, $r(x) \neq 0$. Так как $\deg(r(x)) \leqslant m-1$, то $r(x) \notin (f)$. Поэтому $g + (f)$ — делитель нуля в фактор-кольце $R[x]/(f)$. □

Следствие 4.14.10. *Пусть K — поле, $f(x), g(x) \in K[x]$, $\deg f(x) > 0$, $\deg g(x) > 0$. Тогда многочлены $f(x)$ и $g(x)$ имеют общий неприводимый множитель в том и только в том случае, когда $R(f,g) = 0$.*

Доказательство. 1) Пусть $h(x)$ — общий неприводимый множитель для многочленов $f(x)$ и $g(x)$ в $R[x]$. Тогда $\deg f, \deg g \geqslant \deg h \geqslant 1$. Пусть

$$f(x) = h(x)f_1(x), \quad \deg f_1 < \deg f, \quad g(x) = h(x)g_1(x),$$

поэтому $f_1(x) \notin (f)$.

Так как

$$f_1 g = f_1 h g_1 = g_1 f \in (f),$$

то $g + (f)$ — делитель нуля в $K[x]/(f)$, и следовательно, $R(f, g) = 0$.

2) Если $R(f, g) = 0$, то $g + (f)$ — делитель нуля в $K[x]/(f)$. Таким образом, $g(x)r(x) = k(x)f(x)$ для некоторых многочленов $r(x), k(x) \in K[x]$ и $r(x) \notin (f)$. Так как $K[x]$ — *факториальная область*, то если $f(x)$ и $g(x)$ не имеют общего неприводимого множителя, то из $g(x)r(x) = k(x)f(x)$ следует, что $f \mid r$, и следовательно, $r(x) \in (f)$, что невозможно. Итак, f и g имеют общий неприводимый множитель. □

Следствие 4.14.11. *Пусть R — факториальная область, $f(x), g(x) \in R[x]$, $\deg f(x) > 0$, $\deg g(x) > 0$. Многочлены $f(x)$, $g(x)$ имеют общий неприводимый множитель в $R[x]$ тогда и только тогда, когда $R(f, g) = 0$.*

Доказательство. Пусть $K = Q(R)$ — поле частных кольца R. Тогда по следствию 4.14.11 многочлены $f(x), g(x) \in R[x] \subseteq K[x]$ имеют общий неприводимый множитель в $K[x]$ положительной степени по x тогда и только тогда, когда $R(f, g) = 0$.

Если $f = \varphi f_1, g = \varphi g_1 \in K[x]$, d — общий знаменатель коэффициентов многочленов φ, f_1, g_1. Пусть $\xi = d\varphi, f_2 = df_1, g_2 = dg_1 \in R[x]$. Тогда $d^2 f = \xi f_2, d^2 g = \xi g_2 \in R[x]$. Пусть ψ — неприводимый делитель многочлена ξ положительной степени по x (он существует, поскольку φ имеет положительную степень по x, $\psi = d\varphi$). Тогда ψ делит d^2 или f, поэтому ψ делит f ($d^2 \in R$). Аналогично, ψ делит g. □

Следствие 4.14.12. *Пусть K — поле, $\bar{K}$ — его алгебраическое замыкание, $f(x), g(x) \in K[x]$, $\deg f(x) > 0$, $\deg g(x) > 0$. Тогда многочлены $f(x)$ и $g(x)$ имеют общий корень в поле $\bar{K}$ тогда и только тогда, когда $R(f, g) = 0$.*

Доказательство. Наличие общего корня $c \in \bar{K}$ для $f(x), g(x) \in K[x] \subseteq \bar{K}[x]$ равносильно существованию общего неприводимого множителя в $\bar{K}[x]$ (над алгебраически замкнутым полем все неприводимые многочлены линейны), что, в свою очередь, равносильно условию $R(f, g) = 0 \in K \subseteq \bar{K}$. □

Пример 4.14.13. Пусть $f(x) = 2x^3 - 5x^2 + x + 2, g(x) = x^2 - 3x + 2 \in \mathbb{Q}[x]$. Как мы видели (см. ??), $R(f, g) = 0$, поэтому $f(x)$ и $g(x)$ имеют общий корень в $\bar{\mathbb{Q}}$ (на самом деле 1 и 2 — общие корни многочленов $f(x)$ и $g(x)$ уже в $\mathbb{Z}$).

Однородность и вес определителя Сильвестра

Как мы видели, определитель Сильвестра — ненулевой многочлен,

$$0 \neq S(u_m, u_{m-1}, \ldots, u_0; v_n, v_{n-1}, \ldots, v_0) \in \mathbb{Z}[u_m, \ldots, u_0; v_n, \ldots, v_0].$$

Если имеем ненулевой одночлен

$$0 \neq M = c u_m^{\alpha(m)} \ldots u_0^{\alpha_0} v_n^{\beta(n)} \ldots v_0^{\beta(0)} \in \mathbb{Z}[u_m, \ldots, u_0; v_n, \ldots, v_0],$$
$$c \in \mathbb{Z} \setminus \{0\}, \quad (\alpha(m), \ldots, \alpha(0), \beta(n), \ldots, \beta(0)) \in \mathbb{N}_0^{m+n+2},$$

то определена *степень* одночлена

$$d(M) = \sum_{i=0}^{m} \alpha(i) + \sum_{j=0}^{n} \beta(j),$$

а также его *вес*

$$\omega(M) = \sum_{i=0}^{m} (m-i)\alpha(i) + \sum_{j=0}^{n} (n-j)\beta(j).$$

Пример 4.14.14. Если $M = u_m^2 v_{n-1}^3$, то $d(M) = 5$, $\omega(M) = 0 \cdot 2 + 1 \cdot 3 = 3$; если $M = u_m^2 u_{m-1}^3 u_{m-2}^2 v_n^5 v_{n-1}^2 v_{n-3}^2$, то $d(M) = 16$ $\omega(M) = 0 \cdot 2 + 1 \cdot 3 + 2 \cdot 2 + 0 \cdot 5 + 1 \cdot 2 + 3 \cdot 2 = 15$.

Если каждый одночлен M_i многочлена

$$0 \neq f = r_1 M_1 + \ldots + r_t M_t \in \mathbb{Z}[u_m, \ldots, u_0; v_n, \ldots, v_0], \quad r_i \in \mathbb{Z} \setminus \{0\},$$

имеет один и тот же вес $\omega(M_i) = p$, то многочлен f называется *изобарическим многочленом* веса p.

Примеры 4.14.15.

1) Если $f = u_1^2 v_0^2 + 2u_1 u_0 v_1 v_0 - 3v_0^4 \in \mathbb{Z}[u_1, u_0, v_1, v_0]$, то $d(f) = 4$, f — однородный многочлен, $\omega(u_1^2 v_0^2) = 2$, $\omega(2u_1 u_0 v_1 v_0) = 2$, $\omega(-3v_0^4) = 4$, таким образом, f не является изобарическим многочленом.

2) Пусть

$$f = S(u_2, u_1, u_0, v_2, v_1, v_0) = \begin{vmatrix} u_2 & u_1 & u_0 & 0 \\ 0 & u_2 & u_1 & u_0 \\ v_2 & v_1 & v_0 & 0 \\ 0 & v_2 & v_1 & v_0 \end{vmatrix} =$$
$$= u_2^2 v_0^2 - u_2 u_1 v_1 v_0 + u_2 u_0 v_1^2 - u_2 u_0 v_2 v_0 +$$
$$+ u_1^2 v_2 v_0 - u_2 u_0 v_2 v_0 - u_1 u_0 v_2 v_1 + u_0^2 v_2^2,$$

f — однородный многочлен степени $d(f) = 4$, при этом f — изобарический многочлен веса $\omega(f) = 4$.

Теорема 4.14.16. *Многочлен Сильвестра $S(u_m, \ldots, u_0; v_n, \ldots, v_0)$ является однородным многочленом степени $m+n$ и изобарическим многочленом веса mn.*

Доказательство. 1) Пусть $R = \mathbb{Z}[u_m, \ldots, u_0; v_n, \ldots, v_0]$. Рассмотрим кольцо многочленов $R[y]$. Тогда

$$S(yu_m, \ldots, yu_0; yv_n, \ldots, yv_0) = |yS(u_m, \ldots, u_0; v_n, \ldots, v_0)| =$$
$$= y^{m+n} |S(u_m, \ldots, u_0; v_n, \ldots, v_0)| = y^{m+n} S(u_m, \ldots, u_0; v_n, \ldots, v_0).$$

В силу критерия однородности многочлена 4.4.4 получаем, что $S(u_m, \ldots, u_0; v_n, \ldots, v_0)$ — однородный многочлен степени $m+n$.

2) Пусть

$$\mathcal{S}(u_m,\dots,u_0;v_n,\dots,v_0)=(s_{ij})\in \mathrm{M}_{m+n}(\mathbb{Z}[u_m,\dots,u_0;v_n,\dots,v_0]),$$

тогда

$$S=|\mathcal{S}|=\sum_{\sigma\in \mathrm{S}_{m+n}}\varepsilon(\sigma)\prod_{i=1}^{m+n}s_{i\sigma(i)}.$$

Так как

$$s_{ij}=\begin{cases}u_{m-j+i} & \text{для } i=1,2,\dots,n,\\ v_{-j+i} & \text{для } i=n+1,\dots,m+n\end{cases}$$

(см. ??), то

$$M=\prod_{i=1}^{m+n}s_{i\sigma(i)}=\left(\prod_{i=1}^{n}u_{m-\sigma(i)+i}\right)\left(\prod_{i=n+1}^{m+n}v_{-\sigma(i)+i}\right),$$

и его вес (если этот одночлен ненулевой) равен

$$\omega(M)=\sum_{i=1}^{n}(\sigma(i)-i)+\sum_{i=n+1}^{m+n}(n+\sigma(i)-i)=\sum_{i=1}^{m+n}\sigma(i)-\sum_{i=1}^{m+n}i+\sum_{i=n+1}^{m+n}n=\sum_{i=n+1}^{m+n}n=mn. \quad \square$$

Теорема 4.14.17 (о линейном выражении результанта многочленов). *Пусть* $f(x)=\sum\limits_{i=0}^{m}a_ix^i, g(x)=\sum\limits_{j=0}^{n}b_jx^j\in R[x]$, $a_m\neq 0$, $b_n\neq 0$, $m+n\geqslant 1$. *Тогда существуют* $a(x),b(x)\in R[x]$ *такие, что*

1) $d(a(x))\leqslant n-1$, $d(b(x))\leqslant m-1$;

2) $R(f,g)=a(x)f(x)+b(x)g(x)$.

Доказательство. Пусть $\mathcal{S}=(s_{ij})=\mathcal{S}(a_m,\dots,a_0;b_n,\dots,b_0)$ — $((m+n)\times(m+n))$-матрица Сильвестра, полученная специализацией переменных $u_m\to a_m,\dots,u_0\to a_0$, $v_n\to b_n,\dots,v_0\to b_0$, $R(f,g)=|\mathcal{S}(a_m,\dots,a_0;b_n,\dots,b_0)|$. Ясно, что

$$\mathcal{S}(a_m,\dots,a_0;b_n,\dots,b_0)\begin{pmatrix}x^{n+m-1}\\ x^{n+m-2}\\ \vdots\\ x^m\\ x^{m-1}\\ \vdots\\ x\\ 1\end{pmatrix}=\begin{pmatrix}x^{m-1}f(x)\\ x^{m-2}f(x)\\ \vdots\\ f(x)\\ x^{m-1}g(x)\\ \vdots\\ xg(x)\\ g(x)\end{pmatrix}. \tag{4.6}$$

Пусть $c_i=A_{i,m+n}$ — алгебраическое дополнение к элементу матрицы $\mathcal{S}$, стоящему на пересечении i-й строки и $(m+n)$-го столбца, $i=1,\dots,m+n$. Разложение определителя $R(f,g)$ по $(m+n)$-му столбцу и фальшивые разложения по j-му столбцу, $j\neq m+n$, дают

$$\sum_{i=1}^{m+n}c_is_{i,m+n}=R(f,g),$$

$$\sum_{i=1}^{m+n} c_i s_{i,j} = 0, \quad j \neq m+n.$$

Умножая в матричном равенстве (4.6) каждую i-ю строку на c_i и складывая, получаем

$$\begin{aligned}
&c_1 x^{n-1} f(x) + c_2 x^{n-2} f(x) + \ldots + c_n f(x) + \\
&\quad + c_{n+1} x^{m-1} g(x) + c_{n+2} x^{m-2} g(x) + \ldots + c_{m+n} g(x) = \\
&\quad = \sum_{j=1}^{m+n} c_1 s_{1j} x^{m+n-j} + \sum_{j=1}^{m+n} c_2 s_{2j} x^{m+n-j} + \ldots + \sum_{j=1}^{m+n} c_{m+n} s_{m+n,j} x^{m+n-j} = \\
&\quad = \sum_{i=1}^{m+n} \sum_{j=1}^{m+n} c_i s_{ij} x^{m+n-j} = \sum_{j=1}^{m+n} \left(\sum_{i=1}^{m+n} c_i s_{ij} \right) x^{m+n-j} = R(f,g).
\end{aligned}$$

Итак,

$$a(x) f(x) + b(x) g(x) = R(f,g),$$

где

$$\begin{aligned}
a(x) &= c_1 x^{n-1} + c_2 x^{n-2} + \ldots + c_n, \\
b(x) &= c_{n+1} x^{m-1} + c_{n+2} x^{m-2} + \ldots + c_{m+n}.
\end{aligned}$$

□

Результант двух многочленов многих переменных по фиксированной переменной

Пусть $f = f(x_1, \ldots, x_p), g = g(x_1, \ldots, x_p) \in R[x_1, \ldots, x_p]$, $f \neq 0$, $g \neq 0$. Выберем переменную x_i, $1 \leqslant i \leqslant p$. Пусть $m = \deg_{x_i}(f)$ и $n = \deg_{x_i}(g)$. Тогда

$$f(x_1, \ldots, x_p) = \sum_{k=0}^{m} a_k x_i^k, \quad g(x_1, \ldots, x_p) = \sum_{k=0}^{n} b_k x_i^k,$$

где

$$a_0, \ldots, a_m, b_0, \ldots, b_n \in R[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p], \quad a_m \neq 0, \quad b_n \neq 0.$$

Допустим, что $m + n \geqslant 1$, и тогда мы можем образовать определитель

$$R_{x_i}(f,g) = S(a_m, \ldots, a_0, b_n, \ldots, b_0) \in R[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p].$$

Пример 4.14.18. Пусть

$$f(x,y) = x^2 + y^2 + xy + x + 1, \; g(x,y) = xy + x + y + 1 \in \mathbb{Z}[x,y].$$

Тогда:

а) $f(x,y) = x^2 + (1+y)x + (1+y^2)$, $n = \deg_x(f) = 2$; $g(x,y) = (1+y)x + (1+y)$, $m = \deg_x(g) = 1$; $m + n = 3 \geqslant 1$;

$$R_x(f,g) = \begin{vmatrix} 1 & 1+y & 1+y^2 \\ 1+y & 1+y & 0 \\ 0 & 0 & 1+y \end{vmatrix} = (1+y)^2(1-y+y^2);$$

б) $f(x,y) = y^2 + xy + (x^2 + x + 1)$, $n = \deg_y(f) = 2$; $g(x,y) = (1+x)y + (1+x)$, $m = \deg_y(g) = 1$; $m + n \geqslant 1$;

$$R_y(x,y) = \begin{vmatrix} 1 & x & x^2+x+1 \\ 1+x & 1+x & 0 \\ 0 & 1+x & 1+x \end{vmatrix} = (1+x)^2(2+x^2).$$

Упражнение 4.14.19. Пусть F — поле, $f, g \in F[x,y]$, $m = \deg_x f$, $n = \deg_x g$, $\deg_y f, \deg_y g \leqslant d$. Покажите, что

$$\deg_y\big(R_x(f,g)\big) \leqslant (m+n)d.$$

Упражнение 4.14.20. Пусть рациональная кривая γ на плоскости задана параметрически:

$$\gamma = \left\{ \left(\frac{a(t)}{b(t)}, \frac{c(t)}{d(t)} \right) \in \mathbb{R}^2 \;\middle|\; t \in \mathbb{R} \right\},$$

где $a(t), b(t), c(t), d(t) \in \mathbb{Q}[t]$. Покажите, что кривая γ задаётся уравнением $f(x,y) = 0$, где

$$f(x,y) = \mathrm{Res}_t\big(b(t)\cdot x - a(t),\ d(t)\cdot y - c(t)\big) \in \mathbb{Q}[x,y].$$

Упражнение 4.14.21. Пусть α и β — алгебраические числа над $\mathbb{Q}$, $f, g \in \mathbb{Q}[x]$ — их минимальные неприводимые многочлены. Покажите, что $\mathrm{Res}_t\big(f(x), g(z-x)\big)$ — минимальный многочлен для $\alpha + \beta$, $\mathrm{Res}_x\left(f(x), g\left(\frac{z}{x}\right)\cdot x^{\deg(g)}\right)$ — минимальный многочлен для $\alpha \cdot \beta$.

Результант однородных многочленов от многих переменных

Теорема 4.14.22. *Пусть $f = f(x_1,\ldots,x_p), g = g(x_1,\ldots,x_p) \in R[x_1,\ldots,x_p]$, f и g — однородные многочлены степени m и n соответственно. Предположим, что для переменной x_i, $1 \leqslant i \leqslant p$,*

$$f = a_m x_i^m + a_{m-1}x_i^{m-1} + \ldots + a_1 x_i + a_0,$$
$$g = b_n x_i^n + a_{n-1}x_i^{n-1} + \ldots + b_1 x_i + b_0,$$

где $a_m,\ldots,a_0,b_n,\ldots,b_0 \in R[x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_p]$, $a_m \neq 0$, $b_n \neq 0$, $m+n \geqslant 1$. Тогда $R_{x_i}(f,g)$ либо нулевой многочлен, либо однородный многочлен степени mn в $R[x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_p]$.

Доказательство проведём для $i = p$ (несколько короче запись обозначений).

Итак: $f, g \in R[x_1,\ldots,x_p]$, f — однородный многочлен степени m, g — однородный многочлен степени n, $m+n \geqslant 1$,

$$f = a_m x_p^m + a_{m-1}x_p^{m-1} + \ldots + a_1 x_p + a_0, \quad \deg_{x_p}(f) = m, \quad a_m \neq 0,$$
$$g = b_n x_p^n + b_{n-1}x_p^{n-1} + \ldots + b_1 x_p + b_0, \quad \deg_{x_p}(g) = n, \quad b_n \neq 0,$$

где $a_m,\ldots,a_0; b_n,\ldots,b_0 \in R[x_1,\ldots,x_{p-1}]$, при этом $a_i(x_1,\ldots,x_{p-1})$, $0 \leqslant i \leqslant m$, — однородный многочлен степени $m-i$, $b_j(x_1,\ldots,x_{p-1})$, $0 \leqslant j \leqslant n$, — однородный многочлен степени $n-j$.

Пусть

$$h(x_1,\ldots,x_{p-1}) = R_{x_p}(f,g) = S(a_m,\ldots,a_0;b_n,\ldots,b_0).$$

Тогда для переменной y над кольцом $R[x_1,\ldots,x_p]$ имеем

$$h(yx_1,\ldots,yx_{p-1}) = |C| = \begin{vmatrix} a_m & ya_{m-1} & y^2a_{m-2} & \ldots & y^ma_0 & 0 & 0 & \ldots & 0 \\ 0 & a_m & ya_{m-1} & \ldots & y^{m-1}a_1 & y^ma_0 & 0 & \ldots & 0 \\ \vdots & \vdots & & & \vdots & & & & \vdots \\ 0 & 0 & \ldots & a_m & ya_{m-1} & \ldots & \ldots & \ldots & y^ma_0 \\ b_n & yb_{n-1} & \ldots & \ldots & y^nb_0 & 0 & \ldots & \ldots & 0 \\ 0 & b_n & \ldots & \ldots & y^{n-1}b_1 & y^nb_0 & \ldots & \ldots & 0 \\ \vdots & \vdots & & & \vdots & & & & \vdots \\ 0 & 0 & \ldots & b_n & yb_{n-1} & \ldots & \ldots & \ldots & y^nb_0 \end{vmatrix}.$$

В матрице C умножим первые n строк на $y, y^2, \ldots, y^n$ соответственно, последние m строк умножим на $y^2, \ldots, y^m$ соответственно, получим матрицу

$$L = (l_{ij}) = \begin{pmatrix} a_my & a_{m-1}y^2 & \ldots & a_0y^{m+1} & 0 & 0 & \ldots & 0 \\ 0 & a_my^2 & \ldots & a_1y^{m+1} & a_0y^{m+2} & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ 0 & 0 & \ldots & a_my^n & a_{m-1}y^{n+1} & \ldots & \ldots & a_0y^{m+n} \\ b_ny & b_{n-1}y^2 & \ldots & b_0y^{n+1} & 0 & 0 & \ldots & 0 \\ 0 & b_ny^2 & \ldots & b_1y^{n+1} & b_0y^{n+2} & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ 0 & 0 & \ldots & b_ny^m & b_{n-1}y^{m+1} & \ldots & \ldots & b_0y^{m+n} \end{pmatrix},$$

где

$$l_{ij} = y^ja_{m+i+j} \quad \text{для } 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant m+n;$$
$$l_{n+k,j} = y^jb_{n+i-j} \quad \text{для } 1 \leqslant k \leqslant m,\ 1 \leqslant j \leqslant m+n$$

(здесь $a_s = 0$, если $s < 0$ или $s > m$; $b_t = 0$, если $t < 0$ или $t > n$).

Ясно, что

$$|L| = y^\alpha h(yx_1,\ldots,yx_{p-1}),$$

где

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

С другой стороны, вынося y^j из каждого j-го столбца в $|L|$, получаем

$$|L| = y^\beta h(x_1,\ldots,x_{p-1}),$$

где

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

Следовательно,

$$h(yx_1,\ldots,yx_{p-1}) = y^{\beta-\alpha}h(x_1,\ldots,x_{p-1}),$$

где $\beta - \alpha = mn$. Таким образом, $h(x_1,\ldots,x_{p-1}) = R_{x_p}(f,g)$ — однородный многочлен степени mn (или нулевой многочлен).

□

Пример 4.14.23. Пусть $f, g \in \mathbb{Z}[x_1, x_2, x_3, x_4]$.

а) $f = 2x_1^2 + x_1x_2$ ($n = 2$, $a_2 = 2 \neq 0$, $a_1 = x_2$, $a_0 = 0$), $g = x_1 - x_3$ ($m = 1$, $b_1 = 1 \neq 0$, $b_0 = -x_3$), f и g — однородные многочлены степени 2 и 1 соответственно, $m + n = 3 \geqslant 1$,

$$R_{x_1}(f, g) = \begin{vmatrix} 2 & x_2 & 0 \\ 1 & -x_3 & 0 \\ 0 & 1 & -x_3 \end{vmatrix} = 2x_3^2 + x_2x_3 -$$

однородный многочлен степени $mn = 2$ в $\mathbb{Z}[x_2, x_3, x_4]$.

б) $f = 2x_1^2 + x_1x_2$ ($n = 2$, $a_2 = 2 \neq 0$, $a_1 = x_2$, $a_0 = 0$), $g = x_1^2 + x_3x_4$ ($m = 2$, $b_2 = 1 \neq 0$, $b_1 = 0$, $b_0 = x_3x_4$), f и g — однородные многочлены степени 2, $m + n = 4 \geqslant 1$,

$$R_{x_1}(f, g) = \begin{vmatrix} 2 & x_2 & 0 & 0 \\ 0 & 2 & x_2 & 0 \\ 1 & 0 & x_3x_4 & 0 \\ 0 & 1 & 0 & x_3x_4 \end{vmatrix} = 4x_3^2x_4^2 + x_2^2x_3x_4 -$$

однородный многочлен степени 4 в $\mathbb{Z}[x_2, x_3, x_4]$.

Пример 4.14.24. Пусть $f = x_1x_2, g = x_3x_4 \in \mathbb{Z}[x_1, x_2, x_3, x_4]$, f, g — однородные многочлены степени 2, $\deg_{x_1}(f) = 1$, $\deg_{x_2}(f) = 1$, $\deg_{x_3}(f) = 0$, $\deg_{x_4}(f) = 0$ (в частности, g не удовлетворяет условиям теоремы 4.14.22). Тогда $R_{x_1}(f, g) = x_3x_4$ — однородный многочлен степени 2 (а не степени $mn = 4$) в $\mathbb{Z}[x_2, x_3, x_4]$, что иллюстрирует существенность условий теоремы.

Теорема 4.14.25. *Пусть R — коммутативное кольцо без делителей нуля, $Q = Q(R)$ — его поле частных, $f(x) = a_mx^m + \ldots + a_1x + a_0, g(x) = b_nx^n + \ldots + b_1x + b_0 \in R[x]$, $a_m \neq 0$, $b_n \neq 0$, $m, n \geqslant 1$. Допустим, что $\{\xi_1, \ldots, \xi_m\}$ и $\{\theta_1, \ldots, \theta_n\}$ — корни многочленов f и g соответственно в некотором поле разложения E многочлена $f(x)g(x)$ над полем частных $Q = Q(R)$, $Q \subseteq E$. Тогда*

$$R(f, g) = a_n^m \prod_{i=1}^{n} g(\xi_i) = (-1)^{mn} b_m^n \prod_{j=1}^{m} f(\theta_j).$$

Доказательство (*см. также* ??). Ясно, можно считать, что $R = Q = Q(R)$ — поле.

1) Так как $a_m \neq 0$, $b_n \neq 0$, то запишем $f(x) = a_mf_1(x)$, $g(x) = b_ng_1(x)$, где $f_1(x), g_1(x) \in Q[x]$ — многочлены, старшие коэффициенты которых равны 1, при этом они имеют те же корни, что и многочлены $g(x)$ и $g(x)$ соответственно. Если наша теорема верна для многочленов $f_1(x)$ и $g_1(x)$, то

$$R(f, g) = R_m\big(a_mf_1(x), b_ng_1(x)\big) = a_m^n b_n^m R(f_1, g_1) =$$
$$= a_m^n b_n^m \prod_{i=1}^{m} g_1(\xi_i) = a_m^n \prod_{i=1}^{m} b_ng_1(\xi_i) = a_m^n \prod_{i=1}^{m} g(\xi_i);$$

аналогично,

$$R(f, g) = (-1)^{mn} b_m^n \prod_{j=1}^{m} f(\theta_j).$$

Таким образом, утверждение верно для $f(x)$ и $g(x)$.

2) Итак, можно считать, что $a_m = 1$, $n_n = 1$ для $f, g \in \mathbb{Q}[x]$, тогда

$$f(x) = \prod_{i=1}^{m}(x - \xi_i), g(x) = \prod_{j=1}^{n}(x - \theta_j) \in E[x].$$

Пусть в кольце многочленов $\mathbb{Z}[x, x_1, \ldots, x_m, y_1, \ldots, y_n]$

$$F(x, x_1, \ldots, x_m, y_1, \ldots, y_n) = \prod_{i=1}^{m}(x - x_i);$$

$$G(x, x_1, \ldots, x_m, y_1, \ldots, y_n) = \prod_{j=1}^{n}(x - y_j);$$

$$H(x, x_1, \ldots, x_m, y_1, \ldots, y_n) = \prod_{i=1}^{m}\prod_{j=1}^{n}(x_i - y_j).$$

Ясно, что F, G и H — однородные многочлены степеней соответственно m, n и mn в кольце $\mathbb{Z}[x, x_1, \ldots, x_m, y_1, \ldots, y_n]$. Пусть $P[x_1, \ldots, x_m, y_1, \ldots, y_n] = R_x(F, G)$. В силу леммы ?? $P(x_1, \ldots, x_m, y_1, \ldots, y_n)$ либо нулевой многочлен, либо однородный многочлен степени mn в кольце $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]$.

Ясно, что $P(x_1, \ldots, x_m, y_1, \ldots, y_n) \neq 0$, поскольку можно выбрать такую специализацию для переменных $x_1, \ldots, x_m, y_1, \ldots, y_n$ в $\mathbb{Q}$, что соответствующие многочлены $\bar{F}$ и $\bar{G}$ не будут иметь общих корней, и поэтому в силу следствия ?? $R(\bar{F}, \bar{G}) \neq 0$. Таким образом, $R(\bar{F}, \bar{G})$ является специализацией многочлена $R_x(F, G)$, следовательно, $P(x_1, \ldots, x_m, y_1, \ldots, y_n)$ — ненулевой однородный многочлен степени mn в кольце $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]$.

3) Многочлен $H(x_1, \ldots, x_m, y_1, \ldots, y_n)$ также является ненулевым однородным многочленом степени mn в кольце $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]$. Покажем, что многочлен H делит многочлен P. Действительно, зафиксируем индексы $i \in \{1, \ldots, m\}$ и $j \in \{1, \ldots, n\}$ и рассмотрим одночлен $x_i - y_j$.

Если мы заменим x_i на y_j в многочлене $F(x, x_1, \ldots, x_m, y_1, \ldots, y_n)$, то получим общий множитель $x - y_j$ с многочленом G, таким образом, по следствию ??

$$P(x_1, \ldots, x_{i-1}, y_j, x_{i+1}, \ldots, x_m, y_1, \ldots, y_n) = 0,$$

но это означает, что $x_i - y_j$ делит многочлен $P(x_1, \ldots, x_m, y_1, \ldots, y_n)$ в кольце

$$\mathbb{Z}[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m, y_1, \ldots, y_n][x_i] = \mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]$$

(теорема Безу).

Итак, $(x_i - y_j)$ делит многочлен $P(x_1, \ldots, x_m, y_1, \ldots, y_n)$ для всех $i = 1, \ldots, m$, $j = 1, \ldots, n$.

Так как многочлены $(x_i - y_j)$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, — взаимно простые неприводимые многочлены, то H делит многочлен P в кольце $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]$.

4) Так как H и P — ненулевые однородные многочлены степени mn в кольце $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]$ и H делит P, то

$$P(x_1, \ldots, x_m, y_1, \ldots, y_n) = cH(x_1, \ldots, x_m, y_1, \ldots, y_n)$$

для $0 \neq c \in \mathbb{Z}$. Это означает, что

$$R_x\big(F(x, x_1, \ldots, x_m, y_1, \ldots, y_n), G(x, x_1, \ldots, x_m, y_1, \ldots, y_n)\big) = c\prod_{i=1}^{m}\prod_{j=1}^{n}(x_i - y_j). \qquad (4.7)$$

При замене в (4.7) $x_1, \ldots, x_m$ на $\xi_1, \ldots, \xi_m$, $y_1, \ldots, y_n$ на $\theta_1, \ldots, \theta_n$ соответственно многочлен $F(x, x_1, \ldots, x_m, y_1, \ldots, y_n)$ превращается в $\prod\limits_{i=1}^{m}(x - \xi_i) = f(x)$, многочлен $G(x, x_1, \ldots, x_m, y_1, \ldots, y_n)$ превращается в $\prod\limits_{j=1}^{n}(x - \theta_j) = g(x)$. Таким образом, из равенства (4.7) получаем

$$R(f, g) = c\prod_{i=1}^{m}\prod_{j=1}^{n}(\xi_i - \theta_j). \qquad (4.8)$$

В частности,

$$R(f, g) = c\prod_{i=1}^{m} g(\xi_i).$$

Также мы получаем

$$R(f, g) = c\prod_{i=1}^{n}\prod_{i=1}^{m} -(\theta_j - \xi_i) = c\prod_{j=1}^{n}(-1)^m f(\theta_j) = c(-1)^{mn}\prod_{j=1}^{n} f(\theta_j).$$

5) Покажем, что $c = 1$. Пусть

$$f(x) = (x-1)^m, g(x) = x^n + z_1x^{n-1} + \ldots + z_n \in (\mathbb{Z}[z_1, \ldots, z_n])[x].$$

Так как $\mathbb{Z}[z_1, \ldots, z_n]$ — коммутативная область и определение элемента c не зависит от выбора f и g (c определяется уравнением (4.7)), при этом

$$f(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_1x + a_0 \in \mathbb{Z}[x]$$

и коэффициент при z_n^m в

$$R(f, g) = S(1, a_{m-1}, \ldots, a_0, 1, z_1, \ldots, z_n)$$

равен 1, то

$$R(f, g) = c\prod_{i=1}^{m} g(1) = c\prod_{i=1}^{m}(1 + z_1 + \ldots + z_n).$$

Итак, $c = 1$. □

Теорема 4.14.26. *Пусть R — коммутативное кольцо с 1, $a \in \mathrm{M}_n(R)$, $c_A(x) \in R[x]$ — характеристический многочлен матрицы A, $0 \neq g(x) \in R[x]$. Тогда*

$$R\big(c_A(x), g(x)\big) = |g(A)|.$$

Доказательство. 1) Если $\deg(g) = 0$, то $g(x) = b \in R$, и тогда $R(c_A(x), b) = b^n$; $g(A) = bE_n$, $|g(A)| = |bE_n| = b^n$.

2) Пусть $\deg(g) \geqslant 1$.

2а) Пусть $R = \mathbb{Z}$ и E — поле разложения многочлена $c_A(x)g(x)$ над полем $\mathbb{Q}$, $c_A(x) = \prod\limits_{i=1}^{n}(x - \xi_i)$, $g(x) = b\prod\limits_{j=1}^{m}(x - \theta_j)$, $0 \neq b \in \mathbb{Z}$. В силу теоремы 4.14.25

$$R\big(c_A(x), g(x)\big) = \prod_{i=1}^{n} g(\xi_i).$$

С другой стороны, в кольце $\mathrm{M}_n(E)$ имеем

$$g(A) = b\prod_{j=1}^{m}(A - \theta_j E_n) = b(-1)^m \prod_{j=1}^{m}(\theta_j E_n - A);$$

$$|g(A)| = b^n(-1)^{mn}\prod_{j=1}^{m}|\theta_j E_n - A| =$$

$$= b^n(-1)^{mn}\prod_{j=1}^{m} c_A(\theta_j) = b^n(-1)^{mn}\prod_{j=1}^{m}\prod_{i=1}^{n}(\theta_j - \xi_i) =$$

$$= \prod_{i=1}^{n}\left(b\prod_{j=1}^{m}(\xi_i - \theta_j)\right) = \prod_{i=1}^{n} g(\xi_i) = R\big(c_A(x), g(x)\big).$$

2б) В общем случае (R — любое коммутативное кольцо) применим лемму 4.5.4 (о специализации). Пусть $A = (a_{ij}) \in \mathrm{M}_n(R)$; $g(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0 \in R[x]$, $b_m \neq 0$, $m \geqslant 1$; $\{x_{ij} \mid 1 \leqslant i, j \leqslant n\}$, $\{y_m, \ldots, y_0\}$ и x — независимые переменные над $\mathbb{Z}$, $\bar{R} = \mathbb{Z}[x_{11}, \ldots, x_{nn}, y_m, \ldots, y_0]$, $\bar{A} = (x_{ij}) \in \mathrm{M}_n(\bar{R})$, $\bar{g}(x) = y_m x^m + y_{m-1}x^{m-1} + \ldots + y_1 x + y_0 \in \bar{R}[x]$; $F(x_{11}, \ldots, x_{nn}, y_m, \ldots, y_0) = R_x\big(c_{\bar{A}}(x), \bar{g}(x)\big)$, $G(x_{11}, \ldots, x_{nn}, y_m, \ldots, y_0) = |\bar{g}(\bar{A})|$. Из первой части доказательства в 2а) (при $R = \mathbb{Z}$):

$$F(j_{11}, \ldots, j_{nn}, j_m, \ldots, j_0) = G(j_{11}, \ldots, j_{nn}, j_m, \ldots, j_0)$$

для всех $j_{11}, \ldots, j_{nn}, j_m, \ldots, j_0 \in \mathbb{N}$. В силу леммы 4.5.4 о специализации

$$F(a_{11}, \ldots, a_{nn}, b_m, \ldots, b_0) = G(a_{11}, \ldots, a_{nn}, b_m, \ldots, b_0).$$

В частности,

$$R\big(c_A(x), g(x)\big) = |g(A)|. \qquad \square$$

Примеры 4.14.27.

а) Пусть $R = \mathbb{Z}$, $g(x) = 3x + 2 \in \mathbb{Z}[x]$,

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z}).$$

Тогда $c_A(x) = x^2 - 4x - 1$;

$$R\big(c_A(x), g(x)\big) = \begin{vmatrix} 1 & -4 & -1 \\ 3 & 2 & 0 \\ 0 & 3 & 2 \end{vmatrix} = 19;$$

$$g(A) = 3A + 2E_2 = \begin{pmatrix} 5 & 6 \\ 6 & 11 \end{pmatrix}; |g(A)| = 19 = R\big(c_A(x), g(x)\big).$$

б) $R = \mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$, $g(x) = 3x + 2 \in \mathbb{Z}_6[x]$,

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z}_6).$$

Тогда $c_A(x) = x^2 + 2x + 5$;

$$R(c_A(x), g(x)) = \begin{vmatrix} 1 & 2 & 5 \\ 3 & 2 & 0 \\ 0 & 3 & 2 \end{vmatrix} = 1;$$

$$g(A) = 3A + 2E_2 = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}; \quad |g(A)| = 1 = R(c_A(x), g(x)).$$

Результант и решения систем алгебраических уравнений

Рассмотрим систему алгебраических уравнений над полем K

$$f_i(x_1, \ldots, x_n) = 0, \quad 1 \leqslant i \leqslant m,$$

где $f_i(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$. *Решением* этой системы уравнений называется такая строка $(\alpha_1, \ldots, \alpha_n) \in K^n$, что $f_i(\alpha_1, \ldots, \alpha_n) = 0$ для всех $1 \leqslant i \leqslant m$.

Отметим (см. следствие ??), что если $m = 2$, $\deg_{x_1} f_1 > 0$, $\deg_{x_1} f_2 > 0$, то

$$R_{x_1}(f_1, f_2) \in \langle f_1, f_2 \rangle \cap K[x_2, \ldots, x_n]$$

(здесь: $\langle f_1, f_2 \rangle$ — идеал, порождённый многочленами f_1 и f_2 в $K[x_1, \ldots, x_n]$; идеал $\langle f_1, f_2 \rangle \cap K[x_2, \ldots, x_n]$ кольца $K[x_2, \ldots, x_n]$ называется *первым исключающим идеалом* идеала $\langle f_1, f_2 \rangle$) и

$$R_{x_1}(f_1, f_2) = 0$$

тогда и только тогда, когда f_1 и f_2 имеют общий множитель в кольце $K[x_1, \ldots, x_n]$ положительной степени по x_1.

Рассмотрим подробнее систему двух алгебраических уравнений с двумя неизвестными:

$$f(x, y) = 0, \quad g(x, y) = 0, \quad f(x, y), g(x, y) \in K[x, y], \tag{4.9}$$

K — поле, $m = \deg f(x, y)$, $n = \deg g(x, y)$. Тогда

$$\begin{gathered} F(x) = f(x, y) = a_k(y)x^k + \ldots + a_0(y), \quad a_k(y) \neq 0, \\ G(x) = g(x, y) = b_l(y)x^l + \ldots + b_0(y), \quad b_l(y) \neq 0, \\ a_i(y), b_i(y) \in K[y]. \end{gathered}$$

Таким образом,

$$F(x), G(x) \in K(y)[x],$$

где $K(y) = Q(K[y])$ — поле рациональных функций от y. Тогда

$$R(F, G) = \Phi(y) \in K[y].$$

Пусть система алгебраических уравнений (4.9) удовлетворяет значениям $x = \alpha$, $y = \beta$ из некоторого расширения $\bar{K}$ поля K. Это означает, что многочлены

$$\bar{f}(x) = a_k(\beta)x^k + \ldots + a_0(\beta),$$
$$\bar{g}(x) = b_l(\beta)x^l + \ldots + b_0(\beta),$$

имеют общий корень α, и тогда

$$R(\bar{f}, \bar{g}) = \Phi(\beta) = 0.$$

Пусть теперь β — корень многочлена $\Phi(y)$, т. е. $\Phi(\beta) = 0$. Тогда

$$\Phi(\beta) = R(\bar{f}, \bar{g}) = 0,$$

и поэтому либо многочлены $\bar{f}(x)$, $\bar{g}(x)$ имеют общий корень, либо их старшие коэффициенты оба равны нулю: $a_k(\beta) = 0$, $b_l(\beta) = 0$.

Итак, решение системы (4.9) двух алгебраических уравнений с двумя неизвестными сведено к решению уравнения

$$\Phi(y) = 0 \tag{4.10}$$

от одного неизвестного y (в этом случае говорят, что мы *исключили неизвестное* x из системы уравнений (4.9)).

Поскольку результант (как многочлен Сильвестра) является однородным многочленом степени $m + n$ и изобарическим многочленом веса mn, то степень уравнения (4.10) не превосходит произведения mn.

Пример 4.14.28.

1) Пусть $K = \mathbb{C}$,

$$\begin{cases} f(x,y) = x^2y + x^2 + 2xy + y^3 = 0, \\ g(x,y) = x^2 - 6x - 3y^2 = 0. \end{cases}$$

Тогда $R_x(f,g) = y^4(4y+3)^2$, его корни: $y_1 = 0$, $y_2 = -\dfrac{3}{4}$.

Для многочленов

$$f(x,0) = x^2 \quad \text{и} \quad g(x,0) = x^2 - 6x$$

общий корень $x = 0$, поэтому получаем решение системы $(0,0)$.

Для многочленов

$$f\left(x, -\frac{3}{4}\right) = \frac{1}{4}x^2 - \frac{3}{2}x - \frac{27}{64} \quad \text{и} \quad g\left(x, -\frac{3}{4}\right) = x^2 - 6x - \frac{27}{16}$$

имеем общие корни $3 \pm \dfrac{3}{4}\sqrt{19}$, что даёт решения $\left(3 \pm \dfrac{3}{4}\sqrt{19}, -\dfrac{3}{4}\right)$.

Общий ответ: система имеет решения $(0,0)$, $\left(3 \pm \dfrac{3}{4}\sqrt{19}, -\dfrac{3}{4}\right)$.

2) Пусть $K = \mathbb{C}$,

$$\begin{cases} f(x,y) = xy - x + 2 = 0, \\ g(x,y) = xy - x + 3 = 0. \end{cases}$$

Тогда $R_x(f,g) = y - 1$. Подставляя $y = 1$ в систему, убеждаемся, что система не имеет решений (при этом старшие коэффициенты при x многочленов $f(x,y)$ и $g(x,y)$ одновременно становятся равными нулю).

3) Пусть $K = \mathbb{C}$,

$$\begin{cases} f(x,y) = x^2 + xy + x, \\ g(x,y) = xy + y^2 + y, \end{cases}$$

$R_x(f,g) \equiv 0$ (многочлены f и g имеют общий множитель $x + y + 1$). В этом случае рассматриваемая система имеет бесконечную серию решений $\{(-1-\alpha, \alpha) \mid \alpha \in \mathbb{C}\}$, а также решение $(0,0)$.

Алгебраические кривые или римановы поверхности

Как мы сейчас увидим, многие интересные алгебраические многообразия в $\mathbb{C}^2$ определяются нулями одного многочлена $f(x,y) \in \mathbb{C}[x,y]$. Эти многообразия называются *алгебраическими кривыми* или *римановыми поверхностями*, они обладают весьма тонким строением.

Замечание 4.14.29. Римановы поверхности двумерны, использование названия алгебраической кривой было продиктовано желанием подчеркнуть тот факт, что эти многообразия могут быть описаны аналитически с помощью одного комплексного параметра.

Предложение 4.14.30 (об общем виде алгебраических многообразий в $\mathbb{C}^2$). *Пусть $f(x,y), g(x,y) \in \mathbb{C}[x,y]$ — ненулевые многочлены, не имеющие общего множителя, отличного от константы. Тогда имеется лишь конечное множество общих нулей многочленов f и g (более точное утверждение — оценка Безу: если $m = \deg f$, $n = \deg g$, то число общих нулей ограничено числом mn).*

Доказательство. Действительно, в кольце главных идеалов $\mathbb{C}(x)[y]$, где $\mathbb{C}(x) = Q(\mathbb{C}[x])$ — поле рациональных функций от переменной x, пусть $I = (f,g)$ — идеал, порождённый многочленами f и g. Тогда $I = (f,g) = (h)$, где $h = \text{НОД}(f,g)$ в $\mathbb{C}(x)[y]$. В силу нашего предположения $h = 1$, и поэтому $I = (f,g) = \mathbb{C}(x)[y]$ (здесь мы используем то, что если f и g имеют общий множитель в $\mathbb{C}[x,y]$, не являющийся элементом из $\mathbb{C}(x)$, то многочлены f и g имеют общий множитель в $\mathbb{C}[x,y]$, отличный от константы, см. факториальность кольца многочленов многих переменных), поэтому $1 = rf + sg$, где $r, s \in \mathbb{C}(x)[y]$. Знаменатели многочленов r и s являются многочленами от x. Поэтому найдётся многочлен $p(x) \in \mathbb{C}[x]$ такой, что

$$p(x) = u(x,y)f(x,y) + v(x,y)g(x,y),$$

где $u, v \in \mathbb{C}[x,y]$.

Отсюда следует, что для общего нуля $a = (a_1, a_2) \in \mathbb{C}^2$ многочленов $f(x,y)$ и $g(x,y)$ a_1 является корнем многочлена $p(x) \in \mathbb{C}[x]$, имеющего лишь конечное число корней в $\mathbb{C}$.

Аналогичное соображение имеет место для переменной y. Таким образом, множество общих нулей для f и g конечно. □

Предложение 4.14.31. *Пусть*

$$f(x,y) = u_n(x)y^n + \ldots + u_1(x)y + u_0(x) \in \mathbb{C}[x,y] -$$

неприводимый многочлен, где $u_i(x) \in \mathbb{C}[x]$, *при этом* $u_n(x) \neq 0$, $n = \deg_y f \geqslant 1$, $S \subseteq \mathbb{C}^2$ — *множество нулей многочлена* $f(x,y)$. *Тогда*:

а) *для каждого значения* $a \in \mathbb{C}$ *переменной* x *существуют не более* n *точек в* S, x-*координата которых равна* a;

б) *существует конечное множество* Δ *значений переменной* x *таких, что если* $a \notin \Delta$, *то имеется ровно* n *точек в* S, x-*координата которых равна* a.

Доказательство. Пусть $a \in \mathbb{C}$, рассмотрим многочлен $f(a,y) \in \mathbb{C}[y]$. Тогда включение $(a,b) \in S$ равносильно тому, что $b \in \mathbb{C}$ является корнем многочлена $f(a,y)$.

Заметим, что $f(a,y) \neq 0$ в $\mathbb{C}[y]$, поскольку если бы $f(a,y) = 0$, то все коэффициенты $u_i(x)$ делились бы на $x - a$ в $\mathbb{C}[x]$, и тогда многочлен $f(x,y)$ также делился бы на $x - a$, но он неприводим в силу предположения.

Так как $\deg f(a,y) \leqslant n$, то $f(a,y)$ имеет не более n корней в $\mathbb{C}$. Многочлен $f(a,y)$ может иметь менее чем n корней, если:

а) либо $\deg f(a,y) < n$;

б) либо $\deg f(a,y) = n$, но $f(a,y)$ имеет кратный корень.

Случай а) имеет место, когда $u_n(a) = 0$, и таких значений $a \in \mathbb{C}$ конечно.

Случай б) для многочлена $h(y) = f(a,y) \in \mathbb{C}[y]$ имеет место, когда $h(y)$ и $h'(y) = \partial f/\partial y$ имеют общий корень.

Так как $f(x,y)$ неприводим и $\deg_y \partial f/\partial y < \deg f_y = n$, то f и $\partial f/\partial y$ не имеют общих множителей, отличных от констант. В силу предложения 4.14.30 имеется лишь конечное число общих нулей. □

Замечание 4.14.32. В ситуации нашего предположения обычно говорят, что S является *n-листным накрытием комплексной x-плоскости* P (см. ??). Поскольку существует конечное множество Δ, над которым S имеет менее, чем n листов, то говорят о *разветвлённом накрытии*. Итак, риманова поверхность неприводимого многочлена $f(x,y) \in \mathbb{C}[x,y]$ при $\deg_y f = n > 0$ является n-листным разветвлённым накрытием комплексной плоскости.

Пример 4.14.33. Пусть $f(x,y) = x^2 + xy^2 - 1 = 0$. Если $x \notin \{0, 1, -1\}$, то для каждого из таких x имеется два решения для y; если $x = 0$, то нет решений для y; если $x = 1$ или если $x = -1$, то имеется ровно по одному решению для y. □

4.15. Мономиальные идеалы в $K[x_1, \ldots, x_n]$

Упорядочения мономов в $K[x_1, \ldots, x_n]$

Упорядочение (отношение (частичного или линейного) порядка) на множестве мономов

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \in K[x_1, \ldots, x_n], \quad \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+^n, \quad \alpha_i \geqslant 0,$$

играет ключевую роль во всех алгоритмах с многочленами:

1) $\ldots > x^{m+1} > x^m > \ldots > x^2 > x > 1$ в алгоритме деления многочленов от одной переменной;

2) $x_1 > x_2 > \ldots > x_n$ в алгоритме приведения матрицы к ступенчатому виду;

3) лексикографический порядок в алгоритме представления симметрических многочленов: учитывая биекцию между $\alpha \leftrightarrow x^\alpha$, упорядочение можно рассматривать в $\mathbb{Z}_+^n$ ($\alpha > \beta$ в $\mathbb{Z}_+^n \iff x^\alpha > x^\beta$).

Назовём упорядочение $>$ на $\mathbb{Z}_+^n$ *мономиальным*, если:

1) $>$ является линейным порядком на $\mathbb{Z}_+^n$;

2) если $\alpha > \beta$ и $\gamma \in \mathbb{Z}_+^n$, то $\alpha + \gamma > \beta + \gamma$ (если $x^\alpha > x^\beta$ и $x^\gamma \in K[x_1, \ldots, x_n]$, то $x^\alpha x^\gamma > x^\beta x^\gamma$);

3) $(\mathbb{Z}_+^n, >)$ — вполне упорядоченное множество (т. е. любое непустое подмножество в $\mathbb{Z}_+^n$ имеет наименьший элемент).

Замечание 4.15.1. Условие 3) равносильно условию обрыва строго убывающих последовательностей в $(\mathbb{Z}_+^n, >)$ (см. лемма 4.15.2), а также равносильно тому, что $\alpha \geqslant 0$ для всех $\alpha \in \mathbb{Z}_+^n$ (см. следствие 4.15.8). В список порядков с условиями 1), 2), 3) входят следующие порядки:

а) лексикографические упорядочения $>_{\text{lex}}$; если упорядочение $<_{\text{lex}}$ порождено порядком $x_1 > x_2 > \ldots > x_n$, то

lex

$$\alpha = (\alpha_1, \ldots, \alpha_n) >_{\text{lex}} \beta = (\beta_1, \ldots, \beta_n),$$

если самая левая ненулевая координата строчки $\alpha - \beta \in \mathbb{Z}^n$ положительна ($x^\alpha >_{\text{lex}} x^\beta$ означает $\alpha >_{\text{lex}} \beta$);

deglex

б) градуированное лексикографическое упорядочение deglex (учитывающее сначала степени мономов): если $\alpha, \beta \in \mathbb{Z}_+^n$, то

$$\alpha >_{\text{deglex}} \beta$$

означает, что

$$|\alpha| = \sum_{i=1}^n \alpha_i > |\beta| = \sum_{i=1}^n \beta_i$$

или

$$|\alpha| = |\beta| \quad \text{и} \quad \alpha >_{\text{lex}} \beta;$$

degrevlex

в) градуированное обратное лексикографическое упорядочение degrevlex: $x_1 > \ldots x_n$,

$$\alpha >_{\text{degrevlex}} \beta,$$

если

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i$$

или $|\alpha| = |\beta|$ и самая правая ненулевая координата строчки $\alpha - \beta \in \mathbb{Z}^n$ отрицательна.

Отметим, что упорядочения deglex, degrevlex, как и упорядочение lex, зависят от порядка на переменных $x_1, \ldots, x_n$.

Лемма 4.15.2. *Упорядочение $>$ на непустом множестве M вполне упорядочивает это множество тогда и только тогда, когда каждая строго убывающая последовательность элементов из M обрывается.*

Доказательство. Действительно, если множество $(M, >)$ не является вполне упорядоченным, то существует непустое подмножество $S \subseteq M$, не имеющее минимального элемента. Пусть $s_1 \in M$. Так как s_1 не является минимальным элементом в S, то существует такой элемент $s_2 \in S$, что $s_1 > s_2$, элемент s_2 не является минимальным в S. Продолжая этот процесс, получаем бесконечную убывающую последовательность $s_1 > s_2 > s_3 > \ldots$, $s_i \in S$, $i = 1, 2, \ldots$.

Обратно, если существует бесконечная строго убывающая последовательность $S = \{\alpha_i \mid i \in \mathbb{N}\} \subseteq M$, $\alpha_1 > \alpha_2 > \ldots$, то непустое подмножество S в M не имеет минимального элемента. □

Мономиальные идеалы

Определение 4.15.3. Идеал I кольца $K[x_1, \ldots, x_n]$ называвается мономиальным, если существует такое подмножество $A \subseteq \mathbb{Z}_+^n$, что идеал I порождается множеством $\{x^\alpha,\ \alpha \in A\}$. Таким образом, любой элемент идеала I является конечной суммой $\sum\limits_{\alpha \in A} h_\alpha x^\alpha$, где $h_\alpha \in K[x_1, \ldots, x_n]$.

Лемма 4.15.4. *Пусть $I = \langle x^\alpha,\ \alpha \in A\rangle$ — мономиальный идеал кольца $K[x_1, \ldots, x_n]$. Тогда моном x^β лежит в идеале I тогда и только тогда, когда x^β делится на некоторый моном x^α, $\alpha \in A$.*

Доказательство. Если $x^\beta \in I$, то $x^\beta = \sum\limits_{j=1}^{n} h_j x^{\alpha_j}$, где $h_j \in K[x_1, \ldots, x_n]$, $\alpha_j \in A$, $j = 1, \ldots, n$. Записывая каждый многочлен h_j в виде линейной комбинации мономов и приводя подобные члены в рассматриваемой сумме, получаем, что существует такое i, $1 \leqslant i \leqslant n$, что моном x^β делится на x^{α_i}.

Обратно, если x^β делится на x^α, где $\alpha \in A$, то очевидно, что $x^\beta \in I$. □

Лемма 4.15.5. *Пусть $I = \langle x^\alpha,\ \alpha \in A\rangle$ — мономиальный идеал кольца $K[x_1, \ldots, x_n]$. Тогда следующие условия эквивалентны:*

1) $f \in I$;

2) *каждый одночлен многочлена f лежит в идеале I;*

3) *f является K-линейной комбинацией мономов, лежащих в идеале I.*

Доказательство. 1) $\Longrightarrow$ 2), 1) $\Longrightarrow$ 3). Если $f \in I$, то $f = \sum\limits_{\alpha \in A} h_\alpha x^\alpha$, где $h_\alpha \in K[x_1, \ldots, x_n]$. Записывая каждый многочлен h_α в виде линейной комбинации мономов и приводя подобные члены в $\sum\limits_{\alpha \in A} h_\alpha x^\alpha$, получаем, что каждый одночлен многочлена f лежит в идеале I, при этом многочлен f является K-линейной комбинацией мономов, делящихся на x^α для некоторых $\alpha \in A$.

Импликация 3) $\Longrightarrow$ 2) $\Longrightarrow$ 1) очевидна. □

Следствие 4.15.6. *Два мономиальных идеала кольца $K[x_1, \ldots, x_n]$ совпадают тогда и только тогда, когда совпадают множества мономов, содержащихся в них.*

Теорема 4.15.7 (лемма Диксона). *Любой мономиальный идеал $I = \langle x^\alpha,\ \alpha \in A\rangle$ кольца $K[x_1, \ldots, x_n]$ имеет конечный мономиальный базис ($I = \langle x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_s}\rangle$, где $\alpha_1, \alpha_2, \ldots, \alpha_s \in A$).*

Доказательство проведём индукцией по n. При $n = 1$ идеал I кольца $K[x_1]$ порождён множеством $\{x_1^\alpha \mid \alpha \in A \subseteq \mathbb{Z}_+\}$. Пусть β — наименьший элемент в $\mathbb{Z}_+$. Тогда моном x_1^β делит все мономы x_1^α, $\alpha \in A$. Таким образом, $I = \langle x_1^\beta\rangle$.

Пусть теперь $n > 1$ и утверждение теоремы верно для $n - 1$. Мономы в $K[x_1, \ldots, x_n]$ будем записывать в виде $x^\alpha x_n^m$, где $\alpha \in \mathbb{Z}_+^{n-1}$, $m \in \mathbb{Z}_+$.

Для данного мономиального идеала I кольца $K[x_1, \ldots, x_n]$ рассмотрим мономиальный идеал J кольца $K[x_1, \ldots, x_{n-1}]$, порождённый множеством $\{x^\alpha \mid x^\alpha x_n^m \in I$ для некоторого $m \geqslant 0\}$. По предположению индукции $J = \langle x^{\beta_1}, \ldots, x^{\beta_t}\rangle$ для некоторых $\beta_1, \ldots, \beta_t \in \mathbb{Z}_+^{n-1}$. При этом для каждого i, $1 \leqslant i \leqslant t$, существует такое $m_i \geqslant 0$, что $x^{\beta_i} x^{m_i} \in I$. Пусть $m = \max\limits_{1 \leqslant i \leqslant t}\{m_i\}$. Для каждого l, $0 \geqslant l \geqslant m - 1$, рассмотрим мономиальный идеал $J_l = \langle x^\gamma \mid x^\gamma x_n^l \in I\rangle$ кольца $K[x_1, \ldots, x_n]$. По предположению индукции $J_l = \langle x^{\beta_{l_1}}, \ldots, x^{\beta_{l_{t_l}}}\rangle$ для некоторых $\beta_{l_1}, \ldots, \beta_{l_{t_l}} \in \mathbb{Z}_+^{n-1}$.

Покажем, что мономы

$$\{x^{\beta_1}, \ldots, x^{\beta_t}, x^{\beta_{l_1}}, \ldots, x^{\beta_{l_{t_l}}} \mid 0 \leqslant l \leqslant m - 1\} \tag{4.11}$$

порождают идеал I. Пусть $x^\alpha x_n^p \in I$. Если $p \geqslant m$, то по определению идеала J моном $x^\alpha x_n^p$ делится на некоторый моном $x^{\beta_i} x_n^m$. Если же $p \leqslant m - 1$, то по определению идеала J_p моном $x^\alpha x_n^p$ делится на некоторый моном $x^{\beta_{p_j}} x_n^p$.

По лемме 4.15.4 мономы (4.11) порождают идеал, содержащий те же мономы, которые содержит идеал I. По следствию 4.15.6 эти идеалы совпадают. □

Следствие 4.15.8. *Пусть $>$ — такое линейное упорядочение на $\mathbb{Z}_+^n$, что для $\alpha, \beta, \gamma \in \mathbb{Z}_+^n$, $\alpha > \beta$, имеем $\alpha + \gamma > \beta + \gamma$. Тогда $(\mathbb{Z}_+^n, >)$ является вполне упорядоченным множеством тогда и только тогда, когда $\alpha \geqslant 0$ для всех $\alpha \in \mathbb{Z}_+^n$.*

Доказательство. Пусть $(\mathbb{Z}_+^n, >)$ — вполне упорядоченное множество, β — наименьший элемент в $\mathbb{Z}_+^n$. Достаточно показать, что $\beta \geqslant 0$. Действительно, если $0 > \beta$, то $\beta + 0 > \beta + \beta = 2\beta \in \mathbb{Z}_+^n$, что противоречит тому, что β — наименьший элемент в $\mathbb{Z}_+^n$.

Пусть теперь $\alpha \geqslant 0$ для всех $\alpha \in \mathbb{Z}_+^n$ и A — непустое подмножество в $\mathbb{Z}_+^n$. Докажем, что в A существует наименьший элемент. Для этого рассмотрим мономиальный идеал

$I = \langle x^\alpha,\ \alpha \in A\rangle$ кольца $K[x_1, \ldots, x_n]$. По теореме 4.15.7 существуют такие элементы $\alpha_1, \ldots, \alpha_s$, что $I = \langle x^{\alpha_1}, \ldots, x^{\alpha_s}\rangle$. Можно считать, что $\alpha_1 < \ldots < \alpha_s$.

Пусть $\alpha \in A$. Тогда $x^\alpha \in I = \langle x^{\alpha_1}, \ldots, x^{\alpha_s}\rangle$. По лемме 4.15.4 моном x^α делится на некоторый моном x^{α_i}, где $1 \leqslant i \leqslant s$. Это означает, что $\alpha = \alpha_i + \gamma$, где $\gamma \in \mathbb{Z}_+^n$. По нашему предположению $\gamma \geqslant 0$. Из свойства порядка $>$ имеем

$$\alpha = \alpha_i + \gamma \geqslant \alpha_i + 0 = \alpha_i \geqslant \alpha_1.$$

Таким образом, α_1 — наименьший элемент в A. □

Глава 5

Элементы коммутативной алгебры

5.1. Максимальные идеалы коммутативных колец (максимальный спектр MaxSpec)

Пусть R — коммутативное кольцо с 1. Идеал M кольца R называется *максимальным*, если $M \neq R$ и между M и R нет других идеалов кольца R (т. е. если $J \lhd R$, $M \subseteq J \subseteq R$, то либо $M = J$, либо $J = R$). Совокупность всех максимальных идеалов коммутативного кольца R обозначим через $\mathrm{MaxSpec}(R)$ (*максимальный спектр* кольца R).

Теорема 5.1.1. *Пусть R — коммутативное кольцо с 1. Каждый собственный идеал I кольца R содержится в некотором максимальном идеале M кольца R.*

Доказательство. Рассмотрим частично упорядоченное по включению множество $\mathcal{P}$ всех собственных идеалов кольца R, содержащих идеал I; $\mathcal{P} \neq \varnothing$, поскольку $I \in \mathcal{P}$. К нему применима лемма Цорна (??). Если $J_k \lhd R$, $J_k \neq R$, $k \in \mathbb{N}$,

$$I \subseteq J_1 \subseteq J_2 \subseteq \ldots \subseteq J_k \subseteq J_{k+1} \subseteq \ldots,$$

то для идеала $J = \bigcup\limits_k J_k$ имеем $I \subseteq J$; $J \neq R$ (поскольку если $J = R$, то $1 \in J$, и следовательно, $1 \in J_k$ для некоторого k, но тогда $J_k = R$, что противоречит нашему предположению). Итак, $J_k \subseteq J \in \mathcal{P}$ для всех k. В силу леммы Цорна (??) частично упорядоченное множество $\mathcal{P}$ содержит максимальный элемент M, который, из определения множества $\mathcal{P}$, является максимальным идеалом кольца R, при этом $I \subseteq M$. □

Следствие 5.1.2. *Если кольцо R не содержит максимальных идеалов, то $R = 0$ (нулевое кольцо).*

Лемма 5.1.3. *Пусть R — коммутативное кольцо с 1, M — идеал кольца R. Тогда M — максимальный идеал тогда и только тогда, когда фактор-кольцо R/M — поле.*

Доказательство. Фактор-кольцо $\bar{R} = R/M$ является полем тогда и только тогда, когда идеалы $\bar{I}$ кольца $\bar{R}$ исчерпываются идеалами $\bar{0}$ и $\bar{R}$. В силу соответствия между идеалами $\bar{I}$ кольца $\bar{R}$ и идеалами $I \lhd R$, $M \subseteq I \subseteq R$, получаем наше утверждение. □

Следствие 5.1.4. *Нулевой идеал в R максимальный тогда и только тогда, когда R — поле.* □

Следствие 5.1.5. *Если K — поле и $f\colon R \to K$ — сюръективный гомоморфизм, то* $\operatorname{Ker} f$ *— максимальный идеал кольца R и $R/\operatorname{Ker} \cong K$.* □

Следствие 5.1.6. *Идеал I кольца целых чисел $\mathbb{Z}$ максимальный тогда и только тогда, когда $I = \mathbb{Z}p$, где p — простое число.*

Доказательство. Идеал I имеет вид $I = \mathbb{Z}n$. Фактор-кольцо $\mathbb{Z}/I = \mathbb{Z}/\mathbb{Z}n = \mathbb{Z}_n$ является полем тогда и только тогда, когда n — простое число. □

Следствие 5.1.7. *Идеал I кольца многочленов $K[x]$ над полем K максимальный тогда и только тогда, когда $I = K[x]f(x)$, где $f(x)$ — неприводимый многочлен в $K[x]$.*

Доказательство. Каждый идеал в кольце $K[x]$ имеет вид $I = K[x]f(x)$, где $f(x) \in K[x]$.

Если $f(x)$ — неприводимый многочлен и $0 \neq g(x) + I \in K[x]/I$, то $g(x)$ не делится на $f(x)$. Поэтому НОД$\big(f(x), g(x)\big) = 1$, и следовательно, $1 = f(x)u(x) + g(x)v(x)$ для некоторых $u(x), v(x) \in K[x]$. Таким образом,

$$(g(x) + I)(v(x) + I) = g(x)v(x) + I = \big(1 - f(x)u(x)\big) + I = 1 + I.$$

Итак, R/I — поле, и следовательно I — максимальный идеал.

Если же $f(x) = f_1(x)f_2(x)$, где $f_1(x), f_2(x) \in K[x]$, $\deg f_1 < \deg f$, $\deg f_2 < \deg f$, то

$$0 \neq f_1(x) + I \in K[x]/I, \quad 0 \neq f_2(x) + I \in K[x]/I,$$
$$(f_1(x) + I)(f_2(x) + I) = f(x) + I = I.$$

Следовательно, в фактор-кольце R/I имеются делители нуля, поэтому R/I не является полем. □

Следствие 5.1.8. *Если K — алгебраически замкнутое поле (например, $K = \mathbb{C}$ — поле комплексных чисел), то максимальные идеалы в кольце $K[x]$ имеют вид $M_a = K[x](x-a)$, $a \in K$, при этом $M_a = \operatorname{Ker}(s_a)$, где $s_a\colon K[x] \to K$, $s_a\big(\varphi(x)\big) = \varphi(a)$. Соответствие $a \mapsto M_a$ осуществляет биекцию между элементами a алгебраически замкнутого поля K и максимальными идеалами $M_a = K[x](x-a)$ кольца $K[x]$.*

Доказательство. Так как над алгебраически замкнутым полем K неприводимый многочлен $f(x) \in K[x]$ со старшим коэффициентом 1 имеет вид $x - a$, $a \in K$, то в силу следствия 5.1.7 все максимальные идеалы кольца $K[x]$ имеют вид $M_a = K[x](x-a)$, $a \in K$.

Пусть $\varphi(x) \in K[x]$, тогда $\varphi(a) \in \operatorname{Ker}(s_a)$, если $s_a\big(\varphi(x)\big) = \varphi(a) = 0$, и поэтому это равносильно тому, что $\varphi(x) = q(x)(x-a) \in M_a$. Итак, $\operatorname{Ker}(s_a) = M_a$. □

5.2. Радикал Джекобсона (квазирегулярный идеал) коммутативного кольца с 1

Радикалом Джекобсона (или *квазирегулярным радикалом*) коммутативного кольца R с единицей называется пересечение всех его максимальных идеалов M:

$$J(R) = \bigcap_{M \lhd R} M.$$

Теорема 5.2.1. *Пусть R — коммутативное кольцо с 1. Тогда:*

1) $J(R) \lhd R$, $J(R/J(R)) = 0$;

2) $J(R) = \{x \in R \mid 1 - xy \in U(R)\ \forall y \in R\}$.

Доказательство. 1) Ясно, что $J(R) = \bigcap\limits_{M \lhd R} M \lhd R$. Любой максимальный идеал кольца $\bar{R} = R/J(R)$ имеет вид $\bar{M} = M/J(R)$, где M — максимальный идеал кольца R, содержащий $J(R)$, то есть любой максимальный идеал кольца R. Поэтому $J(R/J(R)) = 0$.

2а) Если $1 - xy \notin U(R)$, то $1 - xy \in M$ для некоторого максимального идеала M кольца R. Так как $x \in J(R) \subseteq M$, то $xy \in M$, и поэтому $1 \in M$, что приводит нас к противоречию. Итак,

$$1 - xy \in U(R) \text{ для всех } y \in R.$$

2б) Если $x \notin M$ для некоторого максимального идеала кольца R, то $M + Rx = R$, поэтому $u + xy = 1$ для некоторых $u \in M$, $y \in R$, следовательно, $1 - xy = u \in M$, и поэтому $1 - xy \notin U(R)$. □

5.3. Простые идеалы коммутативных колец

Пусть R — коммутативное кольцо с 1, идеал I кольца R называется *простым*, если $I \neq R$ и для $x, y \in R$ из $xy \in I$ следует, что или $x \in I$, или $y \in I$ (это равносильно тому, что фактор-кольцо R/I не содержит делителей нуля).

Замечание 5.3.1.

1) Нулевой идеал коммутативного кольца R является простым тогда и только тогда, когда R — кольцо без делителей нуля.

2) Каждый максимальный идеал M коммутативного кольца R является простым (действительно, R/M — поле, и, следовательно, кольцо без делителей нуля). Если через $\operatorname{Spec} R$ обозначить совокупность (спектр) простых идеалов, то

$$\operatorname{MaxSpec} R \subseteq \operatorname{Spec} R.$$

Лемма 5.3.2. *Если P — простой идеал коммутативного кольца R с 1, x — нильпотентный элемент кольца R, то $x \in P$.*

Доказательство. Так как $x^m = 0 \in P$ для некоторого $m \in \mathbb{N}$ и P — простой идеал, то $x \in P$. □

Лемма 5.3.3. *Идеал I коммутативного кольца R является простым тогда и только тогда, когда для любых идеалов J_1 и J_2 кольца R из $J_1J_2 \subseteq I$ следует, что или $J_1 \subseteq I$, или $J_2 \subseteq I$.*

Доказательство. 1) Пусть I — простой идеал. Допустим противное: $J_1 \not\subseteq I$, $J_2 \not\subseteq I$, но $J_1J_2 \subseteq I$. Пусть $j_k \in J_k$, $j_k \notin I$, $k = 1, 2$. Тогда $j_1j_2 \in I$, но $j_1 \notin I$, $j_2 \notin I$, что противоречит простоте идеала I.

2) Пусть для любых идеалов $J_1 \lhd R$, $J_2 \lhd R$ из $J_1J_2 \subseteq I$ следует, что или $J_1 \subseteq I$, или $J_2 \subseteq I$. Если $x, y \in R$ и $xy \in I$, то $RxRy \subseteq Rxy \subseteq I$, поэтому или $x \in Rx \subseteq I$, или $y \in Ry \subseteq I$. □

Замечание 5.3.4. Это соображение явилось основополагающим для введения некоммутативных аналогов простых идеалов и областей целостности — первичных идеалов и первичных колец.

Предложение 5.3.5. *Если R — коммутативное кольцо с 1, P — максимальный элемент частично упорядоченного множества идеалов, не являющихся конечно порождёнными, то P — простой идеал.*

Доказательство. Допустим противное: существуют $a, b \in R$, $ab \in P$, $a \notin P$, $b \notin P$. Тогда $P < P + Ra$, $P < P + Rb$, поэтому $P + Ra$, $P + Rb$ — конечно порождённые идеалы, скажем

$$P + Ra = \sum_{i=1}^{m} R(p_i + r_i a), \quad P + Rb = \sum_{j=1}^{n} R(p'_j + r'_j b),$$

где $p_i, p'_i \in P$, $r_i, r'_i \in R$.

Рассмотрим идеал

$$J = \{r \in R \mid ra \in P\} \lhd R.$$

Так как $ab \in P$, то

$$(p'_j + r'_j b)a = p'_j + r'_j ab \in P \quad \text{для всех} \quad j, \quad 1 \leqslant j \leqslant n,$$

и поэтому

$$P < P + Rb \subseteq J.$$

В силу максимальности идеала P идеал J конечно порождён. Пусть

$$J = R_{j_1} + \ldots + R_{j_k}, \quad j_1, \ldots, j_k \in R.$$

Если $x \in P$, то $x \in P \subseteq P + Ra$, поэтому для подходящих $s_i \in R$ имеем:

$$x = \sum_{s=1}^{m} s_i(p_i + r_i a) = \sum_{i=1}^{m} s_i p_i + \left(\sum_{i=1}^{m} s_i r_i\right) a.$$

Следовательно,

$$\left(\sum_{i=1}^{m} s_i r_i\right) a = x - \sum_{i=1}^{m} s_i p_i \in P,$$

и поэтому

$$\sum_{i=1}^{m} s_i r_i \in J.$$

Тогда для некоторых $t_i \in R$:

$$\sum_{i=1}^{m} s_i r_i = \sum_{i=1}^{k} t_i j_i,$$

и следовательно,

$$x = \sum_{s=1}^{m} s_i p_i + \sum_{i=1}^{k} t_i (j_i a).$$

Таким образом, идеал P порождается конечной системой элементов

$$\{p_1, \ldots, p_m, j_1 a, \ldots, j_k a\},$$

что приводит нас к противоречию. Итак, из $ab \in P$ следует, что либо $a \in P$, либо $b \in P$, то есть P — простой идеал. □

5.4. Мультипликативно замкнутые системы элементов коммутативного кольца

Подмножество $S \subseteq R$ коммутативного кольца R с 1 называется *мультипликативно замкнутым*, если $1 \in S$ и $xy \in S$ для всех $x, y \in S$ (другими словами, $(S, \cdot, 1)$ — моноид).

Лемма 5.4.1. *Пусть I — собственный идеал коммутативного кольца R с 1. Тогда I — простой идеал в том и только в том случае, когда $I^c = R \setminus I$ является мультипликативно замкнутой системой.*

Доказательство. Ясно, что $1 \in I^c$.

1) Если I — простой идеал, $x, y \notin I$, то $xy \notin I$. Итак, $I^c = R \setminus I$ — мультипликативно замкнутая система.

2) Если I^c — мультипликативно замкнутая система, $x \in I^c$ и $y \in I^c$, то $xy \in I^c$. Итак, I — простой идеал. □

Замечание 5.4.2.

1) $S = \{1\}$ — мультипликативно замкнутая система.

2) Если $x \in R$, то $S = \{x^i \mid i \geqslant 0\}$ — мультипликативно замкнутая система.

3) Если $I \lhd R$, то $S = 1 + I = \{1 + x \mid x \in I\}$ — мультипликативно замкнутая система.

4) Множество неделителей нуля коммутативного кольца R с 1 образует мультипликативно замкнутое подмножество.

Теорема 5.4.3. *Пусть S — мультипликативно замкнутое подмножество коммутативного кольца R с 1, I — идеал кольца R, лежащий в $S^c = R \setminus S$. Тогда:*

1) *найдётся идеал $J \lhd R$ такой, что $J \subseteq S^c$, $I \subseteq J$, идеал J максимален относительно S (это означает, что J — максимальный по включению идеал, содержащийся а S^c);*

2) *любой такой идеал J — простой идеал кольца R.*

Доказательство. 1) Пусть

$$\mathcal{L} = \{J \lhd R \mid I \subseteq J \subseteq S^c\}.$$

Множество $\mathcal{L}$ непустое, поскольку $I \in \mathcal{L}$. Если $\Gamma = \{I_i \mid i \in I\}$ — цепь в $(\mathcal{L}, \subseteq)$, то $\bigcup\limits_{i \in I} I_i$ — верхняя грань для цепи Γ в $\mathcal{L}$. Применение леммы Цорна даёт существование максимального элемента J в $(\mathcal{L}, \subseteq)$.

2) Допустим, что идеал J не является простым, пусть $x, y \in J^c$, $xy \in J$. Так как $x \notin J$, то $Rx + J > J$, и поэтому $(Rx + J) \cap S \neq \varnothing$, пусть $s \in (Rx + J) \cap S$. Аналогично, пусть $s' \in (Ry + J) \cap S$. Тогда $ss' \in S$, поскольку подмножество S мультипликативно замкнуто. Так как

$$ss' \in (Rx + J)(Ry + J) \subseteq Rxy + J \subseteq J,$$

то $ss' \in S \cap J$. Но это невозможно, поскольку $J \subseteq S^c$.

Итак, J — простой идеал кольца R. □

Следствие 5.4.4. *Если элемент $x \in R$ не является нильпотентным, то существует простой идеал кольца R, не содержащий элемента x.*

Доказательство. Рассмотрим мультипликативную систему $S = \{x^i \mid i \in \mathbb{N} \cup \{0\}\}$, $I = 0$. Построенный простой идеал $P = J$ таков, что $P \cap S = \varnothing$. Итак, $x \notin P$. □

5.5. Первичный радикал (нильрадикал) коммутативного кольца с 1

Теорема 5.5.1. *Пусть R — коммутативное кольцо с 1. Тогда:*

1) *множество*
$$\operatorname{rad} R = \{r \in R \mid r^n = 0 \text{ для некоторого } n \in \mathbb{N}\}$$
всех нильпотентных элементов кольца R является идеалом кольца R;

2) $\operatorname{rad}(R/\operatorname{rad} R) = 0$ (*другими словами, в фактор-кольце $R/\operatorname{rad} R$ нет ненулевых нильпотентных элементов*).

Доказательство. 1а) Если $x, y \in \operatorname{rad} R$, $x^m = 0$, $y^n = 0$, $m, n \in \mathbb{N}$, то
$$(x+y)^{m+n-1} = \sum_{r+s=m+n-1} C^r_{m+n-1} x^r y^s.$$
Так как одновременно $r < m$, $s < n$ невозможно (тогда $r+s \leqslant m-1+n-1 = (m+n-1)-1$), то или $x^r = 0$, или $y^s = 0$, и поэтому $(x+y)^{m+n-1} = 0$. Следовательно, $x + y \in \operatorname{rad} R$. Ясно, что $(-x)^m = (-1)^m x^m = 0$, поэтому $-x \in \operatorname{rad} R$. Таким образом, $\operatorname{rad} R$ — подгруппа группы $(R, +)$.

1б) Если $x \in \operatorname{rad} R$ и $r \in R$, то $x^m = 0$ для $m \in \mathbb{N}$, поэтому
$$(rx)^m = r^m x^m = r^m \cdot 0 = 0,$$
следовательно, $rx \in \operatorname{rad} R$. Итак, $\operatorname{rad} R$ — идеал кольца R.

2) Если $x + \operatorname{rad} R \in \operatorname{rad}(R/\operatorname{rad} R)$, $x \in R$, то $(x + \operatorname{rad} R)^m = x^m + \operatorname{rad} R = \bar{0}$ в $R/\operatorname{rad} R$. Поэтому $x^m \in \operatorname{rad} R$, следовательно, $(x^m)^n = 0$ для некоторого $n \in \mathbb{N}$. Так как $x^{mn} = 0$ для $mn \in \mathbb{N}$, то $x \in \operatorname{rad} R$, и поэтому $x + \operatorname{rad} R = \bar{0}$ в $R/\operatorname{rad} R$. □

Идеал $\operatorname{rad} R$ коммутативного кольца R называется *нильрадикалом* (или *первичным радикалом*) кольца R.

Теорема 5.5.2. *Нильрадикал $\operatorname{rad} R$ коммутативного кольца R с 1 совпадает с пересечением всех простых идеалов кольца R.*

Доказательство. Пусть $T = \bigcap P$ — пересечение всех простых идеалов P кольца R.

1) Если $x \in \operatorname{rad} R$, то $x^m = 0 \in P$, и поэтому $x \in P$ для любого простого идеала P. Следовательно, $\operatorname{rad} R \subseteq T = \bigcap P$.

2) Если элемент $x \notin \operatorname{rad} R$, то элемент x не является нильпотентным, поэтому (следствие 5.4.4) существует простой идеал P кольца R такой, что $x \notin P$, и следовательно, $x \notin T = \bigcap P$. Таким образом, $T \subseteq \operatorname{rad} R$.

Итак, $\operatorname{rad} R = T$. □

5.6. Радикал идеала коммутативного кольца с 1

Пусть I — идеал кольца R, *радикалом идеала* I назовём следующее подмножество в R:
$$\sqrt{I} = \{x \in R \mid x^n \in I \text{ для некоторого } n \geqslant 1\}.$$

Теорема 5.6.1. *Пусть R — коммутативное кольцо с 1, $I \lhd R$. Тогда:*

1) *$\sqrt{I}$ — идеал кольца R, $I \subseteq \sqrt{I}$;*

2) *радикал $\sqrt{I}$ идеала I совпадает с пересечением всех простых идеалов кольца R, содержащих идеал I.*

Доказательство. 1) Пусть $\varphi\colon R \to R/I$ — канонический сюръективный гомоморфизм. Тогда

$$\sqrt{I} = \varphi^{-1}(\operatorname{rad} R/I),$$

и поэтому радикал $\sqrt{I}$ идеала I является идеалом. Ясно, что $I = \varphi^{-1}(0) \subseteq \varphi^{-1}(\operatorname{rad} R/I) = \sqrt{I}$.

2) Так как $\operatorname{rad}(R/I) = \bigcap \bar{P}$ — пересечение всех простых идеалов $\bar{P}$ фактор-кольца $\bar{R} = R/I$. Поэтому $\sqrt{I} = \varphi^{-1}(\operatorname{rad} R/I) = \bigcap\limits_{I \subseteq P} P$ (пересечение всех простых идеалов P кольца R, содержащих идеал I). □

Следствие 5.6.2. $\sqrt{\sqrt{I}} = \sqrt{I}$.

5.7. Теорема Гильберта о нулях над полем $\mathbb{C}$ комплексных чисел

Наша цель — описать максимальные идеалы в кольце многочленов $\mathbb{C}[x_1, \ldots, x_n]$ от переменных $x_1, \ldots, x_n$ над полем $\mathbb{C}$ комплексных чисел, расширяя только что рассмотренный случай одной переменной.

Если $a = (a_1, \ldots, a_n) \in \mathbb{C}_n$, то рассмотрим гомоморфизм

$$s_a\colon \mathbb{C}[x_1, \ldots, x_n] \to \mathbb{C},$$

где $s_a\big(f(x_1, \ldots, x_n)\big) = f(a_1, \ldots, a_n)$.

Положим

$$M_a = \operatorname{Ker}(s_a) = \{f(x_1, \ldots, x_n) \in \mathbb{C}[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0\}.$$

Теорема 5.7.1 (Гильберта о нулях, 1893 г.). *Максимальные идеалы кольца многочленов $\mathbb{C}[x_1, \ldots, x_n]$ имеют вид $M_a = \operatorname{Ker} s_a$, где $a = (a_1, \ldots, a_n) \in \mathbb{C}_n$; идеал M_a порождён линейными одночленами $x_1 - a_1, \ldots, x_n - a_n$; соответствие $a \mapsto M_a$ является биекцией: $\mathbb{C}_n \to \operatorname{MaxSpec}(\mathbb{C}[x_1, \ldots, x_n])$.*

Доказательство. 1) Пусть $a = (a_1, \ldots, a_n) \in \mathbb{C}_n$. Так как $\mathbb{C}$ — поле и $s_a\colon \mathbb{C}[x_1, \ldots, x_n] \to \mathbb{C}$ — сюръективный гомоморфизм, то $M_a = \operatorname{Ker} s_a$ — максимальный идеал кольца $\mathbb{C}[x_1, \ldots, x_n]$.

2) Покажем, что идеал M_a порождается одночленами $x_1 - a_1, \ldots, x_n - a_n$.

Пусть $f(x_1, \ldots, x_n) \in M_a$. Сделаем замену переменных в $f(x_1, \ldots, x_n)$

$$(x_1, \ldots, x_n) = (a_1, \ldots, a_n) + (u_1, \ldots, u_n),$$

разложим многочлен $f(x_1, \ldots, x_n) = g(u_1, \ldots, u_n)$ по степеням переменных $u_1, \ldots, u_n$, после чего сделаем обратную замену

$$(u_1, \ldots, u_n) = (x_1, \ldots, x_n) - (a_1, \ldots, a_n)$$

в $g(u_1,\dots,u_n)=f(x_1,\dots,x_n)$, в итоге получим

$$f(x_1,\dots,x_n)=f(a_1,\dots,a_n)+\sum_i c_i(x-a_i)+\sum_{i,j} c_{ij}(x_i-a_i)(x_j-a_j)+\dots.$$

Так как $M_a=\operatorname{Ker} s_a$, то $f(a_1,\dots,a_n)=0$, и поэтому многочлен $f(x_1,\dots,x_n)$ лежит в идеале, порождённом одночленами $x-a_1,\dots,x-a_n$. Это показывает, что идеал M_a порождается одночленами $x-a_i$, $1\leqslant i\leqslant n$.

3) Пусть M — максимальный идеал кольца $\mathbb{C}[x_1,\dots,x_n]$, тогда $K=\mathbb{C}[x_1,\dots,x_n]/M$ — поле, канонический гомоморфизм

$$\pi\colon \mathbb{C}[x_1,\dots,x_n]\to K$$

сюръективен. Рассмотрим ограничение гомоморфизма π на подкольцо $\mathbb{C}[x_1]$ кольца $\mathbb{C}[x_1,\dots,x_n]$:

$$\pi_1=\pi|_{\mathbb{C}[x_1]}\colon \mathbb{C}[x_1]\to K.$$

а) Покажем, что $\operatorname{Ker}\pi_1\neq 0$. Действительно, если $\operatorname{Ker}\pi_1=0$, то π_1 — инъективный гомоморфизм, $\mathbb{C}[x_1]\cong\operatorname{Im}\pi_1\subseteq K$. В силу леммы ?? гомоморфизм π_1 единственным образом продолжается на поле частных $\mathbb{C}(x)$ кольца $\mathbb{C}[x]$. Итак, поле K содержит подполе, изоморфное полю рациональных функций $\mathbb{C}(x)$.

Так как мономы $x_1^{i_1}\dots x_n^{i_n}$ образуют счётный базис $\mathbb{C}$-линейного пространства $\mathbb{C}[x_1,\dots,x_n]$, то $\mathbb{C}$-линейное пространство $K=\mathbb{C}[x_1,\dots,x_n]/M$ имеет счётное семейство образующих. В то же время, это противоречит (см. ??) тому, что ${}_{\mathbb{C}}K$ содержит линейное пространство ${}_{\mathbb{C}}\mathbb{C}(x)$, содержащее несчётное линейно независимое множество рациональных функций $\dfrac{1}{x-\alpha}$, $\alpha\in\mathbb{C}$ (действительно, если $\alpha_1,\dots,\alpha_l\in\mathbb{C}$ различны, $c_1,\dots,c_l\in\mathbb{C}$, то из $\dfrac{c_1}{x-\alpha_1}+\dots+\dfrac{c_l}{x-\alpha_l}=0$ следует, в силу единственности разложения в простейшие дроби, что $c_1=c_2=\dots=c_l=0$).

Итак, мы показали, что $\operatorname{Ker}\pi_1\neq 0$.

4) Покажем теперь, что $\operatorname{Ker}\pi_1$ — максимальный идеал в кольце $\mathbb{C}[x_1]$. Пусть $0\neq f\in\operatorname{Ker}\pi_1$. Так как $K\neq 0$ и $\pi_1(1)=1_K$, то $\operatorname{Ker}\pi_1<\mathbb{C}[x_1]$, и поэтому f не является константой, следовательно, $f=(x_1-a_1)g$, $a_1\in\mathbb{C}$. Тогда в поле K имеем

$$\pi_1(x_1-a_1)\pi_1(g)=\pi_1(f)=0,$$

и поэтому либо $\pi_1(x_1-a_1)=0$, либо $\pi_1(g)=0$. Таким образом, один из двух элементов x_1-a_1 или g лежит в $\operatorname{Ker}\pi_1$, $\deg g<\deg f$. Проводя индукцию по степени $\deg f$, убеждаемся в том, что $\operatorname{Ker}\pi_1$ содержит линейный множитель x_1-a_1, $a_1\in\mathbb{C}$, и поэтому $\operatorname{Ker}\pi_1=M_{a_1}$ — максимальный идеал.

5) Рассматривая другой индекс i, $1\leqslant i\leqslant n$, вместо 1, получаем, что M содержит одночлен x_i-a_i. Отсюда следует, что идеал M содержится и, следовательно, равен ядру $\operatorname{Ker} s_a$, $a=(a_1,\dots,a_n)$, гомоморфизма

$$s_a\colon \mathbb{C}[x_1,\dots,x_n]\to\mathbb{C},\quad s_a\bigl(f(x_1,\dots,x_n)\bigr)=f(a_1,\dots,a_n).\qquad\square$$

5.8. Начала алгебраической геометрии

Пусть V — подмножество n-мерного комплексного пространства $\mathbb{C}^n$. Если V можно задать как множество общих нулей конечного числа многочленов от n переменных из $\mathbb{C}[x_1,\dots,x_n]$, то V называется *алгебраическим многообразием*.

Примеры 5.8.1.

1) Точка $(a, b) \in \mathbb{C}^2$ является множеством нулей системы

$$\begin{cases} x - a = 0, \\ y - b = 0. \end{cases}$$

2) Группа $\mathrm{SL}_2(\mathbb{C})$ является совокупностью решений полиномиального уравнения

$$x_{11}x_{22} - x_{12}x_{21} - 1 = 0,$$

и поэтому является алгебраическим многообразием в $\mathbb{C}^4$.

Теорема Гильберта о нулях осуществляет связь между алгеброй и геометрией, устанавливая соответствие между максимальными идеалами кольца многочленов $\mathbb{C}[x_1, \ldots, x_n]$ и точками пространства $\mathbb{C}^n$. Это соответствие можно использовать для нахождения связей алгебраических многообразий с фактор-кольцами колец многочленов.

Теорема 5.8.2. *Пусть $f_1, \ldots, f_r \in \mathbb{C}[x_1, \ldots, x_n]$, V — алгебраическое многообразие, задаваемое системой алгебраических уравнений $f_1(x_1, \ldots, x_n) = 0, \ldots, f_r(x_1, \ldots, x_n) = 0$, $I = (f_1, \ldots, f_r)$ — идеал кольца $\mathbb{C}[x_1, \ldots, x_n]$, порождённый многочленами $f_1, \ldots, f_r$. Тогда существует биекция между максимальными идеалами кольца $R = \mathbb{C}[x_1, \ldots, x_n]/I$ и точками многообразия V:*

$$\mathrm{MaxSpec}(\mathbb{C}[x_1, \ldots, x_n]/I) \to V.$$

Доказательство. В силу теорем об изоморфизме для колец, максимальные идеалы фактор-кольца $R = \mathbb{C}[x_1, \ldots, x_n]/I$ соответствуют тем максимальным идеалам M_a, $a = (a_1, \ldots, a_n) \in \mathbb{C}^n$, кольца $\mathbb{C}[x_1, \ldots, x_n]$, которые содержат идеал I, что равносильно

$$f_1, \ldots, f_r \in M_a.$$

Но при нашем соответствии в теореме Гильберта о нулях:

$$a \leftrightarrow M_a = \operatorname{Ker} s_a, \quad s_a\big(f(x_1, \ldots, x_n)\big) = f(a_1, \ldots, a_n) = 0.$$

Таким образом, $f_1, \ldots, f_r \in M_a$ тогда и только тогда, когда

$$\begin{cases} f_1(a_1, \ldots, a_n) = 0, \\ \quad \ldots \\ f_r(a_1, \ldots, a_n) = 0, \end{cases}$$

но это означает, что $a \in V$. □

Замечание 5.8.3. В силу доказанной теоремы все свойства алгебраического многообразия $V(f_1, \ldots, f_r)$, задаваемого системой алгебраических уравнений

$$f_1(x_1, \ldots, x_n) = f_2(x_1, \ldots, x_n) = \ldots = f_r(x_1, \ldots, x_n) = 0,$$

отражены в строении кольца

$$R = \mathbb{C}[x_1, \ldots, x_n]/(f_1, \ldots, f_r).$$

Часть математики, изучающая алгебраические многообразия и эти связи, называется *алгебраической геометрией*.

Теорема 5.8.4 (следствие из теоремы Гильберта о нулях). *Пусть*

$$f_1, \ldots, f_r \in \mathbb{C}[x_1, \ldots, x_n].$$

Система алгебраических уравнений

$$f_1 = \ldots = f_r = 0$$

не имеет решений $a = (a_1, \ldots, a_n) \in \mathbb{C}^n$ *тогда и только тогда, когда*

$$1 = \sum g_i f_i, \quad \text{где} \quad f_i \in \mathbb{C}[x_1, \ldots, x_n].$$

Доказательство. В силу теоремы Гильберта о нулях (теорема 5.8.2), в кольце $R = \mathbb{C}[x_1, \ldots, x_n]$ каждый максимальный идеал имеет вид $M = M_a$, $a \in \mathbb{C}^n$. Отсутствие решений $a \in \mathbb{C}^n$ нашей системы означает, что не существует максимального идеала $M = M_a$, $a \in \mathbb{C}^n$, такого, что

$$I = (f_1, \ldots, f_r) \subseteq M_a = M$$

(иначе, $f_1(a_1, \ldots, a_n) = \ldots = f_r(a_1, \ldots, a_n) = 0$). В силу теоремы о вложении собственного идеала в максимальный (теорема 5.1.1), это равносильно тому, что

$$I = (f_1, \ldots, f_r) = R = \mathbb{C}[x_1, \ldots, x_n],$$

но это эквивалентно существованию многочленов $g_i \in \mathbb{C}[x_1, \ldots, x_n]$, для которых

$$1 = \sum_{i=1}^{r} g_i f_i. \qquad \square$$

Упражнение 5.8.5. Пусть

$$f_1 = x^2 + y^2 - 1, \; f_2 = x^2 - y + 1, \; f_3 = xy - 1 \in \mathbb{C}[x, y].$$

Докажите существование таких многочленов $g_1, g_2, g_3 \in \mathbb{C}[x, y]$, что

$$1 = g_1 f_1 + g_2 f_2 + g_3 f_3.$$

Указание. Покажите, что система $f_1 = f_2 = f_3 = 0$ не имеет решений в $\mathbb{C}^2$, и примените следствие к теореме Гильберта о нулях.

5.9. Классическая форма теоремы Гильберта о нулях

Теорема Гильберта о нулях имеет много разных формулировок. Исходной формой было следующее утверждение.

Теорема 5.9.1 (классическая форма теоремы о нулях). *Пусть*

$$g, f_1, \ldots, f_r \in \mathbb{C}[x_1, \ldots, x_n],$$

$V = V(f_1, \ldots, f_r)$ *— алгебраическое многообразие общих нулей многочленов* $f_1, \ldots, f_r$, $I = (f_1, \ldots, f_r)$ *— идеал кольца* $\mathbb{C}[x_1, \ldots, x_n]$, *порождённый многочленами* $f_1, \ldots, f_r$. *Если* $g(a) = 0$ *для всех* $a \in V \subseteq \mathbb{C}^n$, *то некоторая степень многочлена* g *принадлежит идеалу* $I = (f_1, \ldots, f_r)$, $g^N \in I$, $N \in \mathbb{N}$, *другими словами,*

$$\{g \in \mathbb{C}[x_1, \ldots, x_n] \mid g(V) = 0\} = \sqrt{I},$$

где $\sqrt{I}$ *— радикал идеала* I.

Лекция N23 (6 декабря 2011 г.)

Доказательство. Рассмотрим кольцо многочленов $\mathbb{C}[x_1, \ldots, x_n, y]$ от переменных $x_1, \ldots, x_n, y$ и $r+1$ многочлен

$$f_1(x_1, \ldots, x_n), \ldots, f_r(x_1, \ldots, x_n), f_{r+1} = g(x_1, \ldots, x_n)y - 1 \in \mathbb{C}[x_1, \ldots, x_n, y].$$

Покажем, что $f_1, \ldots, f_{r+1}$ не имеют общих нулей в $\mathbb{C}^{n+1}$. Действительно, если $(a_1, \ldots, a_n, b) \in \mathbb{C}^{n+1}$ и $(a_1, \ldots, a_n)$ — общий нуль многочленов $f_1, \ldots, f_r \in \mathbb{C}[x_1, \ldots, x_n]$, то, в силу нашего предположения, $g(a_1, \ldots, a_n) = 0$, и поэтому

$$g(a_1, \ldots, a_n)b - 1 = -1 \neq 0.$$

В силу следствия теоремы о нулях, многочлены $f_1, \ldots, f_r, f_{r+1} = gy - 1$ порождают идеал, равный всему кольцу $\mathbb{C}[x_1, \ldots, x_n, y]$, и поэтому

$$1 = \sum_{i=1}^{r} g_i(x_1, \ldots, x_n, y) f_i(x_1, \ldots, x_n) + g_{r+1}(x_1, \ldots, x_n, y)(g(x_1, \ldots, x_n)y - 1).$$

Подставляя $y = \dfrac{1}{g}$ в это равенство (в поле рациональных функций $\mathbb{C}(x_1, \ldots, x_n, y)$), получаем, что

$$1 = \sum_{i=1}^{r} g_i(x_1, \ldots, x_n, g^{-1}) f_i(x_1, \ldots, x_n).$$

Умножая обе части равенства на достаточно большую степень $g(x_1, \ldots, x_n)^N$ многочлена $g(x_1, \ldots, x_n)$, получим

$$g(x_1, \ldots, x_n)^N = \sum_{i=1}^{r} h_i(x_1, \ldots, x_n) f_i(x_1, \ldots, x_n) \in I,$$

поскольку

$$h_i(x_1, \ldots, x_n) = g(x_1, \ldots, x_n)^N g_i(x_1, \ldots, x_n, g^{-1}) \in \mathbb{C}[x_1, \ldots, x_n]. \qquad \square$$

5.10. Нётеровы кольца

Конец лекции N22

Пусть R — коммутативное кольцо. Кольцо R называется *нётеровым*, если всякий идеал $I \lhd R$ является конечно порождённым (т. е. $I = Ri_1 + \ldots + Ri_k$, $i_1, \ldots, i_k \in I$). Это равносильно обрыву возрастающих цепей идеалов, т. е. тому, что не существует бесконечной строго возрастающей цепочки идеалов $I_k \lhd R$

$$I_1 < I_2 < \ldots < I_n < I_{n+1} < \ldots \quad (I_k \neq I_{k+1}).$$

Действительно, в этом случае в нётеровом кольце R имеем $I = \bigcup_k I_k$ — идеал, $I = Ri_1 + \ldots + Ri_k$. Элементы $i_1, \ldots, i_k \in I = \bigcup I_k$ попадают в некоторый идеал I_n. Но тогда $\bigcup_k I_k = I = I_n$, поэтому $I_{n+1} = I_n$, т. е. последовательность не является строго возрастающей.

Если идеал I не порождается конечным числом элементов, то найдутся элементы $i_1, i_2, \ldots \in I$ такие, что последовательность идеалов

$$Ri_1 \subset Ri_1 + Ri_2 \subset \ldots$$

является строго возрастающей. $\square$

Лекция N 23 (6 декабря 2011 г.)

Доказательство. Рассмотрим кольцо многочленов $\mathbb{C}[x_1, \ldots, x_n, y]$ от переменных $x_1, \ldots, x_n, y$ и $r+1$ многочлен

$$f_1(x_1, \ldots, x_n), \ldots, f_r(x_1, \ldots, x_n), f_{r+1} = g(x_1, \ldots, x_n)y - 1 \in \mathbb{C}[x_1, \ldots, x_n, y].$$

Покажем, что $f_1, \ldots, f_{r+1}$ не имеют общих нулей в $\mathbb{C}^{n+1}$. Действительно, если $(a_1, \ldots, a_n, b) \in \mathbb{C}^{n+1}$ и $(a_1, \ldots, a_n)$ — общий нуль многочленов $f_1, \ldots, f_r \in \mathbb{C}[x_1, \ldots, x_n]$, то, в силу нашего предположения, $g(a_1, \ldots, a_n) = 0$, и поэтому

$$g(a_1, \ldots, a_n)b - 1 = -1 \neq 0.$$

В силу следствия теоремы о нулях, многочлены $f_1, \ldots, f_r, f_{r+1} = gy - 1$ порождают идеал, равный всему кольцу $\mathbb{C}[x_1, \ldots, x_n, y]$, и поэтому

$$1 = \sum_{i=1}^{r} g_i(x_1, \ldots, x_n, y) f_i(x_1, \ldots, x_n) + g_{r+1}(x_1, \ldots, x_n, y)(g(x_1, \ldots, x_n)y - 1).$$

Подставляя $y = \dfrac{1}{g}$ в это равенство (в поле рациональных функций $\mathbb{C}(x_1, \ldots, x_n, y)$), получаем, что

$$1 = \sum_{i=1}^{r} g_i(x_1, \ldots, x_n, g^{-1}) f_i(x_1, \ldots, x_n).$$

Умножая обе части равенства на достаточно большую степень $g(x_1, \ldots, x_n)^N$ многочлена $g(x_1, \ldots, x_n)$, получим

$$g(x_1, \ldots, x_n)^N = \sum_{i=1}^{r} h_i(x_1, \ldots, x_n) f_i(x_1, \ldots, x_n) \in I,$$

поскольку

$$h_i(x_1, \ldots, x_n) = g(x_1, \ldots, x_n)^N g_i(x_1, \ldots, x_n, g^{-1}) \in \mathbb{C}[x_1, \ldots, x_n]. \qquad \square$$

5.10. Нётеровы кольца

Пусть R — коммутативное кольцо. Кольцо R называется *нётеровым*, если всякий идеал $I \lhd R$ является конечно порождённым (т. е. $I = Ri_1 + \ldots + Ri_k$, $i_1, \ldots, i_k \in I$). Это равносильно обрыву возрастающих цепей идеалов, т. е. тому, что не существует бесконечной строго возрастающей цепочки идеалов $I_k \lhd R$

$$I_1 < I_2 < \ldots < I_n < I_{n+1} < \ldots \quad (I_k \neq I_{k+1}).$$

Действительно, в этом случае в нётеровом кольце R имеем $I = \bigcup\limits_k I_k$ — идеал, $I = Ri_1 + \ldots + Ri_k$. Элементы $i_1, \ldots, i_k \in I = \bigcup I_k$ попадают в некоторый идеал I_n. Но тогда $\bigcup\limits_k I_k = I = I_n$, поэтому $I_{n+1} = I_n$, т. е. последовательность не является строго возрастающей.

Если идеал I не порождается конечным числом элементов, то найдутся элементы $i_1, i_2, \ldots \in I$ такие, что последовательность идеалов

$$Ri_1 \subset Ri_1 + Ri_2 \subset \ldots$$

является строго возрастающей. $\square$

Примеры 5.10.1 (нётеровых колец).

1) Конечные кольца.

2) $\mathbb{Z}$ и все области главных идеалов.

3) Гомоморфные образы (фактор-кольца) нётеровых колец.

Теорема 5.10.2 (теорема Гильберта о базисе для колец многочленов). *Пусть R — коммутативное нётерово кольцо. Тогда кольцо многочленов $R[x]$ от одной переменной x с коэффициентами из кольца R также является коммутативным нётеровым кольцом.*

Доказательство. Пусть I — любой идеал кольца $R[x]$. Наша цель — доказать, что идеал I конечно порождённый.

Построим следующую последовательность многочленов $f_1, f_2, \ldots \in R[x]$: f_1 — один из многочленов в I наименьшей степени $n_1 = \deg f_1$; f_2 — один из многочленов в $I \setminus \mathrm{Id}(f_1)$ наименьшей степени $n_2 = \deg f_2; \ldots$; f_{i+1} — один из многочленов в $I \setminus \mathrm{Id}(f_1, f_2, \ldots, f_i)$ наименьшей степени n_{i+1} (если $I \setminus \mathrm{Id}(f_1, f_2, \ldots, f_i) \neq \varnothing$, то есть если $\mathrm{Id}(f_1, f_2, \ldots, f_i) < I$) и так далее.

Если на шаге i имеем $\mathrm{Id}(f_1, f_2, \ldots, f_i) = I$, то всё доказано.

Допустим теперь, что получили бесконечную последовательность $\{f_i \in R[x] \mid i \in \mathbb{N}\}$. Ясно, что $n_1 \leqslant n_2 \leqslant \ldots \leqslant n_i \leqslant n_{i+1} \leqslant \ldots$. Пусть в кольце R

$$a_1 = \mathrm{Lt}(f_1),\ a_2 = \mathrm{Lt}(f_2), \ldots,\ a_i = \mathrm{Lt}(f_i), \ldots -$$

последовательность старших коэффициентов многочленов f_i. Так как R — нётерово кольцо, то рассмотрим среди них первый номер $k+1$ такой, что

$$a_{k+1} \in \mathrm{Id}(a_1, \ldots, a_k) \subseteq R.$$

Тогда

$$a_{k+1} = \sum_{j=1}^{k} c_j a_j, \quad c_j \in R.$$

Рассмотрим многочлен

$$g = \sum_{j=1}^{k} c_j x^{n_{k+1} - n_j} f_j \in \mathrm{Id}(f_1, \ldots, f_k) \subseteq R[x],$$

для которого $\mathrm{Lt}(g) = a_{k+1}$. Поэтому

$$\deg(f_{k+1} - g) < n_{k+1}.$$

Так как $f_{k+1} - g \in I \setminus \mathrm{Id}(f_1, \ldots, f_k)$, то приходим к противоречию с тем, что $n_{k+1} = \deg f_{k+1}$ являлось минимальной степенью ненулевых многочленов в $I \setminus \mathrm{Id}(f_1, \ldots, f_k)$. □

Следствие 5.10.3. *Если K — поле, то кольцо многочленов $K[x_1, x_2, \ldots, x_n]$ от n переменных над K является нётеровой областью (то есть коммутативным нётеровым кольцом без делителей нуля).*

Доказательство проводится индукцией по n с учётом того, что

$$K[x_1, x_2, \ldots, x_n] = (K[x_1, x_2, \ldots, x_{n-1}])[x_n]. \qquad \square$$

Замечание 5.10.4. Приведённое доказательство без всяких изменений показывает, что если R — нётерово слева кольцо (условие нётеровости для левых идеалов), то кольцо многочленов $R[x]$ также является нётеровым слева. При этом в кольце $S = R[x]$ для многочленов $f_1, f_2, \ldots, f_k \in R[x]$ следует рассматривать левый идеал в S, порождённый элементами $f_1, f_2, \ldots, f_k$:

$$\text{l-Id}(f_1, \ldots, f_k) = Sf_1 + Sf_2 + \ldots + Sf_k,$$

а в кольце R для элементов $a_1, a_2, \ldots, a_k \in R$ рассматривать левый идеал, ими порождённый:

$$\text{l-Id}(a_1, \ldots, a_k) = Ra_1 + \ldots + Ra_k.$$

Следствие 5.10.5 (в форме, восходящей к Гильберту). *Если R — коммутативное нётерово кольцо, $S \subseteq R[x]$, то найдётся конечное подмножество B в S, $B \subseteq S$, $|B| < \infty$, такое, что $S \subseteq \mathrm{Id}(B)$.*

Это означает, в частности, что если D — нётерова область, $D[x_1, \ldots, x_n]$ — коммутативное нётерово кольцо многочленов без делителей нуля, $D \subseteq D'$, D' — коммутативное кольцо, то $(a_1, a_2, \ldots, a_n) \in (D')^n$ — решение однородной системы полиномиальных уравнений с левыми частями S тогда и только тогда, когда $(a_1, a_2, \ldots, a_n)$ является решением конечной системы полиномиальных уравнений с левыми частями из B.

Определение 5.10.6. Пусть задано мономиальное упорядочение на $K[x_1, \ldots, x_n]$. Тогда любой ненулевой многочлен $f \in K[x_1, \ldots, x_n]$ может быть однозначно записан в виде

$$f = c_1 x^{a_1} + c_2 x^{a_2} + \ldots + c_r x^{a_r},$$

где

$$0 \neq c_i \in K, \quad a_i \in \mathbb{Z}_+^n, \quad 1 \leqslant i \leqslant r, \quad x^{a_1} > x^{a_2} > \ldots > x^{a_r}.$$

Обозначим:

$$\begin{aligned} \mathrm{Lm}(f) &= x^{a_1} && \text{(старший одночлен);} \\ \mathrm{Lc}(f) &= c_1 && \text{(старший коэффициент);} \\ \mathrm{Lt}(f) &= c_1 x^{a_1} && \text{(старший член).} \end{aligned}$$

Ясно, что

$$\mathrm{Lm}(fg) = \mathrm{Lm}(f)\mathrm{Lm}(g),$$
$$\mathrm{Lc}(fg) = \mathrm{Lc}(f)\mathrm{Lc}(g), \quad \mathrm{Lt}(fg) = \mathrm{Lt}(f)\mathrm{Lt}(g).$$

Задача 5.10.7. Пусть $x_1 > x_2 > \ldots > x_n$ и задано упорядочение degrevlex на $K[x_1, \ldots, x_n]$, f — ненулевой однородный многочлен из $K[x_1, \ldots, x_n]$ (все слагаемые-одночлены cx^α, $0 \neq c \in K$, входящие в запись f, имеют одну и ту же полную степень $|\alpha| = \sum\limits_{i=1}^{n} \alpha_i$). Покажите, что x_n делит многочлен f тогда и только тогда, когда x_n делит одночлен $\mathrm{Lt}(f)$. Выведите из этого, что $f \in \langle x_i, \ldots, x_n \rangle$ в том и только в том случае, когда $\mathrm{Lt}(f) \in \langle x_i, \ldots, x_n \rangle$.

5.11. Алгоритм деления в кольце $K[x_1, \ldots, x_n]$

Теорема 5.11.1. *Пусть фиксировано некоторое мономиальное упорядочение $>$ на $\mathbb{Z}_+^n$,*

$$F = \{f_1, \ldots, f_s\} \subset K[x_1, \ldots, x_n] -$$

конечное множество ненулевых многочленов. Тогда любой многочлен $f \in K[x_1, \ldots, x_n]$ можно представить в виде

$$f = \alpha_1 f_1 + \ldots + \alpha_s f_s + r,$$

где $\alpha_i, r \in K[x_1, \ldots, x_n]$, $r = 0$ или r является линейной комбинацией мономов (с коэффициентами из поля K), ни один из которых не делится ни на один из старших одночленов $\mathrm{Lm}(f_1), \ldots, \mathrm{Lm}(f_s)$,

$$\mathrm{Lm}(f) = \max\{\max_{1 \leqslant i \leqslant s}(\mathrm{Lm}(\alpha_i)\mathrm{Lm}(f_i)), \mathrm{Lm}(r)\}.$$

Доказательство. Если многочлен f является линейной комбинацией одночленов, ни один из которых не делится ни на один из старших членов $\mathrm{Lt}(f_1), \ldots, \mathrm{Lt}(f_s)$, то положим $\alpha_1 = \ldots = \alpha_s = 0$, $r = f$.

Иначе, пусть g — старший одночлен среди одночленов, входящих в запись многочлена f, делящийся на один из старших членов $\mathrm{Lt}(f_1), \ldots, \mathrm{Lt}(f_s)$,

$$g = \beta_1 \mathrm{Lt}(f_i) \text{ для некоторого } i,$$

β_1 — одночлен, $\beta_1 \in K[x_1, \ldots, x_n]$. Тогда положим

$$F_1 = f - \beta_1 f_i.$$

Если у многочлена F_1 имеются одночлены, делящиеся на один из $\mathrm{Lt}(f_1), \ldots, \mathrm{Lt}(f_s)$, то старший среди них меньше, чем g. Учитывая, что мономиальное упорядочение обладает условием обрыва убывающих цепей, на некотором шаге получим, что

$$F_l = F_{l-1} - \beta_l f_j$$

и либо $F_l = 0$, либо ни один из одночленов многочлена F_l не делится ни на один из одночленов $\mathrm{Lm}(f_1), \ldots, \mathrm{Lm}(f_s)$. Положим

$$r = F_l.$$

Последовательно выражая $F_{l-1} = r + \beta_l f_j$, затем выражая F_{l-2} и так далее, получим искомое представление для многочлена f. □

Пример 5.11.2 (с $>_{\mathrm{lex}}$, $x_1 > x_2$).

1) $f = x_1 x_2^2 + 1$, $f_1 = x_1 x_2 + 1$, $f_2 = x_2 + 1$:

$$f = x_2 f_1 + (-1) f_2 + 2.$$

2) $f = x_1^2 x_2 + x_1 x_2^2 + x_2^2$, $f_1 = x_1 x_2 - 1$, $f_2 = x_2^2 - 1$:

$$f = (x_1 + x_2) f_1 + 1 \cdot f_2 + x_1 + x_2 + 1.$$

В то же время

$$f = x_1 f_1 + (x_1 + 1) f_2 + 2x_1 + 1,$$

таким образом, остаток определён неоднозначно.

Задача 5.11.3. Пусть K — поле. Покажите, что многочлен $y - g(x)$, где $g(x) \in K[x]$, делит многочлен $f(x, y) \in K[x, y]$ тогда и только тогда, когда $f(x, g(x)) = 0$.

Замечание 5.11.4. Условие $r = 0$ достаточно для того, чтобы многочлен f лежал в идеале, порождённом многочленами $f_1, \ldots, f_s$ ($f \in \langle f_1, \ldots, f_s \rangle$), но не является необходимым условием. Например (при $>_{\text{lex}}$), пусть $f = x_1x_2^2 - x_1$, $f_1 = x_1x_2 + 1$, $f_2 = x_2^2 - 1$. Тогда

$$f = x_2 f_1 + 0 \cdot f_2 + (-x_1 - x_2) = 0 \cdot f_1 + x_1 f_2 + 0 \in \langle f_1, f_2 \rangle,$$

но при этом $r = -x_1 - x_2 \neq 0$.

Возникает вопрос о «хорошей» системе образующих идеала $\langle f_1, \ldots, f_s \rangle$, когда остаток r определён однозначно и условие $r = 0$ равносильно тому, что $f \in \langle f_1, \ldots, f_s \rangle$.

Определение 5.11.5. Пусть $0 \neq I \lhd K[x_1, \ldots, x_n]$ и задано мономиальное упорядочение. Обозначим

$$\mathrm{Lt}(I) = \{cx^\alpha \mid \text{существует } f \in I,\ \mathrm{Lt}(f) = cx^\alpha\}.$$

Пусть $\langle \mathrm{Lt}(I) \rangle$ — идеал кольца $K[x_1, \ldots, x_n]$, порождённый элементами множества $\mathrm{Lt}(I)$.

Замечание 5.11.6. Если идеал I конечно порождён, $I = \langle f_1, \ldots, f_s \rangle$, то может быть $\langle \mathrm{Lt}(f_1), \ldots, \mathrm{Lt}(f_s) \rangle < \mathrm{Lt}(I)$. Например, для идеала $I \lhd R[x_1, x_2]$, $I = \langle f_1, f_2 \rangle$, $f_1 = x_1^3 - 2x_1x_2$, $f_2 = x_1^2x_2 - 2x_2^2 + x_1$, с упорядочением deglex имеем $x_1^2 \in \langle \mathrm{Lt}(I) \rangle$, но $x_1^2 \notin \langle \mathrm{Lt}(f_1), \mathrm{Lt}(f_2) \rangle$.

Определение 5.11.7. При заданном мономиальном упорядочении конечное подмножество $G = \{g_1, \ldots, g_s\}$ идеала I кольца $K[x_1, \ldots, x_n]$ называется базисом Грёбнера (или стандартным базисом) идеала I, если

$$\langle \mathrm{Lt}(I) \rangle = \langle \mathrm{Lt}(g_1), \ldots, \mathrm{Lt}(g_s) \rangle.$$

Замечание 5.11.8. Если $g_1, \ldots, g_s$ — одночлены, $I = \langle g_1, \ldots, g_s \rangle \lhd K[x_1, \ldots, x_n]$, то $G = \{g_1, \ldots, g_s\}$ — базис Грёбнера идеала I. Таким образом, по лемме Диксона, любой мономиальный идеал обладает базисом Грёбнера, состоящим из мономов.

Теорема 5.11.9 (теорема Гильберта о базисе). *Пусть $0 \neq I \lhd K[x_1, \ldots, x_n]$. Тогда идеал I конечно порождён. Более того, при фиксированном мономиальном упорядочении идеал I обладает базисом Грёбнера; базис Грёбнера идеала I порождает идеал I.*

Доказательство. Старшие мономы ненулевых элементов идеала I порождают мономиальный идеал, совпадающий с $\langle \mathrm{Lt}(I) \rangle$. По лемме Диксона (лемма 4.15.7) существуют такие элементы $g_1, \ldots, g_s \in I$, что идеал $\langle \mathrm{Lt}(I) \rangle$ порождён старшими мономами элементов $g_1, \ldots, g_s$. Так как старший моном элемента отличается от старшего одночлена на множитель из поля K, то

$$\langle \mathrm{Lt}(I) \rangle = \langle \mathrm{Lt}(g_1), \ldots, \mathrm{Lt}(g_s) \rangle.$$

Покажем, что $I = \langle g_1, \ldots, g_s \rangle$. Используя алгоритм деления, имеем

$$f = \alpha_1 g_1 + \ldots + \alpha_s g_s + r,$$

где $\alpha_1, \ldots, \alpha_s, r \in K[x_1, \ldots, x_n]$, и если $r \neq 0$, то ни один из одночленов многочлена r не делится ни на один из одночленов $\mathrm{Lt}(g_i)$, $i = 1, \ldots, s$. Тогда

$$r = f - \alpha_1 g_1 - \ldots - \alpha_s g_s \in I.$$

Если $r \neq 0$, то

$$\mathrm{Lt}(r) \in \langle \mathrm{Lt}(I)\rangle = \langle \mathrm{Lt}(g_1),\ldots,\mathrm{Lt}(g_s)\rangle.$$

По лемме 4.15.4 одночлен $\mathrm{Lt}(r)$ делится хотя бы на один из одночленов $\mathrm{Lt}(g_i)$, $1 \leqslant i \leqslant s$. Это противоречит определению остатка r. Поэтому $r = 0$, следовательно, $f \in \langle g_1,\ldots,g_s\rangle$. Таким образом, $I = \langle g_1,\ldots,g_s\rangle$. □

Упражнение 5.11.10.

1) Покажите, что в замечании 5.11.6 множество

$$\{f_1, f_2, x_1^2, 2x_1x_2, -2x_2^2 + x_1\}$$

является базисом Грёбнера идеала I.

2) Пусть I — идеал кольца $\mathbb{R}[x_1,x_2,x_3]$, порождённый элементами $g_1 = x_3 + x_1$, $g_2 = x_2 - x_1$, задано упорядочение lex, $x_1 < x_2 < x_3$. Покажите, что $\{g_1, g_2\}$ — базис Грёбнера идеала I.

Указание. $\langle \mathrm{Lt}(g_1), \mathrm{Lt}(g_2)\rangle = \langle x_3, x_2\rangle$. Пусть существует такой многочлен $f \in I$, что $\mathrm{Lt}(f) \notin \langle \mathrm{Lt}(g_1), \mathrm{Lt}(g_2)\rangle = \langle x_3, x_2\rangle$. Но тогда $\mathrm{Lt}(f)$ не делится ни на x_3, ни на x_2. Так как у нас задано lex-упорядочение, то из этого следует, что $f \in \mathbb{R}[x_1]$. Так как $f \in I$, то существуют такие многочлены $h_1, h_2 \in \mathbb{R}[x_1,x_2,x_3]$, что $f = (x_3+x_1)h_1+(x_2-x_1)h_2$. Но это противоречит тому, что $f \in \mathbb{R}[x_1]$. Покажите, что если задано упорядочение lex с $x_1 > x_2 > x_3$, то $\{g_1, g_2\}$ не является базисом Грёбнера идеала I.

Упражнение 5.11.11. Пусть $A = (a_{ij}) \in \mathrm{M}_{m,n}(\mathbb{R})$ — ступенчатая матрица, I — идеал в $\mathbb{R}[x_1,\ldots,x_n]$, порождённый линейными многочленами $\sum\limits_{j=1}^{n} a_{ij}x_j$, $1 \leqslant i \leqslant m$. Пусть задано lex-упорядочение, при котором главные переменные старше свободных. Покажите, что указанные линейные многочлены образуют базис Грёбнера идеала I.

Предложение 5.11.12. *Пусть $G = \{g_1,\ldots,g_s\}$ — базис Грёбнера идеала I кольца $K[x_1,\ldots,x_n]$ и фиксировано мономиальное упорядочение. Тогда для любого многочлена $f \in K[x_1,\ldots,x_n]$ в алгоритме деления*

$$f = \alpha_1 g_1 + \ldots + \alpha_s g_s + r$$

остаток r определён однозначно ($r = 0$ или r является линейной комбинацией мономов, ни один из которых не делится ни на один из старших членов $\mathrm{Lt}(g_1),\ldots,\mathrm{Lt}(g_s)$). Остаток r называется нормальной формой элемента f: $r = \mathrm{N}_G(f)$. При этом $f \equiv g \pmod{I}$ тогда и только тогда, когда $\mathrm{N}_G(f) = \mathrm{N}_G(g)$.

Доказательство. Пусть

$$f = g + r = g' + r',$$

где $g, g' \in I$, r, r' — остатки. Тогда

$$r - r' = g' - g \in I.$$

Если $r - r' \neq 0$, то

$$\mathrm{Lt}(r - r') \in \langle \mathrm{Lt}(I)\rangle = \langle \mathrm{Lt}(g_1),\ldots,\mathrm{Lt}(g_s)\rangle.$$

Следовательно, $\mathrm{Lt}(r - r')$ делится на $\mathrm{Lt}(g_i)$ для некоторого i. Но это противоречит определению остатков r, r'. □

Замечание 5.11.13. Свойство базиса Грёбнера, изложенное в предложении 5.11.12, эквивалентно определению базиса Грёбнера.

Следствие 5.11.14. *Пусть* $0 \neq I \lhd K[x_1, \ldots, x_n]$, $f \in K[x_1, \ldots, x_n]$ *и фиксировано мономиальное упорядочение. Тогда* $f = g + r$, *где* $g \in I$, *и ни один одночлен многочлена* r *не делится ни на один элемент из* $\langle \mathrm{Lt}(I) \rangle$. *При этом многочлен* r *определён однозначно и называется нормальной формой многочлена* f, *обозначение* $r = \mathrm{N}(f)$.

Замечание 5.11.15. В предложении 5.11.12 однозначно определён только остаток r, но не многочлены $\alpha_1, \ldots, \alpha_s$. Например, пусть

$$I = \langle x_3 + x_1, x_2 - x_1 \rangle \lhd \mathbb{R}[x_1, x_2, x_3]$$

и задано упорядочение lex, $x_1 < x_2 < x_3$. Тогда $\{x_3 + x_1, x_2 - x_1\}$ — базис Грёбнера идеала I (см. упражнение ??). Для монома x_3x_2 имеем

$$x_3x_2 = x_2(x_3 + x_1) - x_1(x_2 - x_1) - x_1^2 = x_1(x_3 + x_1) + x_3(x_2 - x_1) - x_1^2.$$

Следствие 5.11.16. *Пусть* $G = \{g_1, \ldots, g_s\}$ — *базис Грёбнера идеала* I *кольца* $K[x_1, \ldots, x_n]$, $f \in K[x_1, \ldots, x_n]$. *Тогда* $f \in I$ *тогда и только тогда, когда остаток от деления многочлена* f *на* G *равен нулю.*

Замечание 5.11.17. Свойство базиса Грёбнера, отмеченное в следствии 5.11.16, эквивалентно определению базиса Грёбнера. Это даёт возможность решать проблему принадлежности многочлена f идеалу при условии, что известен некоторый базис Грёбнера этого идеала (алгоритм сводится к вычислению остатка в алгоритме деления).

Следствие 5.11.18. *Пусть* $0 \neq I \lhd K[x_1, \ldots, x_n]$ *и фиксировано мономиальное упорядочение на* $K[x_1, \ldots, x_n]$. *Тогда смежные классы, чьи представители — мономы, не лежащие в* $\langle \mathrm{Lt}(I) \rangle$, *образуют* K*-линейный базис фактор-алгебры* $K[x_1, \ldots, x_n]/I$. *Если* $G = \{g_1, \ldots, g_s\}$ — *базис Грёбнера идеала* I, *то элементы* $\{g + I \mid g$ — *моном,* g *не делится ни на какой моном* $\mathrm{Lm}(g_i)$, $1 \leqslant i \leqslant s\}$ *образуют* K*-линейный базис фактор-алгебры* $K[x_1, \ldots, x_n]/I$. *Для* $f \in K[x_1, \ldots, x_n]$ *имеем* $f + I = \mathrm{N}_G(f) + I$. *Операции сложения и умножения в фактор-алгебре* $K[x_1, \ldots, x_n]/I$ *могут быть заданы следующим образом:*

$$(\mathrm{N}_G(f) + I) + (\mathrm{N}_G(g) + I) = \mathrm{N}_G(f + g) + I,$$

при этом $\mathrm{N}_G(f + g) = \mathrm{N}_G(f) + \mathrm{N}_G(g)$,

$$(\mathrm{N}_G(f) + I)(\mathrm{N}_G(g) + I) = \mathrm{N}_G\big(\mathrm{N}_G(f) \cdot \mathrm{N}_G(g)\big) + I.$$

Предложение 5.11.19. *Пусть задано мономиальное упорядочение на* $K[x_1, \ldots, x_n]$, $I \lhd K[x_1, \ldots, x_n]$, $G = \{g_1, \ldots, g_s\} \subseteq I$ — *подмножество ненулевых многочленов идеала* I. *Тогда следующие условия эквивалентны:*

1) G — *базис Грёбнера идеала* I;

2) *для любого* $0 \neq f \in I$ *существует такое* i, $1 \leqslant i \leqslant s$, *что моном* $\mathrm{Lm}(f)$ *делится на* $\mathrm{Lm}(g_i)$;

3) *для* $f \in K[x_1, \ldots, x_n]$ *имеем* $f \in I$ *тогда и только тогда, когда в алгоритме деления* $f = \alpha_1 g_1 + \ldots + \alpha_s g_s + r$ *остаток* r *равен нулю;*

4) *для* $f \in K[x_1,\dots,x_n]$ *имеем* $f \in I$ *тогда и только тогда, когда*

$$f = \sum_{i=1}^{s} h_i g_i, \quad \text{где} \quad h_i \in K[x_1,\dots,x_n], \quad 1 \leqslant i \leqslant s,$$

$$\mathrm{Lm}(f) = \max_{1\leqslant i\leqslant s}\{\mathrm{Lm}(h_i)\mathrm{Lm}(g_i)\}.$$

Доказательство. 1) $\Longrightarrow$ 2). Пусть $f \in I$. Тогда $\mathrm{Lt}(f) \in \langle \mathrm{Lt}(G)\rangle$, $\mathrm{Lt}(f) = \sum\limits_{i=1}^{s} h_i \mathrm{Lt}(g_i)$, $h_i \in K[x_1,\dots,x_n]$, $1 \leqslant i \leqslant s$. Таким образом, каждый из одночленов в правой части (после умножений и приведения подобных членов) делится на некоторый одночлен $\mathrm{Lm}(g_i)$. Но в левой части у нас лишь один член $\mathrm{Lt}(f) = \mathrm{Lc}(f)\mathrm{Lm}(f)$, где $0 \neq \mathrm{Lc}(f) \in K$.

2) $\Longrightarrow$ 3). Пусть в алгоритме деления для $f \in I$ $f = \alpha_1 g_1 + \ldots + \alpha_s g_s + r$. Тогда если $r \neq 0$, то $r \in I$. Тогда существует такое i, $1 \leqslant i \leqslant s$, что моном $\mathrm{Lm}(r)$ делится на $\mathrm{Lm}(g_i)$. Но это противоречит определению остатка в алгоритме деления.

3) $\Longrightarrow$ 4). Имея 3) и теорему 5.11.1 об алгоритме деления, получаем 4).

4) $\Longrightarrow$ 1). Ясно, что $\langle \mathrm{Lt}(G)\rangle \subseteq \langle \mathrm{Lt}(I)\rangle$. Для любого $f \in I$ из 4) имеем

$$\mathrm{Lt}(f) = \sum_{i\in S} \mathrm{Lt}(h_i)\mathrm{Lt}(g_i),$$

где $S = \{i \mid \mathrm{Lm}(f) = \mathrm{Lm}(h_i)\mathrm{Lm}(g_i)\}$. Поэтому $\mathrm{Lt}(f) \in \mathrm{Lt}(G)$. Но идеал $\langle \mathrm{Lt}(I)\rangle$ порождается элементами $\{\mathrm{Lt}(f),\ f \in I\}$. Следовательно, $\langle \mathrm{Lt}(I)\rangle \subseteq \langle \mathrm{Lt}(G)\rangle$. □

Определение 5.11.20. Пусть $0 \neq f, g \in K[x_1,\dots,x_n]$, задано мономиальное упорядочение, $\mathrm{Lm}(f) = x^\alpha$, $\mathrm{Lm}(g) = x^\beta$, $\alpha = (\alpha_1,\dots,\alpha_n), \beta = (\beta_1,\dots,\beta_n) \in \mathbb{Z}_+^n$. Пусть $\gamma = (\gamma_1,\dots,\gamma_n)$, $\gamma_i = \max\{\alpha_i, \beta_i\}$, $i = 1,\dots,n$. Моном x^γ назовём наименьшим общим кратным мономов $\mathrm{Lm}(f)$ и $\mathrm{Lm}(g)$ (обозначение: $x^\gamma = \mathrm{lcm}\big(\mathrm{Lm}(f), \mathrm{Lm}(g)\big)$). S-многочленом от f и g называется многочлен

$$S(f,g) = \frac{x^\gamma}{\mathrm{Lt}(f)} \cdot f - \frac{x^\gamma}{\mathrm{Lt}(g)} \cdot g.$$

Пример 5.11.21.

1) $f = x_2x_1 - x_2, g = x_2^2 - x_1 \in \mathbb{R}[x_1, x_2]$, упорядочение deglex, $x_2 > x_1$. Тогда $\mathrm{Lm}(f) = x_2x_1$, $\mathrm{Lm}(g) = x_2^2$, $\mathrm{lcm}\big(\mathrm{Lm}(f), \mathrm{Lm}(g)\big) = x_2^2x_1$,

$$S(f,g) = \frac{x_2^2x_1}{x_2x_1}(x_2x_1 - x_2) - \frac{x_2^2x_1}{x_2^2}(x_2^2 - x_1) = -x_2^2 + x_1^2.$$

2) $f = x_1^3x_2^2 - x_1^2x_2^3 + x_1, g = 3x_1^4x_2 + x_2^2 \in \mathbb{R}[x_1, x_2]$, упорядочение deglex, $x_1 > x_2$. Тогда $\mathrm{Lm}(f) = x_1^3x_2^2$, $\mathrm{Lm}(g) = x_1^4x_2$, $\mathrm{lcm}\big(\mathrm{Lm}(f), \mathrm{Lm}(g)\big) = x_1^4x_2^2$,

$$S(f,g) = \frac{x_1^4x_2^2}{x_1^3x_2^2} \cdot f - \frac{x_1^4x_2^2}{3x_1^4x_2} \cdot g = -x_1^3x_2^3 + x_1^2 - \left(\frac{1}{3}\right)x_2^3.$$

Лемма 5.11.22. *Пусть* $0 \neq f_1,\dots,f_s \in K[x_1,\dots,x_n]$, $\mathrm{Lm}(f_i) = x^\alpha$, $1 \leqslant i \leqslant s$, $f = \sum\limits_{i=1}^{n} c_i f_i$, $c_i \in K$, $1 \leqslant i \leqslant s$. *Тогда если* $\mathrm{Lm}(f) < x^\alpha$, *то* f *является* K*-линейной комбинацией многочленов* $S(f_i, f_j)$, $1 \leqslant i < j \leqslant s$.

Доказательство. По условию леммы

$$f_i = a_i x^\alpha + f_i^*, \quad \text{где} \quad 0 \neq a_i \in K,$$

f_i^* — линейная комбинация мономов, меньших, чем x^α, $1 \leqslant i \leqslant s$. Так как по условию $\sum\limits_{i=1}^{s} c_i a_i = 0$ и $S(f_i, f_j) = \dfrac{1}{a_i} f_i - \dfrac{1}{a_j} f_j$, то

$$\begin{aligned} f &= \sum_{i=1}^{s} c_i f_i = \sum_{i=1}^{s} c_i a_i \left(\frac{1}{a_i} f_i\right) = \\ &= \sum_{i=1}^{s-1} (c_1 a_1 + \ldots + c_i a_i) \left(\frac{1}{a_i} f_i - \frac{1}{a_{i+1}} f_{i+1}\right) + \left(\sum_{i=1}^{s} c_i a_i\right) \frac{1}{a_s} f_s = \\ &= \sum_{i=1}^{s-1} (c_1 a_1 + \ldots + c_i a_i) S(f_i, f_{i+1}). \end{aligned}$$

□

Теорема 5.11.23. *Пусть $0 \neq g_1, \ldots, g_s \in K[x_1, \ldots, x_n]$ и задано мономиальное упорядочение. Тогда $G = \{g_1, \ldots, g_s\}$ является базисом Грёбнера идеала I, порождённого множеством G, тогда и только тогда, когда для всех $i \neq j$, $1 \leqslant i, j \leqslant s$, остаток от деления многочлена $S(g_i, g_j)$ на G (в любом порядке) равен нулю.*

Доказательство. Если G является базисом Грёбнера идеала $I = \langle g_1, \ldots, g_s \rangle$, то по следствию 5.11.16 для любого многочлена $f \in I$ остаток от деления f на G равен нулю (по определению многочлены $S(f_i, f_j)$ лежат в идеале I).

Пусть теперь остаток от деления любого многочлена $S(g_i, g_j)$, $1 \leqslant i, j \leqslant s$, $i \neq j$, на G (в любом порядке) равен нулю. Для любого многочлена $f \in I$ рассмотрим представление в виде

$$f = \sum_{i=1}^{s} h_i g_i, \quad h_i \in K[x_1, \ldots, x_n], \quad 1 \leqslant i \leqslant n, \tag{5.1}$$

с наименьшим мономом $x^\alpha = \max\limits_{1 \leqslant i \leqslant t} \{\mathrm{Lm}(h_i)\mathrm{Lm}(g_i)\}$, это возможно, так как мономиальное упорядочение является вполне упорядочением. Если $x^\alpha = \mathrm{Lm}(f)$, то по предложению 5.11.19 утверждение доказано. Иначе $\mathrm{Lm}(f) < x^\alpha$.

Пусть $S = \{i \mid \mathrm{Lm}(h_i)\mathrm{Lm}(g_i) = x^\alpha\}$, и для любого $i \in S$ запишем $h_i = c_i \tilde{h}_i + h_i^*$, где $\tilde{h}_i = \mathrm{Lm}(h_i)$, $c_i \in K$, h_i^* — K-линейная комбинация мономов, меньших, чем $\tilde{h}_i$. Положим $g = \sum\limits_{i \in S} c_i \tilde{h}_i g_i$. Тогда $\mathrm{Lm}(\tilde{h}_i g_i) = x^\alpha$ для всех $i \in S$, но $\mathrm{Lm}(f) < x^\alpha$. По лемме 5.11.22 существуют такие элементы $a_{ij} \in K$, $1 \leqslant i, j \leqslant s$, что

$$g = \sum_{\substack{i,j \in S \\ i \neq j}} a_{ij} S(\tilde{h}_i g_i, \tilde{h}_j g_j). \tag{5.2}$$

Так как для $i \in S$ имеем $\mathrm{Lm}(\tilde{h}_i g_i) = x^\alpha$, то

$$\mathrm{lcm}\big(\mathrm{Lm}(\tilde{h}_i g_i), \mathrm{Lm}(\tilde{h}_j g_j)\big) = x^\alpha.$$

Поэтому

$$S(\tilde{h}_i g_i, \tilde{h}_j g_j) = \frac{x^\alpha}{\mathrm{Lt}(\tilde{h}_i g_i)} \tilde{h}_i g_i - \frac{x^\alpha}{\mathrm{Lt}(\tilde{h}_j g_j)} \tilde{h}_j g_j = \frac{x^\alpha}{\mathrm{Lt}(g_i)} g_i - \frac{x^\alpha}{\mathrm{Lt}(g_j)} g_j = \frac{x^\alpha}{f_{ij}} S(g_i, g_j),$$

где $f_{ij} = \operatorname{lcm}(\operatorname{Lm}(g_i), \operatorname{Lm}(g_j))$.

По нашему предположению, остаток от деления любого многочлена $S(g_i, g_j)$ на G (в любом порядке) равен нулю, поэтому остаток от деления любого многочлена $S(\tilde{h}_i g_i, \tilde{h}_j g_j)$ на G равен нулю. Следовательно,

$$S(\tilde{h}_i g_i, \tilde{h}_j g_j) = \sum_{l=1}^{s} \gamma_{ijl} g_l, \tag{5.3}$$

где

$$\max_{1\leqslant l\leqslant s}\{\operatorname{Lm}(\gamma_{ijl})\operatorname{Lm}(g_l)\} = \operatorname{Lm}\big(S(\tilde{h}_i g_i, \tilde{h}_j g_j)\big) < x^\alpha.$$

Подставляя полученные выражения (5.3) в g, (5.2), а (5.2) в f, (5.1), получаем $f = \sum\limits_{i=1}^{s} h'_i g_i$, где $h'_i \in K[x_1,\dots,x_n]$, $1 \leqslant i \leqslant s$, $\max\limits_{1\leqslant i\leqslant t}\{\operatorname{Lm}(h'_i)\operatorname{Lm}(g_i)\} < x^\alpha$, что противоречит минимальности монома x^α в исходном представлении (5.1). □

Следствие 5.11.24. *Пусть $0 \neq g_i \in K[x_1,\dots,x_n]$, $G = \{g_1,\dots,g_s\}$. Тогда G является базисом Грёбнера идеала $I = \langle g_1,\dots,g_s\rangle$ тогда и только тогда, когда для всех $i \neq j$, $1 \leqslant i, j \leqslant s$, имеем*

$$S(g_i, g_j) = \sum_{l=1}^{s} \gamma_{ijl} g_l$$

где $\gamma_{ijl} \in K[x_1,\dots,x_n]$, $\operatorname{Lm}\big(S(g_i,g_j)\big) = \max\limits_{1\leqslant l\leqslant s}\{\operatorname{Lm}(\gamma_{ijl})\operatorname{Lm}(g_l)\}$.

Пример 5.11.25 (идеал скрученной кубики в $\mathbb{R}^3$). Пусть $G = \{x_2 - x_1^2, x_3 - x_1^3\}$, I — идеал кольца многочленов $\mathbb{R}[x_1, x_2, x_3]$, порождённый множеством G, и задано мономиальное упорядочение lex с $x_2 > x_3 > x_1$. Имеем

$$S(x_2 - x_1^2, x_3 - x_1^3) = \frac{x_2x_3}{x_2}(x_2 - x_1^2) - \frac{x_2x_3}{x_3}(x_3 - x_1^3) =$$
$$= x_2x_1^3 - x_3x_1^2 = x_1^3(x_2 - x_1^2) + (-x_1^2)(x_3 - x_1^3) + 0.$$

Поэтому по теореме 5.11.23 множество G — базис Грёбнера идеала I.

Если в lex-упорядочении положить $x_1 > x_2 > x_3$, то множество G не будет базисом Грёбнера идеала I.

Следствие 5.11.26. *Теорема 5.11.23 даёт следующий алгоритм построения базиса Грёбнера идеала. Пусть идеал I порождён подмножеством ненулевых многочленов $G = \{g_1,\dots,g_s\} \subseteq K[x_1,\dots,x_n]$, и задано мономиальное упорядочение. Если для какой-то пары (g_i, g_j), где $i \neq j$, остаток r_{ij} от деления многочлена $S(g_i, g_j)$ на G не равен нулю, то добавим этот остаток к множеству G, $G' = G \cup \{r_{ij}\}$, и рассмотрим S-многочлены множества G'. Продолжая этот процесс, мы построим базис Грёбнера $\tilde{G}$ идеала I, $G \subseteq \tilde{G}$.*

Доказательство. Покажем, что мы построим базис Грёбнера за конечное число шагов. Если $G' \neq G$, то это означает, что $G' = G \cup \{r_{ij}\}$, где $r_{ij} \neq 0$, $\operatorname{Lt}(r_{ij}) \notin \langle \operatorname{Lt}(G)\rangle$. Но $\operatorname{Lt}(r_{ij}) \in \langle \operatorname{Lt}(G')\rangle$. Поэтому $\langle \operatorname{Lt}(G')\rangle \supset \langle \operatorname{Lt}(G)\rangle$. Таким образом, получаем возрастающую цепь идеалов

$$\langle \operatorname{Lt}(G)\rangle \subset \langle \operatorname{Lt}(G')\rangle \subset \langle \operatorname{Lt}(G'')\rangle \subset \dots$$

в кольце $K[x_1, \ldots, x_n]$. По теореме Гильберта о базисе эта цепь стабилизируется, то есть на некотором шаге для $\tilde{G}$ остатки от деления всех S-многочленов на $\tilde{G}$ будут равны нулю. Так как $G \subset \tilde{G}$, то $\tilde{G}$ порождает идеал I. По теореме 5.11.23 $\tilde{G}$ — базис Грёбнера идеала I. □

Упражнение 5.11.27.

1) Пусть задано мономиальное упорядочение deglex на $\mathbb{R}[x_1, x_2]$, $x_2 < x_1$, $g_1 = x_1^3 - 2x_1x_2$, $g_2 = x_1^2x_2 - 2x_2^2 + x_1$, I — идеал в $\mathbb{R}[x_1, x_2]$, порождённый множеством $\{g_1, g_2\}$. Применяя алгоритм, описанный в следствии 5.11.26, постройте базис Грёбнера идеала I.

   Ответ. $\{g_1, g_2, -x_1^2, -2x_1x_2, -2x_2^2 + x_1\}$.

2) Пусть задано lex-упорядочение на $\mathbb{R}[x_1, x_2]$, $x_2 > x_1$, $g_1 = x_2^2 + x_2x_1 + x_1^2$, $g_2 = x_2 + x_1$, $g_3 = x_2$. Постройте базис Грёбнера идеала $I = \langle g_1, g_2, g_3 \rangle \lhd \mathbb{R}[x_1, x_2]$.

   Ответ. $\{g_1, g_2, g_3, x_1^2, x_1\}$.

3) Пусть задано lex-упорядочение на $\mathbb{Z}_5[x_1, x_2]$, $x_1 > x_2$, $g_1 = x_1^2 + x_2^2 + 1$, $g_2 = x_1^2x_2 + 2x_1x_2 + x_1$. Постройте базис Грёбнера идеала $I = \langle g_1, g_2 \rangle \lhd \mathbb{Z}_5[x_1, x_2]$.

   Ответ. $\{g_1, g_2, 3x_1x_2 + 4x_1 + x_2^3 + x_2, 4x_2^5 + 3x_2^4 + x_2^2 + x_2 + 3\}$.

Следствия 5.11.18 и 5.11.26 дают возможность определять структурные константы и производить вычисления в фактор-алгебрах $K[x_1, \ldots, x_n]/I$.

Пример 5.11.28. Пусть $K = \mathbb{Q}$, $f_1 = x_1^2x_2 - x_2 + x_1$, $f_2 = x_1^2x_2 - x_1$, $I = \langle f_1, f_2 \rangle \lhd \mathbb{Q}[x_1, x_2]$ и задано мономиальное упорядочение deglex, $x_1 < x_2$. Применяя алгоритм следствия 5.11.26, получаем базис Грёбнера G идеала I:

$$G = \{x_1^2x_2 - x_2 + x_1, -x_2^2 + x_1x_2 + x_1^2, x_1^3 + x_2 - 2x_1\}.$$

По следствию 5.11.18, смежные классы, чьи представители — многочлены $\{1, x_1, x_2, x_1^2, x_1x_2\}$, образуют линейный базис фактор-алгебры $\mathbb{Q}[x_1, x_2]/I$, $\dim_{\mathbb{Q}}(\mathbb{Q}[x_1, x_2]/I) = 5$. Таблица умножения на базисных элементах:

$\cdot$	1	x_1	x_2	x_1^2	x_1x_2
1	1	x_1	x_2	x_1^2	x_1x_2
x_1	x_1	x_1^2	x_1x_2	$-x_2 + 2x_1$	$x_2 - x_1$
x_2	x_2	x_1x_2	$x_1x_2 + x_1^2$	$x_2 - x_1$	x_1
x_1^2	x_1^2	$-x_2 + 2x_1$	$x_2 - x_1$	$-x_1x_2 + 2x_1^2$	$x_1x_2 - x_1^2$
x_1x_2	x_1x_2	$x_2 - x_1$	x_1	$x_1x_2 - x_1^2$	x_1^2

Пример 5.11.29. Пусть на $\mathbb{R}[x_1, x_2, x_3]$ задано мономиальное упорядочение degrevlex, $x_1 > x_2 > x_3$,

$$I = \langle x_1^2 + x_1x_3, x_1x_2 + x_2^2, x_2x_3 + x_3^2 \rangle \lhd \mathbb{R}[x_1, x_2, x_3].$$

Тогда, применяя алгоритм следствия 5.11.26, получаем базис Грёбнера G идеала I:

$$G = \{x_1^2 + x_1x_3, x_1x_2 + x_2^2, x_2x_3 + x_3^2, x_1x_3^2 - x_3^3, x_2^3 - x_3^3, x_3^4\}.$$

По следствию 5.11.18, смежные классы, чьи представители — многочлены $\{1, x_1, x_2, x_3, x_1x_3, x_2^2, x_3^2, x_3^3\}$, образуют линейный базис фактор-алгебры $\mathbb{R}[x_1, x_2, x_3]/I$, $\dim_{\mathbb{R}}(\mathbb{R}[x_1, x_2, x_3]/I) = 8$. Таблица умножения алгебры $\mathbb{R}[x_1, x_2, x_3]/I$:

$\cdot$	1	x_1	x_2	x_3	x_1x_3	x_2^2	x_3^2	x_3^3
1	1	x_1	x_2	x_3	x_1x_3	x_2^2	x_3^2	x_3^3
x_1	x_1	$-x_1x_3$	$-x_2^2$	x_1x_3	$-x_3^3$	$-x_3^3$	x_3^3	0
x_2	x_2	$-x_2^2$	x_2^2	$-x_3^3$	$-x_3^3$	x_3^3	$-x_3^3$	0
x_3	x_3	x_1x_3	$-x_3^2$	x_3^2	x_3^3	x_3^3	x_3^3	0
x_1x_3	x_1x_3	$-x_3^3$	$-x_3^3$	x_3^3	0	0	0	0
x_2^2	x_2^2	$-x_3^3$	x_3^3	x_3^3	0	0	0	0
x_3^2	x_3^2	x_3^3	$-x_3^3$	x_3^3	0	0	0	0
x_3^3	x_3^3	0	0	0	0	0	0	0

Замечание 5.11.30. Алгоритм, описанный в следствии 5.11.26, не является оптимальным. Можно предложить много его усовершенствований. Во-первых, если остаток r_{ij} от деления многочлена $S(g_i, g_j)$ равен нулю в G, то он будет равен нулю и в G', поэтому нет необходимости вычислять на последующих шагах этот остаток. Существует несколько более сложных условий, которые описывают те S-многочлены, которые нет необходимости рассматривать, их остатки от деления будут нулевыми за счёт алгоритма присоединения остатков других S-многочленов.

Базисы Грёбнера, построенные с помощью алгоритма, описанного в следствии 5.11.26, часто получаются избыточными (из них можно исключить лишние образующие). Пусть, например, G-базис Грёбнера идеала $I \lhd K[x_1, \ldots, x_n]$, $g \in G$, $\mathrm{Lt}(g) \in \langle \mathrm{Lt}(G \setminus \{g\})\rangle$. Тогда $G \setminus \{g\}$ — базис Грёбнера идеала I. Действительно, при данном условии $\langle \mathrm{Lt}(G)\rangle = \langle \mathrm{Lt}(G \setminus \{g\})\rangle$.

Определение 5.11.31. Минимальным базисом Грёбнера идеала $I \lhd K[x_1, \ldots, x_n]$ называется такой базис Грёбнера идеала I, что для всех $g \in G$ имеем:

1) $\mathrm{Lc}(g) = 1$;

2) $\mathrm{Lt}(g) \neq \langle \mathrm{Lt}(G \setminus \{g\})\rangle$.

Замечание 5.11.32. Минимальный базис Грёбнера может быть получен из некоторого базиса Грёбнера с помощью исключения лишних образующих (если $G = \{g_1, \ldots, g_s\}$ — базис Грёбнера идеала I и при $i \neq j$ одночлен $\mathrm{Lm}(g_i)$ делится на $\mathrm{Lm}(g_j)$, то положим $G' = G \setminus \{g_i\}$, и так далее) и домножения на ненулевые константы. Например, в упражнении 5.11.27 1) $\left\{x_1^2, x_1x_2, x_2^2 - \left(\frac{1}{2}\right)x_1\right\}$ — минимальный базис Грёбнера. Минимальный базис Грёбнера определён неоднозначно. Так, в рассмотренном примере $\left\{x_1^2 + cx_1x_2, x_1x_2, x_2^2 - \left(\frac{1}{2}\right)x_1\right\}$ — минимальный базис Грёбнера при любом $c \in \mathbb{R}$.

Упражнение 5.11.33. Пусть G — базис Грёбнера идеала I и $\mathrm{Lc}(g) = 1$ для всех $g \in G$. Покажите, что G является минимальным базисом Грёбнера в том и только в том случае, когда никакое собственное подмножество множества G не является базисом Грёбнера идеала I.

Лемма 5.11.34. *Пусть задано мономиальное упорядочение на* $K[x_1, \ldots, x_n]$, $0 \neq I \lhd K[x_1, \ldots, x_n]$, $G = \{g_1, \ldots, g_s\}$ *и* $F = \{f_1, \ldots, f_t\}$ *— минимальные базисы Грёбнера идеала* I. *Тогда* $t = s$, *и после перенумерации элементов (если необходимо) имеем* $\mathrm{Lt}(f_i) = \mathrm{Lt}(g_i)$, $1 \leqslant i \leqslant s$.

Доказательство. Так как $f_1 \in I$ и G — базис Грёбнера идеала I, то существует такое i, $1 \leqslant i \leqslant s$, что одночлен $\mathrm{Lm}(f_1)$ делится на $\mathrm{Lm}(g_i)$. Перенумеровав множество G, если необходимо, мы можем считать, что $i = 1$. Так как $g_1 \in I$ и F — базис Грёбнера идеала I, то существует такое j, $1 \leqslant j \leqslant t$, что одночлен $\mathrm{Lm}(g_1)$ делится на $\mathrm{Lm}(f_j)$. Следовательно, $\mathrm{Lm}(f_1)$ делится на $\mathrm{Lm}(f_j)$, и поскольку F — минимальный базис Грёбнера, то $j = 1$. Поэтому $\mathrm{Lm}(f_1) = \mathrm{Lm}(g_1)$. Продолжая этот процесс (рассматривая $f_2, f_3, \ldots$), мы получаем, что $t = s$ и $\mathrm{Lm}(f_i) = \mathrm{Lm}(g_i)$, $1 \leqslant i \leqslant s$. □

Определение 5.11.35. Базис Грёбнера G ненулевого идеала I кольца $K[x_1, \ldots, x_n]$ называется редуцированным, если для любого $g \in G$:

1) $\mathrm{Lc}(g) = 1$;

2) никакой одночлен многочлена g не лежит в $\langle \mathrm{Lt}(G \setminus \{g\}) \rangle$.

Предложение 5.11.36. *Пусть* $0 \neq I \lhd K[x_1, \ldots, x_n]$ *и задано мономиальное упорядочение. Тогда существует единственный редуцированный базис Грёбнера идеала* I.

Доказательство. Пусть $G = \{g_1, \ldots, g_s\}$ — минимальный базис Грёбнера идеала I, h_1 — остаток от деления многочлена g_1 на $H_1 = \{g_2, \ldots, g_s\}$, h_2 — остаток от деления многочлена g_2 на $H_2 = \{h_1, g_3, \ldots, g_s\}$, h_3 — остаток от деления многочлена g_3 на $H_3 = \{h_1, h_2, \ldots, g_4, \ldots, g_s\}, \ldots$, h_s — остаток от деления многочлена g_s на $H_s = \{h_1, h_2, \ldots, h_{s-1}\}$. Тогда $H = \{h_1, \ldots, h_s\}$ — редуцированный базис Грёбнера идеала I. Действительно, так как G — минимальный базис Грёбнера идеала I, то $\mathrm{Lm}(h_i) = \mathrm{Lm}(g_i)$, $1 \leqslant i \leqslant s$, и H — также минимальный базис Грёбнера идеала I. По построению множества H h_i — остаток от деления многочлена g_i, при применении алгоритма деления мы уменьшаем одночлены многочлена g_i, используя одночлены $\mathrm{Lm}(h_1), \ldots, \mathrm{Lm}(h_{i-1})$, $\mathrm{Lm}(g_{i+1}), \ldots, \mathrm{Lm}(g_s)$, и так как $\mathrm{Lm}(h_j) = \mathrm{Lm}(g_j)$, $1 \leqslant j \leqslant s$, то H — редуцированный базис Грёбнера.

Пусть теперь $G = \{g_1, \ldots, g_s\}$ и $H = \{h_1, \ldots, h_s\}$ — редуцированные базисы Грёбнера идеала I (так как G и H — минимальные базисы Грёбнера, то по лемме 5.11.34 они содержат одинаковое число элементов и $\mathrm{Lt}(g_i) = \mathrm{Lt}(h_i)$, $1 \leqslant i \leqslant s$). Для любого i, $1 \leqslant i \leqslant s$, если $g_i \neq h_i$, то $0 \neq g_i - h_i \in I$. Поэтому существует такое j, $1 \leqslant j \leqslant s$, что одночлен $\mathrm{Lm}(g_i - h_i)$ делится на $\mathrm{Lm}(h_j)$. Так как $\mathrm{Lm}(g_i - h_i) < \mathrm{Lm}(h_j)$, то $j \neq i$. Следовательно, одночлен $\mathrm{Lm}(h_j) = \mathrm{Lm}(g_j)$ делит один из одночленов многочленов g_i или h_i, что противоречит тому, что G и H — редуцированные базисы Грёбнера идеала I. □

Следствие 5.11.37. *Пусть задано мономиальное упорядочение на* $K[x_1, \ldots, x_n]$, $0 \neq I, J \lhd K[x_1, \ldots, x_n]$. *Тогда* $I = J$ *в том и только в том случае, когда идеалы* I *и* J *имеют один и тот же редуцированный базис Грёбнера.*

Упражнение 5.11.38. Покажите, что идеалы

$$I = \langle x_1 + x_1x_2, x_2 + x_1x_2, x_1^2, x_2^2 \rangle \quad \text{и} \quad \langle x_1, x_2 \rangle$$

кольца $\mathbb{R}[x_1, x_2]$ совпадают.

Замечание 5.11.39. Для ненулевого идеала I кольца $K[x_1, \ldots, x_n]$ существует лишь конечное число редуцированных базисов Грёбнера (при различных мономиальных упорядочениях). Объединение этих базисов — базис идеала I, являющийся базисом Грёбнера идеала I при любом мономиальном упорядочении кольца $K[x_1, \ldots, x_n]$ (универсальным базисом Грёбнера). Например, множество $\{x_1 - x_2^2, x_1x_2 - x_1, x_2^3 - x_2^2, x_1^2 - x_1\}$ является универсальным базисом Грёбнера идеала $I = \langle x_1 - x_2^2, x_1x_2 - x_1\rangle$ кольца многочленов $\mathbb{R}[x_1, x_2]$.

Пример 5.11.40.

1) В упражнении 5.11.27, 2) множество $\{x_2, x_1\}$ является редуцированным базисом Грёбнера идеала
$$I = \langle x_2^2 + x_2x_1 + x_1^2, x_2 + x_1, x_2\rangle \lhd \mathbb{R}[x_1, x_2]$$
с lex-упорядочением, $x_2 > x_1$. Множество $\{x_2 + x_1, x_1\}$ — минимальный базис Грёбнера, не являющийся редуцированным.

2) В упражнении 5.11.27, 3)
$$\{x_1^2 + x_2^2 + 1, x_1x_2 + 3x_1 + 2x_2^3 + 2x_2, x_2^5 + 2x_2^4 + 4x_2^2 + 4x_2 + 2\} -$$
редуцированный базис Грёбнера идеала
$$I = \langle x_1^2 + x_2^2 + 1, x_1^2x_2 + 2x_1x_2 + x_1\rangle \lhd \mathbb{Z}_5[x_1, x_2]$$
(с lex-упорядочением, $x_1 > x_2$).

3) Пусть $n \in \mathbb{N}$, $x_1 > x_2 > x_3 > x_4$ и задано degrevlex-упорядочение на $\mathbb{R}[x_1, x_2, x_3, x_4]$. Покажите, что редуцированный базис Грёбнера идеала
$$I = \langle x_1^{n+1} - x_2x_3^{n-1}x_4, x_1x_2^{n-1} - x_3^n, x_1^nx_3 - x_2^nx_4\rangle \lhd \mathbb{R}[x_1, x_2, x_3, x_4]$$
содержит многочлен $x_3^{n^2+1} - x_2^{n^2}x_4$. Это показывает, что степени промежуточных (и результирующих) многочленов при вычислении базисов Грёбнера могут быть очень большими. Более того, при вычислении базисов Грёбнера коэффициенты могут быть очень сложными рациональными числами, даже если исходные коэффициенты были небольшими целыми числами.

Замечание 5.11.41. Сложность вычисления базисов Грёбнера может зависеть от данного мономиального упорядочения. Например, для идеала
$$I = \langle x_1^5 + x_2^4 + x_3^3 - 1, x_1^3 + x_2^3 + x_3^2 - 1\rangle \lhd \mathbb{R}[x_1, x_2, x_3]$$
$x_1 > x_2 > x_3$, сложность вычисления базиса Грёбнера в lex-упорядочении значительно выше, чем в degrevlex-упорядочении. В упорядочении degrevlex редуцированный базис Грёбнера идеала I состоит из элементов
$$x_2^6 + x_1x_2^4 + 2x_2^3x_3^2 + x_1x_3^3 + x_3^4 - 2x_2^3 - 2x_3^2 - x_1 + 1,$$
$$x_1^2x_2^3 - x_2^4 + x_1^2x_3^2 - x_3^3 - x_1^2 + 1, \quad x_1^3 + x_2^3 + x_3^2 - 1.$$
В упорядочении lex редуцированный базис Грёбнера идеала I имеет больше элементов, элементы имеют высокие степени и большие коэффициенты.

Упражнение 5.11.42.

1) Пусть $K = \mathbb{R}$ и на $\mathbb{R}[x_1, x_2, x_3]$ задано упорядочение lex, $x_1 > x_2 > x_3$,

$$I = \langle x_3^2 x_2 + x_3^2, x_1^3 x_2 + x_1 + x_2 + 1, x_3 + x_1^2 + x_2^3 \rangle \lhd \mathbb{R}[x_1, x_2, x_3].$$

Найдите редуцированный базис Грёбнера идеала I.

Ответ. $\{x_3^4 - x_3^3, x_2^{11} + 3x_2^8 x_3 - 2x_2^7 - 4x_2^4 x_3 + x_2^3 + x_2^2 + 2x_2 + x_3^3 - x_3^2 + x_3 + 1, x_1^2 + x_2^3 + x_3, x_2 x_3^2 + x_3^2, x_1 x_2 + x_1 + x_2^7 + 2x_2^4 x_3 - x_2^4 - 2x_2^3 x_3 + x_2^3 + x_2^2 x_3 - x_2 x_3 + x_2 - x_3^3 + 5x_3^2 + x_3 + 1\}$.

2) Пусть на $\mathbb{R}[x_1, x_2]$ задано упорядочение deglex, $x_1 > x_2$,

$$I = \langle x_1^2 + x_2^2 - 25, x_1^2 - x_2^2 - 1, x_1 x_2 - 1 \rangle \lhd \mathbb{R}[x_1, x_2].$$

Найдите редуцированный базис Грёбнера идеала I.

Ответ. $\{1\}$, это означает, что $I = \mathbb{R}[x_1, x_2]$.

3) Пусть на $\mathbb{R}[x_1, x_2]$ задано упорядочение lex, $x_1 > x_2$,

$$I = \langle x_1^2, x_1 x_2 + x_2^2 \rangle.$$

Найдите редуцированный базис Грёбнера идеала I.

Замечание 5.11.43. Пусть $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$, I — идеал кольца $K[x_1, \ldots, x_n]$, порождённый линейными многочленами $\sum\limits_{j=1}^{n} a_{ij} x_j$, $1 \leqslant i \leqslant m$. Пусть $x_1 > \ldots > x_n$, задано lex-упорядочение на $K[x_1, \ldots, x_n]$, $A_{\text{ст}}$ — ступенчатый вид матрицы A, в котором коэффициенты при лидерах равны 1, $A_{\text{ст}} \in \mathrm{M}_{m,n}(K)$, $r \leqslant m$, $A_{\text{ст}} = (\tilde{a}_{ij})$, $A_{\text{гл ст}}$ — главный ступенчатый вид матрицы A (коэффициенты при лидерах равны 1 и над лидерами стоят нули), $A_{\text{гл ст}} = (\tilde{a}_{ij}) \in \mathrm{M}_{r,n}(K)$, $x_{j_1}, \ldots, x_{j_r}$ — главные неизвестные. Тогда

$$\{x_{j_l} + \tilde{a}_{l,j_l+1} x_{j_l+1} + \ldots + \tilde{a}_{l_n} x_n \mid l = 1, \ldots, r\} -$$

минимальный базис Грёбнера идеала I, а

$$\{x_{j_l} + \tilde{a}_{l,j_l+1} x_{j_l+1} + \ldots + \tilde{a}_{l_n} x_n \mid l = 1, \ldots, r\} -$$

редуцированный базис Грёбнера идеала I. Таким образом, приведение матрицы к ступенчатому виду эквивалентно вычислению базиса Грёбнера соответствующего идеала, порождённого линейными многочленами (и так как редуцированный базис Грёбнера определён однозначно, то это ещё раз показывает, что главный ступенчатый вид матрицы определён однозначно).

Например, пусть $x_1 > x_2 > x_3 > x_4 > x_5 > x_6$ и задано lex-упорядочение на $R = \mathbb{R}[x_1, x_2, x_3, x_4, x_5, x_6]$, I — идеал кольца R, порождённый линейными многочленами $x_1 - 2x_2 + 2x_3 - x_4 + 5x_5 + 8x_6$, $x_3 - x_5 - 2x_6$, $2x_1 - 4x_2 - x_4 + 4x_5 + 8x_6$, $x_1 - 2x_2 - x_4 + 7x_5 + 12x_6$. Приведя соответствующую матрицу к главному ступенчатому виду, имеем

$$\begin{pmatrix} 1 & -2 & 2 & -1 & 5 & 8 \\ 0 & 0 & 1 & 0 & -1 & -2 \\ 2 & -4 & 0 & -1 & 4 & 8 \\ 1 & -2 & 0 & -1 & 7 & 12 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 0 & 0 & -3 & -4 \\ 0 & 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 0 & 1 & -10 & -16 \end{pmatrix}.$$

Таким образом, редуцированный базис Грёбнера идеала I составляют многочлены $x_1 - 2x_2 - 3x_5 - 4x_6$, $x_3 - x_5 - 2x_6$, $x_4 - 10x_5 - 16x_6$.

Замечание 5.11.44. Если $G = \{g_1, \ldots, g_s\}$ — подмножество ненулевых многочленов кольца $K[x_1]$, то построение редуцированного базиса Грёбнера идеала $I = \langle g_1, \ldots, g_s \rangle \lhd K[x_1]$ сводится к нахождению наибольшего общего делителя многочленов $g_1, \ldots, g_s$.

Упражнение 5.11.45.

1) Покажите, что

$$x_2^3 + x_3^3 \in I = \langle x_1^2, x_1x_2 + x_2^2, x_1x_3 + x_3^2 \rangle \lhd \mathbb{R}[x_1, x_2, x_3].$$

Указание. Рассмотрите мономиальное упорядочение degrevlex, $x_1 > x_2 > x_3$, и найдите базис Грёбнера G идеала I,

$$G = \{x_1^2, x_1x_2 + x_2^2, x_1x_3 + x_3^2, x_2^3, x_3^3\}.$$

С помощью алгоритма деления проверьте, что $\mathrm{N}_G(x_2^3 + x_3^3) = 0$.

2) Пусть

$$I = \langle x_1x_3 - x_2^2, x_1^3 - x_3^2 \rangle \lhd \mathbb{C}[x_1, x_2, x_3].$$

а) Покажите, что $f = -4x_1^2x_2^2x_3^2 + x_2^6 + 3x_3^5 \in I$.

Указание. Рассмотрите deglex-упорядочение и найдите редуцированный базис Грёбнера G идеала I,

$$G = \{x_1x_3 - x_2^2, x_1^3 - x_3^2, x_1^2x_2^2 - x_3^3, x_1x_2^4 - x_3^4, x_2^6 - x_3^5\},$$

и с помощью алгоритма деления покажите, что $\mathrm{N}_G(f) = 0$.

б) Покажите, что $g = x_1x_2 - 5x_3^2 + x_1 \notin I$.

Указание. $\mathrm{Lt}(g) = x_1x_2 \notin \langle \mathrm{Lt}(G) \rangle = \langle x_1x_3, x_1^3, x_1^2x_2^2, x_1x_2^4, x_2^6 \rangle$.

3) Пусть $f = x_1^4x_2 - 2x_1^5 + 2x_1^2x_2^2 - 2x_1^3x_2 - 2x_1^4 - 2x_2^3 + 4x_1x_2^2 - 3x_1^2x_2 + 2x_1^3 - x_2 + 2x_1$. Покажите, что $f \in \langle x_1^2x_2 - x_2 + x_1, x_1x_2^2 - x_1 \rangle$.

Задача 5.11.46.

1) Принадлежит ли многочлен $f = -4x_1^2x_2^2x_3^2 + x_2^6 + 3x_3^5$ идеалу $\langle x_1x_3 - x_2^2, x_1^3 - x_3^2 \rangle \lhd \mathbb{C}[x_1, x_2, x_3]$?

2) Принадлежит ли многочлен $f = x_1^3x_3 - 2x_2^2$ идеалу $\langle x_1x_3 - x_2, x_1x_2 + 2x_3^2, x_2 - x_3 \rangle \lhd \mathbb{C}[x_1, x_2, x_3]$?

Существование алгоритмов построения базисов Грёбнера даёт возможность определять, является ли элемент u фактор-алгебры $A = K[x_1, \ldots, x_n]/I$ обратимым, и если да, то находить обратный элемент u^{-1}. В случае, когда $\dim_K A < \infty$, мы можем использовать таблицу умножения алгебры A, и задача сводится к исследованию и решению системы линейных уравнений.

Например, пусть $K = \mathbb{R}$, на $\mathbb{R}[x_1, x_2]$ задано deglex-упорядочение, $x_1 < x_2$, $I = \langle x_1^2x_2 - x_2 + x_1, x_1x_2^2 - x_1 \rangle \lhd \mathbb{R}[x_1, x_2]$. Необходимо определить, обратим ли элемент $f = x_2 + x_1 + 1 + I$ в фактор-алгебре $A = \mathbb{R}[x_1, x_2]/I$, и если обратим, то найти обратный элемент f^{-1}. В примере 5.11.28, 1) мы нашли линейный базис алгебры A и построили

таблицу умножения базисных элементов. Следовательно, если элемент f обратим в A, то существуют такие числа $a, b, c, d, e \in \mathbb{R}$, что

$$(ax_1x_2 + bx_1^2 + cx_2 + dx_1 + e)(x_2 + x_1 + 1) \equiv 1 \pmod{I}.$$

Перемножая скобки и записывая произведение в виде линейной комбинации базисных элементов, после приравнивания коэффициентов при базисных элементах в левой и правой частях, получаем систему линейных уравнений

$$\begin{cases} a + 2c + d = 0 \\ b + c + d = 0 \\ a + c + e = 0 \\ b + d + e = 0 \\ e = 1 \end{cases}$$

Решая эту систему, получаем $a = -2$, $b = -1$, $c = 1$, $d = 0$, $e = 1$. Следовательно,

$$(x_2 + x_1 + 1 + I)^{-1} = (-2x_1x_2 - x_1^2 + x_2 + 1) + I.$$

Отметим, что если исходный элемент не имеет обратного, то соответствующая система линейных уравнений будет несовместной.

В случае, когда $\dim_K K[x_1, \ldots, x_n]/I = \infty$, можно поступать следующим образом. Для многочлена $f \in K[x_1, \ldots, x_n]$ элемент $f + I$ алгебры $A = K[x_1, \ldots, x_n]/I$ обратим тогда и только тогда, когда $\langle I, f\rangle = K[x_1, \ldots, x_n]$ (если $g + I = (f + I)^{-1}$, то $fg - 1 \in I$ эквивалентно тому, что $1 \in \langle I, f\rangle$). Если $I = \langle f_1, \ldots, f_s\rangle$, $f \in K[x_1, \ldots, x_n]$, то для нахождения обратного элемента $(f + I)^{-1}$ алгебры A найдём сначала редуцированный базис Грёбнера G идеала $I = \langle f_1, \ldots, f_n, f\rangle \lhd K[x_1, \ldots, x_n]$. Если $G \neq \{1\}$, то элемент $f + I$ необратим в алгебре A. Если же $G = \{1\}$, то, сохраняя все промежуточные вычисления при нахождении редуцированного базиса Грёбнера G, мы можем записать

$$1 = \alpha_1 f_1 + \ldots + \alpha_s f_s + gf,$$

где $\alpha_1, \ldots, \alpha_s, g \in K[x_1, \ldots, x_n]$. Следовательно, $(f + I)^{-1} = g + I$.

Упражнение 5.11.47. Пусть

$$I = \langle x_1^2 + x_2x_3 - 2, x_2^2 + x_1x_3 - 3, x_1x_2 + x_3^2 - 5\rangle \lhd \mathbb{Q}[x_1, x_2, x_3].$$

Найдите обратный элемент (если он существует) для элемента $x_1 + I$ в фактор-алгебре $A = \mathbb{Q}[x_1, x_2, x_3]/I$.

Указание. Рассмотрите мономиальное упорядочение degrevlex, $x_1 > x_2 > x_3$, и найдите редуцированный базис Грёбнера идеала I:

$$G = \Big\{x_1^2 + x_2x_3 - 2, x_2^2 + x_1x_3 - 3, x_1x_2 + x_3^2 - 5,$$
$$x_1x_3^2 - \frac{5}{2}x_1 + x_2 - \frac{3}{2}x_3, x_2x_3^2 + \frac{3}{2}x_1 - \frac{5}{2}x_2 - x_3,$$
$$x_3^4 + x_1x_3 + \frac{3}{2}x_2x_3 - \frac{15}{2}x_3^2 + \frac{19}{2}\Big\}.$$

Поэтому смежные классы, чьи представители — элементы множества $\{1, x_1, x_2, x_3, x_1x_3, x_2x_3, x_3^2, x_3^3\}$, образуют линейный базис алгебры A. Представляя $(x_1 + I)^{-1}$ в виде линейной комбинации с неопределёнными коэффициентами базисных элементов, запишите $(x_1 + I)(x_1 + I)^{-1} = 1$. После записи левой части в виде линейной комбинации базисных элементов, составьте систему линейных уравнений на коэффициенты (приравнивая левую часть к правой) и решите её. Ответ:

$$(x_1 + I)^{-1} = \frac{2}{11}x_3^3 - \frac{2}{11}x_1 + \frac{3}{11}x_2 - \frac{5}{11}x_3 + I.$$

5.12. Решение систем алгебраических уравнений

Пусть дана следующая система алгебраических уравнений:

$$f_i(x_1, \ldots, x_n) = 0, \quad 0 \neq f_i \in \mathbb{C}[x_1, \ldots, x_n], \quad 1 \leqslant i \leqslant s. \tag{5.4}$$

Из теоремы 5.8.4 получаем следующий критерий совместности системы алгебраических уравнений. Если задано мономиальное упорядочение и G — редуцированный базис Грёбнера идеала $I = \langle f_1, \ldots, f_s \rangle \lhd \mathbb{C}[x_1, \ldots, x_n]$, то система (5.4) несовместна тогда и только тогда, когда $G = \{1\}$.

Если G' — базис Грёбнера идеала I, не обязательно редуцированный, то условие несовместности эквивалентно тому, что множество G содержит ненулевую константу. Например, с помощью этого критерия несложно установить, что система

$$\begin{cases} x^2y + 4y^2 - 17 = 0 \\ 2xy - 3y^3 + 8 = 0 \\ xy^2 - 5xy + 1 = 0 \end{cases}$$

несовместна (над $\mathbb{C}$).

Для идеала I рассмотрим аффинное многообразие

$$V(I) = \{(a_1, \ldots, a_n) \in \mathbb{C}^n \mid f(a_1, \ldots, a_n) = 0 \text{ для всех } f \in I\}.$$

Таким образом, решить систему (5.4) — это значит найти все точки аффинного многообразия $V(I)$.

Определение 5.12.1. Пусть K — поле, $I = \langle f_1, \ldots, f_s \rangle \lhd K[x_1, \ldots, x_n]$. Тогда i-м исключающим идеалом I_l алгебры $K[x_{l+1}, \ldots, x_n]$ называется идеал

$$I_l = I \cap K[x_{l+1}, \ldots, x_n], \quad I_0 = I.$$

Теорема 5.12.2 (об исключении). *Если $0 \neq I = \langle f_1, \ldots, f_s \rangle \lhd K[x_1, \ldots, x_n]$, G — базис Грёбнера идеала I относительно* lex*-упорядочения, $x_1 > x_2 > \ldots > x_n$, то*

$$G_l = G \cap K[x_{l+1}, \ldots, x_n], \quad 0 \leqslant l \leqslant n, \text{ —}$$

базис Грёбнера идеала I_l.

Доказательство. Достаточно показать, что $\langle \mathrm{Lt}(I_l)\rangle = \langle \mathrm{Lt}(G_l)\rangle$. Ясно, что $\langle \mathrm{Lt}(G_l)\rangle \subseteq \langle \mathrm{Lt}(I_l)\rangle$. Пусть $0 \neq f \in I_l$. Покажем, что $\mathrm{Lt}(f)$ делится на $\mathrm{Lt}(g)$ для некоторого $g \in G_l$. Так как $f \in I$ и G — базис Грёбнера идеала I, то $\mathrm{Lt}(f)$ делится на некоторый $\mathrm{Lt}(g)$, $g \in G$. Но $f \in I_l$, и мы используем lex-упорядочение, $x_1 > \ldots > x_n$. Поэтому из того, что $\mathrm{Lt}(g) \in K[x_{l+1}, \ldots, x_n]$, следует, что $g \in K[x_{l+1}, \ldots, x_n]$. Следовательно, $g \in G_l$. □

Замечание 5.12.3. Пусть $0 \neq f_1, \ldots, f_s \in \mathbb{C}[x_1, \ldots, x_n]$, $I = \langle f_1, \ldots, f_s\rangle \lhd \mathbb{C}[x_1, \ldots, x_n]$. Тогда система уравнений $f_i = 0$, $1 \leqslant i \leqslant s$, имеет конечное множество решений тогда и только тогда, когда $\dim \mathbb{C}[x_1, \ldots, x_n]/I < \infty$ (это эквивалентно тому, что если $G = \{g_1, \ldots, g_s\}$ — редуцированный базис Грёбнера идеала I, то для каждого $i = 1, 2, \ldots, n$ существует такое j, $1 \leqslant j \leqslant t$, что $\mathrm{Lt}(g_j) = x_j^{r_j}$, где $r_j \in \mathbb{N}$). Идеал I, обладающий одним из указанных эквивалентных свойств, называется нульмерным. В случае системы линейных алгебраических уравнений указанное свойство эквивалентно определённости системы.

Более сильное утверждение состоит в том, что число различных нулей нульмерного идеала I не превосходит $\dim_{\mathbb{C}} \mathbb{C}[x_1, \ldots, x_n]/I$ (для радикального идеала I эти числа совпадают). Если нули считаются с их кратностями, то число нулей идеала I равно $\dim_{\mathbb{C}} \mathbb{C}[x_1, \ldots, x_n]/I$ (обобщение теоремы Безу). Здесь под кратностью нуля (решения) следует понимать следующее.

Пусть система алгебраических уравнений (5.4) имеет конечное число решений $p_1, \ldots, p_t \in \mathbb{C}^n$, $p_i = (c_{i1}, \ldots, c_{in})$, $1 \leqslant i \leqslant t$, $I = \langle f_1, \ldots, f_s\rangle \lhd \mathbb{C}[x_1, \ldots, x_n]$. Обозначим

$$M_{p_i} = \langle x_1 - c_{i1}, \ldots, x_n - c_{in}\rangle \lhd \mathbb{C}[x_1, \ldots, x_n], \quad 1 \leqslant i \leqslant t.$$

Можно показать, что идеал I допускает единственное представление в виде $I = J_1 \cap J_2 \cap \ldots \cap J_t$, где $J_i \lhd \mathbb{C}[x_1, \ldots, x_n]$, $\sqrt{J_i} = M_{p_i}$, $1 \leqslant i \leqslant t$. Кратностью решения p_i, $1 \leqslant i \leqslant t$, называется число $\dim_{\mathbb{C}} \mathbb{C}[x_1, \ldots, x_n]/J_i$. Например, если система имеет ровно одно решение $(c_1, \ldots, c_n) \in \mathbb{C}^n$, то $\sqrt{I} = \langle x_1 - c_1, \ldots, x_n - c_n\rangle \lhd \mathbb{C}[x_1, \ldots, x_n]$, и кратность решения равна $\dim \mathbb{C}[x_1, \ldots, x_n]/I$.

Например, для системы

$$\begin{cases} x = 0 \\ y^2 = 0 \end{cases}$$

кратность решения $(0, 0)$ равна 2, а для системы

$$\begin{cases} x^2 = 0 \\ y^2 = 0 \end{cases}$$

кратность решения $(0, 0)$ равна 4.

Пусть I — нульмерный идеал алгебры $\mathbb{C}[x_1, \ldots, x_n]$ (многообразие $V(I)$ конечно, $\dim \mathbb{C}[x_1, \ldots, x_n]/I < \infty$). Для любого многочлена $f \in \mathbb{C}[x_1, \ldots, x_n]$ рассмотрим линейное отображение

$$m_f \colon \mathbb{C}[x_1, \ldots, x_n]/I \to \mathbb{C}[x_1, \ldots, x_n]/I,$$

задаваемое правилом

$$m_f(g + I) = fg + I.$$

Тогда для любого i, $1 \leqslant i \leqslant n$, собственные значения оператора m_{x_i} — это в точности x_i-координаты точек многообразия $V(I)$, для любого многочлена $f \in \mathbb{C}[x_1, \ldots, x_n]$ собственные значения оператора m_f совпадают со значениями многочлена f на точках из $V(I)$. Идеал I является радикальным ($I = \sqrt{I}$) в том и только в том случае, когда линейный оператор m_f диагонализируем для любого $f \in \mathbb{C}[x_1, \ldots, x_n]$.

Упражнение 5.12.4.

1) Покажите, что система

$$\begin{cases} x^2 + yz - 2 = 0 \\ y^2 + xz - 3 = 0 \\ xy + z^2 - 5 = 0 \end{cases}$$

имеет решения и число решений конечно.

2) Покажите, что система

$$\begin{cases} zx + yx - x + z^2 - 2 = 0 \\ xy^2 + 2xz - 3x + z + y - 1 = 0 \\ 2z^2 + zy^2 - 3z + 2zy + y^3 - 3y = 0 \end{cases}$$

имеет бесконечное множество решений.

При решении систем алгебраических уравнений, имеющих конечное множество решений, удобно пользоваться мономиальным упорядочением lex. Действительно, пусть $x_1 > \ldots > x_n$. Тогда, используя технику исключения, мы получим, что редуцированный базис Грёбнера G идеала $I = \langle f_1, \ldots, f_s \rangle$ содержит такой элемент g, скажем g_n, что g_n зависит только от x_n. Далее, базис Грёбнера G содержит такой элемент g', скажем g_{n-1}, что $\mathrm{Lt}(g_{n-1}) = x_{n-1}^r$, где $r \in \mathbb{N}$. Так как мы используем мономиальное упорядочение lex, то все члены многочлена g_{n-1} зависят только от x_n и x_{n-1}. Продолжая этот процесс, получаем, что редуцированный базис Грёбнера G содержит такие элементы $g_1, \ldots, g_n$, что $\mathrm{Lt}(g_i) = x_i^{r_i}$, где $r_i \in \mathbb{N}$, а все остальные члены многочлена g_i (кроме $x_i^{r_i}$) зависят только от $x_i, x_{i+1}, \ldots, x_n$ (и не зависят от $x_1, \ldots, x_{i-1}$). При этом может оказаться, что $G \neq \{g_1, \ldots, g_n\}$, что означает, что $|G| > n$. Таким образом, для того чтобы найти все решения нашей системы, достаточно сделать следующее: найти все корни уравнения $g_n(x_n) = 0$; затем для каждого из этих корней λ: подставим в оставшиеся многочлены базиса Грёбнера G (так как у нас имеется элемент $g_{n-1}(x_{n-1}, x_n)$ со старшим x_{n-1}^r, то после подстановки $x_n = \lambda$ возникнет конечное число ненулевых многочленов (от λ) $g_{n-1}(x_{n-1}), h_1(x_{n-1}), \ldots, h_t(x_{n-1})$); найдём $H(x_{n-1}) = \text{НОД}\big(g_{n-1}(x_{n-1}), h_1(x_{n-1}), \ldots, h_t(x_{n-1})\big)$ и для каждого корня μ многочлена H подставим $x_{n-1} = \mu$ в многочлены базиса Грёбнера (в которых уже совершена подстановка $x_n = \lambda$), отличные от $g_{n-1}(x_{n-1}), h_1(x_{n-1}), \ldots, h_t(x_{n-1})$; продолжая этот процесс, находим всё конечное множество решений системы. Отметим, что такой алгоритм эффективен при условии эффективности решения задачи отыскания корней многочлена от одной переменной (и так как такая задача не всегда может быть решена точно, то возникает теория численных методов решения систем алгебраических уравнений).

Например, рассмотрим систему

$$\begin{cases} f_1 = x_1^2 + x_2^2 + x_3^2 - 3 = 0 \\ f_2 = x_1 x_2 x_3 - 1 = 0 \\ f_3 = x_1 x_3 + x_2 - 2 = 0, \end{cases}$$

где $f_1, f_2, f_3 \in \mathbb{C}[x_1, x_2, x_3]$. Полагая $x_1 > x_2 > x_3$ и рассматривая мономиальное упорядочение lex, находим редуцированный базис Грёбнера G идеала $I = \langle f_1, f_2, f_3 \rangle \lhd \mathbb{C}[x_1, x_2, x_3]$:

$$G = \left\{ x_1 + x_2x_3 - \frac{1}{2}x_3^5 + 2x_3^3 - \frac{7}{2}x_3, x_2^2 - 2x_2 + 1, \right.$$
$$\left. x_2x_3^2 - x_2 + \frac{1}{2}x_3^4 - 2x_3^2 + \frac{3}{2}, x_3^6 - 3x_3^4 + 3x_3^2 - 1 \right\}.$$

Полагая $g_3(x_3) = x_3^6 - 3x_3^4 + 3x_3^2 - 1 = 0$, получаем, что $x_3 = \pm 1$. Полагая $g_2(x_2, x_3) = x_2^2 - 2x_2 + 1$, получаем $x_2 = 1$ (вне зависимости от x_3). Подставляя $x_3 = \pm 1$ в многочлен $x_2x_3^2 - x_2 + \frac{1}{2}x_3^4 - 2x_3^2 + \frac{3}{2}$, получаем 0, таким образом, этот многочлен не накладывает ограничений на частичные решения. Подставляя $x_2 = 1$, $x_3 = +1$ в $g_1(x_1, x_2, x_3) = f_1$, имеем $g_1(x_1, 1, 1) = x_1 + x_2 - 2 = 0$ и получаем решение $(1, 1, 1)$. Подставляя $x_2 = 1$, $x_3 = -1$, имеем $g_1(x_1, 1, -1) = x_1 - x_2 + 2 = 0$, решение $(-1, 1, -1)$. Таким образом, рассмотренная система имеет два решения $(1, 1, 1)$ и $(-1, 1, -1)$. Нетрудно убедиться, что кратность каждого из них равна 4.

Отметим, что задача исключения переменных при решении систем алгебраических уравнений может быть также решена с использованием теории результантов (см., например, для двух переменных).

При решении систем алгебраических уравнений, имеющих бесконечное множество решений, ситуация становится намного сложнее. Фактически, необходимо производить исключение переменных (с помощью теории результантов или базисов Грёбнера), а затем пытаться продолжить решение, полученное при исключении переменных, на весь набор переменных. Следующее утверждение показывает, когда существует такое продолжение.

Теорема 5.12.5 (о продолжении). *Пусть $I = \langle f_1, \ldots, f_s \rangle \subseteq \mathbb{C}[x_1, \ldots, x_n]$, I_1 — первый исключающий идеал. Пусть $f_i = g_i(x_2, \ldots, x_n)x_1^{n_i} + \varphi_i$, где многочлен φ_i является суммой одночленов, содержащих x_1 в степени, меньшей чем n_i, $n_i \geqslant 0$, $g_i \in \mathbb{C}[x_2, \ldots, x_n]$, $g_i \neq 0$, $1 \leqslant i \leqslant s$. Пусть $(c_2, \ldots, c_n) \in V(I_1)$ — частичное решение. Тогда, если $(c_2, \ldots, c_n) \notin V(g_1, \ldots, g_s)$, то существует такое число $c_1 \in \mathbb{C}$, что $(c_1, c_2, \ldots, c_n) \in V(I)$.*

Для двух многочленов утверждение теоремы следует из свойств результантов (см. ??). Например, для системы

$$\begin{cases} xy = 1 \\ xz = 1 \end{cases}$$

имеем $I = \langle xy - 1, xz - 1 \rangle$, $I_1 = \langle y - z \rangle$. Частичные решения: (c, c), $c \in \mathbb{C}$. Все эти решения, кроме решения $(0, 0)$, продолжаются до решения системы $\left(\frac{1}{c}, c, c\right)$.

Пример 5.12.6.

$$\begin{cases} f_1 = x^2 - y^2 + x + y - z = 0 \\ f_2 = x^2 + 2y^2 - 2x + y - z = 0 \\ f_3 = x^3 - x^2z - xy^2 - 2y^2z + x^2 + xy + xz - yz + z^2 = 0, \end{cases}$$

где $f_1, f_2, f_3 \in \mathbb{C}[x, y, z]$.

Пусть $x > y > z$ и используется мономиальное упорядочение lex. Базис Грёбнера идеала $I \lhd \mathbb{C}[x, y, z]$: $\{x - y^2, y^4 + y - z\}$. Ясно, что множество решений этой системы — это $\{(t^2, t, t^4 + t) \mid t \in \mathbb{C}\}$.

Последний пример показывает, что часть переменных в системе, имеющей бесконечное множество решений, может играть роль свободных, а оставшиеся — роль главных, по аналогии со свободными и главными переменными в методе Гаусса решения систем линейных уравнений. Фактически, метод исключения и продолжения для систем линейных уравнений совпадает с методом Гаусса.

Разделение переменных на главные и свободные может быть произведено следующим образом (для систем, имеющих бесконечное множество решений). Для фиксированного номера i, $1 \leqslant i \leqslant n$, рассмотрим такое мономиальное упорядочение lex, в котором переменная x_i является младшей, и найдём редуцированный базис Грёбнера идеала $I = \langle f_1, \ldots, f_n \rangle \lhd \mathbb{C}[x_1, \ldots, x_n]$. Если в множестве G имеется многочлен, зависящий только от x_i, то переменная x_i главная (иначе x_i свободная). Если для любого i переменная x_i главная, то система имеет лишь конечное множество решений.

Пусть x_i — свободная переменная. Для фиксированного j, $1 \leqslant j \leqslant n$, $j \neq i$, рассмотрим мономиальное упорядочение, при котором x_j и x_i — младшие переменные, $x_j > x_i$, и находим редуцированный базис Грёбнера идеала I. Если этот базис содержит многочлен, зависящий лишь от x_j и x_i, то переменная x_j главная, иначе x_j — свободная переменная.

Для каждой свободной переменной x_i находим все пары свободных переменных (x_i, x_j) (если такие пары существуют), затем для каждой пары свободных переменных (x_i, x_j) находим все тройки свободных переменных (x_i, x_j, x_k) (если такие тройки существуют) и так далее. Таким образом мы можем построить максимальный набор свободных переменных. Максимально возможное число переменных в наборе свободных переменных системы называется размерностью множества решений системы (это определение обобщает понятие размерности пространства решений системы линейных уравнений).

Определение 5.12.7. Пусть на $K[x_1, \ldots, x_n]$ задано мономиальное упорядочение $<_x$, а на $K[y_1, \ldots, y_m]$ — мономиальное упорядочение $<_y$. Зададим мономиальное упорядочение на $K[y_1, \ldots, y_m, x_1, \ldots, x_n]$, полагая

$$x^\alpha y^\beta < x^\delta y^\gamma \iff \begin{cases} x^\alpha <_x x^\delta \text{ или} \\ x^\alpha = x^\delta \text{ и } y^\beta < y^\gamma. \end{cases}$$

Это упорядочение называется упорядочением исключения, где переменные $x_1, \ldots, x_n$ старше, чем $y_1, \ldots, y_m$.

Задача 5.12.8 (идеалы исключения).

1) Пусть $0 \neq I \lhd K[y_1, \ldots, y_m, x_1, \ldots, x_n]$ и задано мономиальное упорядочение исключения, при котором переменные $x_1, \ldots, x_n$ старше, чем $y_1, \ldots, y_m$. Покажите, что если G — базис Грёбнера идеала I, то $G \cap K[y_1, \ldots, y_m]$ — базис Грёбнера идеала $I \cap K[y_1, \ldots, y_m]$. В частности, если $x_1 < \ldots < x_n$ и задано упорядочение lex на $K[x_1, \ldots, x_n]$, $0 \neq I \lhd K[x_1, \ldots, x_n]$, G — базис Грёбнера идеала I, тогда для любого i, $1 \leqslant i \leqslant n$, $G \cap K[x_1, \ldots, x_i]$ — базис Грёбнера идеала $I \cap K[x_1, \ldots, x_i]$.

2) Пусть $I, J \lhd K[x_1, \ldots, x_n]$, x_{n+1} — новая переменная. Тогда

$$I \cap J = \langle x_{n+1} I, (1 - x_{n+1}) J \rangle \cap K[x_1, \ldots, x_n].$$

Если $I = \langle f_1, \ldots, f_s \rangle$, $J = \langle g_1, \ldots, g_t \rangle$, то

$$\langle x_{n+1} I, (1 - x_{n+1}) I \rangle = \langle x_{n+1} f_1, \ldots, x_{n+1} f_s, (1 - x_{n+1}) g_1, \ldots, (1 - x_{n+1}) g_t \rangle.$$

В частности, для идеалов $I = \langle x_1^2 + x_2^3 - 1, x_1 - x_2x_1 + 3\rangle$ и $J = \langle x_1^2x_2 - 1\rangle$ кольца многочленов $\mathbb{R}[x_1, x_2]$ найдите $I \cap J$.

Указание. Введите новую переменную x_{n+1}, рассмотрите упорядочение deglex на x_1, x_2, $x_1 > x_2$, упорядочение исключения на $\mathbb{R}[x_1, x_2, x_3]$, при котором переменная x_3 старше переменных x_1, x_2, и найдите базис Грёбнера G идеала

$$\langle x_3(x_1^2 + x_2^3 - 1), x_3(x_1 - x_2x_1 + 3), (1 - x_3)(x_1^2x_2 - 1)\rangle \lhd \mathbb{R}[x_1, x_2, x_3].$$

Тогда $G \cap \mathbb{R}[x_1, x_2]$ — базис Грёбнера идеала $I \cap J$ (в нашем случае это все многочлены из G, не содержащие переменной x_3).

3) Покажите, что если $0 \neq f, g \in K[x_1, \ldots, x_n]$, то

$$\langle f\rangle \cap \langle g\rangle = \langle \operatorname{lcm}(f, g)\rangle.$$

Поэтому для того, чтобы найти наименьшее общее кратное $\operatorname{lcm}(f, g)$, достаточно найти редуцированный базис Грёбнера G идеала $\langle x_{n+1}f, (1 - x_{n+1})g\rangle$ кольца $K[x_1, \ldots, x_n, x_{n+1}]$, $x_{n+1} > x_1, \ldots, x_n$, с lex-упорядочением. Тогда $\operatorname{lcm}(f, g)$ — многочлен в G, в котором нет переменной x_{n+1}, $\gcd(f, g) = \dfrac{fg}{\operatorname{lcm}(f, g)}$ (вычисление можно произвести, используя алгоритм деления).

4) Пусть $f = x_1^2x_2^2 - x_2^2 + x_1^2 - 1$, $g = x_1x_2^2 - x_2^2 - x_1 + 1 \in \mathbb{R}[x, y]$. Найдите $\operatorname{lcm}(f, g)$ и $\gcd(f, g)$.

Ответ. $\operatorname{lcm}(f, g) = 1 - x_1^2 - x_2^4 + x_1^2x_2^4$, $\gcd(f, g) = x_1 - 1$.

Замечание 5.12.9. При рассмотрении гомоморфизмов K-алгебр $\varphi\colon K[y_1, \ldots, y_m] \to K[x_1, \ldots, x_n]$ (K-линейных гомоморфизмов колец) возникает естественная задача описания ядра $\operatorname{Ker}\varphi$ и образа $\operatorname{Im}\varphi$. Нетрудно видеть, что если $\varphi(y_i) = f_i \in K[x_1, \ldots, x_n]$, $1 \leqslant i \leqslant m$, $I = \langle y_1 - f_1, \ldots, y_m - f_m\rangle \lhd K[y_1, \ldots, y_m, x_1, \ldots, x_n]$, то $\operatorname{Ker}\varphi = I \cap K[y_1, \ldots, y_m]$.

С использованием упорядочения исключения, при котором переменные $x_1, \ldots, x_n$ старше переменных $y_1, \ldots, y_m$, получаем следующий алгоритм построения базиса Грёбнера идеала $\operatorname{Ker}\varphi$: находим базис Грёбнера идеала

$$I = \langle y_1 - f_1, \ldots, y_m - f_m\rangle \lhd K[y_1, \ldots, y_m, x_1, \ldots, x_n]$$

и выбираем все элементы этого базиса, в которые не входят переменные $x_1, \ldots, x_n$.

Например, пусть $\varphi\colon \mathbb{R}[y_1, y_2, y_3, y_4] \to \mathbb{R}[x_1, x_2]$ — гомоморфизм,

$$f_1 = \varphi(y_1) = x_1^4, \quad f_2 = \varphi(y_2) = x_1^3x_2,$$
$$f_3 = \varphi(y_3) = x_1x_2^3, \quad f_4 = \varphi(y_4) = x_2^4.$$

Вычисляя базис Грёбнера G идеала

$$\langle y_1 - f_1, y_2 - f_2, y_3 - f_3, y_4 - f_4\rangle \lhd K[y_1, y_2, y_3, y_4, x_1, x_2]$$

с упорядочением исключения: упорядочение deglex на x_1, x_2, $x_2 > x_1$; degrevlex-упорядочение на y_1, y_2, y_3, y_4, $y_1 > y_2 > y_3 > y_4$, переменные x_1, x_2 старше, чем y_1, y_2, y_3, y_4, получаем базис Грёбнера идеала $\operatorname{Ker}\varphi$:

$$G \cap \mathbb{R}[y_1, y_2, y_3, y_4] = \{y_2y_3 - y_1y_4, y_3^3 - y_2y_4^2, y_1y_3^2 - y_2^2y_4, y_2^3 - y_1^2y_3\}.$$

Рассмотренные упорядочения исключения можно использовать и для описания образа $\operatorname{Im}\varphi$: пусть G — редуцированный базис Грёбнера идеала

$$I = \langle y_1 - f_1, \ldots, y_m - f_m \rangle \lhd K[y_1, \ldots, y_m, x_1, \ldots, x_n].$$

Тогда для многочлена $f \in K[x_1, \ldots, x_n]$ имеем $f \in \operatorname{Im}\varphi$ в том и только в том случае, когда существует такой многочлен $h \in K[y_1, \ldots, y_m]$, что $h = \mathrm{N}_G(f)$ (остаток от деления f на G), при этом $f = \varphi(h) = h(f_1, \ldots, f_m)$. Это даёт алгоритм для решения задачи о принадлежности многочлена f подалгебре алгебры $K[x_1, \ldots, x_n]$, порождённой многочленами $f_1, \ldots, f_m$.

Например, рассмотрим задачу о том, принадлежит ли многочлен $x^8 - x^2$ подалгебре, порождённой многочленами $x^4 + x^2$ и x^3, или нет. Для этого рассмотрим идеал

$$I = \langle y_1 - x^4 - x^2, y_2 - x^3 \rangle \lhd K[y_1, y_2, x]$$

и упорядочение lex, $x > y_1 > y_2$. Найдём базис Грёбнера G идеала I:

$$G = \{x^2 + xy_2 - y_1, xy_1 + xy_2^2 - y_1y_2 - y_2, xy_2^3 - xy_2 + y_1^2 - y_1y_2^2 - 2y_2^2, y_1^3 - 3y_1y_2^2 - y_2^4 - y_2^2\}.$$

Остаток $\mathrm{N}_G(x^8 - x^2)$ от деления многочлена $x^8 - x^2$ на G равен $y_1^2 - 2y_2^2 - y_1 \in K[y_1, y_2]$. Поэтому

$$x^8 - x^2 = (x^4 + x^2)^2 - 2(x^3)^2 - (x^4 + x^2) \in K[x^4 + x^2, x^3],$$

и многочлен $x^8 - x^2$ принадлежит подалгебре, порождённой многочленами $x^4 + x^2$ и x^3.

В качестве упражнения покажите, что

$$x^{11} + x^{15} \in \mathbb{Z}_2[x^4, x^6 + x^7, x^{10}].$$

Это же соображение может быть применено для выражения симметрических многочленов через элементарные симметрические многочлены. Например, пусть дан симметрический многочлен

$$f = x_1^2x_2 + x_1^2x_3 + x_2^2x_1 + x_2^2x_3 + x_3^2x_1 + x_3^2x_2.$$

Необходимо выразить многочлен f через элементарные симметрические многочлены

$$\sigma_1 = x_1 + x_2 + x_3, \quad \sigma_2 = x_1x_2 + x_1x_3 + x_2x_3, \quad \sigma_3 = x_1x_2x_3.$$

Рассмотрим идеал

$$I = \langle y_1 - (x_1 + x_2 + x_3), y_2 - (x_1x_2 + x_1x_3 + x_2x_3), y_3 - x_1x_2x_3 \rangle \lhd \mathbb{R}[y_1, y_2, y_3, x_1, x_2, x_3],$$

упорядочение lex с $x_1 > x_2 > x_3 > y_1 > y_2 > y_3$. Находим базис Грёбнера G идеала I:

$$G = \{x_1 + x_2 + x_3 - y, x_2^2 + x_2x_3 - x_2y_1 + x_3^2 - x_3y_1 + y_2, x_3^2 - x_3^2y_1 + x_3y_2 - y_3\}.$$

Остаток $\mathrm{N}_G(f)$ от деления многочлена f на G равен $y_1y_2 - 3y_3$. Поэтому $f = \sigma_1\sigma_2 - 3\sigma_3$.

Особый интерес представляет случай, когда $\operatorname{Im}\varphi = K[x_1, \ldots, x_n]$ (отображение «на»). Распознать эту ситуацию можно следующим образом.

Пусть задано упорядочение исключения на $K[y_1, \ldots, y_m, x_1, \ldots, x_n]$, где переменные $x_1, \ldots, x_n$ старше, чем переменные $y_1, \ldots, y_m$, $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$, $I = \langle y_1 - f_1, \ldots, y_m - f_m \rangle \lhd K[y_1, \ldots, y_m, x_1, \ldots, x_n]$, G — редуцированный базис Грёбнера идеала I. Тогда гомоморфизм $\varphi\colon K[y_1, \ldots, y_m] \to K[x_1, \ldots, x_m]$, $\varphi(y_i) = f_i$, $1 \leqslant i \leqslant m$,

отображает алгебру $K[y_1,\dots,y_m]$ на алгебру $K[x_1,\dots,x_n]$ тогда и только тогда, когда для каждого i, $1 \leqslant i \leqslant n$, существует такой многочлен $g_i \in G$, что $g_i = x_i - h_i$, где $h_i \in K[y_1,\dots,y_m]$. При этом $x_i = h_i(f_1,\dots,f_m)$.

В частном случае, когда $m = n$, это даёт возможность распознавать автоморфизмы алгебры многочленов $K[x_1,\dots,x_n]$. Действительно, пусть задан эндоморфизм

$$\varphi\colon K[x_1,\dots,x_n] \to K[x_1,\dots,x_n],$$
$$\varphi(x_i) = f_i \in K[x_1,\dots,x_n], \quad 1 \leqslant i \leqslant n.$$

Рассмотрим алгебру $K[y_1,\dots,y_n,x_1,\dots,x_n]$ с упорядочением исключения, где $x_1,\dots,x_n$ старше, чем $y_1,\dots,y_n$, и идеал $I = \langle y_1 - f_1,\dots,y_n - f_n\rangle \lhd K[y_1,\dots,y_n,x_1,\dots,x_n]$. Пусть G — редуцированный базис Грёбнера идеала I. Тогда для гомоморфизма

$$\tilde\varphi\colon K[y_1,\dots,y_n] \to K[x_1,\dots,x_n], \quad \tilde\varphi(y_i) = f_i, \quad 1 \leqslant i \leqslant n,$$

имеем $\operatorname{Im}\tilde\varphi = K[x_1,\dots,x_n]$ тогда и только тогда, когда для каждого i, $1 \leqslant i \leqslant n$, существует такой элемент $g_i \in G$, что $g_i = x_i - h_i$, $h_i \in K[y_1,\dots,y_n]$, при этом $x_i = h_i(f_1,\dots,f_n)$, $1 \leqslant i \leqslant n$. Но в нашей ситуации условие $\operatorname{Im}\tilde\varphi = \operatorname{Im}\varphi = K[x_1,\dots,x_n]$ эквивалентно тому, что φ — автоморфизм алгебры $K[x_1,\dots,x_n]$ (в частности, $\operatorname{Ker}\varphi = \{0\}$, $G = \{g_1,\dots,g_n\}$). При этом эндоморфизм $\psi\colon K[x_1,\dots,x_n] \to K[x_1,\dots,x_n]$, заданный условиями $\psi(x_i) = h_i(x_1,\dots,x_n)$, $1 \leqslant i \leqslant n$, является автоморфизмом алгебры $K[x_1,\dots,x_n]$, обратным к автоморфизму φ.

Замечание 5.12.10. Базисы Грёбнера могут быть успешно применены в теории символьного интегрирования.

Пусть D — факториальное кольцо, $q = \sum\limits_{i=0}^{n} a_i x^i \in D[x] \setminus \{0\}$, $\deg q = n$. Обозначим через $\operatorname{Cont}(q)$ элемент НОД$(a_0, a_1,\dots,a_n)$ кольца D. Скажем, что многочлен q примитивен, если $\operatorname{Cont}(q) \in D^*$, примитивную часть $\operatorname{pp}(q)$ многочлена q определим как

$$\operatorname{pp}(q) = \frac{q}{\operatorname{Cont}(q)} \in D[x].$$

Пусть K — поле, $f, g \in K[x]$, $\deg(g) > 0$, $\deg(f) < \deg(g)$, многочлен g не имеет кратных корней (в алгебраическом замыкании $\bar K$ поля K), НОД$(f,g) = 1$ и G — редуцированный базис Грёбнера идеала алгебры многочленов $K[t,x]$, порождённого многочленами g и $f - tg'$ (при лексикографическом упорядочении, $x > t$), $G = \{g_1,\dots,g_m\}$, $\operatorname{Lm}(g_{i+1}) > \operatorname{Lm}_i(g_i)$, $i = 1,\dots,m-1$. Тогда

$$\int \frac{f(x)}{g(x)}\,dx = \sum_{i=1}^{m-1} \sum_{a, g_i(a)=0} a \ln\bigl(\operatorname{pp}_x(g_{i+1})(a,x)\bigr),$$

где

$$q_i = \frac{\operatorname{Cont}_x(g_i)}{\operatorname{Cont}_x(g_{i+1})} \in K[t]$$

(алгоритм Чиховского).

Например, пусть $f = x^4 - 3x^2 + 6$, $g = x^6 - 5x^4 + 5x^2 + 4 \in \mathbb{Q}[x]$, как в примере ??. Тогда $G = \{g_1, g_2\}$, $g_1 = t^2 + \dfrac{1}{4}$, $g_2 = x^3 + 2tx^2 - 3x - 4t$, — редуцированный базис Грёбнера идеала алгебры $\mathbb{Q}[t,x]$, порождённого элементами

$$x^6 - 5x^4 + 5x^2 + 4, \quad -6tx^5 + x^4 + 20tx^3 - 3x^2 - 10tx + 6$$

(при лексикографическом упорядочении, $t > x$),

$$\int \frac{x^4 - 3x^2 + 6}{x^6 - 5x^4 + 5x^2 + 4}\,dx = \sum_{a, 4a^2+1=0} a \ln(x^3 + 2ax^2 - 3x - 4a).$$

Алгоритмы теории базисов Грёбнера имеют широкие приложения в коммутативной алгебре и алгебраической геометрии (теория исключения, решение систем алгебраических уравнений, аффинная алгебраическая геометрия), в теории инвариантов, в теории кодирования, в теории систем доказательств, в теории автоматического доказательства геометрических теорем, в роботике, в теории систем, в целочисленной оптимизации, в теории сетей Петри, в статистике, в символьном суммировании, в дифференциальной алгебре. Во многих случаях решение поставленной алгебраической проблемы сводится к построению редуцированного базиса Грёбнера соответствующего идеала.

Алгоритмы построения базисов Грёбнера реализованы в большинстве систем компьютерной алгебры, предназначенных для проведения символьных вычислений на компьютере.

5.13.

Целые гауссовы числа:

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}.$$

$$z \in Z[i],\ z^{-1} \in \mathbb{Z}[i] \iff z = \pm 1, \pm i.$$

Простое гауссово число z нельзя представить в виде $z = z_1 \cdot z_2$, $z_1, z_1 \in \mathbb{Z}[i]$, $z_1, z_2 \neq \pm 1, \pm i$.

Докажите основную теорему арифметики для $\mathbb{Z}[i]$: любое целое гауссово число, отличное от ± 1, $\pm i$, единственным образом разлагается в произведение простых гауссовых чисел (с точностью до перестановки сомножителей и умножения на ± 1, $\pm i$).

$2 = (1+i)(1-i)$ не является простым гауссовыми.

Упражнение 5.13.1. Числа $1+i$, $2 \pm i$, 3 являются простыми гауссовыми.

Задача 5.13.2 (числа и многочлены Бернулли). Пусть $f \in \mathbb{R}[x]$, $\deg f = m$. Тогда существует единственный многочлен $g \in \mathbb{R}[x]$, такой что $\deg g = m + 1$, свободный член многочлена g равен нулю и $g(x) - g(x-1) = f(x)$. В частности, для каждого m существует единственный многочлен $g_m \in \mathbb{R}[x]$ степени $m + 1$ со свободным членом, равным нулю, такой что $g_m(x) - g_m(x-1) = x^m$. Многочлены $g_m(x)$ называются многочленами Бернулли.

Задача 5.13.3. Покажите, что число композиций числа n, в которых нет единиц, равно числу Фибоначчи f_{n-1} (под композицией числа понимается разбиение слагаемых, в котором учитывается порядок слагаемых; например, для числа 4 имеется пять разбиений ($p(4) = 5$): $4 = 4 = 1+3 = 2+2 = 1+1+1$; и восемь композиций: $4 = 4 = 1+3 = 3+1 = 2+2 = 1+1+2 = 1+2+1 = 2+1+1 = 1+1+1+1$).

Конец лекции № 23