МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ»

ЛИНЕЙНАЯ АЛГЕБРА

Часть 2.

Элементы общей алгебры.
Линейные пространства.
Евклидовы и унитарные пространства.
Линейные операторы.
Билинейные и квадратичные формы.

Учебно-методическое пособие для вузов

Составитель:
К.П. Лазарев

Воронеж 2023

Содержание

1. Элементы общей алгебры 5

1.1. Группа. Простейшие свойства группы 5
1.2. Подгруппа. Критерий подгруппы. Пересечение подгрупп 7
1.3. Произведение (сумма) множеств в группе. Произведение (сумма) подгрупп 9
1.4. Смежные классы группы и их свойства 10
1.5. Конечная группа. Теорема Лагранжа о порядке подгруппы 12
1.6. Нормальный делитель. Критерий нормального делителя. Фактор-группа 14
1.7. Изоморфизм групп 15
1.8. Кольцо. Свойства кольца. Подкольцо. Критерий подкольца 17
1.9. Поле 20
1.10. Кольцо классов вычетов 24
1.11. Комплексные числа. Поле комплексных чисел 28
1.12. Алгебраическая форма комплексных чисел. Арифметические операции для комплексных чисел в алгебраической форме. Геометрическое изображение комплексных чисел и операций сложения и вычитания 30
1.13. Тригонометрическая форма комплексного числа. Операции над комплексными числами в тригонометрической форме 33
1.14. Свойства модуля комплексных чисел. Свойства операции сопряжения 38
1.15. Возведение в степень. Извлечение корня 40
1.16. Показательная функция и логарифм от комплексной переменной. Возведение комплексного числа в комплексную степень 50
1.17. Кольцо многочленов от одной переменной 52
1.18. Деление многочленов с остатком. Схема Горнера 57
1.19. Делители многочленов и их свойства 62

1.20. Общий делитель и наибольший общий делитель многочленов. Нахождения наибольшего общего делителя многочленов. Представление наибольшего общего делителя многочленов через эти многочлены 65
1.21. Взаимная простота многочленов. Свойства многочленов, связанные со взаимной простотой 71
1.22. Корень многочлена. Кратность корня 74
1.23. Основная теорема алгебры для многочленов над C 76
1.24. Следствия из основной теоремы алгебры для многочленов над полем C 81
1.25. Многочлены над полем R 85
1.26. Собственные векторы и собственные значения квадратной матрицы 86

2. Линейные пространства над произвольным полем 89

2.1. Линейное пространство. Простейшие свойства пространства. Подпространство. Критерий подпространства 89
2.2. Линейная комбинация векторов. Линейная зависимость и линейная независимость векторов. База и ранг системы векторов 94
2.3. Базис линейного пространства. Координаты вектора в базисе и их свойства 98
2.4. Размерность линейного пространства. Связь понятий базиса и размерности 102
2.5. Матрица перехода от одного базиса к другому. Связь координат вектора в различных базисах 104
2.6. Линейная оболочка векторов 107
2.7. Сумма и пересечение подпространств 110
2.8. Размерность суммы двух подпространств 113
2.9. Прямая сумма подпространств 115

3. Евклидовы и унитарные пространства 117

3.1. Определение евклидова и унитарного пространства. Простейшие свойства скалярного произведения 117

3.2. Неравенство Коши–Буняковского. Длина вектора 121
3.3. Ортогональность 125
3.4. Метод ортогонализации. Существование ортонормированного базиса 127

4. Линейные операторы 129

4.1. Линейный оператор. Свойства линейного оператора 129
4.2. Матрица линейного оператора. Линейное выражение координат образа вектора при линейном отображении. Связь матриц линейного оператора в различных базисах 133
4.3. Собственные векторы и собственные значения линейного преобразования 137
4.4. Условия приведения матрицы линейного преобразования к диагональному виду. Жорданова форма матрицы линейного преобразования 142

5. Билинейные и квадратичные формы 145

5.1. Билинейные формы (функции), их матрицы. Квадратичные формы (функции) 145
5.2. Канонический вид квадратичной формы. Приведение к каноническому виду 148

Список литературы 150

1. Элементы общей алгебры

1.1. Группа. Простейшие свойства группы

Определение 1. *Множество* $\Gamma \neq \emptyset$ *называется группой, если*

(I) *во множестве* Γ *определена алгебраическая операция* $\circ\colon \Gamma \times \Gamma \to \Gamma$ (*т.е.* $\forall\, u, v \in \Gamma\, \exists\, w \in \Gamma$, *обозначаемый* $w = u \circ v$);

(II) *выполнены следующие аксиомы*

АГ1 $\forall\, a, b, c \in \Gamma \quad (a \circ b) \circ c = a \circ (b \circ c)$ — *операция ассоциативна;*

АГ2 $\exists\, e \in \Gamma\, \forall\, a \in \Gamma \quad a \circ e = e \circ a = a$ — *операция имеет нейтральный элемент* e;

АГ3 $\forall\, a \in \Gamma\, \exists\, a' \in \Gamma \quad a \circ a' = a' \circ a = e$ — *для каждого элемента* $a \in \Gamma$ *существует симметричный элемент* $a' \in \Gamma$.

(III) *Группа называется коммутативной или абелевой, если дополнительно выполнена аксиома*

АГ4 $\forall\, a, b \in \Gamma \quad a \circ b = b \circ a$ — *операция коммутативна.*

Обозначение: группа Γ *или* $\langle \Gamma, \circ \rangle$ — *группа* Γ *с операцией* $\circ$.

Группу с операцией сложения обозначают $\langle \Gamma, + \rangle$ *и называют аддитивной, нейтральный элемент называют нулём* (*обозначение* 0), *симметричный для* a *элемент называют противоположным* (*обозначение* $(-a)$).

Группу с операцией умножения обозначают $\langle \Gamma, \cdot \rangle$ *и называют мультипликативной, нейтральный элемент называют единицей* (*обозначение* e *или* 1), *симметричный для* a *элемент называют обратным* (*обозначение* a^{-1}). ■

Замечание. С помощью метода математической индукции доказывается утверждение: *результат вычисления операций для любого количества элементов группы не зависит от расстановки скобок.* Поэтому скобки можно не ставить.

Примеры: 1) $\langle Z, + \rangle$, $\langle Q, + \rangle$, $\langle R, + \rangle$, $\langle R^{m \times n}, + \rangle$ — это аддитивные абелевы группы;

2) $\langle Q \setminus 0, \cdot \rangle$, $\langle R \setminus 0, \cdot \rangle$ — мультипликативные абелевы группы;

3) $\langle \{A \in R^{n\times n} | \det(A) \neq 0\}, \cdot \rangle$, $\langle \{A \in R^{n\times n} | \det(A) = 1\}, \cdot \rangle$ — мультипликативные неабелевы группы.

4) множество $S(X)$ всех биективных отображений $f\colon X \to X$ с операцией суперпозиции отображений — мультипликативная неабелева группа. Частным случаем является множество биективных отображений конечного множества из n элементов в себя. Это группа называется симметрической группой n-го порядка.

Теорема 1.1.1 (Простейшие свойства группы).

Для любой группы $\langle G, \circ \rangle$ справедливы утверждения.

1. *Нейтральный элемент единственен.*
2. *Для любого элемента симметричный к нему элемент единственен.*
3. *Для любых элементов $x, y, z \in G$ из равенства $x \circ z = y \circ z$ или $z \circ x = z \circ y$ следует $x = y$ — сокращение обеих частей равенства на общий элемент.*
4. *Для любых элементов $a, b \in G$ уравнение $a \circ x = b$ имеет единственное решение $x = a' \circ b$ и уравнение $y \circ a = b$ имеет единственное решение $y = b \circ a'$, где a' — симметричный к a элемент.*
5. *Для любых элементов $a, b \in G$ $(a \circ b)' = b' \circ a'$. В частности, в группе $\langle G, \cdot \rangle$ $(ab)^{-1} = b^{-1}a^{-1}$, в группе $\langle G, + \rangle$ $-(a + b) = (-b) + (-a)$.*
6. *Для любого элемента $a \in G$ $(a')' = a$. В частности, в группе $\langle G, \cdot \rangle$ $(a^{-1})^{-1} = a$, в группе $\langle G, + \rangle$ $-(-a) = a$.*

Доказательство.

1. Если e_1 и e_2 — нейтральные элементы, то $e_1 \circ e_2 = e_2 \circ e_1 = e_1$, т.к. e_2 — нейтральный элемент, и $e_1 \circ e_2 = e_2 \circ e_1 = e_2$, т.к. e_1 — нейтральный элемент. Отсюда получаем $e_1 \circ e_2 = e_1$ и $e_1 \circ e_2 = e_2 \Rightarrow e_1 = e_2$.

2. Пусть b_1 и b_2— симметричные элементы к элементу a, т.е. $b_1 \circ a = a \circ b_1 = e$, $b_2 \circ a = a \circ b_2 = e$. Отсюда и из ассоциативности операции следует, что $(b_1 \circ a) \circ b_2 = b_1 \circ (a \circ b_2) \Rightarrow e \circ b_2 = b_1 \circ e \Rightarrow b_2 = b_1$.

3. Пусть $x \circ z = y \circ z$. Возьмём симметричный к z элемент z' и проведём над элементами обеих частей равенства $x \circ z = y \circ z$ справа операцию с z'. Получим $x \circ z \circ z' = y \circ z \circ z' \Rightarrow x \circ e = y \circ e \Rightarrow x = y$.

Аналогично из равенства $z \circ x = z \circ y$ получите $x = y$.

4. Проверим, что $x = a' \circ b$ решение уравнения $a \circ x = b$. Действительно, $a \circ x\big|_{x=a' \circ b} = a \circ a' \circ b = e \circ b = b$. Следовательно, при $x = a' \circ b$ уравнение $a \circ x = b$ превращается в тождество $b = b$.

Пусть $x = c$ — решение уравнения $a \circ x = b$, т.е. $a \circ c = b$. Возьмём симметричный к a элемент a' и для a' проведём операцию с элементами обеих частей равенства $a \circ c = b$. Получим $a' \circ a \circ c = a' \circ b$. Отсюда $e \circ c = a' \circ b$ и $c = a' \circ b$, т.е. решение $a' \circ b$ единственно.

Аналогично докажите утверждение для решения уравнения $y \circ a = b$.

5. Пусть $f = a \circ b$, $g = b' \circ a'$. Докажем, что $f \circ g = g \circ f = e$. Действительно, $f \circ g = (a \circ b) \circ (b' \circ a') = a \circ (b \circ b') \circ a' = a \circ e \circ a' = a \circ a' = e$. Аналогично, $g \circ f = (b' \circ a') \circ (a \circ b) = b' \circ (a' \circ a) \circ b = b' \circ e \circ b = b' \circ b = e$.

6. Из аксиомы группы АГ3 следуют равенства $a \circ a' = a' \circ a = e$, $a' \circ (a')' = (a')' \circ a' = e$. Приравняем $a' \circ (a')' = a' \circ a$ и после сокращения на a' получим $(a')' = a$. ■

1.2. Подгруппа. Критерий подгруппы. Пересечение подгрупп

Определение 1.

Подмножество группы называется её подгруппой, если оно само является группой относительно алгебраической операции в группе.

Итак, если $\langle G, \circ \rangle$ — группа, H — непустое подмножество группы и $\langle H, \circ \rangle$ — группа, то H является подгруппой группы G.■

Теорема 1.2.1 (Критерий подгруппы).

Пусть $\langle G, \cdot \rangle$ — группа и $H \neq \emptyset$ подмножество группы G.
H является подгруппой группы G $\Leftrightarrow$

(CP1) $\forall a, b \in H \;\; ab \in H$;

(CP2) $\forall a \in H \;\; a^{-1} \in H$.

Доказательство.

I. Пусть H является подгруппой группы $\langle G, \cdot \rangle$. Тогда по определению подгруппы $\langle H, \cdot \rangle$ — группа и из определения группы следует выполнение (СР1) и (СР2).

II. Пусть $H \neq \emptyset$ и выполнены (СР1) и (СР2).

Из (СР1) следует, что для любых элементов $a, b \in H$ операция ab определена в H. Тогда выполнена аксиома группы АГ1 $\forall\, a, b, c \in H \quad (ab)c = a(bc)$, поскольку по условию (СР1) все элементы $ab, (ab)c, bc, a(bc)$ из H, $H \subseteq G$ и АГ1 выполнена в группе G.

Покажем, что единица e группы G принадлежит H.

Возьмём любой элемент $a \in H$. Из (СР2) следует, что обратный элемент $a^{-1} \in H$, из (СР1) получаем, что $aa^{-1} \in H$, а по аксиоме АГ3 в группе G $aa^{-1} = e$. Следовательно, $e \in H$.

Отсюда следует выполнение в H аксиом АГ2 и АГ3. Действительно, $\forall\, a \in H$ в H выполнены равенства $ae = ea = a$ и $aa^{-1} = a^{-1}a = e$, так как они выполнены в группе G, $H \subseteq G$ и элементы $a, a^{-1}, e, ae, ea \in H$.■

С помощью критерия подгруппы легко обосновать следующие примеры.

Примеры.

1. $\langle Z, + \rangle$, $\langle Q, + \rangle$ — подгруппы группы $\langle R, + \rangle$.

2. Для произвольного натурального числа m обозначим через mZ все целые числа кратные числу m. Тогда $\langle mZ, + \rangle$ — подгруппа групп $\langle Z, + \rangle, \langle Q, + \rangle$, $\langle R, + \rangle$.■

Теорема 1.2.2 (Пересечение подгрупп).

Пусть $\langle G, \cdot \rangle$ — группа и H_α — семейство её подгрупп, зависящих от индекса α, который пробегает множество A. Тогда пересечение всех подгрупп $\bigcap\limits_{\alpha \in A} H_\alpha$ является подгруппой группы G.

Доказательство. Обозначим $F = \bigcap\limits_{\alpha \in A} H_\alpha$.

Поскольку все подгруппы содержит единицу e группы G, то и их пересечение F содержат e, т. е. F не пусто.

Для доказательства утверждения применим критерий подгруппы — теорему 1.2.1.

СР1: $a, b \in F \Rightarrow ab \in F$.

Пусть $a, b \in F = \bigcap\limits_{\alpha \in A} H_\alpha \Rightarrow \forall \alpha \in A \;\; a, b \in H_\alpha \Rightarrow \forall \alpha \in A \;\; ab \in H_\alpha \Rightarrow ab \in$ $\in \bigcap\limits_{\alpha \in A} H_\alpha = F \Rightarrow$ выполнено CP1.

CP2: $a \in F \Rightarrow a^{-1} \in F$.

Пусть $a \in F = \bigcap\limits_{\alpha \in A} H_\alpha \Rightarrow \forall \alpha \in A \;\; a \in H_\alpha \Rightarrow \forall \alpha \in A \;\; a^{-1} \in H_\alpha \Rightarrow a \in$ $\in \bigcap\limits_{\alpha \in A} H_\alpha = F \Rightarrow$ выполнено CP2.

Следовательно, H — подгруппа. ■

1.3. Произведение (сумма) множеств в группе. Произведение (сумма) подгрупп

Определение 1. *Пусть S и T — два подмножества группы G.*

Для мультипликативной группы произведением множеств S и T называется множество $\{z | z \in G,\, z = ab,\, a \in S,\, b \in T\}$. Это множество обозначается ST.

Для аддитивной группы суммой множеств S и T называется множество $\{z | z \in G,\, z = a + b,\, a \in S,\, b \in T\}$. Это множество обозначается $S + T$. ■

Замечания. 1. Будем использовать мультипликативную терминологию.

2. Для ST будем писать а) aT вместо $\{a\}T$, если $S = \{a\}$ и T содержит более одного элемента, б) Sb вместо $S\{b\}$, если S содержит более одного элемента и $T = \{b\}$, в) ab вместо $\{a\}\{b\}$, если $S = \{a\}$ и $T = \{b\}$. ■

Теорема 1.3.1.

1° *Если H_1, H_2 подгруппы группы $\langle G, \cdot \rangle$ и $H_1H_2 = H_2H_1$, то произведение подгрупп является подгруппой G.*

2° *Произведение подгрупп коммутативной группы $\langle G, \cdot \rangle$ является её подгруппой.*

Доказательство.

1° Обозначим $H = H_1H_2 = H_2H_1$ и применим к H критерий подгруппы.

$\boxed{\text{CP1}}$ $a, b \in H \Rightarrow ab \in H$.

$\begin{cases} a, b \in H \\ H = H_1H_2 \end{cases} \Rightarrow \begin{cases} a = a_1a_2,\, a_1 \in H_1,\, a_2 \in H_2 \\ b = b_1b_2,\, b_1 \in H_1,\, b_2 \in H_2 \end{cases} \Rightarrow ab = (a_1a_2)(b_1b_2)$ {ис-

пользуем ассоциативность$\} \Rightarrow ab = a_1(a_2b_1)b_2$.

Так как $a_2b_1 \in H_2H_1$ и $H_2H_1 = H_1H_2$, то $a_2b_1 \in H_1H_2$ и, следовательно, $a_2b_1 = r_1r_2$, где $r_1 \in H_1$, $r_2 \in H_2$. Тогда $ab = a_1(r_1r_2)b_2$ {используем ассоциативность$\} \Rightarrow ab = (a_1r_1)(r_2b_2)$ {здесь $a_1r_1 \in H_1$, $a_2r_2 \in H_2\} \Rightarrow ab \in$ $\in H_1H_2 = H \Rightarrow ab \in H$.

$\boxed{\text{CP2}}$ $a \in H \Rightarrow a^{-1} \in H$.

$a \in H = H_1H_2 \Rightarrow a = a_1a_2$, $a_1 \in H_1$, $a_2 \in H_2 \Rightarrow a^{-1} = a_2^{-1}a_1^{-1}$, $a_2^{-1} \in$ $\in H_2$, $a_1^{-1} \in H_1 = H \Rightarrow a^{-1} \in H_2H_1 \Rightarrow a^{-1} \in H$.

2° Пусть H_1, H_2 — подгруппы коммутативной группы $\langle G, \cdot \rangle$. Из коммутативности группы получаем $\forall a_1 \in H_1$, $a_2 \in H_2$ $a_1a_2 = a_2a_1$. Тогда $H_1H_2 =$ $= H_2H_1$ и в силу 1° утверждение 2° верно. ■

Замечание. Изменив обозначения произведения на сумму, получим следующее утверждение. *Если H_1, H_2 подгруппы группы $\langle G, + \rangle$ и $H_1 + H_2 =$ $= H_2 + H_1$, то сумма подгрупп является подгруппой.* ■

Задача 1. *Если $H_1, H_2, \ldots, H_n$ подгруппы группы $\langle G, \cdot \rangle$ и при всех $i, j \in \overline{1, n}$ $H_iH_j = H_jH_i$, то произведение подгрупп является подгруппой.*

1.4. Смежные классы группы и их свойства

Определение 1. *Пусть H — подгруппа группы G, a — элемент группы G. Множество aH (множество Ha) называется левым (правым) смежным классом группы G по подгруппе H, порождённым элементом a.* ■

Теорема 1.4.1 (Простейшие свойства смежных классов).
Пусть H — подгруппа группы G. Тогда выполнены следующие свойства.

1° $\forall\, a \in G$ $a \in aH$ *и* $a \in Ha$.

2° *Левый (правый) смежный класс состоит из элементов группы, причём любой элемент группы входит в некоторый левый и в некоторый правый смежный класс.*

3° *Подгруппа H является левым и правым смежным классом.*

4° *В коммутативной группе $aH = Ha$ $\forall\, a \in G$.*

5° *Смежный класс порождается любым своим элементом.*

6° *Любые два левых (правых) смежных класса либо совпадают, либо не пересекаются.*

Доказательство.

1° Поскольку $e \in H$, имеем $a = ae \in aH$ и $a = ea \in Ha$.

2° По определению для $a \in G$ имеем $aH \subseteq G$ и $Ha \subseteq G$. Из 1° следует, что $a \in aH$ и $a \in Ha$.

3° $H = eH = He$.

4° В коммутативной группе G для всех элементов выполнено равенство $ab = ba$, поэтому $aH = Ha \ \forall a \in G$.

5° Для смежного класса aH докажем, что если $b \in aH$, то $aH = bH$.

Из $b \in aH$ следует, что $b = ah_1$, где $h_1 \in H$.

I. Проверим, что $bH \subseteq aH$.

Берём любой элемент $z \in bH \Rightarrow z = bh_2$, где $h_2 \in H$ и $b = ah_1 \Rightarrow$ $\Rightarrow z = a(h_1h_2)$, где $h_1h_2 \in H \Rightarrow z \in aH$.

II. Проверим, что $aH \subseteq bH$.

Берём любой элемент $z \in aH \Rightarrow z = ah_3$, где $h_3 \in H$. Далее из равенства $b = ah_1$ выразим $a = bh_1^{-1}$, где $h_1^{-1} \in H$. Тогда $z = ah_3 = b(h_1^{-1}h_3)$, где $h_1^{-1}h_3 \in H$. Следовательно, $z \in bH$ и $aH \subseteq bH$.

Из $bH \subseteq aH$ и $aH \subseteq bH$ следует равенство $aH = bH$.

Докажите самостоятельно, что при $b \in Ha$ выполнено $Ha = Hb$.

6° Из утверждения 5° следует, что если два левых смежных класса aH и bH имеют общий элемент z, то $zH = aH$, $zH = bH$, т. е. смежные классы совпадают. А если смежные классы не имеют общих элементов, то они не пересекаются. Такое же рассуждение верно для правых смежных классов.

Следствие. Группа разбивается на непересекающиеся левые (правые) смежные классы по заданной подгруппе. Это разбиение называется *левосторонним (правосторонним) разложением группы по подгруппе.* ■

Примеры смежных классов.

Группа	Подгруппа	Смежные классы
$\langle Z, + \rangle$	$H = mZ = \{mk \mid k \in Z\}$, $m \in N, m \rangle 1$	Для $a \in Z : a + H = H + a = \{n \mid n = a + mk,\ k \in Z\}$. Здесь m различных классов: $a + H$, $a = 0, 1, \ldots, m - 1$. Числа из класса $a + H$ при делении на m дают остаток a.

1.5. Конечная группа. Теорема Лагранжа о порядке подгруппы

Определение 1. *Группа, состоящая из конечного числа элементов, называется конечной группой. Число элементов n конечной группы $G = \{a_1, a_2, \ldots, a_n\}$ называется порядком группы и обозначается* $\mathrm{card}(G)$, $|G|$, $\overline{\overline{G}}$. ■

Алгебраическая операция в конечной группе может быть представлена «таблицей умножения».

1. Построим группу 1-го порядка.

Возьмём множество $G = \{e\}$ из одного элемента, который назовём единицей. Операцию определим равенством $ee = e$ или таблицей на рис. 1. Выполнение аксиом очевидно, АГ1: $(ee)e = e(ee)$ (ассоциативность); АГ2: $ee = ee = e$ (единица — это e); АГ3: $ee = ee = e$ (обратный к e — это e); АГ4: $ee = ee$ (коммутативность).

2. Построим группу 2-го порядка.

Возьмём множество из двух элементов $G = \{e, a\}$, e — единица. Определим операцию как в таблице на рис. 2: 1) $ee = e$, 2) $ea = ae = a$, 3) $aa = x$ и выберем x. Для группы в каждой строке (в каждом столбце) таблицы все эле-

1-й \ 2-й	e
e	e

Рис. 1

1-й \ 2-й	e	a
e	e	a
a	a	x

Рис. 2

1-й \ 2-й	e	a
e	e	a
a	a	e

Рис. 3

менты различны, поскольку из равенства $fg = fh\,(gf = hf)$ после сокращения получаем $g = h$, что не верно. Значит, $x = e$ (см. таблицу на рис. 3).

Проверим аксиомы группы.

Из таблицы следует, что $\forall\, x, y \in G \quad xy = yx$, т.е. выполнена АГ4. Равенства 1) – 2) обеспечивают выполнение АГ2. Равенства 1), 3) обеспечивают наличие АГ3: $e^{-1} = e$, $a^{-1} = a$.

Проверим АГ1: $\forall x, y, z \in G \quad (xy)z = x(yz)$.

I. Если среди трёх элементов хотя бы один равен e, то а) $(ef)g = e(fg)$, б) $(fe)g = f(eg)$, в) $(fg)e = f(ge)$, т. к. $(ef)g = fg$, $e(fg) = fg$, $(fe)g = fg$, $f(eg) = fg$, $(fg)e = fg$, $f(ge) = fg$.

II. Если среди трёх элементов ни один не равен e, то все три элемента равны a и $(aa)a = a(aa)$ в силу АГ4.

Итак, выполнена аксиома АГ1.

Следовательно, G — абелева группа.

3. Построим группу 3-го порядка.

Возьмём множество из трёх элементов $G = \{e, a, b\}$, e — единица. Определим операцию как в таблице на рис. 4: 1) $ee = e$, 2) $ea = ae = a$, 3) $eb = be = b$, 4) $aa = x$, 5) $ab = y$, 6) $ba = z$, 7) $bb = u$ и выберем x, y, z, u. Для группы в каждой строке (в каждом столбце) таблицы все элементы различны. Во 2-й строке $x \neq a$, $y \neq a$, в 3-м столбце $y \neq b$, тогда $y = e$, $x = b$ (см. таблицу на рис. 5). Далее во 2-м и 3-м столбцах $z = e$, $v = a$ (см. таблицу на рис. 6).

1-й \ 2-й	e	a	b
e	e	a	b
a	a	x	y
b	b	z	u

Рис. 4

1-й \ 2-й	e	a	b
e	e	a	b
a	a	b	e
b	b	z	v

Рис. 5

1-й \ 2-й	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Рис. 6

Проверим аксиомы группы.

АГ4 выполнена, т. к. из таблицы следует, что $\forall\, x, y \in G \;\; xy = yx$.

АГ2 выполнена в силу равенств 1) – 3).

АГ3 следует из равенств $ee = ab = ba = e \Rightarrow e^{-1} = e$, $a^{-1} = b$, $b^{-1} = a$.

Проверим АГ1: $\forall x, y, z \in G \;\; (xy)z = x(yz)$.

I. Если из трёх элементов хотя бы один равен e, то рассуждаем как в группе 2-го порядка.

II. Если среди трёх элементов нет e, то рассмотрим следующие случаи.

1) Три a или три b : $(aa)a = a(aa)$ и $(bb)b = b(bb)$ в силу коммутативности (АГ4).

2) Два a и один b :

$(aa)b = a(ab)$, т. к. $(aa)b = bb = a$ и $a(ab) = ae = a$;

$(ab)a = a(ba)$, т. к. $(ab)a = ea = a$ и $a(ba) = ae = a$;

$(ba)a = b(aa)$, т. к. $(ba)a = ea = a$ и $b(aa) = bb = a$.

3) Один a и два b :

$(ab)b = a(bb)$, т. к. $(ab)b = eb = b$ и $a(bb) = aa = b$;

$(ba)b = b(ab)$, т. к. $(ba)b = eb = b$ и $b(ab) = be = b$;
$(bb)a = b(ba)$, т. к. $(bb)a = aa = b$ и $b(ba) = be = b$.
Итак, G — абелева группа. ■

Теорема 1.5.1 (Лагранж). *Порядок подгруппы является делителем порядка конечной группы.*

Доказательство. Пусть $H = \{h_1, h_2, \ldots, h_k\}$ — подгруппа группы $G = \{g_1, g_2, \ldots, g_n\}$, причём $h_i \neq h_j\,(i \neq j)$, $g_i \neq g_j\,(i \neq j)$. Нужно доказать, что n делится на k. Рассмотрим левостороннее разложение группы G по подгруппе H. Оно состоит из конечного числа (обозначим его m) непересекающихся левых смежных классов вида a_iH, где $a_i \in G$ — порождающий элемент класса, $i = \overline{1, m}$, причём $m \leqslant n$. Имеем $G = \bigcup\limits_{i=1}^{m} a_iH$ и количество элементов группы равно сумме числа элементов всех классов. Покажем что все k элементов любого класса $a_iH = \{a_ih_1, a_ih_2, \ldots, a_ih_k\}$ различны. Действительно, из равенства $a_ih_p = a_ih_q$ при $p \neq q$ после сокращения получаем $h_p = h_q$, но $h_p \neq h_q$. Таким образом,количество всех элементов группы равно mk, т. е. $n = mk$. ■

1.6. Нормальный делитель. Критерий нормального делителя. Фактор-группа

Определение 1. *Подгруппа H группы G называется нормальным делителем, если для каждого элемента g группы его левый и правый смежные классы по подгруппе H равны, т.е. $gH = Hg$.* ■

Очевидно, что в коммутативной группе, любая подгруппа является нормальным делителем.

Теорема 1.6.1 (Критерий нормального делителя).
Подгруппа H группы G является нормальным делителем тогда и только тогда, когда $g^{-1}hg \in H$ при всех $g \in G$ и $h \in H$.

Доказательство. I. Пусть H — нормальный делитель группы G, т. е. $\forall\, g \in G\ gH = Hg$. Тогда $Hg \subseteq gH$, а это означает, что $\forall\, h \in H\, \exists\, h_1 \in H \mid hg = gh_1$. Умножив слева на обратный g^{-1}, получим $g^{-1}hg = h_1$, где $h_1 \in H$. Следовательно, $\forall\, g \in G,\ h \in H \mid g^{-1}hg \in H$.

II. Пусть $\forall\, g \in G,\ h \in H \mid g^{-1}hg \in H$. Возьмём любые элементы $h \in H$ и $g \in G$, тогда $g^{-1} \in G$, $g^{-1}hg \in H$ и $ghg^{-1} = (g^{-1})^{-1}hg^{-1} \in H$. Отсюда следует, что существуют элементы $h_1, h_2 \in H$ такие, что $g^{-1}hg = h_1$, $ghg^{-1} = h_2 \Leftrightarrow hg = gh_1$, $gh = h_2g$. Тогда $Hg \subseteq gH$ и $gH \subseteq Hg$, т. е. $gH = Hg$.

Теорема 1.6.2 (Фактор-группа).

Все различные левые (правые) смежные классы по нормальному делителю образуют группу относительно умножения подмножеств группы.

Доказательство.

I. а) Операция умножения подмножеств группы ассоциативна. Действительно, если A, B, C — подмножества группы G, то $(AB)C = \{(ab)c \mid a \in A, b \in B, c \in C\} = \{a(bc) \mid a \in A, b \in B, c \in C\} = A(BC)$.

б) Проверим, что произведение левых смежных классов aH и bH по нормальному делителю H является левым смежным классом $(ab)H$.
Преобразуем $(aH)(bH) = \{$ассоциативность произведения$\} = a(Hb)H = \{H-$ нормальный делитель, тогда $Hb = bH\} = a(bH)H = \{$из ассоциативности$\}= =(ab)(HH) = \{H-$ подгруппа, поэтому$HH = H\} = (ab)H$. Получился левый смежный класс, порождённый элементом ab.

II. 1) Ассоциативность умножения классов:

$$((aH)(bH))(cH) = ((ab)H)(cH) = (abc)H = (aH)((bc)H) = (aH)((bH)(cH)).$$

2) Нейтральный элемент для умножения классов $eH = H$:

$$(aH)(eH) = (ae)H = aH,\ (eH)(aH) = (ea)H = aH.$$

3) Обратный классу aH — это класс $a^{-1}H$:

$$(aH)(a^{-1}H) = (aa^{-1})H = eH = H,\ (a^{-1}H)(aH) = (a^{-1}a)H = eH = H. \blacksquare$$

1.7. Изоморфизм групп

Определение 1. *Группа $\langle \Gamma_1, \circ_1 \rangle$ изоморфна группе $\langle \Gamma_2, \circ_2 \rangle$, если существует биективное отображение $f\colon \Gamma_1 \to \Gamma_2$ такое, что $\forall a, b \in \Gamma_1$ выполнено $f(a \circ_1 b) = f(a) \circ_2 f(b)$, т. е. образ «произведения» $a \circ_1 b$ равен «произведению» образов $f(a) \circ_2 f(b)$. Обозначение: $\Gamma_1 \cong \Gamma_2$. При этом говорят, что отображение f — изоморфизм.* $\blacksquare$

Примеры.

1. Группа целых чисел $\langle Z, +\rangle$ и группа чётных чисел $\langle 2Z, +\rangle$ изоморфны.

Отображение $f\colon Z \to 2Z$, $f(n) = 2n$, $n \in Z$ есть изоморфизм: 1) f — биекция, 2) $\forall n_1, n_2 \in Z \;\; f(n_1 + n_2) = 2(n_1 + n_2) = 2n_1 + 2n_2 = f(n_1) + f(n_2)$.

2. Пусть $\Gamma = \{x | x \in R,\, x > 0\}$, тогда $\langle \Gamma, \cdot\rangle \cong \langle R, +\rangle$. Изоморфизм $f\colon \Gamma \to R$, $f(x) = \ln(x)$, т.к. 1) f — биекция, 2) $f(x_1x_2) = \ln(x_1x_2) = \ln(x_1) + \ln(x_2) = f(x_1) + f(x_2)$. ■

Теорема 1.7.1 (Простейшие свойства изоморфизма).

Если группа $\langle \Gamma_1, \circ_1\rangle$ изоморфна группе $\langle \Gamma_2, \circ_2\rangle$ и $f\colon \Gamma_1 \to \Gamma_2$ — изоморфизм, то выполнены следующие свойства.

1° $f^{-1}\colon \Gamma_2 \to \Gamma_1$ — *также изоморфизм.*

2° *Образ единицы $e_1 \in \Gamma_1$ равен единице $e_2 \in \Gamma_2$:* $f(e_1) = e_2$.

3° *Образ обратного элемента для $a \in \Gamma_1$ является обратным для образа $f(a)$, т. е. $f(a^{-1}) = (f(a))^{-1}$.*

Доказательство.

1° Если $f\colon \Gamma_1 \to \Gamma_2$ — биективное отображение, то существует обратное отображение $f^{-1}\colon \Gamma_2 \to \Gamma_1$ также биективное. Покажем, что f^{-1} — изоморфизм, т. е.

$$\forall c, d \in \Gamma_2 \; f^{-1}(c \circ_2 d) = f^{-1}(c) \circ_1 f^{-1}(d).$$

Введём обозначения $a = f^{-1}(c)$, $b = f^{-1}(d)$. Тогда $a, b \in \Gamma_1$, $f(a) = c$, $f(b) = d$. Поскольку f — изоморфизм, то $f(a \circ_1 b) = f(a) \circ_2 f(b) = c \circ_2 d$. Следовательно, $c \circ_2 d = f(a \circ_1 b)$. Применим f^{-1}, и получим $f^{-1}(c \circ_2 d) = f^{-1}(f(a \circ_1 b))$. Поскольку $\forall x \in \Gamma_1 \; f^{-1}(f(x)) = x$, имеем $f^{-1}(c \circ_2 d) = a \circ_1 b$. Подставив $a = f^{-1}(c)$, $b = f^{-1}(d)$, получим $f^{-1}(c \circ_2 d) = f^{-1}(c) \circ_1 f^{-1}(d)$, ч.т.д.

2° Применим f к равенству $e_1 = e_1 \circ_1 e_1$, получим $f(e_1) = f(e_1 \circ_1 e_1)$. Запишем $f(e_1) = e_2 \circ_2 f(e_1)$, т. к. e_2 — единица, и $f(e_1 \circ_1 e_1) = f(e_1) \circ_2 f(e_1)$, т. к. f — изоморфизм. Тогда равенство примет вид $e_2 \circ_2 f(e_1) = f(e_1) \circ_2 f(e_1)$. Сократив на $f(e_1)$, получим $e_2 = f(e_1)$, ч.т.д.

3° Применим к обеим частям равенства $a \circ_1 a^{-1} = e_1$ изоморфизм f, и получим $f(a \circ_1 a^{-1}) = f(e_1) \Leftrightarrow f(a) \circ_2 f(a^{-1}) = e_2$. Отсюда $f(a^{-1}) = (f(a))^{-1}$. ■

1.8. Кольцо. Свойства кольца. Подкольцо. Критерий подкольца

Определение 1. *Непустое множество K называется кольцом, если*

(I) *на K определены алгебраические операции сложения и умножения,*

(II) *выполнены следующие аксиомы:*

АК1 $\forall\, a,\, b, c \in K \quad (a+b)+c = a+(b+c)$,

АК2 $\exists\, 0 \in K\ \forall\, a \in K \quad a+0=a$,

АК3 $\forall\, a \in K\, \exists\, (-a) \in K \quad a+(-a)=0$,

АК4 $\forall\, a, b \in K \quad a+b=b+a$,

АК5 $\forall\, a, b, c \in K \quad (ab)c = a(bc)$,

АК6 $\forall\, a, b, c \in K \quad (a+b)c = ac+bc,\ \ c(a+b) = ca+cb$.

Обозначение: кольцо K или $\langle K, +, \cdot\rangle$.

Группу $\langle K, +\rangle$ называют аддитивной группой кольца.

Если в K существует нейтральный элемент (единица) для умножения, т. е. $\exists e \in K\, \forall a \in K\ \ ae = ea = a$, то K называется кольцом с единицей.

Если операция умножения коммутативна, т. е. $\forall a, b \in K\ \ ab = ba$, то K называется коммутативным кольцом.

Ненулевые элементы a и b кольца K называются делителями нуля, если $ab = 0$ (a — левый, b — правый делитель нуля). ■

Названия аксиом:
АК1 — ассоциативность сложения,
АК2 — наличие для сложения нейтрального элемента 0,
АК3 — наличие для каждого элемента противоположного элемента,
АК4 — коммутативность сложения,
АК5 — ассоциативность умножения,
АК6 — дистрибутивность умножения слева и справа относительно сложения. ■

Примеры. Z, Q, R — коммутативные кольца с единицей. $2Z$ — коммутативное кольцо без единицы. Множество квадратных матриц $R^{n\times n}$ — некоммутативное кольцо с единицей.

Теорема 1.8.1 (Простейшие свойства кольца).

Для любого кольца K справедливы утверждения.

1. *Нуль единственен.*
2. *Для любого элемента противоположный к нему элемент единственен.*
3. *Из условий $a, b, c \in K$, $a + c = b + c$ или $c + a = c + b$ следует равенство $a = b$.*
4. *Для любых элементов $a, b \in K$ уравнения $x + a = b\,(a + x = b)$ имеют единственное решение $x = b + (-a) = (-a) + b$ это решение называется разностью b и a и обозначается $b - a$. Итак, $b - a = b + (-a)$ и эта формула определяет операцию вычитания.*
5. *Умножение дистрибутивно относительно вычитания (справа и слева) $\forall a, b, c \in K \quad (a - b)c = ac - bc$ и $c(a - b) = ca - cb$.*
6. *$\forall a \in K \quad a0 = 0$ и $0a = 0$.*
7. *В кольце с единицей единица единственна (и обозначается 1).*
8. *В кольце K без делителей нуля из условия $a, b \in K$ и $ab = 0$ следует условие $a = 0$ или $b = 0$.*
9. *В кольце K без делителей нуля из условия $a, b, c \in K$, $c \neq 0$ и $ac = bc$ или $ca = cb$ следует равенство $a = b$.*
10. *В кольце с единицей для любого элемента выполнены равенства $(-a) = (-1)a = a(-1)$.*
11. *В кольце для любых элементов выполнены равенства а) $(-a)b = a(-b) = -(ab)$, б) $(-a)(-b) = ab$.*
12. *В кольце с единицей, содержащем не менее двух элементов: $1 \neq 0$.* ■

Доказательство.

Свойства 1 – 4 справедливы, т.к. $\langle K, + \rangle$ — коммутативная группа.

5. Для разности $a - b$ верно $(a - b) + b = a$. Умножим равенство справа на c: $((a - b) + b)c = ac$. Из дистрибутивности умножения справа относительно сложения следует $(a - b)c + bc = ac$. Используя определение разности, получаем $(a - b)c = ac - bc$.

Равенство $c(a-b)=ca-cb$ докажите самостоятельно.

6. Разность равных элементов равна $0:\ ab-ab=0,\ a-a=0,\ b-b=0$. Отсюда, пользуясь дистрибутивностью умножения относительно вычитания, получим $a(b-b)=0,\ \ (a-a)b=0$. Следовательно, $a0=0,\ \ 0b=0$.

7. Если e_1 и e_2 — единицы, то $e_1e_2=e_2e_1=e_1$ и $e_2e_1=e_1e_2=e_2$. Приравняв эти выражения, получаем $e_1=e_2$.

8. Предположим противное: $a\neq 0$ и $b\neq 0$. Тогда $ab=0,\ a\neq 0,\ b\neq 0$, т. е. a и b — делители нуля, что противоречит условию. Следовательно, $a=0$ или $b=0$.

9. Пусть $c\neq 0$ и $ac=bc\Rightarrow ac-bc=0\Rightarrow (a-b)c=0$. Из свойства 8 следует, что $a-b=0$, тогда $a=0+b\Rightarrow a=b$.

Аналогично докажите, что из $c\neq 0$ и $ca=cb$ следует $a=b$.

10. Противоположный элемент единственен. Сложим a и $(-1)a$, получим $a+(-1)a=1a+(-1)a=(1+(-1))a=0a=0$. Отсюда $(-1)a$ является противоположным a, т. е. $(-1)a=(-a)$.

Аналогично докажите, что $a(-1)=(-a)$.

11. а) $ab+(-a)b=(a+(-a))b=0b=0\Rightarrow(-a)b=-(ab)$.

$ab+a(-b)=a(b+(-b))=a0=0\Rightarrow a(-b)=-(ab)$.

б) Из $c+(-c)=0$ следует $c=-(-c)$. Используя это и а), получим

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

12. Пусть $1=0$, тогда существует элемент $a\in K$ такой, что $a\neq 0$. Умножим равенство $1=0$ на a и получим $a\cdot 1=a\cdot 0\Rightarrow a=0$, но $a\neq 0$. ■

Определение 2. *Подмножество M кольца K называется подкольцом K, если M само является кольцом при тех же операциях сложения и умножения, которые заданы в кольце K.* ■

Теорема 1.8.2 (Критерий подкольца).

Пусть K — кольцо и M — непустое подмножество кольца K. M является подкольцом кольца K тогда и только тогда, когда M замкнуто относительно операций сложения, умножения и перехода к противоположному элементу, т. е.

(KK1) $\forall a,b\in M\ \ a+b\in M$;

(KK2) $\forall a,b\in M\ \ ab\in M$;

(КК3) $\forall a \in M \quad (-a) \in M$. ■

Доказательство.

1. Если M — подкольцо K, то M — кольцо и выполнены (КК1) – (КК3).

2. Пусть в M выполнены (КК1) – (КК3). Тогда в M определены операции сложения, умножения и любой элемент $a \in M$ имеет противоположный элемент $(-a) \in M$. Сложив эти элементы получим $a + (-a) \in M$, т. е. $0 \in M$. Тогда аксиомы кольца, выполненные в K, будут выполнены в M. Следовательно M — кольцо. ■

1.9. Поле

Определение 1. *Множество P называется полем, если на множестве P определены две алгебраические операции сложения и умножения и выполнены следующие аксиомы*

АП1 $\forall a, b, c \in P \quad (a + b) + c = a + (b + c)$;

АП2 $\exists 0 \in P \, \forall a \in P \quad a + 0 = a$;

АП3 $\forall a \in P \, \exists (-a) \in P \quad a + (-a) = 0$;

АП4 $\forall a, b \in P \quad a + b = b + a$;

АП5 $\forall a, b, c \in P \quad (ab)c = a(bc)$;

АП6 $\exists (e \in P \; e \neq 0) \, \forall a \in P \quad ae = a$;

АП7 $\forall (a \in P \; a \neq 0) \, \exists \, a^{-1} \in P \quad aa^{-1} = e$;

АП8 $\forall a, b \in P \quad ab = ba$;

АП9 $\forall a, b, c \in P \quad (a + b)c = ac + bc \bigwedge c(a + b) = ca + cb$.

Замечания.

В силу коммутативности сложения и умножения выполнены равенства.

1. $\exists 0 \in P \, \forall a \in P \quad a + 0 = 0 + a = a$.
2. $\forall a \in P \, \exists (-a) \in P \quad a + (-a) = (-a) + a = 0$.
3. $\exists e \in P \setminus \{0\} \, \forall a \in P \quad ae = ea = e$.
4. $(\forall a \in P, a \neq 0) \, \exists a^{-1} \in P \quad aa^{-1} = a^{-1}a = e$. ■

Названия аксиом:
АП1 — ассоциативность сложения,
АП2 — наличие для сложения нейтрального элемента 0,
АП3 — наличие для каждого элемента противоположного элемента,
АП4 — коммутативность сложения,
АП5 — ассоциативность умножения,
АП6 — существование для умножения нейтрального элемента — единицы, отличной от нуля,
АП7 — существование для каждого ненулевого элемента обратного элемента,
АП8 — коммутативность умножения,
АП9 — дистрибутивность умножения относительно сложения справа и слева. ■

Примеры.

1. Множество рациональных чисел.

2. Множество действительных чисел.

3. Множество действительных чисел вида $a + b\sqrt{2}$ с любыми рациональными a и b.

Теорема 1.9.1 (Простейшие свойства поля).

Поле обладает всеми свойствами коммутативного кольца с единицей без делителей нуля.

1. *Нулевой элемент единственен.*

2. *Для любого элемента противоположный к нему элемент единственен.*

3. *Для любых элементов $a, b, c \in P$ из $a+c = b+c$ или $c+a = c+b$ следует $a = b$.*

4. *Для любых элементов $a, b \in P$ существует, и притом единственное, решение $x = b + (-a) = (-a) + b$ уравнения $x + a = b\,(a + x = b)$; это решение называется* **разностью** *b и a, и обозначается $b - a$. Итак, $b - a = b + (-a)$ и эта формула определяет операцию* **вычитания**.

5. *Умножение дистрибутивно относительно вычитания, т.е.*
   $\forall a, b, c \in P \quad (a - b)c = ac - bc$ *и* $c(a - b) = ca - cb$.

6. $\forall a \in P \quad a0 = 0$ *и* $0a = 0$.

7. *Единица единственна (и обозначается 1).*

8. *Обратный элемент для любого ненулевого элемента единственен.*

9. *Для любых элементов* $a, b, c \in P$ *из* $c \neq 0$ *и* $ac = bc$ *или* $ca = cb$ *следует* $a = b$.

10. *Для любых элементов* $a, b \in P$ *из* $ab = 0$ *следует* $a = 0$ *или* $b = 0$, *т.е. в поле нет делителей нуля.*

11. $\forall a \in P \quad (-a) = (-1)a$ (*противоположный элемент для* a *равен* $(-1)a$).

12. $\forall a, b \in P \quad (-a)b = a(-b) = -(ab) = -ab$.

13. $\forall a, b \in P \quad (-a)(-b) = ab$.

14. *Для любых элементов* $a, b \in P$, $a \neq 0$ *существует, и притом единственное, решение* $x = ba^{-1} = a^{-1}b$ *уравнения* $xa = b\,(ax = b)$; *это решение* x *обозначается* $\dfrac{b}{a}$ *или* b/a *и называется* **частным** *или* **дробью**. *Итак,* $\dfrac{b}{a} = ba^{-1}$ *и эта формула определяет операцию* **деления**.

15. *Справедливы следующие правила обращения с дробями:*

    (а) $\forall a, b, c, d \in P,\ b \neq 0,\ d \neq 0 \quad \dfrac{a}{b} = \dfrac{c}{d} \Leftrightarrow ad = bc$;

    (б) $\forall a, b, c \in P,\ b \neq 0,\ c \neq 0 \quad \dfrac{a}{b} = \dfrac{ac}{bc}$;

    (в) $\forall a, b, c, d \in P,\ b \neq 0,\ d \neq 0 \quad \dfrac{a}{b}\dfrac{c}{d} = \dfrac{ac}{bd}$;

    (г) $\forall a, b, c \in P,\ b \neq 0 \quad \dfrac{a}{b} \pm \dfrac{c}{b} = \dfrac{a \pm c}{b}$;

    (д) $\forall a, b, c, d \in P,\ b \neq 0,\ d \neq 0 \quad \dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{ad \pm bc}{bd}$;

    (е) $\forall a, b \in P,\ b \neq 0 \quad \dfrac{-a}{b} = \dfrac{a}{-b} = -\dfrac{a}{b}$.

Доказательство.

8. Пусть b_1 и b_2 — обратные элементы к элементу a, т.е. $b_1 a = 1$, $ab_2 = 1$. Из ассоциативности умножения следует, что $(b_1 a)b_2 = b_1(ab_2)$. Отсюда $1b_2 = b_1 1$ и $b_1 = b_2$.

10. Если $a = 0$, то утверждение верно. При $a \neq 0$ существует обратный a^{-1}. Умножим обе части равенства $ab = 0$ на a^{-1}. Получим $a^{-1}ab = a^{-1}0 \Rightarrow 1b = 0 \Rightarrow b = 0$. Следовательно, утверждение верно.

14. Доказывается как свойство 4.

15 а). Из свойства 14 имеем $\dfrac{a}{b} = b^{-1}a$, $\dfrac{c}{d} = cd^{-1}$. Тогда $\dfrac{a}{b} = \dfrac{c}{d} \Leftrightarrow b^{-1}a =$

$= cd^{-1} \Leftrightarrow b(b^{-1}a)d = b(cd^{-1})d \Leftrightarrow (bb^{-1})(ad) = (bc)(d^{-1}d) \Leftrightarrow 1(ad) = (bc)1 \Leftrightarrow ad = bc.$

15 б). Из свойства 15 а) $\dfrac{a}{b} = \dfrac{ac}{bc} \Leftrightarrow abc = bac$. Равенство $abc = bac$ верно, поэтому равенство $\dfrac{a}{b} = \dfrac{ac}{bc}$ также верно.

15 в). Из свойства 14 имеем $\dfrac{a}{b} = ab^{-1}$, $\dfrac{c}{d} = cd^{-1}$. Тогда $\dfrac{a}{b}\dfrac{c}{d} = ab^{-1}cd^{-1} = acd^{-1}b^{-1} = ac(bd)^{-1} = \dfrac{ac}{bd}$.

15 г). Из свойства 14 имеем $\dfrac{a}{b} = ab^{-1}$, $\dfrac{c}{b} = cb^{-1}$. Тогда $\dfrac{a}{b} \pm \dfrac{c}{b} = ab^{-1} \pm cb^{-1} = (a \pm c)b^{-1} = \dfrac{a \pm c}{b}$.

15 д). Из свойства 15 б) имеем $\dfrac{a}{b} = \dfrac{ad}{bd}$, $\dfrac{c}{d} = \dfrac{cb}{bd}$. Далее из свойства 15 г) получаем $\dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{ad}{bd} \pm \dfrac{cb}{bd} = \dfrac{ad \pm cb}{bd}$.

15 е). $\dfrac{-a}{b} = \{\text{из } 14\} = (-a)b^{-1} = \{\text{из } 12\} = -(ab^{-1}) = \{\text{из } 14\} = -\dfrac{a}{b}$.

Из 13 $(-a)(-b) = ab \Leftrightarrow \{\text{из 15а}\}\dfrac{-a}{b} = \dfrac{a}{-b}$. ■

Определение 2. *Множество M поля P называется подполем P, если оно само является полем при тех же операциях сложения и умножения, которые заданы в поле P.* ■

Теорема 1.9.2 (Критерий подполя). *Пусть заданы поле P и его подмножество M, содержащее не менее двух элементов.*

M является подполем поля P тогда и только тогда, когда M замкнуто относительно операций сложения, умножения, перехода к противоположному элементу и перехода к обратному элементу, т. е.

(КП1) $\forall a, b \in M \quad a + b \in M$;

(КП2) $\forall a, b \in M \quad ab \in M$;

(КП3) $\forall a \in M \quad (-a) \in M$;

(КП4) $\forall (a \in M,\ a \neq 0) \quad a^{-1} \in M$.

Доказательство.

1. Если M — подполе P, то M — поле и выполнены (КП1) – (КП4).

2. Пусть в M выполнены (КП1) – (КП4). Тогда в M определены операции сложения, умножения, любой элемент $a \in M$ имеет противоположный элемент

$(-a) \in M$, любой элемент $a \in M$, $a \neq 0$ имеет обратный элемент $a^{-1} \in M$. Сложив элементы $a + (-a)$, получим $0 \in M$. Перемножив элементы aa^{-1}, получим $1 \in M$. Тогда аксиомы поля, выполненные в P, будут выполнены в M. Следовательно M — поле. ■

Примеры полей.

1. Множество рациональных чисел Q.
2. Множество действительных чисел R.
3. Множество действительных чисел вида $a + b\sqrt{2}$ с любыми $a, b \in Q$.

Определение 3. *Характеристика поля — это наименьшее натуральное число n такое, что сумма n единиц равна нулю:*
$\sum\limits_{m=1}^{n} 1 = \underbrace{1 + \cdots + 1}_{n} = 0$ $(n \cdot 1 = 0)$. *Если такого числа не существует, то характеристика считается равной нулю.* ■

Теорема 1.9.3. *Характеристикой поля может быть либо* 0, *либо простое число.*

Доказательство. Пусть n — характеристика поля и n — составное число, т.е. $\underbrace{1 + 1 + \cdots + 1}_{n} = 0$ и $n = pq$ $(p, q \in N, p < n, q < n)$. Тогда в силу дистрибутивности $(\underbrace{1 + 1 + \cdots + 1}_{p})(\underbrace{1 + 1 + \cdots + 1}_{q}) = \underbrace{1 + 1 + \cdots + 1}_{pq} =$
$= \underbrace{1 + 1 + \cdots + 1}_{n} = 0$, т. е. $(\underbrace{1 + 1 + \cdots + 1}_{p})(\underbrace{1 + 1 + \cdots + 1}_{q}) = 0$.
Поскольку в поле нет делителей нуля, то отсюда следует, что $\underbrace{1 + 1 + \cdots + 1}_{p} =$
$= 0$ при $p < n$ или $\underbrace{1 + 1 + \cdots + 1}_{q} = 0$ при $q < n$, т. е. n — не наименьшее натуральное число, для которого $\underbrace{1 + 1 + \cdots + 1}_{n} = 0$. ■

Задача 2. *Докажите, что пересечение подполей будет подполем.*

1.10. Кольцо классов вычетов

Множество Z с операциями сложения и умножения образуют коммутативное кольцо. Пусть $m > 1$ — фиксированное натуральное число. Для подмножества mZ (это все целые числа кратные m) выполнены условия критерия

(КК1) – (КК3), следовательно, mZ — подкольцо кольца Z. В силу коммутативности умножения mZ — коммутативное подкольцо.

Используя mZ, построим ненулевое кольцо, состоящее из m элементов.

Определение 1. *Числа $p, q \in Z$ называются сравнимыми по модулю m, если при делении на m они дают одинаковые остатки: $p = fm + r$, $q = hm + r$, где $f, h \in Z$. Обозначение: $p = q \pmod m$.* ■

Z образует коммутативную группу по сложению, mZ — это подгруппа Z. Подгруппа mZ является нормальным делителем, т. к. левые и правые смежные классы совпадают: $a + mZ = mZ + a, \forall a \in Z$. Смежный класс $a + mZ = \{a + mk \mid k \in Z\}$, содержащий число a, называется *классом вычетов по модулю m* и обозначается $[a]_m$. Все числа из класса $[a]_m$ сравнимы по модулю m, т. к. они при делении на m дают такой же остаток как и a. Поэтому всего имеется m различных классов вычетов $Z_m = \{[0]_m, [1]_m, \ldots, [m-1]_m\}$. Они образуют группу с операцией сложения классов

$$[a]_m + [b]_m = [a + b]_m. \tag{1}$$

Введём операцию умножения классов

$$[a]_m[b]_m = [ab]_m. \tag{2}$$

Теорема 1.10.1. *Классы вычетов Z_m с операциями (1) и (2) — это коммутативное кольцо с единицей.* ■

Доказательство. Проверим выполнение аксиом.

АК1 Сложение ассоциативно $([a]_m + [b]_m) + [c]_m = [a]_m + ([b]_m + [c]_m)$, т. к. $([a]_m + [b]_m) + [c]_m = [a + b]_m + [c]_m = [(a + b) + c]_m = [a + (b + c)]_m = [a]_m + [b + c]_m = [a]_m + ([b]_m + [c]_m)$.

АК4 Сложение коммутативно $[a + b]_m = [b + a]_m$, т. к. $a + b = b + a$.

АК2 $[0]_m$ — нейтральный элемент для сложения, т. к. $[a]_m + [0]_m = [a + 0]_m = [a]_m$

АК3 $[-a]_m$ — противоположный элемент для $[a]_m$, т. к. $[a]_m + [-a]_m = [a + (-a)]_m = [0]_m$.

АК8 Умножение коммутативно $[a]_m[b]_m = [b]_m[a]_m$, т. к. $[a]_m[b]_m = [ab]_m = [ba]_m = [b]_m[a]_m$.

АК5 Умножение ассоциативно $([a]_m[b]_m)[c]_m = [a]_m([b]_m[c]_m)$,

т. к. $([a]_m[b]_m)[c]_m = [ab]_m[c]_m = [(ab)c]_m = [a(bc)]_m = [a]_m[bc]_m = [a]_m([b]_m[c]_m)$.

АК6 Умножение дистрибутивно $([a]_m + [b]_m)[c]_m = [a]_m[c]_m + [b]_m[c]_m$,
т. к. $([a]_m + [b]_m)[c]_m = [a+b]_m[c]_m = [(a+b)c]_m = [ac+bc]_m = [ac]_m + [bc]_m =$
$= [a]_m[c]_m + [b]_m[c]_m$.

АК7 $[1]_m$ — нейтральный элемент для умножения,
т. к. $[a]_m[1]_m = [a1]_m = [a]_m$. ■

Приведём **три примера конечных колец,** указывая отдельно таблицы сложения и умножения и обозначая классы вычетов их представителями.

$\mathbf{Z_2}$:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

$\mathbf{Z_3}$:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

$\mathbf{Z_4}$:

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

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

В кольце $\mathbf{Z_4}$ есть делители нуля: $[2]_4[2]_4 = [0]_4$ — «$2 \times 2{=}0$».

Теорема 1.10.2. *Кольцо Z_m является полем тогда и только тогда, когда m — простое число. При этом для простого числа m поле Z_m имеет характеристику m.*

Доказательство. I. Пусть кольцо Z_m является полем. Если $m = 2$, то это простое число. При $m > 2$ предположим, что m — составное число, т. е. $m = pq\,(p,\, q \in N,\, 2 \leqslant p < m,\, 2 \leqslant q < m)$. Тогда $[m]_m = [pq]_m = [p]_m[q]_m$, где $[m]_m = [0]_m$ — нулевой класс. Отсюда $[p]_m[q]_m = [0]_m$ для $[p]_m \neq [0]_m$, $[q]_m \neq$ $\neq [0]_m$. Т. е. в поле Z_m есть делители нуля, что невозможно. Следовательно, m — простое число.

II. Пусть m — простое число.

Докажем, что для произвольного ненулевого класса $[a]_m$ $(1 \leqslant a \leqslant m-1)$ найдётся класс $[b]_m$ такой, что $[b]_m[a]_m = [1]_m$. Для этого рассмотрим m клас-

сов $[0]_m[a]_m = [0a]_m$, $[1]_m[a]_m = [1a]_m$, $\ldots$, $[m-1]_m[a]_m = [(m-1)a]_m$ и сначала покажем, что все они различны.

Действительно, из совпадения классов $[pa]_m = [qa]_m$, где p и q различны, $1 \leqslant a \leqslant m-1$, $0 \leqslant p \leqslant m-1$, $0 \leqslant q \leqslant m-1$, следует, что $pa - qa = (p-q)a$ делится на m и $1 \leqslant a \leqslant m-1$, $-(m-1) \leqslant p-q \leqslant (m-1)$. Тогда поскольку a не делится на m, то $p-q$ делится на m и, следовательно, $p = q$, что противоречит условию.

Итак, все m классов $[0a]_m$, $[1a]_m$, $\ldots$, $[(m-1)a]_m$ различны. Тогда среди них найдётся класс $[ba]_m$ равный $[1]_m$, и, следовательно, $[b]_m[a]_m = [1]_m$. Поэтому при простом m у произвольного ненулевого класса $[a]_m$ есть обратный $[b]_m$, и, следовательно, Z_m — поле. Это поле характеристики m, т. к. $\underbrace{[1]_m + \ldots + [1]_m}_{p} = [p]_m \neq [0]_m$ при $0 < p < m$ и $\underbrace{[1]_m + \ldots + [1]_m}_{m} = [m]_m = [0]_m$. ■

Задача 3.

а) Докажите, что $\mathbb{N}$ и $\mathbb{Z}$ с операциями сложения и умножения не являются полями.

б) Докажите, что $\mathbb{C}$ с операциями сложения и умножения является полем.

Задача 4. *Проверьте является ли кольцом или полем данное множество с указанными операциями.*

а) Матрицы $\mathbb{R}^{m\times n}$ (размера $m \times n$ из вещественных элементов) с операциями сложения и умножения.

б) Матрицы $\mathbb{R}^{n\times n}$ с операциями сложения и умножения.

в) Невырожденные матрицы из $\mathbb{R}^{n\times n}$ (определители матриц не равны нулю) с операциями сложения и умножения.

г) Многочлены $\mathbb{R}[x]$ (с действительными коэффициентами от действительной переменной x) с операциями сложения и умножения.

д) Рациональные функции или дроби вида $\dfrac{f(x)}{g(x)}$, где $f(x), g(x) \in \mathbb{R}[x]$, $g(x) \not\equiv 0$, с операциями сложения $\dfrac{f(x)}{g(x)} + \dfrac{h(x)}{k(x)} = \dfrac{f(x)k(x) + g(x)h(x)}{g(x)k(x)}$ и умножения $\dfrac{f(x)}{g(x)} \cdot \dfrac{h(x)}{k(x)} = \dfrac{f(x)h(x)}{g(x)k(x)}$.

1.11. Комплексные числа. Поле комплексных чисел

Определение 1.

1) *Комплексными числами называются упорядоченные пары* (x, y) *вещественных чисел* $x, y \in R$. *Обозначение:* $z = (x, y)$, $z \in C$, *где* C *— множество всех комплексных чисел.*

2) *Два комплексных числа* $z_1 = (x_1, y_1)$ *и* $z_2 = (x_2, y_2)$ *равны, если* $x_1 = x_2$, $y_1 = y_2$; (*обозначение:* $z_1 = z_2$).

3) *Суммой чисел* $z_1 = (x_1, y_1) \in C$ *и* $z_2 = (x_2, y_2) \in C$ *называется число* $(x_1 + x_2, y_1 + y_2) \in C$ (*обозначение:* $z_1 + z_2$). *Операция* $\langle z_1, z_2 \rangle \mapsto z_1 + z_2$ *называется сложением.*

4) *Произведением чисел* $z_1 = (x_1, y_1) \in C$ *и* $z_2 = (x_2, y_2) \in C$ *называется число* $(x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \in C$ (*обозначение:* $z_1 \cdot z_2$ *или* z_1z_2). *Операция* $\langle z_1, z_2 \rangle \mapsto z_1z_2$ *называется умножением.*

5) *Обозначим через* C_R *подмножество комплексных чисел вида* $(x, 0)$. *Каждое число* $(x, 0) \in C_R$ *взаимно-однозначно соответствуют вещественному числу* $x \in R$. *Поэтому числа из* C_R *называют вещественными числами.* ■

Теорема 1.11.1. *Множество комплексных чисел* C *с операциями сложения и умножения являются полем.*

Доказательство.

Из определения 1 следует, что, если $z_1 = (x_1, y_1) \in C$ и $z_2 = (x_2, y_2) \in C$, то $z_1 + z_2 = (x_1 + x_2, y_1 + y_2) \in C$ и $z_1z_2 = (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \in C$, т. е. операции сложения и умножения корректны.

Далее из равенств $z_2 + z_1 = (x_2 + x_1, y_2 + y_1) = (x_1 + x_2, y_1 + y_2) = z_1 + z_2$ и $z_1z_2 = (x_2x_1 - y_2y_1, x_2y_1 + y_2x_1) = (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) = z_2z_1$ следует коммутативность сложения и умножения АП4 и АП8.

Для $z_k = (x_k, y_k) \in C$, $k = \overline{1, 3}$ проверим ассоциативность сложения АП1: $(z_1 + z_2) + z_3 = (x_1 + x_2, y_1 + y_2) + (x_3, y_3) = ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3) = (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)) = (x_1, y_1) + (x_2 + x_3, y_2 + y_3) = z_1 + (z_2 + z_3)$.

АП2: нулём является число $(0, 0)$, т. к. $\forall\, (x, y) \in C$ $(x, y) + (0, 0) = (x, y)$.

АП3: противоположным числу $(x, y) \in C$ является число $(-x, -y) \in C$, т. к. $(x, y) + (-x, -y) = (0, 0)$.

АП5: ассоциативность умножения $(z_1z_2)z_3 = z_1(z_2z_3)$ проверьте самостоятельно.

АП6: единицей является число $(1,0)$, т. к. $\forall\,(x,y) \in C \quad (1,0)(x,y) = (1x - 0y, 1y + 0x) = (x,y)$.

АП7: найдём для $(a,b) \neq (0,0)$ обратный (x,y) такой, что $(a,b)(x,y) = (1,0) \Leftrightarrow (ax - by, ay + bx) = (1,0) \Leftrightarrow \begin{cases} ax - by = 1, \\ bx + ay = 0. \end{cases}$

Определитель системы $\Delta = \begin{vmatrix} a & -b \\ b & a \end{vmatrix} = a^2 + b^2 \neq 0$. Следовательно, система имеет единственное решение. Находим определители $\Delta_1 = \begin{vmatrix} 1 & -b \\ 0 & a \end{vmatrix} = a$, $\Delta_2 = \begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix} = -b$ и решение $x = \dfrac{\Delta_1}{\Delta} = \dfrac{a}{a^2+b^2}$, $y = \dfrac{\Delta_2}{\Delta} = \dfrac{-b}{a^2+b^2}$.

Итак, $(a,b)^{-1} = \left(\dfrac{a}{a^2+b^2}, \dfrac{-b}{a^2+b^2}\right)$.

АП9: дистрибутивность умножения относительно сложения $(z_1 + z_2)z_3 = z_1z_3 + z_2z_3$ проверьте самостоятельно.

Следствия. 1. Для любой пары комплексных чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ существует единственная разность $z_1 - z_2 = (x_1 - x_2, y_1 - y_2)$, и эта формула определяет операцию вычитания.

2. Для любой пары комплексных чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2) \neq (0,0)$ существует единственное частное $\dfrac{z_1}{z_2} = z_1{z_2}^{-1} = \left(\dfrac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2}, \dfrac{-x_1y_2 + y_1x_2}{x_2^2 + y_2^2}\right)$, и эта формула определяет операцию деления.

3. Подмножество комплексных чисел $C_R = \{z \mid z = (x,0),\, x \in R\}$ является подполем поля C с операциями сложения и умножения комплексных чисел и это подполе изоморфно полю вещественных чисел R с операциями сложения и умножения вещественных чисел, причём изоморфизмом является отображение $f((x,0)) = x$, $x \in R$.

Доказательство. Следствие 1 получено из 4-го свойства поля. Для следствия 2 формула $\dfrac{z_1}{z_2} = z_1{z_2}^{-1}$ получена из 14-го свойства поля. Далее значение $z_1{z_2}^{-1} = \left(\dfrac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2}, \dfrac{-x_1y_2 + y_1x_2}{x_2^2 + y_2^2}\right)$ получено при умножении $z_1 = (x_1, y_1)$

на обратное число $z_2^{-1} = \left(\dfrac{x_2}{x_2^2+y_2^2}, \dfrac{-y_2}{x_2^2+y_2^2}\right)$.

3. Действительно, для C_R выполнены критерии подполя.

(КП1) $\forall\, z_1 = (x_1, 0),\ z_2 = (x_2, 0) \in C_R \quad z_1 + z_2 \in C_R$, т. к. $z_1 + z_2 = (x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \in C_R$;

(КП2) $\forall\, z_1 = (x_1, 0),\ z_2 = (x_2, 0) \in C_R \quad z_1 z_2 \in C_R$; т. к. $z_1 z_2 = (x_1, 0)(x_2, 0) = (x_1 x_2 - 0 \cdot 0, 0 \cdot 0 + 0 \cdot 0) = (x_1 x_2, 0) \in C_R$;

(КП3) $\forall\, z = (x, 0) \in C_R \qquad -z = (-x, 0) \in C_R$;

(КП4) $\forall\, z = (x, 0) \in C_R,\ x \neq 0 \qquad z^{-1} = \left(\dfrac{1}{x}, 0\right) \in C_R$,

т. к. обратный для $(a, b) \neq (0, 0)$ — это $(a, b)^{-1} = \left(\dfrac{a}{a^2+b^2}, \dfrac{-b}{a^2+b^2}\right)$, и, следовательно, $(x, 0)^{-1} = \left(\dfrac{x}{x^2+0^2}, \dfrac{-0}{x^2+0^2}\right) = \left(\dfrac{1}{x}, 0\right) \in C_R$.

Отображение $f\colon C_R \to R$, определённое формулой $f((x, 0)) = x,\ x \in R$ является изоморфизмом C_R и R.

I. f биективно.

а) Для любых $(x_1, 0), (x_2, 0) \in C_R$ из $(x_1, 0) \neq (x_2, 0)$ следует $x_1 \neq x_2$, т. е. $f((x_1, 0)) \neq f((x_2, 0))$ и f — инъективно.

б) При любом $x \in R$ для $(x, 0)$ выполняется равенство $f((x, 0)) = x$, т. е. f — сюръективно.

II. Для любых $z_1 = (x_1, 0) \in C_R,\ z_2 = (x_2, 0) \in C_R$ выполнены равенства $f(z_1) = x_1,\ f(z_2) = x_2$.

в) Имеем $z_1 + z_2 = (x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \in C_R$. Тогда $f(z_1 + z_2) = x_1 + x_2$ и выполнено равенство $f(z_1 + z_2) = f(z_1) + f(z_2)$.

г) Имеем $z_1 z_2 = (x_1, 0)(x_2, 0) = (x_1 x_2 - 0 \cdot 0, x_1 \cdot 0 + 0 \cdot x_2) = (x_1 x_2, 0) \in C_R$. Тогда $f(z_1 z_2) = x_1 x_2$ и выполнено равенство $f(z_1 z_2) = f(z_1) f(z_2)$. ■

1.12. Алгебраическая форма комплексных чисел. Арифметические операции для комплексных чисел в алгебраической форме. Геометрическое изображение комплексных чисел и операций сложения и вычитания

Определение 1. *Число $(x, 0)$ $(x \in R)$ называют вещественное число и обозначают просто x.*

Число $(0,1)$ *называют мнимая единица и обозначают* i. ■

Из определения комплексных чисел и операций сложения и умножения получаем следующие факты.

1° $i^2 = -1$.

2° Любое комплексное число $z = (x, y)$ может быть записано в виде

$$z = x + yi.$$

Доказательство.

1° $i^2 = ii = (0,1)\,(0,1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1$.

2° Имеем $z = (x, y) = (x, 0) + (0, y)$. Далее получим $(y, 0)\,(0, 1) = (y \cdot 0 - 0 \cdot 1, y \cdot 1 + 0 \cdot 0) = (0, y)$. Отсюда $(x, 0) + (0, y) = (x, 0) + (y, 0)\,(0, 1) = x + yi$.

Определение 2. *Форма записи комплексного числа* $z = (x, y)$ *в виде* $z = x + yi$ *называется алгебраической формой комплексного числа. При этом* x *и* y *называются соответственно вещественной и мнимой частью числа* z *и обозначаются* $x = \mathrm{Re}(z)$, $y = \mathrm{Im}(z)$. *Отметим, что* $\mathrm{Re}(z) \in R$, $\mathrm{Im}(z) \in R$. *Числа вида* $z = x + 0i$ *называют вещественными и записывают в виде* $z = x$. *Числа вида* $z = 0 + yi$ *называют чисто мнимыми и записывают в виде* $z = yi$.

Два комплексных числа $z_1 = x_1 + y_1 i$ *и* $z_2 = x_2 + y_2 i$ *равны тогда и только тогда, когда равны их действительные и мнимые части:* $x_1 = x_2$ *и* $y_1 = y_2$.

Комплексное число $\overline{z} = x - yi$ *называется сопряжённым к числу* $z = x + yi$.

Преобразование $z \mapsto \overline{z}$ *называется операцией сопряжения.* ■

Арифметические операции для комплексных чисел в алгебраической форме.

Пусть заданы два комплексных числа $z_1 = x_1 + y_1 i$ и $z_2 = x_2 + y_2 i$. При выполнении алгебраических операций используются свойства ассоциативности, коммутативности и дистрибутивности операций.

Операции сложения и вычитания можно производить почленно отдельно для вещественных и мнимых частей:

$z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i$,

$z_1 - z_2 = (x_1 - x_2) + (y_1 - y_2)i.$

Операцию умножения можно произвести так:

$z_1 z_2 = (x_1 + y_1 i)(x_2 + y_2 i) = x_1 x_2 + x_1 y_2 i + y_1 x_2 i + y_1 y_2 i^2$, учитывая равенство $i^2 = -1$, далее получим $z_1 z_2 = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i.$

Для деления $\dfrac{z_1}{z_2}$ удобно числитель и знаменатель дроби умножить на число $\overline{z_2} = x_2 - y_2 i$, сопряжённое к знаменателю:

$$\frac{z_1}{z_2} = \frac{z_1\overline{z_2}}{z_2\overline{z_2}} = \frac{(x_1 + y_1 i)(x_2 - y_2 i)}{(x_2 + y_2 i)(x_2 - y_2 i)} = \frac{(x_1 x_2 + y_1 y_2) + (y_1 x_2 - x_1 y_2)i}{x_2^2 + y_2^2} =$$

$$= \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + \frac{y_1 x_2 - x_1 y_2}{x_2^2 + y_2^2} i. \blacksquare$$

Примеры.

1) $(2 + 4i) + (-3 - 5i) = (2 - 3) + (4 - 5)i = -1 - i;$

2) $(-7 + 3i) - (-3 + 5i) = (-7 + 3) + (3 - 5)i = -4 - 2i;$

3) $(-2 - 3i)(3 + 2i) = -6 - 4i - 9i - 6i^2 = -6 - 13i + 6 = -13i.$

4) $(5 - 4i)(-3 - 2i) = -15 - 8 + (-10 + 12)i = -23 + 2i.$

5) $\dfrac{3 - 7i}{-2 + 3i} = \dfrac{(3 - 7i)(-2 - 3i)}{(-2 + 3i)(-2 - 3i)} = \dfrac{-6 - 21 + (-9 + 14)i}{4 + 9} = \dfrac{-27 + 5i}{13} =$
$= \dfrac{-27}{13} + \dfrac{5}{13} i;$

6) $\dfrac{5 - 10i}{4 - 3i} = \dfrac{5(1 - 2i)}{4 - 3i} = \dfrac{5(1 - 2i)(4 + 3i)}{(4 - 3i)(4 + 3i)} = \dfrac{5(4 + 6 + (3 - 8)i)}{16 + 9} =$
$= \dfrac{5(10 - 5i)}{25} = \dfrac{25(2 - i)}{25} = 2 - i.$

Геометрическое изображение комплексных чисел и операций сложения и вычитания.

Комплексное число $z = x + yi$ изображается на плоскости точкой $M(x, y)$ или вектором $\overline{OM}\{x, y\}$ с декартовыми прямоугольными координатами x, y. (рис. 7 для случая $x > 0$ и $y > 0$). Тогда сопряжённое число $\overline{z} = x - yi$ изображается точкой или вектором с декартовыми прямоугольными координатами $(x, -y)$, противоположное число $-z = -x - yi$ изображается точкой или вектором с декартовыми прямоугольными координатами $(-x, -y)$.

Вещественные числа изображаются точками оси абсцисс, чисто мнимые числа изображаются точками оси ординат. Число ноль изображается началом координат. Сопряжённое число для данного числа изображается точкой, симметричной данной точке относительно оси абсцисс.

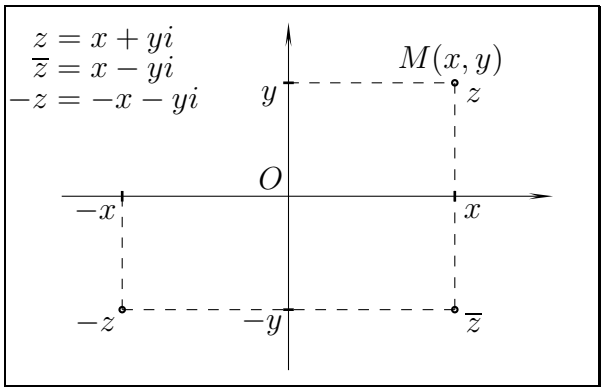

Рис. 7

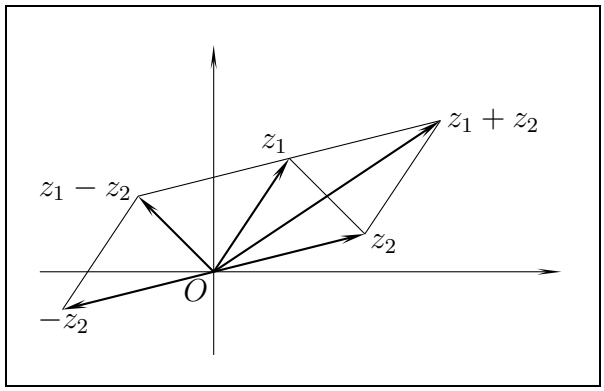

Рис. 8

Плоскость, точками которой изображаются комплексные числа называется *комплексной плоскостью*, её ось абсцисс называется *вещественной осью*, ось ординат — *мнимой осью*.

Сумму $z_1 + z_2$ и разность $z_1 - z_2 = z_1 + (-z_2)$ удобно выражать радиус-векторами, получаемыми по правилу параллелограмма или по правилу треугольника (рис. 8).

1.13. Тригонометрическая форма комплексного числа. Операции над комплексными числами в тригонометрической форме

Определение 1.

Комплексное число $z = x + yi$ изображается на плоскости точкой $M(x, y)$ или вектором $\overline{OM}\{x, y\}$ с декартовыми прямоугольными координатами x, y.

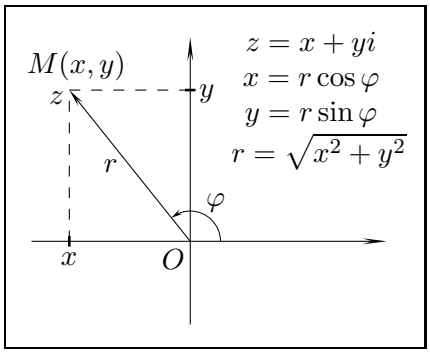

Рис. 9

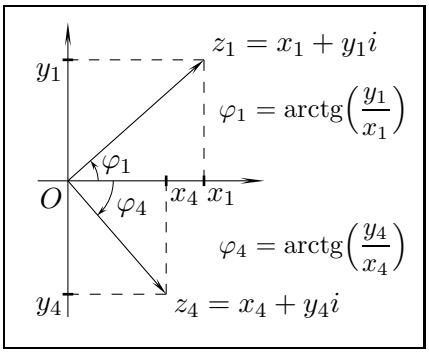

Рис. 10

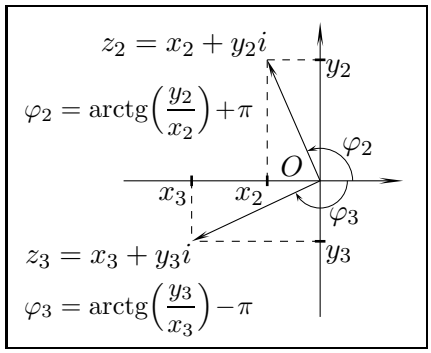

Рис. 11

Положение точки M на плоскости вполне определяется также её полярными координатами: длиной r радиус-вектора $\overline{OM}$ и углом φ между положи-

тельным направлением оси абсцисс и вектором $\overline{OM}$ (рис. 9). Неотрицательное вещественное число $r = \sqrt{x^2 + y^2}$ называется *модулем* числа $z = x + yi$ и обозначается $|z|$. Модуль равен нулю только для числа $z = 0$. Угол φ называется *аргументом* числа $z = x + yi$ и обозначается $\arg(z)$. Аргумент любого числа $z \neq 0$ определён с точностью до числа кратного 2π : $\arg(z) = \varphi_0 + 2\pi k$, $k \in Z$. Аргумент нуля не определён, но удобно считать, что аргументом нуля является произвольное число. Аргумент φ может принимать любые вещественные значения: положительные, отрицательные и ноль, причём положительные углы отсчитываются против часовой стрелки. Если два ненулевых числа имеют равные модули и их аргументы равны или отличаются на число $2\pi k$, $k \in Z$, то эти числа определяют совпадающие точки комплексной плоскости, т. е. эти числа равны. И обратно, если два ненулевых числа равны, то равны их модули и их аргументы равны или отличаются на число $2\pi k$, $k \in Z$.

Между декартовыми прямоугольными координатами и полярными координатами точки существует следующая связь: $x = r\cos\varphi$, $y = r\sin\varphi$.

Применив эти формулы к произвольному комплексному числу $z = x + yi$, получим запись этого числа в виде

$$z = r(\cos\varphi + i\sin\varphi), \quad \text{где } r = |z|,\ \varphi = arg(z) \tag{3}$$

Эту запись называют *тригонометрической формой* комплексного числа.

Для нахождения r, φ по заданным x, y возведём x и y в квадрат, получим $x^2 + y^2 = r^2$ и, следовательно, $r = \sqrt{x^2 + y^2}$. Если $x = 0$ и $y = 0$, то $r = 0$ и угол φ не определён. Если $x \neq 0$ или $y \neq 0$, то $r > 0$ и тогда угол φ определяется из равенств $\cos\varphi = \dfrac{x}{r} = \dfrac{x}{\sqrt{x^2 + y^2}}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{y}{\sqrt{x^2 + y^2}}$. Для получения φ в промежутке $(-\pi, \pi]$ поступим следующим образом. Возьмём $\varphi = \pi/2$ при $x = 0$, $y > 0$, $\varphi = -\pi/2$ при $x = 0$, $y < 0$. При $x \neq 0$ разделим y на x и получим $\operatorname{tg}\varphi = y/x$. Отсюда $\varphi = \operatorname{arctg}(y/x) + \pi k$, $k \in Z$. Далее берём $\varphi = \operatorname{arctg}(y/x)$ при $x > 0$, т. е. для точек (x, y) из 1-й и 4-й четверти, $\varphi = \operatorname{arctg}(y/x) + \pi$ при $x < 0$, $y \geqslant 0$, т. е. для точек (x, y) из 2-й четверти, $\varphi = \operatorname{arctg}(y/x) - \pi$ при $x < 0$, $y < 0$, т. е. для точек (x, y) из 3-й четверти. См. рис. 10–11. ■

Теорема 1.13.1 (Операции в тригонометрической форме).

1° *Пусть* $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$, $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$, *тогда*

$$z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)), \tag{4}$$

т.е. при умножении чисел их модули перемножаются, а аргументы складываются: $|z_1 z_2| = |z_1||z_2|$, $\arg(z_1 z_2) = \arg(z_1) + arg(z_2)$.

2° *Пусть* $z = r(\cos\varphi + i\sin\varphi) \neq 0$, *тогда*

$$z^{-1} = \frac{1}{z} = \frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)). \tag{5}$$

3° *Пусть* $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$, $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2) \neq 0$, *тогда*

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)), \tag{6}$$

т.е. при делении чисел их модули делятся, а аргументы вычитаются: $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$, $\arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - arg(z_2)$.

4° $\forall n \in N$, $z_k = r_k(\cos\varphi_k + i\sin\varphi_k) \in C$, $k = \overline{1,n}$ *справедлива формула*

$$z_1 z_2 \ldots z_n = (r_1 r_2 \ldots r_n)(\cos(\varphi_1 + \varphi_2 + \ldots + \varphi_n) + i\sin(\varphi_1 + \varphi_2 + \ldots + \varphi_n)), \tag{7}$$

т.е. при умножении чисел их модули перемножаются, а аргументы складываются.

5° $\forall n \in N$, $z = r(\cos\varphi + i\sin\varphi) \in C$ *справедлива формула*

$$z^n = r^n(\cos(n\varphi) + i\sin(n\varphi)). \tag{8}$$

6° *Пусть* $z = r(\cos\varphi + i\sin\varphi)$, *тогда*

$$(-z) = r(\cos(\varphi + \pi) + i\sin(\varphi + \pi)), \tag{9}$$

$$\overline{z} = r(\cos(-\varphi) + i\sin(-\varphi)). \tag{10}$$

Доказательство. 1° $z_1 z_2 = r_1(\cos\varphi_1 + i\sin\varphi_1) r_2(\cos\varphi_2 + i\sin\varphi_2) = r_1 r_2(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) = r_1 r_2(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)) = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2))$.

2° Используя (4), умножим $r(\cos\varphi + i\sin\varphi)\frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)) = \frac{r}{r}(\cos(\varphi - \varphi) + i\sin(\varphi - \varphi)) = \cos 0 + i\sin 0 = 1$. Следовательно, обратным для $r(\cos\varphi + i\sin\varphi)$ является $\frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)$.

3° В поле C выполнено равенство $\dfrac{z_1}{z_2} = z_1 z_2^{-1}$. Используя (5) и (4), получим $z_1 z_2^{-1} = r_1(\cos\varphi_1 + i\sin\varphi_1)\dfrac{1}{r_2}(\cos(-\varphi_2) + i\sin(-\varphi_2)) = \dfrac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$.

4° Докажем методом математической индукции.

При $n = 1$ формула $r_1(\cos\varphi_1 + i\sin\varphi_1) = r_1(\cos\varphi_1 + i\sin\varphi_1)$ верна.

Пусть при $n = m$ формула (7) верна.

Докажем формулу (7) при $n = m + 1$:
$z_1 \ldots z_m z_{m+1} = [z_1 \ldots z_m] z_{m+1} = [(r_1 \ldots r_m)(\cos(\varphi_1 + \ldots + \varphi_m) + i\sin(\varphi_1 + \ldots + \varphi_m)](r_{m+1}(\cos\varphi_{m+1} + i\sin\varphi_{m+1})) = (r_1 \ldots r_m r_{m+1})(\cos(\varphi_1 + \ldots + \varphi_m + \varphi_{m+1}) + i\sin(\varphi_1 + \ldots + \varphi_m + \varphi_{m+1}))$. Следовательно, формула (7) верна.

5° Возьмём в формуле (7) равные числа $z_1 = z_2 = \ldots = z_n = z = r(\cos\varphi + i\sin\varphi)$ и получим $z^n = r^n(\cos(n\varphi) + i\sin(n\varphi))$, т.е. формула (8) верна.

6° 1) $r(\cos(\varphi + \pi) + i\sin(\varphi + \pi)) = r(-\cos\varphi - i\sin\varphi) = -z$;

2) $r(\cos(-\varphi) + i\sin(-\varphi)) = r(\cos\varphi - i\sin\varphi) = \overline{z}$. ■

Из теоремы 1.13.1 получим **геометрический смысл умножения и деления** двух комплексных чисел.

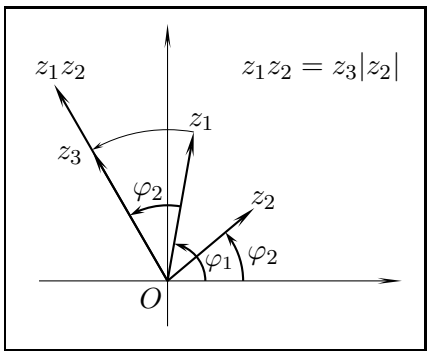

Рис. 12

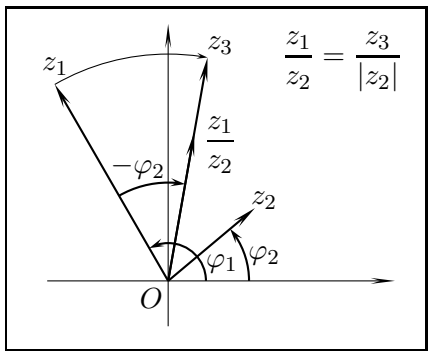

Рис. 13

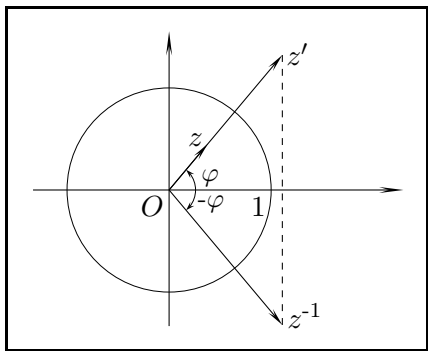

Рис. 14

Пусть заданы числа $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ и $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$, которые изображаются векторами на рис. 12–13.

1) Построение произведения $z_1 z_2$ проведено на рис.12. Вектор z_1 поворачиваем на угол φ_2 (с учётом знака φ_2) и получаем вектор $z_3 = r_1(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2))$. Затем вектор z_3 умножаем на $r_2 = |z_2|$ и получаем вектор, изображающий $r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)) = z_1 z_2$.

2) Построение частного $\dfrac{z_1}{z_2}$ проведено на рис.13. Вектор z_1 поворачиваем на угол $(-\varphi_2)$ и получаем вектор $z_3 = r_1(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$. Затем вектор z_3 делим на $r_2 = |z_2|$ и получаем вектор, изображающий $\dfrac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)) = \dfrac{z_1}{z_2}$.

3) На рис.14 для числа $z = r(\cos\varphi + i\sin\varphi) \neq 0$ проведено построение обратного числа $z^{-1} = \dfrac{1}{r}(\cos(-\varphi) + i\sin(-\varphi))$. От вектора z, переходим к вектору $z' = \dfrac{1}{r}(\cos\varphi + i\sin\varphi)$. Затем от z' преобразованием симметрии относительно вещественной оси переходим к вектору $\dfrac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)) = z^{-1}$. ■

Задача 5. *Запишите в тригонометрической форме следующие числа:* 1) i, 2) $-3i$, 3) -4, 4) 2, 5) $-4-4i$, 6) $-\sqrt{3}+i$, 7) $-3+2i$.

Решение. Для того, чтобы число $z = x + yi \neq 0$, где $x = \mathrm{Re}(z)$, $y = \mathrm{Im}(z)$ записать в тригонометрической форме $z = r(\cos\varphi + i\sin\varphi)$, где $r = |z|$, $\varphi = \arg(z)$, сначала вычисляем модуль $r = |z| = \sqrt{x^2+y^2}$, затем из равенств $\cos\varphi = \dfrac{x}{r}$, $\sin\varphi = \dfrac{y}{r}$ находим аргумент $\varphi = \arg(z)$.

1) Для $z = i = 0 + 1i$ имеем $x = \mathrm{Re}(z) = 0$, $y = \mathrm{Im}(z) = 1$, $r = |z| = \sqrt{x^2+y^2} = \sqrt{0^2+1^2} = 1$, $\cos\varphi = \dfrac{x}{r} = 0$, $\sin\varphi = \dfrac{y}{r} = 1$. Отсюда находим один из возможных аргументов $\varphi = \dfrac{\pi}{2}$ или все возможные аргументы $\varphi = \dfrac{\pi}{2} + 2\pi k$, где $k \in Z$.

Ответ: $i = \cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}$ или $i = \cos\left(\dfrac{\pi}{2} + 2\pi k\right) + i\sin\left(\dfrac{\pi}{2} + 2\pi k\right)$, $k \in Z$.

2) Для $z = -3i = 0 + (-3)i$ имеем $x = 0$, $y = -3$, $r = |z| = \sqrt{0^2 + (-3)^2} = 3$, $\cos\varphi = \dfrac{x}{r} = 0$, $\sin\varphi = \dfrac{y}{r} = -1$. Отсюда находим $\varphi = -\dfrac{\pi}{2}$.

Ответ: $-3i = 3\left(\cos\left(-\dfrac{\pi}{2}\right) + i\sin\left(-\dfrac{\pi}{2}\right)\right)$.

3) Для $z = -4 = -4 + 0i$ имеем $x = -4$, $y = 0$, $r = |z| = \sqrt{(-4)^2 + 0^2} = 4$, $\cos\varphi = \dfrac{x}{r} = -1$, $\sin\varphi = \dfrac{y}{r} = 0$. Отсюда находим $\varphi = \pi$.

Ответ: $-4 = 4(\cos\pi + i\sin\pi)$.

4) Для $z = 2 = 2 + 0i$ имеем $x = 2$, $y = 0$, $r = |z| = \sqrt{2^2 + 0^2} = 2$, $\cos\varphi = \dfrac{x}{r} = 1$, $\sin\varphi = \dfrac{y}{r} = 0$. Отсюда находим $\varphi = 0$.

Ответ: $2 = 2\left(\cos 0 + i \sin 0\right)$.

5) Для $z = -4 - 4i$ имеем $x = -4$, $y = -4$, $r = |z| = \sqrt{(-4)^2 + (-4)^2} = 4\sqrt{2}$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{\sqrt{2}}{2}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{\sqrt{2}}{2}$. Отсюда находим $\varphi = -\dfrac{3\pi}{4}$.

Ответ: $-4 - 4i = 4\sqrt{2}\left(\cos\left(-\dfrac{3\pi}{4}\right) + i\sin\left(-\dfrac{3\pi}{4}\right)\right)$.

6) Для $z = -\sqrt{3} + i$ имеем $x = -\sqrt{3}$, $y = 1$, $r = |z| = \sqrt{(-\sqrt{3})^2 + 1^2} = 2$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{\sqrt{3}}{2}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{1}{2}$. Отсюда находим $\varphi = \dfrac{5\pi}{6}$.

Ответ: $-\sqrt{3} + i = 2\left(\cos\left(\dfrac{5\pi}{6}\right) + i\sin\left(\dfrac{5\pi}{6}\right)\right)$.

7) Для $z = -3 + 2i$ имеем $x = -3$, $y = 2$, $r = |z| = \sqrt{(-3)^2 + 2)^2} = 13$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{3}{\sqrt{13}}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{2}{\sqrt{13}}$. Отсюда находим $\operatorname{tg}\varphi = \dfrac{\sin\varphi}{\cos\varphi} = -\dfrac{2}{3}$ и во второй четверти $\varphi = \operatorname{arctg}\left(-\dfrac{2}{3}\right) + \pi = \pi - \operatorname{arctg}\left(\dfrac{2}{3}\right)$.

Ответ: $-3 + 2i = \sqrt{13}\left(\cos\left(\pi - \operatorname{arctg}\left(\dfrac{2}{3}\right)\right) + i\sin\left(\pi - \operatorname{arctg}\left(\dfrac{2}{3}\right)\right)\right)$.

1.14. Свойства модуля комплексных чисел. Свойства операции сопряжения

Определение 1.

Пусть задано комплексное число $z = x + yi$. Модулем z называется вещественное число $\sqrt{x^2 + y^2}$. Оно обозначается $|z|$.

Комплексное число $\overline{z} = x - yi$ называется сопряжённым к числу z.

Преобразование $z \mapsto \overline{z}$ называется операцией сопряжения.

Замечание. Модуль числа $z = x + 0i$, $x \in R$ равен $\sqrt{x^2}$ и он совпадает с модулем вещественного числа x.

Теорема 1.14.1 (Свойства модуля).

Справедливы следующие утверждения

$$1)\ \forall z \in C \quad |z| = |-z| = |\overline{z}|, \tag{11}$$

$$2)\ \forall z_1,\ z_2 \in C \quad |z_1 z_2| = |z_1||z_2|, \tag{12}$$

$$3)\ \forall z_1,\ z_2 \in C,\ z_2 \neq 0 \quad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \tag{13}$$

$$4)\ \forall z_1,\ z_2 \in C \quad ||z_1| - |z_2|| \leqslant |z_1 \pm z_2| \leqslant |z_1| + |z_2|. \tag{14}$$

Доказательство.

1) Если $z = x + yi$, то значения

$|z| = |x + yi| = \sqrt{x^2 + y^2}$,

$|-z| = |-x - yi| = \sqrt{(-x)^2 + (-y)^2} = \sqrt{x^2 + y^2}$,

$|\overline{z}| = |x - yi| = \sqrt{x^2 + (-y)^2} = \sqrt{x^2 + y^2}$

совпадают, т. е. выполнено (11).

2), 3) Из теоремы 1.13.1 следуют (12) и (13).

4) Пусть числа заданы в тригонометрической форме:

$$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1),\ \text{где}\ r_1 = |z_1|,\ \varphi_1 = \arg(z_1),$$
$$z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2),\ \text{где}\ r_2 = |z_2|,\ \varphi_2 = \arg(z_2).$$

Найдём $z_1 \pm z_2$:

$$z_1 \pm z_2 = r_1(\cos\varphi_1 + i\sin\varphi_1) \pm r_2(\cos\varphi_2 + i\sin\varphi_2) =$$
$$= (r_1\cos\varphi_1 \pm r_2\cos\varphi_2) + (r_1\sin\varphi_1 \pm r_2\sin\varphi_2)i.$$

Далее выразим $|z_1 \pm z_2|^2$:

$$|z_1 \pm z_2|^2 = (r_1\cos\varphi_1 \pm r_2\cos\varphi_2)^2 + (r_1\sin\varphi_1 \pm r_2\sin\varphi_2)^2 =$$
$$= r_1^2\cos^2\varphi_1 \pm 2r_1r_2\cos\varphi_1\cos\varphi_2 + r_2^2\cos^2\varphi_2 +$$
$$+ r_1^2\sin^2\varphi_1 \pm 2r_1r_2\sin\varphi_1\sin\varphi_2 + r_2^2\sin\varphi_2^2 =$$
$$= r_1^2(\cos^2\varphi_1 + \sin^2\varphi_1) \pm 2r_1r_2(\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2) +$$
$$+ r_2^2(\cos^2\varphi_2 + \sin\varphi_2^2) = r_1^2 \pm 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2.$$

Отсюда в силу неравенства $-1 \leqslant \pm\cos(\varphi_1 - \varphi_2) \leqslant 1$ получим

$$r_1^2 - 2r_1r_2 + r_2^2 \leqslant r_1^2 \pm 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2 \leqslant r_1^2 + 2r_1r_2 + r_2^2,\ \text{т. е.}$$
$$(r_1 - r_2)^2 \leqslant r_1^2 \pm 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2 \leqslant (r_1 + r_2)^2\ \text{или}$$
$$|r_1 - r_2|^2 \leqslant |z_1 \pm z_2|^2 \leqslant |r_1 + r_2|^2.$$

Отсюда следует неравенство $|r_1 - r_2| \leqslant |z_1 \pm z_2| \leqslant |r_1 + r_2|$, совпадающее с (14).

Теорема 1.14.2 (Свойства операции сопряжения).

Справедливы следующие утверждения.

$$1)\ \forall z = x + yi \in C \quad \begin{cases} z + \overline{z} = 2x = 2\mathrm{Re}(z), \\ z\overline{z} = x^2 + y^2 = |z|^2; \end{cases} \tag{15}$$

$$2)\ \forall z \in C \quad \overline{\overline{z}} = z, \tag{16}$$

$$3)\ \forall z \in C \quad z = \overline{z} \Leftrightarrow z \in R \Leftrightarrow \mathrm{I}m(z) = 0, \tag{17}$$

$$4)\ \forall z_1,\ z_2 \in C \quad \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \tag{18}$$

$$5)\ \forall z_1,\ z_2 \in C \quad \overline{z_1 - z_2} = \overline{z_1} - \overline{z_2}, \tag{19}$$

$$6)\ \forall z_1,\ z_2 \in C \quad \overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}, \tag{20}$$

$$7)\ \forall z_1,\ z_2 \neq 0 \in C \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}. \tag{21}$$

Доказательство. Пусть $z = x + yi$, тогда $\overline{z} = x - yi$.

1) $z + \overline{z} = 2x = 2\mathrm{R}e(z)$;

$z\overline{z} = (x + yi)(x - yi) = x^2 + y^2 = \left(\sqrt{x^2 + y^2}\right)^2 = |z|^2$.

2) $\overline{\overline{z}} = \overline{x - yi} = x + yi = z$.

3) $z = \overline{z} \Leftrightarrow x + yi = x - yi \Leftrightarrow 2yi = 0 \Leftrightarrow y = 0 \Leftrightarrow \mathrm{I}m(z) = 0$.

Пусть $z_1 = x_1 + y_1 i$, $z_2 = x_2 + y_2 i$, тогда $\overline{z_1} = x_1 - y_1 i$, $\overline{z_2} = x_2 - y_2 i$,

4) $z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i$,

$\overline{z_1} + \overline{z_2} = x_1 - y_1 i + x_2 - y_2 i = (x_1 + x_2) - (y_1 + y_2)i = \overline{z_1 + z_2}$.

5) $z_1 - z_2 = (x_1 - x_2) + (y_1 - y_2)i$,

$\overline{z_1} - \overline{z_2} = x_1 - y_1 i - (x_2 - y_2 i) = (x_1 - x_2) - (y_1 - y_2)i = \overline{z_1 - z_2}$.

6) $z_1 z_2 = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i$,

$\overline{z_1} \cdot \overline{z_2} = (x_1 x_2 - y_1 y_2) - (x_1 y_2 + y_1 x_2)i = \overline{z_1 z_2}$.

7) Обозначим $a = \dfrac{1}{|z_2|^2}$, $a \in R$.

$\dfrac{z_1}{z_2} = \dfrac{z_1 \overline{z_2}}{z_2 \overline{z_2}} = \dfrac{z_1 \overline{z_2}}{|z_2|^2} = z_1 \overline{z_2} a$, $\overline{\left(\dfrac{z_1}{z_2}\right)} = \overline{z_1 \overline{z_2} a} = \overline{z_1}\,\overline{\overline{z_2}}\,\overline{a} = \overline{z_1} z_2 a$,

$\dfrac{\overline{z_1}}{\overline{z_2}} = \dfrac{\overline{z_1}\, z_2}{\overline{z_2} z_2} = \dfrac{\overline{z_1} z_2}{|z_2|^2} = \overline{z_1} z_2 a$. ■

1.15. Возведение в степень. Извлечение корня

Определение 1. *Пусть z — комплексное число и n — целое число. Определим степени z^n следующим образом.*

$$z^n = \begin{cases} \underbrace{zz\ldots z}_{n} & \text{при } n \geqslant 1, \\ 1 \quad \text{при } n = 0,\ z \neq 0, \\ \left(\dfrac{1}{z}\right)^{-n} & \text{при } n \leqslant -1,\ z \neq 0. \end{cases} \quad ■$$

Из определения следуют привычные свойства степеней

$$\begin{aligned}
&\forall n \in Z,\ z_1, z_2 \in C,\ z_1 \neq 0,\ z_2 \neq 0 \quad (z_1 z_2)^n = z_1^n z_2^n, \\
&\forall k \in N,\ n \in Z,\ z_1, z_2, \ldots, z_k \in C,\ z_1 \neq 0,\ z_2 \neq 0, \ldots, z_k \neq 0 \\
&\qquad (z_1 z_2 \ldots z_k)^n = z_1^n z_2^n \ldots z_k^n, \\
&\forall m, n \in Z,\ z \in C,\ z \neq 0 \quad z^m z^n = z^{m+n}, \\
&\forall m, n \in Z,\ z \in C,\ z \neq 0 \quad (z^m)^n = z^{mn}. \blacksquare
\end{aligned} \tag{22}$$

I. Степень числа, заданного в алгебраической форме.

Найдём i^n для $n \geqslant 0:\ i^0 = 1,\ i^1 = i,\ i^2 = i^1 i = -1,\ i^3 = i^2 i = -1i = -i,$ $i^4 = i^3 i = -ii = -i^2 = 1,\ i^5 = i^4 i = 1i = i,\ \ldots$. Понятно, что значения $1,\ i,\ -1,\ -i$ повторяются:

$$i^{4k} = 1,\ i^{4k+1} = i,\ i^{4k+2} = -1,\ i^{4k+3} = -i. \tag{23}$$

Для возведения числа $z = x + yi$ в натуральную степень для небольшого показателя можно применить к выражению $(x + yi)^n$ формулу бинома Ньютона (она справедлива, т. к. её доказательство основано на коммутативности и ассоциативности сложения и умножения, а также на дистрибутивности умножения), а затем воспользоваться формулами (23). Предварительно можно представить степень с использованием формул (22).

Для возведения числа $a+bi$ в отрицательную степень n найдём обратное число $\dfrac{1}{a+bi} = \dfrac{a-bi}{(a+bi)(a-bi)} = \dfrac{a-bi}{a^2+b^2} = \dfrac{a}{a^2+b^2} + \dfrac{-b}{a^2+b^2} i = x + yi$ и возведём полученное число в положительную степень $(x+yi)^{-n}$ так, как было указано выше. $\blacksquare$

II. Степень числа, заданного в тригонометрической форме.

Теорема 1.15.1 (Формула Муавра).

Для произвольного ненулевого комплексного числа $z = r(\cos\varphi + i\sin\varphi)$, $(r = |z|,\ \varphi = \arg(z))$ *при всех* $n \in Z$ *справедлива формула*

$$(r(\cos\varphi + i\sin\varphi))^n = r^n(\cos(n\varphi) + i\sin(n\varphi)), \tag{24}$$

то есть при возведении числа в степень n *его модуль возводится в степень* n, *а его аргумент умножается на* n.

Доказательство. При $n=0$ имеем $(r(\cos\varphi+i\sin\varphi))^0=1$ и $r^0(\cos(0\varphi)+$ $+\,i\sin(0\varphi))=1$, т.е. формула верна. Для $n\in N$ утверждение доказано в теореме 1.13.1 об умножении комплексных чисел в тригонометрической форме.

Далее при $n<0$ имеем $(r(\cos\varphi+i\sin\varphi))^n=\dfrac{1}{(r(\cos\varphi+i\sin\varphi))^{-n}}=$ $=\dfrac{1}{r^{-n}(\cos(-n\varphi)+i\sin(-n\varphi))}=r^n(\cos(n\varphi)+i\sin(n\varphi))$, т. е. формула (24) верна. ■

III. Корень n-ой степени из комплексного числа.

Определение 2. *Пусть $n\in N$. Корнем n-ой степени из комплексного числа z называется комплексное число w такое, что $w^n=z$.*

Существует единственный корень n-ой степени из $z=0$ — это $w=0$, так как $w^n=0\Leftrightarrow|w^n|=0\Leftrightarrow|w|^n=0\Leftrightarrow|w|=0\Leftrightarrow w=0$.

Получим формулу для нахождения корней n-ой степени из комплексного числа, заданного в тригонометрической форме $z=r(\cos\varphi+i\sin\varphi)$.

Теорема 1.15.2 (Корень n-ой степени из комплексного числа). *Для любого числа $z=r(\cos\varphi+i\sin\varphi)\neq0$ существует ровно n различных корней n-ой степени из z:*

$$w_k=\sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n}+i\sin\frac{\varphi+2\pi k}{n}\right),\ k=\overline{0,n-1}. \tag{25}$$

Доказательство. Пусть $w=\rho(\cos\alpha+i\sin\alpha)$ является корнем n-ой степени из $z=r(\cos\varphi+i\sin\varphi)$. Тогда $w^n=z\Rightarrow(\rho(\cos\alpha+i\sin\alpha))^n=r(\cos\varphi+$ $+\,i\sin\varphi)\Rightarrow\rho^n(\cos(n\alpha)+i\sin(n\alpha))=r(\cos\varphi+i\sin\varphi)\Rightarrow\rho^n=r,\ n\alpha=\varphi+$ $+\,2\pi k,\ k\in Z\Rightarrow\rho=\sqrt[n]{r},\ \alpha=\dfrac{\varphi+2\pi k}{n},\ k\in Z$. Отсюда получаем, что w имеет вид $w=\sqrt[n]{r}\left(\cos\dfrac{\varphi+2\pi k}{n}+i\sin\dfrac{\varphi+2\pi k}{n}\right),\ k\in Z$.

Обратно, если мы берём число $w=\sqrt[n]{r}\left(\cos\dfrac{\varphi+2\pi k}{n}+i\sin\dfrac{\varphi+2\pi k}{n}\right)$ при любом $k\in Z$, то $w^n=(\sqrt[n]{r})^n\left(\cos\dfrac{(\varphi+2\pi k)n}{n}+i\sin\dfrac{(\varphi+2\pi k)n}{n}\right)\Rightarrow w^n=$ $=r(\cos\varphi+i\sin\varphi)\Rightarrow w^n=z$, т. е. w является корнем n-ой степени из z.

Следовательно, числа $w_k = \sqrt[n]{r}\left(\cos\dfrac{\varphi+2\pi k}{n} + i\sin\dfrac{\varphi+2\pi k}{n}\right)$, $k \in Z$, и только они, являются корнями n-ой степени из z.

Все w_k имеют модуль $\sqrt[n]{r}$ и на комплексной плоскости они расположены на окружности радиуса $\sqrt[n]{r}$ с центром в 0, аргумент w_k равен $\alpha_k = \dfrac{\varphi+2\pi k}{n}$ или отличается от α_k на $2\pi m$, $m \in Z$. Увеличение k на 1 влечёт увеличение α_k на $\dfrac{2\pi}{n}$. Взяв для k возрастающую последовательность значений от 0 до $n-1$, мы получим возрастающую последовательность n различных значений аргументов $\alpha_0 = \dfrac{\varphi}{n} < \alpha_1 = \dfrac{\varphi}{n} + \dfrac{2\pi}{n} < \alpha_2 = \dfrac{\varphi}{n} + \dfrac{2\pi\cdot 2}{n} < \ldots < \alpha_{n-1} = \dfrac{\varphi}{n} + \dfrac{2\pi\cdot(n-1)}{n}$, которые различаются между собой самое большое на $\dfrac{2\pi\cdot(n-1)}{n} < 2\pi$, т. е. принадлежат промежутку $\left[\dfrac{\varphi}{n}, \dfrac{\varphi}{n} + 2\pi\right)$. Это означает, что на комплексной плоскости числа w_k, $k = \overline{0, n-1}$ расположены на окружности радиуса $\sqrt[n]{r}$ с центром в 0 и они делят эту окружность на n равных частей, поскольку разность аргументов $\alpha_{k+1} - \alpha_k = \dfrac{2\pi}{n}$.

Каждое число w_k для $k \in Z$ совпадает с одним из корней $\{w_0, w_1, w_2, \ldots, w_{n-1}\}$. Действительно, при $k = nm + r$ $(m, r \in Z,\ 0 \leqslant r \leqslant n-1)$ имеем

$$\frac{\varphi+2\pi k}{n} = \frac{\varphi+2\pi nm+2\pi r}{n} = \frac{\varphi+2\pi r}{n} + 2\pi m.$$

Следовательно, $w_k = w_r \in \{w_0, w_1, w_2, \ldots, w_{n-1}\}$.

Следствия. 1. Формулу (25) можно записать в виде равенства двух множеств

$$\sqrt[n]{r(\cos\varphi + i\sin\varphi)} = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n} + i\sin\frac{\varphi+2\pi k}{n}\right),\ k = \overline{0, n-1}. \qquad (26)$$

2. Все значения $\{w_0, w_1, w_2, \ldots, w_{n-1}\}$ корня n-ой степени из числа $z \neq 0$ на комплексной плоскости расположены на окружности радиуса $\sqrt[n]{|z|}$ с центром в нуле и они являются вершинами правильного n-угольника, причём $\arg(w_0) = \dfrac{\arg(z)}{n}$ и $\arg(w_k) = \arg(w_{k-1}) + \dfrac{2\pi}{n}$, $k = \overline{1, n-1}$.

3. Все значения корня n-ой степени из $z \neq 0$ можно получить умножением одного из его значений на все значения корня n-ой степени из единицы $\varepsilon_m = \cos\dfrac{2\pi m}{n} + i\sin\dfrac{2\pi m}{n}$, $m = \overline{0, n-1}$.

Доказательство. 3. Числа $w_k = \sqrt[n]{r}\left(\cos\dfrac{\varphi+2\pi k}{n} + i\sin\dfrac{\varphi+2\pi k}{n}\right)$, $k \in Z$ дают все значения корня n-ой степени из числа $z = r(\cos\varphi + i\sin\varphi)$.

Возьмём одно значение w_p и умножим его на все значения корня n-ой степени из единицы $\varepsilon_m = \cos\dfrac{2\pi m}{n} + i\sin\dfrac{2\pi m}{n}$, $m \in Z$. Получим

$$w_p\varepsilon_m = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi p}{n} + i\sin\frac{\varphi+2\pi p}{n}\right)\left(\cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n}\right) =$$

$= \sqrt[n]{r}\left(\cos\dfrac{\varphi+2\pi(p+m)}{n} + i\sin\dfrac{\varphi+2\pi(p+m)}{n}\right)$. Поскольку $p+m$ принимает любые целые значения последнее выражение содержит все значения корней w_k, $k = \overline{0, n-1}$. ■

IV. Квадратный корень из комплексного числа в алгебраической форме.

Из предыдущего раздела следует существование двух значений квадратного корня из заданного ненулевого комплексного числа $z = a + bi$. Получим эти значения в алгебраической форме $x + yi$.

Пусть $\sqrt{a+bi} = x + yi$, $x, y \in R$. Это означает, что $(x+yi)^2 = a + bi \Leftrightarrow$ $\Leftrightarrow x^2 + 2xyi - y^2 = a + bi \Leftrightarrow (x^2 - y^2) + 2xyi = a + bi \Leftrightarrow$

$$\begin{cases} x^2 - y^2 = a, \\ 2xy = b. \end{cases} \tag{27}$$

Возводя в квадрат обе части каждого из уравнений системы (27), а затем складывая их, получим уравнение, являющееся следствием системы (27):

$$(x^2 - y^2)^2 + 4x^2y^2 = a^2 + b^2 \Leftrightarrow x^4 - 2x^2y^2 + y^4 + 4x^2y^2 = a^2 + b^2 \Leftrightarrow$$
$$\Leftrightarrow x^4 + 2x^2y^2 + y^4 = a^2 + b^2 \Leftrightarrow (x^2 + y^2)^2 = a^2 + b^2 \Leftrightarrow x^2 + y^2 = +\sqrt{a^2 + b^2}.$$

Корень квадратный взят со знаком плюс, т.к. $x^2 + y^2 \in R$ и $x^2 + y^2 \geqslant 0$.

Добавив уравнение $x^2 + y^2 = \sqrt{a^2 + b^2}$ к (27), получим эквивалентную систему

$$\begin{cases} x^2 + y^2 = \sqrt{a^2 + b^2}, \\ x^2 - y^2 = a, \\ 2xy = b. \end{cases} \tag{28}$$

Складывая и вычитая два уравнения в (28), придём к эквивалентной системе

$$\begin{cases} x^2 = \dfrac{a+\sqrt{a^2+b^2}}{2}, \\ y^2 = \dfrac{-a+\sqrt{a^2+b^2}}{2}, \\ 2xy = b. \end{cases} \tag{29}$$

При $b = 0$, $a > 0$ из системы (29) найдём $x^2 = a$, $y^2 = 0$ и получим $x = \pm\sqrt{a}$, $y = 0$. Следовательно,

$$\sqrt{a+0i} = \pm\sqrt{a} + 0i \ \text{ при } \ a > 0.$$

При $b = 0$, $a < 0$ из системы (29) найдём $x^2 = 0$, $y^2 = -a > 0$ и получим $x = 0$, $y = \pm\sqrt{-a}$. Следовательно,

$$\sqrt{a+0i} = 0 \pm \sqrt{-a}i \ \text{ при } \ a < 0.$$

Далее рассмотрим случай $b \neq 0$.

Для любых $a, b \in R$, $b \neq 0$ справедливы неравенства

$$a^2+b^2 > a^2 \Leftrightarrow \sqrt{a^2+b^2} > |a| \Leftrightarrow \begin{cases} \sqrt{a^2+b^2} > a, \\ \sqrt{a^2+b^2} > -a, \end{cases} \Leftrightarrow \begin{cases} -a+\sqrt{a^2+b^2} > 0, \\ a+\sqrt{a^2+b^2} > 0. \end{cases}$$

Тогда из системы (29) найдём

$$x = \pm\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}}, \quad y = \pm\sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}. \tag{30}$$

Причём в силу уравнения $2xy = b$ знаки для x и y в (30) выбираем так, чтобы знак произведения xy совпал со знаком b. При $b > 0$ получаем две пары значений $x > 0$, $y > 0$ и $x < 0$, $y < 0$. При $b < 0$ получаем две пары значений $x > 0$, $y < 0$ и $x < 0$, $y > 0$. Следовательно,

$$\sqrt{a+bi} = \pm\left(\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} + \sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\,i\right) \quad \text{при } b > 0,$$

$$\sqrt{a+bi} = \pm\left(\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} - \sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\,i\right) \quad \text{при } b < 0.$$

Используя функцию $\operatorname{sgn}(x) = \begin{cases} 1 \text{ при } x \geqslant 0, \\ -1 \text{ при } x < 0, \end{cases}$ все результаты можно записать в виде

$$\sqrt{a+bi} = \pm\left(\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} + \operatorname{sgn}(b)\sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\, i\right).$$

Таким образом, из ненулевого комплексного числа существует два значения квадратного корня, которые являются противоположными числами. ■

Задача 6. *Найдите* $(1-i)^{20}$.

Решение. Преобразуем $w = (1-i)^{20} = \left((1-i)^2\right)^{10} = (1-2i+i^2)^{10} = (-2i)^{10} = (-2)^{10}i^{10}$. Поскольку $(-2)^{10} = 2^{10} = 1024$ и $i^{10} = i^2 = -1$, получим $w = -1024$.

Задача 7. *Найдите* $(-3+i)^{-4}$.

Решение. Первый способ. Представим отрицательную степень в виде положительной степени обратного числа $(-3+i)^{-4} = \left(\dfrac{1}{-3+i}\right)^4$. Затем выполним деление $\dfrac{1}{-3+i} = \dfrac{-3-i}{(-3+i)(-3-i)} = \dfrac{-3-i}{(-3)^2-i^2} = \dfrac{-3-i}{10}$. Далее возведём в степень $\left(\dfrac{-3-i}{10}\right)^4 = \dfrac{(-1)^4\left((3+i)^2\right)^2}{10^4} = \dfrac{(9+6i-1)^2}{10^4} = \dfrac{(8+6i)^2}{10^4} = \dfrac{(2(4+3i))^2}{10^4} = \dfrac{2^2(4+3i)^2}{10^4} = \dfrac{4(16+24i-9)}{10^4} = \dfrac{7+24i}{2500} = \dfrac{7}{2500} + \dfrac{6}{625}i$.

Второй способ. Представим отрицательную степень в виде обратного значения к положительной степени $(-3+i)^{-4} = \dfrac{1}{(-3+i)^4}$. Затем возведём в положительную степень $(-3+i)^4 = ((-3+i)^2)^2 = (9-6i-1)^2 = (8-6i)^2 = (2(4-3i))^2 = 4(16-24i-9) = 4(7-24i)$. Далее выполним деление $\dfrac{1}{(-3+i)^4} = \dfrac{1}{4(7-24i)} = \dfrac{7+24i}{4(7-24i)(7+24i)} = \dfrac{7+24i}{4(49+576)} = \dfrac{7+24i}{4\cdot 625} = \dfrac{7}{2500} + \dfrac{6}{625}i$.

Задача 8. *Найдите* $(-1+\sqrt{3}i)^{22}$.

Решение. Представим число $z = -1+\sqrt{3}i$ в тригонометрической форме. Находим $x = \mathrm{Re}(z) = -1$, $y = \mathrm{Im}(z) = \sqrt{3}$, $r = |z| = \sqrt{x^2+y^2} = \sqrt{1+3} = \sqrt{4} = 2$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{1}{2}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{\sqrt{3}}{2}$, $\varphi = \dfrac{2\pi}{3}$. Отсюда $-1+\sqrt{3}i = 2\left(\cos\dfrac{2\pi}{3} + i\sin\dfrac{2\pi}{3}\right)$. Далее $(-1+\sqrt{3}i)^{22} = \left(2\left(\cos\dfrac{2\pi}{3} + i\sin\dfrac{2\pi}{3}\right)\right)^{22} = 2^{22}\left(\cos\dfrac{2\pi\cdot 22}{3} + i\sin\dfrac{2\pi\cdot 22}{3}\right) = 2^{22}\left(\cos\dfrac{42\pi+2\pi}{3} + i\sin\dfrac{42\pi+2\pi}{3}\right) = 2^{22}\left(\cos\left(14\pi+\dfrac{2\pi}{3}\right) + i\sin\left(14\pi+\dfrac{2\pi}{3}\right)\right) = 2^{22}\left(\cos\left(\dfrac{2\pi}{3}\right) + i\sin\left(\dfrac{2\pi}{3}\right)\right) = 2^{22}\left(-\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}\right) = 2^{21}(-1+\sqrt{3}i)$.

Задача 9. *Найдите* $\left(\dfrac{4}{5} - \dfrac{3}{5}i\right)^{-100}$.

Решение. Представим число $z = \dfrac{4}{5} - \dfrac{3}{5}i$ в тригонометрической форме. Находим $x = \mathrm{Re}(z) = \dfrac{4}{5}$, $y = \mathrm{Im}(z) = -\dfrac{3}{5}$, $r = |z| = \sqrt{x^2+y^2} = \sqrt{\left(\dfrac{4}{5}\right)^2 + \left(\dfrac{3}{5}\right)^2} = \sqrt{\dfrac{16}{25} + \dfrac{9}{25}} = 1$, $\cos\varphi = \dfrac{x}{r} = \dfrac{4}{5}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{3}{5}$.
Тогда $\mathrm{tg}\,\varphi = -\dfrac{3}{4}$ и в четвёртой четверти возьмём $\varphi = \mathrm{arctg}\left(-\dfrac{3}{4}\right) = -\,\mathrm{arctg}\,\dfrac{3}{4}$.
Отсюда $\dfrac{4}{5} - \dfrac{3}{5}i = \cos\left(-\,\mathrm{arctg}\,\dfrac{3}{4}\right) + i\sin\left(-\,\mathrm{arctg}\,\dfrac{3}{4}\right)$. И далее по формуле Муавра находим $\left(\dfrac{4}{5} - \dfrac{3}{5}i\right)^{-100} = \left(\cos\left(-\,\mathrm{arctg}\,\dfrac{3}{4}\right) + i\sin\left(-\,\mathrm{arctg}\,\dfrac{3}{4}\right)\right)^{-100} = \cos\left(100\,\mathrm{arctg}\,\dfrac{3}{4}\right) + i\sin\left(100\,\mathrm{arctg}\,\dfrac{3}{4}\right) \approx 0,053 + 0,999i$.

Задача 10. *Выполнить действия* $\dfrac{(-\sqrt{3}-i)^{20}}{(1-\sqrt{3}i)^4}$.

Решение.

Представим числа $-\sqrt{3}-i$, $1-\sqrt{3}i$ в тригонометрической форме и выполним действия.

Для $z = -\sqrt{3}-i$ имеем $x = \mathrm{Re}(z) = -\sqrt{3}$, $y = \mathrm{Im}(z) = -1$, тогда $r = |z| = \sqrt{x^2+y^2} = \sqrt{(-\sqrt{3})^2 + (-1)^2} = \sqrt{4} = 2$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{\sqrt{3}}{2}$, $\sin\varphi =$

$= \dfrac{y}{r} = -\dfrac{1}{2}$. Отсюда $\varphi = \dfrac{7\pi}{6}$, $-\sqrt{3} - i = 2\left(\cos\dfrac{7\pi}{6} + i\sin\dfrac{7\pi}{6}\right)$ и $(-\sqrt{3} - i)^{20} =$
$= 2^{20}\left(\cos\dfrac{7\pi\cdot 20}{6} + i\sin\dfrac{7\pi\cdot 20}{6}\right) = 2^{20}\left(\cos\dfrac{70\pi}{3} + i\sin\dfrac{70\pi}{3}\right)$.

Для $z = 1-\sqrt{3}i$ имеем $x = \mathrm{Re}(z) = 1$, $y = \mathrm{Im}(z) = -\sqrt{3}$, тогда $r = |z| =$ $\sqrt{1+3} = 2$, $\cos\varphi = \dfrac{x}{r} = \dfrac{1}{2}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{\sqrt{3}}{2}$. Отсюда $\varphi = -\dfrac{\pi}{3}$, $1 - \sqrt{3}i =$ $= 2\left(\cos\left(-\dfrac{\pi}{3}\right) + i\sin\left(-\dfrac{\pi}{3}\right)\right)$ и $(1 - \sqrt{3}i)^4 = 2^4\left(\cos\left(-\dfrac{4\pi}{3}\right) + i\sin\left(-\dfrac{4\pi}{3}\right)\right)$.

Далее выполняем деление $\dfrac{(-\sqrt{3} - i)^{20}}{(1 - \sqrt{3}i)^4} = \dfrac{2^{20}\left(\cos\dfrac{70\pi}{3} + i\sin\dfrac{70\pi}{3}\right)}{2^4\left(\cos\left(-\dfrac{4\pi}{3}\right) + i\sin\left(-\dfrac{4\pi}{3}\right)\right)} =$
$= \dfrac{2^{20}}{2^4}\left(\cos\dfrac{74\pi}{3} + i\sin\dfrac{74\pi}{3}\right) = 2^{16}\left(\cos\dfrac{72\pi + 2\pi}{3} + i\sin\dfrac{72\pi + 2\pi}{3}\right) =$
$= 2^{16}\left(\cos\left(24\pi + \dfrac{2\pi}{3}\right) + i\sin\left(24\pi + \dfrac{2\pi}{3}\right)\right) = 2^{16}\left(\cos\left(\dfrac{2\pi}{3}\right) + i\sin\left(\dfrac{2\pi}{3}\right)\right) =$
$= 2^{16}\left(-\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}\right) = 2^{15}(-1 + \sqrt{3}i)$.

Задача 11. *Найдите* $\sqrt{3 + 4i}$.

Решение. Обозначим $\sqrt{3 + 4i} = x + yi$. Тогда $(x + yi)^2 = 3 + 4i \Leftrightarrow (x^2 -$
$- y^2) + 2xyi = 3 + 4i \Leftrightarrow \begin{cases} x^2 - y^2 = 3, \\ 2xy = 4, \end{cases} \Leftrightarrow \begin{cases} (x^2 - y^2)^2 + 4x^2y^2 = 3^2 + 4^2, \\ x^2 - y^2 = 3, \\ 2xy = 4, \end{cases} \Leftrightarrow$

$\Leftrightarrow \begin{cases} (x^2 + y^2)^2 = 25, \\ x^2 - y^2 = 3, \\ 2xy = 4, \end{cases} \Leftrightarrow \begin{cases} x^2 + y^2 = +5, \\ x^2 - y^2 = 3, \\ xy = 2, \end{cases} \Leftrightarrow \begin{cases} 2x^2 = 8, \\ 2y^2 = 2, \\ xy = 2, \end{cases} \Leftrightarrow \begin{cases} x^2 = 4, \\ y^2 = 1, \\ xy = 2, \end{cases} \Leftrightarrow$

$\Leftrightarrow \begin{cases} x = \pm 2, \\ y = \pm 1, \\ xy > 0, \end{cases} \Leftrightarrow x + yi = \pm(2 + i)$. Ответ: $2 + i$, $-2 - i$.

Задача 12. *Найдите* $\sqrt[4]{-1}$.

Решение. Поскольку $w = \sqrt[4]{z} \Leftrightarrow w^4 = z \Leftrightarrow (w^2)^2 = z \Leftrightarrow w^2 = \sqrt{z} \Leftrightarrow w =$ $= \sqrt{\sqrt{z}}$. Следовательно, нужно найти $\sqrt{\sqrt{-1}}$.

В силу $i^2=-1$ имеем $\sqrt{-1}=\pm i$. Далее обозначим $\sqrt{\pm i}=x+yi$.

Тогда $(x+yi)^2=\pm i \Leftrightarrow (x^2-y^2)+2xyi=\pm i \Leftrightarrow \begin{cases} x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow$

$$\Leftrightarrow \begin{cases} (x^2-y^2)^2+4x^2y^2=0+1, \\ x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow \begin{cases} (x^2+y^2)^2=1, \\ x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow \begin{cases} x^2+y^2=1, \\ x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow\Leftrightarrow \begin{cases} x^2=1/2, \\ y^2=1/2, \\ xy=\pm 1/2, \end{cases} \Leftrightarrow \begin{cases} x=\pm 1/\sqrt{2}, \\ y=\pm 1/\sqrt{2}, \\ xy=\pm 1/2. \end{cases}$$ Отсюда находим $x+yi=\pm\dfrac{\sqrt{2}}{2}(1\pm i)$.

Ответ: $\dfrac{\sqrt{2}}{2}(1+i)$, $-\dfrac{\sqrt{2}}{2}(1+i)$, $\dfrac{\sqrt{2}}{2}(1-i)$, $-\dfrac{\sqrt{2}}{2}(1-i)$.

Задача 13. *Найдите* $\sqrt[4]{-8-8\sqrt{3}i}$ *и изобразите результаты на комплексной плоскости.*

Решение. Представим число $z=-8-8\sqrt{3}i$ в тригонометрической форме и извлечём корень.

Для $z=-8-8\sqrt{3}i$ имеем $x=\mathrm{Re}(z)=-8$, $y=\mathrm{Im}(z)=-8\sqrt{3}$, тогда $r=|z|=\sqrt{x^2+y^2}=\sqrt{8^2+8^2\cdot 3}=\sqrt{8^2\cdot 4}=16$, $\cos\varphi=\dfrac{x}{r}=-\dfrac{1}{2}$, $\sin\varphi=\dfrac{y}{r}=-\dfrac{\sqrt{3}}{2}$, $\varphi=\dfrac{4\pi}{3}$. Отсюда $-8-8\sqrt{3}i=16\left(\cos\dfrac{4\pi}{3}+i\sin\dfrac{4\pi}{3}\right)$ и далее

$$\sqrt[4]{16\left(\cos\frac{4\pi}{3}+i\sin\frac{4\pi}{3}\right)}=\sqrt[4]{16}\left(\cos\frac{\frac{4\pi}{3}+2\pi k}{4}+i\sin\frac{\frac{4\pi}{3}+2\pi k}{4}\right)=$$

$$=2\left(\cos\left(\frac{\pi}{3}+\frac{\pi k}{2}\right)+i\sin\left(\frac{\pi}{3}+\frac{\pi k}{2}\right)\right),\ k=\overline{0,3}.$$

Обозначим полученные значения через w_k, тогда ответами в тригонометрической форме являются

$$w_0=2\left(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}\right),\qquad w_1=2\left(\cos\frac{5\pi}{6}+i\sin\frac{5\pi}{6}\right),$$
$$w_2=2\left(\cos\frac{4\pi}{3}+i\sin\frac{4\pi}{3}\right),\ w_3=2\left(\cos\frac{11\pi}{6}+i\sin\frac{11\pi}{6}\right).$$

Ответы изображены на рис.15, где точки, изображающие числа w_0, w_1, w_2, w_3, являются вершинами квадрата, вписанного в окружность радиуса 2

с центром в начале координат. Для w_0 имеем $\arg(w_0) = \arg(z)/4 = \pi/3$ и $\arg(w_1) - \arg(w_0) = \arg(w_2) - \arg(w_1) = \arg(w_3) - \arg(w_2) = \arg(w_0) - \arg(w_3) = = \pi/2$.

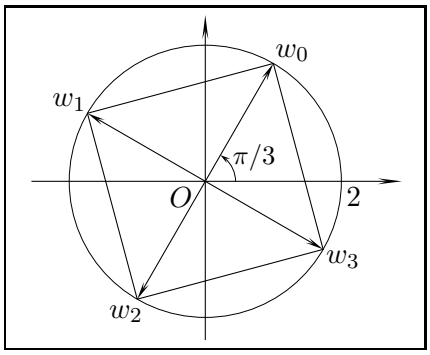

Рис. 15

1.16. Показательная функция и логарифм от комплексной переменной. Возведение комплексного числа в комплексную степень

Определение 1. *Показательную функцию от комплексной переменной* $z = x + yi\,(x,\, y \in R)$ *можно определить формулой Эйлера*

$$e^z = e^x(\cos y + i \sin y). \tag{31}$$

Используя формулу Эйлера, тригонометрическую форму комплексного числа $z = x + yi = r(\cos(\varphi + 2\pi k) + i \sin(\varphi + 2\pi k))$, где $r = |z| = \sqrt{x^2 + y^2}$, $\varphi = = arg(z)$, можно представить в виде $z = re^{i(\varphi+2\pi k)} = e^{\ln r} e^{i(\varphi+2\pi k)} = e^{\ln r + i(\varphi+2\pi k)}$.

Это делает естественным следующее определение.

Определение 2. *Натуральным логарифмом от комплексного числа* $z = x + yi$, *где* $r = |z| = \sqrt{x^2 + y^2}$, $\varphi = arg(z)$ *назовём величину*

$$\ln z = \ln r + i(\varphi + 2\pi k),\ k \in Z. \tag{32}$$

Натуральный логарифм определён с точностью до слагаемого $2\pi ki$, $k \in Z$.

Пусть заданы произвольные $z_1,\, z_2 \in C$ $z_1 \neq 0$ с помощью функции натуральный логарифм определим многозначную функцию возведения комплексного числа в комплексную степень:

$$z_1^{z_2} = e^{z_2 \ln z_1} = e^{z_2(\ln|z_1| + i(\arg z_1 + 2\pi k)}, \ k \in Z. \tag{33}$$

Задача 14. *Докажите или опровергните утверждения:*

1) *Если* $z = x + yi$, *то* $|e^z| = e^x$, $\operatorname{Arg} e^z = y + 2\pi ki$, $k \in Z$.

2) $\forall z_1,\ z_2 \in C\ e^{z_1+z_2} = e^{z_1}e^{z_2}$.

3) $\forall z_1,\ z_2 \in C\ (e^{z_1})^{z_2} = e^{z_1 z_2}$.

4) *Докажите, что* e^z *— периодическая функция.*

5) *Функция* e^z *принимает все комплексные значения кроме* 0, *причём уравнение* $e^z = w$ *при* $w \neq 0$ *имеет бесконечное число решений.*

Задача 15. *Найдите все решения* $z \in C$ *уравнения* $e^z = w$.

Задача 16. *Найдите наименьший по модулю период функция* e^z, $z \in C$.

Задача 17. *Найдите*

1) $x^{iy},\ x > 0, y \in R,$ 2) $(-x)^{iy},\ x > 0, y \in R,$ 3) $(iy)^x,\ x \in R, y > 0,$

4) $(-iy)^x,\ x \in R, y > 0,$ 5) $(iy_1)^{iy_2},\ y_1 > 0, y_2 \in R,$ 6) $(-iy_1)^{iy_2},\ y_1 > 0, y_2 \in R,$

7) $1^i,$ 8) $(-1)1^i,$ 9) $1^{-i},$ 10) $(-1)^{-i},$ 11) $i^i,$ 12) $(-i)1^i,$ 13) $i{-}i,$ 14) $(-i)1^{-i}$.

1.17. Кольцо многочленов от одной переменной

Определение 1. 1) *Пусть дана последовательность чисел из поля* P:

$$f_0,\ f_1,\ f_2,\ \ldots,\ f_k,\ \ldots \tag{34}$$

причём, начиная с некоторого номера $m \geqslant 0$, *все* $f_k = 0$ *при* $k \geqslant m$.

Многочленом над полем P *от переменной* x *называется выражение*

$$f_0 + f_1\,x + f_2\,x^2 + \ldots + f_k\,x^k + \ldots = \sum_{k=0}^{\infty} f_k\,x^k, \tag{35}$$

где (34) — *коэффициенты многочлена. Обозначение многочлена:* f *или* $f(x)$.

Нулевым называем многочлен $\sum\limits_{k=0}^{\infty} 0\,x^k$; *его обозначение:* 0.

Через $P[x]$ *обозначим множество всех многочленов над полем* P *от переменной* x.

2) *Два многочлена* $f(x) = \sum\limits_{k=0}^{\infty} f_k\,x^k$ *и* $g(x) = \sum\limits_{k=0}^{\infty} g_k\,x^k$ *из* $P[x]$ *называются равными, если равны все их соответствующие коэффициенты* $f_k = g_k$, $k = 0, 1, 2, \ldots$.

Обозначение: равные многочлены $f = g$ *или* $f(x) = g(x)$, *неравные многочлены* $f \neq g$ *или* $f(x) \neq g(x)$.

3) *Для ненулевого многочлена* $f = \sum\limits_{k=0}^{\infty} f_k\,x^k$ *степень — это максимальное число* n, *для которого* $f_n \neq 0$.

Степень нулевого многочлена считается равной $(-\infty)$.

Обозначение степени многочлена: $\deg(f)$ *или* $\deg f$. ■

Замечания.

1. При записи ненулевого многочлена принято опускать все члены с нулевыми коэффициентами.

2. Многочлен нулевой степени — это $f(x) = f_0$, причём $f_0 \neq 0$.

3. Многочлен степени n можно записать в виде $f(x) = \sum\limits_{k=0}^{n} f_k\,x^k = f_0 + f_1\,x + f_2\,x^2 + \ldots + f_{n-1}\,x^{n-1} + f_n\,x^n$ (по возрастанию степеней) или $f(x) = \sum\limits_{k=0}^{n} a_k\,x^{n-k} = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ (по убыванию степеней), здесь $a_k = f_{n-k}$, $k = \overline{0, n}$.

4. Операции и неравенства с $(-\infty)$:

$(-\infty) + n = n + (-\infty) = (-\infty)$, если $n \in Z_+$;
$(-\infty) + (-\infty) = (-\infty)$;
$(-\infty) < n$, если $n \in Z_+$. ■

Введём на множестве многочленов $P[x]$ алгебраические операции.

Определение 2. *Пусть* $f \in P[x]$, $g \in P[x]$,
$f = f_0 + f_1\,x + f_2\,x^2 + \ldots + f_k\,x^k + \ldots,$
$g = g_0 + g_1\,x + g_2\,x^2 + \ldots + g_k\,x^k + \ldots.$
I. *Суммой многочленов* f *и* g *называется выражение*

$$(f_0 + g_0) + (f_1 + g_1)\,x + \ldots + (f_k + g_k)\,x^k + \ldots = \sum_{k=0}^{\infty} (f_k + g_k)\,x^k. \tag{36}$$

Обозначение: $f + g$.

II. *Произведением многочленов* f *и* g *называется выражение*[1]

$$f_0\,g_0 + (f_0\,g_1 + f_1\,g_0)\,x + (f_0\,g_2 + f_1\,g_1 + f_2\,g_0)\,x^2 + \ldots + (f_0\,g_k + f_1\,g_{k-1} + \ldots + f_{k-1}\,g_1 + f_k\,g_0)\,x^k + \ldots = \sum_{k=0}^{\infty} \Big(\sum_{r+s=k} f_r\,g_s \Big) x^k. \tag{37}$$

Обозначение fg. ■

Лемма 1.17.1. *Если* f *и* g *многочлены из* $P[x]$, *то сумма* $f+g$ *и произведение* fg *также являются многочленами из* $P[x]$ *и для них выполнены свойства:*

$$\deg(f+g) \leqslant \max\{\deg(f), \deg(g)\}, \tag{38}$$

$$\deg(fg) = \deg(f) + \deg(g). \tag{39}$$

Доказательство.

Пусть f, $g \in P[x]$,

$$f = \sum_{k=0}^{\infty} f_k\,x^k = f_0 + f_1\,x + f_2\,x^2 + \ldots + f_k\,x^k + \ldots\,,$$

$$g = \sum_{k=0}^{\infty} g_k\,x^k = g_0 + g_1\,x + g_2\,x^2 + \ldots + g_k\,x^k + \ldots\,.$$

Для f и g коэффициенты f_k и g_k принадлежат полю P при всех $k \in Z_+$. Поэтому принадлежат полю P суммы и произведения $f_k + g_k$, $f_r\,g_s$, при всех $k, r, s \in Z_+$, а также $\sum\limits_{r+s=k} f_r\,g_s$ при всех $k \in Z_+$.

Обозначим $m = \deg(f)$ и $n = \deg(g)$.

[1] Суммирование в $\sum\limits_{p+r=k} f_p\,g_r$ ведётся по всем целым индексам $p \geqslant 0$, $r \geqslant 0$, для которых $p + r = k$.

1) Рассмотрим случай, когда оба многочлена ненулевые.

Поскольку $f \neq 0$ и $g \neq 0$, то $f = f_0 + f_1 x + \ldots + f_m x^m + \sum\limits_{k=m+1}^{\infty} 0\, x^k$, $g = g_0 + g_1 x + \ldots + g_n x^n + \sum\limits_{k=n+1}^{\infty} 0\, x^k$, где $m \geqslant 0$, $n \geqslant 0$, $f_m \neq 0$, и $g_n \neq 0$.

Для суммы $f+g$ имеем $f_k + g_k = 0$ при $k > \max\{m, n\}$. Следовательно, $f+g$ — многочлен из $P[x]$ и его степень не превосходит $\max\{m, n\}$. Отсюда следует, что $\deg(f+g) \leqslant \max(m, n) = \max\{\deg(f), \deg(g)\}$, т.е. выполнено неравенство (38).

Обозначим $fg = h$, где $h = \sum\limits_{k=0}^{\infty} h_k\, x^k$ и $h_k = \sum\limits_{r+s=k} f_r\, g_s$.

Докажем, что при $k > m+n$ коэффициенты $h_k = 0$, а при $k = m+n$ коэффициент $h_k = f_m\, g_n \neq 0$, т.е. все коэффициенты h_k произведения $h = fg$ равны 0 для индексов $k > m+n$, а $h_{m+n} \neq 0$.

Рассмотрим $h_k = \sum\limits_{r+s=k} f_r\, g_s$ при $k > m+n$. Здесь $r+s = k > m+n$, тогда $r > m$ или $s > n$. В противном случае, из $r \leqslant m$ и $s \leqslant n$ следует неравенство $r+s \leqslant m+n$, что противоречит $r+s > m+n$. При $r > m$ имеем $f_r = 0$, при $s > n$ имеем $g_s = 0$, т.е. все произведения $f_r\, g_s = 0$ и $h_k = \sum\limits_{r+s=k} f_r\, g_s = 0$ при $r+s = k > m+n$. Отсюда следует, что $h = fg$ — многочлен из $P[x]$ и $\deg(fg) \leqslant m+n = \deg(f) + \deg(g)$.

Рассмотрим $h_k = \sum\limits_{r+s=k} f_r\, g_s$ для $k = m+n$. Здесь $r+s = k = m+n$. При $r > m$ или $s > n$ имеем $f_r = 0$ или $g_s = 0$, тогда $f_r\, g_s = 0$. При $r \leqslant m$ и $s \leqslant n$ равенство $r+s = k = m+n$ выполнено только для $r = m$, $s = n$. Тогда $h_{m+n} = \sum\limits_{r+s=m+n} f_r\, g_s = f_m\, g_n \neq 0$. Отсюда и из предыдущего абзаца следует, что $\deg(fg) = m+n = \deg(f) + \deg(g)$, т.е. выполнено равенство (39).

2) Теперь рассмотрим случай, когда хотя бы один многочлен нулевой, например, $g = 0 = \sum\limits_{k=0}^{\infty} 0\, x^k$. Из определения получим $f+g = \sum\limits_{k=0}^{\infty} (f_k + 0)\, x^k = \sum\limits_{k=0}^{\infty} f_k\, x^k = f$, $fg = \sum\limits_{k=0}^{\infty} \Big(\sum\limits_{r+s=k} f_r\, 0 \Big) x^k = \sum\limits_{k=0}^{\infty} 0\, x^k = 0 = g$. Следовательно, $f+g = f \in P[x]$, $fg = g \in P[x]$.

Рассмотрим степени этих многочленов:

$$\deg(f+g) = \deg(f) = \max\{\deg(f), (-\infty)\} = \max\{\deg(f), \deg(g)\},$$
$$\deg(fg) = \deg(g) = (-\infty) = \deg(f) + (-\infty) = \deg(f) + \deg(g).$$

Отсюда следует выполнение (38), (39).

Лемма доказана.

Следствия.

1) *Для многочленов* $f, g \in P[x]$ *отображение* $\langle f, g\rangle \mapsto f+g$ *определяет алгебраическую операцию сложения на* $P[x]$.

2) *Для многочленов* $f, g \in P[x]$ *отображение* $\langle f, g\rangle \mapsto f\,g$ *определяет алгебраическую операцию умножения на* $P[x]$.

3) *Для числа* $\alpha \in P$ *(частный случай многочлена* $f(x) = \alpha$*) и многочлена* $g \in P[x]$ *отображение* $\langle \alpha, g\rangle \mapsto \alpha\, g$ *определяет алгебраическую операцию умножения многочлена на число.* ■

Теорема 1.17.1. I. *Множество многочленов* $P[x]$ *с операциями сложения и умножения многочленов образует коммутативное кольцо с единицей и без делителей нуля, т. е. выполнены следующие свойства:*

1) $\forall f, g, h \in P[x] \quad (f+g)+h = f+(g+h)$,

2) $\exists\, 0 \in P[x]\, \forall f \in P[x] \;\; f+0 = f$,

3) $\forall f \in P[x]\, \exists (-f) \in P[x] \quad f+(-f) = 0$,

4) $\forall f, g \in P[x] \quad f+g = g+f$,

5) $\forall f, g, h \in P[x] \quad (fg)h = f(gh)$,

6) $\exists 1 \in P[x]\, \forall f \in P[x] \;\; f1 = f$,

7) $\forall f, g \in P[x] \quad fg = gf$,

8) $\forall f, g, h \in P[x] \quad (f+g)h = fh+gh$,

9) *из условий* $f, g \in P[x]$, $fg = 0$ *следует* $f = 0$ *или* $g = 0$.

Доказательство.

Докажем некоторые свойства.

Отметим, что нулевым многочленом является $0 = \sum\limits_{k=0}^{\infty} 0\, x^k$, противоположным к многочлену $f = \sum\limits_{k=0}^{\infty} f_k\, x^k$ является многочлен $(-f) = \sum\limits_{k=0}^{\infty} (-f_k)\, x^k$, единичным многочленом служит многочлен $1 = \sum\limits_{k=0}^{\infty} f_k\, x^k$, где $f_0 = 1$, $f_k = 0$ при $k \geqslant 1$.

5) Пусть $f = \sum\limits_{k=0}^{\infty} f_k\, x^k$, $g = \sum\limits_{k=0}^{\infty} g_k\, x^k$, $h = \sum\limits_{k=0}^{\infty} h_k\, x^k$ произвольные многочлены. Докажем ассоциативность умножения $(fg)h = f(gh)$

Обозначим многочлены $fg = s$, $sh = t$, $gh = v$, $fv = w$ с коэффициентами s_k, t_k, v_k, w_k соответственно, и покажем, что коэффициенты $t_k = w_k$ при всех $k \in Z_+$.

Из формулы (37) для операции умножения получим

$$t_k = \sum_{c+r=k} s_c\, h_r = \sum_{c+r=k} \Big(\sum_{j+l=c} f_j\, g_l \Big) h_r = \sum_{c+r=k} \Big(\sum_{j+l=c} f_j\, g_l\, h_r \Big) = \sum_{j+l+r=k} f_j\, g_l\, h_r,$$

$$w_k = \sum_{j+c=k} f_j\, v_c = \sum_{j+c=k} f_j \Big(\sum_{l+r=c} g_l\, h_r \Big) = \sum_{j+c=k} \Big(\sum_{l+r=c} f_j\, g_l\, h_r \Big) = \sum_{j+l+r=k} f_j\, g_l\, h_r.$$

Следовательно, $t_k = w_k$ при всех k, т.е. $(fg)h = f(gh)$.

8) Докажем дистрибутивность умножения относительно сложения.

Пусть $f = \sum\limits_{k=0}^{\infty} f_k\, x^k$, $g = \sum\limits_{k=0}^{\infty} g_k\, x^k$, $h = \sum\limits_{k=0}^{\infty} h_k\, x^k$. Выполняем операции сначала в левой части равенства $f + g = \sum\limits_{k=0}^{\infty} (f_k + g_k)\, x^k$, $(f + g)h = \sum\limits_{k=0}^{\infty} \Big(\sum\limits_{c+r=k} (f_c + g_c) h_r \Big) x^k = \sum\limits_{k=0}^{\infty} \Big(\sum\limits_{c+r=k} (f_c\, h_r + g_c\, h_r) \Big) x^k$, затем в правой части равенства $fh = \sum\limits_{k=0}^{\infty} \Big(\sum\limits_{c+r=k} f_c\, h_r \Big) x^k$, $gh = \sum\limits_{k=0}^{\infty} \Big(\sum\limits_{c+r=k} g_c\, h_r \Big) x^k$, $fh + gh = \sum\limits_{k=0}^{\infty} \Big(\sum\limits_{c+r=k} (f_c\, h_r + g_c\, h_r) \Big) x^k$.

Отсюда следует равенство $(f + g)h = fh + gh$.

9) Докажем отсутствие делителей нуля.

Пусть $f, g \in P[x]$ и $fg = 0$. Из равенства (39) получаем $\deg(f) + \deg(g) = \deg(fg) = \deg(0) = (-\infty)$. Отсюда следует, что $\deg(f) = (-\infty)$ или $\deg(g) = (-\infty)$, т.е. $f = 0$ или $g = 0$.

Следствия.

1) Нуль единственен.

2) Для любого многочлена противоположный к нему единственен.

3) Из условий $f, g, h \in P[x]$, $f + h = g + h$ или $h + f = h + g$ следует равенство $f = g$.

4) Для любых $f, g \in P[x]$ существует, и притом единственное, решение $y = g + (-f) = (-f) + g$ уравнения $y + f = g\,(f + y = g)$; это решение называется **разностью** g и f и обозначается $g - f$. Итак, $g - f = g + (-f)$ и эта формула определяет операцию **вычитания**.

5) Умножение дистрибутивно относительно вычитания, т.е.

$\forall f, g, h \in P[x] \quad (f - g)h = fh - gh$ и $h(f - g) = hf - hg$.

6) Единичный многочлен единственен.

7) Из условий $f, g, h \in P[x]$, $h \neq 0$ и $fh = gh$ или $hf = hg$ следует равенство $f = g$.

Доказательство следствий проведите самостоятельно. ■

Замечания.

1) *Многочлен $f(x) \in P[x]$ имеет обратный $g(x) \in P[x]$ тогда и только тогда, когда $f(x)$ — многочлен нулевой степени.*

2) *Кольцо $P[x]$ не является полем, т.к. не всякий ненулевой многочлен имеет обратный многочлен. Поэтому для умножения многочленов обратная операция — деление — не существует.*

3) *Произведение (37) многочленов $f(x) = f_0 + f_1\,x + f_2\,x^2 + \ldots + f_n\,x^n$ степени n и $g(x) = g_0 + g_1\,x + g_2\,x^2 + \ldots + g_m\,x^m$ степени m можно найти перемножением двух сумм $(f_0 + f_1\,x + f_2\,x^2 + \ldots + f_n\,x^n)(g_0 + g_1\,x + g_2\,x^2 + \ldots + g_m\,x^m)$ с раскрытием скобок и последующей группировкой слагаемых с одинаковыми степенями переменной.*

Доказательство.

1) Если $f(x)$ — многочлен нулевой степени из $P[x]$, т.е. это ненулевое число $f_0 \in P$, то возьмём из $P[x]$ многочлен нулевой степени $g(x) = f_0^{-1}$. Тогда $f(x)g(x) = f_0 f_0^{-1} = 1$, т.е. обратный многочлен существует.

Если для $f(x) \in P[x]$ существует обратный многочлен $g(x) \in P[x]$, то $f(x)g(x) = 1$. Отсюда следует, что $\deg(f) + \deg(g) = \deg(1)$, т.е. $\deg(f) + \deg(g) = 0$. Тогда $\deg(f) = \deg(g) = 0$, т.е. $f(x)$ и $g(x)$ — многочлены нулевой степени.

2) Следует из 1).

3) Это следует из дистрибутивности и из того, что многочлены $f(x)$ и $g(x)$ можно рассматривать как суммы одночленов f_0, f_1x, f_2x^2, $\ldots$, f_nx^n и g_0, g_1x, g_2x^2, $\ldots$, g_mx^m соответственно. ■

Задача 18.

а) Докажите свойства в теореме 1.17.1 и её следствиях.

б) Найдите необходимые и достаточные условия выполнения равенства $\deg(f + g) = \max\{\deg(f), \deg(g)\}$.

1.18. Деление многочленов с остатком. Схема Горнера

Для операции умножения многочленов $P[x]$ над полем P обратная операция — деление — не существует. Кольцо $P[x]$ напоминает множество целых

чисел Z. Эта аналогия проявляется в том, что для многочленов, как и для целых чисел, имеют место понятия деления нацело, деления с остатком, делителя, наибольшего общего делителя и др.

Приведём обоснование *алгоритма деления с остатком.*

Теорема 1.18.1 (Обоснование алгоритма деления с остатком).

В кольце многочленов $P[x]$ для любых многочленов $f(x)$, $g(x)$ $(g(x) \neq 0)$ существует единственная пара многочленов $h(x), r(x)$ такая, что

$$f(x) = g(x)h(x) + r(x), \tag{40}$$

$$\text{где } \deg(r(x)) < \deg(g(x)). \tag{41}$$

Доказательство. I. Сначала докажем существование методом математической индукции по степени многочлена $f(x)$, считая $g(x)$ фиксированным.

Запишем многочлены по убыванию степеней

$f(x) = \sum\limits_{k=0}^{n} a_k x^{n-k} = a_0 x^n + a_1 x^{n-1} + \ldots + a_1 x + a_n, \quad deg(f) = n,$

$g(x) = \sum\limits_{k=0}^{m} b_k x^{m-k} = b_0 x^m + b_1 x^{m-1} + \ldots + b_1 x + b_m, \quad deg(g) = m \geqslant 0.$

1) Для $n = -\infty$ имеем $f = 0$ и $f = g \cdot 0 + f$, т. е. при $h = 0$, $r = f = 0$ выполнено равенство (40). Неравенство (41) верно, т. к. $\deg(r) = \deg(f) = -\infty$, $\deg(g) = m \geqslant 0$.

2) Пусть $n = 0$. Для $m > 0$ имеем $f = g \cdot 0 + f$, т. е. при $h = 0$, $r = f$ выполнено равенство (40). Неравенство (41) верно, т. к. $\deg(r) = \deg(f) = n = 0$, $\deg(g) = m > 0$.

Для $m = 0$ имеем $f(x) = a_0 \neq 0$, $g(x) = b_0 \neq 0$. Отсюда $a_0 = b_0 \dfrac{a_0}{b_0} + 0$, т. е. $f = g\,h + r$, где $h = \dfrac{a_0}{b_0}$, $r = 0$ и $\deg(r) = -\infty < \deg(g) = 0$. Следовательно, выполнены условия (40)–(41).

3) Предположим, что утверждение верно при всех степенях $n < s$ и докажем утверждение при $n = s$. Пусть многочлены имеют вид:

$f(x) = a_0 x^s + a_1 x^{s-1} + \ldots + a_{s-1} x + a_s, \quad deg(f) = s,$

$g(x) = b_0 x^m + b_1 x^{m-1} + \ldots + b_{m-1} x + b_m, \quad 0 \leqslant deg(g) = m.$

Для $s < m$ имеем $f = g \cdot 0 + f$, т. е. при $h = 0$, $r = f$ выполнено равенство (40). Неравенство (41) верно, т. к. $\deg(r) = \deg(f) = s < m = \deg(g)$.

Пусть теперь $s \geqslant m$.

Проведём первый шаг деления многочлена $f(x)$ на многочлен $g(x)$: найдём первое слагаемое частного — одночлен $h_1(x) = \dfrac{a_0}{b_0}x^{s-m}$, вычтем из $f(x)$ произведение $g(x)h_1(x)$, получим многочлен $f_1(x) = f(x) - g(x)h_1(x)$. Далее преобразуем $f_1(x) = (a_0x^s + a_1x^{s-1} + \ldots) - (b_0x^m + b_1x^{m-1} + \ldots)\dfrac{a_0}{b_0}x^{s-m} = (a_0x^s + a_1x^{s-1} + \ldots) - (a_0x^s + \dfrac{a_0b_1}{b_0}x^{s-1} + \ldots) = (a_1 - \dfrac{a_0b_1}{b_0})x^{s-1} + \ldots$. Для $f_1(x)$ имеем а) $f_1(x) = 0$ или б) $f_1(x) \neq 0$ и $\deg(f_1) \leqslant s - 1 < s$. В первом случае из равенства $f_1(x) = f(x) - g(x)h_1(x)$ получим $f(x) = g(x)h_1(x) + 0$ и тогда выполнены условия (40)–(41) при $h(x) = h_1(x)$, $r(x) = 0$. Во втором случае для многочлена $f_1(x)$ степени меньшей s по предположению индукции найдутся многочлены $h_2(x)$ и $r(x)$, для которых $f_1(x) = g(x)h_2(x) + r(x)$ и $\deg(r(x)) < \deg(g(x))$. Тогда из равенств $f_1(x) = f(x) - g(x)h_1(x)$ и $f_1(x) = g(x)h_2(x) + r(x)$ следует, что $f(x) = f_1(x) + g(x)h_1(x) = g(x)h_2(x) + r(x) + g(x)h_1(x) = g(x)(h_1(x) + h_2(x)) + r(x)$, и выполнены условия (40)–(41) при $h(x) = h_1(x) + h_2(x)$. Итак, утверждение доказано при $n = s$ и из метода математической индукции следует существование многочленов $h(x)$, $r(x)$, удовлетворяющих условиям (40)–(41).

II. Докажем единственность многочленов $h(x)$, $r(x)$. Пусть существует ещё одна пара многочленов $h_1(x), r_1(x) \in P[x]$, для которых $f(x) = g(x)h_1(x) + r_1(x)$, где $\deg(r_1(x)) < \deg(g(x))$. Вычтем из равенства $f(x) = g(x)h(x) + r(x)$ почленно равенство $f(x) = g(x)h_1(x) + r_1(x)$. После преобразования получим равенства $0 = g(x)(h(x) - h_1(x)) + r(x) - r_1(x)$ и $g(x)(h_1(x) - h(x)) = r(x) - r_1(x)$. Докажем, что $r(x) - r_1(x) = 0$. Допустим противное: $r(x) - r_1(x) \neq 0$. Тогда $h_1(x) - h(x) \neq 0$. Оценим степени в левой и правой частях равенства $g(x)(h_1(x) - h(x)) = r(x) - r_1(x)$. В левой части имеем $deg(g(x)(h_1(x) - h(x))) = deg(g(x)) + deg(h_1(x) - h(x)) \geqslant deg(g(x))$. А в правой части в силу $deg(r(x)) < deg(g(x))$, $deg(-r_1(x)) = deg(r_1(x)) < deg(g(x))$ и неравенства (38) имеем $\deg(r(x) - r_1(x)) \leqslant \max(\deg(r(x)), \deg(-r_1(x))) < \deg(g(x))$. То есть степени левой и правой частей различны, что невозможно. Следовательно, $r(x) - r_1(x) = 0$ и $g(x)(h_1(x) - h(x)) = 0$, а поскольку в кольце $P[x]$ нет делителей нуля, то $h_1(x) - h(x) = 0$. Следовательно, $r(x) = r_1(x)$ и $h(x) = h_1(x)$.

Теорема доказана. ■

Замечание. В практическом применении алгоритма деления многочлена $f(x) = a_0x^s + \ldots$ на многочлен $g(x) = b_0x^m + \ldots$ на первом шаге находим многочлен $f_1(x) = f(x) - g(x)h_1(x)$, где $h_1(x) = \dfrac{a_0}{b_0}x^{s-m}$, $deg(f_1) = s_1 < s$. Далее проверяем $s_1 < m$ (сюда входит рассмотренный в доказательстве случай $f_1(x) = 0$) или $s_1 \geqslant m$. Для $s_1 < m$ выразим $f(x) = g(x)h_1(x) + f_1(x)$, тогда выполнены (40)–(41) при $h(x) = h_1(x), r(x) = f_1(x)$, и алгоритм завершен. В противном случае, т.е. для $s_1 \geqslant m$, проводим второй шаг алгоритма (аналогичный первому шагу): делим многочлен $f_1(x) = c_0x^{s_1} + \ldots$ на многочлен $g(x)$. Находим $f_2(x) = f(x) - g(x)h_2(x)$, где $h_2(x) = \dfrac{c_0}{b_0}x^{s_1-m}$, $deg(f_2) = s_2 < s_1$. Затем проверяем $s_2 < m$ или $s_2 \geqslant m$ и завершаем или продолжаем выполнение алгоритма. ■

В связи с теоремой 1.18.1 дадим следующее определение.

Определение 1. *Если для многочленов* $f(x), g(x) \in P[x]\,(g(x) \neq 0)$ *существует пара многочленов* $h(x), r(x) \in P[x]$ *такая, что выполнены условия*

$$f(x) = g(x)h(x) + r(x), \tag{40}$$

$$\deg(r(x)) < \deg(g(x)), \tag{41}$$

то многочлены $h(x)$ *и* $r(x)$ *соответственно называют частным и остатком от деления многочлена* $f(x)$ *на многочлен* $g(x)$. ■

Из общего алгоритма деления многочленов для нахождения частного и остатка от деления следует *схема* (*метод*) *Горнера* для деления многочлена на многочлен первой степени (линейный двучлен $x - c$).

Пусть заданы многочлены из $P[x]:\ f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ степени $n \geqslant 1$ и $g(x) = x - c$, $c \in P$.

По теореме 1.18.1 существует единственная пара многочленов $h(x)$, $r(x)$ таких, что выполнены условия $f(x) = (x-c)h(x) + r(x)$ и $deg(r(x)) < deg(x - c) = 1$. Отсюда следует, что $r(x)$ — константа и $deg(h(x)) = n - 1$.

Обозначим $h(x) = b_0x^{n-1} + b_1x^{n-2} + \ldots + b_{n-2}x + b_{n-1}$, $r(x) = r_0$ и выведем рекуррентные формулы вычисления b_j, $j = \overline{0, n-1}$ и r_0.

1) Первый член частного от деления $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ на $g(x) = x - c$ равен a_0x^{n-1} и он совпадает с $h_1(x) = b_0x^{n-1}$ — первым членом многочлена $h(x)$. Тогда $b_0 = a_0$. Выполним 1-й шаг деления $f(x)$ на $g(x)$

$$
\begin{array}{r|l}
- \begin{array}{l} b_0x^n + a_1x^{n-1} + a_2x^{n-2} \ldots + a_{n-1}x + a_n \\ b_0x^n - cb_0x^{n-1} \end{array} & \begin{array}{l} x - c \\ \hline b_0x^{n-1} \end{array} \\
\hline
(cb_0 + a_1)x^{n-1} + a_2x^{n-2} \ldots + a_{n-1}x + a_n &
\end{array}
$$

и полученную разность $f(x) - h_1(x)g(x) = (cb_0 + a_1)x^{n-1} + a_2x^{n-2} \ldots + a_{n-1}x + a_n$ обозначим через $f_1(x)$.

2) Второй член частного от деления $f(x)$ на $g(x)$ совпадает с первым членом частного от деления $f_1(x)$ на $g(x)$, он равен $(cb_0 + a_1)x^{n-2}$ и совпадает с $h_2(x) = b_1x^{n-2}$ — вторым членом многочлена $h(x)$. Тогда $b_1 = cb_0 + a_1$.

Выполним 2-й шаг деления $f(x)$ на $g(x)$ (1-й шаг деления $f_1(x)$ на $g(x)$)

$$
\begin{array}{r|l}
- \begin{array}{l} b_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n \\ b_1x^{n-1} - cb_1x^{n-2} \end{array} & \begin{array}{l} x - c \\ \hline b_1x^{n-2} \end{array} \\
\hline
(cb_1 + a_2)x^{n-2} + \ldots + a_{n-1}x + a_n &
\end{array}
$$

и полученную разность $f_1(x) - h_2(x)g(x) = (cb_1 + a_2)x^{n-2} + \ldots + a_{n-1}x + a_n$ обозначим через $f_2(x)$.

Далее находим

3) $b_2 = cb_1 + a_2$, $f_3(x) = f_2(x) - h_3(x)g(x) = (cb_2 + a_3)x^{n-3} + \ldots + a_{n-1}x + a_n$,

$\ldots$,

$n-1$) $b_{n-2} = cb_{n-3} + a_{n-2}$, $f_{n-1}(x) = f_{n-2}(x) - h_{n-1}(x)g(x) =$

$= (cb_{n-2} + a_{n-1}x + a_n$,

n) $b_{n-1} = cb_{n-2} + a_{n-1}$, $f_n(x) = f_{n-1}(x) - h_n(x)g(x) = cb_{n-1} + a_n$.

Здесь $f_n(x)$ — остаток от деления $f(x)$ на $g(x)$. Тогда

$n+1$) $r_0 = cb_{n-1} + a_n$.

Отсюда получим рекуррентные формулы для последовательного нахождения коэффициентов b_j, $j = \overline{0, n-1}$ и r_0 :

$$
\begin{aligned}
&b_0 = a_0, \\
&b_k = cb_{k-1} + a_k,\ k = \overline{1, n-1}, \\
&r_0 = cb_{n-1} + a_n.
\end{aligned}
\tag{42}
$$

Здесь $b_0 = a_0$ *и каждый следующий коэффициент* b_k *получается умножением* c *на предыдущий коэффициент* b_{k-1} *и прибавлением коэффициента* a_k. Остаток от деления r_0 получается по этому же закону: $r_0 = cb_{n-1} + a_n$.

При вычислении удобно заполнить таблицу (43). В верхнюю строку таблицы располагаем все коэффициенты a_j, $j = \overline{0, n}$ многочлена $f(x)$. Слева сбоку от нижней строки записываем значение c из многочлена $g(x) = x - c$. В ниж-

нюю строку под коэффициентом a_0 записываем равное ему значение b_0, затем под коэффициентом a_1 записываем сумму $cb_0 + a_1$ и, вычислив его, получаем b_1, и так далее заполняем всю строку.

$$\begin{array}{c|c|c|c|c|c|c|}
 & a_0 & a_1 & a_2 & \dots & a_{n-1} & a_n \\
\hline
c & b_0 & cb_0 + a_1 = b_1 & cb_1 + a_2 = b_2 & \dots & cb_{n-2} + a_{n-1} = b_{n-1} & cb_{n-1} + a_n = r_0 \\
\hline
\end{array}. \qquad (43)$$

Задача 19.

Докажите, что если для многочленов $f(x), g(x), h(x), r(x) \in P[x]$ выполнены условия $f(x) = g(x)h(x) + r(x)$ и $deg(r(x)) < deg(g(x))$, то $deg(f(x)) = deg(g(x)) + deg(h(x))$.

1.19. Делители многочленов и их свойства

Определение 1. *Если для многочленов $f(x), g(x) \in P[x]$ существует многочлен $h(x) \in P[x]$ такой, что $f(x) = g(x)h(x)$, то говорят, что $f(x)$ делится на $g(x)$ и $g(x)$ делит $f(x)$, при этом $g(x)$ называют делителем $f(x)$, а $f(x)$ называют кратным $g(x)$. Это записывается следующим образом: $f(x) \vdots g(x)$ и $g(x) \,|\, f(x)$.*

Замечания. 1. Из теоремы 1.18.1 следует, что если для $f(x), g(x) \in P[x]$, где $g(x) \neq 0$, существует $h(x) \in P[x]$ такой, что $f(x) = g(x)h(x)$, то многочлен $h(x)$ единственный.

2. При $f(x) \neq 0$ и $g(x) = 0$ многочлена $h(x)$ такого, что $f(x) = 0h(x)$ не существует.

3. При $f(x) = 0$ и $g(x) = 0$ для любого многочлена $h(x)$ верно $0 = 0h(x)$, т.е. 0 является делителем 0, но результат деления не определён.

Теорема 1.19.1 (Простейшие свойства делителей многочленов).

Для многочленов из $P[x]$ над полем P справедливы следующие свойства.

1. *Если $f(x) \vdots g(x)$, а $g(x) \vdots h(x)$, то $f(x) \vdots h(x)$.*
2. *Если $f(x) \vdots \varphi(x)$ и $g(x) \vdots \varphi(x)$, то $(f(x) + g(x)) \vdots \varphi(x)$.*
3. *Если $f(x)$ делится на $\varphi(x)$, то произведение $f(x)$ на любой многочлен $g(x)$ также делится на $\varphi(x)$.*

4. *Если каждый из многочленов* $f_j(x)$, $j = \overline{1..k}$ *делится на* $\varphi(x)$, *то на* $\varphi(x)$ *делится и многочлен* $f_1(x)g_1(x) + f_2(x)g_2(x) + \ldots + f_k(x)g_k(x)$, *где* $g_j(x)$, $j = \overline{1..k}$ — *произвольные многочлены.*

5. *Всякий многочлен делится на любой многочлен нулевой степени.*

6. *Если* $f(x)$ *делится на* $\varphi(x)$, *то* $f(x)$ *делится и на* $c\varphi(x)$, *где* $c \in P$ — *любое ненулевое число.*

7. *Многочлены* $cf(x)$, *где* $c \in P$, $c \neq 0$, *и только они будут делителями многочлена* $f(x)$, *имеющими такую же степень, что и* $f(x)$.

8. *Многочлены* $f(x)$ *и* $g(x)$ *одновременно делятся друг на друга тогда и только тогда, когда* $g(x) = cf(x)$, *где* $c \in P$, $c \neq 0$.

9. *Всякий делитель одного из двух многочленов* $f(x)$, $cf(x)$, *где* $c \in P, c \neq 0$, *будет являться делителем и для другого многочлена.*

10. *Если* $f(x)$ *делится на* $g(x)$, *то* $deg(f(x)) \geqslant deg(g(x))$.

Доказательство.

1. $f(x) \vdots g(x) \Rightarrow \exists u(x) \in P[x] \mid f(x) = g(x)u(x)$ и
   $g(x) \vdots h(x) \Rightarrow \exists v(x) \in P[x] \mid g(x) = h(x)v(x)$.
   Тогда $f(x) = g(x)u(x) = h(x)v(x)u(x) \vdots h(x)$.
2. $f(x) \vdots \varphi(x) \Rightarrow \exists u(x) \in P[x] \mid f(x) = \varphi(x)u(x)$ и
   $g(x) \vdots \varphi(x) \Rightarrow \exists v(x) \in P[x] \mid g(x) = \varphi(x)v(x)$.
   Тогда $f(x) + g(x) = \varphi(x)u(x) + \varphi(x)v(x) = \varphi(x)(u(x) + v(x)) \vdots \varphi(x)$.
3. $f(x) \vdots \varphi(x) \Rightarrow \exists u(x) \in P[x] \mid f(x) = \varphi(x)u(x)$.
   Тогда $f(x)g(x) = \varphi(x)(u(x)g(x)) \vdots \varphi(x)$.
4. Докажем методом математической индукции по k.

1) При $k = 2$ $f_1(x), f_2(x)$ делятся на $\varphi(x)$. Тогда по свойству 3 $f_1(x)g_1(x)$, $f_2(x)g_2(x)$ делятся на $\varphi(x)$ и по свойству 2 $f_1(x)g_1(x) + f_2(x)g_2(x)$ делится на $\varphi(x)$.

2) Предположим, что утверждение верно при $k = q$ и докажем, что утверждение верно при $k = q + 1$.

Пусть многочлены $f_j(x)$, $j = \overline{1..q+1}$ делятся на $\varphi(x)$. Рассмотрим сумму $\sum\limits_{j=1}^{q+1} f_j(x)g_j(x) = \Big(\sum\limits_{j=1}^{q} f_j(x)g_j(x)\Big) + f_{q+1}(x)g_{q+1}(x)$. На $\varphi(x)$ делится а) первое слагаемое по предположению индукции, б) второе слагаемое по свойству 3, в) их сумма по свойству 2.

5. Если $f(x)$ многочлен и число $c \neq 0$, то $c^{-1}f(x)$ также многочлен и при этом $f(x) = c(c^{-1}f(x)) \Rightarrow f(x) \vdots c$.

6. $f(x) \vdots \varphi(x) \Rightarrow \exists u(x) \in P[x] \mid f(x) = \varphi(x)u(x)$.

Тогда $f(x) = (c\varphi(x))(c^{-1}u(x)) \vdots c\varphi(x)$.

7. а) $f(x) = c^{-1}(cf(x)) \Rightarrow f(x) \vdots cf(x)$ и $deg(cf(x)) = deg(f(x))$.

б) Пусть $f(x) \vdots g(x)$ и $deg(f(x)) = deg(g(x))$.

Тогда $\exists u(x) \in P[x] \mid f(x) = g(x)h(x)$. Отсюда $deg(f(x)) = deg(g(x)) + deg(h(x)) = deg(f(x)) + deg(h(x))$ и $deg(h(x)) = 0$. Следовательно, $h(x) = d \in P$, $d \neq 0$, $f(x) = g(x)d$ и $g(x) = d^{-1}f(x)$.

8. а) $f(x) \vdots g(x) \Rightarrow \exists u(x) \in P[x] \mid f(x) = g(x)u(x)$ и

$g(x) \vdots f(x) \Rightarrow \exists v(x) \in P[x] \mid g(x) = f(x)v(x)$.

Отсюда следует, что 1) $f = 0 \Leftrightarrow g = 0$ и 2) $f \neq 0 \Leftrightarrow g \neq 0$. В первом случае равенство $g(x) = cf(x)$ верно. Во втором случае из равенств $f(x) = g(x)u(x)$ и $g(x) = f(x)v(x)$ следует равенство $g(x) = g(x)u(x)v(x)$. Тогда $u(x)v(x) = 1 \Rightarrow deg(u(x)) = deg(v(x)) = 0$. Следовательно, $v(x) = c \in P$, $c \neq 0$ и $g(x) = f(x)v(x) = cf(x)$, т.е. равенство $g(x) = cf(x)$ верно.

б) Если $g(x) = cf(x)$, $c \in P$, $c \neq 0$, то $f(x) = c^{-1}g(x)$. Тогда $g(x) \vdots f(x)$ и $f(x) \vdots g(x)$. ■

Задача 20.

1) Докажите свойства 9–10 в теореме 1.19.1.

2) Пусть $f(x), g(x) \in \mathbb{R}[x]$, причём $f(x) = x^2 + 2x$, $g(x) = x - b$, $b \in \mathbb{R}$.

Докажите, что многочлен $g(x)$ является делителем многочлена $f(x)$ при $b = 0$ или $b = -2$, а при остальных значениях $b \in \mathbb{R}$ многочлен $f(x)$ не делится на $g(x)$. Найдите все делители многочлена $f(x)$.

1.20. Общий делитель и наибольший общий делитель многочленов. Нахождения наибольшего общего делителя многочленов. Представление наибольшего общего делителя многочленов через эти многочлены

В этом разделе рассматриваются многочлены из $P[x]$ над полем P.

Определение 1. *Пусть заданы многочлены*

$$f_1(x),\ f_2(x),\ \ldots,\ f_k(x). \tag{44}$$

Многочлен $h(x)$ называется общим делителем многочленов (44), если он является делителем каждого из них, т.е. $f_j(x) \vdots h(x)$, $j = \overline{1,k}$.

Обозначение: $h(x) \in$ ОД$(f_1(x), f_2(x), \ldots, f_k(x))$.

Многочлен $d(x)$ называется наибольшим общим делителем заданных многочленов (44), если

1) $d(x)$ — общий делитель этих многочленов,

2) $d(x)$ делится на любой общий делитель $h(x)$ этих многочленов.

Обозначение: $d(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$.

Лемма 1.20.1.

1° *Пусть $d(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$.*

Тогда $cd(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$ для любого числа $c \in P$, $c \neq 0$.

2° *Пусть $d_1(x), d_2(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$.*

Тогда $d_2(x) = cd_1(x)$ для некоторого числа $c \in P$, $c \neq 0$.

Доказательство. 1° Пусть $d(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$.

1) Тогда каждый многочлен $f_j(x) \vdots d(x)$ и из 6-го свойства делителей следует, что каждый многочлен $f_j(x) \vdots cd(x)$ для любого числа $c \in P$, $c \neq 0$, т.е. $cd(x) \in$ ОД$(f_1(x), f_2(x), \ldots, f_k(x))$.

2) Если $h(x)$ общий делитель многочленов (44), то $d(x) \vdots h(x)$. Тогда из 3-го свойства делителей следует, что $cd(x) \vdots h(x)$ для любого числа $c \in P$, $c \neq 0$.

Из 1), 2) получаем $cd(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$ для любого числа $c \in P$, $c \neq 0$.

2° Пусть $d_1(x), d_2(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$. Из определения 1 следует, что $d_1(x)$ и $d_2(x)$ являются общими делителями многочленов (44), при этом $d_1(x) \vdots d_2(x)$ и $d_2(x) \vdots d_1(x)$. Тогда в силу 8-го свойства делителей имеем $d_2(x) = cd_1(x)$ для некоторого числа $c \in P$, $c \neq 0$. ■

Следствие.

Если существует многочлен $d(x)$, являющийся наибольшим общим делителем многочленов (44), то все наибольшие общие делители этих многочленов получаются из формулы $cd(x)$, где $c \in P$, $c \neq 0$. ■

Замечание. При $k = 1$ наибольшим общим делителем многочлена $f_1(x)$ является любой многочлен $cf_1(x)$, где $c \in P$, $c \neq 0$. ■

Покажем, что для двух ненулевых многочленов наибольший общий делитель существует и может быть найден с помощью *алгоритма Евклида* или *алгоритма последовательного деления.*

Теорема 1.20.1. *Для любых двух ненулевых многочленов $f(x), g(x)$ существует наибольший общий делитель.*

Доказательство. Для построение наибольшего общего делителя многочленов $f(x)$, $g(x)$ выполняем *алгоритм Евклида.*

1-й шаг. Делим f на g. Получим $f = gh_1 + r_1$, где $deg(r_1) < deg(g)$. Если $r_1 = 0$, то $f = gh_1$, процесс завершён делением нацело на g. Если $r_1 \neq 0$, то выполняем шаг 2.

2-й шаг. Делим g на r_1. Получим $g = r_1h_2 + r_2$, где $deg(r_2) < deg(r_1)$. Если $r_2 = 0$, то процесс завершён делением нацело на r_1. Если $r_2 \neq 0$, то выполняем шаг 3.

3-й шаг. Делим r_1 на r_2. Получим $r_1 = r_2h_3 + r_3$, где $deg(r_3) < deg(r_2)$. Если $r_2 = 0$, то процесс завершён делением нацело на r_2. Если $r_2 \neq 0$, то выполняем шаг 4 и т. д.

Степени остатков убывают, поэтому процесс оборвётся в тот момент, когда деление выполнится нацело. Пусть это будет $(k+1)$-й шаг и r_k многочлен, на который деление выполнилось нацело. Покажем, что $r_k \in$ НОД(f, g).

Обозначим $r_{-1} = f$, $r_0 = g$ и результаты деления запишем в таблицу.

№ шага	результат деления
1	$r_{-1} = r_0 h_1 + r_1$, $deg(r_1) < deg(r_0)$
2	$r_0 = r_1 h_2 + r_2$, $deg(r_2) < deg(r_1)$
3	$r_1 = r_2 h_3 + r_3$, $deg(r_3) < deg(r_2)$
...	...
$k-1$	$r_{k-3} = r_{k-2} h_{k-1} + r_{k-1}$, $deg(r_{k-1}) < deg(r_{k-2})$
k	$r_{k-2} = r_{k-1} h_k + r_k$, $deg(r_k) < deg(r_{k-1})$
$k+1$	$r_{k-1} = r_k h_{k+1}$

(45)

I. Просмотр таблицы снизу вверх показывает, что $d = r_k$ является делителем r_{k-1}, r_{k-2}, r_{k-3}, ..., r_1, $r_0 = g$, $r_{-1} = f$.

II. Просмотр таблицы сверху вниз показывает, что $r_j = r_{j-2} - r_{j-1} h_j$ для $j = \overline{1,k}$, т. е. все многочлены r_1, r_2, r_3, ..., r_{k-1}, r_k делятся на любой общий делитель многочленов $r_{-1} = f$, $r_0 = g$.

Итак, $d = r_k \in$ НОД(f, g) и теорема доказана. ■

С помощью алгоритма Евклида докажем утверждение о формуле представления наибольшего общего делителя двух многочленов через эти два многочлена.

Теорема 1.20.2. *Если $d(x)$ — наибольший общий делитель многочленов $f(x)$ и $g(x)$, то существуют такие многочлены $u(x)$ и $v(x)$, что*

$$d(x) = f(x)u(x) + g(x)v(x), \tag{46}$$

Если степени многочленов $f(x)$ и $g(x)$ больше нуля, то можно считать при этом, что $deg(u(x)) < deg(g(x))$ и $deg(v(x)) < deg(f(x))$.

Доказательство. I. Пусть $r_k \in$ НОД(f, g). Обозначим $r_{-1} = f$, $r_0 = g$.

Если $k = 0$, то $r_0 = g \in$ НОД(f, g) и $g = f0 + g1$, т.е. (46) выполнено при $u = 0$, $v = 1$.

Если $k = 1$, то $r_1 \in$ НОД(f, g) и $r_1 = r_{-1} - r_0 h_1 = f - g h_1$ т.е. (46) выполнено при $u = 1$, $v = -g$.

Если $k > 0$, то выразим из формулы (45) остатки $r_j = r_{j-2} - r_{j-1} h_j$, $j = \overline{1,k}$. Здесь каждый следующий остаток выражается через два предыдущих. С помощью этих формул будем в порядке возрастания номеров j от 1 до k

последовательно выражать через $r_{-1} = f$, $r_0 = g$ остатки $r_j = fu_j + gv_j$. При $j = k$ получим $r_k = fu_k + gv_k$, т.е. равенство (46) будет выполнено для $d = r_k$ при $u = u_k$, $v = v_k$.

В силу леммы 1.20.1 любой наибольший общий делитель d многочленов f и g выражается через r_k : $d = c \cdot r_k$, $c = const \neq 0$. Тогда $c \cdot r_k = f \cdot (c \cdot u_k) + +g \cdot (c \cdot v_k)$, т.е. $d = f \cdot u + g \cdot v$ при $u = c \cdot u_k$, $v = c \cdot v_k$.

II. Для доказательства второго утверждения теоремы предположим, что равенство (46) выполнено, но, например, $deg(u(x)) \geqslant deg(g(x))$. Разделим $u(x)$ на $g(x)$:

$$u(x) = g(x)w(x) + u_1(x),$$

где $deg(u_1(x)) < deg(g(x))$, и подставим это выражение в (46).

Получим $d(x) = f(x)g(x)w(x) + f(x)u_1(x) + g(x)v(x)$

или $d(x) = f(x)u_1(x) + g(x)(f(x)w(x) + v(x))$.

Отсюда следует другое представление для $d(x)$:

$$d(x) = f(x)u_1(x) + g(x)v_1(x), \quad \text{где} \quad v_1(x) = f(x)w(x) + v(x). \tag{47}$$

Здесь уже выполнено одно требуемое неравенство $deg(u_1(x)) < deg(g(x))$. Докажем второе неравенство $deg(v_1(x)) < deg(f(x))$ методом «от протвного». Предположим противное: $deg(v_1(x)) \geqslant deg(f(x))$. Покажем, что из этого предположения следует вывод: $deg(d) \geqslant deg(f) + deg(g)$, приводящий к противоречию.

Оценим в (47) степень произведения многочленов $g(x)v_1(x)$:

$$deg(g(x)v_1(x)) = deg(g(x)) + deg(v_1(x)) \geqslant deg((g(x)) + deg(f(x)).$$

Оценим в (47) степень произведения многочленов $f(x)u_1(x)$:

$$deg(f(x)u_1(x)) = deg((f(x)) + deg(u_1(x)) < deg(f(x)) + deg(g(x)).$$

Из двух неравенств получаем $deg(g(x)v_1(x)) > deg(f(x)u_1(x))$.

Очевидно, что степень суммы двух многочленов разных степеней равна их максимальной степени. Следовательно, $deg(f(x)u_1(x) + g(x)v_1(x)) = = deg(g(x)v_1(x))$. Тогда из (47) получаем $deg(d(x)) = deg(f(x)u_1(x) + g(x) \times \times v_1(x)) = deg(g(x)v_1(x)) \geqslant deg(g(x)) + deg(f(x))$, т.е. $deg(d(x)) \geqslant deg(g(x)) + + deg(f(x))$.

По условию $d(x) \in$ НОД$(f(x), g(x))$, поэтому $f(x)$ делится на $d(x)$. Следовательно, $deg(f(x)) \geqslant deg(d(x))$ и, учитывая неравенство $deg(d(x)) \geqslant \geqslant deg(g(x)) + deg(f(x))$, получаем $deg(f(x)) \geqslant deg(g(x)) + deg(f(x))$. Отсюда

$deg(g(x)) = 0$, что противоречит условию $deg(g(x)) > 0$.

Полученное противоречие доказывает, что для многочлена $v_1(x)$ в представлении (47) выполнено неравенство $deg(v_1(x)) < deg(f(x))$. И ранее было получено неравенство $deg(u_1(x)) < deg(g(x))$.

Теорема доказана. ■

Существование наибольшего общего делителя для системы многочленов (44) при произвольном $k \geqslant 3$ и способ его вычисления получается из следующей теоремы.

Теорема 1.20.3.

Если заданы любые ненулевые многочлены $f_1(x), f_2(x), \ldots, f_k(x)$ *при произвольном* $k \in \mathbb{N}$, $k \geqslant 2$, *то*

справедливо равенство

$$НОД(f_1(x), f_2(x), \ldots, f_k(x)) = НОД(НОД(f_1(x), \ldots, f_{k-1}(x)), f_k(x)). \qquad (48)$$

Согласно этой теореме для нахождения НОД$(f_1(x), f_2(x), \ldots, f_k(x))$ сначала берём $d_1(x) = f_1(x)$, далее находим $d_2(x) \in$ НОД$(d_1(x), f_2(x))$, после этого определяем $d_j(x) \in$ НОД$(d_{j-1}(x), f_j(x))$, $j = \overline{3, k}$.

Ответом является множество многочленов $cd_k(x)$, где $c = const \neq 0$.

Доказательство теоремы 1.20.3 проведём методом математической индукции.

При $k = 2$ утверждение следует из теоремы 1.20.2.

Далее доказываем, что из предположения о справедливости утверждения при $k = m$ следует его справедливость при $k = m + 1$.

По предположению индукции существует многочлен $d_m(x)$, который является наибольшим общим делителем многочленов $f_1(x), f_2(x), \ldots, f_m(x)$. По теореме 1.20.1 существует многочлен $d_{m+1}(x)$, который является наибольшим общим делителем многочленов $d_m(x), f_{m+1}(x)$. Покажем, что $d_{m+1}(x)$ является наибольшим общим делителем многочленов $f_1(x), f_2(x), \ldots, f_{m+1}(x)$.

Во-первых, все $f_j(x)$, $j = \overline{1, m}$ делятся на $d_m(x)$, а $d_m(x)$ и $f_{m+1}(x)$ делятся на $d_{m+1}(x)$, поэтому все $f_j(x)$, $j = \overline{1, m+1}$ делятся на $d_{m+1}(x)$. Во-вторых, любой общий делитель $h(x)$ многочленов $f_1(x), f_2(x), \ldots, f_{m+1}(x)$ является делителем $d_m(x)$ и $f_{m+1}(x)$, и поэтому делителем $d_{m+1}(x)$. Следовательно, $d_{m+1}(x)$ является наибольшим общим делителем всех многочленов

$f_j(x)$, $j = \overline{1, m+1}$.

В силу леммы 1.20.1 для любой $c_1 = const \neq 0$ НОД$(c_1 d_m(x), f_{m+1}(x)) = \{c_2 d_{m+1}(x) |\, c_2 = const \neq 0\}$ и НОД$(f_1(x), f_2(x), \ldots, f_{m+1}(x)) = \{c_3 d_{m+1}(x) |\, c_3 = const \neq 0$ т.е. для $k = m+1$ справедливо равенство (48). ■

Справедлив аналог теоремы 1.20.2 о формуле представления наибольшего общего делителя нескольких многочленов через эти многочлены.

Теорема 1.20.4. *Если $d(x)$ — наибольший общий делитель многочленов $f_1(x), f_2(x), \ldots, f_k(x)$ при произвольном $k \in \mathbb{N}$, $k \geqslant 2$, то существуют такие многочлены $u_1(x), u_2(x), \ldots, u_k(x)$, что*

$$d(x) = f_1(x)u_1(x) + f_2(x)u_2(x) + \ldots + f_k(x)u_k(x). \qquad (49)$$

Доказательство теоремы 1.20.4 проведём методом математической индукции.

При $k = 2$ утверждение следует из теоремы 1.20.2.

Далее докажем, что из предположения о справедливости утверждения для $D_m(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_m(x))$ следует его справедливость для $D_{m+1}(x) \in$ НОД$(D_m(x), f_{m+1}(x)) =$ НОД$(f_1(x), f_2(x), \ldots, f_{m+1}(x))$ ($D_{m+1}(x)$ строится как в теореме 1.20.3).

По предположению индукции для многочлена $D_m(x)$ существуют многочлены $w_j(x)$, $j = \overline{1, m}$ такие, что $D_m(x) = \sum\limits_{j=1}^{m} f_j(x) w_j(x)$. По теореме 1.20.2 существуют такие многочлены $u(x)$ и $v(x)$, что $D_{m+1}(x) = D_m(x)u(x) + f_{m+1}(x)v(x)$. Тогда $D_{m+1}(x) = \sum\limits_{j=1}^{m} f_j(x)(w_j(x)u(x)) + f_{m+1}(x)v(x)$. Обозначив $u_j(x) = w_j(x)u(x)$, $j = \overline{1, m}$ и $u_{m+1} = v(x)$, получим для $k = m+1$ и $D_{m+1}(x)$ представление (49): $D_{m+1}(x) = \sum\limits_{j=1}^{m+1} f_j(x) u_j(x)$. Поскольку любой наибольший общий делитель $d(x)$ многочленов $f_1(x), f_2(x), \ldots, f_{m+1}(x)$ равен $cD_{m+1}(x)$ для некоторого числа $c \in P$, $c \neq 0$, тогда для $d(x)$ справедливо представление $d(x) = \sum\limits_{j=1}^{m+1} f_j(x)(cu_j(x))$. Теорема доказана. ■

Задача 21.

1) Докажите свойства 9–10 в теореме 1.19.1.

2) Пусть $f(x), g(x) \in \mathbb{R}[x]$, причём $f(x) = x^2 + 2x$, $g(x) = x - b$, $b \in \mathbb{R}$.

Докажите, что многочлен $g(x)$ является делителем многочлена $f(x)$ при $b = 0$ или $b = -2$, а при остальных значениях $b \in \mathbb{R}$ многочлен $f(x)$ не делится на $g(x)$. Найдите все делители многочлена $f(x)$.

1.21. Взаимная простота многочленов. Свойства многочленов, связанные со взаимной простотой

В этом разделе рассматриваются многочлены из $P[x]$ над полем P.

Каждый многочлен делится на любой многочлен нулевой степени

$$d(x) = c,\ c \in P,\ c \neq 0. \tag{50}$$

Поэтому любой многочлен нулевой степени является общим делителем произвольных многочленов.

Определение 1. *Пусть заданы многочлены*

$$f_1(x),\ f_2(x),\ \ldots,\ f_k(x). \tag{51}$$

Многочлены (51) называются взаимно простыми, если их общими делителями являются только многочлены нулевой степени.

Пример 1. Многочлены $f_1(x) = x^2 - 4x$ и $f_2(x) = x^2 - 16$ делятся на $x - 4$ и поэтому они не являются взаимно простыми.

Пример 2. Многочлен $f_1(x) = x^2 + 2x$ делится только на $c_1 \cdot 1$, $c_2 x$, $c_3(x+2)$, $c_4(x^2+2x)$, многочлен $f_2(x) = x^2 - 2x + 1$ делятся только на $c_5 \cdot 1$, $c_6(x-1)$, $c_7(x-1)^2$ и поэтому $f_1(x)$ и $f_2(x)$ являются взаимно простыми.

Пример 3. Возьмём многочлены $f_1(x) = x(x+1)$, $f_2(x) = (x+1)(x+2)$, $f_3(x) = x(x+2)$. Любые два из них не являются взаимно простыми. Действительно, $f_1(x), f_2(x)$ делятся на $x+1$, $f_2(x), f_3(x)$ делятся на $x+2$, $f_1(x), f_3(x)$ делятся на x. Многочлены $f_1(x), f_2(x), f_3(x)$ являются взаимно простыми (докажите!).

Теорема 1.21.1. *(Критерии взаимной простоты)*

Для многочленов (51) следующие условия эквивалентны:

(а) *многочлены взаимно просты;*

(б) *многочлен 1 является наибольшим общим делителем этих многочленов;*

(в) *существуют такие многочлены $u_j(x)$, $j = \overline{1,k}$, что $\sum\limits_{j=1}^{k} f_j(x)u_j(x) = 1$.*

Доказательство.

(а) $\Rightarrow$ (б). Пусть выполнено (а), тогда их общими делителями являются только многочлены нулевой степени. Поэтому наибольшим общим делителем многочленов (51) является любой многочлен $d(x) = c$, $c \in P$, $c \neq 0$, и, в частности, многочлен $d(x) = 1$.

(б) $\Rightarrow$ (в). Если выполнено (б), то из теоремы 1.20.3 следует (в).

(в) $\Rightarrow$ (а). Пусть выполнено (в) и многочлен $d(x)$ является общим делителем многочленов (51). Тогда $(\sum\limits_{j=1}^{k} f_j(x)u_j(x)) \vdots d(x) \Rightarrow 1 \vdots d(x) \Rightarrow deg(1) \geqslant$ $\geqslant deg(d(x)) \Rightarrow deg(d(x)) = 0 \Rightarrow d(x) = c,\ c \in P, c \neq 0 \Rightarrow$ (а).

Итак, (а) $\Leftrightarrow$ (б) $\Leftrightarrow$ (в). ■

Теорема 1.21.2 (Свойства, связанные со взаимной простотой).

1. *Если многочлен $f(x)$ взаимно прост с каждым из многочленов $g_1(x)$ и $g_2(x)$, то $f(x)$ взаимно прост и с их произведением $g_1(x)g_2(x)$.*

2. *Если произведение многочленов $f_1(x)f_2(x)$ делится на многочлен $g(x)$, но $f_1(x)$ и $g(x)$ взаимно просты, то $f_2(x)$ делится на $g(x)$.*

3. *Если многочлен $f(x)$ делится на каждый из многочленов $g_1(x)$ и $g_2(x)$, которые между собой взаимно просты, то $f(x)$ делится и на их произведение $g_1(x)g_2(x)$.*

Доказательство.

1. Многочлены $f(x)$ и $g_1(x)$ взаимно просты $\Rightarrow \exists u(x), v(x) \in P[x]$ такие, что $f(x)u(x) + g_1(x)v(x) = 1$. Умножим обе части равенства на $g_2(x)$, и получим $f(x)(u(x)g_2(x)) + (g_1(x)g_2(x))v(x) = g_2(x)$. Отсюда следует, что любой общий делитель $d(x)$ многочленов $f(x)$ и $g_1(x)g_2(x)$ является делителем $g_2(x)$. Тогда $d(x)$ является общим делителем многочленов $f(x)$ и $g_2(x)$, а они взаимно просты по условию, и поэтому $d(x) = c$, $c \in P$, $c \neq 0$. Итак, только многочлены нулевой степени являются общими делителями многочленов $f(x)$ и $g_1(x)g_2(x)$, т.е. они взаимно просты.

2. Многочлены $f_1(x)$ и $g(x)$ взаимно просты $\Rightarrow \exists u(x), v(x) \in P[x]$ такие, что $f_1(x)u(x)+g(x)v(x) = 1$. Умножим обе части равенства на $f_2(x)$, и получим $(f_1(x)f_2(x))u(x) + g(x)(v(x)f_2(x)) = f_2(x)$. Оба слагаемых левой части делятся на $g(x)$, а тогда на него делится и $f_2(x)$.

3. Многочлен $f(x)$ делится на многочлен $g_1(x)$, тогда $f(x) = g_1(x)h_1(x)$. Многочлен $f(x)$ делится на многочлен $g_2(x)$, следовательно, произведение $g_1(x)h_1(x)$ делится на $g_2(x)$, при этом $g_1(x)$ и $g_2(x)$ между собой взаимно просты. Тогда из свойства 2 следует, что $h_1(x)$ делится на $g_2(x)$, т.е. $h_1(x) = g_2(x)h_2(x)$. Получим $f(x) = g_1(x)h_1(x) = g_1(x)g_2(x)h_2(x)$, откуда следует, что $f(x)$ делится на произведение $g_1(x)g_2(x)$. ■

Следствия. 1а. *Если многочлены $f(x)$ и $g(x)$ взаимно просты, то взаимно просты многочлен $f(x)$ и любая степень многочлена $g^n(x)$, где $n \in \mathbb{N}$.*

1б. *Если многочлены $f(x)$ и $g(x)$ взаимно просты, то взаимно просты любые их степени $f^m(x)$ и $g^n(x)$, где $m, n \in \mathbb{N}$.*

Доказательство. 1а. Докажем методом математической индукции по n.

1) При $n = 1$ утверждение верно.

2) Докажем, что из предположения о справедливости утверждения для $n = k$, следует его справедливость при $n = k + 1$.

Действительно, $f(x)$ и $g(x)$ взаимно просты по условию, $f(x)$ и $g^k(x)$ взаимно просты по предположению индукции. Тогда из свойства 1 следует, что $f(x)$ взаимно прост с произведением $g^k(x)$ и $g(x)$, т.е. $f(x)$ взаимно прост с $g^{k+1}(x)$.

1б. Докажем методом математической индукции по m следующее утверждение: *для каждого $m \in \mathbb{N}$ степень $f^m(x)$ взаимно проста со степенью $g^n(x)$ при всех $n \in \mathbb{N}$.*

1) При $m = 1$ утверждение верно в силу следствия 1а.

2) Пусть утверждение верно при $m = k$, т.е. степень $f^k(x)$ взаимно проста со степенью $g^n(x)$ при всех $n \in \mathbb{N}$.

Представим степень $f^{k+1}(x)$ в виде произведения $f^{k+1}(x) = f^k(x)f(x)$. По предположению индукции степень $f^k(x)$ взаимно проста со степенью $g^n(x)$ при всех $n \in \mathbb{N}$ и из 1) $f(x)$ взаимно прост со степенью $g^n(x)$ при всех $n \in \mathbb{N}$. Тогда отсюда и из свойства 1 следует, что произведение $f^k(x)f(x)$ взаимно просто с $g^n(x)$ при всех $n \in \mathbb{N}$, т.е. утверждение верно при $m = k + 1$. ■

Задача 22. *1) Докажите, что многочлены $f_1(x)$, $f_2(x)$, $f_3(x)$ из примера 3 взаимно просты.*

2) Докажите, что если любые два многочлена из (51) взаимно просты, то все они взаимно просты.

3) Докажите обобщение свойства 1 теоремы 1.21.2 на любое число многочленов $g_j(x), k = \overline{1,k}$.

4) Докажите, что если взаимно просты некоторые степени $f^{m_0}(x)$ *и* $g^{n_0}(x)$ *многочленов* $f(x)$ *и* $g(x)$, *то взаимно просты любые их степени* $f^m(x)$ *и* $g^n(x)$, *где* $m, n \in \mathbb{N}$.

5) Пусть $d(x) = f(x)u(x) + g(x)v(x)$, *где* $d(x)$ — *наибольший общий делитель многочленов* $f(x)$ *и* $g(x)$. *Найдите наибольший общий делитель многочленов* $u(x)$ *и* $v(x)$. *Ответ обоснуйте.*

1.22. Корень многочлена. Кратность корня

В разделе 1.17. мы определили многочлен $f(x) \in P[x]$ над полем P как некоторое формальное выражение, которое вполне определяется набором своих коэффициентов. Однако многочлен можно рассматривать и как функцию от переменной x.

Определение 1. *Если* $f(x) = \sum\limits_{k=0}^{n} f_k x^k$ — *многочлен над полем* P, c — *некоторое число из поля* P, *то число* $f(c) = \sum\limits_{k=0}^{n} f_k\, c^k$ *называется значением многочлена* $f(x)$ *при* $x = c$.

Корнем многочлена $f(x) \in P[x]$ *или уравнения* $f(x) = 0$ *называется такое число* $c \in P$, *что* $f(c) = 0$.

Из теории функций известно, что две функции называются равными, если их значения равны при всех значениях переменной. Очевидно что если многочлены $f(x)$ и $g(x)$ равны как многочлены, т.е. их коэффициенты совпадают (см. определение 35), то они равны как функции. Обратное утверждение требует дополнительного исследования и это будет сделано позже для многочленов $C[x]$. А пока, забегая вперёд, будем иметь в виду, что оба подхода к понятию равенства многочленов совпадают.

Сказанное относится и к операциям над многочленами. Если определены (см. определение 2) сумма и произведение многочленов $f(x)$ и $g(x)$: $\varphi(x) = f(x) + g(x)$, $\psi(x) = f(x)g(x)$, то $\forall\, c \in P$ справедливы равенства $\varphi(c) = f(c) + g(c)$, $\psi(c) = f(c)g(c)$, которые соответствуют определению суммы и

произведения функций. Следовательно, операциям сложения и умножения над многочленами соответствуют операции сложения и умножения над функциями, которые определяются этими многочленами.

Теорема 1.22.1 (Теорема Безу).

Остаток от деления многочлена $f(x)$ *на* $x - c$ $(c \in P)$ *равен* $f(c)$.

Доказательство. Разделим многочлен $f(x)$ на $x - c$. Тогда $f(x) = (x - c)h(x) + r(x)$, где $deg(r(x)) < deg(x - c) = 1$, так что $r(x) = r_0$ — константа. Из $f(x) = (x - c)h(x) + r_0$ получим при $x = c$ значения $f(c) = (c - c)h(c) + r_0$, т.е. остаток $r_0 = f(c)$.

Следствия. 1. *Число* $c \in P$ *является корнем многочлена* $f(x) \in P[x]$ *тогда и только тогда, когда многочлен* $f(x)$ *делится на* $x - c$, *т.е.* $f(x) = (x - c)h(x)$, *где многочлен* $h(x) \in P[x]$.

2. *Если число* $c \in P$ *является корнем ненулевого многочлена* $f(x) \in P[x]$, *то существуют число* $k \in N$ *и многочлен* $s(x) \in P[x]$ *такие, что*

$$f(x) = (x - c)^k s(x), \tag{52}$$

где многочлен $s(x)$ *не делится на* $x - c$, *т.е. число* c *своим корнем не имеет.*

Доказательство.

1. Из теоремы Безу следует, что $f(x) = (x-c)h(x)+r(x)$, где $r(x) = f(c)$.

Если $f(c) = 0$, то $f(x) = (x - c)h(x)$ и $f(x)$ делится на $x - c$.

Если $f(x)$ делится на $x - c$, то остаток от деления равен нулю, то есть $f(c) = 0$.

2. Докажем следствие 2 методом математической индукции по степени m многочлена.

Пусть многочлен $f(x) = a_0x + a_1$ $(a_0 \neq 0)$ степени $m = 1$ имеет корень $x = c$. Из следствия 1 имеем $f(x) = (x - c)h(x)$, где многочлен $h(x) \in P[x]$. Сравнивая левую и правую части равенства заключаем, что $h(x)$ — это ненулевая константа a_0. Отсюда $h(c) \neq 0$ и утверждение верно при $k = 1$, $s(x) = h(x)$.

Покажем, что из справедливости утверждения для степени $m = n$ следует его справедливость для степени $m = n + 1$.

Из следствия 1 имеем $f(x) = (x - c)h(x)$, где многочлен $h(x) \in P[x]$. Очевидно, что степень многочлена $h(x)$ равна n. Если $h(c) \neq 0$, то утвер-

ждение верно при $k = 1$, $s(x) = h(x)$. Если $h(c) = 0$, то по предположению индукции существуют число $k_1 \in N$ и многочлен $h_1(x) \in P[x]$ такие, что $h(x) = (x - c)^{k_1} h_1(x)$, причём $h_1(c) \neq 0$. Тогда $f(x) = (x - c)^{1+k_1} h_1(x)$ и утверждение верно при $k = 1 + k_1$, $s(x) = h_1(x)$.■

Замечание. Следствие 1 из теоремы Безу говорит о том, что задача о нахождении корня $x = c$ многочлена эквивалентна задаче о нахождении делителя $x - c$ этого многочлена. ■

Определение 2. *Число $k \in N$ называется кратностью корня $c \in P$ многочлена $f(x) \in P[x]$, если выполнено равенство*

$$f(x) = (x - c)^k s(x),$$

где многочлен $s(x) \in P[x]$ не делится на $x - c$, т.е. число c своим корнем не имеет. При этом сам корень c называется k-кратным. Для $k = 1$ говорят, что корень c — простой.

Если для многочленов над полем P определено понятие производных, справедливы формулы дифференцирования суммы, произведения и формула Тейлора, то тогда справедливы следующие утверждения о кратности корня многочлена.

Теорема 1.22.2. 1° *Если число c является k-кратным корнем многочлена $f(x) \in P[x]$, то при $k > 1$ оно будет $(k-1)$-кратным корнем первой производной этого многочлена; при $k = 1$ число c не является корнем $f'(x)$.*

2° *Если число c является k-кратным корнем многочлена $f(x) \in P[x]$, то c будет $(k - m)$-кратным корнем в m-й производной этого многочлена при $m < k$ и не будет корнем k-й производной от $f(x)$.*

3° *Число c является k-кратным корнем многочлена $f(x) \in P[x]$ тогда и только тогда, когда $f(c) = f'(c) = \cdots = f^{(k-1)}(c) = 0$ и $f^{(k)}(c) \neq 0$.*

1.23. Основная теорема алгебры для многочленов над C

Теорема 1.23.1. *Любой многочлен $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ с комплексными коэффициентами степени $n \geqslant 1$ имеет хотя бы один корень, в общем случае комплексный.*

Доказательство теоремы. Покажем, что инфимум значений модуля многочлена $|f(z)|$ достигается при некотором $z = z_0$ и при этом $|f(z_0)| = 0$, т. е. z_0 — корень многочлена $f(z)$.

Для этого нам понадобятся дополнительные понятия и факты.

Сначала докажем непрерывность многочлена и его модуля.

Комплексная функция $f(z)$ от комплексной переменной z называется непрерывной в точке z_0, если малому изменению аргумента соответствует малое изменение функции, т.е. для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех z, для которых $|z - z_0| < \delta$, выполняется неравенство $|f(z) - -f(z_0)| < \varepsilon$.

Лемма 1.23.1. (**Непрерывность в нуле многочлена с нулевым свободным членом.**)

Пусть $f(z) = a_1 z + a_2 z^2 + \ldots + a_n z^n$ — многочлен с нулевым свободным членом. Тогда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех z, для которых $|z| < \delta$, выполняется неравенство $|f(z)| < \varepsilon$.

Доказательство леммы 1.23.1. Пусть $|z| < 1$, тогда $|f(z)| = |z||a_1 + +a_2 z^2 + \ldots + a_n z^{n-1}| \leqslant |z|(|a_1| + |a_2||z|^2| + \ldots + |a_n||z|^{n-1}) \leqslant |z|(|a_1| + |a_2| + \ldots + +|a_n|)$. Правая часть мала при малых z. Положим $M = |a_1| + |a_2| + \ldots + |a_n|$ и возьмём $\delta = \min(1, \dfrac{\varepsilon}{M})$. Тогда при $|z| < \delta$ имеем $|f(z)| \leqslant |z|M < \dfrac{\varepsilon M}{M} = \varepsilon$. ■

Лемма 1.23.2. (**Непрерывность многочлена.**)

Многочлен $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ есть непрерывная функция во всех точках комплексной плоскости.

Доказательство леммы 1.23.2. Пусть z_0 — произвольное комплексное число. Разложим $f(z)$ по степеням $z - z_0$: $f(z) = c_0 + c_1(z - z_0) + \ldots + c_n(z - z_0)^n$. Тогда $c_0 = f(z_0)$, так что $f(z) - f(z_0) = c_1(z - z_0) + \ldots + c_n(z - z_0)^n$.

Правая часть представляет собой многочлен от $z - z_0$ с нулевым свободным членом. По лемме 1.23.1 правая часть мала при малых $z - z_0$. Действительно, для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $|f(z) - f(z_0)| < \varepsilon$ для всех z, для которых $|z - z_0| < \delta$. ■

Лемма 1.23.3. (**Непрерывность модуля многочлена.**)

Модуль многочлена есть непрерывная функция.

Доказательство леммы 1.23.3. Утверждение вытекает из леммы 1.23.2 и неравенства $\big||f(z)| - |f(z_0)|\big| \leqslant |f(z) - f(z_0)|$, т.к. правая часть мала при малых $z - z_0$.■

Лемма 1.23.4. (**Значения многочлена $f(z) \to \infty$ при $z \to \infty$.**)

Если $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ многочлен степени $n \geqslant 1$, то для любого $M > 0$ существует $R > 0$ такое что для всех z, для которых $|z| > R$, выполняется неравенство $|f(z)| > M$.

Доказательство леммы 1.23.4. Преобразуем $f(z) = a_0 + a_1 z + \ldots + a_n z^n = a_n z^n (1 + \dfrac{a_{n-1}}{a_n} z^{-1} + \ldots + \dfrac{a_0}{a_n} z^{-n}) = a_n z^n (1 + g(z^{-1}))$, где $g(z^{-1})$ — многочлен от z^{-1} с нулевым свободным членом. Оценим снизу $|f(z)| = |a_n z^n||1 + g(z^{-1})| \geqslant |a_n z^n|\big|1 - |g(z^{-1})|\big|$. Здесь $|a_n z^n|$ сколь угодно большое при достаточно большом $|z|$, $|g(z^{-1})|$ по лемме 1.23.1 сколь угодно малое, например, $|g(z^{-1})| < 1/2$ при достаточно малом $|z^{-1}|$ или при достаточно большом $|z|$. Тогда $|f(z)| > |a_n z^n|/2$ сколь угодно большое при достаточно большом $|z|$.

Получим нужную оценку. По лемме 1.23.1 для $\varepsilon = 1/2$ найдётся $\delta > 0$ такое, что при $|z^{-1}| < \delta$ выполнено неравенство $|g(z^{-1})| < 1/2$. Для произвольного $M > 0$ при $|z| > \sqrt[n]{\dfrac{2M}{|a_n|}}$ справедливо неравенство $|a_n z^n| > 2M$. Выберем $R = max\left(\sqrt[n]{\dfrac{2M}{|a_n|}}, \dfrac{1}{\delta}\right)$. Тогда если $|z| > R$, то $|z^{-1}| < \delta$ и $|z| > \sqrt[n]{\dfrac{2M}{|a_n|}}$, поэтому $|f(z)| = |a_n z^n||1 + g(z^{-1})| \geqslant |a_n z^n|\big|1 - |g(z^{-1})|\big| > 2M\left(1 - \dfrac{1}{2}\right) = M$. ■

Число $z_0 = x_0 + iy_0$ называется пределом последовательности $z_n = x_n + iy_n$, если для любого $\varepsilon > 0$ существует натуральное число n_0 такое, что $|z_n - z_0| < \varepsilon$ для всех n_0. Обозначение: $\lim\limits_{n\to\infty} z_n = z_0$.

Очевидно, что сходимость последовательности комплексных чисел $z_n = x_n + iy_n$ к $z_0 = x_0 + iy_0$ равносильна сходимости двух последовательностей действительных чисел: $\lim\limits_{n\to\infty} x_n = x_0$, $\lim\limits_{n\to\infty} y_n = y_0$.

Последовательность z_n называется ограниченной, если найдётся число $R > 0$ такое, что $|z_n| \leqslant R$ для всех n.

Лемма 1.23.5. (**Лемма Больцано-Вейерштрасса.**)

Из любой ограниченной последовательности можно выделить сходящуюся подпоследовательность.

Доказательство леммы 1.23.5. Пусть $z_n = x_n + iy_n$ и $|z_n| \leqslant R$ для всех n. Тогда $|x_n| \leqslant |z_n| \leqslant R$ и $|y_n| \leqslant |z_n| \leqslant R$ для всех n. По лемме Больцано-Вейерштрасса из ограниченной вещественной последовательности x_n можно выделить сходящуюся подпоследовательность $x_{n_k} \to x_0$. Рассмотрим подпоследовательность y_{n_k}. Она ограничена и из неё можно выделить сходящуюся подпоследовательность $y_{n_{k_m}} \to y_0$. Тогда $x_{n_{k_m}} \to x_0$ и $z_{n_{k_m}} = x_{n_{k_m}} + iy_{n_{k_m}} \to z_0 = x_0 + iy_0$. ■

Лемма 1.23.6. (**Лемма о достижении инфимума модуля многочлена.**)

Точная нижняя грань модуля многочлена $f(z)$ достигается, т.е. существует число z_0 такое, что $|f(z_0)| \leqslant |f(z)|$ при всех комплексных z.

Доказательство леммы 1.23.6. Рассмотрим множество значений модуля многочлена $\{|f(z)| \colon z \in C\}$. Поскольку $|f(z)| \geqslant 0$, то это множество ограничено снизу и, следовательно, имеет инфимум, который обозначим через m. Тогда для любого натурального числа n можно найти комплексное число z_n такое, что

$$m \leqslant |f(z_n)| \leqslant m + \frac{1}{n}. \tag{53}$$

Воспользуемся леммой 1.23.4 и для $M = m + 1$ найдём R такое, что при $|z| > R$ будет $|f(z)| > M$. Поскольку $M = m + 1 \geqslant m + \dfrac{1}{n}$, то при $|z| > R$ выполнено неравенство $|f(z)| > m + \dfrac{1}{n}$ для всех n. Отсюда и из (53) следует, что $|z_n| \leqslant R$, т.е. последовательность z_n ограничена. Согласно лемме 1.23.5 из z_n можно выделить сходящуюся подпоследовательность $z_{n_k} \to z_0$. Тогда в силу непрерывности $|f(z)|$ (лемма 1.23.3)

$$\lim_{k\to\infty} |f(z_{n_k})| = |f(z_0)|. \tag{54}$$

Из неравенства (53) для подпоследовательности z_{n_k} получаем неравенство $m \leqslant |f(z_{n_k})| \leqslant m + \dfrac{1}{n_k}$, поэтому

$$\lim_{k\to\infty} |f(z_{n_k})| = m. \tag{55}$$

В силу единственности предела из (54) и (55) следует, что $|f(z_0)| = m$. ■

Лемма 1.23.7. (**Лемма Даламбера.**)

Если $f(z)$ *— многочлен степени* $n \geqslant 1$ *и* $f(z_0) \neq 0$, *то найдётся число* z_1 *такое, что* $|f(z_1)| < |f(z_0)|$.

Доказательство леммы 1.23.7. Для многочлена $f(z) = a_0 + a_1 z$ утверждение выполнено при $z_1 = -\dfrac{a_0}{a_1}$. При $n > 1$ разложим многочлен $f(z)$ по степеням $z - z_0$:

$$f(z) = c_0 + c_1(z - z_0) + \ldots + c_n(z - z_0)^n. \tag{56}$$

Очевидно, что $c_0 = f(z_0) \neq 0$. Пусть c_k — первый ненулевой коэффициент в (56) после c_0 (такой коэффициент имеется, так как $f(z)$ не константа). Преобразуем многочлен (56)

$$\begin{aligned} f(z) &= c_0 + c_k(z - z_0)^k + c_{k+1}(z - z_0)^{k+1} + \ldots + c_n(z - z_0)^n = \\ &= c_0\left(1 + \frac{c_k}{c_0}(z - z_0)^k + \frac{c_{k+1}}{c_0}(z - z_0)^{k+1} + \ldots + \frac{c_n}{c_0}(z - z_0)^n\right) = \\ &= c_0\left(1 + \frac{c_k}{c_0}(z - z_0)^k + \frac{c_k}{c_0}(z - z_0)^k\left(\frac{c_{k+1}}{c_k}(z - z_0) + \ldots + \frac{c_n}{c_k}(z - z_0)^{n-k}\right)\right) = \\ &= c_0\left(1 + \frac{c_k}{c_0}(z - z_0)^k + \frac{c_k}{c_0}(z - z_0)^k g(z - z_0)\right), \end{aligned} \tag{57}$$

где $g(z - z_0) = \dfrac{c_{k+1}}{c_k}(z - z_0) + \ldots + \dfrac{c_n}{c_k}(z - z_0)^{n-k}$ — многочлен от $z - z_0$ с нулевым свободным членом, предположим сначала, что $g(z - z_0)$ ненулевой. Введём обозначения $h = z - z_0$, $s = \dfrac{c_k}{c_0}$ и запишем равенство (57) в виде $f(z_0 + h) = c_0\left((1 + sh^k) + sh^k g(h)\right)$. Отсюда следует оценка $|f(z_0 + h)| = |c_0|\,|(1 + sh^k) + sh^k g(h)| \leqslant |c_0|(|1 + sh^k| + |sh^k||g(h)|)$. Если найдём такое $h \neq 0$, при котором $0 < |sh^k| < 1$, $|1 + sh^k| = 1 - |sh^k|$ и $0 < |g(h)| < \dfrac{1}{2}$, то $|f(z_0 + h)| < |c_0|(1 - |sh^k| + \dfrac{1}{2}|sh^k|) = |c_0|(1 - \dfrac{1}{2}|sh^k|) < |c_0| = |f(z_0)|$.

Будем искать h в виде $h = r(\cos\varphi + i\sin\varphi)$, $r = |h|$, $\varphi = \arg(h)$. Причём сначала выберем и фиксируем достаточно малое значение $r > 0$ такое, что при всех h, для которых $|h| = r$, выполнены неравенства $0 < |g(h)| < \dfrac{1}{2}$ и $0 < |sh^k| < 1$. Затем выберем $\varphi = \arg(h)$ так, чтобы sh^k являлось вещественным отрицательным числом. Для этого найдём такой $\arg(h)$, для которого $\arg(sh^k) = \pi$. Поскольку $\arg(sh^k) = \arg(s) + \arg(h^k) = \arg(s) + k\arg(h) = \arg(s) + k\varphi$ достаточно взять $\varphi = \varphi_1 = \dfrac{\pi - arg(s)}{k}$, чтобы для

$h_1 = r(\cos\varphi_1 + i\sin\varphi_1)$ при выбранном r выполнялись условия $0 < |g(h_1)| < \dfrac{1}{2}$, $0 < |sh_1^k| < 1$, sh_1^k является вещественным отрицательным числом. Отсюда $-1 < sh_1^k < 0$, $-|sh_1^k| = sh_1^k$, $1 + sh_1^k > 0$, $|1 + sh_1^k| = 1 + sh_1^k = 1 - |sh_1^k|$. Тогда по ранее доказанному $|f(z_0 + h_1)| < |f(z_0)|$.

Доказательство теоремы 1.23.1. Пусть $f(z)$ — произвольный многочлен степени $n \geqslant 1$ над полем C от переменной z. Согласно лемме 1.23.6 множество всех значений $|f(z)|$ имеет инфимум m, который достигается в некоторой точке z_0, т.е. $|f(z_0)| = m$. Значение $m = 0$. Действительно, если $|f(z_0)| \neq 0$, то по лемме 1.23.7 найдётся число z_1 такое, что $|f(z_1)| < |f(z_0)| = \inf\{|f(z)|\}$, что невозможно. Таким образом, z_0 — корень многочлена $f(z)$. ■

1.24. Следствия из основной теоремы алгебры для многочленов над полем C

В этом разделе получим следующие следствия из основной теоремы алгебры для многочленов над C: существование и единственность разложения многочлена на линейные множители, каноническое разложение многочлена на множители, эквивалентность определений равенства двух многочленов по коэффициентам и по значениям, интерполяционный многочлен Лагранжа, формулы Виета.

Теорема 1.24.1. (Существование и единственность разложения многочлена на линейные множители.)

Для любого многочлена $f(z) = \sum\limits_{k=0}^{n} a_k z^k \in C[z]$ *степени* $n \geqslant 1$ *существуют числа* $c_1, c_2, \ldots, c_n \in C$ *такие, что*

$$f(z) = a_n(z - c_1)(z - c_2)\ldots(z - c_n). \tag{58}$$

Это разложение единственно с точностью до порядка сомножителей.

Доказательство. I. Существование докажем ММИ по n.

1) При $n = 1$ имеем $f(z) = a_0 + a_1 z = a_1\left(z + \dfrac{a_0}{a_1}\right)$, т. е. (58) выполнено для $c_1 = -\dfrac{a_0}{a_1}$.

2) Пусть для любого многочлена $f(z) = \sum\limits_{k=0}^{m} a_k z^k$ степени m существуют числа $c_1, c_2, \ldots, c_m \in C$ такие, что $f(z) = a_m(z - c_1)(z - c_2)\ldots(z - c_m)$.

3) Покажем, что утверждение верно для многочлена $f(z) = \sum\limits_{k=0}^{m+1} a_k z^k$ степени $m + 1$.

Из основной теоремы 1.23.1 следует существование корня $z = c_{m+1}$ многочлена $f(z)$ и из следствия 1 теоремы Безу следует разложение $f(z) = h(z)(z - c_{m+1})$. При этом многочлен $h(x)$ имеет степень m и старший коэффициент a_{m+1}. Тогда по предположению индукции существуют существуют числа $c_1, c_2, \ldots, c_m \in C$ такие, что $h(z) = a_{m+1}(z - c_1)(z - c_2)\ldots(z - c_m)$. Следовательно, $f(z) = a_{m+1}(z - c_1)(z - c_2)\ldots(z - c_m)(z - c_{m+1})$.

II. Докажем единственность разложения (58). Предположим, что есть второе разложение $f(z) = a_n(z - d_1)(z - d_2)\ldots(z - d_n)$. Тогда справедливо равенство $a_n(z-c_1)(z-c_2)\ldots(z-c_n) = a_n(z-d_1)(z-d_2)\ldots(z-d_n)$. Отсюда, каждый корень c_j совпадает с некоторым корнем d_k и каждый корень d_q совпадает с некоторым корнем c_r. Пусть произвольный корень $z = \alpha$ встречается в первом разложении s раз, а во втором разложении — t раз. Покажем, что $s = t$. Предположим противное $s > t$. Тогда $a_n(z-\alpha)^s \prod\limits_{c_j \neq \alpha} (z - c_j) = a_n(z-\alpha)^t \prod\limits_{d_q \neq \alpha} (z - d_q)$. После сокращения на $a_n(z - \alpha)^t$ получим равенство $(z - \alpha)^{s-t} \prod\limits_{c_j \neq \alpha} (z - c_j) = \prod\limits_{d_q \neq \alpha} (z - d_q)$, которое не выполнено при $z = \alpha$. ■

Теорема 1.24.2. (Каноническое разложение на множители многочлена над полем C.)

Для любого многочлена $f(z) = \sum\limits_{k=0}^{n} a_k z^k \in C[z]$ степени $n \geqslant 1$ существуют число $m \in N$, $1 \leqslant m \leqslant n$, числа $c_1, c_2, \ldots, c_m \in C$, $c_i \neq c_j$, при $i \neq j$, и числа $k_1, k_2, \ldots, k_m$, $k_1 + k_2 + \ldots + k_m = n$ такие, что

$$f(z) = a_n(z - c_1)^{k_1}(z - c_2)^{k_2}\ldots(z - c_m)^{k_m}. \tag{59}$$

Это разложение единственно с точностью до порядка сомножителей. Числа $c_1, c_2, \ldots, c_m$ являются корнями многочлена $f(z)$ кратностей $k_1, k_2, \ldots, k_m$.

Доказательство. Рассмотрим разложение (58): $f(z) = a_n(z - c_1)(z - c_2)\ldots(z - c_n)$. Среди n чисел $c_1, c_2, \ldots, c_n$ могут быть одинаковые. Перенумеруем эти n чисел так, чтобы $c_1, c_2, \ldots, c_m$ были различными, а каждое

из $c_{m+1}, \ldots, c_n$ совпадало с одним из $c_1, c_2, \ldots, c_m$. Объединив одинаковые сомножители в перенумерованном разложении получим $f(z) = a_n(z-c_1)^{k_1}(z-$ $-c_2)^{k_2}\ldots(z-c_m)^{k_m}$, где $c_i \neq c_j$ при $i \neq j$ по построению, а $k_1+k_2+\ldots+k_m = n$, т. к. $k_1 + k_2 + \ldots + k_m$ — это число сомножителей $z - c_j$ в разложении (58). Единственность разложения следует из теоремы 1.24.1.

Следствие. Любой многочлен $f(z) \in C[z]$ степени $n \geqslant 1$ имеет ровно n корней, если каждый корень считать столько раз, какова его кратность.

Замечание. Это утверждение применимо к многочлену нулевой степени, который не имеет корней, и неприменимо лишь к нулевому многочлену, т.к. такой многочлен имеет корнем любое число $z \in C$. ■

Рассмотрим вопрос о эквивалентности определений равенства двух многочленов по коэффициентам и по значениям.

Теорема 1.24.3. (Равенство двух многочленов.)

Если два многочлена $f(z)$, $g(z) \in C[z]$, степени которых не превосходят $n \geqslant 1$, имеют равные значения при более чем n различных значениях переменной, то они равны.

Доказательство теоремы. Разность $h(z) = f(z) - g(z) \in C[z]$ имеет степень не выше n и имеет больше n различных корней. Следовательно, $h(z) =$ $= 0 \Rightarrow f(z) = g(z)$.

Следствия. 1. *Два многочлена $f(z)$, $g(z) \in C[z]$ равны тогда и только тогда, когда совпадают их значения при всех значениях переменной z.*

2. *Многочлен $f(z) \in C[z]$ степени не выше n однозначно определяется своими значениями в $n+1$ различных точках.*

Доказательство следствий. 1. а) Многочлены $f(z) = \sum\limits_{k=0}^{\infty} f_k z^k$ и $g(z) =$ $= \sum\limits_{k=0}^{\infty} g_k z^k$ равны $\Rightarrow$ коэффициенты равны $\forall k \in Z_+$ $f_k = g_k \Rightarrow$ значения равны $\forall z \in C$ $\sum\limits_{k=0}^{\infty} f_k z^k = \sum\limits_{k=0}^{\infty} g_k z^k$.

б) Если значения многочленов совпадают при всех значениях переменной, то по теореме 1.24.3 они равны.

2. Пусть многочлен $f(z) = \sum\limits_{k=0}^{n} f_k z^k$ степени не выше n вычислен в $n+1$ различных точках. Тогда по теореме 1.24.3 других многочленов с такими же значениями нет. ■

Теорема 1.24.4. (Интерполяционный многочлен.)

Для любых чисел $z_1, z_2, \ldots, z_{n+1} \in C$, $z_i \neq z_j$ *при* $i \neq j$ *и любых чисел* $u_1, u_2, \ldots, u_{n+1} \in C$ *существует единственный многочлен* $f(z) \in C[z]$ *степени не выше* n, *для которого выполнены условия интерполяции*

$$f(z_k) = u_k,\ k = \overline{1, n+1}. \tag{60}$$

Этот многочлен можно записать в форме Лагранжа

$$f(z) = \sum_{k=1}^{n+1} u_k \frac{(z-z_1)\ldots(z-z_{k-1})(z-z_{k+1})\ldots(z-z_{n+1})}{(z_k-z_1)\ldots(z_k-z_{k-1})(z_k-z_{k+1})\ldots(z_k-z_{n+1})} =$$
$$= \sum_{k=1}^{n+1} u_k \prod_{j\neq k} \frac{(z-z_j)}{(z_k-z_j)} \tag{61}$$

или в форме Ньютона

$$f(z) = v_1 + v_2(z-z_1) + v_3(z-z_1)(z-z_2) + \ldots +$$
$$+ v_{n+1}(z-z_1)(z-z_2)\ldots(z-z_n), \tag{62}$$

где v_j, $j = \overline{1, n+1}$ *последовательно вычисляются с использованием условий интерполяции (60).*

Доказательство. Из формулы (61) следует, что это многочлен степени не выше n. При подстановке $z = z_k$ в $u_k \prod\limits_{j\neq k} \dfrac{(z-z_j)}{(z_k-z_j)}$ получим u_k, а в $u_m \prod\limits_{j\neq m} \dfrac{(z-z_j)}{(z_m-z_j)}$ получим 0, тогда выполнены условия (60). В силу следствия 2 теоремы 1.24.3 многочлен с условиями единственен. ■

Теорема 1.24.5. (Формулы Виета.)

Если $c_1, c_2, \ldots, c_n \in C$ *— корни многочлена* $f(z) = \sum\limits_{k=0}^{n} a_k z^k \in C[z]$, *то*

$$\begin{aligned}
&a_{n-1}/a_n = -(c_1 + c_2 + \ldots + c_n),\\
&a_{n-2}/a_n = +(c_1c_2 + c_1c_3 + \ldots + c_1c_n + c_2c_3 + \ldots + c_{n-1}c_n),\\
&a_{n-3}/a_n = (-1)^3 \sum_{i_1<i_2<i_3} c_{i_1}c_{i_2}c_{i_3},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{n-k}/a_n = (-1)^k \sum_{i_1<i_2<\ldots<i_k} c_{i_1}c_{i_2}\ldots c_{i_k},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_0/a_n = (-1)^n c_1c_2\ldots c_n.
\end{aligned} \tag{63}$$

Доказательство. Приравняем коэффициенты при одинаковых степенях $f(z) = \sum\limits_{k=0}^{n} a_k z^k$ и $f(z) = a_n(z-c_1)(z-c_2)\ldots(z-c_n)$:

при $z^{n-1}:\ a_{n-1} = a_n(-1)(c_1 + c_2 + \ldots + c_n)$,

при $z^{n-2}:\ a_{n-2} = a_n(-1)^2(c_1c_2 + c_1c_3 + \ldots + c_1c_n + c_2c_3 + \ldots + c_{n-1}c_n) =$

$$= a_n(-1)^2 \sum_{i_1<i_2} c_{i_1}c_{i_2},$$

. . . . . .

при $z^{n-k}:\ a_{n-k} = a_n(-1)^k \sum\limits_{i_1<i_2<\ldots<i_k} c_{i_1}c_{i_2}\ldots c_{i_k}$,

. . . . . .

при $z^0:\ a_0 = a_n(-1)^n c_1c_2\ldots c_n$. ■

1.25. Многочлены над полем R

Пусть $f(x) = \sum\limits_{k=0}^{n} a_kx^k$ — многочлен с вещественными коэффициентами степени n. Так как $R \subset C$, то $f(x)$ можно рассматривать как многочлен над полем C и тогда для этого многочлена справедливы все утверждения из п.1.24. Уточним некоторые из этих утверждений для многочленов над R.

Теорема 1.25.1. (О сопряжённых корнях многочлена над R.)

Если $c = \alpha + \beta i\,(\beta \neq 0)$ — комплексный корень многочлена $f(x) \in R[x]$ степени $n \geqslant 1$, то $\overline{c} = \alpha - \beta i$ также является корнем этого многочлена, причём кратности корней c и $\overline{c}$ одинаковы.

Доказательство. Если $c = \alpha + \beta i\,(\beta \neq 0)$ — корень многочлена $f(x) = \sum\limits_{k=1}^{n} a_kx^k$, то $a_0 + a_1c + \ldots + a_nc^n = 0$. Возьмём сопряжение от обоих частей этого равенства и получим $\overline{f(c)} = \overline{0}$. Используя свойства сопряжения, преобразуем равенство $\overline{a_0 + a_1c + \ldots + a_nc^n} = \overline{0} \Leftrightarrow \overline{a_0} + \overline{a_1}\,\overline{c} + \ldots + \overline{a_n}\,\overline{c}^n = 0 \Leftrightarrow a_0 + a_1\,\overline{c} + \ldots + a_n\,\overline{c}^n = 0$ и получим $f(\overline{c}) = 0$. Многочлен $f(x)$ делится на $x - c$, на $x - \overline{c}$ и в силу их взаимной простоты делится на их произведение $h_1(x) = (x - c)(x - \overline{c}) = x^2 - (c + \overline{c})x + c\overline{c}$, т.е. $f(x) = h_1(x)f_1(x)$. Здесь $h_1(x)$ имеет вещественные коэффициенты, т.к. $c + \overline{c} = 2Re(c)$, $c\overline{c} = |c|^2$, поэтому согласно алгоритму деления многочлен $f_1(x)$ имеет вещественные коэффициенты, при этом его степень равна $n - 2$. Если $f_1(x)$ не имеет комплексных корней вида $\gamma + \delta i\,(\delta \neq 0)$, то утверждение теоремы верно, т.к. $f(x) = (x-c)(x-\overline{c})f_1(x)$, где c и $\overline{c}$ простые корни. Если $f_1(x)$ имеет комплексный корень $d = \gamma+\delta i\,(\delta \neq 0)$, то по доказанному он имеет комплексный корень

$\overline{d}$. Тогда $f_1(x) = (x-d)(x-\overline{d})f_2(x)$, т. е. $f(x) = (x-c)(x-\overline{c})(x-d)(x-\overline{d})f_2(x)$ и т.д.

Следствие. Любой многочлен $f(x) \in R[x]$ нечётной степени имеет хотя бы один вещественный корень.

Теорема 1.25.2. (Каноническое разложение на множители многочлена над полем R.)

Для любого многочлена $f(x) = \sum\limits_{k=0}^{n} a_k x^k \in R[x]$ степени $n \geqslant 1$ справедливо разложение

$$f(x) = a_n \prod_{j=1}^{m} (x - c_j)^{k_j} \prod_{s=1}^{r} (x^2 + p_s x + q_s)^{t_s}. \tag{64}$$

где числа $m, r \in Z_+$, $m+r > 0$, при $m \neq 0$ числа $c_1, \ldots, c_m$ — это все различные вещественные корни данного многочлена с кратностями $k_1, \ldots, k_m \in N$, при $r \neq 0$ квадратные трёхчлены $x^2 + p_1 x + q_1, \ldots, x^2 + p_r x + q_r$ различны, имеют вещественные коэффициенты и отрицательные дискриминанты, а показатели степеней этих трёхчленов $t_1, \ldots, t_m$ являются натуральным числами, кроме того выполнено равенство $\sum\limits_{j=1}^{m} k_j + 2\sum\limits_{s=1}^{r} t_s = n$.

Разложение (64) единственно с точностью до порядка сомножителей.

1.26. Собственные векторы и собственные значения квадратной матрицы

В этом разделе рассматриваем матрицы из $P^{n\times n}$ над полем P.

Определение 1. *Ненулевой вектор $x = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} \in P^n$ называется собственным вектором матрицы $\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \in P^{n\times n}$, если существует такое число $\lambda \in P$, что*

$$\mathcal{A}x = \lambda x. \tag{65}$$

Число λ называется собственным значением матрицы A, соответствующим собственному вектору x.

При этом говорят также, что собственный вектор x соответствует собственному значению λ.

Множество всех собственных значений матрицы называется её спектром. ■

Примеры. 1. У нулевой матрицы существует собственное значение $\lambda = 0$, которому соответствует любой ненулевой вектор, т. к. $\mathcal{O}x = 0x$.

2. У скалярной матрицы $\beta\mathcal{J}$ существует собственное значение $\lambda = \beta$, которому соответствует любой ненулевой вектор, т. к. $(\beta\mathcal{J})x = \beta x$. ■

Приведём утверждение, которое даёт возможность вычисления собственных векторов матрицы, если известны её собственные значения. Эта теорема не гарантирует наличие собственных значений. Как показывают примеры существуют матрицы, которые не имеют собственных значений.

Теорема 1.26.1. *Матрица $\mathcal{A} = (a_{ij})$ имеет ненулевой собственный вектор x, отвечающий собственному значению $\lambda = \lambda_0$, т.е. выполнено равенство $Ax = \lambda_0 x$, тогда и только тогда, когда λ_0 является корнем характеристического многочлена матрицы $\mathcal{A}$*

$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} \qquad (66).$$

Доказательство. 1) Пусть матрица $\mathcal{A}$ имеет собственный вектор x, отвечающий собственному значению $\lambda = \lambda_0$. Тогда $\mathcal{A}x = \lambda_0 x$, т. е. ненулевой вектор x является нетривиальным решением этого матричного уравнения. Далее $\mathcal{A}x = \lambda_0 x \Leftrightarrow \mathcal{A}x = \lambda_0\mathcal{J}x \Leftrightarrow (\mathcal{A} - \lambda_0\mathcal{J})x = \theta$. В подробной записи это равенство имеет вид

$$\begin{cases} (a_{11} - \lambda_0)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + (a_{22} - \lambda_0)x_2 + \ldots + a_{2n}x_n = 0, \\ \dots \qquad \dots \qquad \dots \qquad \dots \qquad \dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda_0)x_n = 0. \end{cases} \qquad (67)$$

Эта система имеет ненулевое решение тогда и только тогда, когда определитель системы равен нулю, то есть

$$f(\lambda_0) = \begin{vmatrix} a_{11} - \lambda_0 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda_0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda_0 \end{vmatrix} = 0.$$

Последнее означает, что λ_0 является корнем характеристического многочлена $f(\lambda)$.

2) Обратно. Если λ_0 является корнем характеристического многочлена $f(\lambda)$, то система (67) имеет ненулевое решение — вектор x, удовлетворяющий равенствам $(\mathcal{A} - \lambda_0 \mathcal{J})x = \theta \Leftrightarrow \mathcal{A}x = \lambda_0 \mathcal{J}x \Leftrightarrow Ax = \lambda_0 x$ и, следовательно, преобразование A имеет собственный вектор x отвечающий собственному значению $\lambda = \lambda_0$.

Следствия. 1. Любая матрица из $C^{n\times n}$ имеет по крайней мере одно собственное значение и соответствующий ему собственный вектор.

2. Любая матрица из $R^{n\times n}$ при нечётном n имеет по крайней мере одно собственное значение и соответствующий ему собственный вектор.

Доказательство. 1. Пусть $f(\lambda)$ — характеристический многочлен матрицы $\mathcal{A} = (a_{ij})$ преобразования A. Раскрыв этот определитель, получим, что $f(\lambda)$ — это многочлен степени $n \geqslant 1$ от λ. Из основной теоремы алгебры многочленов следует, что этот многочлен имеет корень λ_0, тогда система (67) имеет ненулевое решение — вектор x, который является собственным вектором. ■

Сформулируем *алгоритм поиска собственных векторов и значений.*

Требуется найти собственные векторы и собственные значения матрицы $\mathcal{A} = (a_{ij})$, т.е. $\lambda \in P$ и векторы $x \in P^n$ такие, что $\mathcal{A}x = \lambda x$ и $x \neq \theta$.

Начало.

1. Составляем характеристический многочлен матрицы $\mathcal{A}$:

$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{22} - \lambda \end{vmatrix}.$$

2. Решаем уравнение $f(\lambda) = 0$.

3. Если корней нет, то собственных значений и собственных векторов нет.

4. В противном случае находим корни $\lambda = \lambda_j$, $j = \overline{1,k}$, $1 \leqslant k \leqslant n$, которые являются собственными значениями.

5. Для $j = \overline{1,k}$ при $\lambda = \lambda_j$ записываем систему

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

Находим общее решение $x = (x_1, x_2, \ldots, x_n)^\top \neq (0, 0, \ldots, 0)^\top$. Каждый такой вектор является собственным вектором, соответствующим собственному значению $\lambda = \lambda_j$.

Конец. ■

2. Линейные пространства над произвольным полем

2.1. Линейное пространство. Простейшие свойства пространства. Подпространство. Критерий подпространства

Определение 1.

I. Пусть заданы множество $E \neq \emptyset$, элементы которого называем векторами, и поле P элементы, которого называем числами.

II. Пусть определены две операции:

а) сложение векторов — каждой паре векторов f, g сопоставлен вектор, называемый суммой f и g и обозначаемый $f + g \in E$,

б) умножение вектора на число —каждому вектору f и каждому числу α сопоставлен вектор, называемый произведением f на α и обозначаемый αf.

III. Множество E называется линейным пространством над полем P, если выполняются следующие аксиомы:

L1 $\forall f, g, h \in E \quad (f+g)+h = f+(g+h)$ *(ассоциативность сложения);*

L2 $\exists \theta \in E \colon \forall f \in E \quad f + \theta = f$ *(существование нулевого вектора);*

L3 $\forall f \in E\, \exists(-f) \in E\colon f+(-f) = \theta$ *(существование противоположного вектора);*

L4 $\forall f, g \in E \quad f + g = g + f$ *(коммутативность сложения);*

L5 $\forall \alpha,\, \beta \in P,\, f \in E \quad \alpha(\beta f) = (\alpha\beta)f$;

L6 $\forall f \in E \quad 1f = f$;

L7 $\forall \alpha,\, \beta \in P,\, f \in E \quad (\alpha + \beta)f = \alpha f + \beta f$ *(дистрибутивность умножения вектора на сумму чисел);*

L8 $\forall \alpha \in P,\, f, g \in E \quad \alpha(f + g) = \alpha f + \alpha g$ *(дистрибутивность умножения суммы векторов на число).*

IV. Если $P = \mathbb{R}\,(P = \mathbb{C})$, *то* E *называется вещественным (комплексным) линейным пространством.*

V. Операции а) и б) называются линейными операциями.

Замечания. *В силу аксиомы* L4 *выполнены следующие свойства.*

1) *Для любого вектора* $f \in E \quad \theta + f = f$.

2) *Если вектор* $(-f) \in E$ *является противоположным для* $f \in E$, *то* $(-f) + f = \theta$, *т.е.* $f = -(-f)$ *и вектор* f *является противоположным для вектора* $(-f)$.

Пример. Множество всех геометрических векторов $E = V_3$ с операциями сложения векторов и умножения векторов на вещественные числа является вещественным линейным пространством.

Теорема 2.1.1 (Простейшие свойства линейного пространства).

Для произвольного линейного пространства E *над произвольным полем* P *выполнены следующие свойства.*

1. *Нулевой вектор единственен.*
2. *Для любого вектора противоположный к нему вектор единственен.*
3. *Для любых векторов* $f, g, h \in E$ *из* $f + h = g + h$ *или* $h + f = h + g$ *следует* $f = g$.
4. *Для любых векторов* $f, g \in E$ *существует, и притом единственное, решение* $x = g + (-f)$ *уравнения* $x + f = g$; *это решение называется* **разностью** g *и* f *и обозначается* $g - f$. *Итак,* $g - f = g + (-f)$ *и эта формула определяет операцию* **вычитания**.

5. $\forall \alpha \in P, f, g \in E \quad \alpha(f-g) = \alpha f - \alpha g.$

6. $\forall \alpha, \beta \in P, f \in E \quad (\alpha - \beta) f = \alpha f - \beta f.$

7. $\forall\, \alpha \in P \quad \alpha\theta = \theta.$

8. $\forall\, f \in E \quad 0f = \theta.$

9. *Из равенства* $\alpha f = \theta$, $\alpha \in P$, $f \in E$ *следует* $\alpha = 0$ *или* $f = \theta$.

10. $\forall f \in E \quad (-f) = (-1)f$ (*противоположный вектор для* f *равен* $(-1)f$).

Доказательство.

1. Если θ_1 и θ_2 нулевые векторы, то $\theta_2 + \theta_1 = \theta_2$, т.к. θ_1 нулевой вектор и $\theta_1 + \theta_2 = \theta_1$, т.к. θ_2 нулевой вектор. а из аксиомы $L4$ имеем $\theta_2 + \theta_1 = \theta_1 + \theta_2$. Следовательно, $\theta_1 = \theta_2$.

2. Пусть g_1 и g_2 — противоположные векторы для f, т.е. $g_1 + f = \theta$, $f + g_2 = \theta$. Из ассоциативности сложения следует, что $(g_1 + f) + g_2 = g_1 + (f + g_2)$. Отсюда $\theta + g_2 = g_1 + \theta$ и $g_2 = g_1$.

3. Прибавим $(-h)$ к обеим частям равенства $f + h = g + h$, получим $(f + h) + (-h) = (g + h) + (-h) \Rightarrow f + (h + (-h)) = g + (h + (-h)) \Rightarrow f + \theta = g + \theta \Rightarrow f = g$.

4. Проверим, что $x = g + (-f)$ решение уравнения $x + f = g$. Действительно, $x + f = g + (-f) + f = g + \theta = g$.

Если x_1 и x_2 решения уравнения $x + f = g$, то $x_1 + f = g$ и $x_2 + f = g$. Тогда $x_1 + f = x_2 + f \Rightarrow x_1 = x_2$, т.е. решение единственно.

5. Для разности справедливо равенство $(f - g) + g = f$. Умножим обе части этого равенства на α, получим $\alpha((f - g) + g) = \alpha f$ и $\alpha(f - g) + \alpha g = \alpha f$. Тогда по определению разности получаем $\alpha(f - g) = \alpha f - \alpha g$.

6. Для разности чисел справедливо равенство $(\alpha - \beta) + \beta = \alpha$. Умножим вектор f на это число, получим $((\alpha - \beta) + \beta)f = \alpha f$ и $(\alpha - \beta)f + \beta f = \alpha f$. Тогда по определению разности получаем $(\alpha - \beta)f = \alpha f - \beta f$.

7. Из свойства 5 при $g = f$ получаем $\alpha(f - f) = \alpha f - \alpha f$. Отсюда следует равенство $\alpha\theta = \theta$.

8. Из свойства 6 при $\beta = \alpha$ получаем $(\alpha - \alpha)f = \alpha f - \alpha f$. Отсюда следует равенство $0f = \theta$.

9. Если $\alpha = 0$, то утверждение верно. При $\alpha \neq 0$ существует обратное число α^{-1}. Умножим равенство $\alpha f = \theta$ на α^{-1}, получим $\alpha^{-1}\alpha f = \alpha^{-1}\theta$. Отсюда следует равенство $1f = \theta$, т.е. $f = \theta$.

10. Сложим векторы $f = 1f$ и $(-1)f$. Получим $f+(-1)f = 1f+(-1)f = (1+(-1))f = 0f = \theta$, т.е. $f + (-1)f = \theta$. Следовательно, вектор $(-1)f$ является противоположным к f, т.е. $(-f) = (-1)f$.

Определение 2. *Пусть E — линейное пространство над полем P и L — непустое подмножество в E.*

L называется линейным подпространством линейного пространства E или просто подпространством E, если L само является линейным пространством над полем P относительно тех же операций как и E. ■

Теорема 2.1.2 (Критерий линейного подпространства).

Подмножество $L \neq \emptyset$ линейного пространства E над полем P является подпространством тогда и только тогда, когда L замкнуто относительно линейных операций, т.е.

а) $\forall f, g \in L \quad f + g \in L$;

б) $\forall f \in L, \alpha \in P \quad \alpha f \in L$.

Доказательство. I. Пусть L - подпространство E. Тогда L само является линейным пространством над полем P относительно тех же операций как и E, т. е. выполнены свойства а) и б).

II. Пусть в L выполнены свойства а) и б), следовательно, определены операции сложения векторов из L и умножения векторов из L на числа из P. Тогда в L выполнены аксиомы 1, 4, 5, 6, 7, 8, так как равенства в аксиомах выполнены в E, а $L \subset E$, причём L замкнуто относительно линейных операций.

Поскольку при $f \in L$ в силу свойства б) заключаем, что нулевой вектор $\theta = 0f \in L$, следовательно, аксиома 2 выполнена в L. Аналогично, для $f \in L$ в силу свойства б) заключаем, что противоположный вектор $(-1)f \in L$, и тогда аксиома 3 выполнена в L. ■

Задача 23. *Проверьте является ли линейным пространством над указанным полем P данное множество E с указанными операциями.*

а) $P = \mathbb{R}$, $E = \mathbb{R}^{m\times n}$ *— матрицы размера* $m \times n$ *из вещественных элементов, операции сложения матриц и умножения матриц на числа из* $\mathbb{R}$.

б) $P = \mathbb{C}$, $E = \mathbb{C}^{m\times n}$ *— матрицы размера* $m \times n$ *из комплексных элементов, операции сложения матриц и умножения матриц на числа из* $\mathbb{C}$.

в) $P = \mathbb{R}$, $E = \mathbb{R}[x]$ *— многочлены с действительными коэффициентами от действительной переменной* x, *с операциями сложения многочленов и умножения многочленов на числа из* $\mathbb{R}$.

г) $P = \mathbb{R}$, E *—рациональные функции или рациональные дроби вида* $\dfrac{f(x)}{g(x)}$, *где* $f(x) \in \mathbb{R}[x]$, $g(x) \in \mathbb{R}[x]$, $g(x) \not\equiv 0$, *операции умножения дробей на числа из* $\mathbb{R}$ *и сложения дробей* $\dfrac{f(x)}{g(x)} + \dfrac{h(x)}{k(x)} = \dfrac{f(x)k(x) + g(x)h(x)}{g(x)k(x)}$.

д) $P = \mathbb{R}$, $E = M[\mathbb{R}]$ *— функции, действующие из* $\mathbb{R}$ *в* $\mathbb{R}$, *операции умножения функций на числа из* $\mathbb{R}$ *и сложения функций.*

Задача 24. *С помощью критерия линейного подпространства (теорема 2.1.2) проверьте, является ли линейным подпространством в* E *данное множество* L.

а) $E = \mathbb{R}^{2\times 3}$, L *— подмножество матриц* $A = (a_{ij})$ *из* E, *у которых* $a_{11} = 0$ *и* $a_{21} = 0$.

б) $E = \mathbb{R}^{2\times 2}$, L *подмножество матриц* $A = (a_{ij})$ *из* E, *у которых* $a_{22} = 1$.

в) $E = \mathbb{R}^{2\times 2}$, L *подмножество матриц* $A = (a_{ij})$ *из* E, *у которых* $a_{11} + 2a_{22} = 0$, $2a_{12} - 3a_{21} = 0$.

г) $E = \mathbb{R}[x]$, L *— подмножество многочленов степени не выше* 2.

д) $E = \mathbb{R}[x]$, L *— подмножество многочленов степени равной* 2.

е) $E = \mathbb{R}[x]$, L *— подмножество многочленов, которые в точке* $x = 1$ *равны* 0.

ж) $E = \mathbb{R}[x]$, L *— подмножество многочленов, которые в точке* $x = 2$ *равны* 1.

2.2. Линейная комбинация векторов. Линейная зависимость и линейная независимость векторов. База и ранг системы векторов

В этом разделе будем рассматривать векторы линейного пространства E над полем P.

Определение 1. *Пусть заданы векторы* $e_1, e_2, \ldots, e_k$ *из* E *и числа* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *из* P. *Вектор* $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k = \sum\limits_{i=1}^{k} \alpha_i e_i$ *называется линейной комбинацией векторов* $e_1, e_2, \ldots, e_k$ *с коэффициентами* $\alpha_1, \alpha_2, \ldots, \alpha_k$.

Линейная комбинация называется тривиальной, если все её коэффициенты равны 0, *и нетривиальной, если существует хотя бы один ненулевой коэффициент.*

Если вектор $b = \sum\limits_{i=1}^{k} \alpha_i e_i$, *то говорят, что вектор* b *линейно выражается через векторы* $e_1, e_2, \ldots, e_k$, *и это равенство называют разложением вектора* b *по векторам* $e_1, e_2, \ldots, e_k$.

Замечания.

1. Тривиальная линейная комбинация равна нулевому вектору θ.

2. Линейная комбинация одного вектора e_1 с коэффициентом α_1 — это произведение $\alpha_1 e_1$.

Определение 2. *Система векторов* $e_1, e_2, \ldots, e_k$ *называется линейно зависимой, если найдётся нетривиальная линейная комбинация этих векторов, равная* θ, *т. е. существуют коэффициенты* $\alpha_1, \alpha_2, \ldots, \alpha_k$, *не все равные* 0, *и такие, что* $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$.

Система векторов $e_1, e_2, \ldots, e_k$ *называется линейно независимой, если только тривиальная линейная комбинация этих векторов равна* θ, *т. е. из условия* $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$ *следует* $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.[2]

Приведём свойства, связанные с линейной зависимостью или линейной независимостью векторов.

[2] Также говорят, что векторы $e_1, e_2, \ldots, e_k$ линейно зависимы (независимы) или $e_1, e_2, \ldots, e_k$ — линейно зависимые (независимые) векторы.

Теорема 2.2.1. 1° *При изменении порядка векторов их линейная зависимость (независимость) сохраняется.*

2° *а) Вектор* e_1 *является линейно зависимым (независимым) тогда и только тогда, когда он нулевой (ненулевой).*

б) Векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *являются линейно зависимыми тогда и только тогда, когда хотя бы один из них линейно выражается через остальные векторы, т.е. для некоторого* j *выполнено равенство* $e_j = \sum\limits_{i \neq j} \alpha_i e_i$.

в) Векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *являются линейно независимыми тогда и только тогда, когда ни один из них не является линейной комбинацией остальных векторов, т.е. для каждого* j *выполнено неравенство* $e_j \neq \sum\limits_{i \neq j} \alpha_i e_i$ *при любых числах* α_i.

3° а) *Если часть векторов линейно зависима, то и все векторы линейно зависимы.* б) *Если все векторы линейно независимы, то и любая их часть линейно независима.*

4° *Если векторы* $e_1, e_2, \ldots, e_k$ *линейно независимы, а векторы* $e_1, e_2, \ldots, e_k, x$ *линейно зависимы, то* x *линейно выражается через* $e_1, e_2, \ldots, e_k$.

Доказательство. 1° Если векторы $e_1, e_2, \ldots, e_k$ линейно зависимы, то существует нетривиальная линейная комбинация $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k$ равная θ. Так как при перестановки слагаемых сумма не меняется, то переставленный набор этих векторов образует линейно зависимую систему векторов. Следовательно, при изменении порядка векторов их линейная зависимость сохраняется.

Пусть данные векторы являются линейно независимыми. Докажем, что при изменении порядка векторов их линейная независимость сохраняется.

Предположим противное: переставленный набор этих векторов является линейно зависимым. Тогда, по доказанному выше, векторы в исходном порядке также являются линейно зависимыми, что противоречит условию. Следовательно, при изменении порядка векторов их линейная независимость сохраняется.

2° а) Если вектор e_1 линейно зависимый, то найдётся нетривиальная линейная комбинация $\alpha_1 e_1$ равная θ, т.е. $\alpha_1 e_1 = \theta$, где $\alpha_1 \neq 0$. Отсюда следует, что $e_1 = \theta$.

Обратно. Пусть вектор e_1 — нулевой, т.е. $e_1 = \theta$. Отсюда $1e_1 = \theta$, т.е. нетривиальная линейная комбинация $1e_1$ равна θ. Следовательно, вектор e_1

линейно зависимый.

2° б) Если векторы $e_1, e_2, \ldots, e_k$ линейно зависимы, то найдётся линейная комбинация $\sum\limits_{i=1}^{k} \alpha_i e_i$ равная θ, где некоторый коэффициент $\alpha_j \neq 0$. Тогда имеем $\alpha_j e_j + \sum\limits_{i \neq j} \alpha_i e_i = \theta$ и $\alpha_j e_j = \sum\limits_{i \neq j} (-\alpha_i) e_i = \theta$, откуда $e_j = \sum\limits_{i \neq j} (-\alpha_i/\alpha_j) e_i$, т.е. вектор e_j линейно выражается через остальные векторы.

Обратно. Пусть один из векторов $e_1, e_2, \ldots, e_k$ линейно выражается через остальные векторы: $e_j = \sum\limits_{i \neq j} \alpha_i e_i$, α_i — числа из P. Отсюда $1e_j + \sum\limits_{i \neq j} (-\alpha_i) e_i = \theta$, т.е. имеем нетривиальную линейную комбинацию, равную θ. Следовательно, векторы линейно зависимы.

2° в) эквивалентно 2° б).

3° а) Без ограничения общности считаем, что из m заданных векторов $e_1, \ldots, e_k, e_{k+1}, \ldots, e_m$ линейно зависимыми являются векторы $e_1, e_2, \ldots, e_k$. Тогда существует коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_k$, не все равные 0, и такие, что $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$. Добавим к этому равенству вектор $\sum\limits_{i=k+1}^{m} 0 e_i = \theta$, и получим нетривиальную линейную комбинацию всех векторов $\sum\limits_{i=1}^{k} \alpha_i e_i + \sum\limits_{i=k+1}^{m} 0 e_i$, равную θ. Следовательно, векторы $e_1, \ldots, e_k, \ldots, e_m$ линейно зависимы.

3° б) эквивалентно 3° а).

4° Из линейной зависимости векторов $e_1, \ldots, e_k, x$ следует, что существуют коэффициенты $\alpha_1, \ldots, \alpha_k, \alpha_{k+1}$, не все равные 0, и такие, что

$$\sum_{i=1}^{k} \alpha_i e_i + \alpha_{k+1} x = \theta. \tag{68}$$

Если $\alpha_{k+1} = \theta$, то из (68) имеем $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$, причём среди α_i, $i = \overline{1,k}$ есть ненулевой коэффициент. Тогда векторы $e_1, e_2 \ldots, e_k$ линейно зависимы вопреки условию. Следовательно, $\alpha_{k+1} \neq \theta$ и из (68) можно выразить вектор $x = \sum\limits_{i=1}^{k} \dfrac{-\alpha_i}{\alpha_{k+1}} e_i$. ■

Определение 3. *Базой системы векторов называется её линейно независимая подсистема (часть), через которую линейно выражается любой вектор системы.*

Рангом система векторов $e_1, e_2, \ldots, e_k$ называется число векторов её базы. Обозначение: $rg(e_1, e_2, \ldots, e_k)$.

Теорема 2.2.2. *Подсистема системы векторов образует базу тогда и только тогда, когда является максимальной линейно независимой подсистемой.*[3]

Доказательство. 2-я часть. Переставляя векторы, если нужно, считаем, что первые r векторов $e_1, e_2, \ldots, e_r$ образуют максимальную линейно независимую подсистему в системе $e_1, e_2, \ldots, e_r, \ldots, e_k$.

Покажем, что любой вектор e_j системы $e_1, e_2, \ldots, e_r, \ldots, e_k$ линейно выражается через векторы подсистемы $e_1, e_2, \ldots, e_r$.

При $j \leqslant r$ это очевидно: $e_j = e_j$.

При $j > r$ рассмотрим подсистему из $r+1$ векторов $e_1, e_2, \ldots, e_r, e_j$. Она линейно зависима, т.к. любая линейно независимая подсистема содержит не более r векторов. Кроме того, векторы $e_1, e_2, \ldots, e_r$ линейно независимы. Тогда из утверждения 4° теоремы 2.2.1 следует, что e_j линейно выражается через векторы $e_1, e_2, \ldots, e_r$.

Следовательно, подсистема $e_1, e_2, \ldots, e_r$ образует базу исходной системы $e_1, e_2, \ldots, e_r, \ldots, e_k$.

Следствие. *Все базы данной системы векторов содержат одинаковое число векторов, равное рангу этой системы.* ■

Задача 25. *Докажите следующие утверждения.*

а) Система векторов, содержащая нулевой вектор, линейно зависима.

б) Векторы линейно независимы тогда и только тогда, когда любой вектор, являющийся линейной комбинацией этих векторов, имеет единственное разложение по ним.

в) Ненулевые векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно зависимы тогда и только тогда, когда один из них линейно выражается через предыдущие, т.е.* $\exists j: \ e_j = \sum\limits_{i=1}^{j-1} \alpha_i e_i$.

г) Ненулевые векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно независимы тогда и только тогда, когда ни один из них линейно не выражается через предыдущие, т.е.* $\forall j > 1: \ e_j \neq \sum\limits_{i=1}^{j-1} \alpha_i e_i$ *при любых числах* $\alpha_1, \ldots, \alpha_{j-1}$.

[3] В данной системе векторов подсистема из r векторов называется максимальной линейно независимой подсистемой, если в этой системе нет линейно независимых подсистем более чем из r векторов.

д) Векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *являются линейно зависимыми тогда и только тогда, когда* $e_1 = \theta$ *или один из них линейно выражается через предыдущие, т.е.* $\exists j > 1: \ e_j = \sum\limits_{i=1}^{j-1} \alpha_i e_i$.

е) Векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *являются линейно независимыми тогда и только тогда, когда* $e_1 \neq \theta$ *и ни один из них линейно не выражается через предыдущие, т.е.* $\forall j > 1: \ e_j \neq \sum\limits_{i=1}^{j-1} \alpha_i e_i$ *при любых числах* $\alpha_1, \ldots, \alpha_{j-1}$.

2.3. Базис линейного пространства. Координаты вектора в базисе и их свойства

В этом разделе рассматривается линейное пространство E над полем P.

Определение 1. *Назовём базисом линейного пространства* E *упорядоченный набор векторов* $e_1, e_2, \ldots, e_n$ *таких, что:*

1) $e_1, e_2, \ldots, e_n$ *линейно независимы;*

2) *любой вектор* x *линейно выражается через* $e_1, e_2, \ldots, e_n$, *т. е. найдутся числа* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *такие, что*

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = \sum_{i=1}^{n} \alpha_i e_i, \tag{69}$$

при этом числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ *называются координатами вектора* x *в базисе (или относительно базиса)* $e = (e_1, e_2, \ldots, e_n)$.

Удобно обозначить координатный столбец вектора x так: $[x]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$,

тогда равенство $x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$ можно записать в виде

$$x = (e_1 \ e_2 \ \ldots \ e_n) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} \quad \text{или} \quad x = e[x]_e. \tag{70}$$

Единственность координат вектора в заданном базисе доказана ниже.

Отметим, что для нулевого вектора $x = \theta$ равенство (69) выполнено для $\alpha_i = 0$, $i = \overline{1, n}: \ \theta = \sum\limits_{i=1}^{n} 0 e_i$, т.е. нулевой вектор имеет нулевые координаты в любом базисе.

Для j-го вектора базиса e_j равенство (69) выполнено для $\alpha_i = 0$, $i \neq j$ и $\alpha_j = 1$: $e_j = 0e_1 + \ldots + 1e_j + \ldots + 0e_n$, т.е. все координаты вектора e_j относительно этого базиса равны 0, кроме j-й координаты, которая равна 1.

Теорема 2.3.1 (Свойства координат).

1° *Координаты любого вектора в фиксированном базисе определяются однозначно.*

2° *При сложении векторов складываются их соответствующие координаты*

$$\forall f_1,\, f_2,\, \ldots,\, f_k \in E \quad \left[\sum_{j=1}^{k} f_j\right]_e = \sum_{j=1}^{k} [f_j]_e.$$

3° *При умножении вектора на число все его координаты умножаются на это число*

$$\forall f \in E,\, \beta \in P \quad [\beta f]_e = \beta [f]_e.$$

4° *Каждая координата линейной комбинации векторов равна линейной комбинации соответствующих координат этих векторов с теми же коэффициентами*

$$\forall f_1,\, f_2,\, \ldots,\, f_k \in E,\, \beta_1,\, \beta_2,\, \ldots,\, \beta_k \in P \quad \left[\sum_{j=1}^{k} \beta_j f_j\right]_e = \sum_{j=1}^{k} \beta_j [f_j]_e.$$

5° *Векторы линейно зависимы (независимы) $\Leftrightarrow$ их координатные столбцы линейно зависимы (независимы).*[4]

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ — базис линейного пространства E.

1° Если $x = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $x = \sum\limits_{i=1}^{n} \alpha'_i e_i$, то $x - x = \sum\limits_{i=1}^{n} \alpha_i e_i - \sum\limits_{i=1}^{n} \alpha'_i e_i$ и тогда $\theta = \sum\limits_{i=1}^{n} (\alpha_i - \alpha'_i) e_i$. В силу линейной независимости векторов базиса отсюда следует, что все коэффициенты $(\alpha_i - \alpha'_i) = 0$, т.е. $\alpha_i = \alpha'_i$ для всех i.

2° Применим метод математической индукции.

1) Докажем утверждение при $k = 2$.

Пусть заданы векторы $f_1 = \sum\limits_{i=1}^{n} \alpha_{i1} e_i$ и $f_2 = \sum\limits_{i=1}^{n} \alpha_{i2} e_i$ с координатами $\alpha_{11},\, \alpha_{21},\, \ldots,\, \alpha_{n1}$, и $\alpha_{12},\, \alpha_{22},\, \ldots,\, \alpha_{n2}$. Тогда $f_1 + f_2 = \sum\limits_{i=1}^{n} \alpha_{i1} e_i + \sum\limits_{i=1}^{n} \alpha_{i2} e_i =$

[4] Мы рассматриваем координатные столбцы, состоящие из n чисел из поля P, как элементы P^n или $P^{n\times 1}$ — линейного пространства матриц-столбцов с операциями сложения матриц и умножения матриц на числа из поля P.

$= \sum\limits_{i=1}^{n}(\alpha_{i1}+\alpha_{i2})e_i$. Имеем i-е координаты $f_1,\ f_2,\ f_1+f_2$ — это соответственно $\alpha_{i1},\ \alpha_{i2},\ \alpha_{i1}+\alpha_{i2}$, т.е. при сложении двух векторов складываются их соответствующие координаты. Этот результат можно изобразить в виде координатных столбцов.

Координатные столбцы векторов $f_1,\ f_2,\ f_1+f_2$ соответственно равны:

$$[f_1]_e = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \dots \\ \alpha_{n1} \end{pmatrix},\ [f_2]_e = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \dots \\ \alpha_{n2} \end{pmatrix},\ [f_1+f_2]_e = \begin{pmatrix} \alpha_{11}+\alpha_{12} \\ \alpha_{21}+\alpha_{22} \\ \dots \\ \alpha_{n1}+\alpha_{n2} \end{pmatrix}.$$

Тогда $[f_1+f_2]_e = \begin{pmatrix} \alpha_{11}+\alpha_{12} \\ \alpha_{21}+\alpha_{22} \\ \dots \\ \alpha_{n1}+\alpha_{n2} \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \dots \\ \alpha_{n1} \end{pmatrix} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \dots \\ \alpha_{n2} \end{pmatrix} = [f_1]_e + [f_2]_e,$

поэтому $[f_1+f_2]_e = [f_1]_e + [f_2]_e$, то есть утверждение верно при $k=2$.

2) Предположим, что утверждение верно при $k=m$: $\left[\sum\limits_{j=1}^{m} f_j\right]_e = \sum\limits_{j=1}^{m}[f_j]_e$.

3) Докажем, что утверждение верно при $k=m+1$: $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \sum\limits_{j=1}^{m+1}[f_j]_e$.

Сумму векторов f_j запишем в виде суммы двух векторов $\sum\limits_{j=1}^{m+1} f_j = \sum\limits_{j=1}^{m} f_j + f_{m+1}$. Приравняем их координатные столбцы $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \left[\sum\limits_{j=1}^{m} f_j + f_{m+1}\right]_e$, и применим утверждение при $k=2$. Получим $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \left[\sum\limits_{j=1}^{m} f_j\right]_e + \left[f_{m+1}\right]_e$. Далее, используя предположение индукции $\left[\sum\limits_{j=1}^{m} f_j\right]_e = \sum\limits_{j=1}^{m}[f_j]_e$, запишем $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e =$ $= \sum\limits_{j=1}^{m}[f_j]_e + [f_{m+1}]_e = \sum\limits_{j=1}^{m+1}[f_j]_e$, т. е. $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \sum\limits_{j=1}^{m+1}[f_j]_e$, ч.т.д.

3° Пусть заданы вектор $f = \sum\limits_{i=1}^{n}\alpha_i e_i$ с координатами $\alpha_1, \alpha_2, \dots, \alpha_n$ и число $\beta \in P$. Тогда $\beta f = \beta\sum\limits_{i=1}^{n}\alpha_i e_i = \sum\limits_{i=1}^{n}(\beta\alpha_i)e_i$. Имеем i-е координаты f и βf — это соответственно α_i и $\beta\alpha_i$, т.е. при умножении вектора на число все

его координаты умножаются на это число. Этот результат можно изобразить в виде координатных столбцов: $[\beta f]_e = \beta [f]_e$.

4° Пусть заданы векторы $f_j \in E$, $j = \overline{1,k}$ и их линейная комбинация $\sum\limits_{j=1}^{k} \beta_j f_j$ с коэффициентами $\beta_j \in P$, $j = \overline{1,k}$. Запишем координатный столбец линейной комбинации: $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e$. К этому выражению применим сначала свойство 2°, а затем свойство 3° : $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = \sum\limits_{j=1}^{k} [\beta_j f_j]_e = \sum\limits_{j=1}^{k} \beta_j [f_j]_e$.

Следовательно, $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = \sum\limits_{j=1}^{k} \beta_j [f_j]_e$.

5° Докажем утверждение о линейной зависимости:
векторы $f_j \in E$, $j = \overline{1,k}$ *линейно зависимы* $\Leftrightarrow$ *координатные столбцы векторов* $[f_j]_e$, $j = \overline{1,k}$ *линейно зависимы.*

Пусть векторы f_j, $j = \overline{1,k}$ линейно зависимы. Тогда найдётся нетривиальная линейная комбинация векторов f_j, $j = \overline{1,k}$ равная нулевому вектору, т.е. существуют числа β_j, $j = \overline{1,k}$, среди которых есть $\beta_m \neq 0$, такие, что выполнено равенство $\sum\limits_{j=1}^{k} \beta_j f_j = \theta$. Поскольку равенство векторов эквивалентно равенству координатных столбцов векторов, то для тех же чисел β_j выполнено равенство $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = [\theta]_e$, и, в силу утверждения 4°, справедливо равенство $\sum\limits_{j=1}^{k} \beta_j [f_j]_e = [\theta]_e$ для координатных столбцов векторов. Мы получили нетривиальную линейную комбинацию координатных столбцов векторов, равную нулевому столбцу. Следовательно, координатные столбцы векторов $[f_j]_e$, $j = \overline{1,k}$ линейно зависимы.

Далее проводим рассуждения последнего абзаца в обратном порядке и получим доказательство обратного утверждения.

Пусть координатные столбцы векторов $[f_j]_e$, $j = \overline{1,k}$ линейно зависимы. Тогда найдётся нетривиальная линейная комбинация координатных столбцов

$[f_j]_e$, $j = \overline{1,k}$ равная нулевому столбцу $\begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} = [\theta]_e$, т.е. существуют числа β_j, $j = \overline{1,k}$, среди которых есть $\beta_m \neq 0$, такие, что выполнено равенство $\sum\limits_{j=1}^{k} \beta_j [f_j]_e = [\theta]_e$. Отсюда получаем $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = [\theta]_e$ и далее $\sum\limits_{j=1}^{k} \beta_j f_j = \theta$. Следовательно, нетривиальная линейная комбинация векторов f_j, $j = \overline{1,k}$ равна нулевому вектору, и векторы f_j, $j = \overline{1,k}$ линейно зависимы. ■

2.4. Размерность линейного пространства. Связь понятий базиса и размерности

В этом разделе рассматриваются линейные пространства над полем P.

Определение 1. *Размерностью линейного пространства E называется максимальное число линейно независимых векторов в E. Обозначение: $dim(E)$ или $dim\, E$.*

1) Размерность нулевого пространства $E = \{\theta\}$ считается равной 0 (в нём нет линейно независимых векторов).

2) Число $n \in N$ является размерностью ненулевого пространства, если в нём найдутся n линейно независимых векторов, а любые $n+1$ векторов линейно зависимы.

3) Размерность считается равной ∞, если в пространстве для любого $n \in N$ найдутся n линейно независимых векторов.

Мы говорим, что в случаях 1) и 2) E — n-мерное пространство и E — конечномерное пространство, в случае 3) E — бесконечномерное пространство.

Выясним связь понятий базиса и размерности линейного пространства.

Теорема 2.4.1. 1° *Если размерность линейного пространства E равна $n\,(1 \leqslant n < \infty)$, то в E существует базис из n векторов и любые n линейно независимых векторов из E образуют в нём базис.*

2° *Если некоторый базис линейного пространства* E *состоит из* n *векторов, то* $dim(E) = n$.

Доказательство. 1° Пусть $dim(E) = n$. Из определения размерности следует, что в E существуют n линейно независимых векторов. Докажем, что n произвольных линейно независимых векторов $e_1, e_2, \ldots, e_n$ образуют базис в E. Добавим к ним любой вектор $x \in E$. Из определению размерности следует, что $n+1$ векторов $e_1, e_2, \ldots, e_n, x$ линейно зависимы. В силу теоремы 2.2.1 вектор x линейно выражается через $e_1, e_2, \ldots, e_n$. Следовательно, $e_1, e_2, \ldots, e_n$ — базис в E.

2° Пусть $e = (e_1, e_2, \ldots, e_n)$ — базис в E. Тогда эти n векторов линейно независимы и, следовательно, в E существуют n линейно независимых векторов.

Докажем, что произвольные $n+1$ векторов $f_1, f_2, \ldots, f_{n+1} \in E$ линейно зависимы. Разложим каждый вектор f_j, $j = \overline{1, n+1}$ по базису e : $f_j = \sum\limits_{i=1}^{n} a_{ij} e_i$ и из координатных столбцов $[f_j]_e$ образуем $n \times (n+1)$-матрицу $A = (a_{ij})$. Ранг r_A этой матрицы не больше числа строк n и, следовательно, меньше числа столбцов $n+1$. Из теоремы о базисном миноре следует, что максимальное число её линейно независимых столбцов равно рангу матрицы. Так как $r_A < n+1$, что все $n+1$ столбцов матрицы A линейно зависимы. Следовательно, координатные столбцы $[f_j]_e$, $j = \overline{1, n+1}$ линейно зависимы и по теореме 2.3.1 векторы $f_1, f_2, \ldots, f_{n+1}$ линейно зависимы.

Итак, в E существуют n линейно независимых векторов, а любые $n+1$ векторов линейно зависимы, тогда $dim(E) = n$. ■

Следствия. 1. *Если размерность линейного пространства* E *равна* n $(1 \leqslant n < \infty)$, *то любой базис в* E *состоит из* n *векторов.*

2. $dim(P^n) = n$.

Доказательство. 1. Предположим, что некоторый базис в E содержит k векторов и $k \neq n$. Тогда из утверждения 2° теоремы 2.4.1 следует, что $dim(E) = k$ и $k \neq n$, а это противоречит условию $dim(E) = n$. Следовательно, любой базис в E состоит из n векторов.

2. Базис в P^n образуют n векторов, у которых один элемент равен 1, а остальные элементы равны 0. ■

Теорема 2.4.2. *Если размерность линейного пространства E равна n $(1 \leqslant n < \infty)$, то при любом k меньшем n произвольные k линейно независимых векторов из E можно дополнить до базиса.*

Доказательство.

Пусть $e_1, e_2, \ldots, e_k\,(k < n)$ — данные линейно независимые векторы. Если бы любой вектор линейно выражался через данные векторы, то они образовывали бы базис и тогда $dim(E) = k$, что противоречит условию $dim(E) = n$. Значит, существует вектор e_{k+1}, который линейно не выражается через данные векторы.

Докажем, что векторы $e_1, e_2, \ldots, e_k, e_{k+1}$ линейно независимы. Предположим противное: векторы $e_1, e_2, \ldots, e_k, e_{k+1}$ линейно зависимы. Тогда из утверждения 4° теоремы 2.2.1 следует, что вектор e_{k+1} линейно выражается через векторы $e_1, e_2, \ldots, e_k$, но вектор e_{k+1} линейно не выражается через эти векторы. Следовательно, векторы $e_1, e_2, \ldots, e_k, e_{k+1}$ линейно независимы. Если $k + 1 = n$, то эта система из $k + 1$ векторов образует базис, если же $k + 1 < n$, то аналогичным образом построим линейно независимую систему из $k + 2$ векторов и т. д. За $n - k$ таких шагов построим линейно независимую систему из n векторов $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$, которая образует базис в n-мерном пространстве E. ■

Задача 26. *а) Найдите размерность линейных пространств и линейных подпространств из задач 23 и 24 на стр. 92–93.*

б) Найдите базис линейных пространств и линейных подпространств из задач 23 и 24 на стр. 92–93, если он существует.

2.5. Матрица перехода от одного базиса к другому. Связь координат вектора в различных базисах

В этом разделе рассматриваются различные базисы в n-мерном линейном пространстве над полем P и выясняется связь координат вектора в различных базисах.

Пусть в n-мерном линейном пространстве E над полем P заданы два базиса (71) — «старый» и (72) — «новый»:

$$e = (e_1,\, e_2,\, \ldots,\, e_n) \quad \text{«старый» и} \tag{71}$$

$$e' = (e'_1,\, e'_2,\, \ldots,\, e'_n) \quad \text{«новый».} \tag{72}$$

Разложим векторы нового базиса e' по векторам старого базиса e :

$$\begin{gathered} e'_1 = t_{11}e_1 + t_{21}e_2 + \ldots + t_{n1}e_n, \\ e'_2 = t_{12}e_1 + t_{22}e_2 + \ldots + t_{n12}e_n, \\ \ldots \\ e'_n = t_{1n}e_1 + t_{2n}e_2 + \ldots + t_{nn}e_n, \\ \text{или в общем виде } e'_j = \sum_{i=1}^{n} t_{ij}e_i,\ j = \overline{1,n}, \end{gathered} \tag{73}$$

и получим $[e'_j]_e = \begin{pmatrix} t_{1j} \\ t_{2j} \\ \ldots \\ t_{nj} \end{pmatrix}$ — координатные столбцы векторов e'_j в базисе e.

Из координатных столбцов $[e'_1]_e,\ [e'_2]_e,\ \ldots,\ [e'_n]_e$ составим матрицу

$$\begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}. \tag{74}$$

Определение 1. *Матрица* (74) *называется матрицей перехода от базиса* e *к базису* e'.

Обозначение: T *или* $T_{e\to e'}$.

Итак, $T_{e\to e'} = \Big([e'_1]_e\ [e'_2]_e\ \ldots\ [e'_n]_e\Big)$ *— матрица перехода от базиса* e *к базису* e'.

Замечание. Формулы (73) могут быть записаны в виде произведения

$$\begin{gathered} (e'_1,\, e'_2,\, \ldots,\, e'_n) = (e_1,\, e_2,\, \ldots,\, e_n) \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix} \\ \text{или } e' = eT_{e\to e'}. \end{gathered} \tag{75}$$

Сформулируем и докажем свойства матриц перехода.

Теорема 2.5.1 (Свойства матриц перехода).

1° *Матрица перехода невырождена.*

2° *Если T матрица перехода от базиса e к базису e', то обратная матрица T^{-1} является матрицей перехода от базиса e' к базису e, т.е. $T^{-1}_{e\to e'} = T_{e'\to e}$.*

3° *Произвольная невырожденная матрица из $P^{n\times n}$ является матрицей перехода от одного базиса к другому.*

Доказательство.

1° Матрица перехода $T_{e\to e'}$ состоит из координатных столбцов $[e'_1]_e$, $[e'_2]_e$, $\ldots$, $[e'_n]_e$ векторов базиса. Поскольку векторы базиса линейно независимы, то и их координатные столбцы линейно независимы. Поэтому определитель матрицы перехода не равен нулю и, следовательно, матрица перехода невырождена.

2° Воспользуемся равенством (75), и умножим обе части равенства на обратную матрицу $T^{-1}_{e\to e'}$. Получим $e'T^{-1}_{e\to e'} = eT_{e\to e'}T^{-1}_{e\to e'}$. Поскольку произведение $T_{e\to e'}T^{-1}_{e\to e'} = I_n$, где I_n — единичная матрица, получим $e'T^{-1}_{e\to e'} = eI_n$ или $e'T^{-1}_{e\to e'} = e$. Следовательно, $e = e'T^{-1}_{e\to e'}$ и, сравнив это равенство с (75), получаем, что $T^{-1}_{e\to e'}$ — матрица перехода от базиса e' к базису e.

3° Пусть $T = (t_{ij}) \in P^{n\times n}$ — произвольная невырожденная матрица и $e = (e_1, e_2, \ldots, e_n)$ — произвольный базис в пространстве E. По формулам (73) образуем векторы $e' = (e'_1, e'_2, \ldots, e'_n)$. Матрица T невырождена, поэтому её столбцы линейно независимы, и по построению эти столбцы являются координатными столбцами векторов $e'_1, e'_2, \ldots, e'_n$ в базисе e. Следовательно, векторы $e'_1, e'_2, \ldots, e'_n$ линейно независимы в n-мерном пространстве и тогда они образуют в нём базис e'. Тогда формулы (73) показывают, что T является матрицей перехода от e к базису e'. ■

Рассмотрим базисы (71) и (72) с матрицей перехода (74). Найдём связь между координатами (координатными столбцами) произвольного вектора x в этих базисах.

Теорема 2.5.2 (Связь координат вектора в двух базисах).

Координаты вектора x в базисах e и e' связаны эквивалентными равенствами

$$[x]_e = T_{e\to e'}\,[x]_{e'}\,, \tag{76}$$

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \dots & \dots & \dots & \dots \\ t_{n1} & t_{n2} & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \dots \\ \alpha'_n \end{pmatrix}, \tag{76'}$$

$$\begin{gathered} \alpha_1 = t_{11}\alpha'_1 + t_{12}\alpha'_2 + \ldots + t_{1n}\alpha'_n, \\ \alpha_2 = t_{21}\alpha'_1 + t_{22}\alpha'_2 + \ldots + t_{2n}\alpha'_n, \\ \dots \\ \alpha_n = t_{n1}\alpha'_1 + t_{n2}\alpha'_2 + \ldots + t_{nn}\alpha'_n, \end{gathered} \tag{76''}$$

где $[x]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}$, $[x]_{e'} = \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \dots \\ \alpha'_n \end{pmatrix}$ *— координатные столбцы вектора* x *в базисах* e *и* e' *соответственно,* $T_{e\to e'}$ *— матрица перехода* (74) *от базиса* e *к базису* e'.

Доказательство. Из формулы (70) на стр. 98 следуют два разложения $x = e[x]_e$ и $x = e'[x]_{e'}$ вектора x по базису e и e' соответственно. Подставим во второе равенство выражение (75) базиса e' через базис e, и получим $x = e'[x]_{e'} = (eT_{e\to e'})[x]_{e'} = e(T_{e\to e'}[x]_{e'})$. Итак, получено два разложения вектора x по базису e: $x = e[x]_e$ и $x = e(T_{e\to e'}[x]_{e'})$. В силу единственности координат выполнено равенство координатных столбцов $[x]_e = T_{e\to e'}[x]_{e'}$. Отсюда получаются подробные координатные равенства (76′) и (76″). ■

Равенства (76) – (76)″ называют *формулы преобразования координат.*

Задача 27. *Как изменится матрица перехода от первого ко второму базису, если 1) в первом базисе поменять местами два вектора, 2) во втором базисе поменять местами два вектора, 3) вместо первого вектора первого базиса взять удвоенный вектор первого базиса, 4) заменить первый базис на второй, а второй — на первый.*

2.6. Линейная оболочка векторов

В этом разделе рассматривается линейное пространство E над полем P и указан один из способов построения подпространств этого пространства.

Определение 1. *Линейной оболочкой системы векторов* $e_1, e_2, \ldots, e_k$ *из* E *называется объединение всех линейных комбинаций этих векторов.*

Обозначение: $Lin(e_1, e_2, \ldots, e_k)$.

Итак, $Lin(e_1, e_2, \ldots, e_k) = \{x \mid (\exists\, \alpha_i \in P,\ i = \overline{1,k})\ x = \sum\limits_{i=1}^{k} \alpha_i e_i\}$.

Теорема 2.6.1.

1° *Линейная оболочка* $Lin(e_1, e_2, \ldots, e_k)$ *является подпространством в* E.

2° *Базисом подпространства* $Lin(e_1, e_2, \ldots, e_k) \neq \{\theta\}$ *является любая база системы векторов* $e_1, e_2, \ldots, e_k$.

Доказательство.

1° Очевидно, $Lin(e_1, e_2, \ldots, e_k)$ — это непустое подмножество пространства E. Применим к $Lin(e_1, e_2, \ldots, e_k)$ критерий подпространства.

1) Если $x, y \in Lin(e_1, e_2, \ldots, e_k)$, то $x = \sum\limits_{i=1}^{k} \alpha_i e_i$, $y = \sum\limits_{i=1}^{k} \beta_i e_i$, $\alpha_i, \beta_i \in P$, $i = \overline{1,k}$, тогда $x + y = \sum\limits_{i=1}^{k} (\alpha_i + \beta_i) e_i$, $\alpha_i + \beta_i \in P$, $i = \overline{1,k}$, то есть $x + y \in$ $\in Lin(e_1, e_2, \ldots, e_k)$.

2) Если $x \in Lin(e_1, e_2, \ldots, e_k)$ и $\gamma \in P$, то $x = \sum\limits_{i=1}^{k} \alpha_i e_i$, $\alpha_i \in P, i = \overline{1,k}$, тогда $\gamma x = \sum\limits_{i=1}^{k} (\gamma \alpha_i) e_i$, $\gamma \alpha_i \in P$, $i = \overline{1,k}$, то есть $\gamma x \in Lin(e_1, e_2, \ldots, e_k)$.

Из 1) и 2) следует, что линейная оболочка $Lin(e_1, e_2, \ldots, e_k)$ является подпространством в E.

2° Если векторы $f_1, f_2, \ldots, f_r$ образуют базу системы векторов $e_1, e_2, \ldots, e_k$, то векторы $f_1, f_2, \ldots, f_r$ линейно независимы и каждый вектор e_j линейно выражается через векторы базы. Любой вектор x линейной оболочки $Lin(e_1, e_2, \ldots, e_k)$ линейно выражается через векторы $e_1, e_2, \ldots, e_k$ и, следовательно, через векторы $f_1, f_2, \ldots, f_r$. Итак, векторы базы линейно независимы и любой вектор линейной оболочки $Lin(e_1, e_2, \ldots, e_k)$ линейно выражается через векторы базы, т.е. векторы базы образуют базис линейной оболочки. ■

Теорема 2.6.2 (Свойства размерности). *Пусть* L_1, L_2 *— линейные подпространства линейного пространства* E.

1° *Если* $L_1 \subseteq L_2$, *то* $dim(L_1) \leqslant dim(L_2)$ *(монотонность размерностей).*

2° *Если* $L_1 \subseteq L_2$ *и* $dim(L_1) = dim(L_2) < \infty$, *то* $L_1 = L_2$.

Доказательство. 1° Любая линейно независимая система векторов из L_1 в силу $L_1 \subseteq L_2$ является линейно независимой системой в L_2. Поэтому неравенство $dim(L_1) > dim(L_2)$ невозможно и, следовательно, $dim(L_1) \leqslant dim(L_2)$.

2° При $dim(L_1) = dim(L_2) = 0$ оба подпространства нулевые и утверждение $L_1 = L_2$ верно.

При $dim(L_1) = dim(L_2) = n > 0$ выберем n линейно независимых векторов $e_1, e_2, \ldots, e_n$ из L_1. Поскольку $dim(L_1) = n$ эти векторы образуют базис в L_1. В силу $L_1 \subseteq L_2$ и $dim(L_2) = n$ эти n линейно независимых векторов принадлежат L_2 и, следовательно, они образуют базис в L_2. Используя полученные факты, докажем включение $L_2 \subseteq L_1$. Возьмём любой вектор x из L_2 и разложим его по базису: $x = \sum\limits_{i=1}^{n} \alpha_i e_i$, $\alpha_i \in P$, $i = \overline{1, n}$. Все векторы e_i, $i = \overline{1, n}$ принадлежат L_1, и тогда их линейная комбинация $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in L_1$. Итак, доказано включение $L_2 \subseteq L_1$, что вместе с условием $L_1 \subseteq L_2$ даёт равенство $L_1 = L_2$. ■

Следствие. *Линейная оболочка базиса пространства совпадает с пространством.*

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ базис пространства E и тогда $dim(E) = n$. Из утверждений теоремы 2.6.1 следует, что линейная оболочка $L = Lin(e_1, e_2, \ldots, e_n)$ является подпространством в E и базисом L является базис пространства, поэтому $dim(L) = n$. Итак, $L \subseteq E$, $dim(L) = dim(E)$ и тогда из утверждения 2° теоремы 2.6.2 получаем $L = E$. ■

Задача 28. *Докажите, что если для множеств векторов выполнено включение* $\{e_1, e_2, \ldots, e_k\} \subseteq \{g_1, g_2, \ldots, g_m\}$, *то для линейных оболочек этих множеств выполнено включение* $Lin(e_1, e_2, \ldots, e_k) \subseteq Lin(g_1, g_2, \ldots, g_m)$.

Задача 29. *Докажите, что линейная оболочка системы векторов совпадает с линейной оболочкой её базы.*

2.7. Сумма и пересечение подпространств

В линейном пространстве E над полем P рассмотрим операции пересечения множеств и сложения множеств[5]. С помощь этих операций можно по заданным подпространствам построить другие подпространства.

Определение 1.

1. *Пусть* $L_1, L_2, \ldots, L_k$ *— непустые множества линейного пространства* E. *Суммой этих множеств называется множество всевозможных векторов* x, *представимых в виде*

$$x = x_1 + x_2 + \ldots + x_k, \tag{77}$$

где $x_i \in L_i$, $i = \overline{1,k}$.

Обозначение: $L_1 + L_2 + \ldots + L_k$ *или* $\sum\limits_{i=1}^{k} L_k$.

$$\textit{Итак,} \quad \sum_{i=1}^{k} L_k = \left\{x \,|\, x = x_1 + x_2 + \ldots + x_k,\, x_i \in L_i,\, i = \overline{1,k}\right\}.$$

Представление (77) *называется разложением вектора* x *по множествам* $L_1, L_2, \ldots, L_k$.

2. *Пересечением заданных множеств* L_δ, $\delta \in \Phi$ *называется множество, состоящее из всех таких элементов, которые принадлежат всем заданным множествам* L_δ. *Если таких элементов не найдётся, то пересечением считается пустое множество. Здесь* Φ *— необязательно конечное множество индексов.*

Обозначение: $\bigcap\limits_{\delta \in \Phi} L_\delta$.

В этом разделе будут использоваться сумма и пересечение подпространств линейного пространства E.

Теорема 2.7.1.

1° *Сумма подпространств является подпространством в* E.

2° *Пересечение любого множества подпространств является подпространством в* E.

Доказательство. Нулевой вектор θ принадлежит всем заданным подпространствам и, следовательно, θ принадлежит сумме и пересечению подпространств. Далее используем критерий подпространства.

[5] Не путать её с операцией объединения множеств.

1° Пусть $L_1, L_2, \ldots, L_k$ подпространства в E и $L = \sum\limits_{i=1}^{k} L_k$.

1) Для $x, y \in L$ проверим, что $x + y \in L$.

Если $x, y \in L$, то выполнены условия $x = \sum\limits_{i=1}^{k} x_i$, $y = \sum\limits_{i=1}^{k} y_i$ и при этом $x_i \in L_i$, $y_i \in L_i$, $i = \overline{1,k}$. Тогда для вектора $x+y$ получаем разложение $x+y = \sum\limits_{i=1}^{k} x_i + \sum\limits_{i=1}^{k} y_i = \sum\limits_{i=1}^{k}(x_i + y_i)$ и при этом $x_i + y_i \in L_i$, $i = \overline{1,k}$, поскольку L_i — подпространство. Следовательно, $x + y \in L$.

2) Для $x \in L$, $\alpha \in P$ проверим, что $\alpha x \in L$.

Если $x \in L$, то выполнены условия $x = \sum\limits_{i=1}^{k} x_i$ и при этом $x_i \in L_i$, $i = \overline{1,k}$. Тогда для вектора αx получаем разложение $\alpha x = \alpha \sum\limits_{i=1}^{k} x_i = \sum\limits_{i=1}^{k} \alpha x_i$ и при этом $\alpha x_i \in L_i$, $i = \overline{1,k}$, поскольку L_i — подпространство. Следовательно, $\alpha x \in L$.

Из 1) и 2) следует, что $L = \sum\limits_{i=1}^{k} L_k$ — подпространство.

2° Пусть L_δ подпространство в E при каждом $\delta \in \Phi$ и $L = \bigcap\limits_{\delta \in \Phi} L_\delta$.

1) Для $x, y \in L$ проверим, что $x + y \in L$.

Если $x, y \in L$, то выполнены условия $x, y \in L_\delta$ при всех $\delta \in \Phi$. При каждом $\delta \in \Phi$ L_δ — подпространство, поэтому вектор $x + y \in L_\delta$ при каждом $\delta \in \Phi$. Следовательно, $x + y$ принадлежит пересечению всех подпространств L_δ, т.е. $x + y \in L$.

2) Если $x \in L$, $\alpha \in P$, то выполнены условия $x \in L_\delta$ при всех $\delta \in \Phi$. При каждом $\delta \in \Phi$ L_δ — подпространство, поэтому вектор $\alpha x \in L_\delta$ при каждом $\delta \in \Phi$. Следовательно, αx принадлежит пересечению всех подпространств L_δ, т.е. $\alpha x \in L$.

Из 1) и 2) следует, что $L = \bigcap\limits_{\delta \in \Phi} L_\delta$ — подпространство. ■

Теорема 2.7.2.

Базисом суммы нескольких ненулевых конечномерных подпространств является любая база системы векторов, полученной объединением векторов базисов этих подпространств.

Доказательство. Пусть подпространство L есть сумма подпространств L_i, $i = \overline{1,k}$. Обозначим через F_i, $i = \overline{1,k}$ множество векторов базиса подпространства L_i и через F объединение векторов всех базисов.

Докажем, что подпространство L есть линейная оболочка векторов F, т.е. $L = Lin(F)$.

Сначала покажем, что верно включение $L \subseteq Lin(F)$, т.е. любой вектор x из L есть линейная комбинация векторов из F.

Действительно, если $x \in L$, то выполнены условия $x = \sum\limits_{i=1}^{k} x_i$ и при этом $x_i \in L_i$, $i = \overline{1,k}$. Вектор x_i есть линейная комбинация векторов базиса F_i, тогда при сложении получим, что $x = \sum\limits_{i=1}^{k} x_i$ есть линейная комбинация векторов из объединения всех базисов, т.е из F.

Покажем, что верно и включение $Lin(F) \subseteq L$, т.е. любой вектор x, который равен линейной комбинации векторов из F, принадлежит L.

Пусть вектор x равен линейной комбинации векторов из F. Тогда в его разложении по векторам из F, сгруппируем слагаемые с векторами из F_1 и обозначим их сумму x_1, слагаемые с векторами из F_2 и обозначим их сумму x_2, $\ldots$, слагаемые с векторами из F_k и обозначим их сумму x_k. Получим, разложение $x = x_1 + x_2 + \ldots + x_k$, где $x_i \in L_i$, $i = \overline{1,k}$, так как x_i разложено по базису подпространства L_i.

Далее из теоремы 2.6.1 получаем, что базисом $L = Lin(F)$ является база системы векторов F. ■

Следствие. *Пусть L_1 и L_2 ненулевые конечномерные подпространства линейного пространства E и $L_1 \cap L_2 = \{\theta\}$. Тогда объединение векторов базисов этих подпространств является базисом суммы $L_1 + L_2$.*

Доказательство. Пусть $e_1, e_2, \ldots, e_m$ и $f_1, f_2, \ldots, f_k$ — базисы L_1 и L_2 соответственно.

Покажем, что объединение базисов

$$e_1, e_2, \ldots, e_m, f_1, f_2, \ldots, f_k \tag{78}$$

является базисом $L_1 + L_2$.

Проверим линейную независимость векторов (78). Пусть линейная комбинация векторов (78) равна θ : $\sum\limits_{i=1}^{m} \alpha_i e_i + \sum\limits_{i=1}^{k} \beta_i f_i = \theta$. Тогда $\sum\limits_{i=1}^{m} \alpha_i e_i = -\sum\limits_{i=1}^{k} \beta_i f_i$ и, обозначив $h = \sum\limits_{i=1}^{m} \alpha_i e_i = -\sum\limits_{i=1}^{k} \beta_i f_i$, получаем, что $h \in L_1$ и $h \in L_2$, т.е. $h \in L_1 \cap L_2$. Из условия $L_1 \cap L_2 = \{\theta\}$ следует, что $h = \theta$. Тогда $\sum\limits_{i=1}^{m} \alpha_i e_i = \theta$,

$\sum\limits_{i=1}^{k} \beta_i f_i = \theta$ и из линейной независимости векторов каждого базиса получаем, что все коэффициенты α_i и β_i равны нулю. Итак, векторы (78) линейно независимы и по теореме 2.7.2 образуют базис $L_1 + L_2$. ■

Задача 30. *а) Докажите, что $L_1 \subseteq L_1 + L_2$ для любых подпространств L_1 и L_2.*

б) Пусть L_1, L_2, M_1, M_2 подпространства линейного пространства E и $L_1 \subseteq L_2$, $M_1 \subseteq M_2$. Как связаны подпространства $L_1 + L_2$ и $M_1 + M_2$.

в) Докажите, что

$$Lin(e_1, e_2, \ldots, e_k) + Lin(g_1, g_2, \ldots, g_m) = Lin(e_1, e_2, \ldots, e_k, g_1, g_2, \ldots, g_m).$$

2.8. Размерность суммы двух подпространств

В этом разделе продолжим изучать подпространства линейного пространства E над полем P. Здесь формулируется и доказывается теорема о размерности суммы двух подпространств.

Теорема 2.8.1.

Если L_1 и L_2 любые ненулевые конечномерные подпространства линейного пространства E, то размерность суммы подпространств равна сумме размерностей этих подпространств минус размерность их пересечения, т.е.

$$dim(L_1 + L_2) = dim(L_1) + dim(L_2) - dim(L_1 \cap L_2). \qquad (79)$$

Доказательство. Из определений пересечения и суммы подпространств следует, что $L_1 \cap L_2 \subseteq L_1 \subseteq L_1 + L_2$, $L_1 \cap L_2 \subseteq L_2 \subseteq L_1 + L_2$. Тогда из теорем 2.6.2, 2.7.1 и 2.7.2 получаем, что все указанные множества являются конечномерными подпространствами линейного пространства E.

Если $L_1 \cap L_2 = \{\theta\}$, то $dim(L_1 \cap L_2) = 0$. При этом из следствия теоремы 2.7.2 следует, что $dim(L_1 + L_2) = dim(L_1) + dim(L_2)$, т.е. равенство (79) выполнено.

Рассмотрим случай $L_1 \cap L_2 \neq \{\theta\}$.

При $L_1 = L_1 \cap L_2$ (или аналогично при $L_2 = L_1 \cap L_2$) имеем $L_1 \subseteq L_2$. Сумма подпространств $L_1 + L_2$ состоит из всех векторов $x = x_1 + x_2$, где $x_1 \in L_1$, $x_2 \in L_2$, но в силу $L_1 \subseteq L_2$ вектор $x_1 \in L_2$ и $x = x_1 + x_2 \in L_2$,

тогда $L_1 + L_2 \subseteq L_2$. Поскольку верно и включение $L_2 \subseteq L_1 + L_2$, то получаем $L_2 = L_1 + L_2$. Тогда формула (79) принимает вид $dim(L_2) = dim(L_1 \cap L_2) + dim(L_2) - dim(L_1 \cap L_2)$ и верна.

Остаётся рассмотреть наиболее сложный случай $L_1 \cap L_2 \neq \{\theta\}$, $L_1 \neq \neq L_1 \cap L_2$, $L_2 \neq L_1 \cap L_2$.

Возьмём базис $h_0 = (e_1, e_2, \ldots, e_m)$ в $L_1 \cap L_2$. Тогда эти векторы линейно независимы и в силу $L_1 \cap L_2 \subseteq L_1$ принадлежат L_1. Так как $L_1 \cap L_2 \neq L_1$, векторы из h_0 не образуют базис подпространства L_1, но их можно дополнить до базиса следующим образом. Существует вектор $f_1 \in L_1$, который линейно не выражается через векторы из h_0 и тогда векторы $h_1 = (e_1, e_2, \ldots, e_m, f_1)$ линейно независимы. Если любой вектор из L_1 линейно выражается через векторы из h_1, то h_1 — базис L_1 и процесс закончен. В противном случае существует вектор $f_2 \in L_1$, который линейно не выражается через векторы из h_1. Присоединим вектор f_2 к системе h_1 и получим линейно независимые векторы $h_2 = (e_1, e_2, \ldots, e_m, f_1, f_2)$. Если любой вектор из L_1 линейно выражается через векторы из h_2, то h_2 — базис L_1 и процесс закончен. В противном случае, дополним систему векторов h_2 и т.д. За конечное число шагов $k = dim(L_1) - dim(L_1 \cap L_2)$ получим в L_1 базис $h_k = (e_1, e_2, \ldots, e_m, f_1 f_2, \ldots, f_k)$.

Так как $h_0 \subseteq L_1 \cap L_2 \subseteq L_2$, аналогично дополним векторы h_0 до базиса $p_n = (e_1, e_2, \ldots, e_m, g_1 g_2, \ldots, g_n)$ в L_2.

Объединив базисы в L_1 и L_2, получим совокупность векторов

$$s = (e_1, e_2, \ldots, e_m, f_1, f_2, \ldots, f_k, g_1 g_2, \ldots, g_n). \tag{80}$$

Покажем, что векторы (80) линейно независимы.

Пусть линейная комбинация векторов (80) равна θ :

$$\sum_{i=1}^{m} \alpha_i e_i + \sum_{i=1}^{k} \beta_i f_i + \sum_{i=1}^{n} \gamma_i g_i = \theta. \tag{81}$$

Тогда $\sum\limits_{i=1}^{k} \beta_i f_i = -\sum\limits_{i=1}^{m} \alpha_i e_i - \sum\limits_{i=1}^{n} \gamma_i g_i$ и, обозначив $v = \sum\limits_{i=1}^{k} \beta_i f_i = -\sum\limits_{i=1}^{m} \alpha_i e_i - \sum\limits_{i=1}^{n} \gamma_i g_i$, получаем, что $v \in L_1$ и $v \in L_2$, т.е. $v \in L_1 \cap L_2$. Тогда вектор v имеет разложение по базису в $L_1 \cap L_2$: $v = \sum\limits_{i=1}^{m} \delta_i e_i$. Итак, для вектора v имеем два разложения по базису в L_2 : $v = \sum\limits_{i=1}^{m} \delta_i e_i + \sum\limits_{i=1}^{n} 0 g_i$ и $v = \sum\limits_{i=1}^{m} (-\alpha_i) e_i +$

$+\sum\limits_{i=1}^{n}(-\gamma_i)g_i$. В силу единственности разложения по базису имеем равенства всех соответствующих коэффициентов $\delta_i = -\alpha_i$ и $0 = -\gamma_i$. Поэтому все коэффициенты $\gamma_i = 0$ и равенство (81) принимает вид $\sum\limits_{i=1}^{m}\alpha_i e_i + \sum\limits_{i=1}^{k}\beta_i f_i = \theta$. Отсюда в силу линейной независимости векторов базиса $e_1, e_2, \ldots, e_m, f_1, f_2, \ldots, f_k$ в L_1 получаем, что все коэффициенты $\alpha_i = 0$ и $\beta_i = 0$. Таким образом, равенству (81) удовлетворяет только тривиальная линейная комбинация. Итак, векторы (80) линейно независимы и по теореме 2.7.2 образуют базис $L_1 + L_2$.

Размерность подпространств равна количеству векторов базиса:

$$dim(L_1 + L_2) = m + k + n, \qquad dim(L_1) = m + k,$$
$$dim(L_2) = m + n, \qquad dim(L_1 \cap L_2) = m.$$

Равенство (79) для этих величин выполнено: $m+k+n = (m+k)+(m+n)-m$. ■

2.9. Прямая сумма подпространств

В этом разделе рассматривается прямая сумма подпространств, когда каждый вектор суммы подпространств имеет единственное разложение по подпространствам.

Определение 1.

Говорят, что подпространство L является прямой суммой подпространств $L_1, L_2, \ldots, L_k$, если любой вектор $x \in L$ единственным образом представляется в виде суммы

$$x = x_1 + x_2 + \ldots + x_k, \tag{82}$$

где $x_i \in L_i$, $i = \overline{1,k}$.

Обозначение: $L = L_1 \dot{+} L_2 \dot{+} \ldots \dot{+} L_k$.

Лемма 2.9.1. *Если подпространство L является суммой подпространств $L_1, L_2, \ldots, L_k$, то любой вектор $x \in L$ имеет единственное разложение (82) тогда и только тогда, когда нулевой вектор θ имеет единственное разложение $\theta = \theta + \theta + \ldots + \theta$.*

Доказательство. 1) Очевидно, что если любой вектор $x \in L$ имеет единственное разложение (82), то и нулевой вектор θ имеет единственное разложение.

2) Пусть нулевой вектор θ имеет единственное разложение. Так как L является суммой подпространств L_1, L_2, ..., L_k, то для произвольного вектора $x \in L$ существует разложение (82). Докажем, что это представление единственно. Предположим противное: существует другое разложение $x = x'_1 + x'_2 + \ldots + x'_k$, где $x'_i \in L_i$, $i = \overline{1,k}$, причём некоторая разность $x_j - x'_j \neq 0$. Вычтем $x - x = x_1 + x_2 + \ldots + x_k - x'_1 - x'_2 - \ldots - x'_k$. Отсюда получаем $\theta = (x_1 - x'_1) + (x_2 - x'_2) + \ldots + (x_k - x'_k)$, где $x_i - x'_i \in L_i$, $i = \overline{1,k}$ и $x_j - x'_j \neq 0$, а это противоречит единственности разложения θ. ■

Для двух подпространств приведём критерий прямой суммы.

Теорема 2.9.1. *Пусть линейное подпространство L является суммой подпространств: $L = L_1 + L_2$. Тогда следующие условия эквивалентны.*

1) Сумма $L = L_1 + L_2$ прямая.

2) $L_1 \cap L_2 = \{\theta\}$.

3) Объединение базисов подпространств L_1 и L_2 является базисом суммы подпространств.

4) $dim(L_1 + L_2) = dim(L_1) + dim(L_2)$.

Доказательство. 1) $\Rightarrow$ 2) Пусть сумма $L = L_1 + L_2$ прямая. Возьмём произвольный $x \in L_1 \cap L_2$. Тогда $x \in L_1$, $x \in L_2$ и для него напишем два разложения $x = x + \theta$, $x \in L_1$, $\theta \in L_2$, $x = \theta + x$, $\theta \in L_1$, $x \in L_2$. В силу единственности разложения получаем $x = \theta$. Следовательно, верно 2).

2) $\Rightarrow$ 1) Пусть $L_1 \cap L_2 = \{\theta\}$. Предположим, что сумма не прямая. Тогда из определения 1 и леммы 2.9.1 следует, что θ имеет два разложения: $\theta = \theta + \theta$ и $\theta = x_1 + x_2$, $x_1 \in L_1$, $x_2 \in L_2$, где $x_1 \neq \theta$. Отсюда $x_1 = -x_2$ и обозначим $v = x_1 = -x_2$. Этот ненулевой вектор v принадлежит как $L_1\,(v = x_1)$, так и $L_2\,(v = -x_2)$, т.е. пересечению $L_1 \cap L_2$. Поскольку $L_1 \cap L_2 = \{\theta\}$, получаем $v = \theta$, а тогда $x_1 = \theta$, что противоречит условию $x_1 \neq \theta$.

2) $\Rightarrow$ 3) получается из следствия теоремы 2.7.2.

3) $\Rightarrow$ 4) следует из того, что размерность подпространства совпадает с количеством векторов базиса.

4) $\Rightarrow$ 2) Пусть $dim(L_1 + L_2) = dim(L_1) + dim(L_2)$. Тогда из равенства $dim(L_1 + L_2) = dim(L_1) + dim(L_2) - dim(L_1 \cap L_2)$ получаем $dim(L_1 \cap L_2) = 0$ и, следовательно, верно 2).

$2) \Rightarrow 1)$ уже доказано. ■

Задача 31. *а) Докажите, что* $dim(L_1 + L_2) \leqslant dim(L_1) + dim(L_2)$ *для любых подпространств* L_1 *и* L_2.

б) Докажите, что $dim(L_1 + L_2 + \ldots + L_k) \leqslant dim(L_1) + dim(L_2) + \ldots + dim(L_k)$ *для любых подпространств* $L_1, L_2, \ldots, L_k$, $k > 2$.

3. Евклидовы и унитарные пространства

3.1. Определение евклидова и унитарного пространства. Простейшие свойства скалярного произведения

Определение 1. *Линейное пространство* E *над полем* P *вещественных или комплексных чисел называется соответственно евклидовым или унитарным, если каждой паре векторов* f, g *из* E *поставлено в соответствие число из* P, *обозначаемое* (f, g) *и называемое скалярным произведением* [6] *векторов* f *и* g, *причём выполнены следующие аксиомы:*

A1. $(f, g) = \overline{(g, f)}$,[7]

A2. $(f_1 + f_2, g) = (f_1, g) + (f_2, g)$,

A3. $(\alpha f, g) = \alpha(f, g)$,

A4. $(f, f) > 0$ *при* $f \neq \theta$,

(*здесь* f, f_1, f_2, g *— любые векторы пространства* E, α *— любое число из* P).

Замечание 1. Из аксиомы A1 при $g = f$ следует, что $(f, f) = \overline{(f, f)}$, т.е. (f, f) — вещественное число, причём в силу A4 при $f \neq \theta$ — это положительное вещественное число.

Замечание 2. Для евклидова пространства скалярное произведение принимает только вещественные значения, поэтому аксиому A1 можно записать в виде: $(f, g) = (g, f)$.

[6] Само отображение, которое каждой паре векторов f, g из E ставит в соответствие их скалярное произведение, называют скалярным умножением. А скалярное произведение (f, g) — это результат вычисления отображения на паре векторов f, g.

[7] Число $\overline{z} = x - yi$ называется сопряжённым для комплексного числа $z = x + yi$. Очевидно, что $\overline{x + 0i} = x - 0i$, т.е. сопряжённым для вещественного числа x является само это число.

Пример 1. В пространстве[8] $\mathbb{R}^n$ скалярное произведение векторов $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ можно определить следующим образом:

$$(x, y) = \sum_{j=1}^{n} x_j y_j, \tag{83}$$

а в пространстве $\mathbb{C}^n$ —

$$(x, y) = \sum_{j=1}^{n} x_j \overline{y_j}. \tag{84}$$

Проверим выполнение аксиом для (84). Здесь $x, y, z \in E$, $\alpha \in P$.

1. $(y, x) = \sum\limits_{j=1}^{n} y_j \overline{x_j} = \overline{\sum\limits_{j=1}^{n} x_j \overline{y_j}} = \overline{(x, y)}$.
2. $(x + y, z) = \sum\limits_{j=1}^{n} (x_j + y_j)\overline{z_j} = \sum\limits_{j=1}^{n} x_j \overline{z_j} + \sum\limits_{j=1}^{n} y_j \overline{z_j} = (x, z) + (y, z)$.
3. $(\alpha x, y) = \sum\limits_{j=1}^{n} (\alpha x_j)\overline{y_j} = \alpha \sum\limits_{j=1}^{n} x_j \overline{y_j} = \alpha(x, y)$.
4. $(x, x) = \sum\limits_{j=1}^{n} x_j \overline{x_j} = \sum\limits_{j=1}^{n} |x_j|^2 > 0$ при $x \neq \theta$. ■

Пример 2. Рассмотрим функциональное пространство $C[0, 1]$ (множество функций $f(x)$ таких, что $f\colon [0, 1] \to \mathbb{R}$ и $f(x)$ непрерывна на отрезке $[0, 1]$) с операциями сложения функций и умножения функций на вещественные числа. Здесь скалярное произведение функций $f = f(x)$ и $g = g(x)$ может быть задано так:

$$(f, g) = \int_0^1 f(x) g(x)\, dx. \tag{85}$$

Легко проверяется, что формула (85) определяет скалярное произведение и выполнены аксиомы евклидова пространства. ■

Пример 3. Рассмотрим функциональное пространство $C([0, 1], \mathbb{C})$. Это множество комплекснозначных функций $f(x) = u(x) + iv(x)$, $(u, v \in C[0, 1])$ с операциями сложения функций и умножения функций на комплексные числа. Здесь скалярное произведение функций $f(x) = u(x) + iv(x)$ и $g(x) = p(x) + ir(x)$ может быть задано так:

$$(f, g) = \int_0^1 f(x) \overline{g(x)}\, dx. \tag{86}$$

или после преобразования:

[8] Пространство матриц-строк $\mathbb{R}^{1\times n}$ или матриц-столбцов $\mathbb{R}^{n\times 1}$.

$$(f,g)=\int\limits_0^1(u(x)+iv(x))(p(x)-is(x)\,dx=\int\limits_0^1((p(x)u(x)+s(x)v(x))+i(p(x)v(x)-$$
$$-s(x)u(x)))\,dx=\int\limits_0^1(p(x)u(x)+s(x)v(x))\,dx+i\int\limits_0^1(p(x)v(x)-s(x)u(x))\,dx.$$

Легко проверяется, что формула (86) определяет скалярное произведение и выполнены аксиомы унитарного пространства. ■

Теорема 3.1.1 (Простейшие свойства скалярного умножения).

1. $\forall\, f,\, g_1,\, g_2 \in E \quad (f,\, g_1+g_2)=(f,\, g_1)+(f,\, g_2)$.

2. $\forall\, f, g \in E,\, \beta \in P \quad (f,\, \beta g)=\overline{\beta}(f,\, g)$.

3. Формула скалярного произведения двух линейных комбинаций.

   $\forall\, m, r \in \mathbb{N},\, f_1,\, \ldots,\, f_m,\, g_1,\, \ldots,\, g_r, \in E,\, \alpha_1,\, \ldots,\, \alpha_m,\, \beta_1,\, \ldots,\, \beta_r \in P$
$$\left(\sum_{i=1}^{m}\alpha_i f_i,\, \sum_{j=1}^{r}\beta_j g_j\right)=\sum_{i=1}^{m}\sum_{j=1}^{r}\alpha_i\overline{\beta_j}(f_i,g_j). \tag{87}$$

4. $(\theta,\, g)=(g,\, \theta)=0$ *для всех* $g \in E$ *и, в частности,* $(\theta,\, \theta)=0$.

5. а) $(f,\, f)\geqslant 0$ *для всех* $f \in E$, б) $(f,\, f)=0$ *тогда и только тогда, когда* $f=\theta$.

6. $(f,\, g)=0$ *для всех* $g \in E$ тогда и только тогда, когда $f=\theta$.

7. *Любое подпространство евклидова (унитарного) пространства является евклидовым (унитарным) пространством.*

Доказательство.

1. $(f,\, g_1+g_2)=\overline{(g_1+g_2,\, f)}=\overline{(g_1,\, f)+(g_2,\, f)}=\overline{(g_1,\, f)}+\overline{(g_2,\, f)}=(f,\, g_1)+(f,\, g_2)$. Здесь использована аксиома А1 три раза, аксиома А2 один раз, и один раз утверждение: сопряжённое значение от суммы чисел равно сумме сопряжённых чисел.

2. $(f,\, \beta g)=\overline{(\beta g,\, f)}=\overline{\beta(g,\, f)}=\overline{\beta}\,\overline{(g,\, f)}=\overline{\beta}(f,\, g)$. Здесь использована аксиома А1 два раза, аксиома А3 один раз, и один раз утверждение: сопряжённое значение от произведения чисел равно произведению сопряжённых чисел.

3. Доказывается методом математической индукции.

4. Из аксиомы А3 при $\alpha = 0$ получаем $(0f,\, g) = 0(f,\, g)$. В силу $0f = \theta$ и $0(f,\, g) = 0$ выполнено равенство $(\theta,\, g) = 0$. Тогда по аксиоме А1 $(g,\, \theta) = \overline{(\theta,\, g)} = \overline{0} = 0$.

5. Выполнены условия $(f,\, f) > 0$ при $f \neq \theta$ по аксиоме А4 и $(\theta,\, \theta) = 0$ по утверждению 4. Тогда выполнены утверждение 5а: $(f,\, f) \geqslant 0$ при всех $f \in E$ и утверждение 5б: из $f = \theta$ следует, что $(f,\, f) = 0$, а из условия $(f,\, f) = 0$ следует, что обязательно $f = \theta$.

6. Если $(f,\, g) = 0$ при всех $g \in E$, то для $g = f$ получаем $(f,\, f) = 0$ и из условия 5б) $f = \theta$. Если $f = \theta$, то из утверждения 4 следует $(f,\, g) = 0$ при всех $g \in E$.

7. Скалярное произведение определено на каждой паре векторов пространства и тогда на каждой паре векторов подпространства. Линейное подпространство замкнуто относительно операций сложения и умножения на числа, поэтом аксиомы А1–А4 выполнены на всех векторах подпространства и для всех чисел из соответствующего поля. ■

3.2. Неравенство Коши–Буняковского. Длина вектора

В этом разделе рассматривается евклидово (унитарное) пространство E над полем $P = \mathbb{R}$ ($P = \mathbb{C}$).

Теорема 3.2.1 (Неравенство Коши-Буняковского).

1° *Для всех векторов* $f,\, g \in E$

$$|(f,\, g)|^2 \leqslant (f,\, f)(g,\, g). \tag{88}$$

или, в эквивалентной форме

$$\begin{vmatrix} (f,\, f) & (f,\, g) \\ (g,\, f) & (g,\, g) \end{vmatrix} \geqslant 0.$$

2° *Неравенство (88) обращается в равенство тогда и только тогда, когда векторы* $f,\, g$ *линейно зависимы (пропорциональны).*

Доказательство. 1° Берём любые векторы $f,\, g \in E$.

При $g = \theta$ имеем $(f,\, g) = 0$, $(g, g) = 0$ и неравенство (88) обращается в равенство.

При $g \neq \theta$ и произвольном $\beta \in P$ рассмотрим скалярное произведение $(f - \beta g, f - \beta g)$. В силу утверждения 5а теоремы 3.1.1 это скалярное произведение принимает неотрицательное значение, т.е.

$$(f - \beta g, f - \beta g) \geqslant 0.$$

Преобразуя скалярное произведение по формуле (87), получим неравенство

$$(f,\, f) - \overline{\beta}(f,\, g) - \beta(g,\, f) + \beta\overline{\beta}(g,\, g) \geqslant 0.$$

При $\beta = \dfrac{(f,\, g)}{(g,\, g)}$ в силу $(g,\, g) > 0$ имеем $\overline{\beta} = \dfrac{\overline{(f,\, g)}}{\overline{(g,\, g)}} = \dfrac{\overline{(f,\, g)}}{(g,\, g)}$, тогда

$\overline{\beta}(f,\, g) = \dfrac{\overline{(f,\, g)}(f,\, g)}{(g,\, g)} = \dfrac{|(f,\, g)|^2}{(g,\, g)}$, $\beta(g,\, f) = \beta\overline{(f,\, g)} = \dfrac{(f,\, g)\overline{(f,\, g)}}{(g,\, g)} = \dfrac{|(f,\, g)|^2}{(g,\, g)}$,

$\beta\overline{\beta}(g,\, g) = \dfrac{(f,\, g)\overline{(f,\, g)}(g,\, g)}{(g,\, g)^2} = \dfrac{|(f,\, g)|^2}{(g,\, g)}$, и неравенство примет вид:

$$(f,\, f) - \frac{|(f,\, g)|^2}{(g,\, g)} - \frac{|(f,\, g)|^2}{(g,\, g)} + \frac{|(f,\, g)|^2}{(g,\, g)} \geqslant 0 \quad \text{или}$$

$$(f,\, f) - \frac{|(f,\, g)|^2}{(g,\, g)} \geqslant 0.$$

Последнее неравенство запишем в виде

$$\frac{|(f,\, g)|^2}{(g,\, g)} \leqslant (f,\, f),$$

откуда после умножения на $(g,g)>0$ получается (88).

2° Часть утверждения фактически получается из рассуждений при доказательстве неравенства (88).

Пусть неравенство (88) обратилось в равенство.

При $g=\theta$ векторы f,g линейно зависимы и $g=0f$.

При $g\neq\theta$ выполнено $(g,g)\neq 0$. Тогда проходим неравенства в рассуждениях при доказательстве 1° в обратном порядке, начиная с равенства (88), и все они обращаются в равенства. В итоге получаем равенство $(f-\beta g, f-\beta g)=0$ при $\beta=\dfrac{(f,\,g)}{(g,\,g)}$. Отсюда следует, что $f-\beta g=\theta$, т.е. векторы f,g линейно зависимы и $f=\beta g$.

Пусть f, g линейно зависимы.

При $g=\theta$ или $g=\theta$ неравенство (88) обращается в равенство.

При $f\neq\theta$ и $g\neq\theta$ существует нетривиальная линейная комбинация, равная θ : $\gamma f+\delta g=\theta$. Здесь $\gamma\neq 0$. В противном случае $\gamma=0$, $\delta\neq 0$ и равенство $0f+\delta g=\theta$ противоречит $g\neq\theta$. Аналогично, $\delta\neq 0$. Далее из равенства $\gamma f+\delta g=\theta$ получим $f+\dfrac{\delta}{\gamma}g=\theta$ и, обозначив $\beta=-\dfrac{\delta}{\gamma}$, запишем $f-\beta g=\theta$ и $f=\beta g$.

Найдём значение β. Умножив равенство $f=\beta g$ скалярно на g, получим $(f,g)=(\beta g,g)$ или $(f,g)=\beta(g,g)$. Отсюда выражаем $\beta=\dfrac{(f,g)}{(g,g)}$.

Итак, при $\beta=\dfrac{(f,g)}{(g,g)}$ имеем $f-\beta g=\theta$ и $(f-\beta g,f-\beta g)=0$, то есть неравенство $(f-\beta g,f-\beta g)\geqslant 0$ обратилось в равенство. Тогда все неравенства в рассуждениях при доказательстве утверждения 1°, включая само неравенство (88), обращаются в равенства. ■

Определение 1. *Длиной вектора f называется арифметическое значение квадратного корня из его скалярного квадрата (f,f).*

Обозначение: $\|f\|$ или $|f|$.

Итак, $\|f\|=\sqrt{(f,f)}$.

Длину вектора также называют нормой.

Вектор единичной длины называется нормированным.

Теорема 3.2.2 (Свойства длины).

1° *Для всех векторов* $f \in E$ *существует длина* $\|f\| = \sqrt{(f,f)}$.

2° $\forall f \in E \quad \|f\| = \sqrt{(f,f)} \Leftrightarrow \|f\|^2 = (f,f)$.

3° *Неравенство Коши-Буняковского (88) можно записать в виде*

$$|(f,g)| \leqslant \|f\|\|g\|.$$

4° $\forall f \in E$ а) $\|f\| \geqslant 0$ *и* б) $\|f\| = 0 \Leftrightarrow f = \theta$.

5° $\|\alpha f\| = |\alpha|\|f\|$ *для всех векторов* $f \in E$ *и чисел* $\alpha \in P$.

6° $\|f+g\| \leqslant \|f\| + \|g\|$ *для всех векторов* $f, g \in E$.

Причём знак равенства здесь достается тогда и только тогда, когда $f = \beta g$ *или* $g = \beta f$ *при некотором* $\beta \geqslant 0$.

7° $\big|\|f\| - \|g\|\big| \leqslant \|f+g\|$ *для всех векторов* $f, g \in E$.

Причём знак равенства здесь достается тогда и только тогда, когда $f = \beta g$ *или* $g = \beta f$ *при некотором* $\beta \leqslant 0$.

Доказательство. 1°, 2°, 4° Для любого вектора $f \in E$ в силу утверждения 5 теоремы 3.1.1 $(f,f) \geqslant 0$, причём $(f,f) = 0 \Leftrightarrow f = \theta$. Следовательно, а) существует длина $\|f\| = \sqrt{(f,f)}$, б) $\|f\|^2 = (f,f)$, в) $\|f\| \geqslant 0$ и $\|f\| = 0 \Leftrightarrow f = \theta$.

3° $|(f,g)|^2 \leqslant (f,f)(g,g) \Leftrightarrow |(f,g)|^2 \leqslant \|f\|^2\|g\|^2 \Leftrightarrow |(f,g)| \leqslant \|f\|\|g\|$.

5° $\|\alpha f\|^2 = (\alpha f, \alpha f) = \alpha\overline{\alpha}(f,f) = |\alpha|^2\|f\|^2 \Leftrightarrow \|\alpha f\|^2 = |\alpha|^2\|f\|^2 \Leftrightarrow \|\alpha f\| = |\alpha|\|f\|$.

6° Покажем, что $\|f+g\|^2 \leqslant (\|f\| + \|g\|)^2$, тогда $\|f+g\| \leqslant \|f\| + \|g\|$. Из $\|f+g\|^2 = (f+g, f+g)$ получим $\|f+g\|^2 = (f,f) + (f,g) + (g,f) + (g,g)$. Возьмём модуль левой и правой частей последнего равенства, учитывая, что они неотрицательны, и воспользуемся неравенством модуль суммы не превосходит суммы модулей, тогда $\|f+g\|^2 \leqslant |(f,f)| + |(f,g)| + |(g,f)| + |(g,g)|$. Используя 2° и 3°, придём к неравенству $\|f+g\|^2 \leqslant \|f\|^2 + \|f\|\|g\| + \|g\|\|f\| + \|g\|^2$. Откуда $\|f+g\|^2 \leqslant \|f\|^2 + 2\|f\|\|g\| + \|g\|^2$ или $\|f+g\|^2 \leqslant (\|f\| + \|g\|)^2$.

7° Покажем, что $\|f+g\|^2 \geqslant (\|f\| - \|g\|)^2$, тогда $\|f+g\| \geqslant \big|\|f\| - \|g\|\big|$. Из доказательства 6° возьмём $\|f+g\|^2 = |(f,f) + (g,g) + (f,g) + (g,f)|$, воспользуемся два раза неравенством $|z_1 + z_2| \geqslant |z_1| - |z_2|$ и получим $\|f+g\|^2 \geqslant (f,f) + (g,g) - |(f,g)| - |(g,f)|$. Оценим снизу правую часть последнего неравенства. Из неравенства Коши-Буняковского $|(f,g)| \leqslant \|f\|\|g\|$ и $|(g,f))| \leqslant \|f\|\|g\|$, тогда $-|(f,g)| - |(g,f)| \geqslant -2\|f\|\|g\|$. Следовательно, получаем неравенство $\|f+g\|^2$ $geqslant \|f\|^2 - 2\|f\|\|g\| + \|g\|^2$ или $\|f+g\|^2 \geqslant (\|f\| - \|g\|)^2$. ■

Замечание. Любой ненулевой вектор f можно нормировать $\dfrac{f}{\|f\|}$, разделив его на длину $\|f\|$, или $\dfrac{1}{\|f\|}f$, умножив на $\dfrac{1}{\|f\|}$.

Определение 2. *Неравенства*

$$\big|\|f\| - \|g\|\big| \leqslant \|f+g\| \ \text{ и } \ \|f+g\| \leqslant \|f\| + \|g\|$$

называют неравенствами треугольника.[9]

Определение 3. *Углом между между ненулевыми векторами f и g в евклидовом пространстве назовём тот угол $\varphi\,(0 \leqslant \varphi \leqslant \pi)$, косинус которого равен отношению $\dfrac{(f,g)}{\|f\|\cdot\|g\|}$, т.е. $\cos\varphi = \dfrac{(f,g)}{\|f\|\cdot\|g\|}$.*

Это определение угла корректно, так как скалярное произведение в евклидовом пространстве принимает только вещественные значения и из неравенства Коши-Буняковского следует, что $-1 \leqslant \dfrac{(f,g)}{\|f\|\cdot\|g\|} \leqslant +1$.

В унитарном пространстве скалярное произведение принимает комплексные значения, поэтому понятия угла там не вводится. Рассматривается лишь понятие ортогональности.

Задача 32. *Напишите как выглядит неравенство Коши-Буняковского в пространствах примеров 1-3 на стр. 118–118*

[9] В любом треугольнике ABC выполняются неравенства $\big||\overline{AB}| - |\overline{BC}|\big| \leqslant |\overline{AC}|$ и $|\overline{AC}| \leqslant |\overline{AB}| + |\overline{BC}|$.

3.3. Ортогональность

В этом разделе продолжаем рассматривать евклидово (унитарное) пространство E над полем $P = \mathbb{R}$ ($P = \mathbb{C}$).

Определение 1. *Два вектора f и g пространства E называются ортогональными, если $(f, g) = 0$.*

Система векторов $e_1, e_2, \ldots, e_k \in E$ называется ортогональной системой, если

$$(e_i, e_j) = 0 \quad \text{при } i \neq j. \tag{89}$$

Система векторов $e_1, e_2, \ldots, e_k \in E$ называется ортонормированной системой, если она является ортогональной системой и каждый вектор системы — нормированный, т.е.

$$(e_i, e_j) = \delta_{ij}, \tag{90}$$

где δ_{ij} — символ Кронекера, для которого $\delta_{ij} = \begin{cases} 1, & i = j \\ 0. & i \neq j \end{cases}$

Теорема 3.3.1.

Ортогональная система ненулевых векторов линейно независима.

Доказательство. Пусть $e_1, e_2, \ldots, e_k \in E$ — ортогональная система ненулевых векторов и линейная комбинация векторов этой системы равна θ : $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$. Умножим обе части этого равенства на произвольный вектор e_j заданной ортогональной системы. Получим $\left(\sum\limits_{i=1}^{k} \alpha_i e_i, e_j\right) = (\theta, e_j)$. С помощью формулы скалярного произведения двух линейных комбинаций (87) находим $\sum\limits_{i=1}^{k} \alpha_i (e_i, e_j) = 0$. Здесь все слагаемые при $i \neq j$ равны 0, так как в силу ортогональности системы $(e_i, e_j) = 0$ при $i \neq j$. Следовательно, остаётся слагаемое при $i = j$: $\alpha_j (e_j, e_j) = 0$. Поскольку $e_j \neq \theta$, то $(e_j, e_j) \neq 0$ и тогда $\alpha_j = 0$. Следовательно, все коэффициенты линейной комбинации равны нулю и векторы рассматриваемой системы линейно независимы. ■

Следствия. 1. *Ортонормированная система векторов линейно независима.*

2. *В n-мерном пространстве ортогональная система n ненулевых векторов образует базис.*

3. *В n-мерном пространстве ортонормированная система n векторов образует базис.*

Определение 2.

Базис, векторы которого образуют ортогональную (ортонормированную) систему называется ортогональным (ортонормированным) базисом.

Теорема 3.3.2. *В пространстве E координаты $\alpha_1, \alpha_2, \ldots, \alpha_n$ вектора $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ в базисе $e = (e_1, e_2, \ldots, e_n)$ вычисляются по правилу*

$$\alpha_i = \frac{(f, e_i)}{(e_i, e_i)}, \quad i = \overline{1, n} \tag{91}$$

тогда и только тогда, когда e — ортогональный базис.

Доказательство. Пусть координаты вектора $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ вычисляются по правилу (91) для любого вектора f. Разложение произвольного базисного вектора $f = e_j$ по базису e имеет вид

$$e_j = 0e_1 + \ldots + 0e_{j-1} + 1e_j + 0e_{j+1} + \ldots + 0e_n,$$

где j-я координата $\alpha_j = 1$, а остальные координаты $\alpha_i = 0$ при $i \neq j$. Тогда из (91) получаем $\alpha_i = \dfrac{(e_j, e_i)}{(e_i, e_i)} = 0$ при $i \neq j$, т.е. $(e_j, e_i) = 0$ при $i \neq j$ и e — ортогональный базис.

Докажем обратное утверждение.

Пусть $e = (e_1, e_2, \ldots, e_n)$ — ортогональный базис и вектор f имеет в этом базисе координаты $\alpha_1, \alpha_2, \ldots, \alpha_n$, т.е. $f = \sum\limits_{i=1}^{n} \alpha_i e_i$. Возьмём произвольное $j \in \{1, 2, \ldots, n\}$ и найдём скалярное произведение векторов f и e_j :

$$(f, e_j) = \Big(\sum_{i=1}^{n} \alpha_i e_i, e_j \Big) = \sum_{i=1}^{n} \alpha_i (e_i, e_j).$$

В силу ортогональности системы $(e_i, e_j) = 0$ при $i \neq j$. Следовательно, от суммы остаётся только слагаемое при $i = j$, и мы получаем $(f, e_j) = \alpha_j (e_j, e_j)$ или $\alpha_j = \dfrac{(f, e_j)}{(e_j, e_j)}$ поскольку $(e_j, e_j) \neq 0$. Итак, координаты вычисляются по правилу (91). ■

Следствие. *В пространстве E координаты $\alpha_1, \alpha_2, \ldots, \alpha_n$ вектора $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ в базисе $e = (e_1, e_2, \ldots, e_n)$ вычисляются по правилу*

$$\alpha_i = (f, e_i), \quad i = \overline{1, n} \tag{92}$$

тогда и только тогда, когда e — ортонормированный базис.

Теорема 3.3.3.

В пространстве E скалярное произведение векторов $f = \sum\limits_{i=1}^{n} \alpha_i e_i$, $g = \sum\limits_{i=1}^{n} \beta_i e_i$, заданных своими координатами в базисе $e = (e_1, e_2, \ldots, e_n)$, вычисляется по правилу

$$(f, g) = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}, \tag{93}$$

тогда и только тогда, когда e — ортонормированный базис.

Доказательство. Пусть скалярное произведение любых векторов $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $g = \sum\limits_{i=1}^{n} \beta_i e_i$ вычисляется по правилу (93). Применив это правило для базисных векторов $e_j = \sum\limits_{i=1}^{n} \delta_{ij} e_i$ и $e_k = \sum\limits_{i=1}^{n} \delta_{ik} e_i$ получим $(e_j, e_k) = \sum\limits_{i=1}^{n} \delta_{ij} \overline{\delta_{ik}} = \sum\limits_{i=1}^{n} \delta_{ij} \delta_{ik} = \delta_{jj} \delta_{jk} = \delta_{jk}$, т.е. $e = (e_1, e_2, \ldots, e_n)$ — ортонормированный базис.

Докажем обратное утверждение.

Пусть $e = (e_1, e_2, \ldots, e_n)$ — ортонормированный базис. Используя формулу (3) скалярного произведения двух линейных комбинаций, вычислим скалярное произведение векторов $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $g = \sum\limits_{i=1}^{n} \beta_i e_i$: $\left(\sum\limits_{i=1}^{n} \alpha_i e_i, \sum\limits_{j=1}^{n} \beta_j e_j \right) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \overline{\beta_j} (e_i, e_j) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \overline{\beta_j} \delta_{ij} = \sum\limits_{i=1}^{n} \alpha_i \overline{\beta_i}$. Следовательно, скалярное произведение любых векторов вычисляется по правилу (93). ■

Следствие. Для евклидова пространства формулу (93) можно записать в виде

$$(f, g) = \sum_{i=1}^{n} \alpha_i \beta_i. \qquad .$$

3.4. Метод ортогонализации. Существование ортонормированного базиса

Опишем способ построения из линейно независимой системы векторов ортогональной системы векторов в евклидовом или унитарном пространстве.

Это некоторый способ перехода от заданной линейно независимой системы из m векторов $e = (e_1, e_2, \ldots, e_m)$ к ортогональной системе, состоящей из m ненулевых векторов; эти векторы будем обозначать $f_1, f_2, \ldots, f_m$.

Применяем метод математической индукции.

При $n = 1$ полагаем $f_1 = e_1$.

При $n = 2$ ищем f_2 в виде $f_2 = \alpha_1 f_1 + e_2$.

Так как $f_1 = e_1$, а векторы e_1, e_2 линейно независимы, то вектор f_2 ненулевой при любом числе α_1. Подберём это число так, чтобы вектор f_2 был ортогонален к вектору f_1 :

$$0 = (f_2, f_1) = (\alpha_1 f_1 + e_2, f_1) = \alpha_1 (f_1, f_1) + (e_2, f_1).$$

Поскольку $(f_1, f_1) \neq 0$, отсюда находим $\alpha_1 = -\dfrac{(e_2, f_1)}{(f_1, f_1)}$.

Следовательно, $f_2 = -\dfrac{(e_2, f_1)}{(f_1, f_1)} f_1 + e_2$. Отсюда, ввиду $f_1 = e_1$, следует, что вектор f_2 есть линейная комбинация векторов e_1, e_2.

Пусть уже построена ортогональная система ненулевых векторов f_1, f_2, $\ldots, f_k$, $(k < m)$ дополнительно предположим, что для всякого i, $1 \leqslant i \leqslant k$, вектор f_i является линейной комбинацией векторов $e_1, e_2, \ldots, e_i$. Это предположение будет выполняться тогда и для вектора f_{k+1}, если он будет выбран в виде

$$f_{k+1} = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_k f_k + e_{k+1}.$$

Вектор f_{k+1} ненулевой, т.к. система векторов e линейно независима, а вектор e_{k+1} не входит в записи векторов $f_1, f_2, \ldots, f_k$. Коэффициенты α_i, $i = \overline{1, k}$ найдём из условия, что вектор f_{k+1} должен быть ортогонален ко всем векторам f_i, $i = \overline{1, k}$:

$$0 = (f_{k+1}, f_i) = (\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_k f_k + e_{k+1}, f_i) =$$
$$= \alpha_1 (f_1, f_i) + \alpha_2 (f_2, f_i) + \ldots + \alpha_k (f_k, f_i) + (e_{k+1}, f_i);$$

отсюда, так как векторы f_1, $f_2, \ldots, f_k$, ортогональны между собой

$$\alpha_i (f_i, f_i) + (e_{k+1}, f_i) = 0, \text{ и}$$
$$\alpha_i = -\frac{(e_{k+1}, f_i)}{(f_i, f_i)}, \qquad i = \overline{1, k}.$$

Таким образом, за m шагов мы построим ортогональную систему ненулевых векторов. ■

Алгоритм метод ортогонализации.

Начало. Полагаем $f_1 = e_1$.

Для $k = \overline{1, m-1}$ полагаем

$$f_{k+1} = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_k f_k + e_{k+1}, \text{ где } \alpha_i = -\frac{(e_{k+1}, f_i)}{(f_i, f_i)}, \quad i = \overline{1, k}.$$

Конец. ■

Для построения ортонормированной системы $g_1, g_2, \ldots, g_m$ можно построить сначала ортогональную систему $f_1, f_2, \ldots, f_m$, затем нормировать каждый вектор $g_i = \dfrac{f_i}{\|f_i\|}$, $i = \overline{1, m}$. ■

Теорема 3.4.1.

В ненулевом конечномерном пространстве E существует ортонормированный базис.

Доказательство. Пусть $e = (e_1, e_2, \ldots, e_n)$ базис в E. Применим к системе e линейно независимых векторов метод ортогонализации и получим ортогональную систему, состоящую из n ненулевых векторов $f_1, f_2, \ldots, f_n$. В силу теоремы 3.3.1 эти n векторов линейно независимы и, следовательно, образуют базис. ■

4. Линейные операторы

4.1. Линейный оператор. Свойства линейного оператора

Определение 1. *Пусть E, F линейные пространства над полем P. Отображение*

$$A\colon E \to F \tag{94}$$

называется линейным оператором (линейным отображением), действующим из пространства E в пространство F, если для любых $x, x_1, x_2 \in E$, $\alpha \in P$

1) $A(x_1 + x_2) = A(x_1) + A(x_2)$ (образ суммы векторов равен сумме их образов);

2) $A(\alpha x) = \alpha A(x)$ (образ произведения вектора на число равен произведению образа вектора на это число).

При $E = F$ линейный оператор называется линейным преобразованием пространства E.

При $F = P$ линейный оператор называется линейным функционалом.

Множество всех линейных операторов, действующих из пространства E в пространство F, обозначается $L(E,F)$.

Операторы $A, B \in L(E,F)$ равны, если $A(x) = B(x)$ при всех $x \in E$. ■

Теорема 4.1.1.

Пусть A — линейный оператор (94).

1° *$A(\theta_E) = \theta_F$, где θ_E, θ_F – нулевые векторы в E и F соответственно, т. е. образ нулевого вектора является нулевым вектором.*

2° *$A\big(\sum\limits_{i=1}^{k} \alpha_i x_i\big) = \sum\limits_{i=1}^{k} \alpha_i A(x_i)$ для любой линейной комбинации $\sum\limits_{i=1}^{k} \alpha_i x_i$, т. е. образ линейной комбинации равен линейной комбинации образов с теми же коэффициентами.*

3° *Если векторы $x_1, x_2, \ldots, x_k$ линейно зависимы, то и образы этих векторов $A(x_1), A(x_2), \ldots, A(x_k)$ линейно зависимы, т. е. линейный оператор сохраняет линейную зависимость.*

Доказательство. 1° В определении возьмём число $\alpha = 0$. Получим $A(0x) = 0A(x)$. Здесь $0x = \theta_E$, $0A(x) = \theta_F$, т. е. $A(\theta_E) = \theta_F$.

2° Доказывается по индукции.

При $k = 1$ утверждение $A(\alpha_1 x_1) = \alpha_1 A(x_1)$ верно по определению.

Пусть утверждение верно при $k = m$, т.е. $A\big(\sum\limits_{i=1}^{m} \alpha_i x_i\big) = \sum\limits_{i=1}^{m} \alpha_i A(x_i)$.

Докажем, что утверждение верно при $k = m+1$: $A\big(\sum\limits_{i=1}^{m+1} \alpha_i x_i\big) = \sum\limits_{i=1}^{m+1} \alpha_i A(x_i)$.

$$A\Big(\sum_{i=1}^{m+1} \alpha_i x_i\Big) \boxed{\text{разобьём сумму на две части}} = A\Big(\sum_{i=1}^{m} \alpha_i x_i + \alpha_{m+1} x_{m+1}\Big) \boxed{\text{по определению}} = A\Big(\sum_{i=1}^{m} \alpha_i x_i\Big) +$$

$$+A(\alpha_{m+1} x_{m+1}) \boxed{\text{по предположениюю индукции и определению}} = \sum_{i=1}^{m} \alpha_i A(x_i) + \alpha_{m+1} A(x_{m+1}) \boxed{\text{сложим}} = \sum_{i=1}^{m+1} \alpha_i A(x_i).$$

3° Если $x_1, x_2, \ldots, x_k$ линейно зависимы, то найдётся нетривиальная линейная комбинация этих векторов, равная нулевому вектору: $\sum\limits_{i=1}^{m} \alpha_i x_i = \theta_E$.

Применим оператор A к обеим частям этого равенства: $A\big(\sum\limits_{i=1}^{m} \alpha_i x_i\big) = A(\theta_E)$.

Тогда из 1° и 2° получаем $\sum\limits_{i=1}^{m} \alpha_i A(x_i) = \theta_F$, т.е. нетривиальная линейная ком-

бинация векторов $A(x_1)$, $A(x_2)$, $\ldots$, $A(x_k)$ равна нулевому вектору. Следовательно, векторы $A(x_1)$, $A(x_2)$, $\ldots$, $A(x_k)$ линейно зависимы.

Следствие. Если векторы $A(x_1)$, $A(x_2)$, $\ldots$, $A(x_k)$ линейно независимы, то векторы x_1, x_2, $\ldots$, x_k линейно независимы. ■

Свойство 2° говорит о том, что для нахождения значения линейного оператора $A \in L(E, F)$ на любом векторе достаточно знать его значения на векторах базиса $e_1, e_2, \ldots, e_n$ пространства E, поскольку образ любого вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ можно найти по формуле $A(x) = \sum\limits_{i=1}^{n} \alpha_i A(e_i)$. Реализуем эту идею в следующем утверждении.

Теорема 4.1.2 (Задание линейного оператора).

Пусть e_1, e_2, $\ldots$, e_n — базис пространства E, f_1, f_2, $\ldots$, f_n — любые векторы пространства F. Тогда существует, и притом единственный, линейный оператор $A \in L(E, F)$, который переводит векторы e_1, e_2, $\ldots$, e_n в векторы f_1, f_2, $\ldots$, f_n соответственно, т.е.

$$A(e_j) = f_j,\ j = \overline{1, n}. \tag{95}$$

При этом образом произвольного вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ ($\alpha_i \in P$, $i = \overline{1, n}$) является вектор

$$A(x) = \sum_{i=1}^{n} \alpha_i f_i \in F. \tag{96}$$

Доказательство. Построим искомый оператор A, положив образом произвольного вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ вектор (96). Из единственности разложения вектора x по базису следует, что правило (96) однозначно определяет образ вектора x.

Проверим, что выполнены условия (95). Для вектора $x = e_j$ j-я координата $\alpha_j = 1$, остальные координаты $\alpha_i = 0$, $i \neq j$. Тодгда из (96) получаем $A(e_j) = f_j$.

Проверим линейность оператора A.

1) Пусть $x_1 = \sum\limits_{i=1}^{n} \alpha_i' e_i$, $x_2 = \sum\limits_{i=1}^{n} \alpha_i'' e_i$, тогда $x_1 + x_2 = \sum\limits_{i=1}^{n} (\alpha_i' + \alpha_i'') e_i$. Далее по формуле (96) получаем $A(x_1) = \sum\limits_{i=1}^{n} \alpha_i' f_i$, $A(x_2) = \sum\limits_{i=1}^{n} \alpha_i'' f_i$, $A(x_1 + x_2) = \sum\limits_{i=1}^{n} (\alpha_i' + \alpha_i'') f_i$. Сложив $A(x_1)$ и $A(x_2)$, находим $A(x_1) + A(x_2) = \sum\limits_{i=1}^{n} \alpha_i' f_i + \sum\limits_{i=1}^{n} \alpha_i'' f_i =$

$= \sum\limits_{i=1}^{n}(\alpha_i' + \alpha_i'')f_i = A(x_1 + x_2)$, т.е. выполнено равенство $A(x_1 + x_2) = A(x_1) + A(x_2)$.

2) Пусть $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$, $\beta \in P$, тогда $\beta x = \sum\limits_{i=1}^{n}(\beta\alpha_i)e_i$. Далее по формуле (96) получаем $A(x) = \sum\limits_{i=1}^{n} \alpha_i f_i$, $A(\beta x) = \sum\limits_{i=1}^{n}(\beta\alpha_i)f_i$. Умножив обе части равенства $A(x) = \sum\limits_{i=1}^{n} \alpha_i f_i$ на число β, находим $\beta A(x) = \beta \sum\limits_{i=1}^{n} \alpha_i f_i = \sum\limits_{i=1}^{n}(\beta\alpha_i)f_i = A(\beta x)$, т.е. выполнено равенство $A(\beta x) = \beta A(x)$.

Оператор A единственен. Действительно, если $B \in L(E, F)$ — любой оператор, удовлетворяющий условиям $B(e_j) = f_j$, $j = \overline{1, n}$, то для всех $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$

$$Bx = B\Big(\sum_{i=1}^{n} \alpha_i e_i\Big) = \sum_{i=1}^{n} \alpha_i B(e_i) = \sum_{i=1}^{n} \alpha_i f_i = A(x).$$

Следовательно, $B = A$. ■

4.2. Матрица линейного оператора. Линейное выражение координат образа вектора при линейном отображении. Связь матриц линейного оператора в различных базисах

В этом разделе рассматриваем линейные пространства E, F над полем P и множество линейных операторов $L(E,F)$.

Определение 1. *Пусть E, F — линейные пространства с базисами $e=(e_1,e_2,\ldots,e_n)$, $f=(f_1,f_2,\ldots,f_m)$ соответственно, и A — линейный оператор, действующий из E в F.*

Разложим образы векторов базиса e по базису f :

$$A(e_j)=\sum_{i=1}^{m} a_{ij}f_i,\ j=\overline{1,n} \tag{97}$$

или подробнее

$$\begin{aligned}
A(e_1)&=a_{11}f_1+a_{21}f_2+\ldots+a_{m1}f_m,\\
A(e_2)&=a_{12}f_1+a_{22}f_2+\ldots+a_{m2}f_m,\\
\ldots\ \ \ldots\ \ \ldots\ \ \ldots\ \ \ldots\ \ \ldots\ \ \ldots\ \ \ldots\\
A(e_n)&=a_{1n}f_1+a_{2n}f_2+\ldots+a_{mn}f_m.
\end{aligned}$$

Из координатных столбцов векторов $[A(e_1)]_f,[A(e_2)]_f,\ldots,[A(e_n)]_f$ в базисе f составим матрицу

$$[A]_{f,e}=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{pmatrix}. \tag{98}$$

Матрица $[A]_{f,e}$ называется матрицей линейного оператора A в паре базисов e и f.

Очевидно, что $[A]_{f,e}$ — это $m\times n$-матрица с элементами из поля P, т.е. $[A]_{f,e}\in P^{m\times n}$. ■

Замечание. Из единственности разложения вектора по базису следует, что при фиксированных базисах матрица линейного оператора определена однозначно. Итак, каждому оператору $A\in L(E,F)$ соответствует его матрица $[A]_{f,e}\in P^{m\times n}$.

Таким образом, при фиксированных базисах построено отображение, действующее из множества линейных операторов $L(E,F)$ во множество матриц

$P^{m\times n}$ по правилу

$$A \to [A]_{e,f}. \tag{99}$$

Теорема 4.2.1.

Отображение (99) биективно.

Доказательство. Инъективность. Пусть A, B линейные операторы из $L(A,B)$ и $A \neq B$. В силу теоремы 4.1.2 значения $A(e_i)$ и $B(e_i)$ на всех базисных векторах e_i не могут совпадать, иначе если $A(e_i) = B(e_i)$ при всех $i = \overline{1,n}$, то $A = B$. Тогда координатные столбцы несовпадающих векторов $A(e_i) \neq B(e_i)$ различны, и поэтому матрицы $[A]_{f,e} \neq [B]_{f,e}$.

Сюръективность. Покажем, что любая матрица $H = (h_{ij}) \in P^{m\times n}$ является матрицей некоторого линейного оператора из $L(E,F)$. Возьмём векторы $g_j = \sum\limits_{i=1}^{m} h_{ij} f_i,\ j = \overline{1,n}$ и согласно теореме 4.1.2 построим линейный оператор $G \in L(E,F)$, для которого $G(e_j) = g_j,\ j = \overline{1,n}$. Для него $G(e_j) = g_j = \sum\limits_{i=1}^{m} h_{ij} f_i,\ j = \overline{1,n}$ и из определения 1 следует, что матрица $[G]_{f,e}$ совпадает с H. ■

Установим связь координат вектора и его образа при линейном отображении.

Теорема 4.2.2 (Связь координат вектора и его образа).

Пусть E, F — линейные пространства с базисами $e = (e_1, e_2, \ldots, e_n)$, $f = (f_1, f_2, \ldots, f_m)$ соответственно, A — линейный оператор, действующий из E в F, $x \in E$, $y \in F$.

Тогда следующие условия эквивалентны:

$$y = Ax; \tag{100}$$

$$[y]_f = [A]_{f,e}[x]_e; \tag{101}$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}, \tag{102}$$

где $[x]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}$, $[y]_f = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_m \end{pmatrix}$ *— координатные столбцы вектора* x *в базисе* e *и вектора* y *в базисе* f *соответственно,* $[A]_{f,e} = (a_{ij})$ *— матрица оператора* A *в паре базисов* e, f.

Доказательство. 1) Формула (101) — это краткая запись формулы (102).

2) Пусть верна формула (100): $y = A(x)$, где $x = \sum\limits_{j=1}^{n} \alpha_j e_j$, $y = \sum\limits_{i=1}^{m} \beta_i e_i$.

Тогда $y = A(x) = A\big(\sum\limits_{j=1}^{n} \alpha_j e_j\big) = \sum\limits_{j=1}^{n} \alpha_j A(e_j) = \boxed{\text{подставим (97)}} = \sum\limits_{j=1}^{n} \alpha_j \sum\limits_{i=1}^{m} a_{ij} f_i = \boxed{\text{внесём } \alpha_j} =$

$= \sum\limits_{j=1}^{n} \sum\limits_{i=1}^{m} a_{ij} \alpha_j f_i = \boxed{\text{поменем порядок суммирования}} = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} a_{ij} \alpha_j f_i = \boxed{\text{вынесем } f_i} = \sum\limits_{i=1}^{m} \big(\sum\limits_{j=1}^{n} a_{ij} \alpha_j\big) f_i.$

Итак, $y = \sum\limits_{i=1}^{m} \beta_i e_i$ и $y = \sum\limits_{i=1}^{m} \big(\sum\limits_{j=1}^{n} a_{ij} \alpha_j\big) f_i$. Отсюда в силу единственности разложения по базису имеем $\beta_i = \sum\limits_{j=1}^{n} a_{ij} \alpha_j$, $i = \overline{1, n}$, а эти формулы в матричной записи дают (102).

3) Пусть верна формула (102). Отсюда получаем $\beta_i = \sum\limits_{j=1}^{n} a_{ij} \alpha_j$, $i = \overline{1, m}$. Тогда вектор $y = \sum\limits_{i=1}^{m} \beta_i e_i$ можно записать другим способом $y = \sum\limits_{i=1}^{m} \big(\sum\limits_{j=1}^{n} a_{ij} \alpha_j\big) f_i$. Выполнив преобразования п. 2) в обратном порядке, получим $y = A(x)$. ■

Выясним связь матриц линейного оператора в различных базисах.

Теорема 4.2.3 (Связь матриц оператора в различных базисах).

Пусть в пространстве E *заданы два базиса* $e = (e_1, e_2, \dots, e_n)$, $e' = (e'_1, e'_2, \dots, e'_n)$ *с матрицей перехода* $T_{e \to e'}$, *т.е.* $e' = eT_{e \to e'}$; *в пространстве* F *заданы два базиса* $f = (f_1, f_2, \dots, f_m)$, $f' = (f'_1, f'_2, \dots, f'_m)$ *с матрицей перехода* $T_{f \to f'}$, *т.е.* $f' = fT_{f \to f'}$; *линейному оператору* $A \in L(E, F)$ *в паре базисов* e, f *соответствует матрица* $[A]_{f,e}$, *а в паре базисов* e', f' *— матрица* $[A]_{f',e'}$.

Тогда матрицы $[A]_{f,e}$ *и* $[A]_{f',e'}$ *связаны соотношениями*

$$T_{f \to f'} [A]_{f',e'} = [A]_{f,e} T_{e \to e'}, \tag{103}$$

$$[A]_{f',e'} = T^{-1}_{f\to f'}[A]_{f,e}T_{e\to e'}. \tag{104}$$

Доказательство.

Для произвольного вектора $x \in E$ и его образа $y = A(x)$ в силу (101) имеем

$$[y]_f = [A]_{f,e}[x]_e, \quad [y]_{f'} = [A]_{f',e'}[x]_{e'}. \tag{105}$$

В свою очередь, согласно (76), $[x]_e = T_{e\to e'}[x]_{e'}$, $[y]_f = T_{f\to f'}[y]_{f'}$. Подставив эти соотношения в равенства (105), получим $T_{f\to f'}[y]_{f'} = [A]_{f,e}T_{e\to e'}[x]_{e'}$ или $T_{f\to f'}[A]_{f',e'}[x]_{e'} = [A]_{f,e}T_{e\to e'}[x]_{e'}$. Поскольку последнее равенство выполняется для любых координатных столбцов $[x]_{e'}$, то выполнено равенство $T_{f\to f'}[A]_{f',e'} = [A]_{f,e}T_{e\to e'}$. Откуда после умножения на обратную матрицу $T^{-1}_{f\to f'}$ получаем (104). ■

Замечание. *Для линейного преобразования $A\colon E \to E$ при построении его матрицы часто берут один базис, т.е. $e = f = (e_1, e_2, \ldots, e_n)$, находят матрицу $[A]_{e,e}$, и называют её матрицей линейного преобразования A в базисе e.*

Тогда справедливо следующее утверждение.

Следствие. *Пусть в линейном пространстве E заданы два базиса $e = (e_1, e_2, \ldots, e_n)$ («старый») и $e' = (e'_1, e'_2, \ldots, e'_n)$ («новый») с матрицей перехода $T_{e\to e'}$, т.е. $e' = eT_{e\to e'}$; линейному преобразованию $A \in L(E, E)$ в базисе e соответствует матрица $[A]_{e,e}$, а в базисе e' — матрица $[A]_{e',e'}$.*

Тогда матрицы $[A]_{e,e}$ и $[A]_{e',e'}$ связаны соотношениями

$$T_{e\to e'}[A]_{e',e'} = [A]_{e,e}T_{e\to e'}, \tag{106}$$

$$[A]_{e',e'} = T^{-1}_{e\to e'}[A]_{e,e}T_{e\to e'}. \tag{107}$$

4.3. Собственные векторы и собственные значения линейного преобразования

В этом разделе рассматриваем линейное пространство E над полем P и множество линейных преобразований $L(E, E)$.

Определение 1.

Ненулевой вектор x называется собственным вектором линейного преобразования $A \in L(E, E)$, если существует такое число $\lambda \in P$, что

$$A(x) = \lambda x. \tag{108}$$

Число λ называется собственным значением преобразования A, соответствующим собственному вектору x.

При этом говорят также, что собственный вектор x соответствует собственному значению λ.

Множество всех собственных значений преобразования A называется его спектром. ■

Примеры. 1. В пространстве многочленов $\mathbb{R}[x]$ любой многочлен нулевой степени $f(x) = c \neq 0$ будет собственным вектором преобразования дифференцирования $A(f) = f'$, так как $A(c) = c' = 0$, поэтому $A(c) = 0c$. Итак, многочлену нулевой степени соответствует собственное значение $\lambda = 0$.

2. В пространстве функций $f(x)$ $(f\colon \mathbb{R} \to \mathbb{R})$, имеющих производные всех порядков рассмотрим преобразование $A(f) = f''$ (f'' — вторая производная от функции f). Здесь любое число $\lambda \in \mathbb{R}$ является собственным значением, например, $\lambda = -4$, т.к. $A(\sin(2x)) = -4\sin(2x)$. Придумайте собственные векторы (их называют собственными функциями) для произвольных $\lambda \in \mathbb{R}$.

3. В пространстве V_2 всех геометрических векторов плоскости, у которых начало совпадает с точкой O, рассмотрим преобразование A_α поворота векторов в плоскости вокруг точки O на угол α против часовой стрелки. При $\alpha = \pi(2k+1)$, $k \in Z$ все ненулевые векторы $\overrightarrow{x}$ являются собственными векторами, соответствующие собственному значению $\lambda = -1$, т.к. в этом случае $A_\alpha \overrightarrow{x} = -\overrightarrow{x}$. При $\alpha = 2\pi k$, $k \in Z$ есть собственные векторы и собственные значения (найдите их). При $\alpha \neq \pi k$, $k \in Z$ собственных значений и собственных векторов нет (почему?). ■

Лемма 4.3.1. *Пусть при фиксированном* $\lambda = \lambda_0$ *линейное преобразование* $A \in L(E, E)$ *имеет собственные векторы. Обозначим через* E_{λ_0} *множество, содержащее все собственные векторы, соответствующие* λ_0, *и вектор* θ. *Тогда* E_{λ_0} *— линейное подпространство в* E.

Доказательство. Для любого вектора x из E_{λ_0} верно равенство $A(x) = \lambda_0 x$. Действительно, если $x \neq \theta$, то x — собственный вектор, соответствующий λ_0, и равенство выполнено. Если же $x = \theta$, то равенство $A(\theta) = \lambda_0\theta$ верно, т.к. $A(\theta) = \theta$ и $\lambda_0\theta = \theta$.

1) Проверим, что из $x, y \in E_{\lambda_0}$ следует $x + y \in E_{\lambda_0}$.

$x, y \in E_{\lambda_0} \Rightarrow A(x) = \lambda_0 x,\ A(y) = \lambda_0 y$ $\boxed{\text{сложим}}$ $\Rightarrow A(x) + A(y) = \lambda_0(x + y) \Rightarrow$ $\boxed{\text{Из линейности } A(x+y) = A(x) + A(y)}$ $\Rightarrow A(x + y) = \lambda_0(x + y) \Rightarrow x + y \in E_{\lambda_0}$.

2) Проверим, что из $x \in E_{\lambda_0}$, $\alpha \in P$ следует $\alpha x \in E_{\lambda_0}$.

$x \in E_{\lambda_0}$, $\alpha \in P \Rightarrow A(x) = \lambda_0 x$ $\boxed{\text{Умножим на } \alpha}$ $\Rightarrow \alpha A(x) = \alpha\lambda_0 x$ $\boxed{\text{Из линейности } A(\alpha x) = \alpha A(x) \text{ и перестановка чисел } \alpha \text{ и } \lambda_0}$ $\Rightarrow$ $A(\alpha x) = \lambda_0(\alpha x) \Rightarrow \alpha x \in E_{\lambda_0}$. ■

Теорема 4.3.1.

Собственные векторы $x_1, x_2, \ldots, x_k$, *соответствующие различным собственным значениям* $\lambda_1, \lambda_2, \ldots, \lambda_k$, *линейно независимы.*

Доказательство. Применим индукцию по k. Для $k = 1$ утверждение верно, так как собственный вектор является ненулевым и, следовательно, линейно независим.

Пусть утверждение верно для системы из m векторов. Докажем его для $m+1$ векторов $x_1, \ldots, x_m, x_{m+1}$. Приравняем линейную комбинацию этих векторов нулевому вектору:

$$\alpha_1 x_1 + \ldots + \alpha_m x_m + \alpha_{m+1} x_{m+1} = \theta. \tag{109}$$

Применим преобразование A к обеим частям равенства (109), используя свойство линейного оператора и определение собственного вектора:

$A(\alpha_1 x_1 + \ldots + \alpha_m x_m + \alpha_{m+1} x_{m+1}) = \alpha_1 A(x_1) + \ldots + \alpha_m A(x_m) + \alpha_{m+1} A(x_{m+1}) =$
$= \alpha_1\lambda_1 x_1 + \ldots + \alpha_m\lambda_m x_m + \alpha_{m+1}\lambda_{m+1} x_{m+1}$; $A(\theta) = \theta$, получим

$$\alpha_1\lambda_1 x_1 + \ldots + \alpha_m\lambda_m x_m + \alpha_{m+1}\lambda_{m+1}x_{m+1} = \theta. \qquad (110)$$

Чтобы убрать из линейной комбинации одно слагаемое, умножим обе части (109) на λ_{m+1} и вычтем из (110), получим $\alpha_1(\lambda_1 - \lambda_{m+1})x_1 + \ldots + \alpha_m(\lambda_m - \lambda_{m+1})x_m = \theta$. В силу предположения индукции векторы $x_1, \ldots, x_m$ линейно независимы, следовательно, все коэффициенты полученной комбинации равны нулю: $\alpha_i(\lambda_i - \lambda_{m+1}) = 0, i = \overline{1, m}$. По условию все собственные значения различны, тогда $(\lambda_i - \lambda_{m+1}) \neq 0$ и получаем $\alpha_i = 0, i = \overline{1, m}$. При этом из (109) следует равенство $\alpha_{m+1}x_{m+1} = \theta$ и, поскольку $x_{m+1} \neq \theta$, получаем $\alpha_{m+1} = 0$. Следовательно, все коэффициенты $\alpha_i = 0, i = \overline{1, m+1}$ и векторы $x_1, \ldots, x_m, x_{m+1}$ линейно независимы. ■

Следствия. 1. Линейное преобразование n-мерного пространства не может иметь более чем n различных собственных значений.

2. Если линейное преобразование n-мерного пространства имеет n различных собственных значений, то его собственные векторы, соответствующие различным собственным значениям образуют базис пространства. ■

Приведём утверждение, которое даёт возможность вычисления координат собственных векторов преобразования, если известны его собственные значения. Эта теорема не гарантирует наличие собственных значений. Как показывают примеры в начале этого раздела существуют преобразования, которые не имеют собственных значений.

Теорема 4.3.2. *Пусть в линейном пространстве E над полем P задан базис $e = (e_1, \ldots, e_n)$ и линейное преобразование A с матрицей $[A]_{e,e} = (a_{ij})$ в базисе e.*

Преобразование A имеет собственный вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq \theta$, отвечающий собственному значению $\lambda = \lambda_0$, т.е. выполнено равенство $A(x) = \lambda_0 x$, тогда и только тогда, когда λ_0 является корнем характеристического многочлена матрицы $[A]_{e,e}$

$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} \qquad (111)$$

и

$[x]_e$ *(координатный столбец вектора x в базисе e) является нетривиальным решением матричного уравнения*

$$[A]_{e,e}[x]_e = \lambda_0[x]_e \tag{112}$$

или системы уравнений

$$\begin{cases} (a_{11}-\lambda_0)\beta_1 & + & a_{12}\beta_2 & +\ldots+ & a_{1n}\beta_n & = & 0, \\ a_{21}\beta_1 & + & (a_{22}-\lambda_0)\beta_2 & +\ldots+ & a_{2n}\beta_n & = & 0, \\ \ldots & & \ldots & \ldots & \ldots & & \ldots \\ a_{n1}\beta_1 & + & a_{n2}\beta_2 & +\ldots+ & (a_{nn}-\lambda_0)\beta_n & = & 0. \end{cases} \tag{113}$$

Доказательство. 1) Пусть преобразование A имеет собственный вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq \theta$, отвечающий собственному значению $\lambda = \lambda_0$. Тогда

$$Ax = \lambda_0 x \boxed{\text{приравняем координатные столбцы}} \Rightarrow [Ax]_e = [\lambda_0 x]_e \boxed{\begin{array}{l}\text{применим формулу (101)} \\ [Ax]_e = [A]_{e,e}[x]_e \\ \text{и свойство координат} \\ [\lambda_0 x]_e = \lambda_0[x]_e\end{array}} \Rightarrow [A]_{e,e}[x]_e = \lambda_0[x]_e \Rightarrow$$

координатный столбец вектора $[x]_e$ является нетривиальным решением матричного уравнения $[A]_{e,e}[x]_e = \lambda_0[x]_e$, т.е. выполнено (112). В подробной записи это равенство имеет вид

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix} = \lambda_0 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}.$$

Сделав умножение и приравняв соответствующие элементы левой и правой частей получим систему

$$\begin{cases} a_{11}\beta_1 & + & a_{12}\beta_2 & +\ldots+ & a_{1n}\beta_n & = & \lambda_0\beta_1, \\ a_{21}\beta_1 & + & a_{22}\beta_2 & +\ldots+ & a_{2n}\beta_n & = & \lambda_0\beta_2, \\ \ldots & & \ldots & \ldots & \ldots & & \ldots \\ a_{n1}\beta_1 & + & a_{n2}\beta_2 & +\ldots+ & a_{nn}\beta_n & = & \lambda_0\beta_n, \end{cases}$$

которая эквивалентна (113). Следовательно, столбец $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}$ является нетривиальным решением системы (113).

2) Проведя рассуждения пункта 1) в обратном порядке, получим справедливость следующего утверждения: если столбец $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix}$ является нетривиальным решением системы (113), то вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq \theta$ удовлетворяет равенству $Ax = \lambda_0 x$ и, следовательно, преобразование A имеет собственный вектор x отвечающий собственному значению $\lambda = \lambda_0$. ■

Следствия. 1. В n-мерном комплексном линейном пространстве характеристический многочлен (111) любого линейного преобразования A имеет ровно n корней, если каждый корень считать столько раз какова его кратность.

2. В n-мерном комплексном линейном пространстве любое линейное преобразование имеет по крайней мере один собственный вектор.

Доказательство. Пусть $f(\lambda)$ — характеристический многочлен матрицы $D = (a_{ij})$ преобразования A. Раскрыв этот определитель, получим, что $f(\lambda)$ — это многочлен степени n от λ. Из основной теоремы алгебры следует, что многочлен степени n имеет n корней с учётом их кратностей.

Далее пусть λ_0 — корень характеристического многочлена (111), то есть $f(\lambda_0) = 0$. Рассмотрим систему (113). Матрицу этой системы можно представить с виде $B = D - \lambda_0 I$, где D — матрица преобразования, I — единичная $n \times n$-матрица. Определитель матрицы B равен $f(\lambda_0) = 0$, поэтому система имеет ненулевое решение $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix}$. Тогда по теореме 4.3.2 вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq \theta$ удовлетворяет равенству $Ax = \lambda_0 x$ и, следовательно, преобразование A имеет собственный вектор x отвечающий собственному значению $\lambda = \lambda_0$. . ■

Сформулируем *алгоритм поиска собственных векторов и собственных значений.*

Требуется найти собственные векторы и собственные значения линейного преобразования A, т.е. $\lambda \in P$ и векторы $x \in E$ такие, что $A = \lambda x$ и $x \neq \theta$.

Начало.

1. Находим матрицу преобразования $[A]_{e,e} = (a_{ij})$ в базисе $e = (e_1, \ldots, e_n)$. (Желательно выбирать базис так, чтобы матрица была попроще.)

2. Составляем характеристический многочлен
$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{22} - \lambda \end{vmatrix}$$

3. Решаем уравнение $f(\lambda) = 0$.

4. Если корней нет, то собственных значений собственных векторов нет. Конец.

5. В противном случае находим корни $\lambda = \lambda_j$, $j = \overline{1, k}$, $1 \leqslant k \leqslant n$.

6. Для $j = \overline{1, k}$ при $\lambda = \lambda_j$ записываем систему
$$\begin{cases} (a_{11} - \lambda)\beta_1 + a_{12}\beta_2 + \ldots + a_{1n}\beta_n = 0, \\ a_{21}\beta_1 + (a_{22} - \lambda)\beta_2 + \ldots + a_{2n}\beta_n = 0, \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ a_{n1}\beta_1 + a_{n2}\beta_2 + \ldots + (a_{nn} - \lambda)\beta_n = 0. \end{cases}$$
Находим общее решение $(\beta_1, \beta_2, \ldots, \beta_n) \neq (0, 0, \ldots, 0)$.

   Записываем собственные векторы $x = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n$, соответствующие собственному значению $\lambda = \lambda_j$.

   Конец. ■

4.4. Условия приведения матрицы линейного преобразования к диагональному виду. Жорданова форма матрицы линейного преобразования

Проанализируем возможность приведения матрицы линейного преобразования к диагональному виду, т.е. к наиболее простому виду.

Теорема 4.4.1. *Матрица линейного преобразования* $A \in L(E, E)$ *в некотором базисе* $e = (e_1, e_2, \ldots, e_n)$ *имеет диагональный вид*

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \tag{114}$$

тогда и только тогда, когда базис состоит из собственных векторов e_1, e_2, $\ldots, e_n$, *соответствующих собственным знамениям* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *(собственные значения не обязательно различны).*

Доказательство.

1) Пусть матрица преобразования A в базисе $e = (e_1, e_2, \ldots, e_n)$ имеет диагональный вид (114). По определению матрица состоит из координатных столбцов образов базисных векторов, т.е. из столбцов $[A(e_1)]_e$, $[A(e_2)]_e$, $\ldots$, $[A(e_n)]_e$. Следовательно, из столбцов матрицы (114) видим, что i-я координата вектора $A(e_i)$ равна λ_i, а остальные координаты равны 0, тогда $A(e_i) = 0e_1 + \ldots + \lambda_i e_i + \ldots + 0e_n$, т.е. $A(e_i) = \lambda_i e_i$, $i = \overline{1, n}$. Итак, базис состоит из собственных векторов.

2) Пусть базис $e = (e_1, e_2, \ldots, e_n)$ состоит из собственных векторов и $A(e_i) = \lambda_i e_i$, $i = \overline{1, n}$. Тогда $A(e_i) = 0e_1 + \ldots + \lambda_i e_i + \ldots + 0e_n$. Записав координатные столбцы векторов $[A(e_i)]_e$ в матрицу $[A]_{e,e}$, видим,что она имеет диагональный вид (114). ■

Существуют линейные преобразования, матрица которых не имеют диагонального вида ни в каком базисе. Укажем для них наиболее простой вид матрицы.

Определение 1.

Квадратная $k \times k$*-матрица вида*

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix} \tag{115}$$

называется жордановой клеткой порядка k.

Клеточно-диагональная матрица, у которой все диагональные клетки жордановы, называется жордановой матрицей или матрицей, имеющей жорданову форму. ■

Замечания. 1. При $k = 1$ жордановой клеткой порядка 1 является матрица $(\ \lambda\ )$.

2. Объединение всех диагональных элементов жордановой матрицы даёт все её собственные значения.

3. Диагональная матрица является частным случаем жордановой матрицы, имеющим все жордановы клетки порядка 1.

Теорема 4.4.2 (Жордан).

Для любого линейного преобразования комплексного линейного пространства существует жорданов базис, в котором его матрица имеет жорданову форму.

Жорданова форма матрицы единственна с точностью до порядка расположения диагональных клеток.

5. Билинейные и квадратичные формы

5.1. Билинейные формы (функции), их матрицы. Квадратичные формы (функции)

В этом разделе рассматриваем билинейные и квадратичные формы в линейном пространстве E над полем P.

Определение 1. *Отображение* $B\colon E \times E \to P$ *(от двух переменных) называется билинейной формой, если*

1) $B(f_1 + f_2,\, g) = B(f_1,\, g) + B(f_2,\, g)$,
2) $B(\gamma f,\, g) = \gamma B(f,\, g)$,
3) $B(f,\, g_1 + g_2,) = B(f,\, g_1) + B(f,\, g_2)$,
4) $B(f,\, \delta g) = \delta B(f,\, g)$,

(здесь f, f_1, f_2, g, g_1, g_2 *— любые векторы пространства* E, γ, δ *— любые числа из* P*).*

Билинейная форма называется симметричной, если $B(f,\, g) = B(g,\, f)$ *при всех* $f, g \in E$.

Две формы $B_1(f,\, g)$ *и* $B_2(f,\, g)$ *равны, если* $B_1(f,\, g) = B_2(f,\, g)$ *при всех* $f, g \in E$.

Пример 1. Скалярное произведение (f, g) в евклидовом пространстве является симметричной билинейной формой.

Следующий пример даётся в лемме.

Лемма 5.1.1. *В* n*-мерном пространстве* E *с базисом* $e = (e_1, \ldots, e_n)$ *определим отображение* $B\colon E \times E \to P$ *по правилу*

$$B(f, g) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \alpha_i \beta_j, \tag{116}$$

для произвольных $f = \sum\limits_{i=1}^{n} \alpha_i e_i$, $g = \sum\limits_{i=1}^{n} \beta_i e_i$, *где* b_{ij}, $i, j = \overline{1, n}$ *— элементы произвольной матрицы* D *из* $P^{n \times n}$. *Тогда* $B(f, g)$ *является билинейной формой, а формулу (116) можно записать в матричном виде:*

$$B(f, g) = [f]_e^T D [g]_e, \tag{117}$$

где $[f]_e^T = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}^T = (\alpha_1\, \alpha_2, \dots\, \alpha_n), \quad [g]_e = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix}.$

Доказательство. 1) $g = \sum\limits_{i=1}^{n} \beta_i e_i,\ f_1 = \sum\limits_{i=1}^{n} \alpha_i' e_i,\ f_2 = \sum\limits_{i=1}^{n} \alpha_i'' e_i \Rightarrow f_1 + f_2 = = \sum\limits_{i=1}^{n} (\alpha_i' + \alpha_i'') e_i.$ $B(f_1 + f_2,\, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij}(\alpha_i' + \alpha_i'')\beta_j = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} (b_{ij}\alpha_i'\beta_j + b_{ij}\alpha_i''\beta_j) = = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij}\alpha_i'\beta_j + \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij}\alpha_i''\beta_j = B(f_1,\, g) + B(f_2,\, g).$

2) $f = \sum\limits_{i=1}^{n} \alpha_i e_i,\ g = \sum\limits_{i=1}^{n} \beta_i e_i \Rightarrow \gamma f = \sum\limits_{i=1}^{n} (\gamma\alpha_i) e_i.$

$B(\gamma f,\, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij}\alpha_i(\gamma\beta_j) = \gamma \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij}\alpha_i\beta_j = \gamma B(f,\, g).$

Аналогично проверяются свойства 3) и 4).

Совпадение (117) с (116) проверяется умножением.

Теорема 5.1.1 (Простейшие свойства билинейной формы).

1. $\forall\, m, r \in \mathbb{N},\ f_1,\, \dots,\, f_m,\, g_1,\, \dots,\, g_r, \in E,\ \alpha_1,\, \dots,\, \alpha_m,\, \beta_1,\, \dots,\, \beta_r \in P$
$$B\Big(\sum_{i=1}^{m} \alpha_i f_i, \sum_{j=1}^{r} \beta_j g_j\Big) = \sum_{i=1}^{m} \sum_{j=1}^{r} \alpha_i\beta_j B(f_i, g_j). \tag{118}$$
2. $B(\theta,\, g) = B(g,\, \theta) = 0$ *для всех* $g \in E$ *и, в частности,* $B(\theta,\, \theta) = 0.$

Доказательство. Проведите самостоятельно. ■

Докажем существование и единственность билинейной формы, принимающей заданные значения на парах базисных векторов.

Теорема 5.1.2. *Пусть* E *линейное пространство над полем* P *и* $e = (e_1, \dots, e_n)$ *базис* E.

Для любых чисел $b_{ij},\ i, j = \overline{1, n}$ *из* P *существует, и притом единственная, билинейная форма* $B(f, g)$ *в пространстве* E, *для которой* $B(e_i, e_j) = = b_{ij},\ i, j = \overline{1, n}.$

Доказательство. Отображение $B(f, g)$, определённое формулой (116) является билинейной формой. Для e_i i-я координата 1, для e_j j-я координата 1, остальные координаты равны нулю. Тогда из формулы (116) следует, что

$B(e_i, e_j) = b_{ij}$, $i, j = \overline{1, n}$. Если другая билинейная форма $B_1(f, g)$ удовлетворяет условиям $B_1(e_i, e_j) = b_{ij}$, $i, j = \overline{1, n}$, то из (118) следует, что для произвольных $f = \sum\limits_{i=1}^{n} \alpha_i e_i$, $g = \sum\limits_{j=1}^{n} \beta_j e_j$ значение $B_1(f, g) = B_1\big(\sum\limits_{i=1}^{n} \alpha_i e_i, \sum\limits_{j=1}^{n} \beta_j e_j\big) = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} \alpha_i \beta_j B_1(e_i, e_j) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \alpha_i \beta_j$ совпадает с $B(f, g)$ в формуле (116). Следовательно, формы B и B_1 равны. ■

Таким образом, в лемме 5.1.1 приведён самый общий вид билинейной формы в n-мерном пространстве.

Определение 2. *Пусть $B(f, g)$ — билинейная форма в n-мерном пространстве E и $e = (e_1, \ldots, e_n)$ — базис в E.*

Матрица $[B]_e \in P^{n \times n}$ с элементами $b_{ij} = B(e_i, e_j)$, $i, j = \overline{1, n}$ называется матрицей билинейной формы $B(f, g)$ в базисе e.■

Теорема 5.1.3. *Билинейная форма $B(f, g)$ в n-мерном пространстве E симметрична тогда и только тогда, когда её матрица симметрична в любом базисе.*

Доказательство.

Пусть $B(f, g)$ — билинейная форма в n-мерном пространстве E, $e = (e_1, \ldots, e_n)$ — базис в E и $[B]_e = (b_{ij})$ — матрица билинейной формы $B(f, g)$ в базисе e.

Если билинейная форма симметрична, то при всех векторах f, g выполнено равенство $B(f, g) = B(g, f)$, в частности, при всех $i, j = \overline{1, n}$ $B(e_i, e_j) = B(e_j, e_i)$, т.е. $b_{ij} = b_{ji}$. Значит, матрица симметрична.

Обратное утверждение.

Пусть в базисе e матрица билинейной формы симметрична, т.е. $b_{ij} = b_{ji}$ при всех $i, j = \overline{1, n}$. Для произвольных $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ $g = \sum\limits_{i=1}^{n} \beta_i e_i$ из формулы (116) получаем $B(f, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \alpha_i \beta_j$ и $B(g, f) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \beta_i \alpha_j$, а из симметричности матрицы имеем $B(g, f) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ji} \alpha_j \beta_i$. Так как $\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \alpha_i \beta_j = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ji} \alpha_j \beta_i$, то $B(f, g) = B(g, f)$. ■

Определение 3. *Пусть $B(f,g)$ — симметричная билинейная форма в пространстве E над полем P.*

Квадратичной формой называется отображение $K\colon E \to P$, определённое для произвольного $f \in E$ формулой $K(f) = B(f,f)$.

При этом, форма $B(f,g)$ называется полярной билинейной формой к квадратичной форме $K(f)$.

Матрицей квадратичной формы $K(f)$ в базисе $e = (e_1, \ldots, e_n)$ называется матрица полярной к ней билинейной формы $B(f,g)$ в том же базисе. Обозначим эту матрицу $[K]_e$. Таким образом, $[K]_e = [B]_e$. ■

Запишем в координатном виде квадратичную форму.

Лемма 5.1.2. *Пусть $K(f)$ квадратичная форма с матрицей $[K]_e = (b_{ij})$ в базисе $e = (e_1, \ldots, e_n)$ и произвольный вектор f задан своими координатами $f = \sum\limits_{i=1}^{n} x_i e_i$.*

Тогда

$$K(f) = \sum_{i,j=1}^{n} b_{ij} x_i x_j = \sum_{i=1}^{n} b_{ii} x_i^2 + 2\sum_{i<j} b_{ij} x_i x_j = b_{11}x_1^2 + 2b_{12}x_1x_2 + \ldots + 2b_{1n}x_1x_n +$$
$$+ b_{22}x_2^2 + 2b_{23}x_2x_3 + \ldots + 2b_{2n}x_2x_n + \ldots + 2b_{n-1,n}x_{n-1}x_n + b_{nn}x_n^2. \quad (119)$$

Доказательство. Следует из (116) и симметричности матрицы. ■

5.2. Канонический вид квадратичной формы. Приведение к каноническому виду

Меняя базис пространства, можно изменить вид матрицы квадратичной формы и упростить её.

Определение 1. *Базис $e = (e_1, \ldots, e_n)$ называется каноническим базисом квадратичной формы, если квадратичная форма K имеет в этом базисе диагональную матрицу $[K]_e = diag(d_1, d_2, \ldots, d_n)$.*

В каноническом базисе квадратичная форма согласно (119) имеет вид $K(f) = d_1x_1^2 + d_2x_2^2 + \ldots + d_nx_n^2$, который называется каноническим видом, а числа $d_1, d_2, \ldots, d_n$ — каноническими коэффициентами. Канонический вид называют также суммой квадратов. ■

Количестве ненулевых канонических коэффициентов совпадает с рангом матрицы квадратичной формы.

Теорема 5.2.1 (О приведении к каноническому виду).

Всякую квадратичную форму невырожденным преобразованием координат можно привести к каноническому виду.

Доказательство. Пусть дана квадратичная форма $K = \sum\limits_{i,j=1}^{n} b_{ij}x_ix_j$ от n координат $x_1, x_2, \ldots, x_n$. Постараемся представить K в виде суммы двух слагаемых, первое – это квадрат выражения, содержащего все слагаемые с одной их координат, второе — квадратичная формы от остальных $n-1$ координат. Эта цель легко достигается в случае, когда среди диагональных коэффициентов матрицы $b_{11}, b_{22}, \ldots, b_{nn}$, стоящих при квадратах координат, есть отличные от нуля.

Пусть, например, $b_{11} \neq 0$. Из (119) выделим и преобразуем все слагаемые, содержащие x_1 :

$K = \sum\limits_{i=1}^{n} b_{ii}x_i^2 + 2\sum\limits_{i<j} b_{ij}x_ix_j = (b_{11}x_1^2 + 2\sum\limits_{j=2}^{n} b_{1j}x_1x_j) + \sum\limits_{i,j=2}^{n} b_{ij}x_ix_j = S_1 + S_2.$

$S_1 = b_{11}x_1^2 + 2b_{12}x_1x_2 + \ldots + 2b_{1n}x_1x_n = b_{11}x_1^2 + 2x_1(b_{12}x_2 + \ldots + b_{1n}x_n) = = \dfrac{1}{b_{11}}\Big((b_{11}x_1)^2 + 2(b_{11}x_1)(b_{12}x_2 + \ldots + b_{1n}x_n)\Big) = \dfrac{1}{b_{11}}\Big(\Big(b_{11}x_1)^2 + 2(b_{11}x_1)(b_{12}x_2 + \ldots + b_{1n}x_n) + (b_{12}x_2 + \ldots + b_{1n}x_n)^2\Big) - (b_{12}x_2 + \ldots + b_{1n}x_n)^2\Big) = \dfrac{1}{b_{11}}\Big((b_{11}x_1 + b_{12}x_2 + \ldots + 2b_{1n}x_n)^2 - (b_{12}x_2 + \ldots + b_{1n}x_n)^2\Big) = \dfrac{1}{b_{11}}(b_{11}x_1 + b_{12}x_2 + \ldots + 2b_{1n}x_n)^2 - \dfrac{1}{b_{11}}(b_{12}x_2 + \ldots + b_{1n}x_n)^2$

Итак, $K = \dfrac{1}{b_{11}}(b_{11}x_1 + b_{12}x_2 + \ldots + 2b_{1n}x_n)^2 + \Big(\sum\limits_{i,j=2}^{n} b_{ij}x_ix_j - \dfrac{1}{b_{11}}(b_{12}x_2 + \ldots + 2b_{1n}x_n)^2\Big) = \dfrac{1}{b_{11}}(b_{11}x_1 + b_{12}x_2 + \ldots + 2b_{1n}x_n)^2 + K_1$, где K_1 является квадратичной формой от координат $x_2, \ldots, x_n$.

Если мы теперь введём обозначения

$$y_1 = b_{11}x_1 + b_{12}x_2 + \ldots + 2b_{1n}x_n, \quad y_i = x_i, \ i = \overline{2, n}, \tag{120}$$

то получим

$$K = \frac{1}{b_{11}}y_1^2 + K_1,$$

где теперь K_1 будет теперь квадратичной формой от координат $y_2, \ldots, y_n$. Равенства (120) определяют линейное преобразование координат с невырожденной матрицей, т.к. её определитель есть $b_{11} \neq 0$. В формулах преобразования координат (76) старые координаты выражаются через новые, но в них участвует обратная матрица, которая также невырождена.

Теперь рассмотрим случай, когда все диагональные элементы матрицы квадратичной формы равны нулю $a_{11} = a_{22} = \ldots = a_{nn} = 0$. Если все элементы матрицы равны нулю, то она имеет канонический вид с нулевыми каноническими коэффициентами. Поэтому предположим, что среди элементов матрицы есть ненулевые, например, $b_{12} \neq 0$. Тогда $K = 2b_{12}x_1x_2 + K_2$, где каждое слагаемое в K_2 содержит хотя бы одну координату $x_3, \ldots, x_n$. Совершим линейное преобразование координат

$$y_1 = z_1 - z_2,\ y_2 = z_1 + z_2 \quad y_i = x_i,\ i = \overline{3, n}, \tag{122}$$

Оно будет невырожденным, т.к. определитель

$$\begin{vmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} = 2 \neq 0.$$

В результате член $2b_{12}x_1x_2$ примет вид

$$2b_{12}x_1x_2 = 2b_{12}(z_1 - z_2)(z_1 + z_2) = 2b_{12}z_1^2 - 2b_{12}z_2^2,$$

т.е. в форме появятся два квадрата $b_{12}z_1^2$ и $-2b_{12}z_2^2$ с ненулевыми коэффициентами, причём они не могут сократиться ни с одним из остальных членов, т.к. в каждый из них входит хотя бы одна координата $z_3, \ldots, z_n$. Теперь мы находимся в условиях первого случая и можем ещё одним невырожденным преобразованием привести её к виду (121).

Применяя это рассуждение к квадратичной форме с меньшим числом координат, мы за конечное число шагов получим канонический вид. Выполнив при этом все преобразования координат, мы получим формулы, связывающие начальные и конечные координаты. ■

Список литературы

1. Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры : учебник для студ. вузов / Д.В. Беклемишев. — М. : Физматлит, 2007. — 307 с.

2. Гантмахер Ф.Р. Теория матриц / Ф.Р. Гантмахер. — М. : Физматлит, 2004. — 559 с.

3. Ильин В.А. Линейная алгебра: учебник для студ. физич. специальностей и специальности «Прикладная математика» / В.А. Ильин, Э.Г. Позняк. — М. : Физматлит, 2006. — 278 с.

4. Ильин В.А. Линейная алгебра и аналитическая геометрия : учебник для студ. ун-тов и техн. вузов, обуч. по специальностям «Математика», «Прикладная математика и информатика» / В.А. Ильин, Г.Д. Ким. — М. : Изд-во МГУ, Проспект, 2012. — 388 с.

5. Кострикин А.И. Введение в алгебру. Ч. 2. Линейная алгебра : учебник для студентов университетов, обучающихся по специальности «Математика» и «Прикладная математика» / А.И. Кострикин. — М. : Физматлит, 2004. — 368 с.

6. Курош А.Г. Курс высшей алгебры / А.Г. Курош. — СПб. [и др.] : Лань, 2007. — 431 с.

7. Проскуряков И.В. Сборник задач по линейной алгебре / И.В. Проскуряков. — СПб ; М. ; Краснодар : Лань, 2007. — 475 с.

8. Фаддеев Д.К. Сборник задач по высшей алгебре : учеб. пособие для студ. вузов, обуч. по матем. специальностям / Д.К. Фаддеев, И.С. Соминский. — СПб. : Лань, 2004. — 287 с.