Учебное издание

КАДОМЦЕВ Сергей Борисович

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ И ЛИНЕЙНАЯ АЛГЕБРА

Редакторы *Н.Б. Бартошевич-Жагель*, *В.С. Аролович*
Оригинал-макет: *Д.В. Горбачёв*
Оформление переплета: *А.В. Андросов*

Подписано в печать 13.12.10. Формат 60×90/16. Бумага офсетная.
Печать офсетная. Усл. печ. л. 10,5. Уч.-изд. л. 11,5. Тираж 500 экз.
Заказ №

Издательская фирма «Физико-математическая литература»
МАИК «Наука/Интерпериодика»
117997, Москва, ул. Профсоюзная, 90
E-mail: fizmat@maik.ru, fmlsale@maik.ru;
http://www.fml.ru

Отпечатано в ПФ «Полиграфист»
160001, г. Вологда, ул. Челюскинцев, 3
Тел.: (8172) 72-07-92, 72-61-75, 72-60-63; факс: (8172) 76-00-49, 72-71-11
E-mail: forma@pfpoligrafist.com

ISBN 978-5-9221-1290-1

9

УДК 514
ББК 22.143
К 13

Кадомцев С. Б. **Аналитическая геометрия и линейная алгебра.** — 2-е изд., испр. и доп. — М.: ФИЗМАТЛИТ, 2011. — 168 с. — ISBN 978-5-9221-1290-1.

Настоящее пособие написано на основе курса лекций, читаемого автором на физическом факультете МГУ. Книга состоит из трех частей. В первой из них (аппарат аналитической геометрии и линейной алгебры) рассматриваются действия с матрицами, теория определителей и ее приложения к решению систем линейных уравнений. Во второй части (аналитическая геометрия) помимо традиционного материала подробно обсуждается теория ориентации, строится классификация кривых и поверхностей второго порядка. Третья часть (линейная алгебра) представляет собой систематическое изложение теории линейных, евклидовых и унитарных пространств, основанное на аксиоматике Вейля.

Книга предназначена, прежде всего, студентам физико-математических специальностей.

ISBN 978-5-9221-1290-1

© ФИЗМАТЛИТ, 2001, 2011
© С. Б. Кадомцев, 2001, 2011

СОДЕРЖАНИЕ

Часть I. АППАРАТ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ И ЛИНЕЙНОЙ АЛГЕБРЫ 7

Глава 1
МАТРИЦЫ И ОПРЕДЕЛИТЕЛИ 8

§ 1. Матрицы 8
1. Сложение матриц (8). 2. Умножение матрицы на число (9). 3. Арифметическое пространство (9).

§ 2. Определители 11
1. Предварительные замечания (11). 2. Определитель (12). 3. Разложение определителя по строке (14). 4. Основные свойства определителя (14).

§ 3. Равноправность строк и столбцов определителя 16
1. Перестановки (16). 2. Выражение определителя через его элементы (18). 3. Алгебраическое дополнение (19). 4. Разложение определителя по столбцу (20).

§ 4. Произведение матриц 21
1. Свойства произведения матриц (21). 2. Определитель произведения квадратных матриц (22). 3. Свойства произведения квадратных матриц (23).

§ 5. Базисный минор 24
1. Теорема о базисном миноре (24). 2. Ранг матрицы (26).

Глава 2
СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ 27

§ 1. Существование и единственность решения 27
1. Основные определения (27). 2. Существование решения (28). 3. Однородные системы (28). 4. Единственность решения (28).

§ 2. Нахождение решений 29
1. Формулы Крамера (29). 2. Общий случай (30).

Часть II. АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ 31

Предварительные замечания 32

Глава 1
ВЕКТОРЫ И КООРДИНАТЫ 33

§ 1. Координаты точки 33
1. Ось координат (33). 2. Декартовы координаты (34). 3. Криволинейные координаты на плоскости (34). 4. Криволинейные координаты в пространстве (35).

§ 2. Векторы 36
1. Вектор (36). 2. Равенство векторов (37). 3. Координаты вектора (37). 4. Сумма векторов (38). 5. Произведение вектора на число (38). 6. Отождествление равных векторов (39).

§ 3. Скалярное произведение 39
1. Основные определения (39). 2. Скалярное произведение в координатах (40). 3. Свойства скалярного произведения (40). 4. Площадь параллелограмма (41). 5. Объем параллелепипеда (41).

§ 4. Базис 42
1. Коллинеарные векторы (42). 2. Компланарные векторы (43). 3. Линейная зависимость четырех векторов (43). 4. Базис (43). 5. Аффинные координаты (44).

Глава 2
ПРЕОБРАЗОВАНИЕ КООРДИНАТ 45

§ 1. Преобразование координат на плоскости 45
1. Правые и левые пары (45). 2. Собственные и несобственные преобразования (46). 3. Преобразование координат вектора (47). 4. Правые системы координат (47). 5. Преобразование координат точки (48).

§ 2. Преобразование координат в пространстве 49
1. Преобразование координат вектора (49). 2. Углы Эйлера (50). 3. Преобразование координат точки (51).

§ 3. Векторное произведение 52
1. Определение векторного произведения (52). 2. Смешанное произведение (53). 3. Произведение двух смешанных произведений (53). 4. Скалярное произведение двух векторных произведений (54). 5. Двойное векторное произведение (54).

Глава 3
УРАВНЕНИЯ ПРЯМЫХ И ПЛОСКОСТЕЙ 55

§ 1. Прямая на плоскости 55
1. Уравнение прямой (55). 2. Параметрические уравнения прямой (55). 3. Каноническое уравнение прямой (56). 4. Общее уравнение прямой (56). 5. Нормированное уравнение прямой (56).

§ 2. Плоскость. 57
1. Уравнение плоскости (57). 2. Общее уравнение плоскости (58). 3. Нормированное уравнение плоскости (59).

§ 3. Прямая в пространстве 59
1. Уравнение прямой (59). 2. Параметрические уравнения прямой (60). 3. Канонические уравнения прямой (60). 4. Пересечение двух плоскостей (60).

Глава 4
ЛИНИИ И ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА 62

§ 1. Эллипс, гипербола и парабола 62
1. Эллипс (62). 2. Гипербола (64). 3. Директриса эллипса и гиперболы (66). 4. Парабола (67). 5. Касательная (68). 6. Оптические свойства (69).

§ 2. Кривые второго порядка 71
1. Уравнение кривой второго порядка (71). 2. Классификация (72).

§ 3. Поверхности второго порядка 73
1. Уравнение поверхности второго порядка (73). 2. Цилиндры (76). 3. Конусы (76). 4. Завершение классификации (78).

§ 4. Эллипсоид, гиперболоиды и параболоиды 79
1. Эллипсоид (79). 2. Гиперболоиды (80). 3. Параболоиды (81).

Часть III. ЛИНЕЙНАЯ АЛГЕБРА 83

Глава 1
КОНЕЧНОМЕРНОЕ ЛИНЕЙНОЕ ПРОСТРАНСТВО 84

§ 1. Линейное пространство 84
1. Аксиомы Вейля (84). 2. Линейное пространство (85). 3. Свойства линейного пространства (86). 4. Линейное подпространство (87).

§ 2. n-мерное линейное пространство 88
1. Базис и размерность (88). 2. Примеры (89). 3. Изоморфизм (90). 4. Линейное дополнение (91).

Глава 2
ЛИНЕЙНЫЕ ОПЕРАТОРЫ 94

§ 1. Операторы, действующие из $\mathbf{L}^n$ в $\mathbf{L}^m$ 94
1. Линейный оператор (94). 2. Матрица линейного оператора (95). 3. Образ оператора (96). 4. Ядро оператора (96). 5. Произведение операторов (97).

§ 2. Операторы, действующие из $\mathbf{L}^n$ в $\mathbf{L}^n$ 98
1. Тождественный оператор и обратный оператор (98). 2. Инвариантные подпространства (99). 3. Образ и ядро (100). 4. Структура пространства $\mathbf{L}_0$ (101). 5. Собственные значения и собственные векторы (107). 6. Характеристическое уравнение (108). 7. Жорданова форма матрицы линейного оператора (110).

Глава 3
ПРЕОБРАЗОВАНИЕ БАЗИСОВ И КООРДИНАТ 114

§ 1. Преобразование базисов 114
1. Обозначения (114). 2. Переход к новому базису (114). 3. Последовательные преобразования (115).

§ 2. Преобразование координат 115
1. Преобразование координат вектора (115). 2. Преобразование матрицы линейного оператора (115). 3. Линейная форма (116).

§ 3. Тензоры 117
1. Определение тензора (117). 2. Сумма тензоров одинаковой структуры (118). 3. Прямое произведение тензоров (119). 4. Свертка тензора (119). 5. О билинейной форме (120).

§ 4. Квадратичные формы 121
1. Матрица квадратичной формы (121). 2. Метод Лагранжа (121). 3. Закон инерции (122). 4. Критерий Сильвестра (123).

Глава 4
ЕВКЛИДОВО ПРОСТРАНСТВО 125

§ 1. Длины и углы 125
1. Определение евклидова пространства (125). 2. Неравенство Коши–Буняковского (125). 3. Длина вектора (126). 4. Угол между векторами (126).

§ 2. Ортонормированный базис 127
1. Существование ортонормированного базиса (127). 2. Ортогонализация (127). 3. Ортогональное дополнение (128). 4. Альтернатива Фредгольма (128).

§ 3. Операторы в E^n 129
1. Сопряженный оператор (129). 2. Ортогональный оператор (130). 3. Ортогональные преобразования (132). 4. Самосопряженный оператор (133). 5. Квадратичная форма в E^n (134).

§ 4. Гиперповерхности второго порядка 134
1. Система координат (134). 2. Каноническое уравнение гиперповерхности второго порядка (135). 3. Классификация (136). 4. Инварианты (137).

Глава 5
НЕКОТОРЫЕ ОБОБЩЕНИЯ 139

§ 1. Унитарное пространство 139
1. Основные свойства (139). 2. Нормальный оператор (140). 3. Унитарный оператор (141). 4. Самосопряженный оператор (142).

§ 2. Псевдоевклидово пространство 142
1. Определение (142). 2. Преобразования Лоренца (143).

§ 3. Группы и поля 145
1. Группа (145). 2. Примеры (146). 3. Поле (147).

Заключение 150

Приложение
КОМПЛЕКСНЫЕ ЧИСЛА. ОСНОВНАЯ ТЕОРЕМА АЛГЕБРЫ . . 156
1. Определения (156). 2. Геометрическая интерпретация (157). 3. Многочлен (158). 4. Вещественный многочлен (159). 5. Теорема о модуле многочлена (159). 6. Теорема о вложенных отрезках (160). 7. Основная теорема алгебры (161). 8. Следствия из основной теоремы алгебры (162).

Предметный указатель 164

Часть I

АППАРАТ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ И ЛИНЕЙНОЙ АЛГЕБРЫ

Глава 1
МАТРИЦЫ И ОПРЕДЕЛИТЕЛИ

§ 1. Матрицы

1. Сложение матриц.

Определение 1. *Таблица из чисел* [1] *вида*

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

называется $(m \times n)$*-матрицей.*

Здесь первый индекс у числа a_{ij} — это номер строки, а второй — номер столбца, в котором это число находится.

Иногда мы будем использовать также обозначение $(A)_{ij}$, понимая под ним элемент матрицы A с индексами i, j, т. е. a_{ij}.

Определение 2. *Суммой двух* $(m \times n)$*-матриц* A *и* B *называется такая* $(m \times n)$*-матрица* C, *что* $(C)_{ij} = a_{ij} + b_{ij}$.

Теорема. *Сложение матриц обладает следующими свойствами*:

1° *для любых* $(m \times n)$*-матриц* A *и* B *выполняется равенство* $A + B = B + A$;

2° *для любых* $(m \times n)$*-матриц* A, B *и* C *выполняется равенство* $A + (B + C) = (A + B) + C$;

3° *существует единственная* $(m \times n)$*-матрица* O *такая, что для любой* $(m \times n)$*-матрицы* A *выполняется равенство* $A + O = A$;

4° *для любой* $(m \times n)$*-матрицы* A *существует единственная матрица* $(-A)$ *такая, что* $A + (-A) = O$.

Доказательство. 1°. Выберем в матрице $(A + B)$ произвольный элемент $(A + B)_{ij}$. Имеем: $(A + B)_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = (B + A)_{ij}$. Следовательно, $A + B = B + A$.

2°. $[A+(B+C)]_{ij} = a_{ij}+(b_{ij}+c_{ij}) = (a_{ij}+b_{ij})+c_{ij} = [(A+B)+C]_{ij}$, а значит, $A + (B + C) = (A + B) + C$.

3°. Согласно определению, $(A + B)_{ij} = a_{ij} + b_{ij}$. Это число равно a_{ij} тогда и только тогда, когда b_{ij} равно нулю. Следовательно, единственной матрицей O, удовлетворяющей условию $A + O = A$, является матрица O, все элементы которой равны нулю.

---

[1] Под словом «число» здесь и далее (в части I) понимается, вообще говоря, комплексное число; впрочем, если угодно, можно считать, что речь идет о вещественном числе.

4°. Равенство $(A+B)_{ij} = a_{ij} + b_{ij} = 0$ выполняется тогда и только тогда, когда b_{ij} равно $(-a_{ij})$. Следовательно, единственной матрицей $(-A)$, удовлетворяющей условию $A + (-A) = O$, является матрица $(-A)$, элементы которой соответственно равны $(-a_{ij})$. Теорема доказана.

Замечание. Матрица $A+(-B)$ называется *разностью матриц* A *и* B и обозначается так: $A - B$.

2. Умножение матрицы на число.

Определение. *Произведением* $(m \times n)$*-матрицы* A *на число* λ *называется такая* $(m \times n)$*-матрица* B, что $(B)_{ij} = \lambda a_{ij}$.

Теорема. *Умножение матрицы на число обладает следующими свойствами*:

1° *для любой* $(m \times n)$*-матрицы* A *имеет место равенство* $1 \cdot A = A$;

2° *для любой* $(m \times n)$*-матрицы* A *и любых чисел* λ *и* μ *имеет место равенство* $\lambda(\mu A) = (\lambda\mu)A$;

3° *для любой* $(m \times n)$*-матрицы* A *и любых чисел* λ *и* μ *имеет место равенство* $(\lambda + \mu)A = \lambda A + \mu A$;

4° *для любых* $(m \times n)$*-матриц* A *и* B *и любого числа* λ *выполняется равенство* $\lambda(A + B) = \lambda A + \lambda B$.

Доказательство. Имеем:

1°. $(1 \cdot A)_{ij} = 1 \cdot a_{ij} = a_{ij}$, поэтому $1 \cdot A = A$.

2°. $[\lambda(\mu A)]_{ij} = \lambda(\mu a_{ij}) = (\lambda\mu)a_{ij} = [(\lambda\mu)A]_{ij}$, т. е. $\lambda(\mu A) = (\lambda\mu)A$.

3°. $[(\lambda + \mu)A]_{ij} = (\lambda + \mu)a_{ij} = \lambda a_{ij} + \mu a_{ij} = (\lambda A + \mu A)_{ij}$, т. е. $(\lambda + \mu)A = \lambda A + \mu A$.

4°. $[\lambda(A + B)]_{ij} = \lambda(a_{ij} + b_{ij}) = \lambda a_{ij} + \lambda b_{ij} = (\lambda A + \lambda B)_{ij}$, а значит, $\lambda(A + B) = \lambda A + \lambda B$.

Теорема доказана.

3. Арифметическое пространство.

Определение 1. *Множество всех упорядоченных наборов из* n *чисел* $(a_1, a_2, \ldots, a_n)$, *для которых определены операции сложения и умножения на число по правилам*:

1° $(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$;

2° $\lambda(a_1, a_2, \ldots, a_n) = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n)$,

называется арифметическим пространством.

Если числа, о которых идет речь, вещественные, то пространство обозначается символом $\mathbb{R}^n$, а если комплексные — то символом $\mathbb{C}^n$. Сами элементы арифметического пространства условимся обозначать одной буквой полужирного шрифта: $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ и называть их для краткости *строками*, хотя, конечно, записывать их можно и в виде столбцов.

Поскольку строки складываются и умножаются на число по тем же правилам, что и $(1 \times n)$-матрицы, то сложение строк и умножение их на число удовлетворяет свойствам, указанным в теоремах пп. 1, 2. Это позволяет, в частности, сформулировать следующее определение.

Определение 2. *Сумма вида* $\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_k\mathbf{a}_m$ *называется линейной комбинацией строк* $\mathbf{a}_i$ *с коэффициентами* λ_i.

Если все коэффициенты равны нулю, то линейная комбинация называется *тривиальной*, в противном случае (т. е. если хотя бы один коэффициент отличен от нуля) — *нетривиальной*.

Среди всевозможных срок особую роль играют строки

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, \ldots, 0),\ \ldots,\ \mathbf{e}_n = (0, 0, \ldots, 1) \qquad (1)$$

(т. е. строки $\mathbf{e}_i$, у которых на i-м месте стоит 1, а в остальных местах — 0), поскольку любая строка может быть представлена и притом единственным образом в виде их линейной комбинации:

$$\mathbf{a} = (a_1, a_2, \ldots, a_n) = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \ldots + a_n\mathbf{e}_n.$$

Строки (1) условимся в дальнейшем называть *координатными строками*.

Определение 3. *Строки называются линейно зависимыми, если существует их нетривиальная линейная комбинация, равная нулевой строке* $\mathbf{o} = (0, 0, \ldots, 0)$; *в противном случае они называются линейно независимыми*.

Тем самым, можно сказать так: *строки называются линейно независимыми, если обращение в нулевую строку их линейной комбинации возможно лишь в том случае, когда эта линейная комбинация тривиальна*. Примером линейно независимых строк могут служить, очевидно, координатные строки (1).

Докажем три теоремы о линейной зависимости строк.

Теорема 1. *Если среди строк есть нулевая, то эти строки линейно зависимы*.

Доказательство. Пусть, например, $\mathbf{a}_1 = (0, 0, \ldots, 0)$ (этого всегда можно достичь, занумеровав строки соответствующим образом). Имеем: $1\mathbf{a}_1 + 0\mathbf{a}_2 + \ldots + 0\mathbf{a}_m = (0, 0, \ldots, 0)$, а значит, строки $\mathbf{a}_1, \mathbf{a}_2, \ldots$ $\ldots, \mathbf{a}_m$ линейно зависимы. Теорема доказана.

Теорема 2. *Если какие-нибудь k из m строк линейно зависимы, то и все строки линейно зависимы*.

Доказательство. Пусть, например, строки $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно зависимы, т. е. существуют такие числа $\lambda_1, \lambda_2, \ldots, \lambda_k$, что $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \ldots + \lambda_k\mathbf{a}_k = (0, 0, \ldots, 0)$, причем не все λ_i равны нулю. Имеем: $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \ldots + \lambda_k\mathbf{a}_k + 0\mathbf{a}_{k+1} + \ldots + 0\mathbf{a}_m = (0, 0, \ldots, 0)$. Но это и означает, что строки $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ линейно зависимы. Теорема доказана.

Теорема 3. *Если строки линейно зависимы, то одна из них равна линейной комбинации остальных*.

Доказательство. Если строки $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ линейно зависимы, то существуют такие числа $\lambda_1, \lambda_2, \ldots, \lambda_m$, что $\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_m\mathbf{a}_m = (0, 0, \ldots, 0)$, причем не все λ_i равны нулю. Пусть, например, $\lambda_1 \neq 0$. Имеем: $\mathbf{a}_1 = -\dfrac{\lambda_2}{\lambda_1}\mathbf{a}_2 - \ldots - \dfrac{\lambda_m}{\lambda_1}\mathbf{a}_m$, что и требовалось доказать.

§ 2. Определители

1. Предварительные замечания. Если количество строк матрицы равно количеству ее столбцов, то матрица называется *квадратной*. В этом параграфе речь пойдет о квадратных матрицах.

Начнем с простого примера. Рассмотрим систему двух линейных уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{aligned}\right\}$$

с двумя неизвестными x_1 и x_2. Умножим первое уравнение на a_{22}, второе — на $(-a_{12})$ и сложим их. В результате получим:

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12},$$

или, если выражение в скобках не обращается в нуль,

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Введем обозначение:

$$\begin{vmatrix} a_{11} a_{12} \\ a_{21} a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{1}$$

Тогда полученный результат можно записать так:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Аналогично

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Выражение (1) называется *определителем второго порядка матрицы* $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Определитель матрицы A обозначают также $\det A$.

Нетрудно убедиться в том, что решение системы трех линейных уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{aligned}\right\}$$

с тремя неизвестными x_1, x_2 и x_3 можно записать так:

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A} \quad , \quad x_3 = \frac{\det A_3}{\det A},$$

где

$$A_1 = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix},$$

$$A_3 = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{2}$$

(если, конечно, $\det A \neq 0$). Выражение (2) называется *определителем третьего порядка* матрицы A.

Определители играют весьма важную роль во многих разделах математики. Например, как мы вскоре увидим, определитель второго порядка с точностью до знака равен площади параллелограмма, построенного на векторах, координаты которых являются его строками, а объем параллелепипеда, построенного на данных трех векторах, равен модулю определителя, строками которого являются координаты этих векторов. При этом оказывается, что формулы (1) и (2) имеют простой геометрический смысл: формула (1) выражает тот факт, что площадь параллелограмма равна произведению его основания на высоту, а формула (2) — тот факт, что объем параллелепипеда равен произведению площади его основания на высоту.

Наша ближайшая задача состоит в том, чтобы, исходя из формул (1), (2), обобщить понятие определителя матрицы на случай **квадратной матрицы** с произвольным количеством строк. К решению этой задачи мы и переходим.

2. Определитель. Обратимся к формуле (2). Мы видим, что определитель третьего порядка представляет собой алгебраическую сумму трех слагаемых, знаки в которой чередуются, причем первый знак — плюс. Далее, каждое слагаемое представляет собой произведение элемента первой строки на определитель матрицы, полученной из исходной вычеркиванием первой строки и того столбца, из которого этот элемент взят. Отметим, что если договориться называть определителем первого порядка матрицы $A = (a_{11})$ то единственное число a_{11}, из которого она состоит, то формула (1) примет вид, структурно аналогичный виду формулы (2). Эти наблюдения приводят нас к следующему определению.

Определение. 1° *Определителем первого порядка* (1×1)*-матрицы* A *называется то единственное число* a_{11}, *из которого эта матрица состоит*;

2° *определителем* n*-го порядка* $(n \times n)$*-матрицы* A *при* $n > 1$ *называется число*

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{s=1}^{n} (-1)^{s+1} a_{1s} \Delta_s, \tag{3}$$

где Δ_s — *определитель* $(n-1)$*-го порядка матрицы, получаемой из* A *вычеркиванием первой строки и* s*-го столбца.*

Ясно, что это определение позволяет найти выражение для определителя любого порядка. Например, зная формулу для определителя

третьего порядка, мы можем написать выражение для определителя четвертого порядка, а значит и пятого и т. д. Для краткости будем называть элементами, строками и столбцами определителя элементы, строки и столбцы его матрицы.

Замечание. Как следует из формулы (3), определитель представляет собой функцию n^2 переменных — элементов a_{ij}. Иногда, однако, бывает удобнее рассматривать его как функцию $\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ n переменных: $\mathbf{a}_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $\mathbf{a}_2 = (a_{21}, a_{22}, \ldots, a_{2n})$, $\ldots$, $\mathbf{a}_n = (a_{n1}, a_{n2}, \ldots, a_{nn})$ — строк этого определителя.

Одно из важнейших свойств определителя состоит в следующем.

Теорема. *Если две строки поменять местами, то модуль определителя не изменится, а его знак изменится на противоположный.*

Доказательство. Допустим, что меняются местами i-я и j-я строки, $i < j$. При $n = 2$ (при $n = 1$ нет двух различных строк) справедливость утверждения усматривается непосредственно из формулы (1), поэтому будем считать, что $n \geqslant 3$.

Рассмотрим сначала случай $i = 1$, $j = 2$. Применим формулу (3) к определителю $\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, а затем — к каждому из определителей Δ_{1s}. В результате определитель окажется записанным в виде алгебраической суммы определителей d_{pq} $(p < q)$, полученных из исходного вычеркиванием первых двух строк и столбцов с номерами p, q, с некоторыми коэффициентами. При этом слагаемых с d_{pq} будет два: одно получится в результате вычеркивания первой сроки и p-го столбца, второй строки и q-го столбца:

$$\begin{vmatrix} \cancel{a_{11}} & \cancel{\ldots} & \cancel{a_{1p}} & \cancel{\ldots} & \cancel{a_{1q}} & \cancel{\ldots} \\ \cancel{a_{21}} & \cancel{\ldots} & \cancel{a_{2p}} & \cancel{\ldots} & \cancel{a_{2q}} & \cancel{\ldots} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \ldots & a_{np} & \ldots & a_{nq} & \ldots \end{vmatrix},$$

а другое — в результате вычеркивания первой сроки и q-го столбца, второй строки и p-го столбца:

$$\begin{vmatrix} \cancel{a_{11}} & \cancel{\ldots} & \cancel{a_{1p}} & \cancel{\ldots} & \cancel{a_{1q}} & \cancel{\ldots} \\ \cancel{a_{21}} & \cancel{\ldots} & \cancel{a_{2p}} & \cancel{\ldots} & \cancel{a_{2q}} & \cancel{\ldots} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \ldots & a_{np} & \ldots & a_{nq} & \ldots \end{vmatrix}.$$

Коэффициент при d_{pq} в первом слагаемом будет равен $(-1)^{p+1}a_{1p} \times (-1)^q a_{2q}$, поскольку после вычеркивания p-го столбца q-й столбец окажется на $(q-1)$-м месте, а во втором — равен $(-1)^{q+1}a_{1q}(-1)^{p+1}a_{2p}$, поскольку после вычеркивания q-го столбца p-й столбец останется на p-м месте. Таким образом, при d_{pq} окажется коэффициент

$$(-1)^{p+q+1}(a_{1p}a_{2q} - a_{2p}a_{1q}).$$

Коэффициент при d_{pq} в определителе $\Delta(\mathbf{a}_2, \mathbf{a}_1, \ldots, \mathbf{a}_n)$ можно найти, поменяв местами в полученном выражении индексы 1 и 2:

$$(-1)^{p+q+1}(a_{2p}a_{1q} - a_{1p}a_{2q}).$$

Найденные выражения равны по модулю и противоположны по знаку, что доказывает справедливость первого утверждения.

Перейдем теперь к общему случаю. Воспользуемся методом математической индукции: допустим, что теорема верна при некотором $n = k-1 \geqslant 2$ и докажем, что она верна и при $n = k$. Если $i, j > 1$, то при перестановке строк с номерами i и j в каждом из определителей Δ_{1s} $(k-1)$-го порядка в формуле (3) две строки меняются местами, а значит, по предположению индукции, перед всеми слагаемыми изменяется знак, что и требовалось доказать. Если же $i = 1$, $j > 2$, то, согласно ранее доказанному,

$$\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_j, \ldots) = -\Delta(\mathbf{a}_2, \mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots) =$$
$$= \Delta(\mathbf{a}_2, \mathbf{a}_j, \ldots, \mathbf{a}_1, \ldots) = -\Delta(\mathbf{a}_j, \mathbf{a}_2, \ldots, \mathbf{a}_1, \ldots).$$

Теорема доказана.

3. Разложение определителя по строке.

Определение. *Определитель* Δ_{ij}, *получаемый из* $\det A$ *вычеркиванием* i*-й строки и* j*-го столбца, называется минором, дополнительным к элементу* a_{ij}.

Теорема. *Справедлива следующая формула, называемая формулой разложения определителя по* i*-ой строке*:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}, \tag{4}$$

где Δ_{ij} *— минор, дополнительный к элементу* a_{ij}.

Доказательство. Воспользуемся методом математической индукции. При $i = 1$ формула (4) верна, поскольку совпадает с формулой (3). Допустим, что она верна для строки с номером $i-1$ $(i \leqslant n)$ и докажем, что тогда она верна и для строки с номером i. Поменяем местами i-ю и $(i-1)$-ю строки (при этом у определителя изменится знак) и, в соответствии с предположением индукции, разложим полученный определитель по строке с номером $i-1$:

$$\det A = \Delta(\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_i, \ldots, \mathbf{a}_n) = -\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i-1}, \ldots, \mathbf{a}_n) =$$
$$= -\sum_{j=1}^{n} (-1)^{i-1+j} a_{ij} \Delta_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}.$$

Теорема доказана.

4. Основные свойства определителя. Сформулируем несколько свойств определителя, вытекающих из установленных нами фактов.

1°. *Если все элементы какой-нибудь строки умножить на одно и то же число, то весь определитель умножится на это число.*

В самом деле, если разложить определитель по указанной строке, то в формуле (4) перед каждым слагаемым появится общий множитель. После вынесения его за скобки в скобках останется исходный определитель, что и требовалось доказать.

2°. *Если к строке определителя прибавить какую-нибудь строку* $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, *то его можно будет представить в виде суммы двух определителей: исходного и определителя, в котором указанная строка заменена на прибавленную:* $\Delta(\mathbf{a}_1 + \mathbf{b}, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) + \Delta(\mathbf{b}, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, $\Delta(\mathbf{a}_1, \mathbf{a}_2 + \mathbf{b}, \ldots, \mathbf{a}_n) = \Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) + \Delta(\mathbf{a}_1, \mathbf{b}, \ldots, \mathbf{a}_n), \ldots, \Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n + \mathbf{b}) = \Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) + \Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{b})$. В самом деле,

$$\Delta(\mathbf{a}_1 \ldots, \mathbf{a}_i + \mathbf{b}, \ldots, \mathbf{a}_n) = \sum_{j=1}^{n} (-1)^{i+j} (a_{ij} + b_j) \Delta_{ij} =$$

$$= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij} + \sum_{i=1}^{n} (-1)^{i+j} b_j \Delta_{ij} =$$

$$= \Delta(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n) + \Delta(\mathbf{a}_1, \ldots, \mathbf{b}, \ldots, \mathbf{a}_n),$$

что и требовалось доказать.

3°. *Если в определителе две строки одинаковые, то он равен нулю.*

Действительно, если указанные строки поменять местами, то определитель, с одной стороны, не изменится, а с другой — у него изменится знак. Это возможно лишь в том случае, когда он равен нулю.

4°. *Определитель* $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, *т. е. определитель*

$$\begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{vmatrix},$$

равен единице.

Для $n = 1$ это утверждение очевидно; если же оно доказано для $n = k - 1$, то при $n = k$, раскладывая данный определитель по первой строке, получим: $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) = 1 \cdot 1 = 1$.

Замечание. Свойства 1°, 2°, 3°, 4° иногда называют *основными свойствами определителя*, поскольку из них может быть выведена формула (3), и, следовательно, они могут быть положены в основу аксиоматического определения определителя.

Следствие 1. *Если в определителе две строки пропорциональны (т. е.* $\mathbf{a}_i = \lambda \mathbf{a}_j$ *и* $i \neq j$*), в частности одна из строк состоит из нулей* (случай $\lambda = 0$), *то он равен нулю.*

В самом деле, если, пользуясь свойством 1°, вынести общий множитель, то получится, что две строки в определителе совпадают и, следовательно, он равен нулю.

Следствие 2. *Если одна из строк равна линейной комбинации остальных, то определитель равен нулю.*

Согласно свойству 2°, такой определитель можно представить в виде суммы определителей, в каждом из которых две строки пропорциональны.

Следствие 3. *Если к какой-нибудь строке определителя прибавить линейную комбинацию остальных строк, то определитель не изменится.*

Действительно, согласно свойству 2° он может быть представлен в виде суммы двух определителей: исходного и определителя, в котором одна из строк равна линейной комбинации остальных.

§ 3. Равноправность строк и столбцов определителя

1. Перестановки. Для дальнейшего изучения свойств определителя нам понадобится формула, позволяющая вычислить определитель n-го порядка непосредственно через его элементы. Вывод этой формулы потребует от нас использования некоторых дополнительных фактов, к обсуждению которых мы и переходим.

Определение 1. *Упорядоченная совокупность* $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ *n попарно различных натуральных чисел, не превосходящих n, называется перестановкой из n чисел.*

Так, совокупность чисел (2, 4, 1, 5, 3) является перестановкой из пяти чисел. Совокупность же (1, 3, 4, 1, 5), равно как совокупность (1, 4, 8, 2, 3), перестановкой не является.

Теорема 1. *Количество различных перестановок из n чисел равно $n!$.*

Доказательство. Рассмотрим перестановку $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$. В качестве σ_1 может быть взято любое натуральное число от 1 до n. Поэтому для выбора σ_1 представляется n возможностей. Если число σ_1 уже выбрано, то для выбора числа σ_2 остается $(n-1)$ возможность — числом σ_2 может быть любое натуральное число от 1 до n, кроме числа σ_1. Таким образом, для выбора чисел σ_1 и σ_2 представляется $n(n-1)$ возможностей. Продолжая рассуждать аналогично, мы придем в конце концов к выводу, что всего возможностей $n(n-1)(n-2)\cdot\ldots\cdot 1 = n!$, что и требовалось доказать.

Определение 2. *Говорят, что пара чисел σ_i, σ_j в перестановке $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ образует беспорядок, если $\sigma_i > \sigma_j$, а $i < j$* (т. е. большее число стоит раньше).

Например, перестановка $(1, 2, 3, 4, 5)$ не содержит беспорядков, а в перестановке $(1, 3, 2, 5, 4)$ их два: во-первых, число 3 стоит раньше, чем 2, во-вторых, число 5 стоит раньше числа 4.

Теорема 2. *Если два числа в перестановке поменять местами, то количество беспорядков в ней изменится на некоторое нечетное число.*

Доказательство. Воспользуемся методом математической индукции. Поменяем в перестановке σ числа σ_iи σ_{i+k} местами. Рассмотрим сначала случай $j = i + 1$. Если прежде числа σ_i и σ_{i+1} не образовывали беспорядка, то теперь они будут его образовывать; если же они образовывали беспорядок, то теперь они перестанут его образовывать. При этом все прочие беспорядки, очевидно, сохранятся. Таким образом, общее количество беспорядков изменится ровно на единицу (в ту или в другую сторону), т. е. на нечетное число.

Допустим теперь, что теорема доказана для $k = m - 1$ и докажем, что тогда она справедлива и для $k = m$. Поменяем сначала местами числа σ_{i+k-1} и σ_{i+k}; затем в полученной перестановке $(\ldots, \sigma_i, \ldots$ $\ldots, \sigma_{i+k}, \sigma_{i+k-1}, \ldots)$ поменяем местами σ_i и σ_{i+k}; наконец, в перестановке $(\ldots, \sigma_{i+k}, \ldots, \sigma_i, \sigma_{i+k-1}, \ldots)$ поменяем местами σ_i, и σ_{i+k-1}: $(\ldots, \sigma_{i+k}, \ldots, \sigma_{i+k-1}, \sigma_i, \ldots)$. В результате i-й и $(i + k)$-й элементы поменялись местами, а порядок следования остальных элементов не изменился. При этом количество беспорядков изменялось три раза, причем каждый раз на нечетное число. Следовательно, в результате количество беспорядков изменилось на нечетное число. Теорема доказана.

Вернемся теперь к определителям и поставим такой вопрос. Пусть $\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ — определитель, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ — перестановка. Переставим в определителе строки так, чтобы первой строкой оказалась σ_1-я, второй — σ_2-я и т. д. Поскольку при перестановке строк модуль определителя не меняется, то полученный определитель $\Delta(\mathbf{a}_{\sigma_1}, \mathbf{a}_{\sigma_2}, \ldots, \mathbf{a}_{\sigma_n})$ может отличаться от исходного только знаком. Как определить этот знак? Ответ дает следующая теорема.

Теорема 3.

$$\Delta\left(\mathbf{a}_{\sigma_1}, \mathbf{a}_{\sigma_2}, \ldots, \mathbf{a}_{\sigma_n}\right) = (-1)^{N(\sigma)}\, \Delta\left(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\right) \tag{1}$$

где $N(\sigma)$ — количество беспорядков в перестановке σ.

Доказательство. Выберем из чисел $\sigma_1, \sigma_2, \ldots, \sigma_n$ то σ_k, которое равно 1, и поменяем его местами с σ_1. Одновременно поменяем местами строку $\mathbf{a}_{\sigma_k}$ определителя $\Delta\left(\mathbf{a}_{\sigma_1}, \mathbf{a}_{\sigma_2}, \ldots, \mathbf{a}_{\sigma_n}\right)$ со строкой $\mathbf{a}_{\sigma_1}$. Затем выберем из оставшихся чисел то σ_m, которое равно 2, и поменяем его местами с σ_2, а строку $\mathbf{a}_{\sigma_m}$ определителя — со строкой $\mathbf{a}_{\sigma_2}$, и т. д. Продолжая этот процесс, мы, очевидно, и приведем данную перестановку к виду $(1, 2, \ldots, n)$, а определитель $\Delta\left(\mathbf{a}_{\sigma_1}, \mathbf{a}_{\sigma_2}, \ldots, \mathbf{a}_{s_n}\right)$ — к виду $\Delta\left(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\right)$.

Пусть K — число шагов, необходимых для осуществления этого процесса. Поскольку при каждой перестановке строк знак определителя менялся на противоположный, то $\Delta\left(\mathbf{a}_{s_1}, \mathbf{a}_{\sigma_2}, \ldots, \mathbf{a}_{\sigma_n}\right) =$ $= (-1)^K\, \Delta\left(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\right)$.

С другой стороны, на каждом шаге количество беспорядков в перестановке изменялось на нечетное число. Перестановка $(1, 2, \ldots, n)$ не содержит беспорядков, и, следовательно, общее число беспорядков $N(\sigma)$ в перестановке σ равно алгебраической сумме K нечетных

чисел, которое четно или нечетно в зависимости от того, четно или нечетно число K. Поэтому $(-1)^K = (-1)^{N(\sigma)}$. Теорема доказана.

2. Выражение определителя через его элементы.

Теорема. *Определитель n-го порядка матрицы A выражается формулой*

$$\det A = \sum_{\sigma} (-1)^{N(\sigma)} a_{1\sigma_1} a_{2\sigma_2} \dots a_{n\sigma_n}. \tag{2}$$

Здесь сумма берется по всем перестановкам из n чисел (т. е. каждой перестановке соответствует одно слагаемое), *а $N(\sigma)$ — количество беспорядков в перестановке σ.*

Доказательство. Представим первую строку матрицы A в виде линейной комбинации координатных строк и воспользуемся свойствами 1°, 2⁰ определителя. Имеем

$$\det A = \Delta(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \Delta\Big(\sum_{s_1} a_{1s_1}\mathbf{e}_{s_1}, \mathbf{a}_2, \dots, \mathbf{a}_n\Big) =$$

$$= \sum_{s_1} \Delta\left(a_{1s_1}\mathbf{e}_{s_1}, a_2, \dots, a_n\right) = \sum_{s_1} a_{1s_1}\Delta(\mathbf{e}_{s_1}, \mathbf{a}_2, \dots, \mathbf{a}_n).$$

Аналогичным образом поступим со второй строкой матрицы A, затем — с третьей и т. д. В результате получим:

$$\det A = \sum_{s_1, s_2, \dots, s_n} a_{1s_1} a_{2s_2} \dots a_{ns_n} \Delta\left(\mathbf{e}_{s_1}, \mathbf{e}_{s_2}, \dots, \mathbf{e}_{s_n}\right).$$

Согласно свойству 3° в этой сумме отличны от нуля только те слагаемые, в которых все строки $\mathbf{e}_{s_i}$ попарно различны, т. е. числа s_i образуют перестановку из n чисел. Поэтому полученное выражение можно переписать так:

$$\det A = \sum_{\sigma} a_{1\sigma_1} a_{2\sigma_2} \dots a_{n\sigma_n} \Delta\left(\mathbf{e}_{\sigma_1}\mathbf{e}_{\sigma_2} \dots \mathbf{e}_{\sigma_n}\right).$$

Осталось заметить, что по формуле (1)

$$\Delta\left(\mathbf{e}_{\sigma_1}, \mathbf{e}_{\sigma_2}, \dots, \mathbf{e}_{\sigma_n}\right) = (-1)^{N(\sigma)} \Delta\left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\right),$$

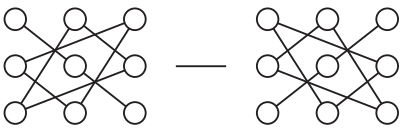
а $\Delta(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$. Теорема доказана.

Замечание. Рассмотрим несколько частных случаев.

При $n = 1$ существует только одна перестановка — (1). Поэтому $\det A = a_{11}$.

При $n = 2$ перестановок две: $(1, 2)$ и $(2, 1)$. В первой из них беспорядков нет, а во второй — один беспорядок. Поэтому выражение для $\det A$ состоит из двух слагаемых: $\det A = a_{11}a_{22} - a_{12}a_{21}$. Символически это можно изобразить так:

При $n = 3$ перестановок шесть. Три из них — $(1, 2, 3)$, $(3, 1, 2)$ и $(2, 3, 1)$ содержат четное количество беспорядков, а три — $(1, 3, 2)$, $(2, 1, 3)$ и $(3, 2, 1)$ — нечетное. Поэтому определитель третьего порядка выглядит так: $\det A = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$. Символически это можно изобразить так:



В общем случае:

а) *определитель представляет собой алгебраическую сумму* $n!$ *слагаемых* (их столько, сколько различных перестановок из n чисел);

б) *каждое слагаемое представляет собой произведение* n *элементов матрицы, взятых из* ***попарно различных*** *строк* (с номерами $1, 2, \ldots, n$) и ***попарно различных*** *столбцов* (с номерами $\sigma_1, \sigma_2, \ldots, \sigma_n$);

в) *знак перед каждым слагаемым определяется четностью количества беспорядков в перестановке номеров столбцов в этом слагаемом.*

3. Алгебраическое дополнение. Рассмотрим формулу (2), выражающую определитель матрицы A через ее элементы. Сгруппируем в ней все те слагаемые, которые содержат в качестве сомножителя элемент a_{ij}, и вынесем общий множитель a_{ij} за скобки. Та сумма, которая останется после этого в скобках, называется *алгебраическим дополнением* A_{ij} *элемента* a_{ij}. Иными словами, A_{ij} — это то, во что превращается правая часть выражения (2) при замене элемента a_{ij} на единицу, а всех остальных элементов i-й строки — на нули.

Теорема. *Алгебраическое дополнение элемента* a_{ij} *равно минору, дополнительному к* a_{ij}, *взятому со знаком «+», если число* $(i+j)$ *четно, и «−» — если нечетно*: $A_{ij} = (-1)^{i+j}\Delta_{ij}$.

Доказательство. Из определения следует, что алгебраическое дополнение A_{ij} представляет собой определитель, полученный из $\det A$ заменой элемента a_{ij} на единицу, а всех остальных элементов i-й строки — на нули. С другой стороны, если такой определитель разложить по i-й строке, то окажется, что он равен $(-1)^{i+j}\Delta_{ij}$. Теорема доказана.

Замечание 1. Доказанная теорема позволяет по-новому записывать формулу разложения определителя по i-й строке:

$$\det A = \sum_{s=1}^{n} a_{is} A_{is}.$$

Замечание 2. Рассмотрим определитель, в котором j-я строка заменена на i-ю ($i \neq j$):

$$\begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} \begin{vmatrix} a_{11} & a_{22} & \dots & \dots & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & \dots & a_{nn} \end{vmatrix}$$

Этот определитель, очевидно, равен нулю — в нем две одинаковые строки. Раскладывая его по j-й строке и учитывая, что алгебраические дополнения к ее элементам не зависят от элементов j-й строки (она вычеркивается), получаем:

$$\sum_{s=1}^{n} a_{is} A_{js} = 0.$$

Объединяя этот результат с утверждением замечания 1, можно написать так:

$$\sum_{s=1}^{n} a_{is} A_{js} = \begin{cases} \det A, & i = j; \\ 0, & i \neq j. \end{cases} \tag{3}$$

4. Разложение определителя по столбцу.

Теорема 1. *Справедлива следующая формула, называемая формулой разложения определителя по j-му столбцу*:

$$\det A = \sum_{s=1}^{n} a_{sj} A_{sj}.$$

Доказательство. Каждое слагаемое в выражении (2) представляет собой произведение n элементов, взятых из различных столбцов, поэтому в каждом слагаемом элемент j-го столбца фигурирует ровно один раз. Следовательно, каждое слагаемое из этого выражения войдет в нашу сумму, и притом только один раз. Теорема доказана.

Определение. *Матрицей, транспонированной по отношению к матрице*

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

называется матрица

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

Таким образом, строками матрицы A^{tr} являются столбцы матрицы A (и наоборот). Например, матрица

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$$

является транспонированной по отношению к матрице

$$\begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}.$$

Теорема 2. *Для любой $(n \times n)$-матрицы A имеет место равенство* $\det A^{tr} = \det A$.

Доказательство. Воспользуемся методом математической индукции.

При $n = 1$ утверждение теоремы очевидно — в этом случае транспонированная матрица совпадает с исходной.

Допустим, что теорема доказана для $n = k$. Тогда разложение определителя $(k+1)$-го порядка матрицы A^{tr} по первой строке совпадает с разложением определителя матрицы A по первому столбцу. Теорема доказана.

Следствие. *Все утверждения о строках определителя справедливы и для его столбцов*. Иными словами, *строки и столбцы в определителе равноправны*.

§ 4. Произведение матриц

1. Свойства произведения матриц.

Определение. *Произведением $(m \times n)$-матрицы A на $(n \times k)$-матрицу B называется $(m \times k)$-матрица C, элементы c_{ij} которой равны* $\sum\limits_{s=1}^{n} a_{is} b_{sj}$.

Таким образом, элемент с индексами i и j матрицы C представляет собой «произведение» i-й строки матрицы A на j-й столбец матрицы B. Обратим внимание на то, что при этом количество столбцов матрицы A должно совпадать с количеством строк матрицы B — иначе произведение матриц A и B не определено.

Из сказанного ясно, что если произведение матриц A и B определено, то произведение матриц B и A, вообще говоря, не определено. Но даже в том случае, когда оба произведения определены (в частности, когда матрицы A и B — квадратные), AB, вообще говоря,

не равно BA. В самом деле, пусть, например, $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$. Тогда $AB = \begin{pmatrix} (1\cdot 0 + 0\cdot 0) & (1\cdot 1 + 0\cdot 1) \\ (1\cdot 0 + 0\cdot 0) & (1\cdot 1 + 0\cdot 1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$, а $BA = \begin{pmatrix} (0\cdot 1 + 1\cdot 1) & (0\cdot 0 + 1\cdot 0) \\ (0\cdot 1 + 1\cdot 1) & (0\cdot 0 + 1\cdot 0) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, т. е. $AB \neq BA$.

Теорема. *Умножение матриц* (*при условии, что оно определено*) *обладает следующими свойствами*:

1° $A(BC) = (AB)C$;

2° $A(B + C) = AB + AC$;

3° $(A + B)C = AC + BC$;

4° $\lambda(AB) = (\lambda A)B = A(\lambda B)$, *где* λ — *любое число.*

Доказательство.

1°. $(A\,(BC))_{ij} = \sum\limits_s a_{is}(BC)_{sj} = \sum\limits_{s,p} a_{is}b_{sp}c_{pj} = \sum\limits_p (AB)_{ip}\,c_{pj} = ((AB)\,C)_{ij}$.

2°. $(A\,(B + C))_{ij} = \sum\limits_s a_{is}\,(b_{sj} + c_{sj}) = \sum\limits_s a_{is}b_{sj} + \sum\limits_s a_{is}c_{sj} = (AB + AC)_{ij}$.

3°. $((A + B)\,C)_{ij} = \sum\limits_s (a_{is} + b_{is})\,c_{sj} = \sum\limits_s a_{is}c_{sj} + \sum\limits_s b_{is}c_{sj} = (AC + BC)_{ij}$.

4°. $(\lambda\,(AB))_{ij} = \lambda \sum\limits_s a_{is}b_{sj} = \sum\limits_s (\lambda a_{is})\,b_{sj} = ((\lambda A)\,B)_{ij} = \sum\limits_s a_{is}\,(\lambda b_{sj}) = (A\,(\lambda B))_{ij}$.

Теорема доказана.

2. Определитель произведения квадратных матриц. Обратимся теперь к квадратным матрицам.

Теорема. $\det(AB) = \det A \det B$.

Доказательство. Пусть $C = AB$. Имеем:

$$\det(AB) = \det C = \Delta(\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n) = \Delta\left(\sum_{s_1} c_{1s_1}\mathbf{e}_{s_1}, \mathbf{c}_2, \ldots, \mathbf{c}_n\right) =$$

$$= \Delta\left(\sum_{s_1,p_1} a_{1p_1}b_{p_1s_1}\mathbf{e}_{s_1}, \mathbf{c}_2, \ldots, \mathbf{c}_n\right) = \Delta\left(\sum_{p_1} a_{1p_1}\mathbf{b}_{p_1}, \mathbf{c}_2, \ldots, \mathbf{c}_n\right) =$$

$$= \sum_{p_1} a_{1p_1}\Delta\left(\mathbf{b}_{p_1}, \mathbf{c}_2, \ldots, \mathbf{c}_n\right) = \ldots$$

$$\ldots = \sum_{p_1,p_2,\ldots,p_n} a_{1p_1}a_{2p_2}\ldots a_{np_n}\Delta\left(\mathbf{b}_{p_1}, \mathbf{b}_{p_2}, \ldots, \mathbf{b}_{p_n}\right).$$

В этой сумме отличными от нуля будут лишь те слагаемые, в которых все $\mathbf{b}_{p_i}$ попарно различны, т. е. числа p_i образуют перестановку из n чисел. Поэтому, с учетом формулы (1) п. 1 § 3, полученное выражение можно переписать так:

$$\det(AB) = \sum_{\sigma} a_{1\sigma_1} a_{2\sigma_2} \ldots a_{n\sigma_n} \Delta\left(\mathbf{b}_{\sigma_1}, \mathbf{b}_{\sigma_2}, \ldots, \mathbf{b}_{\sigma_n}\right) =$$

$$= \sum_{\sigma} (-1)^{N(\sigma)} a_{1\sigma_1} a_{2\sigma_2} \ldots a_{n\sigma_n} \Delta\left(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\right) =$$

$$= \det B \sum_{\sigma} (-1)^{N(\sigma)} a_{1\sigma_1} a_{2\sigma_2} \ldots a_{n\sigma_n} = \det B \det A.$$

Теорема доказана.

Замечание. Подчеркнем особо, что равенства

$$\det(A+B) = \det A + \det B, \quad \det(\lambda A) = \lambda \det A$$

могут выполняться лишь в отдельных (крайне редких) случаях.

3. Свойства произведения квадратных матриц. Произведение квадратных матриц обладает еще двумя важными свойствами.

Теорема. *При любом n:*

1° *существует такая $(n \times n)$-матрица E, что для любой $(n \times \times n)$-матрицы A имеют место равенства $AE = EA = A$;*

2° *для всякой $(n \times n)$-матрицы A с определителем, отличным от нуля, существует такая матрица A^{-1}, что $AA^{-1} = A^{-1}A = E$.*

Доказательство. 1°. Рассмотрим $(n \times n)$-матрицу

$$E = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix},$$

т. е. матрицу, элемент e_{ij} которой равен 1 при $i = j$ и 0 при $i \neq j$.

Пусть A — произвольная $(n \times n)$-матрица. Имеем: $(AE)_{ij} = \sum\limits_{s} a_{is} e_{sj}$. В этой сумме все слагаемые, кроме одного (при $s = j$), равны нулю; единственное отличное от нуля слагаемое равно $a_{ij} e_{jj} = a_{ij} 1 = = a_{ij}$. Но это и означает, что $AE = A$. Аналогично доказывается, что $EA = A$.

2°. Пусть A — произвольная $(n \times n)$-матрица с определителем, отличным от нуля. Рассмотрим матрицу B, элементами которой яв-

ляются числа $b_{ij} = \dfrac{A_{ji}}{\det A}$, где A_{ji} — алгебраическое дополнение элемента a_{ji} [1]. Имеем:

$$(AB)_{ij} = \sum_s a_{is} b_{sj} = \frac{1}{\det A} \sum_s a_{is} A_{js} =$$

$$= \frac{1}{\det A} \begin{cases} \det A, & i = j \\ 0, & i \neq j \end{cases} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} = (E)_{ij}.$$

Аналогично доказывается, что $(BA)_{ij} = (E)_{ij}$. Тем самым, $B = A^{-1}$. Теорема доказана.

Замечание 1. Нетрудно видеть, что матрица E определяется единственным образом: если существует такая матрица E_1, что для любой матрицы A выполняется равенство $AE_1 = A$, то $E_1 = EE_1 = E$. Матрица A^{-1} также определяется единственным образом: если существует такая матрица A_1^{-1}, что $AA_1^{-1} = E$, то $A^{-1}AA_1^{-1} = EA_1^{-1} = A_1^{-1} = A^{-1}E = A^{-1}$. Матрица E называется *единичной*, а матрица A^{-1} — *обратной по отношению к матрице* A.

Замечание 2. В доказанной теореме содержится условие $\det A \neq 0$. Возникает вопрос: а что будет, если это условие нарушено, т. е. $\det A = 0$ — существует ли в этом случае матрица A^{-1}? Ответ оказывается отрицательным. В самом деле, из равенства $AA^{-1} = E$ следует, что $\det A \det A^{-1} = \det E = 1$. Поэтому предположение о том, что $\det A = 0$, приводит к противоречию.

§ 5. Базисный минор

1. Теорема о базисном миноре. Рассмотрим произвольную $(m \times n)$-матрицу A.

Определение 1. *Минором матрицы A, построенным на строках с номерами $i_1, i_2, \ldots, i_k$ и столбцах с номерами $j_1, j_2, \ldots, j_k$, называется определитель, элементами которого являются элементы матрицы A, расположенные на пересечении указанных строк и столбцов.*

Определение 2. *Минор M произвольной матрицы называется базисным, если он отличен от нуля, а любой минор на единицу большего порядка, включающий в себя M, равен нулю (или такого минора не существует); строки и столбцы, на которых построен базисный минор, называются базисными.*

Теорема (о базисном миноре). *Базисные строки (столбцы) линейно независимы; любая строка (столбец) представляет собой линейную комбинацию базисных.*

[1] Обратим внимание на то, что элемент b_{ij} выражается через A_{ji}, а не через A_{ij}!

Доказательство. Доказательство теоремы проведем для строк — для столбцов она доказывается аналогично.

Допустим, что базисные строки линейно зависимы. Тогда одну из них можно представить в виде линейной комбинации остальных и, следовательно, базисный минор равен нулю. Полученное противоречие доказывает первую часть утверждения теоремы.

Осталось доказать, что любая строка равна линейной комбинации базисных строк. Без ограничения общности будем считать, что базисный минор M расположен в левом верхнем углу матрицы (в противном случае строки и столбцы можно соответствующим образом переставить), k — его порядок. Добавим к нему i-ю строку и j-й столбец:

$$\begin{vmatrix} a_{11} & \dots & a_{1k} & a_{1j} \\ \dots & \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} & a_{kj} \\ a_{i1} & \dots & a_{ik} & a_{ij} \end{vmatrix}.$$

Полученный определитель равен нулю. В самом деле, если $i \leqslant k$ или $j \leqslant k$, то в нем две одинаковые строки или два одинаковых столбца; если же $i > k$ и $j > k$, то он является минором $(k+1)$-го порядка, включающим в себя M, а значит, равен нулю по условию теоремы.

Разложим этот определитель по последнему столбцу:

$$a_{1j}A_{1j} + \ldots + a_{kj}A_{kj} + a_{ij}M = 0.$$

Заметим теперь, что числа $A_{1j}, \ldots, A_{kj}$ и M не зависят от j (число M не зависит также и от i). Учитывая, что $M \neq 0$, положим $\lambda_1 = -A_{1j}/M, \ldots, \lambda_k = -A_{kj}/M$. Тогда последнее равенство можно будет переписать так:

$$a_{ij} = \lambda_1 a_{1j} + \lambda_2 a_{2j} + \ldots + \lambda_k a_{kj}.$$

Но это и означает, что i-я строка, выбранная произвольно, равна линейной комбинации базисных строк. Теорема доказана.

Следствие. *Если минор M — базисный, то любой минор на единицу большего порядка (если такой есть) равен нулю.*

Действительно, рассмотрим произвольный минор $N(k+1)$-го порядка. По теореме о базисном миноре любая его строка $\mathbf{n}_i$ равна линейной комбинации строк, на которых построен минор M: $\mathbf{n}_i = \sum\limits_{s=1}^{k} \lambda_{is}\mathbf{a}_s$. Поэтому

$$N = \Delta(\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_{k+1}) =$$

$$= \Delta\Big(\sum_{s_1} \lambda_{1s_1}\mathbf{a}_{s_1}, \sum_{s_2} \lambda_{2s_2}\mathbf{a}_{s_2}, \ldots, \sum_{s_{k+1}} \lambda_{k+1s_{k+1}}\mathbf{a}_{s_1}\Big) =$$

$$= \sum_{s_1 \ldots s_{k+1}} \lambda_{1s_1} \ldots \lambda_{k+1s_{k+1}} \Delta\left(\mathbf{a}_{s_1}, \mathbf{a}_{s_2} \ldots \mathbf{a}_{s_{k+1}}\right).$$

Но каждый из определителей $\Delta(\mathbf{a}_{s_1}, \mathbf{a}_{s_2}, \ldots, \mathbf{a}_{s_{k+1}})$ равен нулю, поскольку из $(k+1)$ его строк различных не более k. Следовательно, минор N равен нулю.

2. Ранг матрицы.

Определение. *Рангом матрицы называется наибольшее число ее линейно независимых строк.*

Ранг матрицы A обозначается так: $\operatorname{rang} A$.

Теорема. *Ранг матрицы равен порядку ее базисного минора* [1].

Доказательство. Пусть k — порядок базисного минора матрицы A. Базисные строки линейно независимы, поэтому $\operatorname{rang} A \geqslant k$.

Допустим, что $\operatorname{rang} A > k$ и, следовательно, среди строк матрицы A найдется $(k+1)$ линейно независимых. Составим из них матрицу. Порядок базисного минора этой матрицы не превосходит k, поскольку любой ее минор является одновременно и минором матрицы A. Поэтому не все ее строки базисные, а значит, одна из них равна линейной комбинации остальных. Полученное противоречие завершает доказательство теоремы.

Замечание. Аналогично доказывается, что наибольшее число линейно независимых столбцов матрицы равно порядку ее базисного минора. Тем самым, *наибольшее число линейно независимых строк равно наибольшему числу линейно независимых столбцов и равно порядку базисного минора*. Это число и называется рангом матрицы.

Следствие. *Определитель равен нулю тогда и только тогда, когда его строки линейно зависимы.*

В самом деле, если строки определителя линейно зависимы, то одна из них равна линейной комбинации остальных, а значит, определитель равен нулю. Обратно, если определитель $(n \times n)$-матрицы A равен нулю, то $\operatorname{rang} A < n$, а значит, ее строки линейно зависимы.

[1] Если матрица состоит из одних нулей, то, согласно определению, базисного минора у нее нет. Условимся считать, что порядок базисного минора в этом случае равен нулю.

Глава 2
СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 1. Существование и единственность решения

1. Основные определения. Рассмотрим систему m линейных уравнений с n неизвестными $x_1, x_2, \ldots, x_n$:

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}\right\}. \qquad (1)$$

Здесь предполагается, что a_{ij} (*коэффициенты системы*) и b_i — известные числа (вещественные или комплексные).

Решением системы (1) называется любой набор чисел $(x_1, x_2, \ldots$ $\ldots, x_n)$, обращающий все уравнения (1) в тождества. Два решения $(x_1, x_2, \ldots, x_n)$ и $(y_1, y_2, \ldots, y_n)$ считаются *различными*, если $x_i \neq y_i$ хотя бы при одном значении i.

Система (1) называется *однородной*, если все $b_i = 0$; в противном случае (если хотя бы одно число b_i отлично от нуля) эта система называется *неоднородной*. При этом система

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{array}\right\}. \qquad (2)$$

называется *однородной системой*, *соответствующей системе* (1).

Введем еще два термина. Матрицу

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

условимся называть *основной матрицей*, а матрицу

$$B = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

— *расширенной матрицей* системы (1).

Заметим, что формулы (1) и (2) можно записать значительно компактнее, если воспользоваться определением произведения матриц:

$$A\mathbf{x} = \mathbf{b}, \tag{1}$$

$$A\mathbf{x} = \mathbf{o}, \tag{2}$$

где

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}, \quad \mathbf{o} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix},$$

A — основная матрица системы (1). Такую запись систем (1) и (2) мы будем для краткости называть *матричной*.

2. Существование решения.

Теорема (Кронекера–Капелли). *Решение системы* (1) *существует тогда и только тогда, когда* $\operatorname{rang} A = \operatorname{rang} B$ (*ранг основной матрицы равен рангу расширенной матрицы*).

Доказательство. Решение системы (1) существует тогда и только тогда, когда последний столбец матрицы B равен линейной комбинации (с коэффициентами $x_1, x_2, \dots, x_n$) столбцов матрицы A, т. е. тогда и только тогда, когда в матрице B линейно независимых столбцов столько же, сколько в A. Но это и означает, что $\operatorname{rang} A = \operatorname{rang} B$. Теорема доказана.

3. Однородные системы. Обратимся к однородной системе (2). Из теоремы Кронекера–Капелли следует, что она всегда имеет решение. Впрочем, это ясно и так: у нее всегда есть решение $x_1 = x_2 = \dots = x_n = 0$ — оно называется *тривиальным*. Любое другое решение системы (2) (для которого хотя бы одно x_i отлично от нуля) называется *нетривиальным*.

Теорема. *Система* (2) *имеет нетривиальные решения тогда и только тогда, когда* $\operatorname{rang} A$ *меньше* n (*количества неизвестных*).

Доказательство. Нетривиальное решение системы (2) существует тогда и только тогда, когда столбцы матрицы A линейно зависимы, т. е. тогда и только тогда, когда $\operatorname{rang} A$ (количество линейно независимых столбцов) меньше n (количества столбцов или, что то же самое, количества неизвестных). Теорема доказана.

Следствие. *Система* (2) *имеет только тривиальное решение тогда и только тогда, когда* $\operatorname{rang} A = n$ (так как $\operatorname{rang} A$ не может быть больше количества столбцов n).

4. Единственность решения.

Теорема 1. *Пусть* $(x_1, x_2, \dots, x_n)$ — *решение системы* (1), $(y_1, y_2, \dots, y_n)$ — *решение системы* (2). *Тогда* $(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$ — *решение системы* (1).

Доказательство. Воспользуемся матричной записью систем (1), (2). Имеем: $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{b} + \mathbf{o} = \mathbf{b}$. Теорема доказана.

Следствие. *Если система* (1) *имеет одно и только одно решение, то система* (2) *имеет только тривиальное решение*, т. е. $\operatorname{rang} A = n$.

Теорема 2. *Пусть* $(x_1, x_2, \ldots, x_n)$ *и* $(y_1, y_2, \ldots, y_n)$ — *два решения системы* (1). *Тогда* $(x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)$ — *решение системы* (2).

Доказательство. Воспользуемся матричной записью систем (1), (2). Имеем: $A(\mathbf{x} - \mathbf{y}) = A\mathbf{x} - A\mathbf{y} = \mathbf{b} - \mathbf{b} = \mathbf{o}$. Теорема доказана.

Следствие 1. *Пусть* $(x_1, x_2, \ldots, x_n)$ — *какое-нибудь решение системы* (1). *Тогда любое другое решение этой системы представляет собой* $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$, *где* $(y_1, y_2, \ldots, y_n)$ — *решение системы* (2).

Следствие 2. *Если система* (2) *имеет только тривиальное решение*, т. е. $\operatorname{rang} A = n$, *то система* (1) *имеет не более одного решения*.

Теорема 3 (существования и единственности решения). *Система* (1) *имеет одно и только одно решение тогда и только тогда, когда* $\operatorname{rang} A = \operatorname{rang} B = n$.

Доказательство. Допустим, что $\operatorname{rang} A = \operatorname{rang} B = n$. Тогда из равенства $\operatorname{rang} A = \operatorname{rang} B$ следует, что решение системы (1) существует; а из равенства $\operatorname{rang} A = n$ следует, что система (1) имеет не более одного решения.

Допустим теперь, что решение системы (1) существует и оно — единственное. Из того что решение существует, следует, что $\operatorname{rang} A = \operatorname{rang} B$. Из единственности этого решения следует, что $\operatorname{rang} A = n$. Теорема доказана.

§ 2. Нахождение решений

1. Формулы Крамера. Рассмотрим систему линейных уравнений

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{array}\right\}, \qquad (1)$$

количество уравнений в которой совпадает с количеством неизвестных n. Специфика этого случая состоит в том, что здесь, очевидно, $\operatorname{rang} A \leqslant \operatorname{rang} B \leqslant n$ (в общем случае $\operatorname{rang} A \leqslant \operatorname{rang} B \leqslant \min(n+1, m)$). Поэтому *система* (3) *имеет одно и только одно решение тогда и только тогда, когда* $\operatorname{rang} A = n$, т. е. $\det A \neq 0$.

Чтобы найти это решение, вспомним, что из условия $\det A \neq 0$ следует существование матрицы A^{-1}, обратной к A. Запишем систему (3) в матричном виде и, предполагая, что $\mathbf{x}$ — решение, умножим обе части полученного тождества слева на эту матрицу: $A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$. Но $A^{-1}A = E$, а $E\mathbf{x} = \mathbf{x}$. Поэтому

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Это и есть ответ.

Замечание. Выведенная нами формула особенно удобна в тех случаях, когда нужно решить систему (1) при неизменной левой части (т. е. матрице A) и нескольких различных значениях правой части,

поскольку наиболее трудоемкую работу — нахождение матрицы A^{-1} — достаточно выполнить один раз.

Встречаются, однако, и другие задачи. Например, иногда требуется найти не все решения x_i, а только часть из них или даже какое-нибудь одно. В этом случае искать A^{-1} не обязательно. В самом деле, имеем:

$$x_k = \sum_{s=1}^{n} \left(A^{-1}\right)_{ks} b_s = \sum_{s=1}^{n} \frac{A_{sk}}{\det A} b_s = \frac{1}{\det A} \sum_{s=1}^{n} b_s A_{sk} =$$

$$= \frac{\begin{vmatrix} a_{11} & \dots & a_{1k-1} & b_1 & a_{1k+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nk-1} & b_n & a_{nk+1} & \dots & a_{nn} \end{vmatrix}}{\det A}.$$

Эти формулы (при разных значениях k) называются *формулами Крамера*.

2. Общий случай. Рассмотрим теперь общий случай: систему m линейных уравнений с n неизвестными. Предположим, что решение этой системы существует, и, следовательно, ранг основной матрицы равен рангу расширенной матрицы: $\operatorname{rang} A = \operatorname{rang} B$. Поставим своей целью его найти.

Без ограничения общности будем считать, что базисный минор порядка r, общий для матриц A и B, расположен в левом верхнем углу матрицы A:

$$\left.\begin{array}{r} a_{11}x_1 + \ldots + a_{1r}x_r + \ldots + a_{1n}x_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + \ldots + a_{r1}x_r + \ldots + a_{rn}x_n = b_r \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + \ldots + a_{mr}x_r + \ldots + a_{mn}x_n = b_m \end{array}\right\}.$$

Поскольку строки матрицы B с номерами, большими r, представляют собой линейные комбинации первых r ее строк, то последние $(m-r)$ уравнений являются следствиями первых r и, следовательно, могут быть отброшены (они не дают ничего нового).

Оставшиеся r уравнений перепишем так:

$$\left.\begin{array}{r} a_{11}x_1 + \ldots + a_{1r}x_r = b_1 - a_{1r+1}x_{r+1} - \ldots - a_{1n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + \ldots + a_{rr}x_r = b_r - a_{rr+1}x_{r+1} - \ldots - a_{rn}x_n \end{array}\right\}.$$

Эту систему можно рассматривать как систему r уравнений с r неизвестными $x_1, x_2, \ldots, x_r$. Ее определитель, будучи базисным минором, отличен от нуля, поэтому, согласно результатам предыдущего пункта, при любой правой части, в частности при любых $x_{r+1}, x_{r+2}, \ldots$ $\ldots, x_n$, она имеет единственное решение. Это означает, что числа $x_{r+1}, x_{r+2}, \ldots, x_n$ можно выбрать произвольно, полагая $x_{r+1} = C_1$, $x_{r+2} = C_2$, $\ldots$, $x_n = C_{n-r}$, а $x_1, x_2, \ldots, x_r$ найти, например, по формулам Крамера. Таким образом, общее решение нашей системы зависит от $(n-r)$ произвольных чисел $C_1, C_2, \ldots, C_{n-r}$.

Часть II

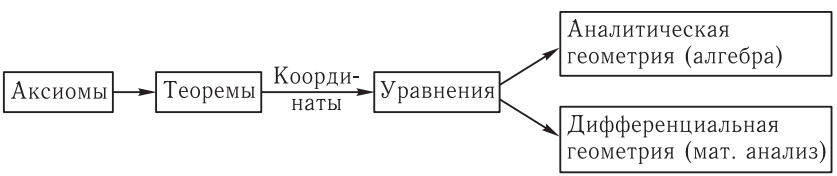
АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ

Предварительные замечания

Аналитическая геометрия — это раздел геометрии, в котором свойства геометрических объектов изучаются методами алгебры.

Поясним эти слова. Геометрия, как и другие разделы математики, строится так: сначала формулируются исходные положения — аксиомы, а затем из них выводятся логические следствия — теоремы. Таким образом, на этом (первом) этапе построение геометрии ведется исключительно на базе собственных средств — аксиом и ранее доказанных теорем. Эта часть геометрии называется *элементарной геометрией*.

Следующий этап в построении геометрии состоит в расширении аппарата путем привлечения средств других разделов математики, в первую очередь, алгебры и математического анализа. Делается это так: вводится система координат, в результате чего каждая точка описывается набором чисел, а геометрические фигуры — уравнениями и неравенствами. Благодаря этому изучение геометрических объектов может быть в ряде случаев сведено к изучению уравнений. Изучение же свойств уравнений осуществляется методами алгебры и математического анализа. Так появляются новые разделы геометрии — аналитическая и *дифференциальная геометрия*:



В дальнейшем мы будем предполагать известными аксиомы и основные теоремы элементарной геометрии; изложение же метода координат (в частности, и материал, входящий в школьную программу) будем снабжать полными доказательствами.

Примем еще несколько соглашений. Прежде всего, раз и на всегда *договоримся считать заданной единицу измерения отрезков*. Тем самым, длина любого отрезка, площадь фигуры и объем тела мы будем представлять себе выражающимися вполне определенными вещественными числами. Далее, под словами «отрезок», «треугольник», «параллелограмм», «параллелепипед» условимся понимать в том числе и *вырожденные объекты*: вырожденный отрезок — это одна точка (она же является и серединой этого отрезка), вырожденный треугольник или параллелограмм — это отрезок (в том числе, вырожденный), вырожденный параллелепипед — это фигура, представляющая собой изображение параллелепипеда на листе бумаги. Наконец, будем считать, что длина вырожденного отрезка равна нулю, площадь вырожденного треугольника или параллелограмма равна нулю, объем вырожденного параллелепипеда равен нулю. Угол между двумя совпадающими лучами мы также будем считать равным нулю.

Глава 1
ВЕКТОРЫ И КООРДИНАТЫ

§ 1. Координаты точки

1. Ось координат. Пусть l — произвольная прямая, O — какая-нибудь ее точка. Согласно аксиомам геометрии, точка O разделяет прямую l на два луча. Выберем один из них и назовем его *положительной полуосью*.

Определение 1. *Прямая, на которой выбрана положительная полуось, называется осью координат; начало положительной полуоси называется началом координат.*

Определение 2. *Координатой точки M, лежащей на оси координат с началом O, называется длина* [1] *отрезка OM, взятая со знаком «+», если точка M лежит на положительной полуоси,* и «−» — *в противном случае.*

Из аксиом геометрии следует, что каждая точка M рассматриваемой оси имеет вполне определенную координату x; обратно, для любого числа x на данной оси существует ровно одна точка M с координатой x. Иными словами, соответствие $M \leftrightarrow x$ является взаимно однозначным. Тот факт, что точка M имеет координату x, условимся обозначать так: $M(x)$.

Теорема 1. *Пусть $A(x_1)$ и $B(x_2)$ — точки, лежащие на оси координат. Тогда $AB = |x_2 - x_1|$.*

Доказательство. Если начало координат O лежит на отрезке AB, то $AB = AO + OB = |x_1| + |x_2| = ||x_1| + |x_2||$, причем знаки чисел x_1 и x_2 разные. Поэтому $AB = |x_2 - x_1|$. В противном случае $AB = |OA - OB| = ||x_1| - |x_2|| = |x_2 - x_1|$, так как знаки чисел x_1 и x_2 совпадают. Теорема доказана.

Теорема 2. *Пусть $A(x_1)$ и $B(x_2)$ — точки, лежащие на оси координат, $M(x)$ — середина отрезка AB. Тогда $x = (x_1 + x_2)/2$.*

Доказательство. По условию теоремы $AM = MB$, поэтому, согласно теореме 1, $|x - x_1| = |x_2 - x|$. Представляются возможными два случая.

1°. $x - x_1 = x_2 - x$, откуда $x = (x_1 + x_2)/2$.

2°. $x - x_1 = -x_2 + x$, откуда $x_1 = x_2$ и, следовательно, $x = x_1 = x_2 = (x_1 + x_2)/2$. Теорема доказана.

---

[1] Напомним, что единицу измерения отрезков мы договорились считать заданной.

2. Декартовы координаты. Условимся называть *проекцией точки M на прямую l* основание перпендикуляра, проведенного из точки M к прямой l.

Определение 1. *Координата проекции точки M на ось координат называется декартовой координатой точки M по этой оси.*

Таким образом, если O — начало координат, то координата точки M равна $OM\cos\varphi$, где φ — угол между лучом OM и положительной полуосью оси координат.

Определение 2. *Упорядоченная совокупность двух* (*трех*) *взаимно перпендикулярных осей координат с общим началом называется декартовой системой координат на плоскости* (*в пространстве*); *координата точки M по i-й оси координат называется i-й декартовой координатой точки M.*

Обычно начало координат обозначают буквой O, а оси координат — Ox, Oy, (Oz). Иногда их обозначают также Ox_1, Ox_2, (Ox_3). Поскольку в дальнейшем речь будет идти главным образом о декартовых координатах, то слово «декартовы» мы будем, как правило, опускать.

Из аксиом геометрии следует, что каждая точка M плоскости (пространства) имеет вполне определенный набор координат x, y (x, y, z); обратно, для любых чисел x, y (x, y, z) на плоскости (в пространстве) существует ровно одна точка M с координатами x, y (x, y, z). Иными словами, соответствие $M \leftrightarrow x, y$ $(M \leftrightarrow x, y, z)$ является взаимно однозначным. Тот факт, что точка M имеет координаты x, y (x, y, z) условимся обозначать так: $M(x, y)$ $(M(x, y, z))$. Часто бывает удобно обозначать координаты точки и саму точку одной и той же буквой: $M(m_1, m_2)$ $(M(m_1, m_2, m_3))$.

Теорема 1. $AB = \sqrt{\sum\limits_i (b_i - a_i)^2}$.

Доказательство. Утверждение теоремы является очевидным следствием теоремы Пифагора, которая, как нетрудно заметить, верна и для вырожденных треугольников.

Теорема 2. *Пусть M — середина отрезка AB. Тогда $m_i = (a_i + b_i)/2$.*

Доказательство. Пусть A_i, B_i и M_i — проекции точек A, B и M на ось Ox_i. Если точки A_i и B_i совпадают, то точка M_i совпадает с A_i и B_i, и, следовательно, является серединой отрезка A_iB_i. Если же точки A_i и B_i различны, то точка M_i является серединой отрезка A_iB_i по теореме Фалеса [1]. И в том, и в другом случае, согласно теореме 2 п. 1, $m_i = (a_i + b_i)/2$.

3. Криволинейные координаты на плоскости. Наряду с декартовой рассматриваются и другие системы координат. Их объединяют

[1] В стереометрии теорема Фалеса формулируется так: если на одной из двух прямых отложить последовательно несколько равных отрезков и через их концы провести параллельные плоскости, пересекающие вторую прямую, то они отсекут на второй прямой равные между собой отрезки.

общим названием *криволинейные координаты*. Любая такая система координат на плоскости задается двумя уравнениями вида

$$\left.\begin{aligned} x &= x\,(u, v) \\ y &= y\,(u, v) \end{aligned}\right\}. \qquad (1)$$

Таким образом, каждому набору чисел u, v соответствует набор чисел x, y. Если верно и обратное утверждение — каждому набору чисел x, y соответствует единственный набор чисел u, v, то числа u, v могут рассматриваться как криволинейные координаты точки. Чтобы найти криволинейные координаты точки M, нужно сперва определить ее декартовы координаты, а затем из формул (1) найти u и v.

Название «криволинейные координаты» объясняется тем, что координатные линии, т. е. линии $u = \text{const}$, и линии $v = \text{const}$, в таких системах координат, вообще говоря, не являются прямыми (как в декартовой системе координат).

Наиболее употребительными криволинейными координатами на плоскости являются *полярные координаты*. Они определяются формулами

$$\left.\begin{aligned} x &= \rho\cos\varphi \\ y &= \rho\sin\varphi \end{aligned}\right\},$$

где $\rho \geqslant 0$, $0 \leqslant \varphi < 2\pi$. Нетрудно видеть, что величины ρ и φ имеют простой геометрический смысл: ρ — это расстояние точки M от начала координат, т. е. OM, а φ — угол между лучом OM и осью Ox, отсчитываемый в направлении оси Oy.

Ясно, что по заданным декартовым координатам (x, y) числа (ρ, φ) определяются однозначно всегда, за исключением единственного случая: $x = y = 0$. В этом случае $\rho = 0$, а число φ может быть произвольным. Таким образом, любая точка, отличная от начала координат O, взаимно однозначно описывается набором чисел (ρ, φ); точка O описывается набором чисел $(0, \varphi)$, где φ — любое число, удовлетворяющее условию $0 \leqslant \varphi < 2\pi$. Отметим, что линии $\varphi = \text{const}$ в полярной системе координат представляют собой лучи с началом O, а линии $\rho = \text{const}$ — окружности с центром O.

Используются и другие криволинейные координаты: *биполярные координаты* (в них линии $u = \text{const}$ представляют собой окружности, проходящие через две данные точки, а линии $v = \text{const}$ — так называемые окружности Аполлония, пересекающие линии $u = \text{const}$ под прямым углом), *эллиптические координаты* (на них мы остановимся в гл. 4) и ряд других. Выбор тех или иных криволинейных координат определяется, как правило, симметриями изучаемых с их помощью геометрических объектов.

4. Криволинейные координаты в пространстве. Криволинейные координаты в пространстве задаются тремя уравнениями вида

$$\left.\begin{aligned} x &= x\,(u, v, w) \\ y &= y\,(u, v, w) \\ z &= z\,(u, v, w) \end{aligned}\right\}.$$

Наиболее употребительными среди них являются цилиндрические и сферические координаты.

Цилиндрические координаты определяются формулами

$$\left.\begin{aligned} x &= \rho\cos\varphi \\ y &= \rho\sin\varphi \\ z &= z \end{aligned}\right\},$$

где $\rho \geqslant 0$, $0 \leqslant \varphi < 2\pi$, а число z — любое. Геометрический смысл цилиндрических координат ясен: число z имеет тот же смысл, что и в декартовых координатах, а ρ и φ — это полярные координаты проекции данной точки на плоскость Oxy. Название «цилиндрические координаты» объясняется тем, что поверхности $\rho = \mathrm{const}$ представляют собой цилиндрические поверхности радиуса ρ с осью Oz.

Сферические координаты определяются формулами [1])

$$\left.\begin{aligned} x &= \rho\cos\varphi\cos\theta \\ y &= \rho\sin\varphi\cos\theta \\ z &= \rho\sin\theta \end{aligned}\right\},$$

где $\rho \geqslant 0$, $0 \leqslant \varphi < 2\pi$, $-\pi/2 \leqslant \theta \leqslant \pi/2$. Название «сферические координаты» объясняется тем, что поверхности $\rho = \mathrm{const}$ представляют собой сферы радиуса ρ; линии $\varphi = \mathrm{const}$ представляют собой меридианы на этих сферах, а линии $\theta = \mathrm{const}$ — параллели (точки $\theta = -\pi/2$ и $\theta = \pi/2$ соответствуют южному и северному полюсу, а линия $\theta = 0$ — экватору).

§ 2. Векторы

1. Вектор.

Определение. *Вектором называется направленный отрезок, т. е. отрезок, для которого один из концов считается первым (началом), а другой — вторым (концом).*

Если начало и конец вектора совпадают, то он называется *нулевым* (в противном случае — *ненулевым*). Условимся обозначать векторы либо двумя заглавными буквами полужирного шрифта, например $\mathbf{AB}$ (A — начало, B — конец), либо одной строчной буквой полужирного шрифта — $\mathbf{a}$; нулевой вектор будем обозначать символом $\mathbf{o}$.

Из определения следует, что любой отрезок AB определяет два вектора: $\mathbf{AB}$ и $\mathbf{BA}$ (если отрезок вырожденный, то эти векторы совпа-

---

[1]) Иногда сферическими координатами называют криволинейные координаты, определяемые формулами

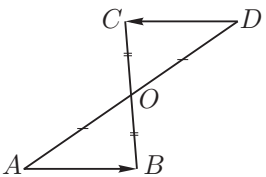
$$\left.\begin{aligned} x &= \rho\cos\varphi\sin\theta \\ y &= \rho\sin\varphi\sin\theta \\ z &= \rho\cos\theta \end{aligned}\right\},$$

где $\rho \geqslant 0$, $0 \leqslant \varphi < 2\pi$, $0 \leqslant \theta \leqslant \pi$. Принципиально эти координаты ни чем не отличаются от рассматриваемых нами — они получаются из них заменой $\theta \to \pi/2 - \theta$.

дают). При этом вектор $\mathbf{BA}$ называется *противоположным* $\mathbf{AB}$ (соответственно $\mathbf{AB}$ — противоположным $\mathbf{BA}$). Вектор, противоположный вектору $\mathbf{a}$, обозначают так: $-\mathbf{a}$.

Длиной или *модулем* вектора $\mathbf{AB}$ называется длина отрезка AB. Длина вектора $\mathbf{AB}$ обозначается так: $|\mathbf{AB}|$ или $|\mathbf{a}|$. Если длина вектора равна единице, то его называют *единичным*.

2. Равенство векторов. Обычно говорят так: векторы называются равными, если их длины равны и они одинаково направлены. Такое определение, при всей своей наглядности, представляется не вполне удачным — им трудно пользоваться. Поэтому поставим своей целью сформулировать другое определение, более удобное с практической точки зрения.



Рассмотрим вектор $\mathbf{AB}$ и произвольную точку O. Пусть $\mathbf{DC}$ — вектор, симметричный $\mathbf{AB}$ относительно точки O (т. е. точка D симметрична точке A, в точка C — точке B). Тогда, очевидно, длины векторов $\mathbf{AB}$ и $\mathbf{DC}$ равны, они лежат на параллельных прямых или на одной прямой, но их направления противоположны (см. рисунок).

Следовательно, в общепринятом смысле векторы $\mathbf{AB}$ и $\mathbf{CD}$ равны. Тем самым, можно сказать так.

Определение. *Два вектора называются равными, если один из них центрально симметричен вектору, противоположному другому.*

3. Координаты вектора.

Определение. *i-й координатой вектора называется разность i-х координат его конца и начала.*

Условимся обозначать через $(m_1, \ldots)$ координату m_1 точки M на прямой, координаты m_1, m_2 точки M на плоскости, или координаты m_1, m_2, m_3 точки M в пространстве. Таким образом, если, например, $A(a_1, \ldots)$, $B(b_1, \ldots)$, $x_1, \ldots$ — координаты вектора $\mathbf{AB}$, то $x_i = b_i - a_i$. Тот факт, что вектор $\mathbf{a}$ имеет координаты $a_1, \ldots$ условимся обозначать так: $\mathbf{a} = \{a_1, \ldots\}$.

Замечание. Из определения следует, что $|\mathbf{a}| = \sqrt{\sum\limits_i a_i^2}$.

Теорема. *Векторы равны тогда и только тогда, когда их координаты совпадают.*

Доказательство. Согласно определению, векторы $\mathbf{AB}$ и $\mathbf{CD}$ равны тогда и только тогда, когда векторы $\mathbf{AB}$ и $\mathbf{DC}$ центрально симметричны, т. е. середины отрезков AD и BC совпадают: $(a_i + d_i)/2 = (b_i + c_i)/2$, или $d_i - c_i = b_i - a_i$, т. е. тогда и только тогда, когда координаты этих векторов совпадают. Теорема доказана.

Следствие 1. *Равные векторы обладают следующими свойствами*:

1° *любой вектор равен самому себе*: $\mathbf{a} = \mathbf{a}$;

2° *если* $\mathbf{a} = \mathbf{b}$, *то* $\mathbf{b} = \mathbf{a}$;

3° *если* $\mathbf{a} = \mathbf{b}$ *и* $\mathbf{b}=\mathbf{c}$, то $\mathbf{a} = \mathbf{c}$.

Следствие 2. *Для любой точки M и любого вектора* $\mathbf{a}$ *существует единственная точка N такая, что* $\mathbf{MN} = \mathbf{a}$ (построение вектора $\mathbf{MN} = \mathbf{a}$ обычно называют *откладыванием вектора* $\mathbf{a}$ *от точки M*).

В самом деле, вектор $\mathbf{MN}$ равен вектору $\mathbf{a}$ тогда и только тогда, когда $n_i - m_i = a_i$, т. е. тогда и только тогда, когда $n_i = a_i + m_i$. Таким образом, положение точки N определяется однозначно.

Следствие 3. *Вектор однозначно определяется своими координатами и началом.*

4. Сумма векторов.

Определение. *Пусть* $\mathbf{AB}$ *и* $\mathbf{CD}$ *— произвольные векторы,* $\mathbf{BE} = \mathbf{CD}$. *Суммой* $\mathbf{AB} + \mathbf{CD}$ *называется вектор* $\mathbf{AE}$.

Таким образом, чтобы сложить два вектора, нужно от конца первого вектора отложить второй; вектор, началом которого является начало первого вектора, а концом — конец отложенного, и есть искомая сумма. Это правило сложения двух векторов называют *правилом треугольника*.

Теорема. *Пусть* $\mathbf{x} = \{x_1, \ldots\}$, $\mathbf{y} = \{y_1, \ldots\}$. *Тогда* $\mathbf{x} + \mathbf{y} = \{x_1 + y_1, \ldots\}$.

Доказательство. Пусть A и B — начало и конец вектора $\mathbf{x}$, E — конец вектора $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Тогда $z_i = e_i - a_i = (e_i - b_i) + (b_i - a_i) = y_i + x_i = x_i + y_i$. Теорема доказана.

5. Произведение вектора на число.

Определение. *Произведением вектора* $\mathbf{AB}$ *на число λ называется такой вектор* $\mathbf{AC}$, *что*:

1° *точки A, B и C лежат на одной прямой*;

2° $AC = |\lambda| AB$;

3° *при* $\lambda > 0$, $AB \neq 0$ *лучи AB и AC совпадают*; *при* $\lambda < 0$, $AB \neq 0$ *лучи AB и AC не совпадают*.

Из этого определения, в частности, следует, что если $\lambda = 0$ или $AB = 0$, то точки A и C совпадают (см. 2°).

Теорема. *Пусть* $\mathbf{x}=\{x_1, \ \ldots \}$. *Тогда* $\lambda\mathbf{x}=\{\lambda x_1, \ \ldots \}$.

Доказательство. Пусть A и B — начало и конец вектора $\mathbf{x}$, C — конец вектора $\mathbf{y} = \lambda\mathbf{x}$, φ — угол между прямой AB [1] и положительной полуосью оси Ox_i $(0 \leqslant \varphi \leqslant \pi/2)$, A_i, B_i и C_i — проекции точек A, B и C на ось Ox_i. Имеем: $A_iC_i = AC\cos\varphi = |\lambda| AB\cos\varphi = |\lambda| A_iB_i$, или $|c_i - a_i| = |\lambda||b_i - a_i|$. Если $(b_i - a_i) = 0$ или $\lambda = 0$, то $(c_i - a_i) = 0$; если $(b_i - a_i) \neq 0$ и $\lambda > 0$, то лучи AB и AC, а значит и лучи A_iB_i и A_iC_i совпадают, поэтому знаки чисел $(b_i - a_i)$ и $(c_i - a_i)$ совпадают; если же $(b_i - a_i) \neq 0$ и $\lambda < 0$, то лучи AB и AC, а значит и лучи A_iB_i и A_iC_i не совпадают, поэтому знаки чисел $(b_i - a_i)$ и $(c_i - a_i)$ разные.

[1] Если точки A и B совпадают, то справедливость утверждения теоремы очевидна.

Таким образом, во всех случаях $(c_i - a_i) = \lambda(b_i - a_i)$, или $y_i = \lambda x_i$, что и требовалось доказать.

6. Отождествление равных векторов. Полученные нами результаты примут существенно более законченный вид, если отождествить все равные друг другу векторы, т. е. принять следующее соглашение: *будем считать, что равные векторы — это один и тот же вектор, но только отложенный от разных точек.*

Предположим, что система координат фиксирована. Тогда каждый вектор $\mathbf{a}$ имеет вполне определенный набор координат $\{a_1, \ldots\}$; обратно, для любых чисел $a_1, \ldots$ существует ровно один вектор $\mathbf{a}$ с координатами $\{a_1, \ldots\}$. Иными словами, *соответствие* $\mathbf{a} \leftrightarrow \{a_1, \ldots\}$ *является взаимно однозначным.*

Это позволяет, в частности, про вектор с координатами $\{a_1, \ldots\}$ говорить так: *вектор* $\{a_1, \ldots\}$. Таким образом, *под словом «вектор» можно понимать не только направленный отрезок, но и упорядоченный набор чисел — координат этого вектора.*

Из теорем п.п. 4 и 5 следует, что векторы, рассматриваемые как упорядоченные наборы чисел, складываются и умножаются точно также, как элементы арифметического пространства. Поэтому *сложение векторов и умножение вектора на число обладают всеми теми же свойствами, что и соответствующие операции в арифметическом пространстве.*

По той же причине к векторам оказываются применимыми все определения и теоремы, связанные с линейными комбинациями и линейной зависимостью элементов арифметического пространства. В частности, любой вектор $\{x_1, x_2, x_3\}$ может быть единственным образом представлен в виде линейной комбинации векторов $\mathbf{e}_1 = \{1, 0, 0\}$, $\mathbf{e}_2 = \{0, 1, 0\}$, $\mathbf{e}_3 = \{0, 0, 1\}$, коэффициентами которой являются координаты этого вектора. Ясно, что векторы $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ — это единичные векторы, лежащие на осях Ox_1, Ox_2 и Ox_3. Как и прежде, будем называть их *координатными векторами.*

§ 3. Скалярное произведение

1. Основные определения.

Определение 1. *Углом между двумя ненулевыми векторами* $\mathbf{AB}$ и $\mathbf{CD}$ *называется угол между лучами* AB *и* CD; *угол между нулевым и любым вектором считается равным* $\pi/2$.

Таким образом, угол φ между любыми двумя векторами удовлетворяет неравенствам $0 \leqslant \varphi \leqslant \pi$.

Определение 2. *Два вектора называются ортогональными, если угол между ними равен* $\pi/2$.

Ортогональность векторов $\mathbf{a}$ и $\mathbf{b}$ условимся обозначать так: $\mathbf{a} \perp \mathbf{b}$. Из определения, в частности, следует, что нулевой вектор ортогонален к любому вектору: $\mathbf{o} \perp \mathbf{b}$.

Определение 3. *Скалярным произведением двух векторов называется произведение их длин на косинус угла между ними.*

Скалярное произведение векторов $\mathbf{a}$ и $\mathbf{b}$ условимся обозначать так: $(\mathbf{ab})$. Из определения, в частности, следует, что $(\mathbf{aa}) = |\mathbf{a}|^2$.

Теорема. $\mathbf{a}\perp\mathbf{b}$ *тогда и только тогда, когда* $(\mathbf{ab}) = 0$.

Доказательство. 1°. Если $\mathbf{a}\perp\mathbf{b}$, то $(\mathbf{ab}) = |\mathbf{a}||\mathbf{b}|\cos(\pi/2) = |\mathbf{a}||\mathbf{b}|0 = 0$.

2°. Если $(\mathbf{ab}) = 0$, т. е. $|\mathbf{a}||\mathbf{b}|\cos\varphi = 0$ (φ — угол между $\mathbf{a}$ и $\mathbf{b}$), то либо $|\mathbf{a}| = 0$, либо $|\mathbf{b}| = 0$, либо $\cos\varphi = 0$. В каждом из этих случаев $\varphi = \pi/2$. Теорема доказана.

2. Скалярное произведение в координатах.

Теорема. *Если* $\mathbf{a} = \{a_1, \ldots\}$, $\mathbf{b} = \{b_1, \ldots\}$, то $(\mathbf{ab}) = \sum\limits_i a_i b_i$.

Доказательство. Пусть $\mathbf{OA} = \mathbf{a}$, $\mathbf{OB} = \mathbf{b}$. Применяя к треугольнику OAB теорему косинусов (она, очевидно, верна и для вырожденного треугольника), получаем:

$$AB^2 = |\mathbf{b} - \mathbf{a}|^2 = OA^2 + OB^2 - 2OA \cdot OB\cos\varphi =$$

$$= |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}|\cos\varphi = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2(\mathbf{ab}),$$

откуда $(\mathbf{ab}) = \frac{1}{2}(|\mathbf{a}|^2 + |\mathbf{b}|^2 - |\mathbf{b} - \mathbf{a}|^2) = \frac{1}{2}\sum\limits_i \left[a_i^2 + b_i^2 - (b_i - a_i)^2\right] = \frac{1}{2}\sum\limits_i 2a_i b_i = \sum\limits_i a_i b_i$. Теорема доказана.

Следствие. *Косинус угла между ненулевыми векторами* $\mathbf{a} = \{a_1, \ldots\}$ и $\mathbf{b} = \{b_1, \ldots\}$ *может быть вычислен по формуле*

$$\cos\varphi = \frac{\sum\limits_i a_i b_i}{\sqrt{\sum\limits_i a_i^2}\sqrt{\sum\limits_i b_i^2}}.$$

Замечание. При доказательстве теоремы мы пользовались теоремой косинусов. Между тем, в некоторых учебниках по элементарной геометрии теорема косинусов выводится из свойств скалярного произведения. Тем самым, у обучавшимся по этим учебникам может возникнуть ощущение «порочного круга». В действительности никакого «порочного круга» здесь нет — в большинстве учебников теорема косинусов доказывается без использования свойств скалярного произведения.

3. Свойства скалярного произведения.

Теорема.

1° *Для любых векторов* $\mathbf{a}$ и $\mathbf{b}$ *выполняется равенство* $(\mathbf{ab}) = (\mathbf{ba})$;

2° *для любых векторов* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *выполняется равенство* $((\mathbf{a} + \mathbf{b})\mathbf{c}) = (\mathbf{ac}) + (\mathbf{bc})$;

3° *для любых векторов* $\mathbf{a}$ и $\mathbf{b}$ *и любого числа* λ *выполняется равенство* $((\lambda\mathbf{a})\mathbf{b}) = \lambda(\mathbf{ab})$;

4° *для любого вектора* $\mathbf{a}$ *произведение* $(\mathbf{aa}) \geqslant 0$, причем $(\mathbf{aa}) = 0$ *тогда и только тогда, когда* $\mathbf{a} = \mathbf{o}$.

Доказательство.

1°. Это следует непосредственно из определения скалярного произведения;

2°. $((\mathbf{a}+\mathbf{b})\mathbf{c}) = \sum\limits_i (a_i + b_i)\, c_i = \sum\limits_i a_i c_i + \sum\limits_i b_i c_i = (\mathbf{ac}) + (\mathbf{bc})$;

3°. $((\lambda\mathbf{a})\mathbf{b}) = \sum\limits_i (\lambda a_i)\, b_i = \lambda \sum\limits_i a_i b_i = \lambda(\mathbf{ab})$;

4°. Это следует из того, что $(\mathbf{aa}) = |\mathbf{a}|^2$.

Теорема доказана.

Следствие. $a_i = (\mathbf{ae}_i)$.

Для доказательства достаточно умножить равенство $\mathbf{a} = \sum\limits_i a_i \mathbf{e}_i$ скалярно на $\mathbf{e}_i$ и учесть, что $(\mathbf{e}_i\mathbf{e}_j) = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j. \end{cases}$

4. Площадь параллелограмма. Условимся обозначать площадь параллелограмма, построенного на векторах $\mathbf{a}$ и $\mathbf{b}$, символом $S(\mathbf{a}, \mathbf{b})$.

Теорема 1. $S(\mathbf{a}, \mathbf{b})^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{ab})^2$.

Доказательство. Пусть φ — угол между векторами $\mathbf{a}$ и $\mathbf{b}$. Имеем

$$S(\mathbf{a}, \mathbf{b})^2 = (|\mathbf{a}||\mathbf{b}| \sin\varphi)^2 = |\mathbf{a}|^2|\mathbf{b}|^2(1 - \cos^2\varphi) = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{ab})^2.$$

Теорема доказана.

Теорема 2. *Площадь параллелограмма, построенного на векторах* $\mathbf{a} = \{a_1, a_2\}$ *и* $\mathbf{b} = \{b_1, b_2\}$, *выражается формулой*

$$S(\mathbf{a}, \mathbf{b}) = \left| \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right| = |\Delta(\mathbf{a}, \mathbf{b})|.$$

Доказательство. Воспользуемся двумя свойствами определителей: а) определитель не меняется при транспонировании; б) определитель произведения матриц равен произведению их определителей. Имеем:

$$\Delta(\mathbf{a}, \mathbf{b})^2 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} (\mathbf{aa}) & (\mathbf{ab}) \\ (\mathbf{ba}) & (\mathbf{bb}) \end{vmatrix} = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{ab})^2 = S(\mathbf{a}, \mathbf{b})^2.$$

Теорема доказана.

5. Объем параллелепипеда. Рассмотрим вектор

$$\mathbf{p}\,(\mathbf{a}, \mathbf{b}) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \{\Delta_{23}, -\Delta_{13}, \Delta_{12}\}, \quad \text{где} \quad \Delta_{ij} = \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix}.$$

Теорема 1. *Вектор* $\mathbf{p}(\mathbf{a}, \mathbf{b})$ *обладает следующими свойствами*:

1°. $\mathbf{p}(\mathbf{a}, \mathbf{b}) \perp \mathbf{a}$, $\mathbf{p}(\mathbf{a}, \mathbf{b}) \perp \mathbf{b}$;

2°. $|\mathbf{p}(\mathbf{a}, \mathbf{b})| = S(\mathbf{a}, \mathbf{b})$.

Доказательство.

1°. $(\mathbf{ap}(\mathbf{a},\mathbf{b}))=a_1\Delta_{23}-a_2\Delta_{13}+a_3\Delta_{12}=\Delta(\mathbf{a},\mathbf{a},\mathbf{b})=0$ [1], $(\mathbf{bp}(\mathbf{a},\mathbf{b}))=$ $=b_1\Delta_{23}-b_2\Delta_{13}+b_3\Delta_{12}=\Delta(\mathbf{b},\mathbf{a},\mathbf{b})=0$.

2°. Обратим внимание на то, что $\Delta_{ij}=-\Delta_{ji}$ (в частности $\Delta_{ii}=0$). Поэтому

$$|\mathbf{p}\,(\mathbf{a},\mathbf{b})\,|^2=\Delta_{23}^2+\Delta_{13}^2+\Delta_{12}^2=\frac{1}{2}\left(\sum_{i,j=1}^{3}\Delta_{ij}^2\right)=\frac{1}{2}\left(\sum_{i,j=1}^{3}(a_ib_j-a_jb_i)^2\right)=$$

$$=\frac{1}{2}\left(\sum_{i,j=1}^{3}a_i^2b_j^2+\sum_{i,j=1}^{3}a_j^2b_i^2-2\sum_{i,j=1}^{3}a_ia_jb_ib_j\right)=$$

$$=\frac{1}{2}\left(\sum_{i=1}^{3}a_i^2\sum_{j=1}^{3}b_j^2+\sum_{j=1}^{3}a_j^2\sum_{i=1}^{3}b_i^2\right)-\sum_{i=1}^{3}a_ib_i\sum_{j=1}^{3}a_jb_j=$$

$$=|\mathbf{a}|^2\,|\mathbf{b}|^2-(\mathbf{ab})^2=S\,(\mathbf{a},\mathbf{b})^2\,.$$

Теорема доказана.

Условимся обозначать объем параллелепипеда, построенного на векторах $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$, символом $V(\mathbf{a},\mathbf{b},\mathbf{c})$.

Теорема 2. *Объем параллелепипеда, построенного на векторах* $\mathbf{a}=\{a_1,a_2,a_3\}$, $\mathbf{b}=\{b_1,b_2,b_3\}$ *и* $\mathbf{c}=\{c_1,c_2,c_3\}$, *выражается формулой*

$$V(\mathbf{a},\mathbf{b},\mathbf{c})=|\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}|=|\Delta(\mathbf{a},\mathbf{b},\mathbf{c})|.$$

Доказательство. Отложим векторы $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ от общего начала. Примем плоскость, содержащую векторы $\mathbf{a}$ и $\mathbf{b}$, за плоскость основания параллелепипеда и обозначим буквой φ угол между векторами $\mathbf{p}(\mathbf{a},\mathbf{b})$ и $\mathbf{c}$. Имеем: $V(\mathbf{a},\mathbf{b},\mathbf{c})=S(\mathbf{a},\mathbf{b})|\mathbf{c}||\cos\varphi|=|\mathbf{p}(\mathbf{a},\mathbf{b})||\mathbf{c}||\cos\varphi|=$ $=|(\mathbf{p}(\mathbf{a},\mathbf{b})\mathbf{c})|=|c_1\Delta_{23}-c_2\Delta_{13}+c_3\Delta_{12}|=|\Delta(\mathbf{c},\mathbf{a},\mathbf{b})|=|\Delta(\mathbf{a},\mathbf{b},\mathbf{c})|$. Теорема доказана.

§ 4. Базис

1. Коллинеарные векторы.

Определение. *Векторы называются коллинеарными, если при откладывании их от общего начала они оказываются лежащими на одной прямой.*

Теорема 1. *Два вектора коллинеарны тогда и только тогда, когда они линейно зависимы.*

Доказательство. Рассмотрим векторы $\mathbf{a}$ и $\mathbf{b}$ с общим началом и введем в образованной ими плоскости систему координат. Векторы $\mathbf{a}$ и $\mathbf{b}$ коллинеарны тогда и только тогда, когда площадь построенного на них параллелограмма $S(\mathbf{a},\mathbf{b})=|\Delta(\mathbf{a},\mathbf{b})|$ равна нулю, т. е. тогда и только тогда, когда строки определителя $\Delta(\mathbf{a},\mathbf{b})$ или, что то же самое, векторы $\mathbf{a}$ и $\mathbf{b}$ линейно зависимы. Теорема доказана.

---

[1] Если в определителе две строки совпадают, то он равен нулю.

Следствие. *Если два вектора неколлинеарны, то они линейно независимы.*

Теорема 2. *Если векторы* $\mathbf{a}$ и $\mathbf{b}$ *коллинеарны, причем вектор* $\mathbf{a}$ *ненулевой, то существует такое число* λ, *что* $\mathbf{b} = \lambda\mathbf{a}$.

Доказательство. Из условия теоремы следует, что ранг матрицы со строками $\mathbf{a}$ и $\mathbf{b}$ равен 1, причем строка $\mathbf{a}$ — базисная. Поэтому, согласно теореме о базисном миноре, строка $\mathbf{b}$ пропорциональна строке $\mathbf{a}$, что и требовалось доказать.

2. Компланарные векторы.

Определение. *Векторы называются компланарными, если при откладывании их от общего начала они оказываются лежащими в одной плоскости.*

Теорема 1. *Три вектора компланарны тогда и только тогда, когда они линейно зависимы.*

Доказательство. Три вектора $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ компланарны тогда и только тогда, когда объем построенного на них параллелепипеда $V(\mathbf{a}, \mathbf{b}, \mathbf{c}) = |\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c})|$ равен нулю, т. е. тогда и только тогда, когда строки определителя $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c})$ или, что то же самое, векторы $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ линейно зависимы. Теорема доказана.

Следствие. *Если три вектора некомпланарны, то они линейно независимы.*

Теорема 2. *Если векторы* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *компланарны, а векторы* $\mathbf{a}$ и $\mathbf{b}$ *неколлинеарны, то существуют такие числа* λ и μ, *что* $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$.

Доказательство. Из условия теоремы следует, что ранг матрицы со строками $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ равен 2, причем строки $\mathbf{a}$ и $\mathbf{b}$ — базисные. Поэтому, согласно теореме о базисном миноре, строка $\mathbf{c}$ равна линейной комбинации строк $\mathbf{a}$ и $\mathbf{b}$. Теорема доказана.

3. Линейная зависимость четырех векторов.

Теорема 1. *Любые четыре вектора линейно зависимы.*

Доказательство. Пусть $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ и $\mathbf{d}$ — произвольные векторы. Составим матрицу A, строками которой являются координаты этих векторов. Ее ранг не превосходит количества ее столбцов, т. е. $\operatorname{rang} A \leqslant 3$. Поэтому ее строки или, что то же самое, векторы $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ и $\mathbf{d}$ линейно зависимы. Теорема доказана.

Теорема 2. *Если векторы* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *некомпланарны, то для любого вектора* $\mathbf{d}$ *существуют такие числа* λ, μ и ν, *что* $\mathbf{d} = \lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{c}$.

Доказательство. Из условия теоремы следует, что ранг матрицы со строками $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ и $\mathbf{d}$ равен 3, причем строки $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ — базисные. Поэтому, согласно теореме о базисном миноре, строка $\mathbf{d}$ равна линейной комбинации строк $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$. Теорема доказана.

4. Базис.

Определение 1. *Говорят, что вектор разложен по каким-то векторам, если он представлен в виде их линейной комбинации.*

Определение 2. *Упорядоченная совокупность векторов* $\{\mathbf{a}_i\}$ *называется базисом, если*:

1° *эти векторы линейно независимы*;

2° *любой вектор может быть по ним разложен*.

Из теорем пп. 1–3 следует, что:

1° *любой ненулевой вектор образует базис на прямой*;

2° *любые два неколлинеарных вектора образуют базис на плоскости*;

3° *любые три некомпланарных вектора образуют базис в пространстве*.

Теорема 1. *Коэффициенты разложения вектора по базису* (т. е. числа, фигурирующие в этом разложении), *определяются единственным образом*.

Доказательство. Допустим, что вектор $\mathbf{b}$ разложен по базису $\{\mathbf{a}_i\}$ двумя способами: $b = \sum\limits_i \lambda_i a_i$ и $b = \sum\limits_i \mu_i a_i$. Вычитая второе равенство из первого, получаем: $\mathbf{o} = \sum\limits_i (\lambda_i - \mu_i)\,\mathbf{a}_i$. Поскольку векторы $\mathbf{a}_i$ линейно независимы, то $\lambda_i = \mu_i$, при всех i. Теорема доказана.

Теорема 2. *Векторы* $\{\mathbf{a}_i\}$ *образуют базис тогда и только тогда, когда любой вектор может быть единственным образом разложен по ним*.

Доказательство. Если векторы $\{\mathbf{a}_i\}$ образуют базис, то любой вектор может быть разложен по ним, причем (согласно теореме 1) единственным образом. Обратно, если любой вектор может быть единственным образом разложен по векторам $\{\mathbf{a}_i\}$, то они образуют базис, так как: 1° любой вектор может быть по ним разложен; 2° они линейно независимы, поскольку вектор $\mathbf{o}$ может быть представлен только в виде их тривиальной линейной комбинации.

5. Аффинные координаты. *Аффинная система координат* определяется заданием произвольного базиса $\{\mathbf{a}_i\}$ и точки O, называемой *началом координат*.

Пусть $\mathbf{OA}_i = \mathbf{a}_i$. Лучи OA_i принимаются за положительные полуоси, а прямые OA_i называются *осями аффинной системы координат*.

Аффинными координатами вектора называются коэффициенты его разложения по базису $\{\mathbf{a}_i\}$. Ясно, что каждому вектору соответствует вполне определенный набор аффинных координат, и обратно — каждому набору аффинных координат соответствует ровно один вектор.

Аффинными координатами точки M называются координаты вектора $\mathbf{OM}$.

В частности, если векторы $\mathbf{a}_i$ — единичные и попарно ортогональные, то, умножая равенство $\mathbf{OM} = \sum\limits_s \lambda_s \mathbf{a}_s$ скалярно на $\mathbf{a}_i$, получаем:

$$\lambda_i = (\mathbf{OM}\mathbf{a}_i) = |\mathbf{OM}||\mathbf{a}_i| \cos\varphi_i = OM \cos\varphi_i,$$

где φ_i — угол между лучами OM и OA_i.

Таким образом, число λ_i представляет собой в этом случае координату проекции точки M на ось OA_i, т. е. i-ю декартову координату точки M. Иными словами, *декартова система координат является частным случаем аффинной*.

Глава 2
ПРЕОБРАЗОВАНИЕ КООРДИНАТ

§ 1. Преобразование координат на плоскости

1. Правые и левые пары. Пусть $\mathbf{a}$ и $\mathbf{b}$ — два неколлинеарных вектора с общим началом. Физики говорят так: упорядоченная пара неколлинеарных векторов $\mathbf{a}$ и $\mathbf{b}$ образуют правую (левую) пару, если кратчайший поворот от $\mathbf{a}$ к $\mathbf{b}$ осуществляется против (по) часовой стрелки(е). На интуитивном уровне эта фраза кажется содержательной — она отвечает нашим представлениям о плоскости. Однако с точки зрения геометрии она лишена смысла, поскольку в аксиомах и теоремах геометрии нет упоминания о каких-либо часовых стрелках. Наша ближайшая задача состоит в том, чтобы придать указанным понятиям ясный геометрический смысл.

Прежде всего, следует понять, что может и чего не может дать геометрия для решения нашей задачи. Судя по всему, интуитивно ясный факт существования правых и левых пар векторов свидетельствует о наличии какого-то геометрического критерия, позволяющего все упорядоченные пары неколлинеарных векторов разделить на два непересекающихся класса. С другой стороны, ясно, что даже если указанное разделение на два класса удастся произвести, то эти классы окажутся эквивалентными — нет и не может быть никакого геометрического признака, по которому один из них можно было бы выделить и назвать классом правых пар. В самом деле, если посмотреть на плоскость с противоположной стороны, то все аксиомы, а значит, и теоремы, сохранятся, но на направление «по часовой стрелке» изменится на противоположное. Поэтому не остается ничего другого, кроме как декларативно объявить один из классов классом правых пар.

Сказанное позволяет сделать первый шаг в решении стоящей перед нами задачи.

Определение 1. *Плоскость называется ориентированной, если на ней задана система координат, которую называют исходной.*

На ориентированной плоскости естественным образом возникает упорядоченная пара неколлинеарных векторов $\mathbf{e}_1$, $\mathbf{e}_2$ — координатных векторов исходной системы координат. Эту пару естественно принять за эталон ориентации.

Следующий шаг состоит в нахождении критерия, позволяющего любую упорядоченную пару неколлинеарных векторов отнести к одному из двух классов. Чтобы это сделать, нужно понять, какими свойствами

должен обладать этот критерий. Вернемся к интуитивным представлениям. Ясно, что если пара $\mathbf{a}$, $\mathbf{b}$ — правая, то пара $\mathbf{a}$, $(-\mathbf{b})$, равно как и $\mathbf{b}$, $\mathbf{a}$, должна быть левой. Вспоминая свойства определителя, мы можем теперь сформулировать такое определение.

Определение 2. *Упорядоченная пара неколлинеарных векторов* $\mathbf{a}$, $\mathbf{b}$ *называется правой* (*левой*), *если в исходной системе координат* $\Delta(\mathbf{a}, \mathbf{b}) > 0$ (< 0).

Ясно, что согласно этому определению любая упорядоченная пара неколлинеарных векторов оказывается либо правой, либо левой, причем пара $\mathbf{e}_1$, $\mathbf{e}_2$ — правая.

Наконец, сформулируем еще одно определение.

Определение 3. *Произвольная декартова система координат на ориентированной плоскости называется правой* (*левой*), *если ее координатные векторы образуют правую* (*левую*) *пару.*

Таким образом, исходная система координат — правая.

2. Собственные и несобственные преобразования. Рассмотрим теперь две декартовых системы координат: Ox_1x_2 и $\widetilde{O}\widetilde{x}_1\widetilde{x}_2$. Первую из них будем называть *старой*, а вторую — *новой*. Разложим каждый из координатных векторов новой системы координат по координатным векторам старой:

$$\left.\begin{aligned} \widetilde{\mathbf{e}}_1 &= a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 \\ \widetilde{\mathbf{e}}_2 &= a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 \end{aligned}\right\},$$

где $a_{ij} = (\widetilde{\mathbf{e}}_i\mathbf{e}_j)$. Поскольку $|\widetilde{\mathbf{e}}_1| = |\widetilde{\mathbf{e}}_2| = 1$ и $(\widetilde{\mathbf{e}}_1\widetilde{\mathbf{e}}_2) = 0$, то

$$\left.\begin{aligned} a_{11}^2 + a_{12}^2 &= 1 \\ a_{21}^2 + a_{22}^2 &= 1 \\ a_{11}a_{21} + a_{12}a_{22} &= 0 \end{aligned}\right\}.$$

Так как $|\widetilde{\mathbf{e}}_1| = |\mathbf{e}_1| = 1$, то $a_{11} = (\widetilde{\mathbf{e}}_1\mathbf{e}_1) = \cos\alpha$, где α — угол между векторами $\widetilde{e}_1$ и e_1 $(0 \leqslant a \leqslant \pi)$. Положим $\varphi = \alpha$, если $a_{12} \geqslant 0$, или $\varphi = 2\pi - \alpha$ в противоположном случае. В результате получим: $a_{11} = \cos\varphi$, $a_{12} = \sin\varphi$. Аналогично, из второго равенства находим: $a_{22} = \cos\theta$, $a_{21} = \sin\theta$. Подставляя эти значения в третье равенство, получаем: $\sin(\varphi + \theta) = 0$, откуда $\varphi + \theta = k\pi$, или $\theta = k\pi - \varphi$, а значит, $a_{21} = -\sin\varphi \cdot \cos k\pi$, $a_{22} = \cos\varphi \cdot \cos k\pi$. Возможны два случая.

1°. Число k — четное, и, следовательно, $a_{21} = -\sin\varphi$, $a_{22} = \cos\varphi$,

$$\Delta\left(\widetilde{\mathbf{e}}_1, \widetilde{\mathbf{e}}_2\right) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{vmatrix} = 1.$$

В этом случае переход от старой системы координат к новой называется *собственным преобразованием*. В частности, в соответствии с нашим определением, переход от исходной системы координат к правой осуществляется при помощи собственного преобразования.

2°. Число k — нечетное, и, следовательно, $a_{21} = \sin\varphi$, $a_{22} = -\cos\varphi$,

$$\Delta\left(\widetilde{\mathbf{e}}_1, \widetilde{\mathbf{e}}_2\right) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{vmatrix} = -1.$$

В этом случае переход от старой системы координат к новой называется *несобственным преобразованием*. В частности, переход от исходной системы координат к левой осуществляется при помощи несобственного преобразования.

Собственное преобразование называют также *поворотом*. Ясно, что несобственное преобразование может рассматриваться как результат последовательного выполнения поворота и симметрии относительно оси $\widetilde{O}\widetilde{x}_1$.

Замечание. То, что $|\Delta\left(\widetilde{e}_1, \widetilde{e}_2\right)| = 1$, было ясно и без вычислений — эта величина представляет собой площадь параллелограмма, построенного на векторах $\widetilde{e}_1, \widetilde{e}_2$, т. е. площадь единичного квадрата.

3. Преобразование координат вектора. Пусть $\mathbf{a}$ — произвольный вектор. Имеем:

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 = \widetilde{a}_1\widetilde{\mathbf{e}}_1 + \widetilde{a}_2\widetilde{\mathbf{e}}_2.$$

Умножая это равенство скалярно сначала на вектор $\widetilde{\mathbf{e}}_1$, а затем на вектор $\widetilde{\mathbf{e}}_2$, получаем:

$$\left.\begin{aligned} \widetilde{a}_1 &= a_{11}a_1 + a_{12}a_2 \\ \widetilde{a}_2 &= a_{21}a_1 + a_{22}a_2 \end{aligned}\right\}.$$

Таким образом, при собственном преобразовании координаты вектора $\mathbf{a}$ преобразуются по формулам

$$\left.\begin{aligned} \widetilde{a}_1 &= a_1\cos\varphi \quad + a_2\sin\varphi \\ \widetilde{a}_2 &= -a_1\sin\varphi + a_2\cos\varphi \end{aligned}\right\},$$

а при несобственном — по формулам

$$\left.\begin{aligned} \widetilde{a}_1 &= a_1\cos\varphi + a_2\sin\varphi \\ \widetilde{a}_2 &= a_1\sin\varphi - a_2\cos\varphi \end{aligned}\right\}.$$

Отметим, что расположение начала новой системы координат относительно старой в этих формулах роли не играет.

4. Правые системы координат.

Теорема. *Для любых векторов* $\mathbf{a}$ и $\mathbf{b}$ *величина* $\Delta(\mathbf{a}, \mathbf{b})$ *одна и та же во всех правых системах координат.*

Доказательство. Пусть Ox_1x_2 — исходная система координат, а $\widetilde{O}\widetilde{x}_1\widetilde{x}_2$ — любая правая. Имеем:

$$\widetilde{\Delta}(\mathbf{a},\mathbf{b}) = \begin{vmatrix} \widetilde{a}_1 & \widetilde{a}_2 \\ \widetilde{b}_1 & \widetilde{b}_2 \end{vmatrix} = \begin{vmatrix} (a_1\cos\varphi + a_2\sin\varphi) & (-a_1\sin\varphi + a_2\cos\varphi) \\ (b_1\cos\varphi + b_2\sin\varphi) & (-b_1\sin\varphi + b_2\cos\varphi) \end{vmatrix} =$$

$$= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \Delta(\mathbf{a},\mathbf{b}).$$

Теорема доказана.

Следствие 1. *Для определения ориентации пары векторов можно пользоваться не только исходной, но и любой правой системой координат*. Иными словами, с этой точки зрения *все правые системы координат равноправны*.

Следствие 2. *Переход от одной правой системы координат к другой осуществляется при помощи собственного преобразования* (*поворота*).

Чтобы привести наши построения в полное соответствие с наглядными представлениями, осталось сделать последний шаг — принять соглашение о правиле изображения правой пары векторов на рисунках.

Соглашение. Исходная система координат изображается на рисунках так, что кратчайший поворот от $\mathbf{e}_1$ к $\mathbf{e}_2$, т. е. поворот на угол $\pi/2$, осуществляется против часовой стрелки.

Из этого следует, что и любая правая система координат изображается по тому же правилу. В самом деле, $\mathbf{e}_1 = \{1,0\} = \{\cos 0, \sin 0\}$, $\mathbf{e}_2 = \{0,1\} = \{\cos(0+\pi/2), \sin(0+\pi/2)\}$, $\widetilde{e}_1 = \{\cos\varphi, \sin\varphi\}$, $\widetilde{e}_2 = \{-\sin\varphi, \cos\varphi\} = \{\cos(\varphi+\pi/2), \sin(\varphi+\pi/2)\}$, поэтому поворот от $\widetilde{e}_1$ к $\widetilde{e}_2$ на угол $\pi/2$ осуществляется в том же направлении, что и от $\mathbf{e}_1$ к $\mathbf{e}_2$.

Наконец, то же правило изображения распространяется и на любую правую пару векторов $\mathbf{a}$ и $\mathbf{b}$. Действительно, выбирая правую систему координат Ox_1x_2 так, чтобы вектор $\mathbf{a}$ имел координаты $\{a,0\}$, $a>0$, из условия $\begin{vmatrix} a & 0 \\ b_1 & b_2 \end{vmatrix} > 0$ найдем: $b_2 > 0$. Это означает, что при откладывании вектора $\mathbf{b}$ от точки O он оказывается лежащим по ту же сторону от оси Ox_1, что и положительная полуось оси Ox_2. Поэтому кратчайший поворот от $\mathbf{a}$ к $\mathbf{b}$ осуществляется также против часовой стрелки.

В дальнейшем мы будем рассматривать только правые системы координат на плоскости.

5. Преобразование координат точки. Пусть, как и прежде, Ox_1x_2 — старая система координат, $\widetilde{O}\widetilde{x}_1\widetilde{x}_2$ — новая. Выразим «новые» координаты произвольной точки M через ее «старые» координаты. Для этого заметим, что «новые» координаты точки M — это координаты вектора $\widetilde{\mathbf{O}}\mathbf{M}$ в новой системе координат. Найдем сначала координаты

этого вектора в старой системе координат, а затем воспользуемся результатом п. 3. Имеем:

$$\widetilde{\mathbf{O}}\mathbf{M} = \widetilde{\mathbf{O}}\mathbf{O} + \mathbf{O}\mathbf{M} = -\mathbf{O}\widetilde{\mathbf{O}} + \mathbf{O}\mathbf{M} = \{m_1 - o_1, m_2 - o_2\}\,.$$

Следовательно,

$$\left.\begin{aligned} \widetilde{m}_1 &= (m_1 - o_1)\cos\varphi + (m_2 - o_2)\sin\varphi \\ \widetilde{m}_2 &= -\,(m_1 - o_1)\sin\varphi + (m_2 - o_2)\cos\varphi \end{aligned}\right\}$$

(напомним, что мы договорились рассматривать только правые системы координат).

Таким образом, переход от старой системы координат к новой осуществляется при помощи параллельного переноса на вектор $\mathbf{O}\widetilde{\mathbf{O}}$ и поворота на угол φ вокруг точки $\widetilde{O}$.

§ 2. Преобразование координат в пространстве

1. Преобразование координат вектора.

Определение 1. *Пространство называется ориентированным, если в нем задана система координат, которую называют исходной.*

Определение 2. *Упорядоченная тройка некомпланарных векторов* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *называется правой* (*левой*), *если в исходной системе координат* $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) > 0$ (< 0).

Определение 3. *Произвольная система координат в ориентированном пространстве называется правой* (*левой*), *если ее координатные векторы образуют правую* (*левую*) *тройку.*

Таким образом, поскольку $\Delta(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1 > 0$, то исходная система координат — правая.

Рассмотрим теперь две системы координат: $Ox_1x_2x_3$ и $\widetilde{O}\widetilde{x}_1\widetilde{x}_2\widetilde{x}_3$. Первую из них будем называть *старой*, а вторую — *новой*. Разложим каждый из координатных векторов новой системы координат по координатным векторам старой:

$$\widetilde{\mathbf{e}}_i = \sum_{s=1}^{3} a_{is}\mathbf{e}_s,$$

где $a_{ij} = (\widetilde{\mathbf{e}}_i\mathbf{e}_j)$ — координаты вектора $\widetilde{\mathbf{e}}_i$ в старой система координат. Определитель матрицы A с точностью до знака равен объему параллелепипеда, построенного на векторах $\widetilde{\mathbf{e}}_1, \widetilde{\mathbf{e}}_2, \widetilde{\mathbf{e}}_3$, т. е. 1 — объему единичного куба. Если $\det A = 1$, преобразование называется *собственным*, а если $\det A = -1$, то *несобственным*. В частности, переход от исходной системы координат к правой осуществляется при помощи

собственного преобразования, а к левой — при помощи несобственного преобразования.

Пусть $\mathbf{a}$ — произвольный вектор. Имеем:

$$\mathbf{a} = \sum_{k=1}^{3} \widetilde{a}_k \widetilde{\mathbf{e}}_k = \sum_{s=1}^{3} a_s \mathbf{e}_s.$$

Умножая это равенство скалярно на вектор $\widetilde{\mathbf{e}}_i$, получаем:

$$\widetilde{a}_i = \sum_{s=1}^{3} a_{is} a_s. \tag{1}$$

Отметим, что расположение начала новой системы координат относительно старой в полученных формулах роли не играет.

Теорема. *Для любых векторов* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *величина* $\Delta(\mathbf{a},\mathbf{b},\mathbf{c})$ *одна и та же во всех правых системах координат.*

Доказательство. Пусть $Ox_1x_2x_3$ — исходная система координат, а $\widetilde{O}\widetilde{x}_1\widetilde{x}_2\widetilde{x}_3$ — любая правая. Пользуясь свойствами произведения матриц, получаем:

$$\widetilde{\Delta}(\mathbf{a},\mathbf{b},\mathbf{c}) = \begin{vmatrix} \widetilde{a}_1 & \widetilde{a}_2 & \widetilde{a}_3 \\ \widetilde{b}_1 & \widetilde{b}_2 & \widetilde{b}_3 \\ \widetilde{c}_1 & \widetilde{c}_2 & \widetilde{c}_3 \end{vmatrix} = \begin{vmatrix} \sum\limits_s a_{1s}a_s & \sum\limits_s a_{2s}a_s & \sum\limits_s a_{3s}a_s \\ \sum\limits_s a_{1s}b_s & \sum\limits_s a_{2s}b_s & \sum\limits_s a_{3s}b_s \\ \sum\limits_s a_{1s}c_s & \sum\limits_s a_{2s}c_s & \sum\limits_s a_{3s}c_s \end{vmatrix} =$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \cdot \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \Delta(\mathbf{a},\mathbf{b},\mathbf{c}).$$

Теорема доказана.

Следствие 1. *Для определения ориентации тройки векторов можно пользоваться не только исходной, но и любой правой системой координат.* Иными словами, с этой точки зрения *все правые системы координат равноправны.*

Следствие 2. *Переход от одной правой системы координат к другой осуществляется при помощи собственного преобразования.*

В дальнейшем *условимся рассматривать только правые системы координат.*

2. Углы Эйлера. На первый взгляд может показаться, что формулы (1) содержат 9 произвольных параметров — элементов матрицы A. На самом деле, как мы сейчас увидим, их не 9, и не 8 ($\det A = 1$), а всего 3. Точнее, оказывается, что любые две правые системы координат с общим началом можно совместить друг с другом, выполняя последовательно три поворота в координатных плоскостях (т. е. вокруг

осей координат). Отметим, что каждый из таких поворотов является собственным преобразованием:

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{vmatrix} = \begin{vmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{vmatrix} = \begin{vmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Пусть $Ox_1x_2x_3$ и $O\widetilde{x}_1\widetilde{x}_2\widetilde{x}_3$ — две правые системы координат с общим началом. Поступим так.

1°. Поворотом в плоскости Ox_1x_2 на угол α перейдем к системе $Ox_1'x_2'x_3$, в которой ось Ox_1' лежит в плоскости $O\widetilde{x}_1\widetilde{x}_2$. Для этого достаточно взять в качестве оси Ox_1' линию пересечения плоскостей Ox_1x_2 и $O\widetilde{x}_1\widetilde{x}_2$ (или положить $\alpha = 0$, если эти плоскости совпадают).

2°. Поворотом в плоскости $Ox_2'x_3$ на угол β перейдем к системе $Ox_1'x_2''\widetilde{x}_3$, т. е. совместим оси Ox_3 и $O\widetilde{x}_3$. Это возможно, поскольку каждая из этих осей перпендикулярна к оси Ox_1'.

3°. Наконец, поворотом в плоскости $Ox_1'x_2''$ на угол γ перейдем к системе $O\widetilde{x}_1\widetilde{x}_2\widetilde{x}_3$. Это можно представить себе так. Совместим оси Ox_1' и $O\widetilde{x}_1$, что возможно, поскольку каждая из них перпендикулярна к оси $O\widetilde{x}_3$. При этом положительная полуось оси Ox_2'' совместится либо с положительной полуосью оси $O\widetilde{x}_2$, либо с ее продолжением. Но все выполненные нами преобразования — собственные, поэтому все рассмотренные системы координат — правые. Значит, положительная полуось оси Ox''_2 совместится с положительной полуосью оси $O\widetilde{x}_2$.

Углы α, β и γ называются *углами Эйлера*. Ясно, что через них можно выразить все коэффициенты a_{ij} в формулах (1).

Соглашение. Исходная система координат изображается на рисунках так, что если смотреть с положительной стороны оси Oz, то кратчайший поворот от $\mathbf{e}_1$ к $\mathbf{e}_2$, т. е. поворот на угол $\pi/2$, осуществляется против часовой стрелки.

Из этого следует, что и любая правая система координат изображается по тому же правилу, поскольку это правило не меняется при поворотах на углы Эйлера.

Наконец, то же правило изображения распространяется и на любую правую тройку векторов $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$. Действительно, выберем правую систему координат так, чтобы векторы $\mathbf{a}$ и $\mathbf{b}$ имели координаты $\{a, 0, 0\}$, $a > 0$, $\mathbf{b} = \{b_1, b_2, 0\}$, $b_2 > 0$. Тогда если смотреть с положительной стороны оси Oz, то кратчайший поворот от $\mathbf{a}$ к $\mathbf{b}$ осуществляется против часовой стрелки. С другой стороны, из условия $\begin{vmatrix} a & 0 & 0 \\ b_1 & b_2 & 0 \\ c_1 & c_2 & c_3 \end{vmatrix} > 0$ найдем: $c_3 > 0$. Это означает, что угол между вектором $\mathbf{c}$ и положительной полуосью оси Oz — острый, т. е. вектор $\mathbf{c}$ направлен в то же полупространство, что и эта полуось.

3. Преобразование координат точки. Пусть, как и прежде, $Ox_1x_2x_3$ — старая система координат, $\widetilde{O}\widetilde{x}_1\widetilde{x}_2\widetilde{x}_3$ — новая, $\widetilde{O}\,(o_1, o_2, o_3)$. Выразим «новые» координаты произвольной точки M через ее «старые»

координаты. Для этого заметим, что «новые» координаты точки M — это координаты вектора $\widetilde{\mathbf{O}}\mathbf{M}$ в новой системе координат. Найдем сначала координаты этого вектора в старой системе координат, а затем воспользуемся формулами (1) п. 1. Имеем:

$$\widetilde{\mathbf{O}}\mathbf{M} = \widetilde{\mathbf{O}}\mathbf{O} + \mathbf{O}\mathbf{M} = -\mathbf{O}\widetilde{\mathbf{O}} + \mathbf{O}\mathbf{M} = \{m_1 - o_1, m_2 - o_2, m_3 - o_3\}.$$

Следовательно,

$$\widetilde{m}_i = \sum_{s=1}^{3} a_{is}\,(m_s - o_s).$$

Таким образом, переход от старой системы координат к новой осуществляется при помощи параллельного переноса на вектор $O\widetilde{O}$ и трех поворотов на углы Эйлера вокруг точки $\widetilde{O}$.

§ 3. Векторное произведение

1. Определение векторного произведения.

Определение. *Векторным произведением векторов* $\mathbf{a}$ и $\mathbf{b}$ *называется вектор* $[\mathbf{ab}]$, *определяемый так*:

1° $[\mathbf{ab}]\perp\mathbf{a}$ и $[\mathbf{ab}]\perp\mathbf{b}$;

2° $|[\mathbf{ab}]| = |\mathbf{a}||\mathbf{b}|\sin\varphi$, *где* φ — *угол между векторами* $\mathbf{a}$ и $\mathbf{b}$;

3° *если* $[\mathbf{ab}] \neq \mathbf{o}$, *то тройка* $(\mathbf{a}, \mathbf{b}, [\mathbf{ab}])$ — *правая*.

Замечание. Из определения следует, что *модуль векторного произведения двух векторов равен площади параллелограмма, построенного на этих векторах*. Поэтому, в частности, *векторное произведение двух векторов равно нулевому вектору тогда и только тогда, когда эти векторы коллинеарны*.

Теорема. *Векторное произведение векторов* $\mathbf{a} = \{a_1, a_2, a_3\}$ и $\mathbf{b} = \{b_1, b_2, b_3\}$ *выражается формулой*:

$$[\mathbf{ab}] = \mathbf{p} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \left\{ \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right\} = \{\Delta_{23}, -\Delta_{13}, \Delta_{12}\}.$$

Доказательство. Как мы помним (теорема 1 п. 5 § 3 гл. 1), вектор $\mathbf{p}$ ортогонален к векторам $\mathbf{a}$ и $\mathbf{b}$, а его длина равна площади параллелограмма, построенного на этих векторах, т. е. равна произведению их длин на синус угла между ними. Тем самым, осталось доказать, что если $[\mathbf{ab}] \neq \mathbf{o}$, то тройка $(\mathbf{a}, \mathbf{b}, \mathbf{p})$ — правая, т. е. что $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{p}) > 0$. Имеем:

$$\Delta(\mathbf{a}, \mathbf{b}, \mathbf{p}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \Delta_{23} & -\Delta_{13} & \Delta_{12} \end{vmatrix}.$$

Разлагая этот определитель по последней строке, получаем: $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{p}) = \Delta_{23}^2 + \Delta_{13}^2 + \Delta_{12}^2 = |p|^2 = |\,[ab]\,|^2 > 0$. Теорема доказана.

Из доказанной теоремы и свойств определителей следует, что:

1° *для любых векторов* $\mathbf{a}$ и $\mathbf{b}$ *выполняется равенство* $[\mathbf{ab}] = -[\mathbf{ba}]$;

2° *для любых векторов* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *выполняется равенство* $[(\mathbf{a}(\mathbf{b} + \mathbf{c})] = [\mathbf{ab}] + [\mathbf{ac}]$;

3° *для любых векторов* $\mathbf{a}$ и $\mathbf{b}$ *и любого числа* λ *выполняется равенство* $[(\lambda\mathbf{a})\mathbf{b}] = \lambda[\mathbf{ab}]$.

2. Смешанное произведение.

Определение. *Смешанным произведением трех векторов называется скалярное произведение первого вектора на векторное произведение второго и третьего*: $(\mathbf{abc}) = (\mathbf{a}[\mathbf{bc}])$.

Сформулируем ряд свойств смешанного произведения.

1°. Из теоремы п. 1 следует, что

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

2°. Из 1° следует, что *смешанное произведение трех векторов равно объему построенного на них параллелепипеда, взятому со знаком «+», если они образуют правую тройку, и «−» — если левую.* В частности, *смешанное произведение трех векторов равно нулю тогда и только тогда, когда эти векторы компланарны.*

3°. Из 1° следует также, что *если в смешанном произведении поменять местами два сомножителя, то его модуль не изменится, а знак изменится на противоположный.*

4°. Из 3° следует, что $(\mathbf{abc}) = -(\mathbf{acb}) = (\mathbf{cab})$. Такая перестановка объектов, при которой последний из них ставится на первое место, а порядок следования остальных не меняется, называется *циклической перестановкой* этих объектов. Можно сказать, тем самым, что *смешанное произведение не меняется при циклической перестановке сомножителей.*

5°. Наконец, из 4° следует, что $(\mathbf{a}[\mathbf{bc}]) = (\mathbf{abc}) = (\mathbf{cab}) = (\mathbf{c}[\mathbf{ab}]) = ([\mathbf{ab}]\mathbf{c})$. Таким образом, *безразлично, какие именно сомножители в смешанном произведении — первый и второй или второй и третий — перемножаются векторно.*

3. Произведение двух смешанных произведений. Пусть $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\mathbf{e}$ и $\mathbf{f}$ — произвольные векторы. Имеем:

$$(\mathbf{abc})\,(\mathbf{def}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \cdot \begin{vmatrix} d_1 & e_1 & f_1 \\ d_2 & e_2 & f_2 \\ d_3 & e_3 & f_3 \end{vmatrix} = \begin{vmatrix} (\mathbf{ad}) & (\mathbf{ae}) & (\mathbf{af}) \\ (\mathbf{bd}) & (\mathbf{be}) & (\mathbf{bf}) \\ (\mathbf{cd}) & (\mathbf{ce}) & (\mathbf{cf}) \end{vmatrix}.$$

В частности,

$$(\mathbf{abc})^2 = \begin{vmatrix} |\mathbf{a}|^2 & (\mathbf{ab}) & (\mathbf{ac}) \\ (\mathbf{ab}) & |\mathbf{b}|^2 & (\mathbf{bc}) \\ (\mathbf{ac}) & (\mathbf{bc}) & |\mathbf{c}|^2 \end{vmatrix}.$$

Этот определитель называется *определителем Грама векторов* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$. Ясно, что *определитель Грама данных векторов неотрицателен и равен нулю тогда и только тогда, когда эти векторы линейно зависимы.*

4. Скалярное произведение двух векторных произведений.

Теорема. *Для любых векторов* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ и $\mathbf{d}$ *имеет место равенство*

$$([\mathbf{ab}]\,[\mathbf{cd}]) = \begin{vmatrix} (\mathbf{ac}) & (\mathbf{ad}) \\ (\mathbf{bc}) & (\mathbf{bd}) \end{vmatrix}.$$

Доказательство. Воспользуемся тем, что левая и правая части рассматриваемого равенства не зависят от выбора системы координат, и зададим ее так, чтобы векторы $\mathbf{a}$ и $\mathbf{b}$ имели координаты $\mathbf{a} = \{a, 0, 0\}$, $\mathbf{b} = \{b_1, b_2, 0\}$. Поскольку $[\mathbf{ab}] = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a & 0 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = ab_2\mathbf{e}_3$, то

$$([\mathbf{ab}][\mathbf{cd}]) = ab_2(\mathbf{e}_3\mathbf{cd}) = ab_2 \begin{vmatrix} 0 & 0 & 1 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{vmatrix} = ab_2 \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix}.$$

С другой стороны,

$$\begin{vmatrix} (\mathbf{ac}) & (\mathbf{ad}) \\ (\mathbf{bc}) & (\mathbf{bd}) \end{vmatrix} = \begin{vmatrix} ac_1 & ad_1 \\ b_1c_1 + b_2c_2 & b_1d_1 + b_2d_2 \end{vmatrix} = a\left(\begin{vmatrix} c_1 & d_1 \\ b_1c_1 & b_1d_1 \end{vmatrix} + \begin{vmatrix} c_1 & d_1 \\ b_2c_2 & b_2d_2 \end{vmatrix} \right) =$$

$$= 0 + ab_2 \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix}.$$

Теорема доказана.

5. Двойное векторное произведение.

Определение. *Двойным векторным произведением векторов* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *называется вектор* $[\mathbf{a}[\mathbf{bc}]]$.

Теорема. *Для любых векторов* $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ *имеет место равенство*

$$[\mathbf{a}[\mathbf{bc}]] = \mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab}).$$

Доказательство. Докажем, что i-е координаты векторов $[\mathbf{a}[\mathbf{bc}]]$ и $(\mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab}))$ совпадают при всех i. Воспользуемся свойством 5° п. 2. Имеем:

$$([\mathbf{a}[\mathbf{bc}]])_i = (\mathbf{e}_i[\mathbf{a}[\mathbf{bc}]]) = ([\mathbf{e}_i\mathbf{a}][\mathbf{bc}]) = \begin{vmatrix} (\mathbf{e}_i\mathbf{b}) & (\mathbf{e}_i\mathbf{c}) \\ (\mathbf{ab}) & (\mathbf{ac}) \end{vmatrix} \left(\mathbf{e}_i \begin{vmatrix} \mathbf{b} & \mathbf{c} \\ (\mathbf{ab}) & (\mathbf{ac}) \end{vmatrix} \right) =$$

$$= (\mathbf{e}_i(\mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab})) = (\mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab}))_i.$$

Таким образом, i-е координаты векторов $[\mathbf{a}[\mathbf{bc}]]$ и $\mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab})$ совпадают при всех i, а значит совпадают и сами эти векторы. Теорема доказана.

Глава 3
УРАВНЕНИЯ ПРЯМЫХ И ПЛОСКОСТЕЙ

§ 1. Прямая на плоскости

1. Уравнение прямой. Пусть на плоскости задана декартова система координат Oxy. Найдем уравнение прямой m, проходящей через две данные точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$. Точка $M(x, y)$ принадлежит прямой m тогда и только тогда, когда площадь параллелограмма, построенного на векторах $\mathbf{M}_1\mathbf{M}$ и $\mathbf{M}_1\mathbf{M}_2$, равна нулю:

$$\begin{vmatrix} (x - x_1) & (y - y_1) \\ (x_2 - x_1) & (y_2 - y_1) \end{vmatrix} = 0. \tag{1}$$

Это и есть *уравнение прямой, проходящей через две данные точки*.

В частности, уравнение прямой, пересекающей оси координат в точках $M_1(x_0,0)$ и $M_2(0,y_0)$, имеет вид

$$\begin{vmatrix} (x - x_0) & y \\ -x_0 & y_0 \end{vmatrix} = xy_0 - x_0y_0 + yx_0 = 0,$$

или, если произведение x_0y_0 отлично от нуля,

$$\frac{x}{x_0} + \frac{y}{y_0} = 1.$$

Это уравнение называется *уравнением прямой в отрезках*.

2. Параметрические уравнения прямой. Если вторую строку определителя в уравнении (1) умножить на произвольное число $\lambda \neq 0$, то получится уравнение, эквивалентное исходному. Иными словами, вектор $\mathbf{M}_1\mathbf{M}_2$ можно заменить любым ненулевым вектором $\mathbf{p} = \{a, b\}$, коллинеарным вектору $\mathbf{M}_1\mathbf{M}_2$, т. е. параллельным прямой m. Это означает, что уравнение

$$\begin{vmatrix} (x - x_1) & (y - y_1) \\ a & b \end{vmatrix} = 0 \tag{2}$$

представляет собой *уравнение прямой, проходящей через точку* M_1 *и параллельной вектору* $\mathbf{p} = \{a, b\}$. Из условия $\mathbf{p} \neq \mathbf{o}$ следует, что равенство (2) выполняется тогда и только тогда, когда существует такое число t, что $\mathbf{M}_1\mathbf{M} = \mathbf{p}t$, или $\mathbf{OM} = \mathbf{OM}_1 + \mathbf{p}t$, т. е.

$$\left.\begin{aligned} x &= x_1 + at \\ y &= y_1 + bt \end{aligned}\right\}.$$

Полученные уравнения называются *параметрическими уравнения прямой*. Параметр t в них меняется от $-\infty$ до ∞. *Если* интерпретировать t как время, а вектор $\mathbf{p} = \{a, b\}$ — как скорость, то полученные уравнения можно рассматривать как закон прямолинейного и равномерного движения материальной точки.

3. Каноническое уравнение прямой. Перепишем уравнение (2) в виде

$$b(x - x_1) - a(y - y_1) = 0. \tag{3}$$

Если произведение ab отлично от нуля, то это уравнение можно записать так:

$$\frac{x - x_1}{a} = \frac{y - y_1}{b}. \tag{4}$$

Из условия $\mathbf{p} = \{a, b\} \neq \mathbf{o}$ следует, что числа a и b не могут обращаться в нуль одновременно. Поэтому, если, например, $a = 0$, то из уравнения (3) находим: $x - x_1 = 0$, а y — любое число. Примем такое соглашение: *если один из знаменателей в уравнении* (4) *обращается в нуль, то соответствующий числитель также обращается в нуль*. Тогда уравнение (4) станет эквивалентным уравнению (3).

Уравнение (4) вместе с принятым соглашением называется *каноническим уравнением прямой*.

4. Общее уравнение прямой. Перепишем уравнение (3) так:

$$A(x - x_1) + B(y - y_1) = 0, \tag{5}$$

где $A = b$, $B = -a$, и рассмотрим вектор $\mathbf{n} = \{A, B\}$. Ясно, что $|\mathbf{n}| = |\mathbf{p}| \neq 0$ и $\mathbf{n} \perp \mathbf{p}$. Таким образом, уравнение (5) при условии $A^2 + B^2 \neq 0$ является *уравнением прямой, проходящей через точку* $M_1(x_1, y_1)$ *и перпендикулярной к вектору* $\mathbf{n} = \{A, B\}$.

Полагая в уравнении (5) $C = -Ax_1 - By_1$, получаем:

$$Ax + By + C = 0. \tag{6}$$

Итак, уравнение любой прямой может быть записано в виде (6), где $A^2 + B^2 \neq 0$. Верно и обратное утверждение: уравнение (6) при условии $A^2 + B^2 \neq 0$ является уравнением некоторой прямой. В самом деле, из условия $A^2 + B^2 \neq 0$ следует, что $\operatorname{rang}(AB) = \operatorname{rang}(AB - C) = 1$, поэтому уравнение (6) имеет по крайней мере одно решение: $Ax_1 + By_1 + C = 0$. Следовательно, уравнение (6) эквивалентно уравнению (5), т. е. представляет собой уравнение прямой, перпендикулярной к вектору $\mathbf{n} = \{A, B\}$.

Уравнение (6) при условии $A^2 + B^2 \neq 0$ называется *общим уравнением прямой*.

5. Нормированное уравнение прямой. Без ограничения общности будем считать, что число C в уравнении (6) неположительно (в противном случае обе части уравнения можно умножить на -1). Разделив

обе части этого уравнения на $\sqrt{A^2+B^2}$ и положив $\dfrac{A}{\sqrt{A^2+B^2}} = \cos\theta$, $\dfrac{B}{\sqrt{A^2+B^2}} = \sin\theta$, $\dfrac{-C}{\sqrt{A^2+B^2}} = d \geqslant 0$, получим:

$$x\cos\theta + y\sin\theta = d.$$

Уравнение (7) при условии $d \geqslant 0$ называется *нормированным уравнением прямой*. В нем $\mathbf{n} = \{\cos\theta, \sin\theta\}$ — единичный вектор, перпендикулярный к прямой m, θ — угол, который он образует с положительной полуосью оси Ox, отсчитываемый против часовой стрелки.

Заметим, что уравнение (7) можно записать также в виде

$$(\mathbf{OMn}) = d,$$

где O — начало координат, а M — точка данной прямой m. Следовательно, d — это координата точки M по оси с началом O и координатным вектором $\mathbf{n}$. Прямая m перпендикулярна к этой оси, поэтому все ее точки имеют по ней одну и ту же координату. Поскольку $d \geqslant 0$, то: а) d — это расстояние от прямой m до начала координат; б) вектор $\mathbf{n}$ направлен от начала координат к прямой.

Итак, в нормированном уравнении прямой:

1° $\mathbf{n} = \{\cos\theta, \sin\theta\}$ — единичный вектор, перпендикулярный к прямой и направленный от начала координат к этой прямой;

2° θ — угол, который вектор $\mathbf{n}$ образует с положительной полуосью оси Ox, отсчитываемый против часовой стрелки;

3° d — расстояние от прямой до начала координат.

Замечание. Пусть $M_1(x_1, y_1)$ — произвольная точка плоскости. Величина

$$d_1 = x_1\cos\theta + y_1\sin\theta = (\mathbf{OM_1n})$$

представляет собой координату точки M_1 по оси с началом O и координатным вектором $\mathbf{n}$. Поэтому величина $d_1 - d$, называемая *отклонением точки M_1 от прямой m*, равна расстоянию от точки M_1 до прямой m, взятому со знаком «$-$», если точки O и M_1 лежат по одну сторону от прямой m, и «$+$» — если по разные. Таким образом, нормированным уравнением прямой удобно пользоваться в тех случаях, когда требуется найти расстояние от точки до прямой или определить, по какую сторону от прямой находится эта точка.

§ 2. Плоскость

1. Уравнение плоскости. Пусть в пространстве задана декартова система координат $Oxyz$. Найдем уравнение плоскости, проходящей через три данные точки $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, и $M_3(x_3, y_3, z_3)$, не лежащие на одной прямой. Точка $M(x, y, z)$ принадлежит этой плоскости тогда и только тогда, когда смешанное произведение векторов $\mathbf{M_1M}$, $\mathbf{M_1M_2}$ и $\mathbf{M_1M_3}$ равно нулю:

$$(\mathbf{M_1M} \cdot \mathbf{M_1M_2} \cdot \mathbf{M_1M_3}) = 0. \qquad (1)$$

Это и есть *уравнение плоскости, проходящей через три данные точки*. В координатах оно выглядит так:

$$\begin{vmatrix} (x-x_1) & (y-y_1) & (z-z_1) \\ (x_2-x_1) & (y_2-y_1) & (z_2-z_1) \\ (x_3-x_1) & (y_3-y_1) & (z_3-z_1) \end{vmatrix} = 0.$$

В частности, уравнение плоскости, пересекающей оси координат в точках $M_1(x_0, 0, 0)$, $M_2(0, y_0, 0)$ и $M_3(0, 0, z_0)$, имеет вид

$$\begin{vmatrix} (x-x_0) & y & z \\ -x_0 & y_0 & 0 \\ -x_0 & 0 & z_0 \end{vmatrix} = xy_0z_0 - x_0y_0z_0 + zx_0y_0 + yz_0x_0 = 0,$$

или, если произведение $x_0y_0z_0$ отлично от нуля,

$$\frac{x}{x_0} + \frac{y}{y_0} + \frac{z}{z_0} = 1.$$

Это уравнение называется *уравнением плоскости в отрезках*.

2. Общее уравнение плоскости. Перепишем уравнение (1) так:

$$(\mathbf{M}_1\mathbf{M}[\mathbf{M}_1\mathbf{M}_2 \cdot \mathbf{M}_1\mathbf{M}_3]) = 0.$$

Если в этом уравнении умножить вектор $[\mathbf{M}_1\mathbf{M}_2 \cdot \mathbf{M}_1\mathbf{M}_3]$, который, очевидно, перпендикулярен к плоскости $M_1M_2M_3$, на произвольное число $\lambda \neq 0$, то получится уравнение, эквивалентное исходному. Иными словами, вектор $[\mathbf{M}_1\mathbf{M}_2 \cdot \mathbf{M}_1\mathbf{M}_3]$ можно заменить любым ненулевым вектором $\mathbf{n} = \{A, B, C\}$, коллинеарным вектору $[\mathbf{M}_1\mathbf{M}_2 \cdot \mathbf{M}_1\mathbf{M}_3]$, т. е. перпендикулярным к плоскости $M_1M_2M_3$. Это означает, что уравнение

$$A(x-x_1) + B(y-y_1) + C(z-z_1) = 0 \tag{2}$$

при условии $A^2+B^2+C^2 \neq 0$ представляет собой *уравнение плоскости, проходящей через точку* $M_1(x_1, y_1, z_1)$ *и перпендикулярной к вектору* $\mathbf{n} = \{A, B, C\}$.

Полагая в уравнении (2) $D = -Ax_1 - By_1 - Cz_1$, получаем:

$$Ax + By + Cz + D = 0. \tag{3}$$

Итак, уравнение любой плоскости может быть записано в виде (3), где $A^2 + B^2 + C^2 \neq 0$. Верно и обратное утверждение: уравнение (3) при условии $A^2 + B^2 + C^2 \neq 0$ является уравнением некоторой плоскости. В самом деле, из условия $A^2 + B^2 + C^2 \neq 0$ следует, что $\operatorname{rang}(A\ B\ C) = \operatorname{rang}(A\ B\ C\ \ -D) = 1$, поэтому уравнение (3) имеет по крайней мере одно решение: $Ax_1 + By_1 + Cz_1 + D = 0$. Следовательно, уравнение (3) эквивалентно уравнению (2), т. е. представляет собой уравнение плоскости, перпендикулярной к вектору $\mathbf{n} = \{A, B, C\}$.

Уравнение (3) при условии $A^2 + B^2 + C^2 \neq 0$ называется *общим уравнением плоскости*.

3. Нормированное уравнение плоскости. Без ограничения общности будем считать, что число D в уравнении (3) неположительно (в противном случае обе части уравнения можно умножить на -1). Разделив обе части этого уравнения на $\sqrt{A^2+B^2+C^2}$ и положив $A\big/\sqrt{A^2+B^2+C^2}=\cos\alpha$, $B\big/\sqrt{A^2+B^2+C^2}=\cos\beta$, $C\big/\sqrt{A^2+B^2+C^2}=$ $=\cos\gamma$, $-D\big/\sqrt{A^2+B^2+C^2}=d\geqslant 0$, получим:

$$x\cos\alpha+y\cos\beta+z\cos\gamma=d. \tag{4}$$

Уравнение (4) при условии $d\geqslant 0$ называется *нормированным уравнением плоскости*. В нем $\mathbf{n}=\{\cos\alpha,\cos\beta,\cos\gamma\}$ — единичный вектор, перпендикулярный к плоскости, α, β, γ — углы, которые он образует с положительными полуосями осей координат.

Заметим, что уравнение (4) можно записать также в виде

$$(\mathbf{OMn})=d,$$

где O — начало координат, а M — точка данной плоскости. Следовательно, d — это координата точки M по оси с началом O и координатным вектором $\mathbf{n}$. Рассматриваемая плоскость перпендикулярна к этой оси, поэтому все ее точки имеют по ней одну и ту же координату. Поскольку $d\geqslant 0$, то: а) d — это расстояние от данной плоскости до начала координат; б) вектор $\mathbf{n}$ направлен от начала координат к плоскости.

Итак, в нормированном уравнении плоскости:

1° $\mathbf{n}=\{\cos\alpha,\cos\beta,\cos\gamma\}$ — единичный вектор, перпендикулярный к плоскости и направленный от начала координат к этой плоскости;

2° α, β, γ — углы, которые вектор $\mathbf{n}$ образует с положительными полуосями осей координат;

3° d — расстояние от данной плоскости до начала координат.

Замечание. Пусть $M_1(x_1,y_1,z_1)$ — произвольная точка пространства. Величина

$$d_1=x_1\cos\alpha+y_1\cos\beta+z_1\cos\gamma=(\mathbf{OM}_1\cdot\mathbf{n})$$

представляет собой координату точки M_1 по оси с началом O и координатным вектором $\mathbf{n}$. Поэтому величина d_1-d, называемая *отклонением точки M_1 от плоскости*, равна расстоянию от точки M_1 до данной плоскости, взятому со знаком «$-$», если точки O и M_1 лежат по одну сторону от плоскости, и «$+$» — если по разные.

§ 3. Прямая в пространстве

1. Уравнение прямой. Пусть в пространстве задана декартова система координат $Oxyz$. Найдем уравнение прямой m, проходящей через две данные точки $M_1(x_1,y_1,z_1)$ и $M_2(x_2,y_2,z_2)$. Точка $M(x,y,z)$ принадлежит прямой m тогда и только тогда, когда векторное произведение векторов $\mathbf{M}_1\mathbf{M}$ и $\mathbf{M}_1\mathbf{M}_2$ равно нулю:

$$[\mathbf{M}_1\mathbf{M}\cdot\mathbf{M}_1\mathbf{M}_2]=\mathbf{o}. \tag{1}$$

Это и есть *уравнение прямой, проходящей через две данные точки*. В координатах оно выглядит так:

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ (x-x_1) & (y-y_1) & (z-z_1) \\ (x_2-x_1) & (y_2-y_1) & (z_2-z_1) \end{vmatrix} = 0.$$

2. Параметрические уравнения прямой. Если вектор $\mathbf{M}_1\mathbf{M}_2$ в уравнении (1) умножить на произвольное число $\lambda \neq 0$, то получится уравнение, эквивалентное исходному. Иными словами, вектор $\mathbf{M}_1\mathbf{M}_2$ можно заменить любым ненулевым вектором $\mathbf{p} = \{a, b, c\}$, коллинеарным вектору $\mathbf{M}_1\mathbf{M}_2$, т. е. параллельным прямой m:

$$[\mathbf{M}_1\mathbf{M} \cdot \mathbf{p}] = \mathbf{o}.$$

Полученное уравнение представляет собой *уравнение прямой, проходящей через точку M_1 и параллельной вектору* $\mathbf{p} = \{a, b, c\}$. Из условия $\mathbf{p} \neq 0$ следует, что это равенство выполняется тогда и только тогда, когда существует такое число t, что $\mathbf{M}_1\mathbf{M} = \mathbf{p}t$, или $\mathbf{OM} = \mathbf{OM}_1 + \mathbf{p}t$, т. е.

$$\left.\begin{aligned} x &= x_1 + at \\ y &= y_1 + bt \\ z &= z_1 + ct \end{aligned}\right\}. \qquad (2)$$

Эти уравнения называются *параметрическими уравнениями прямой*. Параметр t в них меняется от $-\infty$ до ∞. Если интерпретировать t как время, а вектор $\mathbf{p} = \{a, b, c\}$ — как скорость, то полученные уравнения можно рассматривать как закон прямолинейного и равномерного движения материальной точки.

3. Канонические уравнения прямой. Если произведение abc отлично от нуля, то параметр t из уравнений (2) можно исключить, переписав их так:

$$\frac{x-x_1}{a} = \frac{y-y_1}{b} = \frac{z-z_1}{c}. \qquad (3)$$

Из условия $\mathbf{p} = \{a, b, c\} \neq \mathbf{o}$ следует, что числа a, b и c не могут обращаться в нуль одновременно. Поэтому если, например, $a = 0$ (и $b = 0$), то из уравнений (2) находим: $x - x_1 = 0$ (и $y - y_1 = 0$, а z — любое число). Примем такое соглашение: *если один или два из знаменателей в уравнениях* (3) *обращаются в нуль, то соответствующие числители также обращаются в нуль*. Тогда уравнения (3) станут эквивалентными уравнениям (2).

Уравнения (3) вместе с принятым соглашением называются *каноническими уравнениями прямой*.

4. Пересечение двух плоскостей. Рассмотрим две плоскости, заданные уравнениями $A_1x+B_1y+C_1z+D_1 = 0$ и $A_2x+B_2y+C_2z+D_2 =$

$= 0$. Чтобы найти все их общие точки, необходимо решить систему уравнений

$$\left.\begin{aligned} A_1x + B_1y + C_1z = -D_1 \\ A_2x + B_2y + C_2z = -D_2 \end{aligned}\right\}.$$

Поскольку $A_i^2 + B_i^2 + C_i^2 \neq 0$, то возможны три случая.

1°. $\operatorname{rang}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = \operatorname{rang}\begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \end{pmatrix} = 1$. В этом случае одно из уравнений является следствием другого, а значит, данные плоскости совпадают.

2°. $\operatorname{rang}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = 1$, $\operatorname{rang}\begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \end{pmatrix} = 2$. По теореме Кронекера–Капелли в этом случае решений нет и, следовательно, данные плоскости параллельны.

3°. $\operatorname{rang}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = \operatorname{rang}\begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \end{pmatrix} = 2$. Чтобы найти решение, необходимо найти базисный минор, общий для основной и расширенной матрицы. Пусть, например, таким минором является минор $\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}$. Перепишем данную систему так:

$$\left.\begin{aligned} A_1x + B_1y = -D_1 - C_1z \\ A_2x + B_2y = -D_2 - C_2z \end{aligned}\right\}.$$

Теперь, полагая z равным произвольному числу $\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} t$, найдем x и y по формулам Крамера:

$$\left.\begin{aligned} x &= -\begin{vmatrix} D_1 & B_1 \\ D_2 & B_2 \end{vmatrix} \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}^{-1} - \begin{vmatrix} C_1 & B_1 \\ C_2 & B_2 \end{vmatrix} t \\ y &= -\begin{vmatrix} A_1 & D_1 \\ A_2 & D_2 \end{vmatrix} \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}^{-1} - \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix} t \\ z &= \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} t \end{aligned}\right\}.$$

Таким образом, пересечением данных плоскостей является прямая

$$\left.\begin{aligned} x &= x_1 + at \\ y &= y_1 + bt \\ z &= ct \end{aligned}\right\},$$

параллельная вектору

$$\{a, b, c\} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix},$$

что и понятно — линия пересечения данных плоскостей должна быть перпендикулярной к каждому из перпендикуляров к этим плоскостям.

Глава 4

ЛИНИИ И ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА

§ 1. Эллипс, гипербола и парабола

1. Эллипс. Эллипс, по-видимому, был известен еще в глубокой древности, когда облик геометрии соответствовал дословному переводу ее названия. В те времена основными инструментами для выполнения построений на местности были колья и веревки, позволявшие проводить прямые и окружности, а значит, и выполнять все те построения, которые теперь называют построениями с помощью циркуля и линейки.

Ясно, как с помощью указанных инструментов построить окружность: нужно закрепить один из концов веревки и в натянутом состоянии прочертить вторым концом линию. Напрашивается вопрос: а что получится, если закрепить оба конца ненатянутой веревки, а затем в натянутом состоянии прочертить линию? Получится эллипс. Таким образом, мы приходим к следующему определению.

Определение. *Эллипсом называется множество всех таких точек плоскости, для которых сумма расстояний до двух фиксированных точек постоянна.*

Фиксированные точки называются *фокусами* эллипса.

Пусть $2c$ — расстояние между фокусами, $2a$ — сумма расстояний от точки эллипса до фокусов. Введем декартову систему координат Oxy так, чтобы фокусы F_1 и F_2 имели координаты $F_1(-c,0)$ и $F_2(c,0)$, и выведем в ней уравнение эллипса. Стоящую перед нами задачу можно сформулировать так: найти множество всех таких точек $M(x,y)$, для которых

$$MF_1 + MF_2 = 2a.$$

Из неравенства треугольника следует, что $MF_1 + MF_2 \geqslant F_1F_2$, т. е. $a \geqslant c$. При $a = c$ эллипс вырождается в отрезок F_1F_2, поэтому будем считать, что $a > c$.

В координатах уравнение эллипса принимает вид

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

Умножим обе части этого равенства на разность фигурирующих в нем радикалов, а затем разделим на $2a$. В результате получим:

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \frac{1}{2a}\left((x+c)^2 - (x-c)^2\right) = \frac{2cx}{a}.$$

Теперь можно выразить каждый из радикалов:

$$\sqrt{(x \pm c)^2 + y^2} = a \pm \frac{cx}{a}.$$

Возведем это выражение в квадрат и преобразуем его к виду

$$\frac{a^2 - c^2}{a^2} x^2 + y^2 = a^2 - c^2.$$

Поскольку $a \neq c$, то полученное равенство можно переписать так:

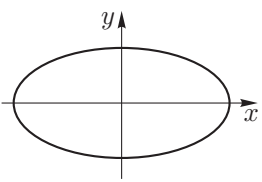
$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1, \tag{1}$$

или, с учетом условия $a > c$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{2}$$

где $b^2 = a^2 - c^2 \leqslant a^2$.

При возведении в квадрат могли появиться лишние корни, соответствующие случаю $a \pm cx/a < 0$. Но этого не происходит: как видно из уравнения (2), $|x| \leqslant a$, поэтому $|cx/a| \leqslant c < a$. Таким образом, уравнение (2) эквивалентно исходному. Оно называется *каноническим уравнением эллипса*.



Уравнение (2) позволяет обнаружить следующие свойства эллипса.

1°. Эллипс имеет две взаимно перпендикулярные оси симметрии (оси Ox и Oy), а значит, и центр симметрии (начало координат O). Оси симметрии эллипса называются его *полуосями*; та из них, на которой лежат фокусы, называется *большой* полуосью, а другая — *малой*; числа a и b иногда также называют *полуосями*.

2°. Поскольку $x^2/a^2 = 1 - y^2/b^2 \leqslant 1$ и $y^2/b^2 = 1 - x^2/a^2 \leqslant 1$, то эллипс целиком содержится в прямоугольнике ($|x| \leqslant a$, $|y| \leqslant b$), стороны которого параллельны его осям.

3°. При $x \geqslant 0$, $y \geqslant 0$ уравнение (2) может быть записано в виде

$$y = b\sqrt{1 - \frac{x^2}{a^2}}.$$

Функция $y(x)$ монотонно убывает от $y = b$ при $x = 0$ до $y = 0$ при $x = a \geqslant b$. С учетом установленных нами симметрий, это позволяет изобразить эллипс на рисунке.

4°. Уравнение (2) может быть записано в параметрическом виде

$$\left.\begin{aligned} x &= a\cos\varphi \\ y &= b\sin\varphi \end{aligned}\right\},$$

где $0 \leqslant \varphi < 2\pi$.

5°. Как из уравнения (2), так и из параметрических уравнений легко усматривается, что эллипс может быть получен из окружности (радиуса b) равномерным растяжением (в a/b раз) вдоль прямой (оси Ox, т. е. преобразованием $x \to (b/a)x$).

2. Гипербола. Возникает естественный вопрос: что получится, если в определении эллипса сумму расстояний заменить их разностью? Получится гипербола. Таким образом, мы приходим к следующему определению.

Определение. *Гиперболой называется множество всех таких точек плоскости, для которых модуль разности расстояний до двух фиксированных точек есть постоянная положительная величина.*

Фиксированные точки называются *фокусами гиперболы*.

Пусть $2c$ — расстояние между фокусами, $2a$ — модуль разности расстояний от точки гиперболы до фокусов. Введем декартову систему координат Oxy так, чтобы фокусы F_1 и F_2 имели координаты $F_1(-c, 0)$ и $F_2(c, 0)$, и выведем в ней уравнение гиперболы. Стоящую перед нами задачу можно сформулировать так: найти множество всех таких точек $M(x, y)$, для которых

$$MF_1 - MF_2 = \pm 2a.$$

Из неравенства треугольника следует, что $|MF_1 - MF_2| \leqslant F_1F_2$, т. е. $a \leqslant c$. При $a = c$ гипербола вырождается в два луча прямой F_1F_2, поэтому будем считать, что $a < c$. В координатах уравнение гиперболы принимает вид

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \pm 2a.$$

Умножим обе части этого равенства на сумму фигурирующих в нем радикалов, а затем разделим на $\pm 2a$. В результате получим:

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = \pm \frac{2cx}{a}.$$

Теперь можно выразить каждый из радикалов:

$$\sqrt{(x \pm c)^2 + y^2} = \left| a \pm \frac{cx}{a} \right|.$$

Возведем это выражение в квадрат (при этом лишних корней, очевидно, не появится) и преобразуем его к виду

$$\frac{a^2 - c^2}{a^2} x^2 + y^2 = a^2 - c^2.$$

Поскольку $a \neq c$, то полученное равенство можно переписать так:

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1.$$

Это и есть искомое уравнение гиперболы. Сравнивая полученный результат с уравнением (1), мы приходим к весьма неожиданному выводу: *гипербола имеет точно такое же уравнение, как и эллипс*!

Различие состоит только в том, что для эллипса $a > c$, а для гиперболы $a < c$. С учетом этого условия уравнение гиперболы можно переписать так:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \tag{3}$$

где $b^2 = c^2 - a^2$.

Уравнение (3) называется *каноническим уравнением гиперболы*. Оно позволяет обнаружить следующие свойства гиперболы.

1°. Гипербола имеет две взаимно перпендикулярные оси симметрии (оси Ox и Oy), а значит, и центр симметрии (начало координат O). Оси симметрии гиперболы называются ее *полуосями*; та из них, на которой лежат фокусы, называется *вещественной* полуосью, а другая — *мнимой*; числа a и b иногда также называют *полуосями*.

2°. Поскольку $\dfrac{x^2}{a^2} = 1 + \dfrac{y^2}{b^2} \geqslant 1$, то в полосе ($|x|<a$), содержащей мнимую ось гиперболы, точек гиперболы нет.

3°. Поскольку $\dfrac{y^2}{b^2} = \dfrac{x^2}{a^2} - 1 < \dfrac{x^2}{a^2}$, то в области между двумя пересекающимися прямыми ($|y| \geqslant \dfrac{b}{a}|x|$) точек гиперболы также нет.

4°. При $x \geqslant 0$, $y \geqslant 0$ уравнение (3) может быть записано в виде

$$y = b\sqrt{\frac{x^2}{a^2} - 1}.$$

Функция $y(x)$ монотонно и неограниченно возрастает от $y = 0$ при $x = a$. С учетом установленных нами симметрий, это позволяет изобразить гиперболу на рисунке. Однако прежде чем это сделать, установим еще одно ее свойство.

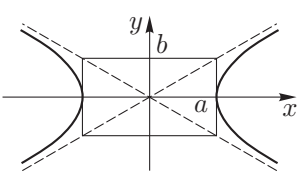
5°. Ясно, что при $x \to \infty$ функция $y(x)$ становится приближенно равной $(b/a)x$. Уточним это свойство. Имеем:

$$0 < \frac{b}{a}x - \frac{b}{a}\sqrt{x^2 - a^2} = \frac{b}{a} \cdot \frac{x^2 - x^2 + a^2}{x + \sqrt{x^2 - a^2}} = \frac{ab}{x + \sqrt{x^2 - a^2}} \leqslant \frac{ab}{x} \to 0$$

Таким образом, гипербола имеет асимптоту ($y = (b/a)x$).

Теперь можно изобразить гиперболу на рисунке.

Мы видим, в частности, что гипербола имеет две ветви.



6°. Уравнение (3) может быть записано в параметрическом виде

$$\left.\begin{aligned} x &= \pm a \operatorname{ch} t \\ y &= b \operatorname{sh} t \end{aligned}\right\},$$

где $\operatorname{ch} t = (e^t + e^{-t})/2$ — гиперболические косинус, а $\operatorname{sh} t = (e^t - e^{-t})/2$ — гиперболические синус. При этом знак «+» соответствует одной ветви гиперболы, а знак «−» — другой.

3. Директриса эллипса и гиперболы. Вернемся к уравнению (1), которое, как отмечалось, является общим уравнением эллипса и гиперболы. Умножая это уравнение на $a^2 - c^2$, перенося слагаемое $(c^2/a^2)x^2$ в правую часть, $-c^2$ — в левую часть и вычитая из обеих частей $2xc$, получаем (при $c \neq 0$):

$$x^2 + c^2 - 2xc + y^2 = a^2 + \frac{c^2}{a^2}x^2 - 2xc = \frac{c^2}{a^2}\left(x^2 - 2x\frac{a^2}{c} + \frac{a^4}{c^2}\right),$$

откуда

$$\sqrt{(x-c)^2 + y^2} = \frac{c}{a}\left|x - \frac{a^2}{c}\right|. \tag{4}$$

Если предположить, что правая, а значит и левая, части равенства (4) обращаются в нуль, то получится противоречие: $x = c$, $y = 0$, $c = a^2/c$, т. е. $c = a$, что не выполняется ни для эллипса, ни для гиперболы. Поэтому равенство (4) можно переписать так:

$$\frac{\sqrt{(x-c)^2 + y^2}}{\left|x - a^2/c\right|} = \frac{c}{a}. \tag{5}$$

В этом уравнении числитель представляет собой расстояние от точки $M(x, y)$ до фокуса F_2, а знаменатель — расстояние от нее до прямой d_2, определяемой уравнением $x = a^2/c$. Поскольку уравнение (5) при $c \neq 0$ эквивалентно уравнению (1), то мы приходим к следующему выводу: *эллипс*, *отличный от окружности*, (*гипербола*) *является множеством всех таких точек плоскости*, *для которых отношение расстояния до фиксированной точки к расстоянию до фиксированной прямой постоянно и меньше* (*больше*) *единицы*.

Фиксированную точку, как и прежде, будем называть фокусом, фиксированную прямую — *директрисой*, а указанное отношение — *эксцентриситетом*.

Замечание 1. Как отмечалось выше, правая часть равенства (4) не может обращаться в нуль. Поэтому *директриса эллипса* (*гиперболы*) *не имеет с ним* (*с ней*) *общих точек*.

Замечание 2. Поскольку эллипс и гипербола симметричны относительно оси Oy, то каждая из этих кривых имеет *две директрисы* d_1 и d_2, соответствующие фокусам F_1 и F_2.

Замечание 3. Ясно, что каждый эллипс (каждая гипербола) однозначно определяется заданием двух чисел: экцентриситета e и расстояния p между фокусом и директрисой. Введем новую систему координат Oxy (для удобства мы сохраняем прежние обозначения для координат) с началом в фокусе F_2. Тогда соответствующая ему директриса будет иметь уравнение $x = -p$. Точка $M(x, y)$ принадлежит нашей кривой тогда и только тогда, когда

$$\sqrt{x^2 + y^2} = e|x + p|,$$

или

$$x^2 + y^2 = e^2 (x + p)^2 . \qquad (6)$$

Тем самым, мы получили еще одно уравнение, общее для эллипса и гиперболы.

Замечание 4. Поскольку экцентриситет e равен c/a, то $b = a\sqrt{1 - e^2}$ для эллипса и $b = a\sqrt{e^2 - 1}$ для гиперболы. Поэтому *любые два эллипса* (*две гиперболы*) *с одинаковыми эксцентриситетами подобны*. Тем самым можно сказать, что число e определяет форму кривой, а число p — ее размеры.

Для эллипса с одинаковыми полуосями, т. е. для окружности, $e = 0$. При увеличении числа e эллипс становится все более «вытянутым». Этим и объясняется название «эксцентриситет».

4. Парабола. Сам собой напрашивается вопрос: какая кривая соответствует эксцентриситету, равному 1? Эта кривая называется параболой. Таким образом, мы приходим к следующему определению.

Определение. *Параболой называется множеством всех таких точек плоскости, для которых расстояние до фиксированной точки равно расстоянию до фиксированной прямой.*

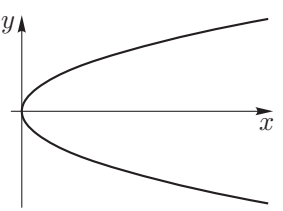
Как и прежде, фиксированную точку мы будем называть *фокусом*, фиксированную прямую — *директрисой параболы*.

Полагая в уравнении (6) $e = 1$, получаем уравнение параболы:

$$y^2 = 2px + p = 2p\left(x + \frac{p}{2}\right).$$

Перенося начало координат в точку $x = -p/2$ и возвращаясь к старым обозначениям для координат, получаем:

$$y^2 = 2px.$$



Это уравнение называется *каноническим уравнением параболы*. Оно позволяет обнаружить следующие свойства параболы.

1°. Парабола имеет одну ось симметрии (ось Ox). Эта ось называется *ось параболы*.

2°. Поскольку $x = y^2/(2p) \geqslant 0$, то парабола целиком содержится в полуплоскости ($x \geqslant 0$), граница которой перпендикулярна оси параболы.

3°. При $y \geqslant 0$ полученное уравнение может быть записано в виде

$$y = \sqrt{2px}.$$

Функция $y(x)$ монотонно и неограниченно возрастает от $y = 0$ при $x = 0$. С учетом симметрии, это позволяет изобразить параболу на рисунке.

Замечание. Ясно, что уравнение (6) является общим для всех трех кривых — эллипса, гиперболы и параболы. Иногда его записывают в полярных координатах.

Полагая в уравнении (6) $x = \rho\cos\varphi$, $y = \rho\sin\varphi$ и извлекая корень из обеих частей, получаем:

$$\rho = \pm e(\rho cos\varphi + p),$$

откуда

$$\rho(1 \mp e\cos\varphi) = \pm ep,$$

$$\left.\begin{aligned} \rho &= \frac{ep}{1 - e\cos\varphi} \\ \rho &= \frac{-ep}{1 + e\cos\varphi} \end{aligned}\right\}.$$

Поскольку $\rho \geqslant 0$, то первое из уравнений этой пары является уравнением эллипса, параболы и одной из ветвей гиперболы («правой»), а второе — уравнением другой ветви гиперболы («левой»).

5. Касательная. Рассмотрим произвольную линию L и возьмем на ней какие-нибудь точки M_0 и M_1. Прямую M_0M_1 естественно назвать *секущей*. Представим себе, что точка M_1, двигаясь по линии L, приближается к M_0. Предельное положение секущей M_0M_1 при стремлении точки M_1 к M_0 называется *касательной* к линии L в точке M_0.

Пусть теперь L — одна из кривых: эллипс, гипербола или парабола, уравнение которой записано в виде (6). Рассмотрим прямую

$$\left.\begin{aligned} x &= x_0 + t\cos\theta \\ y &= y_0 + t\sin\theta \end{aligned}\right\}$$

($0 \leqslant \theta < \pi$), проходящую через произвольную точку $M_0(x_0, y_0)$ кривой L, и выясним, есть ли у этой прямой и кривой L другие общие точки. Для этого подставим x, y из уравнений прямой в уравнение (6) и установим, будет ли полученное уравнение иметь решения, отличные от $t = 0$ (соответствующего точке M_0). Имеем:

$$x_0^2 + y_0^2 + 2t\,(x_0\cos\theta + y_0\sin\theta) + t^2 =$$

$$= e^2\,(x_0 + p)^2 + 2te^2\,(x_0 + p)\cos\theta + t^2e^2\cos^2\theta.$$

Учитывая, что $M_0(x_0, y_0)$ — точка кривой L, и сокращая на t, получаем уравнение вида

$$t(1 - e^2\cos^2\theta) = A\cos\theta + B\sin\theta, \tag{7}$$

где A и B — некоторые числа, зависящие от x_0, y_0, p и e.

Замечание 1. Отметим, что в этом уравнении *коэффициент при t и правая часть не могут обращаться в нуль одновременно* — иначе равенство (7) было бы справедливо при всех значениях t и, следовательно, прямая целиком принадлежала бы кривой L, чего не может быть ни для эллипса, ни для гиперболы, ни для параболы.

Замечание 2. Обратим особое внимание на то, что числа A и B в формуле (7) не обращаются в нуль одновременно. В самом деле, если $A = B = 0$, то при $e \geqslant 1$ и $\cos\theta = 1/e$ правая часть и коэффициент при t в формуле (7) обращаются в нуль одновременно, чего, как только

что отмечалось, быть не может; при $e < 1$ получаем, что при любом θ прямая и кривая имеют единственную общую точку M_0, т. е. вся кривая состоит из единственной точки M_0, чего также не может быть.

Итак, в соответствии с замечанием 1, возможны три случая.

1°. Коэффициент при t в уравнении (7) обращается в нуль, и, следовательно, правая часть этого уравнения отлична от нуля. В этом случае уравнение решений не имеет, поэтому прямая и кривая L имеют единственную общую точку $M_0(x_0, y_0)$.

Для эллипса эта возможность, очевидно, не реализуется.

В случае параболы коэффициент при t обращается в нуль при $\cos\theta = 1$, т. е. тогда, когда прямая параллельна оси параболы.

В случае гиперболы коэффициент при t обращается в нуль при $\cos\theta = \pm 1/e = \pm a/c$, $\sin\theta = \sqrt{1 - a^2/c^2} = b/c$ [1], $\operatorname{tg}\theta = \pm b/a$, т. е. тогда, когда прямая параллельна асимптоте гиперболы.

2°. Коэффициент при t и правая часть в уравнении (7) отличны от нуля. В этом случае уравнение (7) имеет решение $t \neq 0$, поэтому прямая и кривая L имеют еще одну общую точку M_1. Тем самым, прямая M_0M_1 является секущей.

3°. Правая часть уравнения (7) обращается в нуль, и, следовательно, коэффициент при t этого уравнения отличен от нуля. В этом случае уравнение (7) не имеет решений $t \neq 0$, поэтому кривая L и прямая не имеют общих точек, отличных от точки M_0. Этот случай можно рассматривать как предельный случай случая 2°: точки M_0 и M_1 «сливаются», а прямая M_0M_1 становится касательной. Как мы только что установили, кривая L и касательная не имеют общих точек, отличных от точки касания M_0.

Следствие 1. *Через любую точку эллипса, гиперболы или параболы проходит одна и только одна касательная* (поскольку уравнение $A\cos\theta + B\sin\theta = 0$ при любых A и B, не обращающихся в нуль одновременно, имеет единственное решение θ, удовлетворяющее условию $0 \leqslant \theta < \pi$).

Следствие 2.

1° *Если прямая имеет единственную общую точку с эллипсом, то эта прямая — касательная*;

2° *если прямая имеет единственную общую точку с параболой и не параллельна оси параболы, то эта прямая — касательная*;

3° *если прямая имеет единственную общую точку с гиперболой и не параллельна асимптоте гиперболы, то эта прямая — касательная.*

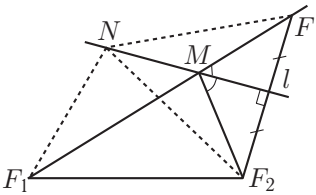
6. Оптические свойства.

Теорема 1. *Касательная к эллипсу с фокусами F_1 и F_2 в точке M является биссектрисой внешнего угла треугольника F_1F_2M.*

Доказательство. Пусть l — биссектриса внешнего угла треугольника F_1F_2M при вершине M, F — точка, симметричная F_2

[1] $\sin\theta \geqslant 0$, поскольку $0 \leqslant \theta < \pi$.

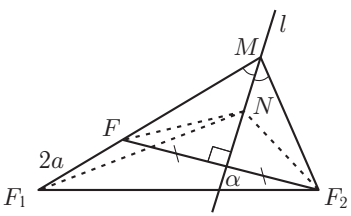
относительно прямой l. Поскольку точки F_1, M и F лежат на одной прямой, то $F_1F = F_1M + FM = F_1M + F_2M = 2a$, так как точка M лежит на эллипсе. Для любой другой точки N прямой l, в силу неравенства треугольника, $F_1N + F_2N = F_1N + FN > F_1F = 2a$, поэтому точка N не лежит на эллипсе. Итак, прямая l и эллипс имеют единственную общую точку, поэтому эта прямая — касательная. Теорема доказана.

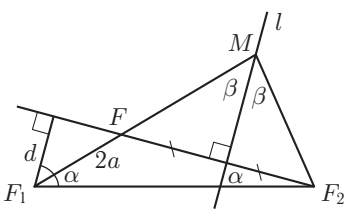


Следствие (*оптическое свойство эллипса*). *Луч света, выпущенный из фокуса эллиптического зеркала, после отражения пройдет через другой фокус.*

Теорема 2. *Касательная к гиперболе с фокусами F_1 и F_2 в точке M является биссектрисой угла треугольника F_1F_2M.*

Доказательство. Пусть, например, $F_1M > F_2M$, l — биссектриса угла треугольника F_1F_2M при вершине M, F — точка, симметричная F_2 относительно прямой l. Поскольку точки F_1, F и M лежат на одной прямой, то $F_1F = F_1M - FM = F_1M - F_2M = 2a$, так





как точка M лежит на гиперболе. Для любой другой точки N прямой l, в силу неравенства треугольника, $|F_1N - F_2N| = |F_1N - FN| < F_1F = 2a$, поэтому точка N не лежит на гиперболе. Итак, прямая l и гипербола имеют единственную общую точку.

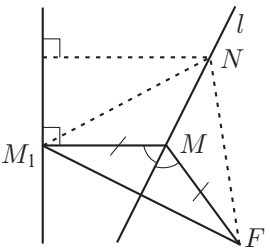
Осталось доказать, что прямая l не параллельна асимптоте гиперболы. Выразим угол α между прямыми l и F_1F_2 через угол 2β при вершине M треугольника F_1F_2M. Для этого заметим, что расстояние d от точки F_1 до прямой FF_2 может быть вычислено двумя способами: с одной стороны, $d = F_1F\cos\beta = 2a\cos\beta$, с другой стороны, $d = F_1F_2\cos\alpha = 2c\cos\alpha$. Приравнивая эти два выражения, получаем: $\cos\alpha = (a/c)\cos\beta \neq a/c$ (так как $0 < \beta \leqslant \pi/2$), поэтому прямая l не параллельна асимптоте гиперболы. Теорема доказана.

Следствие (*оптическое свойство гиперболы*). *Луч света, выпущенный из фокуса гиперболического зеркала, после отражения пойдет так, как если бы он вышел из другого фокуса.*

Теорема 3. *Касательная к параболе с фокусом F в точке M является биссектрисой угла между прямой FM и прямой, параллельной оси параболы.*

Доказательство. Пусть MM_1 — перпендикуляр, проведенный из точки M к директрисе, l — биссектриса угла FMM_1. Поскольку

точка M лежит на параболе, то $FM = MM_1$, а значит, прямая l является биссектрисой равнобедренного треугольника FMM_1 и, следовательно, серединным перпендикуляром к отрезку FM_1. Поэтому для любой другой точки N прямой l $FN = M_1N$. Но отрезок M_1N, будучи наклонной, больше расстояния от точки N до директрисы, поэтому точка N не лежит на параболе. Итак, прямая l и парабола имеют единственную общую точку. Очевидно также, что прямая l, будучи серединным перпендикуляром к отрезку FM_1, не параллельна оси параболы, поэтому эта прямая — касательная. Теорема доказана.



Следствие (*оптическое свойство параболы*). *Луч света, выпущенный из фокуса параболического зеркала, после отражения пойдет параллельно оси.*

Замечание 1. Наличием оптических свойств объясняется название «фокус». Например, в фокусе параболического зеркала фокусируется после отражения пучок лучей, параллельных оси.

Замечание 2. Пусть F_1 и F_2 — фиксированные точки. Тогда через любую точку M плоскости проходит один и только один эллипс с фокусами F_1 и F_2, поскольку сумма $F_1M + F_2M$ имеет вполне определенное значение (при этом отрезок F_1F_2 считается вырожденным эллипсом). Аналогично, через точку M проходит одна и только одна гипербола с фокусами F_1 и F_2. Из установленных нами свойств следует, что эллипс и гипербола пересекаются в точке M под прямым углом.

На этом свойстве основаны *эллиптические координаты*. Координатные линии в них представляют собой софокусные эллипсы и гиперболы.

§ 2. Кривые второго порядка

1. Уравнение кривой второго порядка.

Определение. *Кривой второго порядка называется множество всех точек $M(x, y)$ плоскости, координаты которых удовлетворяют уравнению*

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0 \qquad (1)$$

при $a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$.

Примерами кривых второго порядка могут служить эллипс, гипербола и парабола.

Замечание. Ясно, что при растяжении плоскости в каком-либо направлении, т. е. при преобразовании вида $x_1 \to px_1$, $x_2 \to qx_2$ $(p, q \neq 0)$, кривая второго порядка переходит в кривую второго порядка.

Имея своей целью классифицировать все кривые второго порядка, постараемся, перейдя к новой системе координат, максимально упростить уравнение (1). Как мы помним, такой переход осуществляется

при помощи поворота и параллельного переноса. Начнем с поворота на угол φ. Имеем:

$$\left.\begin{aligned} x_1 &= \widetilde{x}_1 \cos\varphi - \widetilde{x}_2 \sin\varphi \\ x_2 &= \widetilde{x}_1 \sin\varphi + \widetilde{x}_2 \cos\varphi \end{aligned}\right\}$$

(здесь старые координаты выражены через новые, поэтому роль φ играет $-\varphi$). Подставим эти выражения в формулу (1) и постараемся подобрать φ так, чтобы коэффициент при $\widetilde{x}_1\widetilde{x}_2$ обратился в нуль:

$$(a_{22} - a_{11}) \sin 2\varphi + 2a_{12} \cos 2\varphi = 0.$$

Если $a_{12} = 0$, то положим $\varphi = 0$; если же $a_{12} \neq 0$, то положим $\operatorname{ctg} 2\varphi = (a_{11} - a_{22})/2a_{12}$. Таким образом, коэффициент при $\widetilde{x}_1\widetilde{x}_2$ обратится в нуль, а уравнение (1) окажется приведенным к виду

$$a_1\widetilde{x}_1^2 + a_2\widetilde{x}_2^2 + 2\widetilde{b}_1\widetilde{x}_1 + 2\widetilde{b}_2\widetilde{x}_2 + \widetilde{c} = 0$$

(нетрудно заметить, что $\widetilde{c} = c$, однако для наших целей это не важно). В этом уравнении $a_1^2 + a_2^2 \neq 0$, поскольку новые и старые координаты связаны между собой линейно. Без ограничения общности будем считать, что $a_1 \neq 0$ (в противном случае оси координат можно переобозначить).

Сделаем теперь параллельный перенос вдоль оси Ox_1, полагая $\widetilde{x}_1 = x - \widetilde{b}_1/a_1$. В результате уравнение примет вид

$$a_1x^2 + a_2\widetilde{x}_2^2 + 2\widetilde{b}_2\widetilde{x}_2 + c' = 0. \tag{2}$$

Возможны два случая.

I. $a_2 \neq 0$. Полагая $\widetilde{x}_2 = y - \widetilde{b}_2/a_2$, приведем уравнение (2) к виду

$$a_1x^2 + a_2y^2 + c'' = 0,$$

где $a_1, a_2 \neq 0$. Таким образом, этот случай распадается на два:

I_1 при $c'' \neq 0$ уравнение (2) принимает вид $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 1$;

I_2 при $c'' = 0$ уравнение (2) принимает вид $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 0$.

II. $a_2 = 0$, и, следовательно, уравнение (2) имеет вид

$$a_1x^2 + 2\widetilde{b}_2\widetilde{x}_2 + c' = 0.$$

Этот случай также распадается на два:

II_1 при $\widetilde{b}_2 \neq 0$, полагая $\widetilde{x}_2 = y - c'/(2\widetilde{b}_2)$, приведем уравнение (2) к виду $x^2 = ay$, где $a \neq 0$;

II_2 при $\widetilde{b}_2 = 0$ уравнение (2) принимает вид $x^2 = a$.

Итак, уравнение (1) распадается на четыре случая.

2. Классификация. Теперь нетрудно классифицировать все кривые второго порядка. Имеем:

I_1 $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 1$:

1° при $a > 0$, $b > 0$ это уравнение описывает эллипс;

2° при $a > 0$, $b < 0$ или при $a < 0$, $b > 0$ это уравнение описывает гиперболу;

3° при $a < 0$, $b < 0$ это уравнение описывает пустое множество.

I_2 $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 0$:

4° при $a > 0$, $b > 0$ или при $a < 0$, $b < 0$ это уравнение описывает точку $x = y = 0$;

5° при $a > 0$, $b < 0$ или при $a < 0$, $b > 0$ это уравнение описывает две прямые $y = \pm\sqrt{-b/a}\,x$, пересекающиеся в точке $x = y = 0$;

II_1 $x^2 = ay$, где $a \neq 0$:

6° это уравнение описывает параболу.

II_2 $x^2 = a$:

7° при $a > 0$ это уравнение описывает две параллельные прямые $x = \pm\sqrt{a}$;

8° при $a = 0$ это уравнение описывает одну прямую $x = 0$.

Наконец, при $a < 0$ это уравнение описывает пустое множество, которое уже вошло в нашу классификацию.

Итак, *всякая кривая второго порядка представляет собой одно из восьми множеств точек*: *эллипс*, *гипербола*, *парабола*, *две пересекающиеся прямые*, *две параллельные прямые*, *одна прямая*, *точка*, *пустое множество*.

Замечание. Ясно, что при растяжении плоскости в каком-либо направлении тип каждой из указанных восьми кривых не меняется.

§ 3. Поверхности второго порядка

1. Уравнение поверхности второго порядка.

Определение. *Поверхностью второго порядка называется множество всех точек $M(x_1, x_2, x_3)$, координаты которых удовлетворяют уравнению*

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \\ + 2a_{23}x_2x_3 + 2b_1x_1 + 2b_2x_2 + 2b_3x_3 + c = 0$$

при $a_{11}^2 + a_{22}^2 + a_{33}^2 + a_{12}^2 + a_{13}^2 + a_{23}^2 \neq 0$.

Замечание. Нетрудно видеть, что сечением поверхности второго порядка плоскостью может быть только кривая второго порядка или сама эта плоскость. В самом деле, при соответствующем выборе системы координат уравнение указанной плоскости примет вид $x_3 = 0$, а уравнение сечения поверхности — вид

$$a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2 + 2b_1x_1 + 2b_2x_2 + c = 0.$$

Если $a_{11}^2 + a_{22}^2 + a_{12}^2 \neq 0$, то сечением будет кривая второго порядка. В противном случае — либо прямая, либо пустое множество, либо вся

плоскость. Это и доказывает утверждение, поскольку прямая и пустое множество — кривые второго порядка.

Наша цель состоит в том, чтобы классифицировать все поверхности второго порядка. Для этого постараемся, перейдя к новой системе координат, максимально упростить данное уравнение. Как мы помним, такой переход осуществляется при помощи поворотов на углы Эйлера и параллельного переноса. Начнем с поворотов.

В нашем распоряжении имеются три угла Эйлера. Поэтому представляется весьма правдоподобным, что подбором этих углов удастся обратить в нуль три коэффициента при произведениях разноименных координат, т. е. привести данное уравнение к виду

$$a_1\widetilde{x}_1^2 + a_2\widetilde{x}_2^2 + a_3\widetilde{x}_3^2 + 2\widetilde{b}_1\widetilde{x}_1 + 2\widetilde{b}_2\widetilde{x}_2 + 2\widetilde{b}_3\widetilde{x}_3 + \widetilde{c} = 0. \qquad (1)$$

Это утверждение и на самом деле оказывается верным, но доказать его путем подбора углов Эйлера очень трудно. Поступим иначе. Полагая $a_{ij} = a_{ji}$, заметим, прежде всего, что выражение

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 = \sum_{i,j=1}^{3} a_{ij}x_ix_j$$

можно переписать так:

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j = \sum_{i=1}^{3}\left(\sum_{j=1}^{3} a_{ij}x_j\right)x_i = (A\mathbf{x}\cdot\mathbf{x}).$$

Далее, поскольку $a_{ij} = a_{ji}$, то

$$(A\mathbf{x}\cdot\mathbf{y}) = \sum_{i,j=1}^{3} a_{ij}x_jy_i = \sum_{i,j=1}^{3} a_{ji}y_ix_j = (A\mathbf{y}\cdot\mathbf{x}).$$

Перейдем теперь к решению нашей задачи. Попробуем сначала ответить на такой вопрос: как должен быть направлен вектор $\widetilde{\mathbf{e}}_1$, чтобы в системе координат $O\widetilde{x}_1x'_2x'_3$ выражение $\sum\limits_{i,j=1}^{3} a_{ij}x_ix_j$ приняло вид

$$a_1\widetilde{x}_1^2 + \left(a'_{22}x_2'^2 + a'_{33}x_3'^2 + 2a'_{23}x'_2x'_3\right)? \qquad (2)$$

Представим вектор $\mathbf{x}$ в виде $\mathbf{x} = \widetilde{x}_1\widetilde{\mathbf{e}}_1 + x'_2\mathbf{e}'_2 + x'_3\mathbf{e}'_3$. Поскольку $A\mathbf{x} = \widetilde{x}_1A\widetilde{\mathbf{e}}_1 + x'_2A\mathbf{e}'_2 + x'_3A\mathbf{e}'_3$, то

$$(A\mathbf{x}\cdot\mathbf{x}) = (A\widetilde{\mathbf{e}}_1\cdot\widetilde{\mathbf{e}}_1)\,\widetilde{x}_1^2 + 2\,(A\widetilde{\mathbf{e}}_1\cdot\mathbf{e}'_2)\,\widetilde{x}_1x'_2 + 2\,(A\widetilde{\mathbf{e}}_1\cdot\mathbf{e}'_3)\,\widetilde{x}_1x'_3 +$$
$$+ (A\mathbf{e}'_2\cdot\mathbf{e}'_2)\,x_2'^2 + 2\,(A\mathbf{e}'_2\cdot\mathbf{e}'_3)\,x'_2x'_3 + (A\mathbf{e}'_3\cdot\mathbf{e}'_3)\,x_3'^2.$$

Мы хотим, чтобы выражения $\left(A\widetilde{\mathbf{e}}_1\cdot\mathbf{e}'_2\right)$ и $\left(A\widetilde{\mathbf{e}}_1\cdot\mathbf{e}'_3\right)$ обратились в нуль. Для этого векторы $A\widetilde{\mathbf{e}}_1$ и $\widetilde{\mathbf{e}}_1$ должны быть коллинеарными, т. е. $A\widetilde{\mathbf{e}}_1 = \lambda\widetilde{\mathbf{e}}_1$, или, что то же самое, $(A - \lambda E)\widetilde{\mathbf{e}}_1 = \mathbf{o}$. Эта однородная система имеет нетривиальное решение $\widetilde{\mathbf{e}}_1$ тогда и только тогда, когда ее

определитель равен нулю, т. е. когда число λ удовлетворяет кубическому уравнению $\det(A - \lambda E) = 0$. Поскольку кубическое уравнение всегда имеет по крайней мере один вещественный корень λ_0, то при $\lambda = \lambda_0$ рассматриваемая система имеет нетривиальное решение $\widetilde{\mathbf{e}}_1$. Принимая вектор $\widetilde{\mathbf{e}}_1$ за первый координатный вектор новой системы координат, мы и приведем выражение $\sum\limits_{i,j=1}^{3} a_{ij}x_ix_j$ к виду (2). Теперь (точно так же, как это делалось применительно к кривым второго порядка) поворотом в плоскости $Ox_2'x_3'$ можно привести исходное уравнение к виду (1) (нетрудно заметить, что $\widetilde{c} = c$, однако для наших целей это не важно).

В нашем распоряжении остается еще параллельный перенос. Без ограничения общности будем считать, что $|a_1| \geqslant |a_2| \geqslant |a_3|$ (в противном случае оси координат можно переобозначить). Ясно, что $a_1^2 + a_2^2 + a_3^2 \neq 0$. Поэтому представляются возможными три случая.

I. a_1, a_2, $a_3 \neq 0$. Полагая $\widetilde{x}_1 = x - \widetilde{b}_1/a_1$, $\widetilde{x}_2 = y - \widetilde{b}_2/a_2$, $\widetilde{x}_3 = z - \widetilde{b}_3/a_3$, приведем уравнение (1) к виду

$$a_1x^2 + a_2y^2 + a_3z^2 + c' = 0.$$

Таким образом, этот случай распадается на два:

I_1 при $c' \neq 0$ уравнение (1) принимает вид $\dfrac{x^2}{a} + \dfrac{y^2}{b} + \dfrac{z^2}{c} = 1$;

I_2 при $c' = 0$ уравнение (1) принимает вид $\dfrac{x^2}{a} + \dfrac{y^2}{b} + \dfrac{z^2}{c} = 0$.

II. $a_1, a_2 \neq 0$, $a_3 = 0$. Полагая $\widetilde{x}_1 = x - \dfrac{\widetilde{b}_1}{a_1}$, $\widetilde{x}_2 = y - \dfrac{\widetilde{b}_2}{a_2}$, приведем уравнение (1) к виду

$$a_1x^2 + a_2y^2 + 2\widetilde{b}_3\widetilde{x}_3 + c' = 0.$$

Этот случай распадается на три:

II_1 при $\widetilde{b}_3 \neq 0$, полагая $\widetilde{x}_3 = z - \dfrac{c'}{2\widetilde{b}_3}$, получим уравнение вида $\dfrac{x^2}{a} + \dfrac{y^2}{b} = z$;

II_2 при $\widetilde{b}_3 = 0$, $c' \neq 0$ уравнение приводится к виду $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 1$;

II_3 при $\widetilde{b}_3 = 0$, $c' = 0$ уравнение принимает вид $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 0$.

III. $a_1 \neq 0$, $a_2 = a_3 = 0$. Полагая $\widetilde{x}_1 = x - \dfrac{\widetilde{b}_1}{a_1}$, приведем уравнение (1) к виду

$$a_1x^2 + 2\widetilde{b}_2\widetilde{x}_2 + 2\widetilde{b}_3\widetilde{x}_3 + c' = 0.$$

Этот случай распадается на два:

III_1 при $\widetilde{b}_2^2 + \widetilde{b}_3^2 \neq 0$ сделаем сначала поворот

$$\left.\begin{aligned} \widetilde{x}_2 &= y' \cos\varphi - z' \sin\varphi \\ \widetilde{x}_3 &= y' \sin\varphi + z' \cos\varphi \end{aligned}\right\}$$

и подберем угол φ так, чтобы коэффициент при z' обратился в нуль. Для этого, как нетрудно видеть, достаточно положить $\varphi = 0$ при $\widetilde{b}_3 = 0$ или $\operatorname{ctg}\varphi = \widetilde{b}_2/\widetilde{b}_3$ при $\widetilde{b}_3 \neq 0$. В результате получится уравнение $a_1x^2 + by' + c' = 0$, где $b \neq 0$. Параллельным переносом $y' = y - c'/b$ оно приводится к виду $x^2 = ay$, где $a \neq 0$;

III_2 при $\widetilde{b}_2 = \widetilde{b}_3 = 0$ уравнение принимает вид $x^2 = a$.

Итак, уравнение (1) распадается на семь случаев.

2. Цилиндры. Рассмотрим кривую, лежащую в плоскости, и через каждую ее точку проведем прямую, перпендикулярную к этой плоскости. Поверхность, образованная проведенными прямыми, называется *цилиндром*. Указанные прямые называются *образующими* цилиндра, а исходная кривая — его *направляющей*.

Можно сказать иначе: направляющая цилиндра — это его сечение плоскостью, перпендикулярной к образующей. Из этого следует, что если среди поверхностей второго порядка есть цилиндры, то их не более, чем кривых второго порядка, т. е. не более 8.

С другой стороны, все уравнения, в которые не входит переменная z (случаи II_2, II_3, III_1, III_2), являются, очевидно, уравнениями цилиндров с образующими, параллельными оси Oz (поскольку переменная z может принимать произвольные значения). По внешнему виду указанные уравнения идентичны уравнениям всех кривых второго порядка (случаи I_1, I_2, II_1, II_2 предыдущего параграфа). Поэтому цилиндров второго порядка ровно восемь:

1° *эллиптический цилиндр* (направляющая — эллипс);

2° *гиперболический цилиндр* (направляющая — гипербола);

3° *параболический цилиндр* (направляющая — парабола);

4° *две пересекающиеся плоскости* (направляющая — две пересекающиеся прямые);

5° *две параллельные плоскости* (направляющая — две параллельные прямые);

6° *одна плоскость* (направляющая — одна прямая);

7° *прямая* (направляющая — точка);

8° *пустое множество* (направляющая — пустое множество).

3. Конусы. Рассмотрим кривую, лежащую в плоскости, и точку, не лежащую в этой плоскости. Через эту точку и каждую точку кривой проведем прямую. Поверхность, образованная проведенными прямыми, называется *конусом*. Указанные прямые называются *образующими* конуса, исходная кривая — его *направляющей*, а исходная точка — *вершиной* конуса.

Ясно, что среди поверхностей второго порядка конусов не более, чем кривых второго порядка, т. е. не более 8. С другой стороны, все поверхности, описываемые уравнением

$$\frac{x^2}{a}+\frac{y^2}{b}+\frac{z^2}{c}=0$$

(случай I_2), являются конусами с вершиной в начале координат. В самом деле, как видно из уравнения, вместе с каждой точкой (x_0, y_0, z_0), отличной от начала координат, рассматриваемая поверхность целиком содержит прямую

$$\left.\begin{aligned} x &= tx_0 \\ y &= ty_0 \\ z &= tz_0 \end{aligned}\right\},$$

проходящую через эту точку (при $t=1$) и начало координат (при $t=0$).

Представляются возможными два случая:

а) числа a, b и c имеют один и тот же знак; в этом случае нашему уравнению удовлетворяют координаты единственной точки — начала координат;

б) знак одного из чисел a, b и c противоположен знаку двух других (без ограничения общности можно считать, что $a, b > 0$, а $c < 0$) и, следовательно, наше уравнение имеет вид

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z^2.$$

Поверхность, описываемая этим уравнением, называется *конусом второго порядка*.

Итак, мы нашли еще две поверхности второго порядка:

9° *одна точка*;

10° *конус второго порядка*.

Первая из этих поверхностей представляет собой конус, направляющей которого является пустое множество.

Образующей конуса второго порядка является, например, сечение плоскостью $z=1$, представляющее собой эллипс с полуосями a и b. Поэтому конус второго порядка можно назвать также эллиптическим конусом. При $a=b=1$ он представляет собой прямой круговой конус с прямым углом при вершине:

$$x^2+y^2=z^2.$$

Ясно, что в общем случае конус второго порядка может быть получен из прямого кругового конуса с прямым углом при вершине равномерным растяжением в a раз вдоль оси Ox и в b раз вдоль оси Oy (т. е. преобразованием $x \to (1/a)x,\ \ y \to (1/b)y$).

Выясним, что представляет собой сечение прямого кругового конуса с прямым углом при вершине плоскостью, не проходящей через вершину. Для этого поступим так: сначала повернем систему координат на угол φ, полагая

$$\left.\begin{aligned} x &= \widetilde{x}\cos\varphi - \widetilde{z}\sin\varphi \\ z &= \widetilde{x}\sin\varphi + \widetilde{z}\cos\varphi \end{aligned}\right\},$$

а затем рассмотрим сечение плоскостью $\widetilde{z} = z_0$. В результате в плоскости $\widetilde{z} = z_0$ получим уравнение

$$\widetilde{x}^2 \cos 2\varphi - 2\widetilde{x} z_0 \sin 2\varphi + y^2 = z_0^2 \cos 2\varphi.$$

При $\cos 2\varphi = 0$ это уравнение представляет собой уравнение параболы. Если же $\cos 2\varphi \neq 0$, то, разделив на него, приведем полученное уравнение к виду

$$(\widetilde{x} - z_0 \operatorname{tg} 2\varphi)^2 + \frac{y^2}{\cos 2\varphi} = \left(\frac{z_0}{\cos 2\varphi}\right)^2.$$

Ясно, что при $\cos 2\varphi > 0$ полученное уравнение представляет собой уравнение эллипса, а при $\cos 2\varphi < 0$ — уравнение гиперболы. Итак, *сечение прямого кругового конуса с прямым углом при вершине плоскостью, не проходящей через вершину, представляет собой либо эллипс, либо гиперболу, либо параболу*. Поскольку любой конус второго порядка может быть получен из прямого кругового конуса с прямым углом при вершине равномерным растяжением вдоль осей Ox и Oy, то его сечения также обладают указанным свойством. По этой причине эллипс, гиперболу и параболу иногда называют *коническими сечениями*.

Замечание. Сказанное позволяет заключить, что конус второго порядка, будучи эллиптическим, в определенном смысле может рассматриваться также как параболический или гиперболический — отсюда и название «конус второго порядка».

Конусы, направляющими которых являются две пересекающиеся или две параллельные прямые, одна прямая и точка, представляют собой соответственно две пересекающиеся плоскости, одну плоскость и прямую, т. е. поверхности, которые мы отнесли к числу цилиндров. Следовательно, *среди поверхностей второго порядка конусов, отличных от цилиндров, ровно два*.

4. Завершение классификации. Осталось рассмотреть два случая.

I_1 $\dfrac{x^2}{a} + \dfrac{y^2}{b} + \dfrac{z^2}{c} = 1.$

11°. Если числа a, b и c положительны, то уравнение поверхности имеет вид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Поверхность, описываемая этим уравнением, называется *эллипсоидом*.

12°. Если из чисел a, b и c два положительны, а одно отрицательно, то уравнение поверхности имеет вид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} + 1.$$

Поверхность, описываемая этим уравнением, называется *однополостным гиперболоидом*.

13°. Если одно из чисел a, b и c положительно, а два отрицательны, то уравнение поверхности имеет вид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} - 1.$$

Поверхность, описываемая этим уравнением, называется *двуполостным гиперболоидом*.

Наконец, если числа a, b и c отрицательны, то это уравнение описывает пустое множество, которое уже вошло в нашу классификацию.

II_1 $\dfrac{x^2}{a} + \dfrac{y^2}{b} = z.$

14°. Если знаки чисел a и b совпадают, то уравнение поверхности имеет вид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z.$$

Поверхность, описываемая этим уравнением, называется *эллиптическим параболоидом*.

15°. Если знаки чисел a и b не совпадают, то уравнение поверхности имеет вид

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z.$$

Поверхность, описываемая этим уравнением, называется *гиперболическим параболоидом*.

Итак, *существует ровно 15 типов поверхностей второго порядка*: *8 цилиндров*, *2 конуса*, *эллипсоид*, *2 гиперболоида и 2 параболоида*.

§ 4. Эллипсоид, гиперболоиды и параболоиды

1. Эллипсоид. Чтобы лучше представить себе форму эллипсоида, заданного уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

допустим сначала, что $b = a$. Выясним, что представляет собой сечение такого эллипсоида плоскостью $z = z_0$. Для этого перепишем уравнение нашего эллипсоида так:

$$x^2 + y^2 = \left(1 - \frac{z^2}{c^2}\right) a^2.$$

Полагая в этом уравнении $z = z_0$, обнаружим, что при $-c \leqslant z_0 \leqslant c$ сечением является окружность с центром на оси Oz (нулевого радиуса при $z = \pm c$), а при остальных значениях z_0 — пустое множество. Если вращать наш эллипсоид вокруг оси Oz, то каждая точка указанной

окружности будет переходить в точку этой же окружности, а значит сам эллипсоид будет переходит в себя. Ясно также, что весь эллипсоид может быть получен вращением вокруг оси Oz своего сечения плоскостью $y = 0$, т. е. эллипса

$$\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1.$$

Поэтому такой эллипсоид называют *эллипсоидом вращения*.

В общем случае эллипсоид получается из эллипсоида вращения равномерным растяжением вдоль оси, перпендикулярной к оси вращения (в нашем случае — растяжением в b/a раз вдоль оси Oy), что позволяет отчетливо представить себе его форму.

2. Гиперболоиды. Рассмотрим сначала двуполостный гиперболоид, заданный уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} - 1.$$

Ясно, что эта поверхность не имеет точек в слое $-c < z_0 < c$ и, следовательно, состоит из двух частей, в одной из которых $z_0 < -c$, а в другой $z_0 > c$. Отсюда и название — «двуполостный» гиперболоид.

При $b = a$ уравнение двуполостного гиперболоида может быть записано так:

$$x^2 + y^2 = \left(\frac{z^2}{c^2} - 1\right) a^2.$$

Поскольку сечения такого гиперболоида плоскостями $z = \mathrm{const}$ представляют собой окружности с центром на оси Oz, то, если вращать его вокруг оси Oz, он будет переходит в себя. Таким образом, этот гиперболоид может быть получен вращением вокруг оси Oz своего сечения $y = 0$, т. е. вращением гиперболы

$$\frac{z^2}{c^2} - \frac{x^2}{a^2} = 1$$

вокруг ее вещественной оси. Поэтому такой гиперболоид называют *двуполостным гиперболоидом вращения*. В общем случае двуполостный гиперболоид получается из двуполостного гиперболоида вращения равномерным растяжением вдоль оси, перпендикулярной к оси вращения.

Совершенно аналогично однополостный гиперболоид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} + 1$$

получается из *однополостного гиперболоида вращения* равномерным растяжением вдоль оси, перпендикулярной к оси вращения. Последний может быть получен вращением гиперболы

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$$

вокруг ее мнимой оси. По форме однополостный гиперболоид напоминает трубку с раструбами. Он состоит из одной части, поэтому и называется «однополостный».

Замечание 1. Рассмотрим сечение однополостного гиперболоида вращения

$$\frac{x^2}{a^2}+\frac{y^2}{a^2}=\frac{z^2}{c^2}+1$$

плоскостью $x=a$. Это сечение, очевидно, представляет собой две пересекающиеся прямые: прямую m_1, определяемую уравнениями

$$\left.\begin{aligned} x&=a\\ \frac{y}{a}&=\frac{z}{c}\end{aligned}\right\},$$

и прямую m_2, определяемую уравнениями

$$\left.\begin{aligned} x&=a\\ \frac{y}{a}&=-\frac{z}{c}\end{aligned}\right\}.$$

Рассмотрим одну из них, например прямую m_1. Ясно, что если мы будем вращать эту прямую вокруг оси Oz, то она «заметет» весь гиперболоид. Иными словами, *однополостный гиперболоид вращения может быть получен вращением прямой вокруг оси, скрещивающейся с этой прямой*.

Замечание 2. Мысль, высказанная в замечании 1, допускает дальнейшее развитие. Ясно, что однополостный гиперболоид вращения может быть получен вращением вокруг оси Oz не только прямой m_1, но и прямой m_2. Это означает, что через каждую точку нашего гиперболоида проходят две прямые, целиком ему принадлежащие. Поскольку любой однополостный гиперболоид может быть получен из нашего равномерным растяжением, то он также обладают указанным свойством. Итак, *через каждую точку однополостного гиперболоида проходят две прямые, целиком ему принадлежащие*.

3. Параболоиды. Рассмотрим сначала эллиптический параболоид, заданный уравнением

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z.$$

Ясно, что при $b=a$ он представляет собой поверхность, получаемую вращением параболы

$$x^2=a^2z$$

вокруг оси Oz, т. е. вокруг оси параболы. Эта поверхность называется *параболоидом вращения*. В общем случае эллиптический параболоид получается из параболоида вращения равномерным растяжением вдоль оси, перпендикулярной к оси вращения.

Рассмотрим теперь гиперболический параболоид, заданный уравнением

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=z.$$

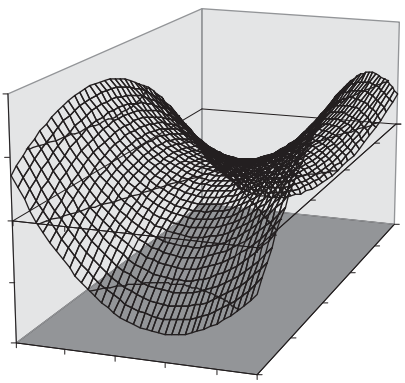
Эта поверхность симметрична относительно плоскостей Oxz, Oyz и оси Oz. Ее сечения плоскостями $x = x_0$ представляют собой параболы

$$z - \frac{x_0^2}{a^2} = -\frac{y^2}{b^2},$$

выпуклые «вверх», а сечения плоскостями $y = y_0$ — параболы

$$z + \frac{y_0^2}{b^2} = \frac{x^2}{a^2},$$

выпуклые «вниз». Ясно, что такая поверхность не имеет никакого сходства с поверхностью вращения. По своей форме она напоминает седло.



При $a = b = 1$ уравнение гиперболического параболоида принимает вид $x^2 - y^2 = z$. Полагая $x = (\widetilde{x} + \widetilde{y})/\sqrt{2}$, $y = (-\widetilde{x} + \widetilde{y})/\sqrt{2}$ (т. е. осуществляя в плоскости Oxy поворот на $-\pi/4$), приведем это уравнение к виду $2\widetilde{x}\widetilde{y} = z$. Нетрудно заметить, что через каждую точку $(x_0, y_0, 2x_0y_0)$ этой поверхности проходят две прямые: $\left.\begin{array}{l}\widetilde{x} = x_0\\ z = 2x_0\widetilde{y}\end{array}\right\}$ и $\left.\begin{array}{l}\widetilde{y} = y_0\\ z = 2y_0\widetilde{x}\end{array}\right\}$, целиком ей принадлежащие. Поскольку любой гиперболический параболоид может быть получен из нашего равномерным растяжением, то он также обладают указанным свойством. Итак, *через каждую точку гиперболического параболоида проходят две прямые, целиком ему принадлежащие.*

Замечание. Таким образом, среди поверхностей второго порядка мы нашли две поверхности — однополостный гиперболоид и гиперболический параболоид, каждая из которых покрыта двумя различными семействами прямых. На первый взгляд может показаться, что среди произвольных поверхностей могут быть и другие поверхности, обладающие этим свойством. Оказывается, однако, что это не так. В качестве самостоятельного упражнения предлагается доказать, что если поверхность обладает указанным свойством, то эта поверхность — либо плоскость, либо однополостный гиперболоид, либо гиперболический параболоид.

Часть III

ЛИНЕЙНАЯ АЛГЕБРА

Глава 1
КОНЕЧНОМЕРНОЕ ЛИНЕЙНОЕ ПРОСТРАНСТВО

§ 1. Линейное пространство

1. Аксиомы Вейля. Как известно, одна из первых попыток логически обосновать геометрию, в частности сформулировать систему аксиом геометрии, была предпринята Евклидом в его знаменитых «Началах» (ок. 300 г. до н. э.). Полностью вопрос об аксиоматизации геометрии был решен великим немецким математиком Давидом Гильбертом (1899 г.). В несколько видоизмененном виде система аксиом Гильберта фигурирует в современных школьных учебниках геометрии, поэтому здесь мы не будем на ней останавливаться подробно. Подчеркнем лишь, что основными объектами в ней являются *точки*, *прямые*, *плоскости*, «*лежать между*» (для трех точек одной прямой) и ряд других.

В 1918 году немецкий математик Герман Вейль предложил принципиально новую систему аксиом геометрии. В ней основных объектов два: *векторы* и *точки*. Эти объекты не определяются, а их свойства описываются аксиомами, список которых удобно разделить на четыре группы.

I. Имеются две операции: *операция сложения*, ставящая в соответствие любой упорядоченной паре векторов $\mathbf{x}$, $\mathbf{y}$ единственный вектор $\mathbf{x}+\mathbf{y}$, и *операция умножения вектора на число*, ставящая в соответствие любому вектору $\mathbf{x}$ и любому числу λ единственный вектор $\lambda\mathbf{x}$. При этом:

1° *для любых векторов* $\mathbf{x}$ и $\mathbf{y}$ *выполняется равенство* $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$;

2° *для любых векторов* $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{z}$ *выполняется равенство* $\mathbf{x} + (\mathbf{y}+\mathbf{z}) = (\mathbf{x}+\mathbf{y})+\mathbf{z}$;

3° *существует такой вектор* $\mathbf{o}$, *что для любого вектора* $\mathbf{x}$ *выполняется равенство* $\mathbf{x}+\mathbf{o}=\mathbf{x}$;

4° *для любого вектора* $\mathbf{x}$ *существует такой вектор* $(-\mathbf{x})$, *что* $\mathbf{x}+(-\mathbf{x})=\mathbf{o}$;

5° *для любого вектора* $\mathbf{x}$ *имеет место равенство* $1\mathbf{x}=\mathbf{x}$;

6° *для любых векторов* $\mathbf{x}$ и $\mathbf{y}$ *и любого числа* λ *выполняется равенство* $\lambda(\mathbf{x}+\mathbf{y}) = \lambda\mathbf{x}+\lambda\mathbf{y}$;

7° *для любого вектора* $\mathbf{x}$ *и любых чисел* λ *и* μ *имеет место равенство* $(\lambda+\mu)\mathbf{x} = \lambda\mathbf{x}+\mu\mathbf{x}$;

8° *для любого вектора* $\mathbf{x}$ *и любых чисел* λ *и* μ *имеет место равенство* $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$.

II. Пользуясь приведенными аксиомами, можно определить понятия линейной комбинации векторов (тривиальной, или нетривиальной), линейной зависимости и линейной независимости так, как это делалось в п. 3 §1 гл. 1 ч. 1. Это позволяет сформулировать еще две аксиомы:

1° *существует n линейно независимых векторов*;

2° *любые* $(n+1)$ *векторов линейно зависимы*.

Для планиметрии $n = 2$, а для стереометрии $n = 3$.

III. Предполагается, что имеется *операция скалярного умножения*, ставящая в соответствие любой упорядоченной паре векторов $\mathbf{x}$, $\mathbf{y}$ единственное число $(\mathbf{x}, \mathbf{y})$. При этом:

1° *для любых векторов* $\mathbf{x}$ и $\mathbf{y}$ *выполняется равенство* $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$;

2° *для любых векторов* $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{z}$ *выполняется равенство* $(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$;

3° *для любых векторов* $\mathbf{x}$ и $\mathbf{y}$ *и любого числа* λ *выполняется равенство* $(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$;

4° *для любого вектора* $\mathbf{x}$ *имеет место неравенство* $(\mathbf{x}, \mathbf{x}) \geqslant 0$, причем $(\mathbf{x}, \mathbf{x}) = 0$ *тогда и только тогда, когда* $\mathbf{x} = \mathbf{o}$.

IV. Имеются точки и правило, по которому любой упорядоченной паре точек A, B ставится в соответствие единственный вектор $\mathbf{AB}$. При этом:

1° *для любого вектора* $\mathbf{x}$ *и любой точки* O *существует единственная точка* M *такая*, что $\mathbf{OM} = \mathbf{x}$ (откладывание вектора от данной точки);

2° *для любых точек* A, B *и* C *имеет место равенство* $\mathbf{AB} + \mathbf{BC} = \mathbf{AC}$ (правило треугольника).

Как мы знаем, если в рамках аксиоматики Гильберта назвать вектором направленный отрезок, а сумму векторов, произведение вектора на число и скалярное произведение векторов определить так, как это было сделано выше, то все аксиомы Вейля окажутся теоремами — мы их доказали. Оказывается, что верно и обратное: если, исходя из аксиом Вейля, определить прямые, плоскости и другие основные понятия системы аксиом Гильберта, то все аксиомы Гильберта станут теоремами — их можно будет доказать. Таким образом, системы аксиом Гильберта и Вейля эквивалентны.

Из рассмотрения аксиом Вейля становится ясным, что основную роль в них играют векторы — точки появляются лишь в последней группе аксиом, причем от точек требуется лишь, чтобы они были определенным образом связаны с векторами. Поэтому наиболее содержательная часть евклидовой геометрии должна описываться векторами. Это наблюдение наводит на мысль о целесообразности создания отдельного раздела математики, посвященного изучению свойств векторов, описываемых аксиомами Вейля (I, II и III группы аксиом). Таким разделом математики и является *линейная алгебра*.

2. Линейное пространство.

Определение. *Множество* $\mathbf{L}$ *называется линейным пространством, если для всех его элементов определены операции сложения двух элементов и умножения элемента на число, удовлетворяющие*

I группе аксиом Вейля. Элементы линейного пространства называются векторами.

Замечание. Числа, фигурирующие в этом определении, понимаются, вообще говоря, в весьма широком смысле (на этом вопросе мы остановимся в конце курса). Мы же будем понимать под числами либо вещественные числа (в этом случае говорят о линейном пространстве над полем $\mathbb{R}$ вещественных чисел), либо комплексные числа (в этом случае говорят о линейном пространстве над полем $\mathbb{C}$ комплексных чисел). Как правило, о каких именно числах — вещественных или комплексных — идет речь, для нас будет несущественно. В тех случаях, когда это будет принципиально, мы будем специально оговаривать, что в данном случае понимается под словом «число».

Пример 1. Очевидным примером линейного пространства над полем $\mathbb{R}$ является множество векторов — направленных отрезков на прямой, на плоскости или в пространстве.

Пример 2. Не менее очевидный пример линейного пространства — арифметическое пространство ($\mathbb{R}^n$, или $\mathbb{C}^n$).

Пример 3. Еще одним очевидным примером линейного пространства является множество всех $(m \times n)$-матриц.

Пример 4. Рассмотрим множество $C[a, b]$ всех непрерывных на отрезке $[a, b]$ функций. Если сумму функций и умножение функции на число понимать в общепринятом смысле, то это множество также окажется примером линейного пространства, поскольку при фиксированном значении аргумента складываются и умножаются на число обычные числа, причем в результате таких операций получаются функции, непрерывные на отрезке $[a, b]$.

Можно привести очень большое количество самых разнообразных примеров линейного пространства. Но и приведенных примеров достаточно, чтобы понять, что класс линейных пространств весьма широк.

3. Свойства линейного пространства.

Теорема 1. *В линейном пространстве*:

1° *вектор* $\mathbf{o}$, *существование которого гарантируется аксиомой* 3°, *определяется единственным образом*;

2° *вектор* $(-\mathbf{x})$, *существование которого гарантируется аксиомой* 4°, *определяется единственным образом.*

Доказательство. 1°. Допустим, что существуют два вектора — $\mathbf{o}_1$ и $\mathbf{o}_2$ — такие, что для любого вектора $\mathbf{x}$ имеют место равенства: $\mathbf{x} + \mathbf{o}_1 = \mathbf{x} + \mathbf{o}_2 = \mathbf{x}$. Тогда, в частности, $\mathbf{o}_1 + \mathbf{o}_2 = \mathbf{o}_1 = \mathbf{o}_2 + \mathbf{o}_1 = \mathbf{o}_2$, т. е. $\mathbf{o}_1 = \mathbf{o}_2$.

2°. Допустим, что для некоторого вектора $\mathbf{x}$ существуют два вектора — $(-\mathbf{x})_1$ и $(-\mathbf{x})_2$ — такие, что имеют место равенства: $\mathbf{x} + (-\mathbf{x})_1 = \mathbf{x} + (-\mathbf{x})_2 = \mathbf{o}$. Тогда $(-\mathbf{x})_1 + (\mathbf{x} + (-\mathbf{x})_2) = (-\mathbf{x})_1 + \mathbf{o} = (-\mathbf{x})_1 = (-\mathbf{x})_1 + \mathbf{x}) + (-\mathbf{x})_2 = \mathbf{o} + (-\mathbf{x})_2 = (-\mathbf{x})_2 + \mathbf{o} = (-\mathbf{x})_2$, т. е. $(-\mathbf{x})_1 = (-\mathbf{x})_2$. Теорема доказана.

Теорема 2. *Для любого* $\mathbf{x} \in \mathbf{L}$ *имеют место равенства*:

1° $0\mathbf{x} = \mathbf{o}$;

2° $(-1)\mathbf{x} = (-\mathbf{x})$.

Доказательство. 1°. Имеем: $0\mathbf{x} + \mathbf{x} = (0+1)\mathbf{x} = \mathbf{x}$. Прибавляя к левой и правой частям этого равенства вектор $(-\mathbf{x})$, получаем: $0\mathbf{x} + (\mathbf{x} + (-\mathbf{x})) = 0\mathbf{x} + \mathbf{o} = 0\mathbf{x} = \mathbf{x} + (-\mathbf{x}) = \mathbf{o}$, откуда $0\mathbf{x} = \mathbf{o}$.

2°. Имеем: $\mathbf{x} + (-1)\mathbf{x} = (1-1)\mathbf{x} = 0\mathbf{x} = \mathbf{o}$. Теорема доказана.

Теорема 3. *Для любых двух векторов* $\mathbf{a}, \mathbf{b} \in \mathbf{L}$ *существует единственный вектор* $\mathbf{x} \in \mathbf{L}$, *являющийся решением уравнения* $\mathbf{a} + \mathbf{x} = \mathbf{b}$ *и называемый разностью* $\mathbf{b} - \mathbf{a}$ *векторов* $\mathbf{b}$ и $\mathbf{a}$.

Доказательство. Допустим сначала, что существует такой вектор $\mathbf{x}$, что $\mathbf{a} + \mathbf{x} = \mathbf{b}$. Прибавим к обеим частям этого тождества вектор $(-\mathbf{a})$: $(-\mathbf{a}+\mathbf{a})+\mathbf{x} = (-\mathbf{a})+\mathbf{b}$, откуда $\mathbf{x} = (-\mathbf{a})+\mathbf{b}$. Таким образом, если решение уравнения $\mathbf{a} + \mathbf{x} = \mathbf{b}$ существует, то оно определяется единственным образом: $\mathbf{x} = (-\mathbf{a}) + \mathbf{b}$. Осталось доказать, что $\mathbf{x} = (-\mathbf{a}) + \mathbf{b}$ действительно является решением указанного уравнения. Имеем: $\mathbf{a} + (-\mathbf{a}) + \mathbf{b} = \mathbf{o} + \mathbf{b} = \mathbf{b}$, что и требовалось доказать.

Следствие. $\mathbf{b} - \mathbf{a} = \mathbf{b} + (-\mathbf{a}) = \mathbf{b} + (-1)\mathbf{a}$.

4. Линейное подпространство.

Определение. *Подмножество* $\mathbf{L}'$ *линейного пространства* $\mathbf{L}$ *называется линейным подпространством, если оно само является линейным пространством, в котором сумма векторов и произведение вектора на число определяются так же, как в* $\mathbf{L}$.

Теорема. *Непустое подмножество* $\mathbf{L}'$ *линейного пространства является линейным подпространством тогда и только тогда, когда оно обладает двумя свойствами*:

1° *для любых двух векторов* $\mathbf{x}$ *и* $\mathbf{y}$ *множества* $\mathbf{L}'$ *вектор* $\mathbf{x} + \mathbf{y}$ *принадлежит* $\mathbf{L}'$;

2° *для любого вектора* $\mathbf{x}$ *множества* $\mathbf{L}'$ *и любого числа* λ *вектор* $\lambda\mathbf{x}$ *принадлежит* $\mathbf{L}'$.

Доказательство. Ясно, что если множество $\mathbf{L}'$ является линейным подпространством, то свойствами 1° и 2° оно обладает — это следует непосредственно из определения.

Допустим, что непустое подмножество $\mathbf{L}'$ линейного пространства $\mathbf{L}$ обладает свойствами 1° и 2°. Тогда для всех элементов множества $\mathbf{L}'$ определены операции сложения двух элементов и умножения элемента на число, причем они осуществляются по тем же правилам, что и в пространстве $\mathbf{L}$. Элемент $\mathbf{o}$ пространства $\mathbf{L}$ принадлежит $\mathbf{L}'$, так как он равен $0\mathbf{x}$, где $\mathbf{x}$ — произвольный вектор $\mathbf{L}'$; для любого элемента $\mathbf{x}$ множества $\mathbf{L}'$ вектор $(-\mathbf{x})$ также принадлежит $\mathbf{L}'$, так как он равен $(-1)\mathbf{x}$. Наконец, операции сложения и умножения на число удовлетворяют и всем остальным аксиомам линейного пространства, поскольку все векторы множества $\mathbf{L}'$ являются в то же время векторами линейного пространства $\mathbf{L}$. Следовательно, множество $\mathbf{L}'$ — линейное подпространство. Теорема доказана.

Пример 1. Рассмотрим множество всех решений однородной системы линейных уравнений $A\mathbf{x} = \mathbf{o}$. Это множество является, очевидно, подмножеством арифметического пространства. С другой стороны, если $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ и $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ — два решения, то $\mathbf{x} + \mathbf{y}$ и $\lambda\mathbf{x}$ при любом λ — также решения этой системы. В самом

деле, $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{o} + \mathbf{o} = \mathbf{o}$, $A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda\mathbf{o} = \mathbf{o}$. Следовательно, множество всех решений системы $A\mathbf{x} = \mathbf{o}$ является линейным подпространством арифметического пространства.

Пример 2. *Линейной оболочкой* $\mathbf{L}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *векторов* $\mathbf{x}_1$, $\mathbf{x}_2, \ldots, \mathbf{x}_k$ называется множество всех линейных комбинаций этих векторов.

Нетрудно видеть, что:

линейная оболочка является линейным подпространством, поскольку сумма любых двух ее элементов и произведение любого ее элемента на число ей принадлежат;

линейная оболочка является минимальным линейным подпространством, *содержащим векторы* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, так как любое линейное подпространство, содержащее указанные векторы, содержит также и любую их линейную комбинацию.

Отметим, что если какие-нибудь m $(m < k)$ из данных векторов (без ограничения общности — $\mathbf{x}_1, \ldots, \mathbf{x}_m$) представляют собой линейные комбинации остальных, то $\mathbf{L}(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \mathbf{L}(\mathbf{x}_{m+1}, \ldots, \mathbf{x}_k)$, поскольку любая линейная комбинация векторов $\mathbf{x}_1, \ldots, \mathbf{x}_k$ является в то же время и линейной комбинацией векторов $\mathbf{x}_{m+1}, \ldots, \mathbf{x}_k$.

§ 2. n-мерное линейное пространство

1. Базис и размерность. Наличие двух операций в линейном пространстве позволяет определить понятия линейной комбинации и линейной зависимости векторов (п. 3 § 1 гл. 1 ч. 1), а также ввести понятие базиса так, как это делалось в аналитической геометрии (п. 4 § 4 гл. 1 ч. 2).

Определение 1. *Линейное пространство называется n-мерным* (и обозначается $\mathbf{L}^n$), *если оно удовлетворяет II группе системы аксиом Вейля*.

Число n при этом называют *размерностью* этого пространства и пишут: $\dim \mathbf{L} = n$.

Теорема 1. *Любые n линейно независимых векторов в пространстве $\mathbf{L}^n$ образуют базис*.

Доказательство. Допустим, что векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно независимы, и рассмотрим произвольный вектор $\mathbf{x}$. Поскольку размерность пространства равна n, то векторы $\mathbf{x}, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно зависимы. Следовательно, существуют такие числа $\lambda_0, \lambda_1, \ldots, \lambda_n$, не обращающиеся в нуль одновременно, что $\lambda_0\mathbf{x} + \lambda_1\mathbf{e}_1 + \ldots + \lambda_n\mathbf{e}_n = \mathbf{o}$, причем λ_0 отлично от нуля (иначе оказалось бы, что векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно зависимы). Но тогда обе части этого равенства можно умножить на $1/\lambda_0$ и получить разложение вектора $\mathbf{x}$ по векторам $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$.

Итак, векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ линейно независимы, и любой вектор $\mathbf{x}$ пространства $\mathbf{L}$ может быть по ним разложен. Теорема доказана.

Теорема 2. *Если в линейном пространстве $\mathbf{L}$ существует базис из n векторов, то это пространство — n-мерное*.

Доказательство. Допустим, что векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ образуют базис линейного пространства $\mathbf{L}$. Тем самым в этом пространстве существует n линейно независимых векторов (векторы $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$). Осталось доказать, что любые $(n+1)$ векторов линейно зависимы.

Пусть $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{n+1}$ — произвольные векторы пространства $\mathbf{L}$. Разложим каждый из них по базису $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$:

$$\left.\begin{array}{l} \mathbf{g}_1 = g_{11}\mathbf{e}_1 + \ldots + g_{1n}\mathbf{e}_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{g}_{n+1} = g_{(n+1)1}\mathbf{e}_1 + \ldots + g_{(n+1)n}\mathbf{e}_n \end{array}\right\}.$$

Ранг матрицы G не превосходит количества n ее столбцов, поэтому ее строки (их $n+1$) линейно зависимы. Следовательно, векторы $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{n+1}$ также линейно зависимы. Теорема доказана.

Определение 2. *Коэффициенты разложения вектора по базису* (т. е. числа, фигурирующие в этом разложении), *называются координатами вектора в этом базисе. Координаты вектора в данном базисе определяются единственным образом* — это утверждение доказывается точно так же, как в аналитической геометрии (п. 4 § 4 гл. 1 ч. 2). Тот факт, что вектор $\mathbf{x}$ имеет координаты $x_1, x_2, \ldots, x_n$, условимся, как и прежде, обозначать так: $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$.

Теорема 3.

1° *Если* $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$, *то* $\mathbf{x} + \mathbf{y} = \{x_1 + y_1, x_2 + y_1, \ldots, x_n + y_n\}$;

2° *если* $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, то $\lambda\mathbf{x} = \{\lambda x_1, \lambda x_2, \ldots, \lambda x_n\}$.

Доказательство. Имеем: 1° $\mathbf{x} + \mathbf{y} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n + y_1\mathbf{e}_1 + \ldots + y_n\mathbf{e}_n = (x_1 + y_1)\mathbf{e}_1 + \ldots + (x_n + y_n)\mathbf{e}_n$;

2° $\lambda\mathbf{x} = \lambda(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = (\lambda x_1)\mathbf{e}_1 + \ldots + (\lambda x_n)\mathbf{e}_n$. Теорема доказана.

2. Примеры.

Пример 1. Ясно, что размерность пространства векторов — направленных отрезков на прямой, на плоскости и в пространстве равна соответственно 1, 2 и 3.

Пример 2. Ясно также, что $\dim \mathbb{R}^n = \dim \mathbb{C}^n = n$.

Пример 3. Размерность линейного пространства — множества всех $(m \times n)$-матриц равна mn, поскольку оно отличается от $\mathbb{R}^{m\times n}$ или $\mathbb{C}^{m\times n}$ лишь формой записи элементов. Базисом в нем являются, например, $(m \times n)$-матрицы, в каждой из которых один элемент равен 1, а все остальные — 0.

Пример 4. Если векторы $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ линейно независимы, то размерность их линейной оболочки равна k, поскольку эти векторы образуют базис в $\mathbf{L}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$.

Пример 5. Множество $C[a, b]$ всех непрерывных на отрезке $[a, b]$ функций не имеет конечной размерности, поскольку, например, степенные функции $1, x, x^2, \ldots, x^n, \ldots$ линейно независимы (напомним, что многочлен тождественно равен нулю тогда и только тогда, когда все его коэффициенты равны нулю), а их бесконечно много. Линейные

пространства, не имеющие конечной размерности, изучаются в специальном разделе математики, называемом «*функциональный анализ*».

Пример 6. Найдем размерность пространства решений однородной системы линейных уравнений. Напомним, что каждое решение представляет собой набор из n чисел $\{x_1, x_2, \ldots, x_r, c_1, c_2, \ldots, c_{n-r}\}$ (здесь r — ранг матрицы системы, n — количество неизвестных), где числа $c_1, c_2, \ldots, c_{n-r}$ могут быть выбраны произвольно, а числа $x_1, x_2, \ldots, x_r$ по выбранным $c_1, c_2, \ldots, c_{n-r}$ находятся однозначно. Полагая сначала $c_1 = 1$, $c_2 = 0, \ldots, c_{n-r} = 0$, затем $c_1 = 0$, $c_2 = 1, \ldots$ $\ldots, c_{n-r} = 0$ и т. д., мы получим $(n - r)$ решений. При этом любое решение $\{x_1, x_2, \ldots, x_r, c_1, c_2, \ldots, c_{n-r}\}$ данной системы представляет собой их линейную комбинацию с коэффициентами $c_1, c_2, \ldots, c_{n-r}$, и, сверх того, они линейно независимы (поскольку равенство $\{x_1, x_2, \ldots$ $\ldots, x_r, c_1, c_2, \ldots, c_{n-r}\} = \{0, \ldots, 0\}$ возможно лишь при $c_1 = c_2 = \ldots$ $\ldots = c_{n-r} = 0$). Таким образом, указанные решения образуют базис в пространстве решений, а значит размерность этого пространства равна $n - r$.

Любой базис в пространстве решений однородной системы линейных уравнений называется *фундаментальной совокупностью решений* этой системы.

3. Изоморфизм.

Определение. *Линейные пространства* $\mathbf{L}$ и $\mathbf{L}'$ *называются изоморфными, если между их элементами можно установить такое взаимно однозначное соответствие* $\mathbf{x} \leftrightarrow \mathbf{x}'$, что:

1° *если* $\mathbf{x} \leftrightarrow \mathbf{x}'$, $\mathbf{y} \leftrightarrow \mathbf{y}'$, то $\mathbf{x} + \mathbf{y} \leftrightarrow \mathbf{x}' + \mathbf{y}'$;

2° *если* $\mathbf{x} \leftrightarrow \mathbf{x}'$, *то* $\lambda\mathbf{x} \leftrightarrow \lambda\mathbf{x}'$.

Само это соответствие называется *изоморфизмом*.

Замечание. При изоморфизме нулевой элемент пространства $\mathbf{L}$ переходит в нулевой элемент пространства $\mathbf{L}'$. В самом деле, если $\mathbf{x} \leftrightarrow \mathbf{x}'$, то $0\mathbf{x} = \mathbf{o} \leftrightarrow 0\mathbf{x}' = \mathbf{o}'$.

Теорема 1. *Если пространства* $\mathbf{L}$ и $\mathbf{L}'$ *изоморфны, то* $\dim \mathbf{L} = \dim \mathbf{L}'$.

Доказательство. Допустим, что пространства $\mathbf{L}$ и $\mathbf{L}'$ изоморфны, $\dim \mathbf{L} = n$. Пусть $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ — произвольные линейно независимые векторы пространства $\mathbf{L}$. Тогда соответствующие им векторы $\mathbf{x}'_1, \mathbf{x}'_2, \ldots, \mathbf{x}'_n$ также линейно независимы, поскольку если их линейная комбинация равна $\mathbf{o}'$, то соответствующая ей линейная комбинация векторов $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ равна $\mathbf{o}$. С другой стороны, любые $(n + 1)$ векторов пространства $\mathbf{L}'$ линейно зависимы, так как соответствующие им векторы пространства $\mathbf{L}$ линейно зависимы. Следовательно, $\dim \mathbf{L}' = n$. Теорема доказана.

Теорема 2. *Любые два линейных пространства над одним и тем же полем и одинаковой размерности изоморфны.*

Доказательство. Пусть $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ и $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ — базисы в пространствах $\mathbf{L}$ и $\mathbf{L}'$ соответственно. Поставим в соответствие произвольному элементу $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ пространства $\mathbf{L}$ элемент $\mathbf{x}' = x_1\mathbf{e}'_1 + x_2\mathbf{e}'_2 + \ldots + x_n\mathbf{e}'_n$ пространства $\mathbf{L}'$. Это соответствие,

очевидно, взаимно однозначное, а значит, как следует из теоремы 3 п. 1, является изоморфизмом. Теорема доказана.

Замечание 1. Эта теорема позволяет вместо абстрактного конечномерного линейного пространства рассматривать любую его модель, т. е. любой конкретный пример линейного пространства (например, $\mathbb{R}^n$ или $\mathbb{C}^n$). При этом все результаты, полученные в рамках этой модели, будут автоматически переноситься на абстрактное линейное пространство той же размерности или любую его модель.

Замечание 2. К любой системе аксиом обычно предъявляют три требования: *непротиворечивость*, *полнота* и *независимость*.

Непротиворечивость означает, что ни одна из аксиом не противоречит остальным аксиомам. Применительно к аксиомам Вейля это требование, очевидно, выполнено, иначе ей не могли бы удовлетворять векторы — направленные отрезки.

Полнота означает, что все модели данной системы аксиом изоморфны, т. е. между любыми основными объектами любых двух из них можно установить взаимно однозначное соответствие, сохраняющее основные отношения между ними. Тем самым теорему 2 можно сформулировать так: *система аксиом n-мерного линейного пространства полна*.

Наконец, система аксиом называется независимой, если ни одна из аксиом не может быть выведена из других как теорема. Этому требованию аксиомы Вейля не удовлетворяют. В самом деле, если к полной системе аксиом добавить новые аксиомы, не описывающие новых основных понятий, то она станет либо противоречивой, либо зависимой — в противном случае можно было бы указать две неизоморфные друг другу модели исходной, полной системы аксиом. Из этого, в частности, следует, что аксиомы III группы не могут быть аксиомами, поскольку новых основных понятий они не описывают. Более того, даже вопрос о независимости аксиом I группы (линейного пространства) оказывается не вполне ясным. Как мы увидим в конце курса, при наиболее естественной трактовке аксиомы 4° аксиома 1° является теоремой, а не аксиомой!

4. Линейное дополнение.

Определение 1. *Пересечением $\mathbf{L}_1 \cap \mathbf{L}_2$ двух линейных подпространств $\mathbf{L}_1$ и $\mathbf{L}_2$ линейного пространства $\mathbf{L}^n$ называется множество всех векторов пространства $\mathbf{L}^n$, принадлежащих как $\mathbf{L}_1$, так и $\mathbf{L}_2$.*

Теорема 1. *Пересечение двух линейных подпространств является линейным подпространством.*

Доказательство. Если $\mathbf{x}, \mathbf{y} \in \mathbf{L}_1 \cap \mathbf{L}_2$, то $\mathbf{x}, \mathbf{y} \in \mathbf{L}_1$ и $\mathbf{x}, \mathbf{y} \in \mathbf{L}_2$, поэтому $\mathbf{x}+\mathbf{y} \in \mathbf{L}_1$ и $\mathbf{x}+\mathbf{y} \in \mathbf{L}_2$, т. е. $\mathbf{x}+\mathbf{y} \in \mathbf{L}_1 \cap \mathbf{L}_2$; если $\mathbf{x} \in \mathbf{L}_1 \cap \mathbf{L}_2$, то $\mathbf{x} \in \mathbf{L}_1$ и $\mathbf{x} \in \mathbf{L}_2$, поэтому $\lambda\mathbf{x} \in \mathbf{L}_1$ и $\lambda\mathbf{x} \in L_2$, т. е. $\lambda\mathbf{x} \in \mathbf{L}_1 \cap \mathbf{L}_2$. Теорема доказана.

Теорема 2. *Пусть* $\mathbf{e}_1, \ldots, \mathbf{e}_k$ — *базис подпространства* $\mathbf{L}_1$, $\mathbf{g}_1, \ldots, \mathbf{g}_m$ — *базис подпространства* $\mathbf{L}_2$. *Векторы* $\mathbf{e}_1, \ldots, \mathbf{e}_k$, $\mathbf{g}_1, \ldots$ $\ldots, \mathbf{g}_m$ *линейно независимы тогда и только тогда, когда* $\mathbf{L}_1 \cap \mathbf{L}_2 = \mathbf{o}$.

Доказательство. Рассмотрим равенство

$$\lambda_1 \mathbf{e}_1 + \ldots + \lambda_k \mathbf{e}_k = \mu_1 \mathbf{g}_1 + \ldots + \mu_m \mathbf{g}_m. \tag{1}$$

Левая часть этого равенства представляет собой вектор $\mathbf{x}_1$ пространства $\mathbf{L}_1$, а правая — вектор $\mathbf{x}_2$ пространства $\mathbf{L}_2$. Если векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{g}_1, \ldots, \mathbf{g}_m$ линейно независимы, то равенство (1) возможно лишь при $\lambda_i = \mu_j = 0$, т. е. когда $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{o}$, а значит, $\mathbf{L}_1 \cap \mathbf{L}_2 = \mathbf{o}$. Обратно, если $\mathbf{L}_1 \cap \mathbf{L}_2 = \mathbf{o}$, то равенство $\mathbf{x}_1 = \mathbf{x}_2$ возможно лишь при $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{o}$, т. е. равенство (1) возможно лишь при $\lambda_i = \mu_j = 0$. Следовательно, векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{g}_1, \ldots, \mathbf{g}_m$ линейно независимы. Теорема доказана.

Определение 2. *Говорят, что пространство* $\mathbf{L}^n$ *представляет собой прямую сумму* $\mathbf{L}_1 \oplus \mathbf{L}_2$ *двух своих линейных подпространств* $\mathbf{L}_1$ *и* $\mathbf{L}_2$, *если любой вектор* $\mathbf{x}$ *пространства* $\mathbf{L}^n$ *может быть единственным образом представлен в виде суммы* $\mathbf{x}_1 + \mathbf{x}_2$, *где* $\mathbf{x}_1 \in \mathbf{L}_1$, $\mathbf{x}_2 \in \mathbf{L}_2$.

Теорема 3. $\mathbf{L}^n = \mathbf{L}_1 \oplus \mathbf{L}_2$ *тогда и только тогда, когда совокупность базисов подпространств* $\mathbf{L}_1$ *и* $\mathbf{L}_2$ *образует базис пространства* $\mathbf{L}^n$.

Доказательство. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_k$ — базис пространства $\mathbf{L}_1$, $\mathbf{g}_1, \ldots, \mathbf{g}_m$ — базис пространства $\mathbf{L}_2$.

Если $\mathbf{L}^n = \mathbf{L}_1 \oplus \mathbf{L}_2$, то любой вектор $\mathbf{x}$ пространства $\mathbf{L}^n$ может быть единственным образом представлен в виде суммы $\mathbf{x}_1 + \mathbf{x}_2$, где $\mathbf{x}_1 \in \mathbf{L}_1, \mathbf{x}_2 \in \mathbf{L}_2$, т. е. единственным образом разложен по векторам $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{g}_1, \ldots, \mathbf{g}_m$. Следовательно, эти векторы образуют базис пространства $\mathbf{L}^n$ (п. 4 § 4 гл. 1 ч. 2).

Обратно, если векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{g}_1, \ldots, \mathbf{g}_m$ образуют базис пространства $\mathbf{L}^n$, то любой вектор $\mathbf{x}$ этого пространства может быть единственным образом по ним разложен, т. е. представлен в виде суммы $\mathbf{x}_1 + \mathbf{x}_2$, где $\mathbf{x}_1 \in \mathbf{L}_1, \mathbf{x}_2 \in \mathbf{L}_2$. Теорема доказана.

Следствие. $\mathbf{L}^n = \mathbf{L}_1 \oplus \mathbf{L}_2$ *тогда и только тогда, когда* $\mathbf{L}_1$ *и* $\mathbf{L}_2$ *обладают двумя свойствами*: 1° $\dim \mathbf{L}_1 + \dim \mathbf{L}_2 = n$; 2° $\mathbf{L}_1 \cap \mathbf{L}_2 = \mathbf{o}$.

В самом деле, если $\mathbf{L}^n = \mathbf{L}_1 \oplus \mathbf{L}_2$, то векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{g}_1, \ldots, \mathbf{g}_m$ образуют базис пространства $\mathbf{L}^n$, поэтому их количество равно n и они линейно независимы, т. е. $\mathbf{L}_1 \cap \mathbf{L}_2 = \mathbf{o}$. Обратно, если $\mathbf{L}_1 \cap \mathbf{L}_2 = \mathbf{o}$ и $\dim \mathbf{L}_1 + \dim \mathbf{L}_2 = n$, то векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{g}_1, \ldots, \mathbf{g}_m$ линейно независимы и их количество равно n, поэтому они образуют базис пространства $\mathbf{L}^n$. Следовательно, $\mathbf{L}^n = \mathbf{L}_1 \oplus \mathbf{L}_2$.

Определение 3. *Пространство* $\overline{L}$ *называется линейным дополнением подпространства* $\mathbf{L}$ *пространства* $\mathbf{L}^n$, *если* $\mathbf{L}^n = \mathbf{L} \oplus \overline{\mathbf{L}}$.

Теорема 4. *Любое подпространство* $\mathbf{L}$ *пространства* $\mathbf{L}^n$ *имеет линейное дополнение.*

Доказательство. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_k$ — базис пространства $\mathbf{L}$. Воспользуемся методом математической индукции. Если $k = n$, то $\overline{L} = \mathbf{o}$. Допустим, что теорема уже доказана для случая $k = m + 1 \leqslant n$ и докажем, что тогда она верна и в случае $k = m < n$.

Поскольку $k < n$, то в пространстве $\mathbf{L}^n$ существует такой вектор $\mathbf{e}_{k+1}$, что векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}$ линейно независимы (в противном случае $\dim \mathbf{L}^n = k < n$) и, следовательно, образуют базис в линейной оболочке $\mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1})$. Так как $\dim \mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}) = k + 1 = m+1$, то по предположению индукции в $\mathbf{L}^n$ существует линейное дополнение пространства $\mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1})$. Пусть $\mathbf{e}_{k+2}, \ldots, \mathbf{e}_n$ — базис в нем. Тогда по теореме 3: а) $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в $\mathbf{L}^n$; б) $\mathbf{L}^n = \mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_k) \oplus \mathbf{L}(\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$, т. е. $\mathbf{L}(\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$ — линейное дополнение пространства $\mathbf{L} = \mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$. Теорема доказана.

Следствие. *Любую совокупность линейно независимых векторов можно дополнить до базиса во всем пространстве* (для этого достаточно построить линейное дополнение к линейной оболочке этих векторов и найти в нем какой-нибудь базис).

Замечание. Обратим особое внимание на то, что *линейное дополнение подпространства* $\mathbf{L}$ *определяется неоднозначно*. В самом деле, если, например, $\mathbf{e}_1$, $\mathbf{e}_2$ — базис пространства $\mathbf{L}^2$, то при любом λ имеем $\mathbf{L}^2 = \mathbf{L}(\mathbf{e}_1) \oplus \mathbf{L}(\mathbf{e}_2 + \lambda\mathbf{e}_1)$ (поскольку векторы $\mathbf{e}_1$ и $\mathbf{e}_2 + \lambda\mathbf{e}_1$ линейно независимы), т. е. $\mathbf{L}(\mathbf{e}_2 + \lambda\mathbf{e}_1)$ — линейное дополнение пространства $\mathbf{L}(\mathbf{e}_1)$. В связи с этим возникает вопрос: существует ли простой алгоритм построения какого-нибудь из линейных дополнений данного подпространства? Такой алгоритм существует, однако описать его мы пока не можем. К этому вопросу мы вернемся в гл. 4.

Глава 2
ЛИНЕЙНЫЕ ОПЕРАТОРЫ

§ 1. Операторы, действующие из $\mathbf{L}^n$ в $\mathbf{L}^m$

1. Линейный оператор.

Определение 1. *Линейным оператором A, действующим из линейного пространства $\mathbf{L}^n$ в линейное пространство $\mathbf{L}^m$, называется такое отображение $\mathbf{L}^n$ в $\mathbf{L}^m$, при котором*:

1° *для любых $\mathbf{x}, \mathbf{y} \in \mathbf{L}^n$ имеет место равенство $A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$*;

2° *для любого $\mathbf{x} \in \mathbf{L}^n$ и любого числа λ имеет место равенство $A(\lambda\mathbf{x}) = \lambda A(\mathbf{x})$.*

Из этого определения, очевидно, следует, что для любых $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{L}^n$ и любых чисел $\lambda_1, \ldots, \lambda_k$ имеет место равенство $A(\lambda_1\mathbf{x}_1 + \ldots + \lambda_k\mathbf{x}_k) = \lambda_1 A(\mathbf{x}_1) + \ldots + \lambda_k A(\mathbf{x}_k)$.

Приведем два примера линейных операторов.

Пример 1. Линейная функция $\mathbf{y} = a\mathbf{x}$ является, очевидно, примером линейного оператора, действующего из $\mathbb{R}^1$ в $\mathbb{R}^1$ (или из $\mathbb{C}^1$ в $\mathbb{C}^1$).

Пример 2. Рассмотрим произвольную $(m \times n)$-матрицу A. Совокупность равенств $y_i = \sum\limits_{s=1}^{n} a_{is}x_s$ $(i = 1, \ldots, m)$ определяет линейный оператор, действующий из $\mathbb{R}^n$ в $\mathbb{R}^m$ (или из $\mathbb{C}^n$ в $\mathbb{C}^m$). В самом деле, $\sum\limits_{s=1}^{n} a_{is}(x_s + \widetilde{x}_s) = \sum\limits_{s=1}^{n} a_{is}x_s + \sum\limits_{s=1}^{n} a_{is}\widetilde{x}_s$ и $\sum\limits_{s=1}^{n} a_{is}\lambda x_s = \lambda \sum\limits_{s=1}^{n} a_{is}x_s$. Как мы вскоре увидим, это пример является в определенном смысле универсальным.

Определение 2. *Суммой линейных операторов A и B, действующих из $\mathbf{L}^n$ в $\mathbf{L}^m$, называется отображение C $\mathbf{L}^n$ в $\mathbf{L}^m$ вида*: $C(\mathbf{x}) = A(\mathbf{x}) + B(\mathbf{x})$.

Поскольку $C(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) + B(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + B(\mathbf{x}) + A(\mathbf{y}) + B(\mathbf{y}) = C(\mathbf{x}) + C(\mathbf{y})$ и $C(\lambda\mathbf{x}) = A(\lambda\mathbf{x}) + B(\lambda\mathbf{x}) = \lambda A(\mathbf{x}) + \lambda B(\mathbf{x}) = \lambda C(\mathbf{x})$, то $C(\mathbf{x})$ — линейный оператор.

Определение 3. *Произведением оператора A, действующего из $\mathbf{L}^n$ в $\mathbf{L}^m$, на число λ, называется отображение B $\mathbf{L}^n$ в $\mathbf{L}^m$ вида*: $B(\mathbf{x}) = \lambda A(\mathbf{x})$.

Так как $B(\mathbf{x}+\mathbf{y}) = \lambda A(\mathbf{x}+\mathbf{y}) = \lambda A(\mathbf{x}) + \lambda A(\mathbf{y}) = B(\mathbf{x}) + B(\mathbf{y})$ и $B(\mu\mathbf{x}) = \lambda A(\mu\mathbf{x}) = \lambda\mu A(\mathbf{x}) = \mu B(\mathbf{x})$, то $B(\mathbf{x})$ — линейный оператор.

Теорема. *Множество $A(\mathbf{L}^n, \mathbf{L}^m)$ всех линейных операторов, действующих из $\mathbf{L}^n$ в $\mathbf{L}^m$, является линейным пространством.*

Доказательство. Роль $\mathbf{o}$ в множестве $A(\mathbf{L}^n, \mathbf{L}^m)$ играет *нулевой оператор* O, действующий по правилу $O(\mathbf{x}) = \mathbf{o}$, так как $A(\mathbf{x}) + O(\mathbf{x}) = A(\mathbf{x})$. Роль $(-A)$ играет оператор $(-1)A$, так как $A(\mathbf{x}) + (-1)A(\mathbf{x}) = \mathbf{o} = O(\mathbf{x})$. Все остальные аксиомы линейного пространства также выполнены, так как при фиксированном значении аргумента складываются и умножаются на число векторы из $\mathbf{L}^m$.

2. Матрица линейного оператора. Пусть A — линейный оператор, действующий из $\mathbf{L}^n$ в $\mathbf{L}^m$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис пространства $\mathbf{L}^n$, $\mathbf{g}_1, \ldots, \mathbf{g}_m$ — базис пространства $\mathbf{L}^m$, $\mathbf{x}$ — произвольный вектор пространства $\mathbf{L}^n$, $\mathbf{y} = A(\mathbf{x})$. Имеем: $\mathbf{y} = A\left(\sum\limits_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum\limits_{j=1}^{n} A(\mathbf{e}_j)\, x_j$. Каждый из векторов $A(\mathbf{e}_j)$, будучи элементом пространства $\mathbf{L}^m$, может быть разложен по базису этого пространства:

$$A(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij} \mathbf{g}_i.$$

Матрица, состоящая из чисел a_{ij}, называется *матрицей линейного оператора* A *в базисах* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{g}_1, \ldots, \mathbf{g}_m$.

Итак, $\mathbf{y} = \sum\limits_{i,j} a_{ij} x_j \mathbf{g}_i = \sum\limits_{i=1}^{m} y_i \mathbf{g}_i$, откуда $\sum\limits_{i=1}^{m} \left(y_i - \sum\limits_{j=1}^{n} a_{ij} x_j\right) \mathbf{g}_i = \mathbf{o}$. Поскольку векторы $\mathbf{g}_1, \ldots, \mathbf{g}_m$ линейно независимы, то это равенство эквивалентно равенствам

$$y_i = \sum_{j=1}^{n} a_{ij} x_j. \tag{1}$$

Замечание 1. Из определения матрицы линейного оператора следует, что ее первый столбец — это координаты вектора $A(\mathbf{e}_1)$, второй — это координаты вектора $A(\mathbf{e}_2)$ и т. д.

Замечание 2. Мы установили, что при выбранных базисах $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{g}_1, \ldots, \mathbf{g}_m$ каждому линейному оператору соответствует $(m \times n)$-матрица, столбцами которой являются координаты векторов $A(\mathbf{e}_i)$. Ранее (см. пример 2 п. 1) было установлено, что каждой $(m \times n)$-матрице соответствует линейный оператор, определяемый формулой (1). Таким образом, при выбранных базисах $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{g}_1, \ldots, \mathbf{g}_m$ соответствие между линейными операторами и $(m \times n)$-матрицами является взаимно однозначным.

Замечание 3. Если договориться записывать координаты векторов в виде столбцов, то формулу (1) можно записать так: $\mathbf{y} = A\mathbf{x}$, где $A\mathbf{x}$ — произведение матриц A и $\mathbf{x}$. С учетом замечания 2 это дает возможность обозначать сам оператор и его матрицу одной и той же буквой, а вместо записи $A(\mathbf{x})$ использовать запись $A\mathbf{x}$.

Теорема. *Пространство* $A(\mathbf{L}^n, \mathbf{L}^m)$ *и множество всех* $(m \times n)$*-матриц изоморфны.*

Доказательство. Выберем какой-нибудь базис пространства $\mathbf{L}^n$ и какой-нибудь базис пространства $\mathbf{L}^m$. Тогда, согласно замечанию 2,

будет установлено взаимно однозначное соответствие между линейными операторами из пространства $A(\mathbf{L}^n, \mathbf{L}^m)$ и $(m \times n)$-матрицами. При этом сумме матриц A и B соответствует оператор C, действующий по правилу: $C\mathbf{x} = (A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$, т. е. оператор $A+B$. Матрице λA соответствует оператор B, действующий по правилу: $B\mathbf{x} = (\lambda A)\mathbf{x} = \lambda A\mathbf{x}$, т. е. оператор λA. Следовательно, указанное соответствие является изоморфизмом. Теорема доказана.

Следствие. $\dim A(\mathbf{L}^n, \mathbf{L}^m) = mn$.

3. Образ оператора.

Определение 1. *Множество значений линейного оператора A называется образом этого оператора и обозначается так*: $\operatorname{im} A$.

Таким образом, $\operatorname{im} A = A(\mathbf{L}^n)$.

Замечание. Ясно, что если A — линейный оператор, действующий из $\mathbf{L}^n$ в $\mathbf{L}^m$, то $\operatorname{im} A \subseteq \mathbf{L}^m$ (т. е. $\operatorname{im} A$ является подмножеством $\mathbf{L}^m$). Более того, $\operatorname{im} A$ — это линейное подпространство $\mathbf{L}^m$. В самом деле:

1° если $\mathbf{y}_1, \mathbf{y}_2 \in \operatorname{im} A$, т. е. $\mathbf{y}_1 = A\mathbf{x}_1$ и $\mathbf{y}_2 = A\mathbf{x}_2$ при некоторых $\mathbf{x}_1, \mathbf{x}_2$, то $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{y}_1 + \mathbf{y}_2$, т. е. $\mathbf{y}_1 + \mathbf{y}_2 \in \operatorname{im} A$;

2° если $\mathbf{y} \in \operatorname{im} A$, т. е. $\mathbf{y} = A\mathbf{x}$, то $A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda\mathbf{y}$, т. е. $\lambda\mathbf{y} \in \operatorname{im} A$.

Теорема. $\dim \operatorname{im} A = \operatorname{rang} A$.

Доказательство. Пусть A — линейный оператор, действующий из $\mathbf{L}^n$ в $\mathbf{L}^m$. Имеем: $A(x) = A\Big(\sum\limits_{i=1}^{n} x_i e_i\Big) = \sum\limits_{i=1}^{n} x_i A(e_i)$, поэтому $\operatorname{im} A = \mathbf{L}(A(\mathbf{e}_1), \ldots, A(\mathbf{e}_n))$, т. е. $\operatorname{im} A$ — это линейная оболочка столбцов матрицы A, которая, очевидно, совпадает с линейной оболочкой базисных столбцов. Поскольку базисные столбцы линейно независимы и их количество равно рангу матрицы A, то $\dim \operatorname{im} A = \operatorname{rang} A$. Теорема доказана.

Следствие. $\dim \operatorname{im} A \leqslant n$.

Определение 2. *Рангом линейного оператора называется ранг его матрицы.*

Из доказанной теоремы следует, что ранг оператора, будучи равным размерности его образа, не зависит от выбора базисов.

Замечание. Пусть $\overline{\operatorname{im}} A$ — линейное дополнение образа оператора A, действующего из $\mathbf{L}^n$ в $\mathbf{L}^m$. Каждый ненулевой вектор $\mathbf{y}$ этого пространства не имеет прообраза. Иными словами, система уравнений $A\mathbf{x} = \mathbf{y}$ не имеет решений (в противном случае оказалось бы, что $\mathbf{y} \in \operatorname{im} A$, чего не может быть, поскольку $\overline{\operatorname{im}} A \cap \operatorname{im} A = \mathbf{o}$). Ясно, что $\dim \overline{\operatorname{im}} A + \dim \operatorname{im} A = m$.

4. Ядро оператора.

Определение. *Множество всех аргументов $\mathbf{x}$ линейного оператора $A\mathbf{x}$, для которых $A\mathbf{x} = \mathbf{o}$, называется ядром этого оператора и обозначается так*: $\ker A$.

Замечание. В арифметическом пространстве $\ker A$ — это множество решений однородной системы линейных уравнений $A\mathbf{x} = \mathbf{o}$. Следовательно, $\ker A$ — это линейное подпространство того пространства, из которого действует оператор A.

Теорема. *Пусть A — линейный оператор, действующий из $\mathbf{L}^n$ в $\mathbf{L}^m$. Тогда* $\dim \operatorname{im} A + \dim \ker A = n$.

Доказательство. $\dim \ker A$ — это размерность пространства решений однородной системы линейных уравнений $A\mathbf{x} = \mathbf{o}$. Следовательно, $\dim \ker A = n - \operatorname{rang} A = n - \dim \operatorname{im} A$. Теорема доказана.

Замечание. Пусть $\overline{\ker} A$ — линейное дополнение ядра оператора A, действующего из $\mathbf{L}^n$ в $\mathbf{L}^m$. Для каждого ненулевого вектора $\mathbf{x}$ этого пространства имеет место неравенство $A\mathbf{x} \neq \mathbf{o}$ (в противном случае оказалось бы, что $\mathbf{x} \in \ker A$, чего не может быть, поскольку $\overline{\ker} A \cap \ker A = \mathbf{o}$). Ясно, что $\dim \overline{\ker} A + \dim \ker A = n$, поэтому из доказанной теоремы следует равенство $\dim \overline{\ker} A = \dim \operatorname{im} A$.

5. Произведение операторов.

Определение. *Пусть B — линейный оператор, действующий из $\mathbf{L}^n$ в $\mathbf{L}^k$, A — линейный оператор, действующий из $\mathbf{L}^k$ в $\mathbf{L}^m$. Произведением AB называется отображение C $\mathbf{L}^n$ в $\mathbf{L}^m$ вида*: $C(\mathbf{x}) = A(B\mathbf{x})$.

Поскольку $AB(\mathbf{x}+\mathbf{y}) = A(B\mathbf{x}+B\mathbf{y}) = AB(\mathbf{x}) + AB(\mathbf{y})$ и $AB(\lambda\mathbf{x}) = A(\lambda B\mathbf{x}) = \lambda AB(\mathbf{x})$, то AB — линейный оператор.

Теорема 1. *Матрица оператора AB равна AB.*

Доказательство. Пусть $\mathbf{y} = AB(\mathbf{x})$. Имеем:

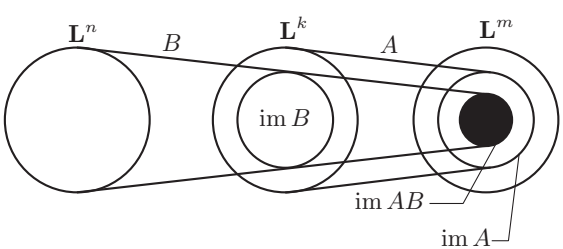
$$y_i = \sum_{s=1}^{k} a_{is}\left(\sum_{r=1}^{n} b_{sr}x_r\right) = \sum_{r=1}^{n}\left(\sum_{s=1}^{k} a_{is}b_{sr}\right)x_r.$$

Теорема доказана.

Следствие. *Произведение операторов обладает всеми свойствами произведения матриц.*

Теорема 2. $\operatorname{rang} AB \leqslant \operatorname{rang} A$ и $\operatorname{rang} AB \leqslant \operatorname{rang} B$.

Доказательство. Справедливость этого утверждения легко усматривается из следующего схематического рисунка.



В самом деле, $\operatorname{im} AB$ — это то, во что оператор A переводит $\operatorname{im} B$. Следовательно, $\dim \operatorname{im} AB \leqslant \dim \operatorname{im} B$ (см. следствие из теоремы п. 3),

т. е. $\operatorname{rang} AB \leqslant \operatorname{rang} B$. Далее, $\operatorname{im} B \subseteq \mathbf{L}^k$, а значит, $\operatorname{im} AB \subseteq \operatorname{im} A$. Поэтому $\dim \operatorname{im} AB \leqslant \dim \operatorname{im} A$, т. е. $\operatorname{rang} AB \leqslant \operatorname{rang} A$. Теорема доказана.

Замечание. Пользуясь рисунком, нетрудно найти точное значение разности $\operatorname{rang} B - \operatorname{rang} AB$. Действительно, если рассматривать оператор A как действующий из $\operatorname{im} B$ в $\operatorname{im} AB$, то (согласно теореме п. 4) можно написать так: $\dim \operatorname{im} A|_{\operatorname{im} B} + \dim \ker A|_{\operatorname{im} B} = \dim \operatorname{im} B$. Но $\dim \operatorname{im} A|_{\operatorname{im} B} = \dim \operatorname{im} AB = \operatorname{rang} AB$, $\dim \ker A|_{\operatorname{im} B} = \dim(\ker A \cap \operatorname{im} B)$, $\dim \operatorname{im} B = \operatorname{rang} B$. Таким образом,

$$\operatorname{rang} B - \operatorname{rang} AB = \dim(\ker A \cap \operatorname{im} B).$$

§ 2. Операторы, действующие из $\mathbf{L}^n$ в $\mathbf{L}^n$

1. Тождественный оператор и обратный оператор. Особого внимания заслуживает случай, когда размерность того пространства, из которого действует линейный оператор, совпадает с размерностью того пространства, в которое он действует. В этом случае обычно считают, что указанные пространства совпадают [1]. Иными словами, под линейным оператором, действующим из $\mathbf{L}^n$ в $\mathbf{L}^n$, понимают оператор, переводящий элементы пространства $\mathbf{L}^n$ в элементы этого же пространства.

Напомним, что для определения матрицы линейного оператора, действующего из $\mathbf{L}^n$ в $\mathbf{L}^m$, необходимо задать два базиса: один — в пространстве $\mathbf{L}^n$, а другой — в пространстве $\mathbf{L}^m$. Если же эти пространства совпадают, то необходимость введения двух базисов отпадает — достаточно одного базиса. Тем самым, матрица линейного оператора, действующего из $\mathbf{L}^n$ в $\mathbf{L}^n$, определяется равенствами: $A(e_j) = \sum\limits_{i=1}^{n} a_{ij} e_i$.

Определение 1. *Отображение E $\mathbf{L}^n$ в $\mathbf{L}^n$ вида $E(\mathbf{x}) = \mathbf{x}$ называется тождественным оператором.*

Ясно, что $E(\mathbf{x})$ — линейный оператор.

Замечание. *Матрица тождественного оператора в любом базисе единичная.* В самом деле, из равенства $A\mathbf{e}_i = \mathbf{e}_i$ следует, что у i-го столбца этой матрицы i-й элемент равен 1, а все остальные элементы равны 0.

Определение 2. *Линейный оператор A^{-1}, действующий из $\mathbf{L}^n$ в $\mathbf{L}^n$, называется обратным по отношению к оператору A, если $AA^{-1} = E$.*

Замечание 1. *Линейным оператором, обратным по отношению к оператору A, является оператор с матрицей A^{-1}*, поскольку произведение этой матрицы на матрицу A должно быть равно E. Из этого следует, что *оператор A^{-1} существует тогда и только тогда, когда* $\det A \neq 0$, т. е. *тогда и только тогда, когда* $\operatorname{rang} A = n$ (или $\ker A = \mathbf{o}$, так как $\dim \ker A = n - \operatorname{rang} A$).

---

[1] Это предположение не ограничивает общности рассмотрения, поскольку все n-мерные пространства изоморфны друг другу.

Замечание 2. Если оператор A действует из $\mathbf{L}^n$ в $\mathbf{L}^n$, то и операторы $A^2 = AA$, $A^3 = AAA$, ... действуют из $\mathbf{L}^n$ в $\mathbf{L}^n$. При этом, очевидно, $A^{p+q} = A^p A^q$. Считают также, что $A^0 = E$ и $A^{-p} = (A^{-1})^p$ (если, конечно, оператор A^{-1} существует).

2. Инвариантные подпространства.

Определение. *Подпространство* $\mathbf{L} \subseteq \mathbf{L}^n$ *называется инвариантным подпространством линейного оператора* A, *действующего из* $\mathbf{L}^n$ *в* $\mathbf{L}^n$, *если* $A\mathbf{x} \in \mathbf{L}$ *для любого* $\mathbf{x} \in \mathbf{L}$.

Иными словами, подпространство $\mathbf{L}$ называется инвариантным подпространством оператора A, если $A(\mathbf{L}) \subseteq \mathbf{L}$. Примерами инвариантных подпространств могут служить $\mathbf{L}^n$ и $\mathbf{o}$.

Теорема 1. *Для любого натурального* k *подпространства* $\ker A^k$, $\operatorname{im} A^k$ *являются инвариантными подпространствами оператора* A.

Доказательство. Если $\mathbf{x} \in \ker A^k$, т. е. $A^k\mathbf{x} = \mathbf{o}$, то $A^k(A\mathbf{x}) = A(A^k\mathbf{x}) = A\mathbf{o} = \mathbf{o}$, поэтому $A\mathbf{x} \in \ker A^k$.

Если $\mathbf{x} \in \operatorname{im} A^k$, т. е. $\mathbf{x} = A^k\mathbf{y}$ для некоторого $\mathbf{y}$, то $A\mathbf{x} = A(A^k\mathbf{y}) = A^k(A\mathbf{y})$, и, следовательно, $A\mathbf{x} \in \operatorname{im} A^k$. Теорема доказана.

Следствие. $\operatorname{im} A^{k+1} \subseteq \operatorname{im} A^k$, поскольку $\operatorname{im} A^{k+1} = A(\operatorname{im} A^k) \subseteq \operatorname{im} A^k$.

Теорема 2. *Если* $\mathbf{L}^n = \mathbf{L}_1 \oplus \mathbf{L}_2$, *причем* $\mathbf{L}_1$ *и* $\mathbf{L}_2$ — *инвариантные подпространства оператора* A, *то в некотором базисе матрица оператора* A *имеет вид*

$$A = \begin{pmatrix} B & O_1 \\ O_2 & C \end{pmatrix},$$

где B — *матрица оператора, действующего из* $\mathbf{L}_1$ *в* $\mathbf{L}_1$, C — *матрица оператора, действующего из* $\mathbf{L}_2$ *в* $\mathbf{L}_2$, *а матрицы* O_1 *и* O_2 *состоят из нулей.*

Доказательство. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_k$ — базис пространства $\mathbf{L}_1$, $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ — базис пространства $\mathbf{L}_2$, и, следовательно, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис пространства $\mathbf{L}^n$. Рассмотрим матрицу оператора A в этом базисе. Поскольку $A(\mathbf{e}_j) = \sum\limits_{i=1}^{n} a_{ij}e_i \in \mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ при всех $j \leqslant k$, то $a_{ij} = 0$ при всех $i > k$ и $j \leqslant k$. Полагая $b_{ij} = a_{ij}$ при $i, j \leqslant k$, получаем матрицу оператора B, действующего из $\mathbf{L}_1$ в $\mathbf{L}_1$, в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_k$. Аналогично находим: $a_{ij} = 0$ при всех $i \leqslant k$ и $j > k$. Полагая $c_{ij} = a_{(i-k)\ (j-k)}$ при $i, j > k$, получаем матрицу оператора C, действующего из $\mathbf{L}_2$ в $\mathbf{L}_2$, в базисе $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$. Теорема доказана.

Замечание. Пусть $\mathbf{L}$ — инвариантное подпространство оператора A. Линейный оператор B, действующий из $\mathbf{L}$ в $\mathbf{L}$ и совпадающий с A на подпространстве $\mathbf{L}$, называется *сужением оператора* A *на инвариантное подпространство* $\mathbf{L}$. Обращаясь к доказанной теореме, можно сказать, что операторы B и C представляют собой сужения оператора A на инвариантные подпространства $\mathbf{L}_1$ и $\mathbf{L}_2$.

3. Образ и ядро. Специфика рассматриваемого случая состоит в том, что ядро и образ оператора A, а значит и линейные дополнения к ним, оказываются лежащими в одном и том же пространстве $\mathbf{L}^n$. При этом $\dim\operatorname{im} A + \dim\ker A = \dim\overline{\operatorname{im}}A + \dim\overline{\ker}A = n$, $\dim\operatorname{im} A = \dim\overline{\ker}A$ и, следовательно, $\dim\ker A = \dim\overline{\operatorname{im}}A$. Это, однако, не означает, что, например, пространство $\mathbf{L}^n$ представляет собой прямую сумму образа и ядра. Изучение взаимного расположения указанных четырех подпространств составит основную цель этого и следующего пунктов. Существенную роль в наших рассуждениях будет играть формула, выведенная в конце п. 5 § 1:

$$\operatorname{rang} B - \operatorname{rang} AB = \dim(\ker A \cap \operatorname{im} B). \tag{1}$$

Теорема 1. $\mathbf{L}^n = \operatorname{im} A \oplus \ker A$ тогда и только тогда, когда $\operatorname{rang} A^2 = \operatorname{rang} A$.

Доказательство. Поскольку $\dim\operatorname{im} A + \dim\ker A = n$, то $\mathbf{L}^n = \operatorname{im} A \oplus \ker A$ тогда и только тогда, когда $\operatorname{im} A \cap \ker A = \mathbf{o}$, т. е. $\dim(\ker A \cap \operatorname{im} A) = \operatorname{rang} A - \operatorname{rang} A^2 = 0$. Теорема доказана.

Теорема 2. *Если при некотором натуральном* k $\operatorname{im} A^{k+1} = \operatorname{im} A^k$, *то* $\operatorname{im} A^{k+2} = \operatorname{im} A^{k+1}$.

Доказательство. Воспользуемся формулой (1). Если $\operatorname{im} A^{k+1} = \operatorname{im} A^k$, то $\dim\operatorname{im} A^k - \dim\operatorname{im} A^{k+1} = \dim(\operatorname{im} A^k \cap \ker A) = 0$. Поэтому $\dim\operatorname{im} A^{k+1} - \dim\operatorname{im} A^{k+2} = \dim(\operatorname{im} A^{k+1} \cap \ker A) = \dim(\operatorname{im} A^k \cap \ker A) = 0$, причем $\operatorname{im} A^{k+2} \subseteq \operatorname{im} A^{k+1}$. Теорема доказана.

Теорема 3. *Если* $\dim\ker A \neq 0$, *то существует такое натуральное число* $m \leqslant n$, что $\dim\operatorname{im} A^{m-1} > \dim\operatorname{im} A^m = \dim\operatorname{im} A^{m+1}$.

Доказательство. Положим $d_i = \dim\operatorname{im} A^i$ ($d_0 = \dim\operatorname{im} E = n$). Поскольку $\operatorname{im} A^{i+1} \subseteq \operatorname{im} A^i$ и $\dim\ker A \neq 0$, то $n = d_0 > d_1 \geqslant d_2 \geqslant \ldots \geqslant 0$. Поэтому существует такое натуральное число $m \leqslant n$, что $d_{m-1} > d_m = d_{m+1}$. Теорема доказана.

Следствие 1. $\mathbf{L}^n = \mathbf{L}_0 \oplus \mathbf{L}$, *где* $\mathbf{L}_0 = \ker A^m$, $\mathbf{L} = \operatorname{im} A^m$.

В самом деле, из равенства $\dim\operatorname{im} A^m = \dim\operatorname{im} A^{m+1}$ и теоремы 2 следует, что $\dim\operatorname{im} A^{m+m} = \dim\operatorname{im} A^m$, т. е. $\operatorname{rang} A^{2m} = \operatorname{rang} A^m$, откуда, согласно теореме 1, и вытекает справедливость этого утверждения.

Следствие 2. *Существует такой базис пространства* $\mathbf{L}^n$, *в котором матрица оператора* A *имеет вид*

$$A = \begin{pmatrix} B & O_1 \\ O_2 & C \end{pmatrix},$$

где B *и* C *— матрицы операторов, являющихся сужениями оператора* A *на инвариантные подпространства* $\mathbf{L}_0$ *и* $\mathbf{L}$ *соответственно, а матрицы* O_1 *и* O_2 *состоят из нулей.*

Это утверждение является следствием только что сделанного наблюдения и двух теорем п. 2.

Замечание 1. Поскольку $\mathbf{L}_0 = \ker A^m$, то для любого вектора $\mathbf{x}$ пространства $\mathbf{L}_0$ $A^m\mathbf{x} = B^m\mathbf{x} = \mathbf{o}$. Следовательно, *оператор* B^m — *нулевой*. С другой стороны, *оператор* B^{m-1} — *ненулевой*. В самом

деле, предполагая противное, получаем: $\dim\ker A^{m-1} \geqslant \dim \mathbf{L}_0 = \dim\ker A^m$, т. е. $\dim\operatorname{im} A^{m-1} \leqslant \dim\operatorname{im} A^m$, что противоречит способу выбора числа m.

Замечание 2. Так как $\mathbf{L} = \operatorname{im} A^m \subseteq \operatorname{im} A$, то *пространство* $\mathbf{L}$ *не содержит векторов дополнения к образу* A. Далее, из равенства $\dim\operatorname{im} A^m = \dim\operatorname{im} A^{m+1}$ и формулы (1) следует, что $\dim(\operatorname{im} A^m \cap \ker A) = \operatorname{rang} A^m - \operatorname{rang} A^{m+1} = 0$, т. е. *пространство* $\mathbf{L}$ *не содержит векторов ядра оператора* A. Это означает, что:

1° $\ker C = \mathbf{o}$ (поскольку на множестве $\mathbf{L}$ операторы A и C совпадают), и, следовательно, $\operatorname{rang} C = \dim \mathbf{L}$, т. е. $\det C \neq 0$;

2° $\ker A = \ker B$.

4. Структура пространства $\mathbf{L}_0$. Обратимся теперь к оператору B, действующему из $\mathbf{L}_0$ в $\mathbf{L}_0$. Поскольку $B^m = O$, то для любого вектора $\mathbf{x}$ пространства $\mathbf{L}_0$ $B^m\mathbf{x} = \mathbf{o}$. Можно сказать иначе: для любого ненулевого вектора $\mathbf{x}$ пространства $\mathbf{L}_0$ существует такое натуральное число $k \leqslant m$, что $B^{k-1}\mathbf{x} \neq \mathbf{o}$, а $B^k\mathbf{x} = \mathbf{o}$. Тем самым каждый ненулевой вектор $\mathbf{x}^k$ порождает *серию ненулевых векторов* $\mathbf{x}^{k-1} = B\mathbf{x}^k$, $\mathbf{x}^{k-2} = B\mathbf{x}^{k-1} = B^2\mathbf{x}^k, \ldots, \mathbf{x}^1 = B\mathbf{x}^2 = B^{k-1}\mathbf{x}^k$ (векторы серии принято нумеровать верхними индексами). При этом $B\mathbf{x}^1 = B^k\mathbf{x}^k = \mathbf{o}$, т. е. *последний из векторов серии принадлежит ядру оператора* B. Если $\mathbf{x}^k \in \operatorname{im} B$, т. е. существует такой вектор $\mathbf{x}^{k+1}$, что $B\mathbf{x}^{k+1} = \mathbf{x}^k$, то серия векторов $\mathbf{x}^i$ может быть им пополнена. Если же $\mathbf{x}^k \notin \operatorname{im} B$, то *вектор* $\mathbf{x}^k$ *называется старшим вектором серии*, *серия — непополнимой*, *а число* k *— длиной серии*, *порожденной старшим вектором* $\mathbf{x}^k$. В частности, серия длины 1 состоит из вектора, принадлежащего ядру и не имеющего прообраза.

Рассмотрим совокупность серий (непополнимых или пополнимых — не имеет значения) векторов x_i^j. Справедливо следующее утверждение.

Теорема 1. *Векторы* x_i^j *линейно независимы тогда и только тогда*, *когда векторы* x_i^1 *— последние векторы серий — линейно независимы.*

Доказательство. Если часть из векторов x_i^j — векторы x_i^1 — линейно зависимы, то и все векторы x_i^j линейно зависимы.

Допустим, что векторы x_i^1 линейно независимы, и некоторая линейная комбинация векторов x_i^j равна $\mathbf{o}$. Действуя на нее оператором B^{m-1} и тем самым превращая ее в некоторую линейную комбинацию векторов x_i^1, мы обнаружим, что все коэффициенты при x_i^m в исходной линейной комбинации равны 0. Далее, действуя на нее оператором B^{m-2}, обнаружим, что и все коэффициенты при x_i^{m-1} равны нулю, и т. д. Таким образом мы установим, что исходная линейная комбинация тривиальна и, следовательно, векторы x_i^j линейно независимы. Теорема доказана.

Основываясь на доказанной теореме, условимся называть серии x_i^j *различными*, если последние векторы этих серий линейно независимы.

Теорема 2. *Количество различных серий длины не меньшей k равно* $\dim(\ker B \cap \operatorname{im} B^{k-1})$.

Доказательство. Каждый ненулевой вектор $\mathbf{x}$ пространства $(\ker B \cap \operatorname{im} B^{k-1})$, будучи вектором $\ker B$, является последним вектором некоторой серии и представляется в виде $\mathbf{x} = B^{k-1}\mathbf{y}$, поэтому длина этой серии не меньше k. С другой стороны, в каждой серии длины не меньшей k есть вектор $\mathbf{x}^k$, для которого $B^{k-1}\mathbf{x}^k = \mathbf{x}^1$ — последний вектор серии — является ненулевым вектором пространства $(\ker B \cap \operatorname{im} B^{k-1})$. Таким образом, это пространство представляет собой множество последних векторов указанных серий. Следовательно, количество различных серий длины не меньшей k, равно максимальному количеству его линейно независимых векторов, т. е. его размерности. Теорема доказана.

Следствие 1. *Количество различных серий длины не меньшей k равно* $\operatorname{rang} B^{k-1} - \operatorname{rang} B^k$ (см. формулу (1)).

Следствие 2. *Количество различных серий длины k равно*

$$(\operatorname{rang} B^{k-1} - \operatorname{rang} B^k) - (\operatorname{rang} B^k - \operatorname{rang} B^{k+1}) = \\ = \operatorname{rang} B^{k+1} + \operatorname{rang} B^{k-1} - 2\operatorname{rang} B^k.$$

Теорема 3. *Векторы всех различных непополнимых серий $\mathbf{e}_i^j$ образуют базис пространства $\mathbf{L}_0$.*

Доказательство. Найдем общее количество N векторов, составляющих всевозможные различные непополнимые серии. Согласно следствию 2 имеем:

$$N = \sum_{k=1}^{m} k\left(\operatorname{rang} B^{k+1} + \operatorname{rang} B^{k-1} - 2\operatorname{rang} B^k\right).$$

Разобьем эту сумму на три суммы и, учитывая, что $B^m = B^{m+1} = O$, $B^0 = E$, преобразуем первые две из них так:

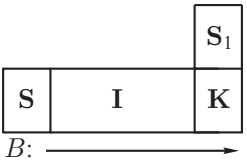
$$\sum_{k=1}^{m} k \operatorname{rang} B^{k+1} = \sum_{k=1}^{m-1} (k-1) \operatorname{rang} B^k,$$

$$\sum_{k=1}^{m} k \operatorname{rang} B^{k-1} = \operatorname{rang} E + \sum_{k=1}^{m-1} (k+1) \operatorname{rang} B^k.$$

Ясно также, что в третьей сумме можно ограничиться суммированием до $m-1$. Вновь объединяя три суммы в одну, получаем:

$$N = \operatorname{rang} E + \sum_{k=1}^{m} \operatorname{rang} B^k\,(k-1+k+1-2k) = \operatorname{rang} E = \dim \mathbf{L}_0.$$

Таким образом, общее количество векторов, составляющих всевозможные различные непополнимые серии, равно размерности пространства $\mathbf{L}_0$, причем, согласно теореме 1, эти векторы линейно независимы. Следовательно, они образуют базис пространства $\mathbf{L}_0$. Теорема доказана.



Замечание 1. В базисе $\mathbf{e}_i^j$ отчетливо видна структура $\operatorname{im} B$, $\ker B$, $\overline{\operatorname{im}} B$ и $\overline{\ker} B$. Действительно, пусть $\mathbf{S}_1$ — линейная оболочка старших векторов серий длины 1, $\mathbf{S}$ — линейная оболочка остальных старших векторов, $\mathbf{K}$ — линейная оболочка векторов ядра, соответствующая сериям длины не меньшей 2, $\mathbf{I}$ — линейная оболочка всех остальных векторов базиса $\mathbf{e}_i^j$.

Тогда $\overline{\operatorname{im}} B = \mathbf{S}_1 \oplus \mathbf{S}$, $\ker B = \mathbf{S}_1 \oplus \mathbf{K}$, $\operatorname{im} B = \mathbf{I} \oplus \mathbf{K}$, $\overline{\ker} B = \mathbf{S} \oplus \mathbf{I}$. С учетом замечания 2 предыдущего пункта можно утверждать также, что $\overline{\operatorname{im}} A = \overline{\operatorname{im}} B$, $\ker A = \ker B$, $\operatorname{im} A = \operatorname{im} B \oplus \mathbf{L}$, $\overline{\ker} A = \overline{\ker} B \oplus \mathbf{L}$. Отметим, что в случае $m = 1$ (длины всех серий равны 1) $\mathbf{S} = \mathbf{I} = \mathbf{K} = \mathbf{o}$ и $\mathbf{L}^n = \operatorname{im} A \oplus \ker A$.

Замечание 2. В построенном базисе действие оператора B задается формулами: $B\mathbf{e}_i^1 = \mathbf{o}$, $B\mathbf{e}_i^j = \mathbf{e}_i^{j-1}$. Поэтому если расположим базисные векторы в порядке $\mathbf{e}_1^1, \mathbf{e}_1^2, \ldots, \mathbf{e}_2^1, \mathbf{e}_2^2, \ldots$, то матрица оператора B в этом базисе примет вид

$$\begin{pmatrix} 0 & \mu_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \mu_2 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & \mu_{k-1} \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix},$$

где числа μ_i равны либо 1, либо 0.

Замечание 3. Теорема 3 не дает конструктивного метода построения базиса из непополнимых серий. Однако такой метод можно указать, опираясь на теорему 2. Он состоит в следующем. Найдем сначала самые длинные серии — серии длины m. С этой целью рассмотрим пространство $(\ker B \cap \operatorname{im} B^{m-1}) = \operatorname{im} B^{m-1}$ (поскольку все векторы $\operatorname{im} B^{m-1}$ лежат в $\ker B$), в котором лежат последние векторы всех серий длины не меньшей m, а значит ровно m, так как серий длины большей m нет. Выберем в $\operatorname{im} B^{m-1}$ произвольным образом базис и выпишем выбранные базисные векторы. Эти векторы являются последними векторами всех различных серий длины m. Далее найдем те векторы пространства $\overline{\ker} B^{m-1}$, которые оператор B^{m-1} переводит в выписанные векторы, т. е. старшие векторы серий длины m. Применяя к ним различные степени оператора B, выпишем все эти серии.

Рассмотрим теперь пространство $(\ker B \cap \operatorname{im} B^{m-2})$. В нем, наряду с найденными последними векторами серий длины m, содержатся последние векторы серий длины $m-1$. Дополнив уже известную нам

часть базиса этого пространства до базиса в нем, мы найдем, тем самым, последние векторы всех различных серий длины $m-1$. Соответствующие им векторы в $\overline{\ker} B^{m-2}$ являются старшими векторами этих серий, а различные степени оператора B, примененные к ним, дают сами серии.

Продолжая этот процесс, мы построим, в конце концов, все различные непополнимые серии, векторы которых, согласно теореме 3, образуют базис пространства $\mathbf{L}_0$. Рассмотрим два примера.

Пример 1. Построить (если это возможно) базис из непополнимых серий оператора

$$B = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & -3 \\ 4 & -1 & 3 & -3 \end{pmatrix}.$$

Решение. Нетрудно проверить, что $B^2 = O$, поэтому искомый базис существует. Пространство $\operatorname{im} B$ состоит из векторов вида

$$B \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 - x_2 \\ 3x_1 + 3x_3 - 3x_4 \\ 4x_1 - x_2 + 3x_3 - 3x_4 \end{pmatrix}.$$

Это пространство двумерно (если из четвертой координаты вычесть третью, то получится вторая, равная первой). Для выбора базиса в нем достаточно, например, взять в первом случае $x_1 = 1$, $x_2 = x_3 = x_4 = 0$, а во втором — $x_1 = x_2 = x_3 = 0$, $x_4 = 1$. В соответствии с этим получаем:

$$\mathbf{e}_1^2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_1^1 = B\mathbf{e}_1^2 = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 4 \end{pmatrix},$$

$$\mathbf{e}_2^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{e}_2^1 = B\mathbf{e}_2^2 = \begin{pmatrix} 0 \\ 0 \\ -3 \\ -3 \end{pmatrix}.$$

В построенном таким образом базисе $\mathbf{e}_1^1$, $\mathbf{e}_1^2$, $\mathbf{e}_2^1$, $\mathbf{e}_2^2$ матрица данного оператора имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Пример 2. Построить базис из непополнимых серий оператора

$$B = \begin{pmatrix} 0 & 0 & 0 & -\frac{\sqrt{6}}{4} & 0 & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\sqrt{6}}{4} & 0 & \frac{\sqrt{2}}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Решение. Имеем:

$$B^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B^3 = O.$$

Пространство $\operatorname{im} B^2$ состоит из векторов вида

$$B^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\sqrt{3}}{2}x_4 + \frac{1}{2}x_6 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Это пространство одномерно, и для выбора базиса в нем достаточно взять $x_6 \neq -\sqrt{3}x_4$, например, положить $x_1 = x_2 = x_3 = x_4 = x_5 = 0$, $x_6{=}1$. В соответствии с этим получаем:

$$\mathbf{e}_1^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{e}_1^2 = B\mathbf{e}_1^3 = \begin{pmatrix} -\sqrt{2}/4 \\ 0 \\ \sqrt{2}/4 \\ 0 \\ \sqrt{3}/2 \\ 0 \end{pmatrix}, \quad \mathbf{e}_1^1 = B^2\mathbf{e}_1^3 = \begin{pmatrix} 0 \\ 1/2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Далее: пространство $(\ker B \cap \operatorname{im} B)$ состоит из векторов вида $B\mathbf{x}$ при условии, что $x_6 = -\sqrt{3}x_4$, т. е. из векторов вида

$$\begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_3 \\ 0 \\ 0 \\ -2x_4 \\ 0 \end{pmatrix}.$$

Чтобы дополнить известную нам часть базиса этого пространства — вектор e_1^1 — до базиса пространства $(\ker B \cap \operatorname{im} B)$, достаточно выбрать $x_4 \neq 0$, например, положить $x_1 = x_2 = x_3 = x_5 = 0$, $x_4 = 1$ (и, следовательно, $x_6 = -\sqrt{3}x_4 = -\sqrt{3}$). Таким образом,

$$\mathbf{e}_2^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ -\sqrt{3} \end{pmatrix}, \quad \mathbf{e}_2^1 = B\mathbf{e}_2^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -2 \\ 0 \end{pmatrix}.$$

Наконец, пространство $(\ker B \cap \operatorname{im} E)$ представляет собой $\ker B$, т. е. определяется системой уравнений

$$B\mathbf{x} = \begin{pmatrix} -\frac{\sqrt{6}}{4}x_4 - \frac{\sqrt{2}}{4}x_6 \\ -\frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_3 \\ \frac{\sqrt{6}}{4}x_4 + \frac{\sqrt{2}}{4}x_6 \\ 0 \\ -\frac{1}{2}x_4 + \frac{\sqrt{3}}{2}x_6 \\ 0 \end{pmatrix} = \mathbf{o}.$$

Отсюда находим $x_4 = x_6 = 0$, $x_1 = x_3$. Таким образом, пространство $\ker B$ состоит из векторов вида

$$\begin{pmatrix} x_1 \\ x_2 \\ x_1 \\ 0 \\ x_5 \\ 0 \end{pmatrix}.$$

Чтобы дополнить известную нам часть базиса этого пространства — векторы e_1^1 и e_2^1 — до базиса $\ker B$, достаточно выбрать $x_1 \neq 0$, например, положить, $x_1 = 1$, $x_2 = x_5 = 0$. Таким образом,

$$\mathbf{e}_3^1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

В построенном базисе $\mathbf{e}_1^1$, $\mathbf{e}_1^2$, $\mathbf{e}_1^3$, $\mathbf{e}_2^1$, $\mathbf{e}_2^2$, $\mathbf{e}_3^1$ матрица данного оператора имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

5. Собственные значения и собственные векторы. Из всевозможных инвариантных подпространств оператора A особый интерес представляют его одномерные инвариантные подпространства — в них $A\mathbf{x} = \lambda\mathbf{x}$.

Определение. *Число* λ *называется собственным значением оператора* A, *если существует ненулевой вектор* $\mathbf{x}$, *для которого* $A\mathbf{x} = \lambda\mathbf{x}$; *вектор* $\mathbf{x}$ *при этом называется собственным вектором оператора* A, *соответствующим собственному значению* λ. *Множество всех собственных значений оператора называется его спектром.*

Примером собственного вектора оператора может служить ненулевой вектор его ядра — он соответствует собственному значению $\lambda = 0$.

Замечание 1. Если $\mathbf{x}$ — собственный вектор, соответствующий собственному значению λ, то $\mu\mathbf{x}$ при $\mu \neq 0$ — также собственный вектор, соответствующий собственному значению λ, поскольку $A\mu\mathbf{x} = \mu A\mathbf{x} = \lambda\mu\mathbf{x}$. Тем самым, собственных векторов, соответствующих одному и тому же собственному значению, бесконечно много.

Замечание 2. Если λ — собственное значение оператора A, то $\lambda + \mu$ — собственное значение оператора $A + \mu E$. В самом деле, пусть $\mathbf{x}$ — собственный вектор оператора A, соответствующий собственному значению λ. Тогда $(A + \mu E)\mathbf{x} = A\mathbf{x} + \mu\mathbf{x} = (\lambda + \mu)\mathbf{x}$, что и требовалось доказать. Таким образом, можно сказать, что *спектр оператора* $A + \mu E$ *представляет собой спектр оператора* A, *«сдвинутый» на* μ.

Теорема. *Собственные векторы, соответствующие различным собственным значениям, линейно независимы.*

Доказательство. Воспользуемся методом математической индукции. Если речь идет об одном собственном значении, то утверждение теоремы очевидно. Допустим, что теорема доказана для k собственных значений, и докажем, что тогда она верна и для $(k+1)$

собственных значений. Предположим, что некоторая линейная комбинация соответствующих этим собственным значениям собственных векторов равна $\mathbf{o}$:

$$\alpha_1\mathbf{x}_1 + \ldots + \alpha_k\mathbf{x}_k + \alpha_{k+1}\mathbf{x}_{k+1} = \mathbf{o}. \tag{2}$$

Докажем, что все $\alpha_i = 0$. Имеем:

$$A(\alpha_1\mathbf{x}_1 + \ldots + \alpha_k\mathbf{x}_k + \alpha_{k+1}\mathbf{x}_{k+1}) =$$
$$= \alpha_1 A(\mathbf{x}_1) + \ldots + \alpha_k A(\mathbf{x}_k) + \alpha_{k+1}A(\mathbf{x}_{k+1}) = \mathbf{o},$$

или $\alpha_1\lambda_1\mathbf{x}_1 + \ldots + \alpha_k\lambda_k\mathbf{x}_k + \alpha_{k+1}\lambda_{k+1}\mathbf{x}_{k+1} = \mathbf{o}$. Вычитая из этого равенства равенство (2), умноженное на λ_{k+1}, получаем:

$$\alpha_1(\lambda_1 - \lambda_{k+1})\mathbf{x}_1 + \ldots + \alpha_k(\lambda_k - \lambda_{k+1})\mathbf{x}_k = \mathbf{o}.$$

По предположению индукции, $\alpha_1 = \ldots = \alpha_k = 0$ ($\lambda_i - \lambda_{k+1} \neq 0$, поскольку все λ_j попарно различны). Но тогда из равенства (2) следует, что и $\alpha_{k+1} = 0$. Теорема доказана.

Следствие. *Количество различных собственных значений оператора не превышает размерности того пространства, в котором он действует* (так как количество соответствующих им линейно независимых собственных векторов не превосходит размерности этого пространства).

Замечание. *Если в пространстве $\mathbf{L}^n$ существует базис из собственных векторов оператора A, то матрица оператора A в этом базисе диагональна*, т. е. $a_{ij} = 0$ при всех $i \neq j$, *причем ее диагональные элементы являются собственными значениями*. В самом деле, в базисе из собственных векторов k-й столбец матрицы A равен $A(\mathbf{e}_k) = \lambda_k\mathbf{e}_k$, т. е. представляет собой вектор, k-я координата которого равна λ_k, а все остальные — нулю.

Обратно, *если в некотором базисе матрица оператора A диагональна, то этот базис состоит из его собственных векторов, а диагональные элементы матрицы являются собственными значениями*, поскольку в этом случае $A(\mathbf{e}_k) = a_{kk}\mathbf{e}_k$.

6. Характеристическое уравнение. Итак, число λ является собственным значением оператора A тогда и только тогда, когда существует вектор $\mathbf{x} \neq \mathbf{o}$, для которого $A\mathbf{x} = \lambda\mathbf{x}$, т. е. $(A - \lambda E)\mathbf{x} = \mathbf{o}$. Однородная система линейных уравнений $(A - \lambda E)\mathbf{x} = \mathbf{o}$ имеет нетривиальные решения $\mathbf{x}$ тогда и только тогда, когда ее определитель равен нулю:

$$\begin{vmatrix} (a_{11} - \lambda) & a_{12} & \ldots & a_{1n} \\ a_{21} & (a_{22} - \lambda) & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & (a_{nn} - \lambda) \end{vmatrix} = 0.$$

Это уравнение n-й степени относительно λ называется *характеристическим уравнением*. Ему удовлетворяет любое собственное значение λ, и, обратно, любое его решение λ является собственным значением.

Замечание 1. Если речь идет о линейном пространстве над полем $\mathbb{R}$, то последняя фраза требует уточнения: любое **вещественное** решение характеристического уравнения является собственным значением. Поскольку, согласно основной теореме алгебры, всякое уравнение n-й степени имеет ровно n корней, в том числе комплексных и совпадающих, то в линейном пространстве над полем $\mathbb{R}$ линейный оператор может вообще не иметь собственных значений, но в n-мерном линейном пространстве над полем $\mathbb{C}$ у него всегда ровно n собственных значений (среди них, конечно, могут быть и совпадающие). В этом состоит самое главное различие между указанными пространствами.

Замечание 2. Отметим, что *в линейном пространстве над полем* $\mathbb{R}$ *каждому комплексному корню характеристического уравнения* (не являющемуся вещественным) *соответствует двумерное инвариантное подпространство соответствующего оператора.*

В самом деле, пусть $\lambda + i\mu$ — один из таких корней ($\mu \neq 0$). Поскольку $\det(A-(\lambda+i\mu)E) = 0$, то однородная система $A\mathbf{z} = (\lambda+i\mu)\mathbf{z}$ имеет нетривиальное решение $\mathbf{z} = \mathbf{x} + i\mathbf{y}$: $A(\mathbf{x} + i\mathbf{y}) = (\lambda + i\mu)(\mathbf{x} + i\mathbf{y})$, или $A\mathbf{x} + iA\mathbf{y} = \lambda\mathbf{x} - \mu\mathbf{y} + i(\lambda\mathbf{y} + \mu\mathbf{x})$, откуда находим:

$$\left.\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} - \mu\mathbf{y} \\ A\mathbf{y} &= \lambda\mathbf{y} + \mu\mathbf{x} \end{aligned}\right\}.$$

Тем самым, $\mathbf{L}(\mathbf{x}, \mathbf{y})$ — инвариантное подпространство оператора A. Осталось доказать, что это подпространство — двумерное, т. е. доказать, что векторы $\mathbf{x}$ и $\mathbf{y}$ линейно независимы. Заметим, что из условия $\mu \neq 0$ следует, что $\mathbf{y} \neq \mathbf{0}$ (иначе из второго уравнения получилось бы, что $\mathbf{x} = \mathbf{o}$, в то время как $|\mathbf{x}|^2 + |\mathbf{y}|^2 \neq 0$). Поэтому из линейной зависимости векторов $\mathbf{x}$ и $\mathbf{y}$ следует существование вещественного числа ν, для которого $\mathbf{x} = \nu\mathbf{y}$. В этом случае наши уравнения преобразуются так:

$$\left.\begin{aligned} \nu A\mathbf{y} &= \lambda\nu\mathbf{y} - m\mathbf{y} \\ A\mathbf{y} &= \lambda\mathbf{y} + \mu\nu\mathbf{y} \end{aligned}\right\},$$

откуда, исключая $A\mathbf{y}$, получаем $\mu+\mu\nu^2 = 0$, чего не может быть ($\mu \neq 0$, а число ν — вещественное). Следовательно, векторы $\mathbf{x}$ и $\mathbf{y}$ линейно независимы.

Замечание 3. Обратимся к результатам пп. 3, 4. Представим пространство $\mathbf{L}^n$, в котором действует оператор A, в виде прямой суммы $\mathbf{L}_0 \oplus \mathbf{L}$, где $\mathbf{L}_0 = \ker\operatorname{im} A^m$, $\mathbf{L} = \operatorname{im} A^m$. Выберем такой базис пространства $\mathbf{L}^n$, в котором матрица оператора A имеет вид

$$A = \begin{pmatrix} B & O_1 \\ O_2 & C \end{pmatrix},$$

где B и C — матрицы операторов, являющихся сужениями оператора A на инвариантные подпространства $\mathbf{L}_0$ и $\mathbf{L}$ соответственно, а матрицы O_1 и O_2 состоят из нулей. При этом $B^m = O$, $\det C \neq 0$. Сверх того, без ограничения общности будем считать, что часть выбранного базиса,

содержащаяся в пространстве $\mathbf{L}_0$, состоит из непополнимых серий и, следовательно, матрица B имеет вид

$$B = \begin{pmatrix} 0 & \mu_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \mu_2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & \mu_{k-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

где числа μ_i равны либо 1, либо 0. Напишем характеристическое уравнение $\det(A - \lambda E) = 0$ и разложим $\det(A - \lambda E)$ сперва по первому столбцу, затем — по второму и т. д., вплоть до k-го. В результате получим: $(-\lambda)^k \det(C - \lambda E) = 0$. Учитывая, что $\det C \neq 0$, мы приходим к следующим выводам:

1° *размерность k пространства $\mathbf{L}_0$ равна кратности корня $\lambda = 0$ характеристического уравнения для оператора A*;

2° *пространство $\mathbf{L}^n$ представляет собой прямую сумму двух инвариантных подпространств $\mathbf{L}_0$ и $\mathbf{L}$, сужение оператора A на первое из которых имеет единственное собственное значение $\lambda = 0$, а на второе — не имеет собственного значения $\lambda = 0$.*

7. Жорданова форма матрицы линейного оператора.

Определение 1. *Матрица Λ_k вида*

$$\Lambda_k = \begin{pmatrix} \lambda_k & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_k & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda_k & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_k \end{pmatrix}$$

называется жордановой клеткой.

В частности, жордановой клеткой является матрица, состоящая из единственного элемента λ_k.

Определение 2. *Говорят, что матрица A имеет жорданову форму, если она имеет вид*

$$A = \begin{pmatrix} \Lambda_1 & & 0 \\ & \Lambda_2 & \\ \dots & \dots & \dots \\ 0 & & \Lambda_m \end{pmatrix},$$

где Λ_i — *жордановы клетки.*

Замечание 1. Можно сказать так: у матрицы A, имеющей жорданову форму, отличными от нуля могут быть только элементы вида a_{ii} и элементы вида a_{ii+1}, причем отличные от нуля элементы вида a_{ii+1} равны 1 и, сверх того, если $a_{ii+1} = 1$, то $a_{i+1i+1} = a_{ii}$. Поэтому, в частности, на диагоналях различных жордановых клеток могут находится и совпадающие числа λ_i.

Замечание 2. Нетрудно видеть, что матрица оператора B, упомянутого в замечании 3 предыдущего пункта, имеет жорданову форму.

Замечание 3. Ясно, что если матрица A имеет жорданову форму, то и матрица $A + \lambda E$ имеет жорданову форму, поскольку она получается из матрицы A прибавлением ко всем ее диагональным элементам одного и того же числа λ.

Замечание 4. Характеристическое уравнение для матрицы линейного оператора, имеющей жорданову форму, записывается так:

$$\prod_{i=1}^{n} (\lambda_i - \lambda) = 0$$

(некоторые из чисел λ_i могут совпадать). Следовательно, числа λ_i являются собственными значениями этого оператора.

Теорема. *Для любого линейного оператора в линейном пространстве* $\mathbf{L}^n$ *над полем* $\mathbb{C}$ *существует такой базис, в котором матрица этого оператора имеет жорданову форму.*

Доказательство. Воспользуемся методом математической индукции. Допустим сначала, что оператор A имеет единственное собственное значение λ. Тогда оператор $(A - \lambda E)$, спектр которого «сдвинут» на λ (см. замечание 2 п. 5), имеет единственное собственное значение 0 кратности n. Следовательно, пространство $\mathbf{L}_0$ для него совпадает с пространством $\mathbf{L}^n$ (см. замечание 3 п. 6), т. е. $(A - \lambda E)^n = O$. В базисе из непополнимых серий матрица оператора $(A - \lambda E)$ имеет жорданову форму. Поэтому и матрица оператора A в этом базисе имеет жорданову форму.

Допустим теперь, что теорема доказана для оператора, имеющего k попарно различных собственных значений, и докажем, что тогда она верна и для оператора A, имеющего $(k+1)$ попарно различных собственных значений. Пусть λ — какое-нибудь его собственное значение. Рассмотрим оператор $(A - \lambda E)$. Пространство $\mathbf{L}^n$ представляет собой прямую сумму двух инвариантных относительно оператора $(A - \lambda E)$ подпространств $\mathbf{L}_0$ и $\mathbf{L}$, сужение этого оператора на первое из которых имеет единственное собственное значение 0, а на второе — не имеет собственного значения 0. Подпространства $\mathbf{L}_0$ и $\mathbf{L}$, будучи инвариантным относительно оператора $(A - \lambda E)$, инвариантны и относительно оператора A, так как $A\mathbf{x} = (A - \lambda E)\mathbf{x} + \lambda\mathbf{x}$. Спектр оператора A представляет собой спектр оператора $(A - \lambda E)$, «сдвинутый» на λ (см. замечание 2 п. 5). Поэтому матрица A при соответствующем выборе базиса может быть записана в виде

$$A = \begin{pmatrix} B & O_1 \\ O_2 & C \end{pmatrix},$$

причем все собственные значения оператора B равны λ, а все собственные значения оператора C отличны от λ, что, в частности, означает, что количество попарно различных из них равно k. Тем самым, матрица оператора B может быть приведена к жордановой форме в силу ранее доказанного, а матрица оператора C — по предположению индукции. Теорема доказана.

Замечание 1. Доказанная теорема верна и в линейном пространстве над полем $\mathbb{R}$, если все корни характеристического уравнения для оператора A вещественны. Чтобы убедиться в этом, достаточно посмотреть на приведенное доказательство под соответствующим углом зрения.

Замечание 2. Вектор $\mathbf{x}^k$, для которого при некотором натуральном k выполняются два соотношения:

1° $(A-\lambda E)^k\mathbf{x}^k \neq \mathbf{o}$,

2° $(A-\lambda E)^{k+1}\mathbf{x}^k = \mathbf{o}$

— называется *присоединенным вектором оператора* A k-го *порядка*, *соответствующим собственному значению* λ. Этот вектор, тем самым, является вектором некоторой серии оператора $(A-\lambda E)$ длины не меньшей $(k+1)$, а собственный вектор оператора A является «присоединенным вектором нулевого порядка». Приведенное доказательство теоремы показывает, что *базис, в котором матрица оператора* A *имеет жорданову форму, состоит из его собственных и присоединенных векторов*. В частности, если оператор A не имеет присоединенных векторов, т. е. если при каждом собственном значении λ длины всех серий оператора $(A-\lambda E)$ равны 1, то указанный базис состоит лишь из собственных векторов оператора A. В таком базисе матрица A диагональна.

Замечание 3. Идею доказательства этой теоремы можно использовать на практике. В самом деле, указанным способом можно сперва построить часть искомого базиса для одного собственного значения, затем — для другого и т. д. Приведем пример. Предположим, что требуется привести матрицу оператора

$$A = \begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$$

к жордановой форме. Будем рассуждать следующим образом. Характеристическое уравнение имеет вид

$$(\lambda-1)^2(\lambda+1)^2 = 0.$$

Оно имеет два корня: $\lambda = 1$ и $\lambda = -1$. Кратность каждого из этих корней равна 2.

Рассмотрим сначала корень $\lambda = 1$. Имеем:

$$A - E = \begin{pmatrix} 2 & -4 & 0 & 2 \\ 4 & -6 & -2 & 4 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 2 & -2 \end{pmatrix}.$$

Ранг этого оператора равен 3 (базисный минор расположен в левом верхнем углу), а значит, размерность его ядра равна 1. Далее,

$$(A-E)^2 = \begin{pmatrix} -12 & 16 & 12 & -16 \\ -16 & 20 & 16 & -20 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 4\begin{pmatrix} -3 & 4 & 3 & -4 \\ -4 & 5 & 4 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

и, следовательно, величина $\dim\ker(A-E)^2$ равна 2 — кратности собственного значения $\lambda=0$ оператора $A-E$. Находим $\ker(A-E)^2$ как решение системы $(A-E)^2\mathbf{x}=\mathbf{o}$:

$$\mathbf{x}=\begin{pmatrix}C_1\\C_2\\C_1\\C_2\end{pmatrix}.$$

Значит, пространство $\operatorname{im}(A-E)\cap\ker(A-E)^2$ состоит из векторов вида

$$(A-E)\begin{pmatrix}C_1\\C_2\\C_1\\C_2\end{pmatrix}=2(C_1-C_2)\begin{pmatrix}1\\1\\1\\1\end{pmatrix}.$$

Это пространства одномерно. Для того чтобы вектор $(A-E)\mathbf{x}$ образовывал в нем базис, достаточно взять C_1 отличным от C_2, например положить $C_1=1$, $C_2=0$. Таким образом,

$$\mathbf{e}_1^2=\begin{pmatrix}1\\0\\1\\0\end{pmatrix},\quad \mathbf{e}_1^1=(A-E)\begin{pmatrix}1\\0\\1\\0\end{pmatrix}=\begin{pmatrix}2\\2\\2\\2\end{pmatrix}.$$

Используя аналогичные рассуждения применительно к корню $\lambda=-1$, получаем:

$$\mathbf{e}_2^2=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \mathbf{e}_2^1=(A+E)\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=\begin{pmatrix}4\\4\\0\\0\end{pmatrix}.$$

В построенном базисе $\mathbf{e}_1^1$, $\mathbf{e}_1^2$, $\mathbf{e}_2^1$, $\mathbf{e}_2^2$ матрица оператора A имеет вид

$$\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&-1&1\\0&0&0&-1\end{pmatrix}.$$

З а м е ч а н и е 4. Если не требуется искать базис, в котором матрица оператора A имеет жорданову форму, а требуется лишь найти эту форму, то можно воспользоваться следующими соображениями: согласно следствию 2 п. 4 количество различных серий из собственных и присоединенных векторов длины k для данного собственного значения λ равно

$$\operatorname{rang}(A-\lambda E)^{k+1}+\operatorname{rang}(A-\lambda E)^{k-1}-2\operatorname{rang}(A-\lambda E)^k.$$

В частности, в рассмотренном примере при $\lambda=\pm1$ $\operatorname{rang}(A-\lambda E)^k$ при $k\geqslant 2$ равен 2, $\operatorname{rang}(A-\lambda E)$ равен 3, $\operatorname{rang}(A-\lambda E)^0$ равен 4, поэтому при каждом из этих значений λ есть лишь одна серия длины 2, что мы и получили.

Глава 3

ПРЕОБРАЗОВАНИЕ БАЗИСОВ И КООРДИНАТ

§ 1. Преобразование базисов

1. Обозначения. При изучении вопросов, связанных с преобразованием координат, часто возникают весьма громоздкие выражения, содержащие разнообразные суммы. Чтобы сделать несколько более компактной запись таких выражений, целесообразно принять следующие соглашения:

1° если какой-либо индекс (например, i) встречается дважды, причем один раз он записан сверху, а другой раз — снизу (например, $a^i b_i$), то предполагается, что по этому индексу производится суммирование, хотя сам знак суммы $\left(\sum\limits_i\right)$ может отсутствовать;

2° элементы базиса обозначаются буквами с нижними индексами.

Тем самым, если мы желаем обойтись без знака суммы, то, например, координаты вектора следует обозначить буквами с верхними индексами — тогда разложение вектора по базису можно будет записать так: $\mathbf{x} = x^i \mathbf{e}_i$. Другой пример — матрица линейного оператора. Если ее элементы обозначать буквами с двумя индексами, один из которых верхний, а другой — нижний, то результат действия оператора можно будет записать так: $y^i = a^i_j x^j$.

Перепишем последнее равенство в виде произведения матриц:

$$\begin{pmatrix} y^1 \\ \dots \\ y^n \end{pmatrix} = \begin{pmatrix} a^1_1 & \dots & a^1_n \\ \dots & \dots & \dots \\ a^n_1 & \dots & a^n_n \end{pmatrix} \begin{pmatrix} x^1 \\ \dots \\ x^n \end{pmatrix}.$$

Таким образом, в этих обозначениях *верхний индекс — это номер строки, а нижний — номер столбца*.

2. Переход к новому базису. Пусть $\{\mathbf{e}_i\}$ и $\{\widetilde{\mathbf{e}}_i\}$ — два базиса. Условимся называть базис $\{\mathbf{e}_i\}$ старым, а базис $\{\widetilde{\mathbf{e}}_i\}$ — новым. Каждый вектор нового базиса можно разложить по старому базису:

$$\widetilde{\mathbf{e}}_i = \alpha^s_i \mathbf{e}_s \tag{1}$$

(напомним, что по индексу s производится суммирование). В этих формулах $\det \alpha \neq 0$, так как столбцы матрицы α — координаты базисных векторов $\widetilde{\mathbf{e}}_i$ — линейно независимы (поскольку эти векторы линейно независимы).

Обратно, если $\det\alpha \neq 0$ и $\{\mathbf{e}_i\}$ — базис, то векторы $\widetilde{\mathbf{e}}_i = \alpha_i^s \mathbf{e}_s$, координаты которых являются линейно независимыми столбцами матрицы α, также линейно независимы и, следовательно, образуют базис.

Таким образом, формулы (1) являются формулами перехода к новому базису (от старого базиса $\{\mathbf{e}_i\}$) тогда и только тогда, когда $\det\alpha \neq 0$. Матрица α называется *матрицей перехода* от базиса $\{\mathbf{e}_i\}$ к базису $\{\widetilde{\mathbf{e}}_i\}$.

Переход от старого базиса к новому иногда называют также *преобразованием базиса*.

3. Последовательные преобразования. Перейдем сначала от базиса $\{\mathbf{e}_i\}$ к базису $\{\widetilde{\mathbf{e}}_i\}$ с помощью матрицы α, а затем от базиса $\{\widetilde{\mathbf{e}}_i\}$ — к базису $\{\mathbf{e}'_i\}$ с помощью матрицы β. Найдем матрицу перехода от базиса $\{\mathbf{e}_i\}$ к базису $\{\mathbf{e}'_i\}$. Имеем: $\mathbf{e}'_i = \beta_i^s \widetilde{\mathbf{e}}_s = \beta_i^s \alpha_s^j \mathbf{e}_j = \left(\alpha_s^j \beta_i^s\right) \mathbf{e}_j$. Таким образом, искомая матрица представляет собой произведение $\alpha\beta$.

В частности, если базисы $\{\mathbf{e}_i\}$ и $\{\mathbf{e}'_i\}$ совпадают, то $\alpha_s^j \beta_i^s = \delta_i^j = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j, \end{cases}$ т.е. $\alpha\beta = E$, а значит, $\beta = \alpha^{-1}$. Итак, *если переход от старого базиса к новому осуществляется с помощью матрицы α, то переход от нового базиса к старому осуществляется с помощью матрицы $\beta = \alpha^{-1}$*:

$$\mathbf{e}_i = \beta_i^s \widetilde{\mathbf{e}}_s. \tag{2}$$

Замечание. Выражение $\delta_i^j = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j \end{cases}$ часто называют *символом Кронекера*.

§ 2. Преобразование координат

1. Преобразование координат вектора. Пусть $\{\mathbf{e}_i\}$ — старый базис, $\{\widetilde{\mathbf{e}}_i\}$ — новый базис, $\widetilde{\mathbf{e}}_i = a_i^s \mathbf{e}_s$, $\mathbf{e}_i = \beta_i^s \widetilde{\mathbf{e}}_s$, где $\beta = \alpha^{-1}$. Выразим координаты произвольного вектора $\mathbf{x}$ в новом базисе через его координаты в старом базисе. Имеем: $\mathbf{x} = x^s \mathbf{e}_s = x^s \beta_s^i \widetilde{\mathbf{e}}_i = \widetilde{x}^i \widetilde{\mathbf{e}}_i$, откуда (в силу линейной независимости векторов $\widetilde{\mathbf{e}}_i$)

$$\widetilde{x}^i = \beta_s^i x^s. \tag{1}$$

Таким образом, координаты вектора преобразуются с помощью матрицы β.

Замечание. Ясно, что $x^i = \alpha_s^i \widetilde{x}^s$.

2. Преобразование матрицы линейного оператора. Выразим теперь матрицу линейного оператора A, действующего из $\mathbf{L}^n$ в $\mathbf{L}^n$, в новом базисе через его матрицу в старом базисе. Имеем:

$$A\left(\widetilde{\mathbf{e}}_i\right) = A\left(\alpha_i^s \mathbf{e}_s\right) = \alpha_i^s A\left(\mathbf{e}_s\right) = \alpha_i^s a_s^r \mathbf{e}_r = \alpha_i^s a_s^r \beta_r^j \widetilde{\mathbf{e}}_j = \widetilde{a}_i^j \widetilde{\mathbf{e}}_j,$$

откуда

$$\widetilde{a}_i^j = \alpha_i^s \beta_r^j a_s^r. \tag{2}$$

Замечание. Полученную формулу можно записать так: $\widetilde{A} = \beta A \alpha$. Следовательно, $\det \widetilde{A} = \det \beta \det A \det \alpha = \det A$, т. е. *определитель матрицы линейного оператора не зависит от выбора базиса.*

3. Линейная форма.

Определение 1. *Числовая функция* (т. е. функция, значениями которой являются числа) *векторного аргумента* $A(\mathbf{x})$ *называется линейной формой, если*:

1° *для любых* $\mathbf{x}$, $\mathbf{y}$ $A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$;

2° *для любого* $\mathbf{x}$ *и любого числа* λ $A(\lambda\mathbf{x}) = \lambda A(\mathbf{x})$.

Пусть $A(\mathbf{x})$ — линейная форма в $\mathbf{L}^n$. Имеем:

$$A(\mathbf{x}) = A(x^i \mathbf{e}_i) = A(\mathbf{e}_i) x^i = a_i x^i.$$

Числа $a_i = A(\mathbf{e}_i)$ называются *коэффициентами линейной формы* в базисе $\{\mathbf{e}_i\}$. Выясним, как они преобразуются при переходе к новому базису. Имеем $\widetilde{a}_i = A(\widetilde{\mathbf{e}}_i) = A(\alpha_i^s \mathbf{e}_s) = \alpha_i^s A(\mathbf{e}_s) = \alpha_i^s a_s$. Итак,

$$\widetilde{a}_i = \alpha_i^s a_s, \tag{3}$$

т. е. коэффициенты линейной формы преобразуются с помощью матрицы α (как базис), а не с помощью матрицы β (как координаты вектора).

Определение 2. *Числовая функция нескольких векторных аргументов называется линейной по одному из них, если при фиксированных значениях остальных аргументов она является линейной формой.*

Определение 3. *Числовая функция двух векторных аргументов называется билинейной формой, если она линейна по каждому из них.*

Пусть $B(\mathbf{x}, \mathbf{y})$ — билинейная форма. Имеем:

$$B(\mathbf{x}, \mathbf{y}) = B(x^i \mathbf{e}_i, y^j \mathbf{e}_j) = B(\mathbf{e}_i, y^j \mathbf{e}_j) x^i = B(\mathbf{e}_i, \mathbf{e}_j) x^i y^j = b_{ij} x^i y^j.$$

Матрица с элементами $b_{ij} = B(\mathbf{e}_i, \mathbf{e}_j)$ называется *матрицей билинейной формы* в базисе $\{\mathbf{e}_i\}$. Выясним, как она преобразуется при переходе к новому базису. Имеем:

$$\widetilde{b}_{ij} = B(\widetilde{\mathbf{e}}_i, \widetilde{\mathbf{e}}_j) = B\left(\alpha_i^s \mathbf{e}_s, \alpha_j^r \mathbf{e}_r\right) = \alpha_i^s B\left(\mathbf{e}_s, \alpha_j^r \mathbf{e}_r\right) =$$
$$= \alpha_i^s \alpha_j^r B(\mathbf{e}_s, \mathbf{e}_r) = \alpha_i^s \alpha_j^r b_{sr}.$$

Итак,

$$\widetilde{b}_{ij} = \alpha_i^s \alpha_j^r b_{sr}. \tag{4}$$

Эта формула похожа на формулу (2) преобразования матрицы линейного оператора, но в ней матрица α фигурирует дважды, а в формуле (2) фигурируют матрицы α и β.

Замечание. Точно так же можно ввести понятие *полилинейной формы* — числовой функции нескольких векторных аргументов, линейной по каждому из них. Рассуждая аналогично, можно назвать числа $p_{i_1\ldots i_k} = P(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_k})$ координатами полилинейной формы в базисе $\{\mathbf{e}_i\}$, а саму форму записать так:

$$P(\mathbf{x}_1, \ldots, \mathbf{x}_k) = p_{i_1\ldots i_k} x^{i_1} \ldots x^{i_k}.$$

Ясно, что при переходе к новому базису координаты полилинейной формы будут преобразовываться следующим образом:

$$\widetilde{p}_{i_1\ldots i_k} = \alpha_{i_1}^{s_1} \ldots \alpha_{i_k}^{s_k} p_{s_1\ldots s_k}. \tag{5}$$

§ 3. Тензоры

1. Определение тензора. Прежде чем дать определение тензора, обсудим несколько вопросов общего характера.

Ранее отмечалось, что теорема об изоморфизме двух линейных пространств одинаковой размерности позволяет при изучении свойств абстрактного линейного пространства реально рассматривать вместо него любую из его интерпретаций, в частности геометрическую (вектор — направленный отрезок) и алгебраическую (вектор — набор чисел). Именно параллельное рассмотрение указанных интерпретаций в наибольшей мере способствует эффективному построению линейной алгебры. Например, то, что размерность образа линейного оператора не превосходит размерности его области определения, отнюдь не очевидно геометрически, но очевидно алгебраически — ранг матрицы не превосходит количества ее строк. С другой стороны, то, что ранг произведения операторов не превосходит рангов сомножителей, почти очевидно геометрически (вспомним рисунок), но далеко не очевидно алгебраически. Идея параллельного рассмотрения геометрической и алгебраической интерпретаций вектора как раз и лежит в основе понятия тензора.

При изучении геометрического объекта методом координат мы фактически заменяем его новым объектом: «геометрический объект + система координат». Поэтому итогом нашего исследования является информация об этом новом объекте, а не об исходном. Тем самым, на последнем этапе всегда возникает задача о разделении полученной информации на чисто геометрическую и ту, которая привнесена специальным выбором системы координат. Иногда эта задача оказывается весьма сложной, порой даже сложнее исходной. Один из способов ее решения состоит в том, чтобы все время оперировать только такими объектами (их и называют тензорами), которые, в определенном смысле, не зависят от выбора системы координат.

Поясним, о чем идет речь. Если интерпретировать вектор как набор чисел, то он, конечно, будет зависеть от выбора системы координат.

Однако вектор — направленный отрезок — от выбора системы координат, очевидно, не зависит. Но с этой точки зрения такие объекты, как линейный оператор, линейная, билинейная и полилинейная форма также не зависят от выбора системы координат — их можно интерпретировать как функции от направленных отрезков. Все эти объекты и называются тензорами.

С алгебраической точки зрения вектор, линейный оператор, линейная, билинейная и полилинейная форма — это наборы чисел, преобразующиеся при переходе к новому базису по формулам (3)–(7). Рассмотрение этих формул приводит нас к следующему определению.

Определение. *Геометрический объект, задаваемый в каждом базисе пространства* $\mathbf{L}^n$ *совокупностью* n^{p+q} *чисел, называется* p *раз ковариантным и* q *раз контравариантным тензором, если при переходе к новому базису эти числа преобразуются так*:

$$\widetilde{t}^{\,j_1\ldots j_q}_{i_1\ldots i_p} = \alpha^{s_1}_{i_1}\ldots\alpha^{s_p}_{i_p}\beta^{j_1}_{r_1}\ldots\beta^{j_q}_{r_q}t^{r_1\ldots r_q}_{\sigma_1\ldots s_p}. \tag{1}$$

В этих формулах матрица α встречается p раз, а матрица β — q раз, чем и объясняется название «p раз ковариантный и q раз контравариантный тензор». Иногда этот тензор называют *тензором типа* (p,q), а число $p+q$ — его *валентностью* или *рангом*. Обратим внимание на то, что числа p и q представляют собой количества нижних и верхних индексов t, что является формальным следствием принятого нами соглашения о суммировании.

Тем самым, можно сказать, что вектор — это тензор типа $(0,1)$, линейный оператор — тензор типа $(1,1)$, линейная форма — тензор типа $(1,0)$, билинейная форма — тензор типа $(2,0)$, полилинейная форма — тензор типа $(k,0)$.

Замечание. Полезно в качестве самостоятельного упражнения убедиться в корректности определения тензора: если сначала перейти от базиса $\{\mathbf{e}_i\}$ к базису $\{\mathbf{e}'_i\}$, а затем от базиса $\{\mathbf{e}'_i\}$ — к базису $\{\widetilde{\mathbf{e}}_i\}$, то результат будет таким же, как при переходе от базиса $\{\mathbf{e}_i\}$ непосредственно к базису $\{\widetilde{\mathbf{e}}_i\}$.

2. Сумма тензоров одинаковой структуры.

Определение. *Суммой* $A+B$ *двух тензоров типа* (p,q) *называется объект* C, *определяемый в каждом базисе пространства* $\mathbf{L}^n$ *совокупностью* n^{p+q} *чисел* $c^{j_1\ldots j_q}_{i_1\ldots i_p} = a^{j_1\ldots j_q}_{i_1\ldots i_p} + b^{j_1\ldots j_q}_{i_1\ldots i_p}$.

Теорема. *Сумма двух тензоров типа* (p,q) *является тензором типа* (p,q).

Доказательство. Имеем:

$$\widetilde{c}^{\,j_1\ldots j_q}_{i_1\ldots i_p} = \widetilde{a}^{\,r_1\ldots r_q}_{s_1\ldots s_p} + \widetilde{b}^{\,r_1\ldots r_q}_{s_1\ldots s_p} =$$

$$= \alpha^{s_1}_{i_1}\ldots\alpha^{s_p}_{i_p}\beta^{j_1}_{r_1}\ldots\beta^{j_q}_{r_q}a^{r_1\ldots r_q}_{s_1\ldots s_p} + \alpha^{s_1}_{i_1}\ldots\alpha^{s_p}_{i_p}\beta^{j_1}_{r_1}\ldots\beta^{j_q}_{r_q}b^{r_1\ldots r_q}_{s_1\ldots s_p} =$$

$$= \alpha^{s_1}_{i_1}\ldots\alpha^{s_p}_{i_p}\beta^{j_1}_{r_1}\ldots\beta^{j_q}_{r_q}\left(a^{r_1\ldots r_q}_{s_1\ldots s_p} + b^{r_1\ldots r_q}_{s_1\ldots s_p}\right) = \alpha^{s_1}_{i_1}\ldots\alpha^{s_p}_{i_p}\beta^{j_1}_{r_1}\ldots\beta^{j_q}_{r_q}c^{r_1\ldots r_q}_{s_1\ldots s_p}.$$

Теорема доказана.

3. Прямое произведение тензоров.

Определение. *Прямым произведением тензора A типа (p_1, q_1) на тензор B типа (p_2, q_2) называется объект C, определяемый в каждом базисе пространства $\mathbf{L}^n$ совокупностью $n^{p_1+q_1+p_2+q_2}$ чисел*

$$c^{j_1\ldots j_{q_1}\ldots j_{q_1+q_2}}_{i_1\ldots i_{p_1}\ldots i_{p_1+p_2}} = a^{j_1\ldots j_{q_1}}_{i_1\ldots i_{p_1}} b^{j_{q_1+1}\ldots j_{q_1+q_2}}_{i_{p_1+1}\ldots i_{p_1+p_2}}.$$

Теорема. *Прямое произведение тензора A типа (p_1, q_1) на тензор B типа (p_2, q_2) является тензором типа $(p_1 + p_2, q_1 + q_2)$.*

Доказательство. $\widetilde{c}^{j_1\ldots j_{q_1}\ldots j_{q_1+q_2}}_{i_1\ldots i_{p_1}\ldots i_{p_1+p_2}} = \widetilde{a}^{j_1\ldots j_{q_1}}_{i_1\ldots i_{p_1}} \widetilde{b}^{j_{q_1+1}\ldots j_{q_1+q_2}}_{i_{p_1+1}\ldots i_{p_1+p_2}} =$

$$= \alpha^{s_1}_{i_1}\ldots\alpha^{s_{p_1}}_{i_{p_1}}\beta^{j_1}_{r_1}\ldots\beta^{j_{q_1}}_{r_{q_1}} a^{r_1\ldots r_{q_1}}_{s_1\ldots s_{p_1}}\alpha^{s_{p_1+1}}_{i_{p_1+1}}\ldots$$

$$\ldots\alpha^{s_{p_1+p_2}}_{i_{p_1+p_2}}\beta^{j_{q_1+1}}_{r_{q_1+1}}\ldots\beta^{j_{q_1+q_2}}_{r_{q_1+q_2}} b^{r_{q_1+1}\ldots r_{q_1+q_2}}_{s_{p_1+1}\ldots s_{p_1+p_2}} =$$

$$= \alpha^{s_1}_{i_1}\ldots\alpha^{s_{p_1}}_{i_{p_1}}\alpha^{s_{p_1+1}}_{i_{p_1+1}}\ldots\alpha^{s_{p_1+p_2}}_{i_{p_1+p_2}}\beta^{j_1}_{r_1}\ldots$$

$$\ldots\beta^{j_{q_1}}_{r_{q_1}}\beta^{j_{q_1+1}}_{r_{q_1+1}}\ldots\beta^{j_{q_1+q_2}}_{r_{q_1+q_2}} a^{r_1\ldots r_{q_1}}_{s_1\ldots s_{p_1}} b^{j_{q_1+1}\ldots j_{q_1+q_2}}_{i_{p_1+1}\ldots i_{p_1+p_2}} =$$

$$= \alpha^{s_1}_{i_1}\ldots\alpha^{s_{p_1}}_{i_{p_1}}\alpha^{s_{p_1+1}}_{i_{p_1+1}}\ldots\alpha^{s_{p_1+p_2}}_{i_{p_1+p_2}}\beta^{j_1}_{r_1}\ldots\beta^{j_{q_1}}_{r_{q_1}}\beta^{j_{q_1+1}}_{r_{q_1+1}}\ldots\beta^{j_{q_1+q_2}}_{r_{q_1+q_2}} c^{r_1\ldots r_{q_1}\ldots r_{q_1+q_2}}_{s_1\ldots s_{p_1}\ldots s_{p_1+p_2}},$$

что и требовалось доказать.

Замечание. Если все компоненты тензора типа (p, q) в каждом базисе умножить на одно и то же число λ, т. е. составить прямое произведение данного тензора и тензора типа $(0, 0)$, то получится тензор типа (p, q). Ясно, что множество всех тензоров типа (p, q) в $\mathbf{L}^n$ образует линейное пространство, так как при фиксированном наборе индексов складываются и умножаются на число обычные числа. Его размерность равна n^{p+q}, поскольку координаты тензора в одном базисе могут быть выбраны произвольно, а в любом другом базисе — вычислены по формулам (9).

4. Свертка тензора.

Определение. *Сверткой тензора типа (p, q), где p, $q \geqslant 1$, по двум индексам, один из которых нижний, а другой — верхний, называется объект, определяемый в каждом базисе пространства $\mathbf{L}^n$ совокупностью n^{p+q-2} чисел, которые строятся так: указанные индексы переобозначаются одной буквой, в результате чего сверху и снизу оказывается один и тот же индекс и, следовательно, по нему производится суммирование.*

Теорема. *Свертка тензора A типа (p, q) является тензором типа $(p - 1, q - 1)$.*

Доказательство. Пусть, например, свертка осуществляется по первым индексам (в остальных случаях доказательство аналогичное). Имеем:

$$\widetilde{b}^{j_2\ldots j_q}_{i_2\ldots i_p} = \widetilde{a}^{sj_2\ldots j_q}_{si_2\ldots i_p} = \alpha^{s_1}_{s}\ldots\alpha^{s_p}_{i_p}\beta^{s}_{r_1}\ldots\beta^{j_q}_{r_q} a^{r_1\ldots r_q}_{s_1\ldots s_p}.$$

Поскольку $\alpha_s^{s_1}\beta_{r_1}^s = \delta_{r_1}^{s_1}$, то из суммы по s_1 остается одно слагаемое — то, в котором $s_1 = r_1$. Тем самым,

$$\widetilde{b}_{i_2\ldots i_p}^{j_2\ldots j_q} = \alpha_{i_2}^{s_2}\ldots\alpha_{i_p}^{s_p}\beta_{r_2}^{s_2}\ldots\beta_{r_q}^{j_q}a_{r_1 s_2\ldots s_p}^{r_1 r_2\ldots r_q} = \alpha_{i_2}^{s_2}\ldots\alpha_{i_p}^{s_p}\beta_{r_2}^{s_2}\ldots\beta_{r_q}^{j_q}b_{s_2\ldots s_p}^{r_2\ldots r_q},$$

что и требовалось доказать.

Замечание 1. Теперь при желании можно полностью перейти на язык тензоров, отказавшись от употребления слов «вектор», «линейная форма», «линейный оператор» и т. д. Приведем несколько примеров.

Рассмотрим тензор a_i типа $(1,0)$ (линейную форму) и тензор x^i типа $(0,1)$ (вектор). Их прямое произведение $a_i x^j$ — это тензор типа $(1,1)$. Следовательно, его свертка $a_i x^i$ — это тензор типа $(0,0)$, т. е. число, не зависящее от выбора базиса (оно представляет собой значение линейной формы A при аргументе $\mathbf{x}$).

Другой пример: если b_{ij} — тензор типа $(2,0)$ (билинейная форма), x^i и y^i — тензоры типа $(0,1)$ (векторы), то $b_{ij}x^i y^j$ — это тензор типа $(0,0)$, т. е. число, не зависящее от выбора базиса (оно представляет собой значение билинейной формы при аргументах $\mathbf{x}$,$\mathbf{y}$).

Аналогично, если a_i^j — тензор типа $(1,1)$ (линейный оператор), x^i — тензор типа $(0,1)$ (вектор), то $a_i^j x^i$ — это тензор типа $(0,1)$ (вектор, представляющий собой результат действия линейного оператора на вектор $\mathbf{x}$).

Замечание 2. Линейный оператор — это тензор типа $(1,1)$, поэтому его свертка a_s^s — это тензор типа $(0,0)$, т. е. число, не зависящее от выбора базиса. Это число обозначается $\operatorname{sp} A$ и называется *следом оператора* A. Оно представляет собой сумму диагональных элементов матрицы оператора. Тем самым, $\operatorname{sp} A$, равно как и $\det A$ (п. 2 § 2), не зависит от выбора базиса.

Это ясно и из других соображений. Собственные значения оператора, очевидно, не зависят от выбора базиса. Следовательно, от выбора базиса не зависят и коэффициенты характеристического уравнения, поскольку они выражаются через корни этого уравнения. Но $\operatorname{sp} A$ — это коэффициент при $\lambda^{(n-1)}$, а $\det A$ — это коэффициент при λ^0.

5. О билинейной форме.

Теорема 1. *Ранг матрицы билинейной формы не зависит от выбора базиса.*

Доказательство. Поскольку $\widetilde{b}_{ij} = a_i^s a_j^r b_{sr}$, т. е. $\widetilde{B} = a^{tr}Ba$, то $\operatorname{rang}\widetilde{B} \leqslant \operatorname{rang} B$. С другой стороны, $B = \beta^{tr}\widetilde{B}\beta$, откуда $\operatorname{rang} B \leqslant \operatorname{rang}\widetilde{B}$. Следовательно, $\operatorname{rang}\widetilde{B} = \operatorname{rang} B$. Теорема доказана.

Определение 1. *Рангом билинейной формы называется ранг ее матрицы.*

Теорема 2. *В линейном пространстве над полем $\mathbb{R}$ знак определителя матрицы билинейной формы не зависит от выбора базиса.*

Доказательство. Поскольку $\widetilde{B} = a^{tr}Ba$, то $\det\widetilde{B} = \det\alpha^2\det B$, причем $\det\alpha \neq 0$. Теорема доказана.

Определение 2. *Билинейная форма* B $(\mathbf{x},\mathbf{y})$ *называется симметричной, если при всех* $\mathbf{x},\mathbf{y}$ *имеет место равенство* $B(\mathbf{y},\mathbf{x}) = B(\mathbf{x},\mathbf{y})$.

Теорема 3. *Билинейная форма симметрична тогда и только тогда, когда ее матрица симметрична*, т. е. $b_{ij} = b_{ji}$.

Доказательство. Если билинейная форма $B(\mathbf{x},\mathbf{y})$ симметрична, то $b_{ij} = B(\mathbf{e}_i,\mathbf{e}_j) = B(\mathbf{e}_j,\mathbf{e}_i) = b_{ji}$. Обратно, если $b_{ij} = b_{ji}$, то $B(\mathbf{y},\mathbf{x}) = b_{ij}y^ix^j = b_{ij}x^jy^i = b_{ji}x^iy^j = b_{ij}x^iy^j = B(\mathbf{x},\mathbf{y})$. Теорема доказана.

Следствие. *Если матрица тензора типа* $(2,0)$ (билинейной формы) *симметрична в одном базисе* (т. е. эта билинейная форма симметрична), *то она симметрична и в любом другом базисе*. Такой тензор называется *симметричным тензором типа* $(2,0)$. Отметим, что *матрица линейного оператора аналогичным свойством не обладает*: если она и окажется в каком-то базисе симметричной, то в другом базисе она, вообще говоря, уже не будет симметричной.

§ 4. Квадратичные формы

1. Матрица квадратичной формы.

Определение 1. *Квадратичной формой называется выражение* $B(\mathbf{x},\mathbf{x})$, *где* $B(\mathbf{x},\mathbf{y})$ *— билинейная форма*.

Теорема. *Каждая квадратичная форма может быть получена из симметричной билинейной формы, причем такая билинейная форма определяется единственным образом*.

Доказательство. Допустим, что данная квадратичная форма $B(\mathbf{x},\mathbf{x})$ получена из билинейной формы $B(\mathbf{x},\mathbf{y})$. Тогда эта же квадратичная форма может быть получена и из симметричной билинейной формы $A(\mathbf{x},\mathbf{y}) = \dfrac{1}{2}(B(\mathbf{x},\mathbf{y}) + B(\mathbf{y},\mathbf{x}))$.

С другой стороны, если одна и та же квадратичная форма получена из двух симметричных билинейных форм, т. е. $A(\mathbf{x},\mathbf{x}) = B(\mathbf{x},\mathbf{x})$, то $A(\mathbf{x},\mathbf{y}) = B(\mathbf{x},\mathbf{y})$. В самом деле, из равенства $A(\mathbf{e}_i + \mathbf{e}_j, \mathbf{e}_i + \mathbf{e}_j) = B(\mathbf{e}_i + \mathbf{e}_j, \mathbf{e}_i + \mathbf{e}_j)$ следует, что $A(\mathbf{e}_i,\mathbf{e}_i) + 2A(\mathbf{e}_i,\mathbf{e}_j) + A(\mathbf{e}_j,\mathbf{e}_j) = B(\mathbf{e}_i,\mathbf{e}_i) + 2B(\mathbf{e}_i,\mathbf{e}_j) + B(\mathbf{e}_j,\mathbf{e}_j)$, откуда $A(\mathbf{e}_i,\mathbf{e}_j) = B(\mathbf{e}_i,\mathbf{e}_j)$, т. е. $a_{ij} = b_{ij}$. Теорема доказана.

Определение 2. *Матрицей квадратичной формы называется матрица той симметричной билинейной формы, из которой она получена*.

Замечание. Так как матрица квадратичной формы в любом базисе совпадает с матрицей соответствующей ей симметричной билинейной формы, то *квадратичная форма — это симметричный тензор типа* $(2,0)$.

2. Метод Лагранжа.

Определение. *Базис линейного пространства над полем* $\mathbb{R}$ $(\mathbb{C})$, *в котором квадратичная форма имеет вид* $A(\mathbf{x},\mathbf{x}) = \lambda_i(x^i)^2$, *где* λ_i *— одно из чисел* 1, -1, 0 $(1,0)$, *называется каноническим*.

Теорема. *Для любой квадратичной формы существует канонический базис.*

Доказательство. Запишем выражение $A(\mathbf{x},\mathbf{x}) = a_{ij}x^ix^j$ в каком-нибудь базисе и воспользуемся методом математической индукции. Если в выражении $a_{ij}x^ix^j$ реально не фигурирует ни одной переменной x^i, т. е. $A(\mathbf{x},\mathbf{x}) \equiv 0$, то данный базис (как, впрочем, и любой другой) является каноническим.

Допустим, что теорема доказана для случая, когда в выражении $a_{ij}x^ix^j$ реально фигурируют $(n-1)$ переменных, и докажем, что тогда она верна и для n реально фигурирующих переменных $x^1, \ldots, x^n$. Возможны два случая.

1°. Коэффициент при $(x^1)^2$, т. е. a_{11}, отличен от нуля. Сгруппируем все слагаемые, в которые входит x^1, и дополним их до полного квадрата. В результате получим

$$A(x,x) = a_{11}\left(x^1 + \frac{a_{12}}{a_{11}}x^2 + \ldots + \frac{a_{1n}}{a_{11}}x^n\right)^2 + B(x,x) = a_{11}(y)^2 + B(x,x),$$

где $B(\mathbf{x},\mathbf{x})$ — квадратичная форма, не содержащая x^1 и, следовательно, по предположению индукции, имеющая канонический базис. Для завершения доказательства, тем самым, осталось положить

$$\widetilde{x}^1 = \begin{cases} \sqrt{a_{11}}\,y & \text{при} \quad a_{11} > 0, \\ \sqrt{-a_{11}}\,y & \text{при} \quad a_{11} < 0 \end{cases} \qquad (\text{или просто } \widetilde{x}^1 = \sqrt{a_{11}}y).$$

2°. $a_{11} = 0$. В этом случае при некотором i $a_{1i} \neq 0$ (иначе переменная x^1 не входила бы в выражение $a_{ij}x^ix^j$). Полагая $x^1 = y^1 + y^i$, $x^i = y^1 - y^i$, мы сведем стоящую перед нами задачу к случаю 1°, поскольку слагаемое $2a_{1i}x^1x^i$ превратится в разность $2a_{1i}(y^1)^2 - 2a_{1i}(y^i)^2$. Теорема доказана.

Замечание 1. Идею доказательства этой теоремы можно использовать на практике. В самом деле, указанным способом можно сперва привести данную квадратичную форму к виду $A(x,x) = l_1\left(\widetilde{x}^1\right)^2 + B(x,x)$, где $B(\mathbf{x},\mathbf{x})$ — квадратичная форма, не содержащая x^1, затем применить ту же идею к квадратичной форме $B(\mathbf{x},\mathbf{x})$ и т. д. Такой метод нахождения канонического базиса квадратичной формы называется *методом Лагранжа*.

Замечание 2. Название «канонический базис» может ввести в заблуждение — может показаться, что такой базис только один. На самом деле это не так. Например, если искать канонический базис методом Лагранжа, начиная с x^2, а не с x^1, то результат окажется, вообще говоря, другим.

3. Закон инерции. Начиная с этого места и до конца параграфа мы будем рассматривать ***только линейные пространства над полем*** $\mathbb{R}$.

Определение 1. *Квадратичная или билинейная форма называется положительно* (*отрицательно*) *определенной*, *если для любого* $\mathbf{x} \neq \mathbf{o}$ $A(\mathbf{x}, \mathbf{x}) > 0$ (< 0).

Определение 2. *Положительным* (*отрицательным*, *нулевым*) *индексом инерции квадратичной формы в каком-либо ее каноническом базисе называется количество положительных* (*отрицательных*, *нулевых*) *диагональных элементов* (т. е. элементов a_{ii}) *ее матрицы в этом базисе*.

Замечание. *Нулевой индекс инерции квадратичной формы не зависит от выбора канонического базиса* (поскольку он равен $n - \operatorname{rang} A$).

Теорема. *Если положительный индекс инерции квадратичной формы отличен от* 0, *то он равен максимальной размерности подпространства*, *в котором она является положительно определенной*.

Доказательство. Без ограничения общности будем считать, что в некотором каноническом базисе $\{\mathbf{e}_i\}$ квадратичная форма имеет вид

$$A(\mathbf{x}, \mathbf{x}) = (x^1)^2 + \ldots + (x^{k_+})^2 - (x^{k_+ + 1})^2 - \ldots - (x^{k_+ + k_-})^2$$

(в противном случае базисные векторы можно перенумеровать). Ясно, что в подпространстве $\mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_{k_+})$ эта квадратичная форма является положительно определенной. Если $k_+ = n$, то справедливость утверждения очевидна. Пусть $k_+ < n$.

Допустим, что существует подпространство $\mathbf{L}^{k_+ + 1}$, в котором она также положительно определена. Разложим каждый базисный вектор $\mathbf{g}_i$ этого подпространства по базису $\{\mathbf{e}_i\}$ и представим его в виде $\mathbf{g}_i = \mathbf{g}'_i + \mathbf{g}''_i$, где $\mathbf{g}'_i \in \mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_{k_+})$, $\mathbf{g}''_i \in \mathbf{L}(\mathbf{e}_{k_+ + 1}, \ldots, \mathbf{e}_n)$. Векторы $\mathbf{g}'_i$ линейно зависимы — их $(k_+ + 1)$ в k_+-мерном пространстве. Поэтому существует их нетривиальная линейная комбинация, равная $\mathbf{o}$. Следовательно, линейная комбинация векторов $\mathbf{g}_i$ с теми же коэффициентами дает ненулевой вектор (поскольку векторы $\mathbf{g}_i$ линейно независимы) $\mathbf{x}$ пространства $\mathbf{L}(\mathbf{e}_{k_+ + 1}, \ldots, \mathbf{e}_n)$, а значит, $A(\mathbf{x}, \mathbf{x}) \leqslant 0$. Тем самым, мы пришли к противоречию — нашли ненулевой вектор $\mathbf{x} \in \mathbf{L}^{k_+ + 1}$, для которого $A(\mathbf{x}, \mathbf{x}) \leqslant 0$. Теорема доказана.

Следствие 1. *Положительный индекс инерции не зависит от выбора канонического базиса*.

Следствие 2. *Отрицательный индекс инерции не зависит от выбора канонического базиса*, так как от него не зависит величина $k_+ + k_- = \operatorname{rang} A$.

Итак, *все три индекса инерции не зависят от выбора канонического базиса*. Это свойство, называемое *законом инерции квадратичных форм*, дает основание для следующего определения.

Определение 3. *Вид*, *который принимает квадратичная форма в своем каноническом базисе*, *называется ее каноническим видом*.

4. Критерий Сильвестра. Условимся называть *угловым минором* матрицы ее минор, расположенный в левом верхнем углу.

Теорема (*критерий Сильвестра*). *Квадратичная форма является положительно определенной тогда и только тогда, когда все угловые миноры ее матрицы положительны.*

Доказательство. Допустим сначала, что квадратичная форма является положительно определенной, и докажем, что все угловые миноры ее матрицы положительны. Воспользуемся методом математической индукции. В одномерном пространстве справедливость утверждения очевидна. Предположим, что оно справедливо в n-мерном пространстве, и докажем его справедливость в $(n+1)$-мерном пространстве. Поскольку квадратичная форма положительно определена, в частности для векторов подпространства $\mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ $\left(\sum\limits_{i,j=1}^{n} a_{ij}x^ix^j \geqslant 0\right)$, то все угловые миноры ее матрицы, вплоть до n-го порядка, положительны по предположению индукции. Но и ее определитель больше нуля — в каноническом базисе он равен 1 (так как $k_+ = n+1$), а знак определителя матрицы билинейной, а значит и квадратичной, формы не зависит от выбора базиса.

Допустим теперь, что все угловые миноры матрицы квадратичной формы положительны, и докажем, что эта квадратичная форма является положительно определенной. Вновь воспользуемся методом математической индукции. В одномерном пространстве справедливость утверждения очевидна. Предположим, что оно справедливо в n-мерном пространстве, и докажем его справедливость в $(n+1)$-мерном пространстве. По предположению индукции данная квадратичная форма положительно определена для векторов подпространства $\mathbf{L}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ и, следовательно, ее положительный индекс инерции не меньше n. Но если он равен n, то в каноническом базисе определитель ее матрицы неположительный, а знак определителя матрицы квадратичной формы не зависит от выбора базиса. Поэтому он равен $n+1$, что и требовалось доказать.

Следствие. *Квадратичная форма является отрицательно определенной тогда и только тогда, когда угловые миноры ее матрицы четного порядка положительны, а нечетного — отрицательны.*

В самом деле, квадратичная форма $a_{ij}x^ix^j$ является отрицательно определенной тогда и только тогда, когда квадратичная форма $(-a_{ij})x^ix^j$ положительно определена.

Глава 4
ЕВКЛИДОВО ПРОСТРАНСТВО

§ 1. Длины и углы

1. Определение евклидова пространства.

Определение. *Линейное пространство над полем* $\mathbb{R}$ *называется евклидовым, если в нем зафиксирована некоторая симметричная положительно определенная билинейная форма* $G(\mathbf{x}, \mathbf{y})$, *называемая скалярным произведением* $(\mathbf{x}, \mathbf{y})$.

Билинейную форму $(\mathbf{x}, \mathbf{y}) = g_{ij}x^i y^j$ иногда называют также *метрическим тензором*; n-мерное евклидово пространство обозначается так: $\mathbb{E}^n$.

Замечание. Определение евклидова пространства в развернутом виде можно сформулировать так: *линейное пространство над полем* $\mathbb{R}$ *называется евклидовым, если для любой упорядоченной пары его векторов* $\mathbf{x}$ *и* $\mathbf{y}$ *определена операция скалярного умножения, ставящая ей в соответствие вещественное число* $(\mathbf{x}, \mathbf{y})$. *При этом*:

1° *для любых* $\mathbf{x}$, $\mathbf{y}$ *имеет место равенство* $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$;

2° *для любых* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ *имеет место равенство* $(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$;

3° *для любых* $\mathbf{x}$, $\mathbf{y}$ *и любого числа* λ *имеет место равенство* $(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$;

4° *для любого* $\mathbf{x}$ *имеет место неравенство* $(\mathbf{x}, \mathbf{x}) \geqslant 0$, *причем* $(\mathbf{x}, \mathbf{x}) = 0$ *тогда и только тогда, когда* $\mathbf{x} = \mathbf{o}$.

Перечисленные утверждения составляют III группу аксиом Вейля. Теперь отчетливо видно, что эти «аксиомы» в действительности таковыми не являются. Ясно также, что поскольку в линейном пространстве имеется бесконечное количество симметричных, положительно определенных билинейных форм (их столько же, сколько симметричных матриц, удовлетворяющих критерию Сильвестра), то *каждое линейное пространство может быть сделано евклидовым, причем бесконечным числом способов*.

2. Неравенство Коши–Буняковского.

Теорема. *Для любых двух элементов* $\mathbf{x}$, $\mathbf{y}$ *евклидова пространства имеет место неравенство* $(\mathbf{x}, \mathbf{y})^2 \leqslant (\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})$.

Доказательство. Из определения скалярного произведения следует, что для любых векторов $\mathbf{x}$, $\mathbf{y}$ и любого числа λ имеет место неравенство $(\lambda\mathbf{x} - \mathbf{y}, \lambda\mathbf{x} - \mathbf{y}) \geqslant 0$, или $\lambda^2(\mathbf{x}, \mathbf{x}) - 2\lambda(\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{y}) \geqslant 0$.

Из произвольности λ следует, что квадратное уравнение $\lambda^2(\mathbf{x},\mathbf{x}) - 2\lambda(\mathbf{x},\mathbf{y})+(\mathbf{y},\mathbf{y}) = 0$ имеет не более одного корня λ (иначе в интервале между корнями $\lambda^2(\mathbf{x},\mathbf{x}) - 2\lambda(\mathbf{x},\mathbf{y}) + (\mathbf{y},\mathbf{y}) < 0$), т. е. его дискриминант неположителен: $(\mathbf{x},\mathbf{y})^2 - (\mathbf{x},\mathbf{x})(\mathbf{y},\mathbf{y}) \leqslant 0$, что и требовалось доказать.

Замечание. Установленное нами неравенство в отечественной литературе называется *неравенством Коши–Буняковского*, хотя, видимо, правильнее было бы называть его неравенством Коши–Буняковского–Шварца, поскольку в мировой литературе в указанном контексте можно встретить любые комбинации из этих трех фамилий.

3. Длина вектора.

Определение. *Длиной вектора* $\mathbf{x}$ *называется число* $|\mathbf{x}| = \sqrt{(x,x)}$. Ясно, что число $|\mathbf{x}|$ — вещественное.

Замечание. В новых обозначениях неравенство Коши–Буняковского можно записать так: $|(\mathbf{x},\mathbf{y})| \leqslant |\mathbf{x}||\mathbf{y}|$.

Теорема. 1° *Для любого* $\mathbf{x}$ *имеет место неравенство* $|\mathbf{x}| \geqslant 0$, *причем* $|\mathbf{x}| = 0$ *тогда и только тогда, когда* $\mathbf{x} = \mathbf{o}$;

2° *для любого вектора* $\mathbf{x}$ *и любого числа* λ *имеет место равенство* $|\lambda\mathbf{x}| = |\lambda||\mathbf{x}|$;

3° *для любых* $\mathbf{x}$, $\mathbf{y}$ *имеет место неравенство* $|\mathbf{x}+\mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|$ (*неравенство треугольника*).

Доказательство. 1°. Справедливость этого неравенства вытекает непосредственно из определения.

2°. Имеем: $|\lambda x| = \sqrt{(\lambda x, \lambda x)} = \sqrt{\lambda^2\,(x,x)} = |\lambda||x|$.

3°. Имеем: $|\mathbf{x}+\mathbf{y}|^2 = (\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) = |\mathbf{x}|^2+|\mathbf{y}|^2+2(\mathbf{x},\mathbf{y}) \leqslant |\mathbf{x}|^2+|\mathbf{y}|^2+ 2|(\mathbf{x},\mathbf{y})| \leqslant |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}||\mathbf{y}| = (|\mathbf{x}| + |\mathbf{y}|)^2$, откуда $|\mathbf{x}+\mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|$. Теорема доказана.

Замечание. Линейное пространство, каждому элементу $\mathbf{x}$ которого ставится в соответствие вещественное число $|\mathbf{x}|$ (или $||\mathbf{x}||$) так, что имеют место свойства 1°–3°, называется *нормированным пространством*, а число $|\mathbf{x}|$ (или $||\mathbf{x}||$) — *нормой* $\mathbf{x}$. Поэтому можно сказать, что евклидово пространство является нормированным пространством с нормой $|\mathbf{x}| = \sqrt{(x,x)}$.

4. Угол между векторами.

Определение 1. *Углом между ненулевыми векторами* $\mathbf{x}$ *и* $\mathbf{y}$ *называется величина* φ $(0 \leqslant \varphi \leqslant \pi)$, *определяемая из уравнения* $\cos\varphi = \dfrac{(x,y)}{|x||y|}$.

Из неравенства Коши–Буняковского следует, что угол определен для любых двух ненулевых векторов.

Определение 2. *Векторы* $\mathbf{x}$ *и* $\mathbf{y}$ *называются ортогональными* ($\mathbf{x}\perp\mathbf{y}$), *если* $(\mathbf{x},\mathbf{y}) = 0$.

Сделаем несколько очевидных замечаний.

Замечание 1. $\mathbf{x}\perp\mathbf{x}$ *тогда и только тогда, когда* $\mathbf{x} = \mathbf{o}$.

Замечание 2. *Если вектор ортогонален нескольким векторам, то он ортогонален и любой их линейной комбинации.*

Замечание 3. *Если* $\mathbf{x} \perp \mathbf{y}$, *то* $|\mathbf{x}+\mathbf{y}|^2 = (\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2$ (*теорема Пифагора*).

Замечание 4. *Ненулевые попарно ортогональные векторы линейно независимы.*

В самом деле, пусть $\mathbf{x}_1, \ldots, \mathbf{x}_k$ — ненулевые попарно ортогональные векторы. Допустим, что $\lambda^s \mathbf{x}_s = \mathbf{o}$. Умножая это равенство скалярно на $\mathbf{x}_i$ $(i = 1, \ldots, k)$, получаем $\lambda^i = 0$, что и требовалось доказать.

§ 2. Ортонормированный базис

1. Существование ортонормированного базиса. Из последнего замечания предыдущего пункта следует, что упорядоченная совокупность n попарно ортогональных единичных векторов в $\mathbb{E}^n$ (если, конечно, она существует) образует базис. Он называется *ортонормированным базисом*.

Теорема. *В* $\mathbb{E}^n$ *существует ортонормированный базис.*

Доказательство. Пусть в некотором базисе скалярное произведение выражается формулой $(\mathbf{x}, \mathbf{y}) = g_{ij} x^i y^j$. Этой симметричной билинейной форме соответствует положительно определенная квадратичная форма $(\mathbf{x}, \mathbf{x})$, матрица которой в каноническом базисе единичная. Следовательно, в этом базисе матрица билинейной формы $(\mathbf{x}, \mathbf{y})$ также единичная: $\widetilde{g}_{ij} = (\widetilde{e}_i, \widetilde{e}_j) = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j. \end{cases}$ Теорема доказана.

Замечание 1. Поскольку в ортонормированном базисе $g_{ij} = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j, \end{cases}$ то $(\mathbf{x}, \mathbf{y}) = x^1 y^1 + \ldots + x^n y^n$. В частности, $|\mathbf{x}|^2 = (x^1)^2 + \ldots + (x^n)^2$.

Замечание 2. Разложение вектора по ортонормированному базису выглядит так: $\mathbf{x} = (\mathbf{x}, \mathbf{e}_1)\mathbf{e}_1 + \ldots + (\mathbf{x}, \mathbf{e}_n)\mathbf{e}_n$ (чтобы убедиться в справедливости этого равенства, достаточно скалярно умножить обе его части на $\mathbf{e}_i$, полагая $i = 1, \ldots, n$).

Определение. *Пусть* $\mathbf{L}$ — *линейное подпространство пространства* $\mathbb{E}^n$, $\mathbf{e}_1, \ldots, \mathbf{e}_k$ — *ортонормированный базис в этом пространстве. Вектор* $\mathbf{y} = (\mathbf{x}, \mathbf{e}_1)\mathbf{e}_1 + \ldots + (\mathbf{x}, \mathbf{e}_k)\mathbf{e}_k$ *называется проекцией вектора* $\mathbf{x}$ *на подпространство* $\mathbf{L}$.

Замечание. Ясно, что вектор $\mathbf{z} = \mathbf{x} - \mathbf{y}$ ортогонален каждому из векторов $\mathbf{e}_1, \ldots, \mathbf{e}_k$, а значит и любой их линейной комбинации, т. е. ортогонален любому вектору подпространства $\mathbf{L}$.

2. Ортогонализация. Рассмотрим такую задачу: *как, зная какой-нибудь базис* $\{\mathbf{e}_i\}$ *пространства* $\mathbb{E}^n$, *построить ортонормированный базис этого пространства*? Можно, конечно, поступить так: найти матрицу $g_{ij} = (\mathbf{e}_i, \mathbf{e}_j)$, а затем построить канонический базис квадратичной формы $A(\mathbf{x}, \mathbf{x}) = g_{ij} x^i x^j$ — он и будет ортонормированным. На практике, однако, чаще поступают иначе.

Положим $\widetilde{\mathbf{e}}_1 = \mathbf{e}_1/|\mathbf{e}_1|$. Затем, вычитая из вектора $\mathbf{e}_2$ его проекцию на подпространство $\mathbf{L}(\widetilde{\mathbf{e}}_1)$ и обозначая результат через $\mathbf{g}_2$, положим $\widetilde{\mathbf{e}}_2 = \mathbf{g}_2/|\mathbf{g}_2|$. Далее, вычитая из вектора $\mathbf{e}_3$ его проекцию на подпространство $\mathbf{L}(\widetilde{\mathbf{e}}_1, \widetilde{\mathbf{e}}_2)$ и обозначая результат через $\mathbf{g}_3$, положим $\widetilde{\mathbf{e}}_3 = \mathbf{g}_3/|\mathbf{g}_3|$, и т. д. Продолжая этот процесс, мы получим, в конце концов, ортонормированный базис $\{\widetilde{\mathbf{e}}_i\}$. Этот прием называется *ортогонализацией*.

3. Ортогональное дополнение.

Определение. *Пусть* $\mathbf{L}$ *— линейное подпространство пространства* $\mathbb{E}^n$. *Множество* $\mathbf{L}_\perp$ *всех векторов пространства* $\mathbb{E}^n$, *ортогональных всем векторам подпространства* $\mathbf{L}$, *называется ортогональным дополнением подпространства* $\mathbf{L}$.

Теорема. *Ортогональное дополнение линейного подпространства* $\mathbf{L}^k$ *пространства* $\mathbb{E}^n$ *является линейным подпространством размерности* $n-k$.

Доказательство. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — ортонормированный базис пространства $\mathbb{E}^n$, $\mathbf{g}_1, \ldots, \mathbf{g}_k$ — произвольный базис пространства $\mathbf{L}^k$. Разложим каждый из векторов $\mathbf{g}_i$ по базису $\{\mathbf{e}_i\}$: $\mathbf{g}_i = g_i^j \mathbf{e}_j$. Столбцы матрицы G, будучи координатами линейно независимых векторов $\mathbf{g}_i$, линейно независимы, поэтому $\operatorname{rang} G = k$. Ортогональное дополнение подпространства $\mathbf{L}^k$ представляет собой множество всех решений $\mathbf{x}$ однородной системы

$$(\mathbf{g}_i, \mathbf{x}) = \sum_{j=1}^{n} g_i^j x^j = 0 \qquad (i = 1, \ldots, k).$$

Следовательно, оно является линейным подпространством размерности $n-k$. Теорема доказана.

Следствие 1. *Для любого линейного подпространства* $\mathbf{L}$ *пространства* $\mathbb{E}^n$ *имеет место равенство* $\mathbb{E}^n = \mathbf{L} \oplus \mathbf{L}_\perp$ (так как совокупность ортонормированных базисов этих подпространств представляет собой ортонормированную совокупность n векторов, т. е. базис пространства $\mathbb{E}^n$).

Следствие 2. *Подпространство* $\mathbf{L}$ *является ортогональным дополнением подпространства* $\mathbf{L}_\perp$, поскольку вектор $\mathbf{x} = \mathbf{y} + \mathbf{z}$ $(\mathbf{y} \in \mathbf{L},\ \mathbf{z} \in \mathbf{L}_\perp)$ ортогонален всем векторам подпространства $\mathbf{L}_\perp$ тогда и только тогда, когда $\mathbf{z} = \mathbf{o}$, а значит, $\mathbf{x} = \mathbf{y} \in \mathbf{L}$.

Замечание. В ходе доказательства теоремы нам пришлось воспользоваться символом Σ, поскольку возникла какая-то «путаница» с верхними и нижними индексами. Это явление, причина которого вскоре выяснится, типично при рассмотрении ортонормированных базисов.

4. Альтернатива Фредгольма. Как мы помним, линейное дополнение подпространства $\mathbf{L}$ линейного пространства $\mathbf{L}^n$ определяется неоднозначно. В линейном пространстве над полем $\mathbb{R}$ с фиксированным

базисом существует простой алгоритм построения одного из линейных дополнений подпространства $\mathbf{L}$. Он состоит в том, что сначала вводят скалярное произведение, считая данный базис ортонормированным, а затем строят ортогональное дополнение $\mathbf{L}_\perp$ подпространства $\mathbf{L}$ — оно и является искомым линейным дополнением. Воспользуемся этой идеей для доказательства следующей теоремы.

Теорема. *Система m линейных уравнений с n неизвестными $A\mathbf{x} = \mathbf{b}$, коэффициенты которой вещественны, имеет решение при любых вещественных правых частях $\mathbf{b}$ тогда и только тогда, когда система $A^{tr}\mathbf{x} = \mathbf{o}$ имеет лишь тривиальное решение.*

Доказательство. Рассмотрим линейный оператор с матрицей A, действующий из $\mathbb{R}^n$ в $\mathbb{R}^m$. Считая тот базис пространства $\mathbb{R}^m$, в котором числа a_{ij} и b_i являются координатами столбцов $\mathbf{a}_j$ и $\mathbf{b}$, ортонормированным, представим пространство $\mathbb{R}^m$ в виде прямой суммы $\operatorname{im} A = \mathbf{L}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ и $(\operatorname{im} A)_\perp$ — множества всех решений однородной системы $(\mathbf{a}_i, \mathbf{x}) = \sum\limits_{j=1}^{m} a_{ji}x_j = 0$ $(i = 1, \ldots, n)$, т. е. системы $A^{tr}\mathbf{x} = \mathbf{o}$. Система $A\mathbf{x} = \mathbf{b}$ имеет решение при любых $\mathbf{b}$ тогда и только тогда, когда любой вектор $\mathbf{b}$ пространства $\mathbb{R}^m$ принадлежит $\operatorname{im} A$, т. е. тогда и только тогда, когда $(\operatorname{im} A)_\perp = \mathbf{o}$, и, следовательно, система $A^{tr}\mathbf{x} = \mathbf{o}$ имеет лишь тривиальное решение. Теорема доказана.

Следствие (альтернатива Фредгольма). *Либо система линейных уравнений $A\mathbf{x} = \mathbf{b}$ с вещественными коэффициентами имеет решение при любых вещественных правых частях $\mathbf{b}$, либо система $A^{tr}\mathbf{x} = \mathbf{o}$ имеет нетривиальное решение.*

Замечание 1. Ясно, что система $A\mathbf{x} = \mathbf{b}$ имеет решение тогда и только тогда, когда $\mathbf{b} \in \operatorname{im} A$ или, что то же самое, вектор $\mathbf{b}$ ортогонален к любому вектору пространства $(\operatorname{im} A)_\perp$. Учитывая, что это пространство является множеством всех решений системы $A^{tr}\mathbf{x} = \mathbf{o}$, мы приходим к следующему выводу: *система $A\mathbf{x} = \mathbf{b}$ имеет решение тогда и только тогда, когда вектор $\mathbf{b}$ ортогонален к любому вектору ядра оператора с матрицей A^{tr}.*

Замечание 2. Рассмотрим систему $A\mathbf{x} = \mathbf{b}$ и запишем ее в виде $A\mathbf{x} = \mathbf{b}_1 + \mathbf{b}_2$, где $\mathbf{b}_1 \in \operatorname{im} A, \mathbf{b}_2 \in (\operatorname{im} A)_\perp$. Умножая обе части этого равенства слева на матрицу A^{tr} и учитывая, что $A^{tr}\mathbf{b}_2 = \mathbf{o}$, получим: $A^{tr}A\mathbf{x} = A^{tr}\mathbf{b}_1$. Но система $A\mathbf{x} = \mathbf{b}_1$ всегда имеет решение, поскольку $\mathbf{b}_1 \in \operatorname{im} A$. Следовательно, и полученная система, т. е. *система $A^{tr}A\mathbf{x} = A^{tr}\mathbf{b}$, имеет решение при всех $\mathbf{b}$.*

§ 3. Операторы в E^n

1. Сопряженный оператор.

Определение. *Оператор A^* называется сопряженным по отношению к оператору A, если для любых векторов $\mathbf{x}$, $\mathbf{y}$ имеет место равенство $(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$.*

Теорема. *Оператор A^* является сопряженным по отношению к оператору A тогда и только тогда, когда в ортонормированном базисе матрица A^* равна A^{tr}.*

Доказательство. Оператор B является сопряженным по отношению к оператору A тогда и только тогда, когда в ортонормированном базисе $(A\mathbf{x}, \mathbf{y}) = \sum\limits_{i,s} a_s^i x^s y^i = (\mathbf{x}, B\mathbf{y}) = \sum\limits_{i,s} x^s b_i^s y^i$, или $\sum\limits_{i,s} \left(a_s^i - b_i^s\right) x^s y^i \equiv 0$. Но билинейная форма тождественно равна нулю тогда и только тогда, когда все ее коэффициенты (равные ее значениям на базисных векторах) равны нулю. Следовательно, полученное тождество эквивалентно равенству $B = A^{tr}$. Теорема доказана.

Следствие 1. *Для любого оператора A существует единственный оператор A^** (чтобы найти его матрицу, нужно взять матрицу оператора A в каком-нибудь ортонормированном базисе и транспонировать ее).

Следствие 2. $(A^*)^* = A$.

2. Ортогональный оператор.

Определение 1. *Оператор A называется ортогональным, если для любых векторов $\mathbf{x}$, $\mathbf{y}$ имеет место равенство $(A\mathbf{x}, A\mathbf{y}) = (\mathbf{x}, \mathbf{y})$.*

Замечание. Поскольку ортогональный оператор не меняет скалярного произведения векторов, то он не меняет длины векторов и углы между ними. В частности, любой ортонормированный базис он переводит в ортонормированный базис. Геометрически это соответствует повороту и отражению относительно координатных плоскостей.

Теорема 1. *Если $\mathbf{L}$ — инвариантное подпространство ортогонального оператора A, то и $\mathbf{L}_\perp$ — его инвариантное подпространство.*

Доказательство. Оператор A переводит ортонормированный базис пространства $\mathbf{L}$ в ортонормированную совокупность векторов пространства $\mathbf{L}$, количество которых равно размерности этого пространства, т. е. в ортонормированный базис пространства $\mathbf{L}$. Следовательно, $A(\mathbf{L}) = \mathbf{L}$. Поэтому для любого вектора $\mathbf{x} \in \mathbf{L}$ существует такой вектор $\mathbf{y} \in L$, что $A\mathbf{y} = \mathbf{x}$.

Если $\mathbf{z} \in \mathbf{L}_\perp$, то для любого $\mathbf{x} \in \mathbf{L}$ имеет место равенство $(A\mathbf{z}, \mathbf{x}) = (A\mathbf{z}, A\mathbf{y}) = (\mathbf{z}, \mathbf{y}) = 0$, поскольку $\mathbf{y} \in \mathbf{L}$. Теорема доказана.

Теорема 2. *Оператор A является ортогональным тогда и только тогда, когда $A^* = A^{-1}$.*

Доказательство. Оператор A является ортогональным тогда и только тогда, когда $(A\mathbf{x}, A\mathbf{y}) \equiv (\mathbf{x}, A^*A\mathbf{y}) \equiv (\mathbf{x}, \mathbf{y})$, или $(\mathbf{x}, A^*A\mathbf{y} - \mathbf{y}) \equiv 0$, т. е. $A^*A\mathbf{y} \equiv \mathbf{y}$ или $A^*A = E$. Теорема доказана.

Следствие. *Оператор A является ортогональным тогда и только тогда, когда в ортонормированном базисе матрица A^{-1} равна A^{tr}.*

Определение 2. *Матрица A называется ортогональной, если $A^{-1} = A^{tr}$.*

Таким образом, можно сказать, что *оператор является ортогональным тогда и только тогда, когда в ортонормированном базисе его матрица — ортогональная.*

Замечание 1. По определению ортогональной матрицы $AA^{tr} = A^{tr}A = E$, т. е. $\sum\limits_s a_{is}a_{js} = \sum\limits_s a_{si}a_{sj} = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j. \end{cases}$ Таким образом, строки (столбцы) ортогональной матрицы — это координаты векторов, образующих ортонормированный базис.

Замечание 2. Поскольку для ортогональной матрицы $AA^{tr} = E$, то $(\det A)^2 = 1$. Если $\det A = 1$, то матрица A называется *собственной*, а если $\det A = (-1)$, то *несобственной* ортогональной матрицей.

Отметим также, что поскольку определитель матрицы линейного оператора не зависит от выбора базиса, то определитель матрицы ортогонального оператора в любом базисе равен ± 1.

Теорема 3. *Для любого ортогонального оператора существует такой ортонормированный базис, в котором его матрица имеет вид*

$$\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 & & & & & & \\ \sin\varphi_1 & \cos\varphi_1 & & & & & & \\ & & \ddots & & & & 0 & \\ & & & \cos\varphi_k & -\sin\varphi_k & & & \\ & & & \sin\varphi_k & \cos\varphi_k & & & \\ & & & & & \lambda_1 & & \\ & 0 & & & & & \ddots & \\ & & & & & & & \lambda_{n-2k} \end{pmatrix},$$

где $\lambda_i = \pm 1$.

Доказательство. Воспользуемся теоремой 1 и методом математической индукции. При $n = 1$ справедливость утверждения очевидна. Допустим, что теорема доказана для $n \leqslant m-1$, и докажем, что в таком случае она верна и для $n = m$.

Если характеристическое уравнение для данного оператора имеет по крайней мере один вещественный корень λ_m, то $\lambda_m = \pm 1$, поскольку из равенства $A\mathbf{x} = \lambda_m \mathbf{x}$ следует, что $(A\mathbf{x}, A\mathbf{x}) = (\mathbf{x}, \mathbf{x}) = \lambda_m^2(\mathbf{x}, \mathbf{x})$. В этом случае для доказательства теоремы достаточно принять собственный вектор, соответствующий λ_m, за базисный вектор $\mathbf{e}_m$, а затем в ортогональном дополнении к $\mathbf{L}(\mathbf{e}_m)$ привести матрицу оператора к требуемому виду.

Возможен, однако, и такой случай: характеристическое уравнение не имеет вещественных корней. Пусть $\lambda + i\mu$ — один из его комплексных корней. Ему соответствует двумерное инвариантное подпространство $\mathbf{L}$ оператора A (см. п. 6 § 2 гл. 2). Выберем ортонормированный базис $\{\mathbf{e}_i\}$

так, что $\mathbf{L} = \mathbf{L}(\mathbf{e}_1, \mathbf{e}_2)$. Поскольку $\mathbf{L}_\perp$ — также инвариантное подпространство, то матрица оператора в базисе $\{\mathbf{e}_i\}$ имеет «блочный» вид

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ 0 & 0 & & & \\ \dots & \dots & & \widetilde{A} & \\ 0 & 0 & & & \end{pmatrix}.$$

Строки ортогональной матрицы — это координаты векторов, образующих ортонормированный базис, поэтому

$$\left.\begin{aligned} a_{11}^2 + a_{12}^2 = 1 \\ a_{21}^2 + a_{22}^2 = 1 \\ a_{11}a_{21} + a_{12}a_{22} = 0 \end{aligned}\right\}.$$

Полагая $a_{11} = \cos\varphi$, $a_{12} = -\sin\varphi$, получаем:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \pm\sin\varphi & \pm\cos\varphi \end{pmatrix}.$$

Однако один из этих случаев отпадает, поскольку характеристическое уравнение $\begin{pmatrix} \cos\varphi - \lambda & -\sin\varphi \\ -\sin\varphi & -\cos\varphi - \lambda \end{pmatrix} = \lambda^2 - 1 = 0$, а значит и характеристическое уравнение для матрицы A, имеют вещественные корни, что противоречит предположению. Следовательно, $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. Для завершения доказательства теоремы осталось привести матрицу $\widetilde{A}$ к требуемому виду, что возможно в силу предположения индукции.

3. Ортогональные преобразования. Пусть $\{\mathbf{e}_i\}$ — ортонормированный базис, $\{\widetilde{\mathbf{e}}_i\}$ — произвольный базис, $\widetilde{\mathbf{e}}_i = \alpha_i^s \mathbf{e}_s$.

Теорема. *Базис $\{\widetilde{\mathbf{e}}_i\}$ является ортонормированным тогда и только тогда, когда матрица α — ортогональная.*

Доказательство. Базис $\{\widetilde{\mathbf{e}}_i\}$ является ортонормированным тогда и только тогда, когда $(\widetilde{\mathbf{e}}_i, \widetilde{\mathbf{e}}_j) = \delta_{ij}$, где $\delta_{ij} = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j. \end{cases}$ Имеем:

$$(\widetilde{\mathbf{e}}_i, \widetilde{\mathbf{e}}_j) = (\alpha_i^s \mathbf{e}_s, \alpha_j^r \mathbf{e}_r) = \alpha_i^s \alpha_j^r (\mathbf{e}_s, \mathbf{e}_r) = \alpha_i^s \alpha_j^r \delta_{sr} = \sum_s \alpha_i^s \alpha_j^s = \delta_{ij}.$$

Итак, $(\widetilde{\mathbf{e}}_i, \widetilde{\mathbf{e}}_j) = \delta_{ij}$ тогда и только тогда, когда $\sum\limits_s \alpha_i^s \alpha_j^s = \delta_{ij}$, т. е. $\alpha^{tr}\alpha = E$, или $\alpha^{-1} = \alpha^{tr}$. Теорема доказана.

Следствие. *Произведение ортогональных матриц является ортогональной матрицей.*

Замечание. *При ортогональных преобразованиях*, т. е. при переходах от ортонормированных базисов к ортонормированным базисам, *исчезает различие между ковариантными и контравариантными тензорами.* В самом деле, рассмотрим, например, тензор типа (1,1)

(линейный оператор). Его матрица при переходе к новому базису преобразуется так: $\hat{A} = \beta A \alpha$. Матрица тензора типа $(2,0)$ (билинейной или квадратичной формы) преобразуется так: $\widetilde{B} = \alpha^{tr} B \alpha$. Но при ортогональном преобразовании $\beta = \alpha^{tr}$, поэтому эти две формулы становятся одинаковыми. С этим, в частности, связана «путаница» с верхними и нижними индексами, иногда возникающая при рассмотрении ортонормированных базисов.

4. Самосопряженный оператор.

Определение. *Линейный оператор A называется самосопряженным, если $A^* = A$.*

Тем самым, можно сказать, что *в любом ортонормированном базисе матрица самосопряженного оператора симметрична*: $a_{ij} = a_{ji}$.

Теорема 1. *Все корни характеристического уравнения для самосопряженного оператора вещественны.*

Доказательство. Допустим, что указанное характеристическое уравнение имеет корень $\lambda + i\mu$, т. е. $\det(A - (\lambda + i\mu)E) = 0$. Тогда однородная система $A\mathbf{z} = (\lambda + i\mu)\mathbf{z}$ имеет нетривиальное решение $\mathbf{z} = \mathbf{x} + i\mathbf{y}$: $A(\mathbf{x}+i\mathbf{y}) = (\lambda+i\mu)(\mathbf{x}+i\mathbf{y})$, или $A\mathbf{x}+iA\mathbf{y} = \lambda\mathbf{x}-\mu\mathbf{y}+i(\lambda\mathbf{y}+\mu\mathbf{x})$, откуда

$$\left.\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} - \mu\mathbf{y} \\ Ay &= \lambda\mathbf{y} + \mu\mathbf{x} \end{aligned}\right\}.$$

Умножая скалярно первое уравнение на $(-\mathbf{y})$, второе — на $\mathbf{x}$ и складывая их, получаем: $-(A\mathbf{x}, \mathbf{y}) + (\mathbf{x}, A\mathbf{y}) = \mu(|\mathbf{x}|^2 + |\mathbf{y}|^2)$. Но левая часть этого равенства равна нулю, поскольку оператор A — самосопряженный. Следовательно, $\mu = 0$. Теорема доказана.

Теорема 2. *Для любого самосопряженного оператора существует ортонормированный базис из его собственных векторов.*

Доказательство. Заметим сначала, что если $\mathbf{x}$ — собственный вектор самосопряженного оператора A, $\mathbf{y} \perp \mathbf{x}$, то $A\mathbf{y} \perp \mathbf{x}$. В самом деле, $(\mathbf{x}, A\mathbf{y}) = (A\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y}) = 0$. Таким образом, ортогональное дополнение подпространства $\mathbf{L}(\mathbf{x})$ является инвариантным подпространством оператора A.

Перейдем теперь непосредственно к доказательству теоремы. Найдем какое-нибудь собственное значение λ_1 (его существование гарантирует теорема 1) и соответствующий ему единичный собственный вектор $\mathbf{e}_1$. Ортогональное дополнение подпространства $\mathbf{L}(\mathbf{e}_1)$ является инвариантным подпространством оператора A, поэтому теперь можно рассматривать оператор A только на нем. Найдем какое-нибудь собственное значение λ_2 и соответствующий ему единичный собственный вектор $\mathbf{e}_2$, после чего будем рассматривать оператор A только на ортогональном дополнении подпространства $\mathbf{L}(\mathbf{e}_1, \mathbf{e}_2)$ и т. д. В результате мы построим ортонормированный базис из собственных векторов $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Теорема доказана.

Следствие. *Для любого самосопряженного оператора существует такой ортонормированный базис, в котором его матрица диагональна.*

Замечание 1. Справедливо и обратное утверждение: *если существует ортонормированный базис, в котором матрица оператора диагональна, то этот оператор — самосопряженный*, поскольку в указанном базисе матрица этого оператора симметрична.

Замечание 2. Отметим, что в том базисе, в котором матрица ортогонального оператора A имеет вид, указанный в теореме 3 п. 2, матрица $A + A^{tr}$ диагональна. Попробуйте объяснить этот факт и, основываясь на Вашем объяснении, привести еще одно доказательство упомянутой теоремы.

5. Квадратичная форма в E^n.

Теорема. *Для любой квадратичной или симметричной билинейной формы существует такой ортонормированный базис, в котором ее матрица диагональна.*

Доказательство. Рассмотрим сначала произвольный ортонормированный базис и запишем в нем матрицу данной квадратичной или симметричной билинейной формы. Рассмотрим теперь самосопряженный оператор с такой же матрицей и построим ортонормированный базис из его собственных векторов. В нем матрица оператора диагональна. Но переход от одного ортонормированного базиса к другому ортонормированному базису — это ортогональное преобразование, а при таких преобразованиях матрицы оператора и квадратичной или билинейной формы преобразуются одинаково. Следовательно, построенный ортонормированный базис — искомый. Теорема доказана.

Следствие. *В линейном пространстве над полем $\mathbb{R}$ для любых двух квадратичных или симметричных билинейных форм, одна из которых — положительно определенная, существует такой базис, в котором матрица одной из них — единичная, а другой — диагональная.* В самом деле, принимая положительно определенную симметричную билинейную форму (данную, или ту, из которой получена данная положительно определенная квадратичная форма) за скалярное произведение, мы сможем построить такой ортонормированный базис, в котором матрица второй квадратичной или симметричной билинейной формы диагональна.

§ 4. Гиперповерхности второго порядка

1. Система координат. Добавим теперь последнюю, IVгруппу аксиом Вейля. Будем считать, что наряду с векторами в $\mathbb{E}^n$ имеются точки, а также правило, по которому любой упорядоченной паре точек A, B ставится в соответствие вектор $\mathbf{AB}$. При этом:

1° *для любого вектора* $\mathbf{x}$ *и любой точки* O *существует единственная точка* M *такая, что* $\mathbf{OM} = \mathbf{x}$ (откладывание вектора от данной точки);

2° *для любых точек A, B и C имеет место равенство* $\mathbf{AB} + \mathbf{BC} = \mathbf{AC}$ (правило треугольника).

Следствие 1. *Для любой точки A имеет место равенство* $\mathbf{AA} = \mathbf{o}$. В самом деле, прибавляя к обеим частям равенства $\mathbf{AA} + \mathbf{AA} = \mathbf{AA}$ вектор $(-\mathbf{AA})$, получим: $\mathbf{AA} + \mathbf{o} = \mathbf{o}$, откуда $\mathbf{AA} = \mathbf{o}$.

Следствие 2. *Для любых точек A и B имеет место равенство* $\mathbf{BA} = -\mathbf{AB}$. Действительно, $\mathbf{AB} + \mathbf{BA} = \mathbf{AA} = \mathbf{o}$.

Определение. *Системой координат $Ox^1 \ldots x^n$ называется совокупность ортонормированного базиса $\{\mathbf{e}_i\}$ и точки O (начала координат), а координатами точки M — координаты вектора* $\mathbf{OM}$ *в базисе $\{\mathbf{e}_i\}$.*

Ясно, что каждая точка M имеет вполне определенный набор координат $x^1, \ldots, x^n$; обратно, для любых чисел $x^1, \ldots, x^n$ существует ровно одна точка M с координатами $x^1, \ldots, x^n$ (это следует из аксиомы 1^0). Иными словами, соответствие $M \leftrightarrow (x^1, \ldots, x^n)$ является взаимно однозначным. Тот факт, что точка M имеет координаты $x^1, \ldots, x^n$ будем обозначать так: $M(x^1, \ldots, x^n)$.

Пусть $Ox^1, \ldots, x^n$ — «старая» система координат, $\widetilde{O}\widetilde{x}^1, \ldots, \widetilde{x}^n$ — «новая», $\widetilde{O}\left(o^1, \ldots, o^n\right)$. Выразим «новые» координаты точки M через ее «старые» координаты. Для этого заметим, что «новые» координаты точки M — это координаты вектора $\widetilde{\mathbf{OM}}$ в «новой» системе координат. Найдем сначала координаты этого вектора в «старой» системе координат. Согласно аксиоме 2° и следствию 2 имеем: $\widetilde{\mathbf{OM}} = \widetilde{\mathbf{OO}} + \mathbf{OM} = -\mathbf{O}\widetilde{\mathbf{O}} + \mathbf{OM} = \left\{x^1 - o^1, \ldots, x^n - o^n\right\}$. Следовательно, $\widetilde{x}^i = \beta^i_s\left(x^s - o^s\right)$.

Таким образом, переход от «старой» системы координат к «новой» осуществляется при помощи параллельного переноса на вектор $\mathbf{O}\widetilde{\mathbf{O}}$ и ортогонального преобразования, поскольку «старый» и «новый» базисы — ортонормированные. Из полученных формул следует, что старые координаты выражаются через новые так: $x^i = \alpha^i_s \widetilde{x}^s + o^i$, где $\alpha = \beta^{tr}$.

2. Каноническое уравнение гиперповерхности второго порядка. Множество всех точек $M(x^1, \ldots, x^n)$, координаты которых удовлетворяют уравнению $f(x^1, \ldots, x^n) = 0$, называется *гиперповерхностью*. В частности, если $f(x^1, \ldots, x^n) = a_i x^i + b$ и $\sum a_i^2 \neq 0$, то гиперповерхность называется *гиперплоскостью*. Примерами гиперплоскостей могут служить прямая на плоскости и плоскость в пространстве.

Множество всех точек $M(x^1, \ldots, x^n)$, координаты которых удовлетворяют уравнению $a_{ij}x^i x^j + 2b_i x^i + c = 0$, где $\sum a_{ij}^2 \neq 0$, называется *гиперповехностью второго порядка*. При переходе к новому базису по формулам $x^i = \alpha^i_s \widetilde{x}^s$ это уравнение преобразуется так: $\left(\alpha^s_i \alpha^r_j a_{sr}\right)\widetilde{x}^i \widetilde{x}^j + 2\left(\alpha^s_i b_s\right)\widetilde{x}^i + c = 0$. Следовательно, $A(\mathbf{x}, \mathbf{x}) = a_{ij}x^i x^j$ — квадратичная форма, $B(\mathbf{x}) = b_i x^i$ — линейная форма. Поскольку мы ограничиваемся

рассмотрением ортонормированных базисов и, следовательно, не делаем различия между ковариантными и контравариантными тензорами, то линейную форму $B(\mathbf{x})$ можно рассматривать как скалярное произведение вектора $\mathbf{b}$ на вектор $\mathbf{x}$: $B(\mathbf{x}) = (\mathbf{b}, \mathbf{x})$.

Уравнение гиперповерхности второго порядка можно существенно упростить, если перейти к новой системе координат. Делается это, вообще говоря, в три этапа.

1°. Не меняя начала координат, выполним такое ортогональное преобразование, при котором матрица A примет диагональный вид. Само уравнение при этом перепишется так:

$$\sum_{i=1}^{k} \lambda_i \left(\widetilde{x}^i\right)^2 + 2\sum_{i=1}^{n} \widetilde{b}_i \widetilde{x}^i + c = 0.$$

Здесь $k = \operatorname{rang} A \leqslant n$, $\lambda_i \neq 0$.

2°. Сделаем теперь параллельный перенос, полагая при $i \leqslant k$ $\widetilde{x}^i = y^i - b_i/\lambda_i$. В результате уравнение примет вид

$$\sum_{i=1}^{k} \lambda_i \left(y^i\right)^2 + 2\sum_{i=k+1}^{n} \widetilde{b}_i \widetilde{x}^i + \widetilde{c} = 0$$

(при $k = n$ вторая группа слагаемых отсутствует).

3°. В случае $k + 1 < n$ в нашем распоряжении остается еще ортогональное преобразование в пространстве $L\left(\widetilde{e}_{k+1}, \ldots, \widetilde{e}_n\right)$ — оно, очевидно, не меняет квадратичной формы. Выберем базис в этом пространстве так, чтобы новый $(k+1)$-й базисный вектор оказался направленным вдоль вектора $\mathbf{b}$. Тогда в новой системе координат вектор $\mathbf{b}$ будет иметь координаты $\{b, 0, \ldots, 0\}$. Возвращаясь к старым обозначениям, получим окончательно:

$$\sum_{i=1}^{k} \lambda_i \left(x^i\right)^2 + 2bx^{k+1} + c = 0.$$

Это уравнение называется *каноническим уравнением гиперповерхности второго порядка*.

3. Классификация. Все гиперповерхности второго порядка, канонические уравнения которых не содержат хотя бы одной переменной (т. е. либо $k+1 < n$, либо $k+1 = n$, но $b = 0$), называются *цилиндрами*.

Если же каноническое уравнение гиперповерхности второго порядка содержит все переменные, то возможны 3 случая.

1. Если $b = 0$, $c \neq 0$, то уравнение принимает вид

$$\sum_{i=1}^{n} a_i \left(x^i\right)^2 = 1.$$

Гиперповерхности, уравнение которых имеет такой вид, называются *эллипсоидами и гиперболоидами*.

2. Если $b = c = 0$, то уравнение принимает вид

$$\sum_{i=1}^{n} a_i \left(x^i\right)^2 = 0.$$

Гиперповерхности, описываемые уравнениями такого вида, называются *конусами*.

3. Наконец, если $b \neq 0$, то, полагая $x^{k+1} = x^n - c/(2b)$, приведем каноническое уравнение к виду

$$\sum_{i=1}^{n-1} a_i \left(x^i\right)^2 = x^n.$$

Гиперповерхности, описываемые такими уравнениями, называются *параболоидами*.

Замечание. Нетрудно убедиться в том, что в n-мерном пространстве имеется $n^2 + 3n - 1$ различных типов гиперповерхностей второго порядка, n из которых — *мнимые*, т. е. их уравнения описывают пустое множество. При этом цилиндров — n^2+n-3 ($n-1$ из них — мнимые), эллипсоидов и гиперболоидов — $n+1$ (1 — мнимый), конусов — $n/2+1$ при четном n и $(n+1)/2$ — при нечетном, а параболоидов — $n/2$ при четном n и $(n+1)/2$ — при нечетном.

4. Инварианты. Вернемся к общему уравнению гиперповерхности второго порядка:

$$a_{ij}x^i x^j + 2b_i x^i + c = 0.$$

Говорят, что функция $f(a_{11}, \ldots, a_{nn}, b_1, \ldots, b_n, c)$ является *инвариантом* этого уравнения, если при переходе к любой другой системе координат ее значение не изменяется:

$$f(\widetilde{a}_{11}, \ldots, \widetilde{a}_{nn}, \widetilde{b}_1, \ldots, \widetilde{b}_n, \widetilde{c}) = f(a_{11}, \ldots, a_{nn}, b_1, \ldots, b_n, c).$$

Рассмотрим две матрицы:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix} \text{ и } B = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{12} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} & b_n \\ b_1 & b_2 & \ldots & b_n & c \end{pmatrix}.$$

Теорема. *Все коэффициенты характеристического уравнения для матрицы A и $\det B$ являются инвариантами уравнения гиперповерхности второго порядка.*

Доказательство. Переход от одной системы координат к другой — это последовательное выполнение ортогонального преобразования и параллельного переноса. При параллельном переносе матрица A

вообще не изменяется, поэтому не изменяются и коэффициенты указанного характеристического уравнения. При ортогональном преобразовании коэффициенты матрицы A изменяются, но не меняются собственные значения самосопряженного оператора A, а значит и коэффициенты указанного характеристического уравнения, поскольку последние через них выражаются:

$$\begin{vmatrix} (a_{11}-\lambda) & a_{12} & \dots & a_{1n} \\ a_{21} & (a_{22}-\lambda) & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & (a_{nn}-\lambda) \end{vmatrix} \equiv (\lambda_1-\lambda)(\lambda_2-\lambda)(\lambda_n-\lambda).$$

Осталось доказать инвариантность $\det B$. Заметим, прежде всего, что гиперповерхность, описываемая данным уравнением, может рассматриваться как сечение конуса $B(\mathbf{x},\mathbf{x}) = 0$ в $(n+1)$-мерном пространстве гиперплоскостью $x^{n+1} = 1$, т. е. как направляющая этого конуса. В самом деле,

$$B(\mathbf{x},\mathbf{x})|_{x^{k+1}=1} = \sum_{i,j=1}^{n} a_{ij}x^i x^j + 2\sum_{i=1}^{n} b_i x^i \cdot 1 + c\cdot 1^2.$$

Далее, переходу к новой системе координат в n-мерном пространстве по формулам $x^i = \alpha^i_s \widetilde{x}^s + o^i$ соответствует преобразование в $(n+1)$-мерном пространстве с матрицей $\overline{\alpha} = \begin{pmatrix} \alpha^1_1 & \dots & \alpha^1_n & o^1 \\ \dots & \dots & \dots & \dots \\ \alpha^n_1 & \dots & \alpha^n_n & o^n \\ 0 & \dots & 0 & 1 \end{pmatrix}$, поскольку

$$\begin{pmatrix} \alpha^1_1 & \dots & \alpha^1_n & o^1 \\ \dots & \dots & \dots & \dots \\ \alpha^n_1 & \dots & \alpha^n_n & o^n \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} \widetilde{x}^1 \\ \dots \\ \widetilde{x}^n \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha^1_s \widetilde{x}^s + o^1 \\ \dots \\ \alpha^n_s \widetilde{x}^s + o^n \\ 1 \end{pmatrix}.$$

Следовательно, при переходе к новой системе координат матрица B преобразуется так: $\widetilde{B} = \overline{\alpha}^{tr} B \overline{\alpha}$, откуда

$$\det \widetilde{B} = (\det \overline{\alpha})^2 \det B = (\det \alpha)^2 \det B = \det B,$$

так как матрица α — ортогональная. Теорема доказана.

Глава 5
НЕКОТОРЫЕ ОБОБЩЕНИЯ

§ 1. Унитарное пространство

1. Основные свойства. Унитарное пространство — это прямое обобщение евклидова пространства на случай линейного пространства над полем $\mathbb{C}$. Поэтому иногда его называют *комплексным евклидовым пространством*; n-мерное унитарное пространство обозначают так: $\mathbf{U}^n$. Бесконечномерное унитарное пространство играет фундаментальную роль в квантовой механике.

Определение. *Линейное пространство над полем* $\mathbb{C}$ *называется унитарным, если для любой упорядоченной пары его векторов* $\mathbf{x}$ *и* $\mathbf{y}$ *определена операция скалярного умножения, ставящая ей в соответствие комплексное число* $(\mathbf{x}, \mathbf{y})$. *При этом*:

1° *для любых* $\mathbf{x}$, $\mathbf{y}$ *имеет место равенство*

$$(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$$

(черта означает комплексное сопряжение);

2° *для любых* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ *имеет место равенство*

$$(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z});$$

3° *для любых* $\mathbf{x}$, $\mathbf{y}$ *и любого числа* λ *имеет место равенство*

$$(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y});$$

4° *для любого* $\mathbf{x}$ *имеет место неравенство*

$$(\mathbf{x}, \mathbf{x}) \geqslant 0,$$

причем $(\mathbf{x}, \mathbf{x}) = 0$ *тогда и только тогда, когда* $\mathbf{x} = \mathbf{o}$.

Замечание 1. Знак «$\geqslant$» в утверждении 4° может вызвать недоумение — не ясно, как его понимать применительно к комплексным числам. В действительности речь здесь и не идет о комплексных числах. В самом деле, согласно 1°, $(\mathbf{x}, \mathbf{x}) = \overline{(\mathbf{x}, \mathbf{x})}$, поэтому число $(\mathbf{x}, \mathbf{x})$ — вещественное.

Замечание 2. Определенное указанным способом скалярное произведение уже не является билинейной формой. Действительно, $(\mathbf{x}, \mathbf{y} + \mathbf{z}) = \overline{(\mathbf{y} + \mathbf{z}, \mathbf{x})} = \overline{(\mathbf{y}, \mathbf{x})} + \overline{(\mathbf{z}, \mathbf{x})} = (\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{z})$, однако $(\mathbf{x}, \lambda\mathbf{y}) = \overline{(\lambda\mathbf{y}, \mathbf{x})} = \overline{\lambda}\,\overline{(\mathbf{y}, \mathbf{x})} = \overline{\lambda}(\mathbf{x}, \mathbf{y}) \neq \lambda(\mathbf{x}, \mathbf{y})$ (вообще говоря).

Теорема. *Для любых векторов* $\mathbf{x}$, $\mathbf{y}$ *имеет место неравенство* $|(\mathbf{x}, \mathbf{y})|^2 \leqslant (\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})$ (*неравенство Коши–Буняковского*).

Доказательство. Согласно 4° для любых векторов $\mathbf{x}$, $\mathbf{y}$ и любого числа λ имеет место неравенство $(\lambda\mathbf{x}-\mathbf{y}, \lambda\mathbf{x}-\mathbf{y}) \geqslant 0$. В соответствии с замечанием 2, преобразуем это неравенство так:

$$(\lambda\mathbf{x}-\mathbf{y}, \lambda\mathbf{x}-\mathbf{y}) = \lambda(\mathbf{x}, \lambda\mathbf{x}-\mathbf{y}) - (\mathbf{y}, \lambda\mathbf{x}-\mathbf{y}) =$$
$$= \lambda\overline{\lambda}(\mathbf{x},\mathbf{x}) - \lambda(\mathbf{x},\mathbf{y}) - \overline{\lambda}(\mathbf{y},\mathbf{x}) + (\mathbf{y},\mathbf{y}) =$$
$$= |\lambda|^2(\mathbf{x},\mathbf{x}) - \lambda(\mathbf{x},\mathbf{y}) - \overline{\lambda}\,\overline{(\mathbf{x},\mathbf{y})} + (\mathbf{y},\mathbf{y}) \geqslant 0.$$

Полагая $\lambda = \dfrac{\overline{(\mathbf{x},\mathbf{y})}}{|(\mathbf{x},\mathbf{y})|}t$ (при $(\mathbf{x},\mathbf{y}) = 0$ справедливость неравенства Коши–Буняковского не вызывает сомнения), где t — произвольное вещественное число, получаем: $(\mathbf{x},\mathbf{x})t^2 - 2|(\mathbf{x},\mathbf{y})|t + (\mathbf{y},\mathbf{y}) \geqslant 0$. Из произвольности t следует, что дискриминант левой части неположителен: $|(\mathbf{x},\mathbf{y})|^2 - (\mathbf{x},\mathbf{x})(\mathbf{y},\mathbf{y}) \leqslant 0$, что и требовалось доказать.

Замечание. В унитарном пространстве можно ввести понятия длины вектора ($|\mathbf{x}| = \sqrt{(\mathbf{x},\mathbf{x})}$), ортонормированного базиса в $\mathbf{U}^n$ и ортогонального дополнения подпространства $\mathbf{L}$ пространства $\mathbf{U}^n$. Правда, доказательство существования ортонормированного базиса здесь выглядит иначе — ведь скалярное произведение в этом пространстве уже не является билинейной формой. Это, впрочем, не мешает провести доказательство, применяя к произвольному базису процесс ортогонализации (п. 2 § 2 гл. 4).

Не составляет труда перенести на случай унитарного пространства и результаты, полученные в пп. 3 и 4 § 2 гл. 4. При этом следует иметь в виду, что в роли матрицы A^{tr} здесь выступает матрица $\overline{A}^{tr}$. Так, например, альтернатива Фредгольма для системы линейных уравнений с комплексными коэффициентами формулируется следующим образом: *либо система линейных уравнений* $A\mathbf{x} = \mathbf{b}$ *имеет решение при любых* $\mathbf{b}$, *либо система* $\overline{A}^{tr}\mathbf{x} = \mathbf{o}$ *имеет нетривиальное решение.*

2. Нормальный оператор. Как и прежде, будем говорить, что оператор A^* *называется сопряженным по отношению к оператору* A, если для любых векторов $\mathbf{x}$, $\mathbf{y}$ имеет место равенство $(A\mathbf{x},\mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$. Нетрудно доказать, что *оператор* A^* *является сопряженным по отношению к оператору* A *тогда и только тогда, когда в ортонормированном базисе матрица* A^* *равна* $\overline{A}^{tr}$ — это утверждение доказывается точно так же, как для евклидова пространства (п. 1 § 3 гл. 4). Из этого, в частности, следует, что *для любого оператора* A *существует единственный оператор* A^*, причем $(A^*)^* = A$.

Введем теперь новое понятие.

Определение. *Оператор* A *называется нормальным, если* $AA^* = A^*A$.

Теорема 1. *Если оператор* A *— нормальный*, и $A\mathbf{x} = \lambda\mathbf{x}$, то $A^*\mathbf{x} = \overline{\lambda}\mathbf{x}$.

Доказательство. Равенство $A\mathbf{x} = \lambda\mathbf{x}$ можно переписать так: $(A - \lambda E)\mathbf{x} = \mathbf{o}$ или $((A - \lambda E)\mathbf{x}, (A - \lambda E)\mathbf{x})) = 0$. Но

$$((A - \lambda E)\mathbf{x}, (A-\lambda E)\mathbf{x})) = ((A^* - \overline{\lambda}E)(A-\lambda E)\mathbf{x}, \mathbf{x}) =$$
$$= ((A - \lambda E)(A^*-\overline{\lambda}E)\mathbf{x}, \mathbf{x}) = ((A^*-\overline{\lambda}E)\mathbf{x}, (A^*-\overline{\lambda}E)\mathbf{x})).$$

Следовательно, $A^*\mathbf{x} = \overline{\lambda}\mathbf{x}$. Теорема доказана.

Теорема 2. *Для любого нормального оператора существует ортонормированный базис из его собственных векторов.*

Доказательство. Доказательство этой теоремы аналогично доказательству теоремы о собственных векторах самосопряженного оператора в $\mathbb{E}^n$. Прежде всего заметим, что если $\mathbf{x}$ — собственный вектор нормального оператора A, $\mathbf{y}\perp\mathbf{x}$, то $A\mathbf{y}\perp\mathbf{x}$. В самом деле, $(\mathbf{x}, A\mathbf{y}) = (A^*\mathbf{x}, \mathbf{y}) = (\overline{\lambda}\mathbf{x}, \mathbf{y}) = \overline{\lambda}(\mathbf{x}, \mathbf{y}) = 0$. Таким образом, ортогональное дополнение к $\mathbf{L}(\mathbf{x})$ является инвариантным подпространством оператора A.

Дальнейшие рассуждения дословно совпадают с соответствующей частью доказательства теоремы о собственных векторах самосопряженного оператора в $\mathbb{E}^n$ (т. 2 п. 4 § 3 гл. 4).

Следствие. *Для любого нормального оператора A существует такой ортонормированный базис, в котором его матрица диагональна, причем диагональными элементами являются его собственные значения.*

Замечание. Справедливо и обратное утверждение: *если существует ортонормированный базис, в котором матрица оператора диагональна, то этот оператор — нормальный*, поскольку в указанном базисе и AA^*, и A^*A представляют собой диагональную матрицу с диагональными элементами $|a_{ii}|^2$.

3. Унитарный оператор.

Определение. *Оператор A называется унитарным, если для любых векторов $\mathbf{x}$, $\mathbf{y}$ имеет место равенство* $(A\mathbf{x}, A\mathbf{y}) = (\mathbf{x}, \mathbf{y})$.

Замечание. Поскольку унитарный оператор не меняет скалярного произведения, то любой ортонормированный базис он переводит в ортонормированный базис. Опираясь на аналогию с ортогональным оператором, можно сказать, что это в каком-то смысле соответствует повороту и отражению относительно координатных плоскостей.

Точно так же, как для ортогонального оператора в $\mathbb{E}^n$, доказывается, что *оператор A является унитарным тогда и только тогда, когда* $A^* = A^{-1}$. Из этого, в свою очередь, следует, что:

1° *любой унитарный оператор является нормальным*;

2° *в ортонормированном базисе матрица унитарного оператора удовлетворяет равенству* $A^{-1} = \overline{A}^{tr}$.

Отметим, что матрица, удовлетворяющая равенству $A^{-1} = \overline{A}^{tr}$, называется *унитарной*. *При помощи унитарных матриц осуществляется переход от одного ортонормированного базиса к другому*

(доказательство этого утверждения полностью аналогично доказательству соответствующего утверждения для ортогональных матриц в $\mathbb{E}^n$). Из этого, в частности, следует, что *произведение двух унитарных матриц является унитарной матрицей*.

Теорема. *Все собственные значения унитарного оператора по модулю равны* 1.

Доказательство. Пусть $A^* = A^{-1}$, $A\mathbf{x} = \lambda\mathbf{x}$. Имеем: $A^*A\mathbf{x} = E\mathbf{x} = \mathbf{x}$. С другой стороны, согласно теореме 1 п. 2, $A^*A\mathbf{x} = A^*\lambda\mathbf{x} = \lambda A^*\mathbf{x} = \lambda\overline{\lambda}\mathbf{x}$. Тем самым, $\lambda\overline{\lambda} = |\lambda|^2 = 1$, что и требовалось доказать.

Следствие. *В базисе из собственных векторов матрица унитарного оператора имеет вид*

$$\begin{pmatrix} e^{i\varphi_1} & 0 & \dots & 0 \\ 0 & e^{i\varphi_2} & \dots & 0 \\ 0 & 0 & \dots & e^{i\varphi_n} \end{pmatrix},$$

поскольку любое комплексное число λ можно представить в виде $\lambda = |\lambda|e^{i\varphi}$.

4. Самосопряженный оператор. Оператор A называется *самосопряженным* или *эрмитовым*, если $A^* = A$. Ясно, что в ортонормированном базисе его матрица удовлетворяет равенству: $A = \overline{A}^{tr}$. Отметим, что матрица A, обладающая этим свойством, называется *эрмитовой*.

Непосредственно из определения следует, что *самосопряженный оператор является нормальным оператором*.

Теорема. *Все собственные значения самосопряженного оператора вещественные*.

Доказательство. Пусть $A^* = A$, $A\mathbf{x} = \lambda\mathbf{x}$. Согласно теореме 1 п. 2 $A^*\mathbf{x} = \overline{\lambda}\mathbf{x}$. Тем самым, $\lambda = \overline{\lambda}$, что и требовалось доказать.

§ 2. Псевдоевклидово пространство

1. Определение. Снова вернемся к линейному пространству над полем $\mathbb{R}$, но введем теперь скалярное произведение иначе — без требования положительной определенности.

Определение 1. *Линейное пространство* $\mathbf{L}^n$ *над полем* $\mathbb{R}$ *называется псевдоевклидовым, если в нем фиксирована некоторая симметричная билинейная форма* $(\mathbf{x}, \mathbf{y})$ *ранга* n, *называемая скалярным произведением*.

Определение 2. *Пусть* k — *положительный индекс инерции квадратичной формы* $(\mathbf{x}, \mathbf{x})$, *и, следовательно,* $(n-k)$ — *ее отрицательный индекс инерции. Пара чисел* $(k, n-k)$ *называется сигнатурой псевдоевклидова пространства*.

Псевдоевклидово пространство с сигнатурой $(k, n-k)$ обозначается так: $\mathbb{E}^n_{k,n-k}$. В частности, $\mathbb{E}^n_{n,0}$ — это обычное евклидово пространство $\mathbb{E}^n$.

Для скалярного произведения, как и для всякой симметричной билинейной формы, существует канонический базис $\{\mathbf{e}_i\}$ (канонический базис квадратичной формы $(\mathbf{x},\mathbf{x})$), в котором

$$(\mathbf{e}_i,\mathbf{e}_j)=\begin{cases}1, & i=j\leqslant k,\\ -1, & i=j>k,\\ 0, & i\neq j.\end{cases}$$

а скалярное произведение имеет вид $(\mathbf{x},\mathbf{y}) = x^1y^1+\ldots+x^ky^k - x^{k+1}y^{k+1}-\ldots-x^ny^n$. Такой базис называется *галилеевым базисом*. В нем, в частности, $(\mathbf{x},\mathbf{x})=(x^1)^2+\ldots+(x^k)^2-(x^{k+1})^2-\ldots-(x^n)^2$. Нетрудно видеть, что величина $(\mathbf{x},\mathbf{x})$ может быть как положительной, так и неположительной.

Пусть O — произвольная точка, $\mathbf{OM}=\mathbf{x}$. Множество всех векторов $\mathbf{x}$, для которых $(\mathbf{x},\mathbf{x})=0$, называется *световым конусом* (почему этот конус называется световым, станет ясно чуть позже).

Пространство $\mathbb{E}^{n+1}_{1,n}$ называется *пространством Минковского*. Обычно галилеев базис пространства Минковского обозначают так: $\mathbf{e}_0,\mathbf{e}_1,\ldots,\mathbf{e}_n$. Уравнение светового конуса в этом базисе имеет вид $\left(x^0\right)^2 = \sum\limits_{i=1}^{n}\left(x^i\right)^2$, и, следовательно, направляющей этого конуса является n-мерная сфера $\sum\limits_{i=1}^{n}\left(x^i\right)^2=\mathrm{const}$.

Особый интерес представляет пространство Минковского при $n=3$, поскольку оно является *пространством событий* в специальной теории относительности. В этой теории под x^1, x^2, x^3 понимают координаты частицы, а под x^0-ct, где c — скорость света, а t — время. Каждая частица в пространстве событий движется по некоторой траектории (движется всегда, так как меняется t), называемой *мировой линией*. Фотон движется прямолинейно со скоростью c: $x^i=v^it$, где $(v^1)^2+(v^2)^2+(v^3)^2=c^2$, т. е. по образующей светового конуса. Отсюда и название «световой конус» — по нему движется свет.

2. Преобразования Лоренца.

Определение. *Переход от одного галилеева базиса к другому галилееву базису называется преобразованием Лоренца.*

Уточним вид преобразования Лоренца в пространстве Минковского. Прежде всего, заметим, что переход от галилеева базиса $\mathbf{e}_0,\mathbf{e}_1,\ldots,\mathbf{e}_n$ к галилееву базису $\mathbf{e}_0,\widetilde{\mathbf{e}}_1,\ldots,\widetilde{\mathbf{e}}_n$ — это обычное ортогональное преобразование в пространстве $\mathbf{L}(\mathbf{e}_1,\ldots,\mathbf{e}_n)$, т. е. преобразование с ортогональной матрицей α. В самом деле, $\widetilde{\mathbf{e}}_i=\alpha_i^s\mathbf{e}_s$, $\widetilde{\mathbf{e}}_j=\alpha_j^r\mathbf{e}_r$, причем $(\widetilde{\mathbf{e}}_i,\widetilde{\mathbf{e}}_j)=(\mathbf{e}_i,\mathbf{e}_j)=\begin{cases}-1 & \text{при} \quad i=j,\\ 0 & \text{при} \quad i\neq j.\end{cases}$ Имеем: $(\widetilde{\mathbf{e}}_i,\widetilde{\mathbf{e}}_j)=(\alpha_i^s\mathbf{e}_s,\alpha_j^r\mathbf{e}_r)=\alpha_i^s\alpha_j^r(\mathbf{e}_s,\mathbf{e}_r)=\sum\limits_s\alpha_i^s\alpha_j^s(-1)$, откуда $\alpha^{tr}\alpha=E$, т. е. $\alpha^{-1}=\alpha^{tr}$. Ясно, что переход от базиса $\mathbf{e}_0,\mathbf{e}_1,\ldots,\mathbf{e}_n$ к базису $-\mathbf{e}_0,\mathbf{e}_1,\ldots,\mathbf{e}_n$ —

это также ортогональное преобразование, поскольку его матрица $\alpha = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ ортогональна.

Рассмотрим теперь общий случай.

Пусть $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$ и $\widetilde{\mathbf{e}}_0, \widetilde{\mathbf{e}}_1, \dots, \widetilde{\mathbf{e}}_n$ — два галилеевых базиса. Будем считать, что $(\mathbf{e}_0, \widetilde{\mathbf{e}}_0) \geqslant 0$ (в противном случае перейдем от базиса $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$ к базису $-\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$ ортогональным преобразованием). Поступим следующим образом.

1°. Выберем в пространстве $L(\mathbf{e}_0, \widetilde{\mathbf{e}}_0)$ вектор $\mathbf{e}'_1$, для которого $(\mathbf{e}'_1, \mathbf{e}_0) = 0$, т. е. $\mathbf{e}'_1 \in L(\mathbf{e}_1, \dots, \mathbf{e}_n)$, и перейдем от базиса $\mathbf{e}_0, \mathbf{e}_1, \dots$ $\dots, \mathbf{e}_n$ ортогональным преобразованием к базису $\mathbf{e}_0, \mathbf{e}'_1, \dots, \mathbf{e}'_n$. В нем $\mathbf{e}'_1 \in L(\mathbf{e}_0, \widetilde{\mathbf{e}}_0)$.

2°. Преобразованием в пространстве $L(\mathbf{e}_0, \widetilde{\mathbf{e}}_0)$ перейдем от базиса $\mathbf{e}_0, \mathbf{e}'_1$ к базису $\widetilde{\mathbf{e}}_0, \mathbf{e}''_1$, в котором $(\mathbf{e}''_1 \widetilde{\mathbf{e}}_0) = 0$, $(\mathbf{e}''_1, \mathbf{e}''_1) = -1$ и $(\mathbf{e}'_1, \mathbf{e}''_1) \geqslant 0$.

3°. Наконец, перейдем от базиса $\widetilde{\mathbf{e}}_0, \mathbf{e}''_1, \dots, \mathbf{e}'_n$ ортогональным преобразованием к базису $\widetilde{\mathbf{e}}_0, \widetilde{\mathbf{e}}_1, \dots, \widetilde{\mathbf{e}}_n$.

Выясним, как осуществляется переход 2°. Имеем:

$$\left.\begin{aligned} \widetilde{\mathbf{e}}_0 &= \alpha_0^0 \mathbf{e}_0 + \alpha_0^1 \mathbf{e}'_1 \\ \mathbf{e}''_1 &= \alpha_1^0 \mathbf{e}_0 + \alpha_1^1 \mathbf{e}'_1 \end{aligned}\right\},$$

где $\alpha_0^0 = (\mathbf{e}_0, \widetilde{\mathbf{e}}_0) \geqslant 0$ и $\alpha_1^1 = (\mathbf{e}'_1, \mathbf{e}''_1) \geqslant 0$. Далее, поскольку $(\widetilde{\mathbf{e}}_0, \widetilde{\mathbf{e}}_0) = 1$, $(\widetilde{\mathbf{e}}_0, \mathbf{e}''_1) = 0$ и $(\mathbf{e}''_1, \mathbf{e}''_1) = -1$, то

$$\left.\begin{aligned} \left(\alpha_0^0\right)^2 - \left(\alpha_0^1\right)^2 &= 1 \\ \alpha_0^0 \alpha_1^0 - \alpha_0^1 \alpha_1^1 &= 0 \\ \left(\alpha_1^1\right)^2 - \left(\alpha_1^0\right)^2 &= 1 \end{aligned}\right\}.$$

Полагая $\alpha_0^1 = \operatorname{sh}\varphi$, $\alpha_1^0 = \operatorname{sh}\theta$, из первого и третьего уравнения находим: $\alpha_0^0 = \operatorname{ch}\varphi$, $\alpha_1^1 = \operatorname{ch}\theta$. Подставляя эти значения во второе уравнение, получаем: $\operatorname{ch}\varphi \operatorname{sh}\theta - \operatorname{sh}\varphi \operatorname{ch}\theta = \operatorname{sh}(\varphi - \theta) = 0$, откуда $\varphi = \theta$. Итак, переход 2° осуществляется с помощью матрицы

$$\alpha = \begin{pmatrix} \operatorname{ch}\varphi & \operatorname{sh}\varphi \\ \operatorname{sh}\varphi & \operatorname{ch}\varphi \end{pmatrix}.$$

Это преобразование называется *гиперболическим поворотом*. В специальной теории относительности его обычно записывают несколько иначе, полагая $\operatorname{th}\varphi = a$:

$$\alpha = \begin{pmatrix} \dfrac{1}{\sqrt{1-a^2}} & \dfrac{a}{\sqrt{1-a^2}} \\ \dfrac{a}{\sqrt{1-a^2}} & \dfrac{1}{\sqrt{1-a^2}} \end{pmatrix}.$$

Итак, *любое преобразование Лоренца в пространстве Минковского представляет собой суперпозицию* (т. е. последовательное выполнение) *ортогональных преобразований и гиперболического поворота*.

§ 3. Группы и поля

1. Группа. Сделанные нами обобщения до сих пор касались только евклидова пространства. В этом параграфе мы рассмотрим некоторые обобщения понятия линейного пространства.

Определение 1. *Множество* $\mathbf{G}$ *называется группой, если для любой упорядоченной пары его элементов* $\mathbf{a}$ *и* $\mathbf{b}$ *определена операция, ставящая ей в соответствие единственный элемент* $\mathbf{ab} \in \mathbf{G}$. *При этом*:

1° *для любых* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *имеет место равенство* $\mathbf{a}(\mathbf{bc}) = (\mathbf{ab})\mathbf{c}$;

2° *существует такой элемент* $\mathbf{e} \in \mathbf{G}$, что *для любого* $\mathbf{a} \in \mathbf{G}$ *имеет место равенство* $\mathbf{ae} = \mathbf{a}$;

3° *среди всех элементов* $\mathbf{e}$, *обладающих свойством* 2°, *существует такой элемент* $\mathbf{e}_0$, *что для любого* $\mathbf{a} \in \mathbf{G}$ *существует элемент* $\mathbf{a}^{-1} \in \mathbf{G}$, *удовлетворяющий равенству* $\mathbf{aa}^{-1} = \mathbf{e}_0$.

Если, сверх того, для любых $\mathbf{a}$, $\mathbf{b}$ имеет место равенство $\mathbf{ab} = \mathbf{ba}$, то группа называется *коммутативной* или *абелевой*.

Замечание. Если операцию обозначать знаком «+», то вместо $\mathbf{e}$ естественно писать $\mathbf{o}$, а вместо $\mathbf{a}^{-1}$ – $(-\mathbf{a})$, хотя, конечно, это только естественная форма записи, поскольку операция в группе абстрактна, она не имеет отношения к обычному умножению или сложению.

Рассмотрим несколько свойств группы.

Свойство 1. *Элемент* $\mathbf{e}$, *обладающий свойством* 2°, *определяется единственным образом и, следовательно, совпадает с* $\mathbf{e}_0$. В самом деле, «умножая» обе части равенства $\mathbf{ee}^{-1} = \mathbf{e}_0$ слева на $\mathbf{e}$, получаем: $\mathbf{eee}^{-1} = \mathbf{ee}_0$ или $\mathbf{ee}^{-1} = \mathbf{e}$, откуда $\mathbf{e}_0 = \mathbf{e}$.

Это свойство позволяет в дальнейшем вместо $\mathbf{e}_0$ писать просто $\mathbf{e}$.

Замечание. В большинстве учебников свойство 3° формулируют кратко: *для любого* $\mathbf{a} \in \mathbf{G}$ *существует элемент* $\mathbf{a}^{-1} \in \mathbf{G}$, *удовлетворяющий равенству* $\mathbf{aa}^{-1} = \mathbf{e}$. Такая формулировка, однако, не вполне корректна, поскольку она допускает неоднозначную трактовку. В самом деле, наряду с нашей трактовкой (она, по-видимому, наиболее естественна с точки зрения норм русского языка) возможна и, например, такая: *для любого* $\mathbf{a} \in \mathbf{G}$ *существует такой элемент* $\mathbf{a}^{-1} \in \mathbf{G}$, *что элемент* $\mathbf{e} = \mathbf{aa}^{-1}$ *обладает свойством* 2°. При такой трактовке установленное нами свойство 1 уже не будет иметь места. Например, в множестве с операцией $\mathbf{ab} = \mathbf{a}$ любой элемент можно принять за $\mathbf{e}$, а значит и за $\mathbf{a}^{-1}$. И хотя формально эта операция удовлетворяет всем свойствам 1°–3°, группой в общепринятом смысле такое множество не является.

Свойство 2. Имеем: $\mathbf{a}^{-1}\mathbf{a} = \mathbf{a}^{-1}\mathbf{a}\mathbf{e} = \mathbf{a}^{-1}\mathbf{a}\mathbf{a}^{-1}(\mathbf{a}^{-1})^{-1} = \mathbf{a}^{-1}\mathbf{e}(\mathbf{a}^{-1})^{-1} = \mathbf{a}^{-1}(\mathbf{a}^{-1})^{-1} = \mathbf{e}$, т. е. $\mathbf{a}^{-1}\mathbf{a} = \mathbf{e}$.

Свойство 3. Имеем: $\mathbf{e}\mathbf{a} = \mathbf{a}\mathbf{a}^{-1}\mathbf{a} = \mathbf{a}\mathbf{e} = \mathbf{a}$, т. е. $\mathbf{e}\mathbf{a} = \mathbf{a}$.

Свойство 4. *Элемент* $\mathbf{a}^{-1}$, *обладающий свойством* 3°, *определяется единственным образом*. В самом деле, допустим, что есть два таких элемента: a_1^{-1} и a_2^{-1}. Имеем: $a_1^{-1}\mathbf{a}a_2^{-1} = a_1^{-1}\mathbf{e} = a_1^{-1}$. С другой стороны, $a_1^{-1}\mathbf{a}a_2^{-1} = \mathbf{e}a_2^{-1} = a_2^{-1}$. Следовательно, $a_1^{-1} = a_2^{-1}$.

Свойство 5. *Для любых* $\mathbf{a}, \mathbf{b} \in \mathbf{G}$ *существует единственное решение* $\mathbf{x}$ *уравнения* $\mathbf{a}\mathbf{x} = \mathbf{b}$ *и единственное решение* $\mathbf{y}$ *уравнения* $\mathbf{y}\mathbf{a} = \mathbf{b}$. В самом деле, рассмотрим, например, первое уравнение (для второго доказательство аналогичное). Если решение $\mathbf{x}$ существует, то, «умножая» обе части уравнения слева на $\mathbf{a}^{-1}$, получим: $\mathbf{x} = \mathbf{a}^{-1}\mathbf{b}$, т. е. $\mathbf{x}$ определяется единственным образом. С другой стороны, $\mathbf{x} = \mathbf{a}^{-1}\mathbf{b}$ действительно является решением, поскольку $\mathbf{a}\mathbf{x} = \mathbf{a}\mathbf{a}^{-1}\mathbf{b} = \mathbf{e}\mathbf{b} = \mathbf{b}$, т. е. решение существует.

Замечание. В абелевой группе $\mathbf{x} = \mathbf{a}^{-1}\mathbf{b} = \mathbf{b}\mathbf{a}^{-1} = \mathbf{y}$, что дает возможность говорить о наличии в ней операции «деления»: $\mathbf{x} = \mathbf{y} = \mathbf{b}/\mathbf{a}$.

Определение 2. *Подмножество* $\mathbf{G}'$ *группы* $\mathbf{G}$ *называется подгруппой, если* $\mathbf{G}'$ *— группа с той же операцией, что и в* $\mathbf{G}$.

Теорема. *Подмножество* $\mathbf{G}'$ *группы* $\mathbf{G}$ *является подгруппой тогда и только тогда, когда оно обладает двумя свойствами*:

1° *для любых двух элементов* $\mathbf{a}$ *и* $\mathbf{b}$ *множества* $\mathbf{G}'$ *элемент* $\mathbf{a}\mathbf{b}$ *принадлежит* $\mathbf{G}'$;

2° *для любого* $\mathbf{a} \in \mathbf{G}'$ *элемент* $\mathbf{a}^{-1}$ *принадлежит* $\mathbf{G}'$.

Доказательство. Если множество $\mathbf{G}'$ является подгруппой, то оно обладает свойствами 1° и 2° — это следует из определения группы. Допустим, что для некоторого подмножества $\mathbf{G}'$ группы $\mathbf{G}$ выполнены требования 1° и 2°. Тогда для любых элементов $\mathbf{a}$, $\mathbf{b}$ множества $\mathbf{G}'$определена операция, ставящая им в соответствие элемент $\mathbf{a}\mathbf{b} \in \mathbf{G}'$ и обладающая свойством 1° группы. Кроме того, для любого элемента $\mathbf{a}$ множества $\mathbf{G}'$ элемент $\mathbf{a}^{-1}$, а значит и элемент $\mathbf{e} = \mathbf{a}\mathbf{a}^{-1}$, принадлежат $\mathbf{G}'$, поэтому множество $\mathbf{G}'$ обладает и свойствами 2°, 3° группы. Теорема доказана.

2. Примеры.

Пример 1. Линейное пространство является абелевой группой, поскольку операция сложения в нем обладает следующими свойствами:

1° для любых векторов $\mathbf{a}$ и $\mathbf{b}$ $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$;

2° для любых векторов $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$;

3° существует такой вектор $\mathbf{o}$, что для любого вектора $\mathbf{a}$ $\mathbf{a} + \mathbf{o} = \mathbf{a}$;

4° для любого вектора $\mathbf{a}$ существует вектор $(-\mathbf{a})$ такой, что $\mathbf{a} + (-\mathbf{a}) = \mathbf{o}$.

Если отказаться от свойства 1°, а свойство 4° трактовать так, как его трактуют в определении группы (см. замечание предыдущего пункта), то и в этом случае линейное пространство останется группой.

Более того, оно останется абелевой группой, т. е. свойством 1° оно попрежнему будет обладать. В самом деле, имеем: $(1+1)(\mathbf{a}+\mathbf{b}) = 1(\mathbf{a}+\mathbf{b})+1(\mathbf{a}+\mathbf{b}) = \mathbf{a}+\mathbf{b}+\mathbf{a}+\mathbf{b}$. С другой стороны, $(1+1)(\mathbf{a}+\mathbf{b}) = (1+1)\mathbf{a}+(1+1)\mathbf{b} = \mathbf{a}+\mathbf{a}+\mathbf{b}+\mathbf{b}$. Итак, $\mathbf{a}+\mathbf{b}+\mathbf{a}+\mathbf{b} = \mathbf{a}+\mathbf{a}+\mathbf{b}+\mathbf{b}$. Прибавляя к обеим частям этого равенства справа $(-\mathbf{b})$, а слева — $(-\mathbf{a})$, получим: $\mathbf{b}+\mathbf{a} = \mathbf{a}+\mathbf{b}$. Таким образом, при наиболее естественной трактовке аксиомы 4°, аксиома 1° оказывается лишней.

Пример 2. Комплексные $(n \times n)$-матрицы с отличным от нуля определителем образуют группу относительно операции умножения. Эту группу называют *полной линейной группой* (general linear group) и обозначают так: $\mathrm{GL}(n)$.

Полная линейная группа имеет разнообразные подгруппы. Рассмотрим некоторые из них.

Пример 3. Комплексные $(n \times n)$-матрицы с определителем, равным 1, образуют, очевидно, подгруппу группы $\mathrm{GL}(n)$. Эту группу называют *унимодулярной группой* и обозначают так: $\mathrm{SL}(n)$.

Пример 4. Ортогональные $(n \times n)$-матрицы также образуют подгруппу группы $\mathrm{GL}(n)$. В самом деле, с помощью ортогональных матриц в евклидовом пространстве осуществляется переход от одного ортонормированного базиса к другому. Поэтому если матрицы A и B — ортогональные, то матрицы AB и A^{-1} также ортогональные. Эту группу называют *ортогональной группой* и обозначают так: $\mathrm{O}(n)$.

Примеры 5, 6. По аналогичной причине подгруппами группы $\mathrm{GL}(n)$ являются множество унитарных матриц и множество матриц, с помощью которых осуществляются преобразования Лоренца в пространстве $\mathbf{E}^n_{p,q}$. Первую группу называют *унитарной группой* $(\mathrm{U}(n))$, а вторую — *псевдоортогональной группой* или *группой Лоренца* $(\mathrm{O}(p,q))$.

Примеры 7, 8, 9. Мы получим еще три примера подгрупп группы $\mathrm{GL}(n)$, если в каждой из трех последних групп выделим лишь те матрицы, определители которых равны 1. Полученные таким образом группы обозначаются соответственно $\mathrm{SO}(n)$, $\mathrm{SU}(n)$ и $\mathrm{SO}(p,q)$.

3. Поле.

Определение. *Множество* $\mathbf{K}$ *называется полем, если для любой упорядоченной пары его элементов* $\mathbf{a}$ *и* $\mathbf{b}$ *определены две операции, называемые сложением и умножением, ставящие ей в соответствие элементы* $\mathbf{a}+\mathbf{b} \in \mathbf{K}$ и $\mathbf{ab} \in K$. *При этом*:

1° *множество* $\mathbf{K}$ *является абелевой группой относительно операции сложения*;

2° *множество всех элементов* $\mathbf{K}$, *отличных от* $\mathbf{o}$, *является абелевой группой относительно операции умножения*;

3° *для любых* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{K}$ *имеет место равенство* $\mathbf{a}(\mathbf{b}+\mathbf{c}) = \mathbf{ab}+\mathbf{ac}$;

4° *множество* $\mathbf{K}$ *содержит не менее двух элементов*.

Рассмотрим несколько свойств поля.

Свойство 1. *Для любых* $\mathbf{a}, \mathbf{b} \in \mathbf{K}$ *существует единственный элемент* $\mathbf{x} \in K$ *— его называют разностью* $\mathbf{b} - \mathbf{a}$, *являющийся решением уравнения* $\mathbf{a} + \mathbf{x} = \mathbf{x} + \mathbf{a} = \mathbf{b}$.

Свойство 2. Имеем: $\mathbf{a}(\mathbf{b} - \mathbf{c}) + \mathbf{ac} = \mathbf{a}(\mathbf{b} - \mathbf{c} + \mathbf{c}) = \mathbf{ab}$, откуда $\mathbf{a}(\mathbf{b} - \mathbf{c}) = \mathbf{ab} - \mathbf{ac}$.

Свойство 3. Имеем: $\mathbf{ao} = \mathbf{a}(\mathbf{a} - \mathbf{a}) = \mathbf{aa} - \mathbf{aa} = \mathbf{o}$, т. е. $\mathbf{ao} = \mathbf{oa} = \mathbf{o}$.

Свойство 4. *Для любых* $\mathbf{a}, \mathbf{b} \in \mathbf{K}$ при $\mathbf{a} \neq \mathbf{o}$ *существует единственный элемент* $\mathbf{x} \in \mathbf{K}$ *— его называют отношением* $\mathbf{b}/\mathbf{a}$, *являющийся решением уравнения* $\mathbf{ax} = \mathbf{xa} = \mathbf{b}$. В самом деле, при $\mathbf{b} \neq \mathbf{o}$ это следует из определения поля и свойств абелевой группы; при $\mathbf{b} = \mathbf{o}$ рассматриваемое уравнение эквивалентно уравнению $\mathbf{x} = \mathbf{a}^{-1}\mathbf{o}$, т. е. уравнению $\mathbf{x} = \mathbf{o}$.

Следствие. *Равенство* $\mathbf{ab} = \mathbf{o}$ *возможно только в тех случаях, когда либо* $\mathbf{a} = \mathbf{o}$, *либо* $\mathbf{b} = \mathbf{o}$ (поскольку если $\mathbf{a} \neq \mathbf{o}$, то $\mathbf{b} = \mathbf{o}$, и наоборот). Обычно это утверждение формулируют так: *в поле отсутствуют делители нуля*.

Приведем теперь несколько примеров полей.

Пример 1. Множество $\mathbb{C}$ комплексных чисел.

Примеры 2, 3. Множество $\mathbb{R}$ вещественных чисел и множество $\mathbb{Q}$ рациональных чисел являются *подполями* поля $\mathbb{C}$. Подполя поля $\mathbb{C}$ комплексных чисел называют *числовыми полями*. Отметим, что числовое поле содержит 1, поэтому оно содержит все целые числа, а значит и все рациональные числа. Тем самым, *любое числовое поле содержит в себе поле* $\mathbb{Q}$.

Примеры 4, 5. Существуют ли числовые поля, отличные от $\mathbb{C}$, $\mathbb{R}$ и $\mathbb{Q}$? Конечно! Приведем два примера. Первый пример — множество всех вещественных чисел вида $a+b\sqrt{2}$, где числа a и b — рациональные. Другой пример — множество всех комплексных чисел вида $a + ib$, где числа a и b — также рациональные.

Пример 6. Поскольку любое числовое поле содержит в себе поле $\mathbb{Q}$, то каждое из них содержит бесконечное количество элементов. А существуют ли поля, состоящие из конечного числа элементов? Оказывается, существуют. Приведем пример.

Условимся под формулой $b = a \pmod p$ понимать следующее: $b - a = kp$, где k — целое число. Например, в этих обозначениях формула $\sin x = 1$ эквивалентна формуле $x = \dfrac{\pi}{2} \pmod{2\pi}$.

Пусть p — натуральное число большее 1. Рассмотрим множество $\mathbf{Z}_p$ чисел: $a_0 = 0$, $a_1 = 1, \ldots, a_{p-1} = p - 1$. Ясно, что для любого целого числа b существует единственное число $a \in \mathbf{Z}_p$, для которого $a = b \pmod p$. Определим на множестве $\mathbf{Z}_p$ две операции: $a \oplus b = (a + b) \pmod p$ и $a \otimes b = (ab) \pmod p$ так, чтобы $a \oplus b \in \mathbf{Z}_p$ и $a \otimes b \in \mathbf{Z}_p$. Тогда окажется, что:

множество $\mathbf{Z}_p$ представляет собой абелеву группу относительно операции $\oplus$ (роль $\mathbf{o}$ играет a_0, а роль $(-a_i)$ — число a_{p-i});

$\mathbf{a} \otimes (\mathbf{b} \oplus \mathbf{c}) = \mathbf{a} \otimes \mathbf{b} \oplus \mathbf{a} \otimes \mathbf{c}$;

$\mathbf{a} \otimes \mathbf{b} = \mathbf{b} \otimes \mathbf{a}$;

$\mathbf{a} \otimes (\mathbf{b} \otimes \mathbf{c}) = (\mathbf{a} \otimes \mathbf{b}) \otimes \mathbf{c}$;

для любого $a \in \mathbf{Z}_p$ $a \otimes a_1 = a$ (поскольку $a_1 = 1$), т. е. роль $\mathbf{e}$ играет $a_1 = 1$.

Тем не менее, множество $\mathbf{Z}_p$ может и не быть полем. В самом деле, если число p — составное, т. е. $p = mn$, то $a_m \otimes a_n = \mathbf{o}$, в то время как в поле делители нуля отсутствуют.

Если же число p — простое, то множество $\mathbf{Z}_p$ — поле (оно называется *полем сравнений по модулю* p). Действительно, нам осталось доказать, что для любого $a \in \mathbf{Z}_p$ $(a \neq \mathbf{o})$ существует такой элемент a^{-1}, что $a \otimes a^{-1} = \mathbf{e} = 1$. Докажем это. Рассмотрим числа a, $2 \otimes a$, ..., $(p-1) \otimes a$. Все эти числа отличны от $\mathbf{o}$, поскольку произведение двух натуральных чисел меньших простого числа p не может быть кратно p. Более того, все они попарно различны (из равенства $i \otimes a = j \otimes a$ следовало бы, что $|i - j|a = \mathbf{o}$, чего, как только что отмечалось, быть не может) и меньше p. Следовательно, среди них обязательно есть число $k \otimes a = 1$. Но тогда $k = a^{-1}$.

Таким образом, возникает целая серия полей, состоящих из конечного числа элементов: $\mathbf{Z}_2$, $\mathbf{Z}_3$, $\mathbf{Z}_5$, $\mathbf{Z}_7$ и т. д.

Замечание. Возвращаясь к определению линейного пространства, мы можем сказать теперь, что под числами в нем можно понимать элементы произвольного поля. При этом, правда, изменится теория квадратичных форм (а также доказательство свойства 3° в п. 4 § 2 гл. 1 ч. I).

Заключение

В школьном курсе математики алгебра и геометрия выступают как два независимых раздела, имеющих между собой мало общего. В противоположность этому аналитическая геометрия и линейная алгебра находятся как раз на стыке этих наук, причем в первой из них превалирует геометрия, а во второй — алгебра. Образно говоря, аналитическая геометрия — это алгебраизированная геометрия, а линейная алгебра — это геометризированная алгебра. Связь между алгеброй и геометрией устанавливается в значительной мере на базе тех алгебраических фактов, которые изложены в первой части книги — «Аппарат аналитической геометрии и линейной алгебры», посвященной действиям с матрицами, теории определителей квадратных матриц и приложениям этой теории к решению систем линейных уравнений. Важную роль в последующих рассуждениях играет также введенное на первых страницах книги понятие арифметического пространства. Указанный круг вопросов, конечно же, вплотную примыкает как к аналитической геометрии, так и к линейной алгебре, однако не является, строго говоря, предметом ни той, ни другой науки.

Вторая часть, «Аналитическая геометрия», в значительной мере известна из курса геометрии средней школы (метод координат, векторы и линейные операции над ними, скалярное произведение). Принципиально новым здесь является, главным образом, осознание роли определителей. Оказывается, что определитель второго порядка с точностью до знака равен площади параллелограмма, построенного на векторах, координаты которых являются его строками, а объем параллелепипеда, построенного на данных трех векторах, равен модулю определителя, строками которого являются координаты этих векторов. Не менее важную роль играет знак определителя — он позволяет судить об ориентации упорядоченных пар или троек векторов, в частности, придать ясный геометрический смысл словам «по часовой стрелке».

Осознание роли определителей в геометрии, в свою очередь, позволяет ввести понятия векторного произведения двух векторов и смешанного произведения трех векторов. Эти понятия, при всем своем внешнем различии (векторное произведение — это вектор, а смешанное произведение — это число), тесно связаны друг с другом. Более того, с определенной точки зрения это почти одно и то же. Поясним, что имеется в виду.

Назовем смешанным произведением двух векторов на плоскости определитель, строками которого являются координаты этих векторов в правой системе координат. Ясно, что так определенное смешанное произведение двух векторов представляет собой площадь построенного на них параллелограмма, взятую со знаком «+», если они образуют правую пару, и «−» — если левую. Будем теперь рассматривать те же векторы, но не на плоскости, а в пространстве. Тогда нашему

смешанному произведению можно приписать направление — считать, что это вектор, ортогональный данным и образующий с ними правую тройку. В результате мы приходим к понятию векторного произведения двух векторов. Далее, определим смешанное произведение трех векторов в пространстве как определитель, строками которого являются координаты этих векторов в правой системе координат. Тогда оно окажется равным объему построенного на этих векторах параллелепипеда, взятому со знаком «+», если они образуют правую тройку, и «−» — если левую. Это, как мы помним, следует из формулы для объема параллелепипеда (произведение площади основания на высоту). Если теперь рассматривать те же векторы не в трехмерном, а в четырехмерном пространстве, то можно определить их векторное произведение, приписав смешанному произведению направление и т. д. Ясно, что таким способом можно определить векторное произведение и смешанное произведение n векторов соответственно в $(n+1)$-мерном и n-мерном пространстве. При этом они будут отличаться друг от друга только тем, что первому из них приписано определенное направление, а второму — нет.

Наличие векторного и смешанного произведения позволяет, как мы видели, достаточно просто вывести различные виды уравнений прямых и плоскостей. Отметим также, что указанные произведения широко используются в физике. Например, момент приложенной к точке M силы $\mathbf{F}$ относительно точки O представляет собой векторное произведение $[\mathbf{OM} \cdot \mathbf{F}]$.

Особое место в аналитической геометрии занимает раздел «Линии и поверхности второго порядка». Определения эллипса и гиперболы похожи друг на друга. Их можно даже объединить, сказав: эллипсом (гиперболой) называется множество всех таких точек плоскости, для которых сумма (модуль разности) расстояний до двух фиксированных точек есть постоянная положительная величина. Есть и другое общее определение этих кривых, включающее заодно и параболу: эллипсом, отличным от окружности (гиперболой, параболой), называется множество всех таких точек плоскости, для которых отношение расстояния до фиксированной точки к расстоянию до фиксированной прямой постоянно и меньше единицы (больше единицы, равно единице). Это сходство в определениях проявляется, в частности, в том, что эллипс, гипербола и парабола могут быть получены в качестве сечения конуса плоскостью, не проходящей через вершину. Опираясь на этот факт, можно наглядно проследить, как эти кривые переходят друг в друга при изменении угла наклона секущей плоскости.

Эллипс, гипербола и парабола играют фундаментальную роль в физике. Так, брошенный камень движется по параболе, а движение небесных тел (планет, комет, метеоритов и т. д.) под действием притяжения Солнца происходит по эллипсу или гиперболе. Широко используются и оптические свойства этих кривых. Но не это главное. Рассмотрим некоторую функцию $f(x, y)$ и поставим такой вопрос: как приближенно описать поведение этой функции в окрестности точки $(0, 0)$? Естественно воспользоваться формулой Тейлора. В качестве первого, линейного,

приближения получаем: $f(x,y) \approx f(0,0) + f_x(0,0)x + f_y(0,0)y$. В этом приближении график нашей функции, т. е. поверхность $z = f(x,y)$, представляет собой плоскость T — ее называют касательной плоскостью к поверхности $z = f(x,y)$ в точке (0,0). Как правило, линейного приближения бывает недостаточно. Второе приближение выглядит так: $f(x,y) \approx f(0,0) + f_x(0,0)x + f_y(0,0)y + \frac{1}{2}\big[f_{xx}(0,0)x^2 + 2f_{xy}(0,0)xy + f_{yy}(0,0)y^2\big]$. В этом приближении поверхность $z = f(x,y)$ представляет собой поверхность второго порядка (если, конечно, хотя бы одна из вторых производных отлична от нуля) — либо параболоид, либо параболический цилиндр. Квадратичная часть в этом уравнении, или, как ее называют в линейной алгебре, квадратичная форма, описывает, очевидно, отклонение поверхности от касательной плоскости T. Всякая квадратичная форма $a_{11}x^2 + 2a_{12}xy + ay^2$ (в нашем случае $a_{11} = \frac{1}{2}f_{xx}(0,0)$, $a_{12} = f_{xy}(0,0)$, $a_{22} = \frac{1}{2}f_{yy}(0,0)$), как мы знаем, при помощи поворота может быть приведена к виду $a_1\widetilde{x}^2 + a_2\widetilde{y}^2$, т. е. к сумме квадратов координат с некоторыми коэффициентами, причем $a_1a_2 = a_{11}a_{22} - a_{12}^2$ (убедитесь в этом). Поэтому если $a_{11}a_{22} - a_{12}^2 = a_1a_2 > 0$, т. е. числа a_1 и a_2 имеют одинаковые знаки, то поверхность $z = f(x,y)$ в окрестности точки (0,0) лежит по одну сторону от плоскости T — в этом случае точка $(0,0)$ называется *эллиптической*. При $a_{11}a_{22} - a_{12}^2 < 0$ числа a_1 и a_2 имеют разные знаки, поэтому часть поверхности $z = f(x,y)$ лежит по одну сторону от плоскости T, а часть — по другую. Точка $(0,0)$ при этом называется *гиперболической*. Наконец, если $a_{11}a_{22} - a_{12}^2 = 0$, т. е. одно из чисел a_1 и a_2 равно нулю, то точка $(0,0)$ называется *параболической*. В этом случае поведение поверхности $z = f(x,y)$ по отношению к плоскости T в окрестности точки $(0,0)$ определяется производными функции f более высоких порядков.

По-существу, та же идея лежит в основе классификации дифференциальных уравнений в частных производных вида

$$a_{11}f_{xx} + 2a_{12}f_{xy} + a_{22}f_{yy} + F(x,y,f,f_x,f_y) = 0.$$

При $a_{11}a_{22} - a_{12}^2 > 0$ это уравнение называется *эллиптическим*, при $a_{11}a_{22} - a_{12}^2 < 0$ — *гиперболическим*, а при $a_{11}a_{22} - a_{12}^2 = 0$ — *параболическим*. Поведение решений эллиптических, гиперболических и параболических уравнений весьма различно, что проявляется, в частности, в различии тех физических процессов, которые этими уравнениями описываются. Эллиптические уравнения описывают, например, стационарное электрическое или тепловое поле, потенциальное течение жидкости. Гиперболическими уравнениями описываются разнообразные колебательные и волновые процессы, а параболическими — диффузионные процессы, например, процесс распространения тепла.

Обратим внимание и на то, что идея классификации квадратичных форм по количеству положительных и отрицательных элементов их матрицы, приведенной к диагональному виду (т. е. по индексам

инерции), лежит в основе построения псевдоевклидовых пространств, а затем и специальной теории относительности.

Обратимся теперь к нашим исследованиям, связанным с классификацией кривых и поверхностей второго порядка. Как в случае кривых, так и в случае поверхностей мы начинали с того, что с помощью поворота системы координат приводили входящую в уравнение квадратичную форму к сумме квадратов с некоторыми коэффициентами. Для кривых эта задача решается просто — получается одно уравнение для тангенса угла поворота, которое, разумеется, всегда имеет решение. Сложнее обстоит дело с поверхностями. Путем весьма нетривиальных рассуждений мы установили, что стоящая перед нами задача сводится к решению уравнения $\det(A - \lambda E) = 0$ (это уравнение появилось у нас позже и в линейной алгебре, правда, из совершенно других соображений). Спасительным для нас оказалось то, что это — кубическое уравнение относительно λ, а кубическое уравнение всегда имеет вещественный корень. Если бы мы попытались применить ту же идею к уравнению гиперповерхности второго порядка в четырехмерном пространстве, то ничего бы не получилось: уравнение $\det(A - \lambda E) = 0$ в этом случае является уравнением четвертой степени, а такое уравнение может, вообще говоря, и не иметь вещественных корней.

Между тем, представляется весьма правдоподобным, что квадратичная форма в любом пространстве может быть при помощи поворота (или, как говорят в линейной алгебре, ортогонального преобразования) приведена к сумме квадратов с некоторыми коэффициентами. В самом деле, переход от старой системы координат в n-мерном пространстве к новой осуществляется по формулам $\widetilde{e}_i = \sum\limits_{s=1}^{n} a_{is} e_s$. Поскольку столбцы матрицы A — это координаты новых базисных векторов в старом базисе и обе системы координат — декартовы, то на матрицу A накладываются следующие ограничения: ее столбцы должны быть координатами единичных (n условий) и попарно ортогональных ($n(n-1)/2$ условий) векторов. Такие матрицы в линейной алгебре называются ортогональными. Следовательно, из n^2 элементов матрицы A произвольно выбранными могут быть $n^2 - n - n(n-1)/2 = n(n-1)/2$ элементов. Но и количество коэффициентов квадратичной формы, которые мы хотим обратить в нуль, равно $n(n-1)/2$. Таким образом, количество уравнений совпадает с количеством неизвестных, что и дает основания для нашей гипотезы.

Вопрос о возможности приведения квадратичной формы к сумме квадратов (с некоторыми коэффициентами) ортогональным преобразованием является, пожалуй, центральным вопросом линейной алгебры — науки о векторах, описываемых аксиоматикой Вейля. Геометрия подобна зоопарку, имеющему несколько входов (различных систем аксиом), где, например, шансы увидеть верблюда во многом зависят от того, через какой вход мы войдем. С этой точки зрения система аксиом Вейля значительно удобнее для исследования указанного вопроса, нежели

система аксиом Гильберта. Она удобна и в плане возможных обобщений — пространство можно считать n-мерным, а под числами понимать элементы произвольного поля.

Основным объектом линейной алгебры является n-мерное линейное пространство. Это — абстрактное пространство, но, благодаря теореме об изоморфизме линейных пространств одинаковой размерности, вместо него можно рассматривать любую его конкретную интерпретацию, в частности геометрическую интерпретацию (вектор — направленный отрезок) и алгебраическую интерпретацию (вектор — набор чисел). Геометрическая интерпретация трактует вектор так, как он вводится в аналитической геометрии. Вектор здесь не связан с какой-либо системой координат. В рамках алгебраической интерпретации набор чисел, задающих вектор, конечно же зависит от выбора системы координат, но эта зависимость не произвольна — она описывается вполне определенными формулами. На первый взгляд алгебраическая интерпретация вектора кажется надуманной. Но в действительности именно она адекватно соответствует физической трактовке вектора. В самом деле, в природе нет какой-то абсолютной системы координат, все инерциальные системы отсчета равноправны. Нет и каких-либо выделенных направлений. Поэтому понятие «направление», а значит и привычная геометрическая интерпретация вектора (при всей своей наглядности), не имеют, строго говоря, физического смысла.

Поскольку единственное основное понятие линейной алгебры — это вектор, то здесь естественным образом возникают объекты, непосредственно связанные с вектором: линейный оператор, билинейная форма, получаемая из нее квадратичная форма, полилинейная форма и ряд других. Все эти объекты, включая и вектор, объединяются общим названием «тензоры» — они не зависят от выбора системы координат в том же смысле, что и вектор (вспомним геометрическую интерпретацию вектора). По этой причине тензоры широко используются в современной теоретической физике (например, в общей теории относительности).

С алгебраической точки зрения квадратичная форма задается $(n \times \times n)$-матрицей. Линейный оператор, действующий из $\mathbf{L}^n$ в $\mathbf{L}^n$, также задается $(n \times n)$-матрицей. Но квадратичная форма — это дважды ковариантный тензор, а линейный оператор — один раз ковариантный и один раз контравариантный тензор. Это означает, что при переходе к новому базису указанные матрицы преобразуются по-разному. Иными словами, если в каком-то базисе эти матрицы совпадали, то в другом базисе они, вообще говоря, совпадать не будут. Оказывается, однако, что в евклидовом пространстве при ортогональных преобразованиях — а именно эта ситуация нас и интересует с точки зрения задачи о приведении квадратичной формы к сумме квадратов — различие между ковариантностью и контравариантностью исчезает. Это дает возможность сформулировать нашу задачу на языке операторов и, благодаря этому, решить ее.

Мы кратко описали центральную линию раздела «Линейная алгебра». Подобно стволу дерева, она имеет многочисленные ответвления. В роли одного из них выступают наши исследования, связанные со структурой линейного оператора. Мы установили, что с каждым оператором, действующим из $\mathbf{L}^n$ в $\mathbf{L}^n$, связаны два инвариантных подпространства — ядро и образ, сумма размерностей которых равна n, но пространство $\mathbf{L}^n$ не является, вообще говоря, их прямой суммой. Путем весьма непростых рассуждений нам удалось описать взаимное расположение указанных подпространств в терминах базиса из непополнимых серий. Обратим особое внимание на то, что эти рассуждения не предполагали какой-либо конкретизации понятия числа — они верны для любого поля. Лишь при рассмотрении собственных векторов, собственных значений и связанного с ними характеристического уравнения для нас стало принципиально, о каком поле идет речь. Это объясняется тем, что характеристическое уравнение, будучи уравнением n-й степени, имеет корни далеко не в любом поле. В поле комплексных чисел, благодаря наличию основной теоремы алгебры, оно имеет ровно n корней (включая кратные), поэтому матрица любого оператора может быть приведена к жордановой форме.

Линейная алгебра имеет многочисленные физические приложения. Например, как мы говорили, унитарное пространство играет фундаментальную роль в квантовой механике, где физическая величина интерпретируется как самосопряженный оператор, а процесс измерения случайным образом фиксирует один из собственных векторов этого оператора. Обширны и приложения теории групп, которая используется как в самых абстрактных разделах современной теоретической физики (например, в квантовой теории поля), так и во всех без исключения разделах естествознания, в живописи, в архитектуре и даже в быту.

Приложение
КОМПЛЕКСНЫЕ ЧИСЛА. ОСНОВНАЯ ТЕОРЕМА АЛГЕБРЫ

1. Определения.

Определение 1. *Комплексными числами называются упорядоченные пары $z = (x, y)$ вещественных чисел, для которых определены операции сложения и умножения по правилам*:

I $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$;

II $(x_1, y_1)(x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)$.

Первое из вещественных чисел пары называется *вещественной частью* комплексного числа, а второе — его *мнимой частью*. Вещественная и мнимая части комплексного числа z обозначаются соответственно так: $\operatorname{Re} z$, $\operatorname{Im} z$.

Из определения 1, в частности, следует, что любое комплексных число (x, y) можно представить в следующем виде:

$$(x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0). \tag{1}$$

В связи с этим представляется важным выяснить, какими свойствами обладают числа вида $(x, 0)$. Согласно определению 1

$$(x, 0) + (y, 0) = (x + y, 0), \quad (x, 0)(y, 0) = (xy, 0).$$

Мы видим, что:

1) между числами вида $(x, 0)$ и вещественными числами x имеется взаимно однозначное соответствие $(x, 0) \leftrightarrow x$;

2) $(x, 0) + (y, 0) \leftrightarrow x + y$;

3) $(x, 0)(y, 0) \leftrightarrow xy$.

Это дает возможность отождествить числа указанного вида с вещественными числами, т. е. считать, что комплексное число $(x, 0)$ — это вещественное число x. Напомним, что подобная ситуация возникает и в теории вещественных чисел, где периодические десятичные дроби отождествляются с рациональными числами.

Итак, мы принимаем следующее соглашение: $(x, 0) \equiv x$. Теперь формула (1) принимает вид

$$(x, y) = x + iy,$$

где $i = (0, 1)$. Отметим также, что

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

а формулы из определения 1 эквивалентны формулам сложения и умножения многочленов $x_1 + iy_1$ и $x_2 + iy_2$ с учетом условия $i^2 = -1$. Из этого следует, что:

1° *для любого z имеют место равенства*: $z + 0 = z$, $z \cdot 1 = z$;

2° *для любых* z_1 *и* z_2 *имеют место равенства*: $z_1 + z_2 = z_2 + z_1$, $z_1 z_2 = z_2 z_1$;

3° *для любых* z_1, z_2 *и* z_3 *имеют место равенства*: $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$, $z_1(z_2 z_3) = (z_1 z_2) z_3$, $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$;

4° *для любого* $z = x + iy$ *существует такое число* $-z = -x - iy$, *что* $z + (-z) = 0$.

Определение 2. Число $\overline{z} = x - iy$ *называется комплексно сопряженным по отношению к числу* $z = x + iy$.

Из этого определения следует, что:

5° $\overline{z_1} + \overline{z_2} = x_1 - iy_1 + x_2 - iy_2 = (x_1 + x_2) - i(y_1 + y_2) = \overline{z_1 + z_2}$;

6° $\overline{z_1}\,\overline{z_2} = (x_1 - iy_1)(x_2 - iy_2) = (x_1 x_2 - y_1 y_2) - i(x_1 y_2 + y_1 x_2) = \overline{z_1 z_2}$;

7° $\overline{z} = z$ *тогда и только тогда, когда число* z — *вещественное*;

8° *число* $z + \overline{z} = 2x$ — *вещественное*;

9° $z\overline{z} = x^2 + y^2$ *и, следовательно, число* $z\overline{z}$ — *вещественное*, $z\overline{z} \geqslant 0$, *причем* $z\overline{z} = 0$ *тогда и только тогда, когда* $z = 0$;

10° *для любого* $z = x + iy \neq 0$ *существует такое число*

$$z^{-1} = \frac{\overline{z}}{z\overline{z}} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2},$$

что $zz^{-1} = z\dfrac{\overline{z}}{z\overline{z}} = 1$.

11°. $\overline{\overline{z}} = z$.

Замечание 1. При формулировке свойства 9° мы воспользовались символом $\geqslant$. Это оказалось возможным сделать благодаря тому, что число $z\overline{z}$ — вещественное. Применительно к комплексным числам символы $\geqslant$, $\leqslant$, $>$, $<$ не используются.

Замечание 2. Мы видим, что множество комплексных чисел образует поле. Следовательно, для любых z_1 и z_2 определена разность $z_1 - z_2$ и, при $z_2 \neq 0$, определено отношение $\dfrac{z_1}{z_2}$.

2. Геометрическая интерпретация. Комплексное число $z = x + iy$ удобно представлять себе как точку на плоскости, имеющую декартовы координаты (x, y). При этом числу 0 соответствует начало координат O, вещественным числам – ось Ox, а *мнимым числам*, т. е. числам вида iy — ось Oy.

Ясно, что отображение плоскости на себя вида $z \to z + z_0$ является параллельным переносом.

Рассмотрим точки $A(x_1, y_1)$, $B(x_2, y_2)$, соответствующие двум комплексным числам, и точку $C(x_1 + x_2, y_1 + y_2)$, соответствующую их сумме. Середина отрезка OC имеет координаты $\left(\dfrac{x_1 + x_2}{2}, \dfrac{y_1 + y_2}{2}\right)$, т. е. совпадает с серединой отрезка AB. Следовательно, четырехугольник $AOBC$ — параллелограмм с диагональю OC (если, конечно, точки A, B и O не лежат на одной прямой). Мы видим, что с геометрической точки зрения комплексные числа складываются по *правилу параллелограмма* (как векторы $\vec{OA}$ и $\vec{OB}$).

Точки на плоскости можно задавать не только декартовыми, но и, например, полярными координатами. С этой точки зрения комплексное число $z = x + iy$ можно представить в виде

$$z = \rho(\cos\varphi + i\sin\varphi).$$

Числа ρ и φ называются соответственно *модулем* и *аргументом* числа z и обозначаются так: $|z|$ и $\arg z$ (аргумент нуля не определен). Ясно, что

$$|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}, \quad 0 \leqslant \arg z < 2\pi.$$

Из правила параллелограмма и неравенства треугольника следует, что

$$|z_1| - |z_2| \leqslant |z_1 + z_2| \leqslant |z_1| + |z_2|.$$

Указанная форма записи комплексного числа удобна для описания операции умножения. В самом деле,

$$\rho_1(\cos\varphi_1 + i\sin\varphi_1)\rho_2(\cos\varphi_2 + i\sin\varphi_2) =$$
$$= \rho_1\rho_2(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)) =$$
$$= \rho_1\rho_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)).$$

Следствие. $z^n = |z|^n(\cos n\varphi + i\sin n\varphi)$ (*формула Муавра*).

Замечание. Иногда говорят, что при умножении двух комплексных чисел их модули перемножаются, а аргументы складываются. Здесь, однако, требуется уточнение. Дело в том, что сумма $\varphi_1 + \varphi_2$ может оказаться, например, больше 2π, в то время как аргумент комплексного числа меньше 2π. Поэтому правильнее писать так: $|z_1z_2| = |z_1||z_2|$, $\arg(z_1z_2) = (\arg z_1 + \arg z_2) \pmod{2\pi}$ (см. п. 3 § 3 гл. 5 ч. III).

3. Многочлен. *Многочленом* (точнее, многочленом-функцией) называется функция

$$f(z) = a_nz^n + a_{n-1}z^{n-1} + \ldots + a_0,$$

где $a_n \neq 0$, n — целое неотрицательное число, называемое *степенью* многочлена, числа a_k (*коэффициенты* многочлена) и z — комплексные. Число z_0 называется *корнем* многочлена $f(z)$, если $f(z_0) = 0$. Стоящая перед нами цель состоит в доказательстве следующего утверждения, получившего название «*Основная теорема алгебры*»:

любой многочлен ненулевой степени имеет по крайней мере один корень.

Для достижения этой цели нам потребуется доказать ряд вспомогательных теорем.

Теорема. *Для любого многочлена $f(z)$ существует такое положительное число m, что при всех z, удовлетворяющих неравенствам $0 < |z| < 1/2$, имеет место неравенство $|f(z) - a_0| < m|z|$.*

Доказательство. Если степень n многочлена $f(z)$ равна нулю, т. е. $f(z) = a_0$, то справедливость утверждения очевидна. Пусть $n \neq 0$, $a = \max\{|a_i|\} \geqslant |a_n| > 0$. Воспользуемся равенством $|z_1z_2| = |z_1||z_2|$,

неравенством $|z_1 + z_2| \leqslant |z_1| + |z_2|$ и формулой суммы членов геометрической прогрессии:

$$|f(z) - a_0| = |a_n z^n + \ldots + a_1 z| \leqslant |a_n||z|^n + \ldots + |a_1||z| \leqslant$$
$$\leqslant a|z|\frac{1-|z|^n}{1-|z|} < a|z|\frac{1}{1-|z|} < 2a|z|.$$

Таким образом, $m = 2a$. Теорема доказана.

Следствие. *Для любого многочлена $f(z)$ существует такое положительное число m, что при всех z, удовлетворяющих неравенствам $0 < |z| < 1/2$, имеет место неравенство $|f(z)| > |a_0| - m|z|$.*

В самом деле, воспользуемся доказанной теоремой и неравенством $|z_1 + z_2| \geqslant |z_1| - |z_2|$:

$$m|z| > |f(z) - a_0| = |a_0 - f(z)| = |a_0 + (-f(z)| \geqslant |a_0| - |f(z)|,$$

откуда $|f(z)| > |a_0| - m|z|$.

4. Вещественный многочлен. Рассмотрим *вещественный многочлен* $r(x) = a_n x^n + \ldots + a_0$, т. е. многочлен, в котором числа a_k и x — вещественные. Согласно теореме п. 3 при $0 < |x| < 1/2$ и $a_0 \neq 0$ получаем

$$|r(x) - a_0| = \left|\frac{r(x)}{a_0} - 1\right||a_0| < m|x|,$$

поэтому при $0 < |x| < \min\left(\frac{1}{2}, \frac{|a_0|}{2m}\right)$ имеем

$$\left|\frac{r(x)}{a_0} - 1\right| < \frac{m}{|a_0|}|x| < \frac{1}{2}$$

и, следовательно, $\frac{r(x)}{a_0} > \frac{1}{2}$. Это означает, что при достаточно малых $|x|$ знак $r(x)$ совпадает со знаком числа a_0. Воспользуемся этим наблюдением для доказательства следующей теоремы.

Теорема. *Если вещественный многочлен $r(x) = a_n x^n + \ldots + a_0$ степени $n \neq 0$ имеет в точке $x = 0$ нестрогий минимум* (т. е. при всех достаточно малых по модулю x имеет место неравенство $r(x) - r(0) \geqslant 0$), *то первый из отличных от нуля коэффициентов $a_1, \ldots, a_n$ имеет четный номер и положителен.*

Доказательство. Пусть $a_k \neq 0$, а при всех положительных $i < k$ выполняется равенство $a_i = 0$. По условию теоремы при всех достаточно малых по модулю x имеет место неравенство $r(x) - r(0) = x^k(a_n x^{n-k} + \ldots + a_k) \geqslant 0$. Поскольку при достаточно малых по модулю x знак выписанной разности совпадает со знаком числа $x^k a_k$, то число k — четное и $a_k > 0$. Теорема доказана.

5. Теорема о модуле многочлена. Ясно, что если многочлен $f(z)$ имеет корень $z = 0$, то в точке $z = 0$ вещественная функция $|f(z)|$ имеет минимум, равный 0. Оказывается, что верно и обратное утверждение.

Теорема. *Если вещественная функция $|f(z)|$, где $f(z)$ — многочлен степени $n \neq 0$, имеет нестрогий минимум в точке $z = 0$, то $z = 0$ — корень этого многочлена.*

Доказательство. Положим $z = \rho(\cos\varphi + i\sin\varphi)$, $a_k = \rho_k \times \times (\cos\varphi_k + i\sin\varphi_k)$ и воспользуемся формулой Муавра:

$$f(z) = \sum_{k=0}^{n} \rho_k \rho^k [\cos(\varphi_k + k\varphi) + i\sin(\varphi_k + k\varphi)].$$

Следовательно,

$$\left.\begin{aligned} u = \operatorname{Re} f(z) = \rho_0 \cos\varphi_0 + \sum_{k=1}^{n} [\rho_k \cos(\varphi_k + k\varphi)]\rho^k \\ v = \operatorname{Im} f(z) = \rho_0 \sin\varphi_0 + \sum_{k=1}^{n} [\rho_k \sin(\varphi_k + k\varphi)]\rho^k \end{aligned}\right\}.$$

По условию теоремы функция $|f(z)|$, а значит и функция $|f(z)|^2 = = u^2 + v^2$, имеет нестрогий минимум в точке $z = 0$. Допустим, что число $z = 0$ не является корнем многочлена $f(z)$, т. е. $\rho_0 \neq 0$.

Выражение u^2+v^2 при каждом фиксированном значении φ представляет собой многочлен от переменной ρ, имеющий нестрогий минимум в точке $\rho = 0$. Следовательно:

коэффициент при ρ, равный

$$2\rho_0\rho_1[\cos\varphi_0\cos(\varphi_1 + \varphi) + \sin\varphi_0\sin(\varphi_1 + \varphi)] = 2\rho_0\rho_1\cos(\varphi_0 - \varphi_1 - \varphi),$$

равен 0 при всех φ, что возможно только в случае $\rho_1 = 0$;

коэффициент при ρ^2, равный (с учетом равенства $\rho_1 = 0$)

$$2\rho_0\rho_2\cos(\varphi_0 - \varphi_2 - 2\varphi),$$

неотрицателен при всех φ, что возможно только в случае $\rho_2 = 0$, и т. д. Таким образом, $a_1 = \ldots = a_n = 0$, т. е. наш многочлен является многочленом нулевой степени, а это противоречит условию теоремы. Поэтому $\rho_0 = 0$ и, следовательно, число $z = 0$ — корень многочлена $f(z)$. Теорема доказана.

6. Теорема о вложенных отрезках. До сих пор все наши рассуждения относились в равной мере как к полю комплексных чисел, так и к полю комплексных чисел вида $a + ib$, где числа a и b — рациональные. Между тем, во втором поле теорема, аналогичная основной теореме алгебры, неверна даже для многочленов второй степени. В связи с этим возникает необходимость уточнить постановку задачи.

Напомним, что отличие между множеством рациональных чисел и множеством вещественных чисел состоит в том, что второе из них обладает следующим свойством полноты: *любое непустое множество вещественных чисел*, *ограниченное сверху* (*снизу*), *имеет точную верхнюю* (*нижнюю*) *грань*. В зависимости от способа введения вещественных чисел это утверждение выступает либо в роли простой

теоремы, либо в роли аксиомы. Воспользуемся им для доказательства так называемой *теоремы о вложенных отрезках*.

Теорема. *Пусть $a_n \leqslant x \leqslant b_n$ — бесконечная последовательность вложенных друг в друга отрезков, т. е. $a_{n+1} \geqslant a_n$ и $b_{n+1} \leqslant b_n$ при всех n (каждый следующий отрезок целиком содержится в предыдущем). Тогда существует точка x_0, принадлежащая одновременно всем этим отрезкам.*

Доказательство. Поскольку $a_n \leqslant b_n \leqslant b_1$ при всех n, то множество чисел a_n имеет точную верхнюю грань a. Аналогично множество чисел b_n имеет точную нижнюю грань b. Тем самым, достаточно доказать, что $a \leqslant b$ — тогда любая точка x из непустого множества $a \leqslant x \leqslant b$ будет искомой точкой x_0. Докажем это. Предположим противное, т. е. допустим, что $a > b$. Точка $\dfrac{a+b}{2}$, очевидно, представляет собой середину отрезка $[b, a]$. Поэтому (согласно определению точной грани) существуют такие натуральные числа i и j, что $a_i > \dfrac{a+b}{2}$, $b_j < \dfrac{a+b}{2}$. Полагая $k = \max\{i, j\}$, получаем: $a_k \geqslant a_i > \dfrac{a+b}{2}$, $b_k \leqslant b_j < \dfrac{a+b}{2}$, т. е. $b_k < a_k$ и, следовательно, отрезка $a_k \leqslant x \leqslant b_k$ не существует, а это противоречит условию теоремы. Поэтому $a \leqslant b$, что и требовалось доказать.

Следствие (теорема о вложенных квадратах). *Для любой последовательности вложенных друг в друга квадратов, определяемых неравенствами* $\begin{cases} a_n \leqslant x \leqslant b_n \\ c_n \leqslant y \leqslant d_n \end{cases}$ *($a_n \leqslant x \leqslant b_n$ и $c_n \leqslant y \leqslant d_n$ — две последовательности вложенных друг в друга отрезков), существует точка (x_0, y_0), принадлежащая всем этим квадратам.*

Ясно, что в роли x_0 выступает общая точка отрезков $a_n \leqslant x \leqslant b_n$, а в роли y_0 — общая точка отрезков $c_n \leqslant x \leqslant d_n$.

7. Основная теорема алгебры.

Теорема. *Любой многочлен ненулевой степени с комплексными коэффициентами имеет по крайней мере один комплексный корень.*

Доказательство. Не нарушая общности, можно считать, что $a_n = 1$ (поскольку $a_n \neq 0$, то уравнение $f(z) = 0$ можно разделить на a_n). Далее, в случае $a_0 = 0$ утверждение теоремы очевидно (корнем многочлена является число $z = 0$), поэтому будем считать, что $a_0 \neq 0$.

Поскольку $|f(z)| \geqslant 0$, то множество чисел $|f(z)|$ имеет точную нижнюю грань I. В соответствии с результатами п.5 достаточно доказать, что существует такая точка z_0, что $|f(z_0)| = I$. Докажем это.

Ясно, что $|a_0| = |f(0)| \geqslant I$. С другой стороны, согласно следствию из теоремы п. 3 существует такое $m > 0$, что при всех z, удовлетворяющих неравенству $|z| > 2$,

$$|f(z)| = |z^n||1 + a_{n-1}\left(\frac{1}{z}\right) + \ldots + a_0\left(\frac{1}{z}\right)^n| > |z^n|\left(1 - \frac{m}{|z|}\right).$$

Следовательно, при $|z| > \max\{2, 2m\}$ выполняются неравенства $|f(z)| > |z|^n/2 \geqslant |z|/2$, а при $|z| > \max\{2, 2m, 4|a_0|\}$ — неравенства $|f(z)| > 2|a_0| \geqslant |a_0| + I$. Поэтому вне круга $|z| \leqslant R = \max\{2, 2m, 4|a_0|\}$, а значит, и вне квадрата S_0 со стороной $2R$, определяемого неравенствами $|x| \leqslant R$, $|y| \leqslant R$, справедливы неравенства $\inf|f(z)| \geqslant |a_0| + I > I$. Это означает, что $\inf_{S_0}|f(z)| = I$.

Разобьем теперь квадрат S_0 на четыре равных квадрата. По крайней мере на одном из них — назовем его S_1 — точная нижняя грань чисел $|f(z)|$ равна I (в противном случае точная нижняя грань чисел $|f(z)|$ на квадрате S_0 была бы отлична от I). Разобьем квадрат S_1 на четыре равных квадрата и выберем тот из них (квадрат S_2), на котором точная нижняя грань чисел $|f(z)|$ равна I. Продолжая аналогичные рассуждения, мы построим последовательность вложенных друг в друга квадратов S_k, на каждом из которых точная нижняя грань чисел $|f(z)|$ равна I. Все эти квадраты имеют общую точку z_0. Осталось доказать, что $|f(z_0)| = I$.

Без ограничения общности будем считать, что $z_0 = 0$ (в противном случае можно сделать параллельный перенос, положив $z = \widetilde{z} + z_0$).

Для всех точек $z = x + iy$ квадрата S_k $(k > 0)$ справедливы неравенства $|x| \leqslant \dfrac{2R}{2^k}$, $|y| \leqslant \dfrac{2R}{2^k}$, поэтому $|z| \leqslant |x| + |y| \leqslant \dfrac{4R}{2^k} < \dfrac{4R}{k}$. С другой стороны, согласно следствию из теоремы п. 3, существует такое положительное число m, что при всех z, удовлетворяющих неравенствам $0 < |z| < \min\left(\dfrac{1}{2}, \dfrac{4R}{k}\right)$, имеют место неравенства

$$|f(z)| > |f(0)| - m|z| > |f(0)| - \frac{4mR}{k}.$$

Следовательно, число I, равное точной нижней грани чисел $|f(z)|$ на каждом из квадратов S_k, при каждом k удовлетворяет неравенству $I \geqslant |f(0)| - \dfrac{4mR}{k}$, что возможно только в случае $I \geqslant |f(0)|$. Но по определению точной нижней грани $I \leqslant |f(0)|$, поэтому $|f(0)| = |f(z_0)| = I$.

Таким образом, в точке z_0 функция $|f(z)|$ имеет минимум и, следовательно, $f(z_0) = 0$. Теорема доказана.

8. Следствия из основной теоремы алгебры. .

Теорема 1. *Любой многочлен $f(z)$ степени $n \neq 0$ можно представить в виде $f(z) = (z - z_0)g(z)$, где z_0 — его корень, а $g(z)$ — многочлен степени $(n-1)$.*

Доказательство. Если z_0 — корень многочлена $f(z)$, то многочлен $f(\widetilde{z} + z_0)$ имеет корень $\widetilde{z} = 0$ и, следовательно, имеет вид

$$f(\widetilde{z} + z_0) = b_n\widetilde{z}^n + b_{n-1}\widetilde{z}^{n-1} + \ldots + b_1\widetilde{z} = \widetilde{z} \cdot h(\widetilde{z}),$$

где $h(\widetilde{z})$ — многочлен степени $(n-1)$. Полагая $\widetilde{z} = z - z_0$, получаем:

$$f(z) = (z - z_0)h(z - z_0) = (z - z_0)g(z),$$

где $g(z)$ — многочлен степени $(n-1)$. Теорема доказана.

Следствие. *Любой многочлен $f(z)$ степени $n > 0$ можно представить в виде*

$$f(z) = a_n(z - z_1)\ldots(z - z_n), \tag{1}$$

где z_k — его корни.

Для доказательства этого утверждения достаточно n раз воспользоваться только что доказанной теоремой.

Замечание. Имея в виду формулу (1), часто говорят так: *любой многочлен степени n имеет ровно n корней*. Здесь, однако, требуется уточнение. Дело в том, что среди корней z_k в формуле (1) могут быть совпадающие, или, как говорят, *кратные корни*. Натуральное число α называется *кратностью корня* z_k, если этот корень фигурирует в формуле (1) ровно α раз. Тем самым, правильнее было бы сказать так:

сумма кратностей всех корней многочлена равна степени этого многочлена.

Теорема 2. *Если многочлен с вещественными коэффициентами имеет своим корнем кратности α число z_0, то и комплексно сопряженное по отношению к z_0 число $\overline{z}_0$ является корнем кратности α этого многочлена.*

Доказательство. Будем считать, что $\operatorname{Im} z_0 \neq 0$ — в противном случае утверждение теоремы тривиально. Поскольку коэффициенты данного многочлена — вещественные числа, то равенство

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 = 0$$

эквивалентно равенству

$$a_n \overline{z}^n + a_{n-1} \overline{z}^{n-1} + \ldots + a_0 = 0.$$

(см. свойства 5°, 6° и 11° п. 1). Поэтому число z_0 является корнем данного многочлена тогда и только тогда, когда $\overline{z}_0$ — корень этого многочлена. Следовательно, данный многочлен можно представить в виде

$$(z - z_0)(z - \overline{z}_0)g(z) = (z^2 - (z_0 + \overline{z}_0) + z_0\overline{z}_0)g(z).$$

Многочлен $g(z)$ степени (n-2) является частным от деления данного многочлена на многочлен второй степени с вещественными коэффициентами, поэтому его коэффициенты — вещественные числа. Из этого, как мы уже знаем, следует, что число z_0 является корнем многочлена $g(z)$ тогда и только тогда, когда $\overline{z}_0$ — его корень. Поэтому в этом случае многочлен $g(z)$ можно записать так:

$$g(z) = (z^2 - (z_0 + \overline{z}_0) + z_0\overline{z}_0)h(z).$$

Продолжая рассуждать аналогичным образом, мы за α шагов получим доказательство теоремы.

Следствие. *Любой многочлен нечетной степени с вещественными коэффициентами имеет по крайней мере один вещественный корень.*

В самом деле, если предположить, что все корни такого многочлена не являются вещественными, то сумма их кратностей окажется четной и, следовательно, неравной степени многочлена, чего не может быть.

Предметный указатель

Алгебраическое дополнение 19
Аналитическая геометрия 32
Аффинная система координат 44
Аффинные координаты вектора 44
— — точки 44

Базис 43
— галилеев 143
— канонический квадратичной формы 122
— ортонормированный 127
Беспорядок 16
Билинейная форма 116
— — положительно (отрицательно) определенная 123
— — симметричная 121

Валентность тензора 118
Вейля аксиомы 84
Вектор 36, 84, 86
— единичный 37
— нулевой 36
—, противоположный данному вектору 37
—, разложенный по векторам 43
Векторное произведение двух векторов 52
Векторы коллинеарные 42
— координатные 39
— компланарные 43
— ортогональные 39, 126
— равные 37

Гипербола 64
Гиперболоид двуполостный 79
— — вращения 80
— однополостный 78
— — вращения 80
Гиперповерхность 135
— второго порядка 135
— — — мнимая 137
Грама определитель 54
Группа 145
— коммутативная (абелева) 145
Группа ортогональная ($\mathrm{O}(n)$) 147
— полная линейная ($\mathrm{GL}(n)$) 147
— псевдоортогональная ($\mathrm{O}(p,q)$) 147
— унимодулярная ($\mathrm{SL}(n)$) 147
— унитарная ($\mathrm{U}(n)$) 147

Двойное векторное произведение 54
Декартова система координат 34
Делители нуля 148
Директриса 66, 67
Дифференциальная геометрия 32
Длина вектора 37, 126
— непополнимой серии 101

Жорданова клетка 110
— форма матрицы 110

Закон инерции квадратичных форм 123

Изоморфизм 90
Инвариант уравнения гиперповерхности второго порядка 137
Инвариантное подпространство оператора 99
Индекс инерции квадратичной формы 123

Канонический вид квадратичной формы 123
Касательная 68
Квадратичная форма 121
— —, положительно (отрицательно) определенная 123
Конические сечения 78
Конус 76
— второго порядка 77
— световой 143
Конусы 137
Координата точки на оси 33
— — по оси 34
Координаты биполярные 35
— вектора 37, 89

Координаты криволинейные 35
— полярные 35
— сферические 36
— цилиндрические 36
— эллиптические 35, 71
Коши–Буняковского неравенство 125, 139
Коэффициенты линейной системы 27
— — формы 116
Крамера формулы 29
Кривая второго порядка 71
Кронекера символ 115
Кронекера–Капелли теорема 28

Лагранжа метод 122
Линейная алгебра 85
— комбинация 10
— — нетривиальная 10
— — тривиальная 10
— оболочка 88
— форма 116
Линейное дополнение 92
Линейный оператор 94
— — нормальный 140
— —, обратный по отношению к данному 98
— — ортогональный 130
— — самосопряженный 133, 142
— —, сопряженный по отношению к данному 129, 140
— — тождественный 98
— — унитарный 141
— — эрмитов 142
Линия мировая 143
Лоренца группа 147
— преобразование 143

Матрица 8
— билинейной формы 116
— диагональная 108
— единичная 24
— квадратичной формы 121
— квадратная 11
— линейного оператора 95
— обратная 24
— ортогональная 131
— — собственная (несобственная) 131
Матрица основная 27
— перехода к новому базису 115
— расширенная 27
— транспонированная 20
— унитарная 141
Матричная запись системы 28
Минор 24
— базисный 24
—, дополнительный к элементу 14
— угловой 123
Модуль вектора 37

Направляющая конуса 76
— цилиндра 76
Начало аффинной системы координат 44
— координат 33
Независимость системы аксиом 91
Непротиворечивость системы аксиом 91
Неравенство треугольника 126
Норма 126

Образ оператора 96
Образующая конуса 76
— цилиндра 76
Определитель 11
— произведения матриц 22
Определителя основные свойства 14
Оптическое свойство гиперболы 70
— — параболы 71
— — эллипса 70
Ортогонализация 128
Ортогональное дополнение 128
Ось аффинной системы координат 44
— координат 33
Откладывание вектора от точки 38, 85
Отклонение точки от плоскости 59
— — — прямой 57
Отношение 148

Пара векторов левая 46
— — правая 46

Парабола 67
Параболоид вращения 81
— гиперболический 79
— эллиптический 79
Параболоиды 137
Пересечение двух линейных подпространств 91
Перестановка 16
Пифагора теорема 127
Плоскость ориентированная 45
Поверхность второго порядка 73
Поворот 47
— гиперболический 144
Подгруппа 146
Подпространство линейное 87
Полнота системы аксиом 91
Положительная полуось 33
Полуоси гиперболы 65
— эллипса 63
Поле 147
— сравнений по модулю p 149
— числовое 148
Полилинейная форма 117
Последний вектор серии 101
Правило треугольника 38, 85
Преобразование базиса 115
— — несобственное 47, 49
— — ортогональное 132
— — собственное 46, 49
Присоединенный вектор 112
Проекция вектора на подпространство 127
— точки на прямую 34
Произведение вектора на число 38, 84
— двух смешанных произведений 53
— линейного оператора на число 94
— матриц 21
— матрицы на число 9
— операторов 97
Пространства изоморфные 90
Пространство арифметическое 9
— евклидово 125
— — комплексное 139
— линейное 85
Пространство линейное n-мерное 88
— Минковского 143
— нормированное 126
— ориентированное 49
— псевдоевклидово 142
— событий 143
— унитарное 139
Прямая сумма двух линейных подпространств 92
Прямое произведение тензоров 119

Разложение определителя по столбцу 20
— — — строке 14
Размерность 88
Разность 148
— векторов 87
— матриц 9
Ранг билинейной формы 120
— матрицы 26
— оператора 96
— тензора 118
Решение системы 27
— — нетривиальное 28
— — тривиальное 28

Свертка тензора 119
Секущая 68
Серии непополнимые различные 102
Серия ненулевых векторов 101
— — — непополнимая 101
Сигнатура псевдоевклидова пространства 142
Сильвестра критерий 124
Система координат 135
— — исходная 45, 49
— — левая 46, 49
— — правая 46, 49
— неоднородная 27
— однородная 27
— —, соответствующая данной системе 27
Скалярное произведение двух векторных произведений 54
— — векторов 39, 85, 125, 142
След оператора 120

Смешанное произведение трех векторов 53
Собственный вектор 107
Собственное значение 107
Спектр оператора 107
— — «сдвинутый» 107
Старший вектор серии 101
Столбцы базисные 24
Строки 9
— базисные 24
— координатные 10
— линейно зависимые 10
— — независимые 10
Сужение оператора на инвариантное подпространство 99
Сумма векторов 38, 84
— линейных операторов 94
— матриц 8
— тензоров 118

Тензор 118
— метрический 125
— симметричный типа (2,0) 121
— типа (p, q) 118
Точка 84
Тройка векторов левая 49
— — правая 49

Угол между векторами 39, 126
Уравнение гиперболы каноническое 65
— гиперповерхности второго порядка каноническое 136
— параболы каноническое 67
— плоскости в отрезках 58
— — нормированное 59
— — общее 58
— —, проходящей через данную точку и перпендикулярной к данному вектору 58
— —, проходящей через три данные точки 58
Уравнение прямой в отрезках 55
— — каноническое 56
— — нормированное 57
— — общее 56
— —, проходящей через данную точку и параллельной данному вектору 56, 60
— —, — — — — — перпендикулярной к данному вектору 55
— —, проходящей через две данные точки 55, 60
— эллипса каноническое 63
Уравнения прямой канонические 60
— — параметрические 56, 60

Фокус гиперболы 64
— параболы 67
— эллипса 62
Формула разложения определителя по строке 14
Фредгольма альтернатива 129, 140
Фундаментальная совокупность решений 90
Функциональный анализ 90

Характеристическое уравнение 108

Циклическая перестановка 53
Цилиндр 76
Цилиндры 136

Эйлера углы 51
Эксцентриситет 66
Элементарная геометрия 32
Эллипс 62
Эллипсоид 78
— вращения 80
Эллипсоиды и гиперболоиды 137

Ядро оператора 96