ВМ
И С

Л. И. ГОЛОВИНА

ЛИНЕЙНАЯ АЛГЕБРА И НЕКОТОРЫЕ ЕЕ ПРИЛОЖЕНИЯ

ИЗБРАННЫЕ ГЛАВЫ ВЫСШЕЙ МАТЕМАТИКИ ДЛЯ ИНЖЕНЕРОВ И СТУДЕНТОВ ВТУЗОВ

Л. И. ГОЛОВИНА

ЛИНЕЙНАЯ АЛГЕБРА И НЕКОТОРЫЕ ЕЕ ПРИЛОЖЕНИЯ

ИЗДАНИЕ ЧЕТВЕРТОЕ, ИСПРАВЛЕННОЕ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов высших технических учебных заведений

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1985

22.143
Г 81
УДК 512.86

Головина Л. И. **Линейная алгебра и некоторые ее приложения:** Учебное пособие для вузов.— 4-е изд., испр.— М.: Наука, Главная редакция физико-математической литературы, 1985.— 392 с.

Основное содержание книги составляют теория определителей и краткий курс собственно линейной алгебры. В качестве «приложений» линейной алгебры рассматриваются самые разные вопросы: дается краткое изложение общей теории кривых и поверхностей второго порядка, вводятся основные понятия тензорной алгебры, излагаются основные понятия теории групп и элементы теории представлений групп. В одной из глав книги методы линейной алгебры применяются к основным понятиям физики — принципам относительности, классическому и релятивистскому.

Ил. 39. Библиогр. 34 назв.

Рецензенты:

член-корреспондент АН СССР профессор *А. И. Кострикин*
кандидат физико-математических наук доцент *Д. В. Беклемишев*

Лидия Ивановна Головина

ЛИНЕЙНАЯ АЛГЕБРА И НЕКОТОРЫЕ ЕЕ ПРИЛОЖЕНИЯ

Редакторы *И. М. Яглом, И. Е. Морозова*
Техн. редактор *С. Я. Шкляр*
Корректоры *Е. В. Сидоркина, В. П. Сорокина*

ИБ № 12629

Печать с матриц. Подписано к печати 07.08.85. Формат 84×108 1/32. Бумага тип. № 3. Гарнитура литературная. Печать высокая. Усл. печ. л. 20,58. Усл. кр.-отт. 20,79. Уч.-изд. л. 19,53. Тираж 18000 экз. Заказ № 514. Цена 85 коп.

Ордена Трудового Красного Знамени издательство «Наука»
Главная редакция физико-математической литературы
117071 Москва В-71, Ленинский проспект, 15

4-я типография издательства «Наука»
630077 г. Новосибирск 77, Станиславского, 25

Г $\frac{1702030000-132}{053(02)-85}$ 72-85

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Эта книга представляет собой учебное пособие по линейной алгебре, рассчитанное на студентов втузов и естественно-научных факультетов университетов. Она может быть полезной и читателю, желающему самостоятельно познакомиться с основными понятиями линейной алгебры.

Глава I является вводной; она содержит необходимые для дальнейшего сведения из теории определителей и систем линейных уравнений. Основными в книге являются главы II—VI, в которых излагается собственно курс линейной алгебры. Остальные главы, по существу, не относятся к линейной алгебре, но их результаты основаны на предыдущем материале (...«некоторые ее приложения»); эти главы могут читаться и не подряд (см. ниже схему зависимости глав).

Глава VII посвящена общей теории кривых и поверхностей второго порядка; она имеет целью дополнить и углубить соответствующую часть курса аналитической геометрии, не претендуя на ее замену.

Глава VIII, посвященная общим понятиям тензорной алгебры, является довольно конспективной и может служить введением в более обстоятельные изложения той же темы, из числа которых назовем, например, указанные в списке литературы книги [15] и [16].

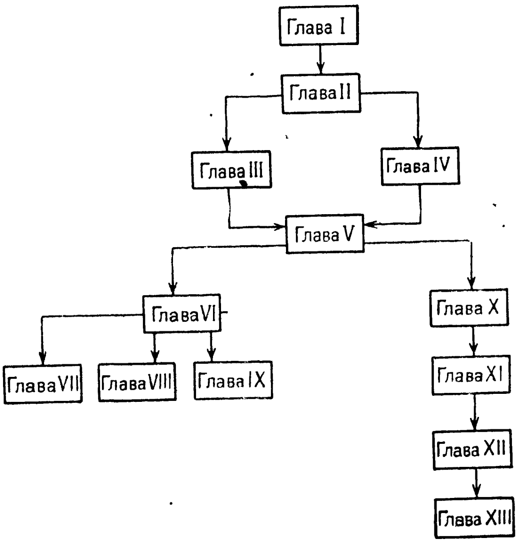
Несколько необычной для учебника линейной алгебры является глава IX, посвященная специальной теории относительности. При изучении линейной алгебры эта глава может быть и опущена, но опыт преподавания показывает, что обычно она вызывает у слушателей большой интерес.

Главы X—XI содержат самые общие сведения из теории групп и описание групп симметрии геометрических фигур и тел. Глава XII—XIII краткое, но достаточно строгое, изложение основных понятий теории представлений групп и теории характеров. Конечно, все эти вопросы уже не относятся к линейной алгебре, но методы теории групп и, в частности, основанная на линейной алгебре теория представлений групп, играют все большую роль в современной физике и химии, так что учебное пособие по линейной алгебре для втузов не может не затронуть и этих разделов.

В настоящее издание книги внесены некоторые добавления, в частности,— параграф о жордановой форме матрицы. От читателя не требуется почти никаких предварительных сведений из высшей математики, предполагается лишь, что он знаком с элементами аналитической геометрии. Используемые здесь понятия математического анализа (производная, интеграл) встречаются только в примерах, которые при чтении книги могут быть пропущены.

Содержание настоящей книги составляет несколько расширенный курс лекций, неоднократно читавшийся автором на отделении физхимии химического факультета МГУ.

Схема зависимости глав



Москва, Август, 1978 г. *Л. И. Головина*

ГЛАВА I

ОПРЕДЕЛИТЕЛИ И СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

В этой главе содержится вспомогательный материал, относящийся к решению систем линейных уравнений (т. е. уравнений первой степени). Для исследования таких систем вводится важное понятие определителя. Результаты этой главы, интересные и сами по себе, и в приложениях к аналитической геометрии, необходимы для понимания дальнейших глав книги.

§ 1. Системы уравнений с двумя и тремя неизвестными

При решении *одного уравнения первой степени с одним неизвестным*

$$ax = b$$

возможны три случая:

1. Если $a \neq 0$, уравнение имеет единственное решение $x = \frac{b}{a}$.

2. Если $a = 0$ и $b = 0$, уравнение имеет бесчисленное множество решений: любое число x удовлетворяет уравнению $ax = b$ (так как $0 \cdot x = 0$) и, значит, является его решением.

3. Если $a = 0$, но $b \neq 0$, уравнение не имеет решений, так как при подстановке вместо x любого числа в левой части получается нуль, в то время как правая часть отлична от нуля.

Из дальнейшего будет видно, что аналогичные три случая имеют место и при решении произвольной системы линейных уравнений.

Рассмотрим *систему двух уравнений с двумя неизвестными:*

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2. \end{cases} \qquad (1)$$

Решением такой системы называется каждая пара значений $x=\alpha$, $y=\beta$, подстановка которых вместо x и y обращает оба уравнения в тождества. Чтобы решить эту систему, умножим первое уравнение на b_2, второе — на $-b_1$ и сложим их; мы получим

$$x(a_1b_2-a_2b_1)=c_1b_2-c_2b_1.$$

Отсюда, если $a_1b_2-a_2b_1\neq 0$, будем иметь

$$x=\frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1}. \tag{2}$$

Аналогично находим, что

$$y=\frac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}. \tag{3}$$

Таким образом, в случае, когда $a_1b_2-a_2b_1\neq 0$, система (1) имеет единственное решение.

Выражения, стоящие в числителях и знаменателях правых частей равенств (2) и (3), устроены одинаково. А именно, рассмотрим квадратную таблицу чисел

$$A=\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}.$$

Такие таблицы называются *матрицами*. Горизонтальные ряды образующих матрицу чисел называются ее *строками*, вертикальные — *столбцами*. Числа a_1, b_1, a_2, b_2, составляющие матрицу, называются ее *элементами*. В нашем примере мы имеем квадратную матрицу второго порядка. Диагональ, идущая из левого верхнего угла матрицы в правый нижний, называется ее *главной диагональю*. Знаменатели дробей, стоящих в правых частях равенств (2) и (3), устроены следующим образом: из произведения элементов, стоящих по главной диагонали матрицы A, вычитается произведение элементов, стоящих во второй, или *побочной*, ее диагонали:

$$a_1b_2-a_2b_1.$$

Полученное выражение называется *определителем* матрицы A (определителем второго порядка) и обозначается так:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Таким образом, по определению,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1.$$

В этих обозначениях числитель дроби, стоящей в правой части равенства (2), представляет собой определитель

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} = c_1b_2 - c_2b_1,$$

получающийся из знаменателя заменой первого столбца столбцом свободных членов, а числитель дроби, стоящей в правой части равенства (3), — определитель

$$\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} = a_1c_2 - a_2c_1,$$

получающийся из знаменателя заменой второго столбца столбцом свободных членов уравнений системы (1).

Итак, мы нашли, что если $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$, то

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}.$$

Это — формулы Крамера для решения системы двух уравнений с двумя неизвестными.

Пример. *Пользуясь формулами Крамера, решить систему уравнений*

$$\begin{cases} 2x + 5y = 8, \\ 3x + y = -1. \end{cases}$$

Решение.

$$x = \frac{\begin{vmatrix} 8 & 5 \\ -1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 5 \\ 3 & 1 \end{vmatrix}} = \frac{8+5}{2-15} = -1, \quad y = \frac{\begin{vmatrix} 2 & 8 \\ 3 & -1 \end{vmatrix}}{-13} = \frac{-2-24}{-13} = 2.$$

Рассмотрим теперь случай, когда

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1 = 0. \tag{4}$$

Равенство (4) можно переписать так*):

$$\frac{a_1}{a_2}=\frac{b_1}{b_2},$$

т. е. в этом случае коэффициенты при неизвестных пропорциональны. Если, кроме того, и

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}=0, \quad \text{т. е.} \quad \frac{c_1}{c_2}=\frac{b_1}{b_2},$$

то и свободные члены пропорциональны коэффициентам при неизвестных, и мы имеем на самом деле *одно уравнение с двумя неизвестными* — оно допускает бесчисленное множество решений.

Наконец, если

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=0, \quad \text{но} \quad \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}\neq 0,$$

т. е. если

$$\frac{a_1}{a_2}=\frac{b_1}{b_2}\neq\frac{c_1}{c_2},$$

то уравнения, очевидно, противоречат друг другу и система не имеет ни одного решения.

Рассмотрим теперь систему *трех линейных уравнений с тремя неизвестными*:

$$\begin{cases} a_1x+b_1y+c_1z=d_1, \\ a_2x+b_2y+c_2z=d_2, \\ a_3x+b_3y+c_3z=d_3. \end{cases} \tag{5}$$

Решением этой системы называется каждая такая тройка чисел $x=\alpha$, $y=\beta$, $z=\gamma$, при подстановке которых все три уравнения обращаются в тождества. Умножив первое уравнение на $\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}=b_2c_3-b_3c_2$, второе — на $-\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}=b_3c_1-b_1c_3$, третье — на $\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}=b_1c_2-b_2c_1$

*) Здесь и далее мы считаем, что знаменатели отличны от нуля; случай, когда это не так, рассмотрите сами.

и сложив их все, мы получим

$$x(a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1)=$$
$$=d_1b_2c_3-d_1b_3c_2+d_2b_3c_1-d_2b_1c_3+d_3b_1c_2-d_3b_2c_1$$

(коэффициенты при y и z, как легко видеть, будут равны нулю). Отсюда, если коэффициент при x отличен от нуля, получаем

$$x=\frac{d_1b_2c_3-d_1b_3c_2+d_2b_3c_1-d_2b_1c_3+d_3b_1c_2-d_3b_2c_1}{a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1}. \tag{6}$$

Посмотрим, как устроено выражение, стоящее в знаменателе правой части равенства (6). Для этого рассмотрим квадратную таблицу (*матрицу третьего порядка*)

$$A=\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}.$$

Будем снова называть *главной диагональю* диагональ, идущую из левого верхнего угла этой матрицы в правый нижний, и *побочной* — диагональ, идущую из левого нижнего угла в правый верхний.

Знаменатель в формуле (6) представляет собой алгебраическую сумму шести членов, каждый из которых

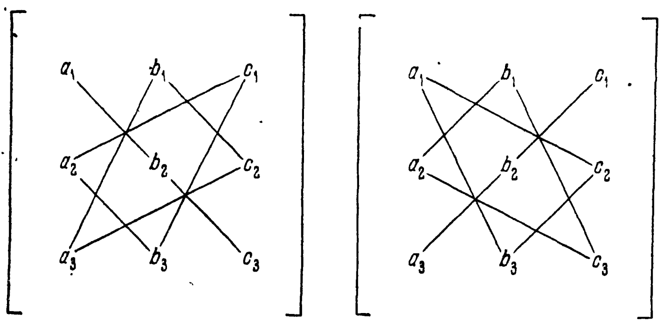


Рис. 1.

является произведением трех элементов, взятых по одному из каждой строки и каждого столбца матрицы A, причем знак **плюс** имеет произведение элементов, принадлежащих главной диагонали, и два произведения элементов, образующих в матрице (равнобедренные) треугольники с основаниями, параллельными главной диагонали (рис. 1, *а*), а знак минус имеет произведение элементов, принадлежащих побочной диагонали, и два произведения элементов, образующих треугольники с основаниями, параллельными побочной диагонали (рис. 1, *б*).

Такое выражение называется *определителем*, составленным из матрицы A (определителем третьего порядка), и обозначается так:

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

Таким образом, по определению,

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

$$= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2.$$

Выражение, стоящее в числителе правой части формулы (6), получается из знаменателя, если каждую букву a заменить буквой d с тем же номером, т. е.

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}} = \frac{D_1}{D}. \qquad (6a)$$

Аналогично можно показать, что при $D \neq 0$ из системы (5) следуют равенства

$$y = \frac{D_2}{D}, \quad z = \frac{D_3}{D}, \qquad (7)$$

где D_i, $i = 1, 2, 3$ — определитель, получающийся из определителя D заменой i-го столбца столбцом свободных членов. Это — формулы Крамера для системы трех уравнений с тремя неизвестными.

Пример. *Решить по формулам Крамера систему уравнений*

$$\begin{cases} x + 2y + 3z = 7, \\ x - 3y + 2z = 5, \\ x + y + z = 3. \end{cases}$$

Решение.

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 1 & -3 & 2 \\ 1 & 1 & 1 \end{vmatrix} = -3 + 3 + 4 + 9 - 2 - 2 = 9 \neq 0;$$

$$D_1 = \begin{vmatrix} 7 & 2 & 3 \\ 5 & -3 & 2 \\ 3 & 1 & 1 \end{vmatrix} = -21 + 15 + 12 + 27 - 10 - 14 = 9,$$

$$D_2 = \begin{vmatrix} 1 & 7 & 3 \\ 1 & 5 & 2 \\ 1 & 3 & 1 \end{vmatrix} = 5 + 14 + 9 - 15 - 7 - 6 = 0,$$

$$D_3 = \begin{vmatrix} 1 & 2 & 7 \\ 1 & -3 & 5 \\ 1 & 1 & 3 \end{vmatrix} = -9 + 7 + 10 + 21 - 6 - 5 = 18.$$

Следовательно,

$$x = \frac{D_1}{D} = 1, \quad y = \frac{D_2}{D} = 0, \quad z = \frac{D_3}{D} = 2.$$

Для того чтобы понять, что такое *определитель n-го порядка*, рассмотрим снова определители второго и третьего порядков:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 \tag{8}$$

и

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

$$= a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1 - a_2 b_1 c_3 - a_1 b_3 c_2. \quad (9$$

Мы видим, что определитель есть алгебраическая сумма всевозможных произведений его элементов, взятых по одному из каждой строки и каждого столбца.

Каждое такое произведение называется *членом определителя*. В каждом члене определителя второго порядка расположим множители в порядке следования столбцов:

$$a_1b_2 - a_2b_1$$

и рассмотрим соответствующие расположения (*перестановки*) нижних индексов (указывающих номера строк):

$$1,\ 2 \text{ и } 2,\ 1.$$

В первом произведении эти индексы расположены по возрастанию, и соответствующее произведение входит в определитель со знаком плюс; во втором они, как говорят, образуют беспорядок, или *инверсию*, 2, 1, и соответствующий член входит в определитель со знаком минус.

В определителе третьего порядка шесть членов. Если в каждом из них расположить множители в порядке следования столбцов, то в членах, входящих со знаком *плюс*, нижние индексы образуют перестановки

$$1,\ 2,\ 3;\ 2,\ 3,\ 1 \text{ и } 3,\ 1,\ 2.$$

Рассмотрим три пары индексов 1, 2; 1, 3 и 2, 3 из первой перестановки 1, 2, 3; числа каждой пары расположены по возрастанию — в этой перестановке нуль инверсий. Во второй перестановке 2, 3, 1 три пары индексов: 2, 3; 2, 1 и 3, 1, две из которых — 2, 1 и 3, 1, образуют инверсии. В третьей перестановке 3, 1, 2 — три пары индексов 3, 1; 1, 2 и 3, 2, из которых две — 3, 1 и 3, 2, образуют инверсии.

Произведениям, входящим со знаком *минус*, соответствуют три перестановки индексов

$$3,\ 2,\ 1;\ 2,\ 1,\ 3 \text{ и } 1,\ 3,\ 2,$$

причем в первой, как нетрудно видеть, три инверсии: 3, 2; 3, 1 и 2, 1, а во второй и третьей — по одной; соответственно 2, 1 и 3, 2. Таким образом, со знаком плюс входят те члены, у которых в перестановке индексов четное число инверсий, а со знаком минус — те, у которых это число нечетно.

Для дальнейшего нам будет удобно ввести для определителей второго и третьего порядков новые обозначения:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

где все элементы определителя обозначены одной и той же буквой a с двумя индексами, первый из которых указывает номер строки, в которой стоит этот элемент, а второй — номер соответствующего столбца. (Элементы, например, первого определителя читаются так: a один один, a один два, a два один, a два два.) Тогда

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

и

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum \pm a_{i_1 1} a_{i_2 2} a_{i_3 3},$$

где знак плюс стоит перед теми произведениями, в которых перестановка i_1, i_2, i_3 четная (т. е. имеет четное число инверсий), и знак минус — перед теми, где она нечетна. Это можно записать еще и так:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum (-1)^{\alpha} a_{i_1 1} a_{i_2 2} a_{i_3 3},$$

где α есть число инверсий в перестановке первых индексов, i_1, i_2, i_3 (вторые индексы расположены в порядке возрастания), а суммирование распространяется на все шесть перестановок i_1, i_2, i_3 из трех чисел 1, 2, 3.

§ 2. Перестановки и транспозиции. Определитель n-го порядка

Пусть даны n элементов $a_1, a_2, \ldots, a_n$ (например, это могут быть числа $1, 2, 3, \ldots, n$). Как известно, всевозможные расположения этих элементов называются *перестановками* из n элементов. Всего из n элементов можно составить $n!$ перестановок (докажите это).

Если какая-нибудь пара a_i, a_k элементов перестановки расположена в ней так, что элемент с бо́льшим номером стоит раньше элемента с меньшим номером, то говорят, что эти элементы образуют *инверсию*. Пусть нам надо сосчитать число инверсий в какой-то перестановке, образованной числами 1, 2, 3, ..., n (это могут быть номера элементов $a_1, a_2, \ldots, a_n$). Сделать это можно следующим образом. Сосчитаем сначала число элементов, стоящих впереди единицы — все эти элементы и только они образуют инверсии с единицей. Вычеркнем затем единицу и сосчитаем число элементов, стоящих впереди двойки — это будут все те элементы, которые образуют инверсии с двойкой (не считая уже вычеркнутой единицы, которая тоже может образовывать инверсию с двойкой, но в таком случае эту инверсию мы уже учли раньше). Затем вычеркнем двойку и сосчитаем число элементов, стоящих впереди тройки, и т. д. Все полученные числа сложим — эта сумма и будет равна общему числу инверсий. Число инверсий в перестановке $i_1, i_2, \ldots, i_n$ обозначается так: $[i_1, i_2, \ldots, i_n]$. Например,

$$[2, 5, 1, 4, 7, 3, 6] = 2 + 0 + 3 + 1 + 0 + 1 = 7.$$

Перестановки с четным числом инверсий называются *четными*, перестановки с нечетным числом инверсий — *нечетными* перестановками.

Пусть дана перестановка из n элементов $a_1, a_2, \ldots$ $\ldots, a_i, \ldots, a_k, \ldots, a_n$. Поменяем местами два ее элемента a_i и a_k; при этом мы получим перестановку $a_1, a_2, \ldots$ $\ldots, a_k, \ldots, a_i, \ldots, a_n$. Такая операция перемещения двух элементов перестановки называется *транспозицией*.

Теорема 1. *От одной транспозиции четность перестановки меняется* (т. е. четная перестановка становится нечетной, а нечетная — четной).

Доказательство. Рассмотрим сначала случай, когда меняются местами два соседних элемента α и β перестановки

$$a_1, a_2, \ldots, a_l, \alpha, \beta, b_1, b_2, \ldots, b_m. \tag{10}$$

После транспозиции элементов α и β получим перестановку

$$a_1, a_2, \ldots, a_l, \beta, \alpha, b_1, b_2, \ldots, b_m. \tag{11}$$

Так как перестановки (10) и (11) отличаются друг от друга только взаимным расположением элементов α и β (а взаимное расположение каждого из этих элементов и какого-то другого, так же как и взаимное расположение любых двух из остальных элементов, остались прежними), то число инверсий в перестановке (11) на единицу больше или на единицу меньше числа инверсий в перестановке (10), и значит, одна из этих перестановок четная, а другая — нечетная.

Рассмотрим теперь общий случай. Пусть меняются местами элементы α и β перестановки

$$a_1, \ldots, a_l, \alpha, c_1, \ldots, c_k, \beta, b_1, \ldots, b_m,$$

между которыми стоят еще k элементов $c_1, c_2, \ldots, c_k$. Мы можем выполнить транспозицию элементов α и β посредством нескольких транспозиций рядом стоящих элементов: поменяем местами α сначала с c_1, затем с c_2, и т. д., наконец, с c_k (при этом мы сделаем k транспозиций рядом стоящих элементов); затем поменяем местами α и β (еще одна транспозиция) и, наконец, поменяем местами β последовательно с c_k, с c_{k-1}, и т. д. до c_1 (еще k транспозиций рядом стоящих элементов). В конечном счете β станет на место α (и наоборот). При каждой такой транспозиции четность перестановки, как мы уже видели, меняется. А так как она изменится $2k+1$, т. е. нечетное число раз, то окончательно нечетная перестановка сделается четной, а четная — нечетной, что и требовалось доказать.

Следствие. *Число нечетных перестановок из n элементов равно числу четных перестановок* (и равно, следовательно, $n!/2$).

Доказательство. Пусть из $n!$ перестановок из n элементов p перестановок четны и q нечетны. Сделаем в каждой четной перестановке *одну и ту же* транспозицию, например, поменяем местами первые два элемента. Тогда каждая четная перестановка превратится в нечетную, причем ясно, что все p полученных при этом нечетных перестановок будут разными. А так как общее число нечетных перестановок из n элементов, по предположению, равно q, то $p \leqslant q$. Точно так же можно убедиться в том, что, наоборот, $q \leqslant p$. Следовательно, $p = q$.

Дадим теперь общее определение определителя n-го порядка. Пусть имеется квадратная таблица, состоящая из n строк и n столбцов (*матрица n-го порядка*):

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

Числа a_{ik} называются ее *элементами*, горизонтальные ряды элементов матрицы называются ее *строками*, вертикальными — *столбцами*. ***Определителем***, составленным из этой матрицы (*определителем n-го порядка*), *называется алгебраическая сумма всевозможных произведений элементов, взятых по одному из каждого столбца и каждой строки матрицы A. Если в каждом таком произведении* (члене определителя) *множители расположены в порядке следования столбцов, то со знаком* плюс *берутся те произведения, у которых перестановка первых индексов* четная, *а со знаком* минус — *те, у которых она* нечетная. Короче:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots a_{i_n n},$$

где суммирование распространяется на всевозможные перестановки $i_1, i_2, \dots, i_n$ из n чисел $1, 2, 3, \dots, n$. Так как число перестановок из n элементов равно $n!$, то определитель n-го порядка состоит из $n!$ членов. Ввиду следствия из теоремы 1 ровно половина из них, т. е. $n!/2$, входит в определитель со знаком плюс и столько же — со знаком минус.

§ 3. Свойства определителей

С увеличением порядка определителя число его членов очень быстро растет. Так, определитель четвертого порядка состоит из 24 членов, определитель пятого порядка — из 120, определитель шестого порядка — из 720 членов, и т. д. Поэтому вычислить определитель порядка выше трех, пользуясь только его определением, практически невозможно. Для того чтобы вычислять такие определители, нам придется изучить их свойства.

Прежде всего мы докажем одно вспомогательное предложение.

Лемма (о знаке члена определителя). *Произведение* $a_{i_1k_1}a_{i_2k_2}\ldots a_{i_nk_n}$ *входит в определитель n-го порядка со знаком, определяемым выражением*

$$(-1)^{[i_1,i_2,\ldots,i_n]+[k_1,k_2,\ldots,k_n]};$$

мы будем говорить в таком случае короче: *входит со знаком*

$$(-1)^{[i_1,i_2,\ldots,i_n]+[k_1,k_2,\ldots,k_n]}.$$

Доказательство. Заметим прежде всего, что если поменять местами два множителя произведения $a_{i_1k_1}a_{i_2k_2}\ldots a_{i_nk_n}$, то как в первых, так и во вторых его индексах произойдет по одной транспозиции, и значит, четность каждого из чисел

$$[i_1, i_2, \ldots, i_n] \text{ и } [k_1, k_2, \ldots, k_n]$$

изменится, а четность их суммы останется прежней.

Пусть нам дано произведение $a_{i_1k_1}a_{i_2k_2}\ldots a_{i_nk_n}$. С помощью нескольких транспозиций этих множителей расположим их так, чтобы вторые индексы шли в порядке возрастания. Для этого сначала сделаем транспозицию, при которой на первое место станет элемент из первого столбца, затем такую, чтобы на второе место попал элемент из второго столбца, и т. д. (Так, например, произведение $a_{45}a_{14}a_{52}a_{21}a_{33}$ последовательно преобразуется в $a_{21}a_{14}a_{52}a_{45}a_{33}$, затем в $a_{21}a_{52}a_{14}a_{45}a_{33}$, в $a_{21}a_{52}a_{33}a_{45}a_{14}$, и, наконец, в $a_{21}a_{52}a_{33}a_{14}a_{45}$.) Если в конечном счете, когда вторые индексы расположатся по возрастанию, первые образуют перестановку $[m_1, m_2, \ldots, m_n]$, то рассматриваемый член, по определению, входит в определитель со знаком $(-1)^{[m_1,m_2,\ldots,m_n]}$. Но так как четность суммы $[i_1, i_2, \ldots, i_n]+[k_1, k_2, \ldots, k_n]$ числа инверсий в первых и числа инверсий во вторых индексах ри транспозициях множителей не менялась, то четность той суммы в первоначальном расположении множитеей совпадает с четностью числа $[m_1, m_2, \ldots, m_n]$ — исла инверсий в перестановке первых индексов окончательного расположения: в нем вторые индексы образуют нуль инверсий. Следовательно,

$$(-1)^{[m_1, m_2, \ldots, m_n]} = (-1)^{[i_1, i_2, \ldots, i_n]+[k_1, k_2, \ldots, k_n]},$$

что и доказывает наше утверждение.

Пример. *Найти, с каким знаком произведение* $a_{32}a_{43}a_{51}a_{15}a_{24}$ *входит в определитель пятого порядка:*

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}.$$

Решение.

$$[3, 4, 5, 1, 2] = 3 + 3 = 6,$$
$$[2, 3, 1, 5, 4] = 2 + 1 = 3;$$
$$(-1)^{6+3} = (-1)^9 = -1.$$

Рассматриваемое произведение входит в определитель D со знаком *минус.*

Свойство 1 («равноправие» строк и столбцов определителя). *При транспонировании, т. е. при замене каждой строки определителя столбцом с тем же номером, определитель не меняется.*

Доказательство. Рассмотрим определитель

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

и транспонированный определитель

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

строками которого служат столбцы определителя D. Надо показать, что $D' = D$.

Каждый член определителя D является членом и определителя D', так как его множители и в определителе D' находятся в разных строках и разных столбцах; обратно, каждый член определителя D' будет членом и определителя D. Таким образом, оба определителя представляют собой «алгебраическую сумму» (т. е. сумму, в которой некоторые слагаемые берутся со знаком минус) одних и тех же членов вида $a_{i_1k_1}a_{i_2k_2}\ldots a_{i_nk_n}$. Различие заключается только в том, что в определителе D первые индексы — это номера строк, а вторые — номера столбцов, а в определителе D' — наоборот. Но так как по лемме о знаке члена определителя знак такого произведения как в первом, так и во втором определителе будет одним и тем же:

$$(-1)^{[i_1, i_2, \ldots, i_n]+[k_1, k_2, \ldots, k_n]},$$

то $D' = D$.

Свойство 2. *Если поменять местами две строки или два столбца определителя, то определитель изменит знак, а по абсолютной величине не изменится.*

Докажем это утверждение, например, для столбцов. Поменяв в определителе

$$D_1 = \begin{vmatrix} a_{11} & a_{12} \ldots a_{1p} \ldots a_{1q} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2p} \ldots a_{2q} \ldots a_{2n} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} \ldots a_{np} \ldots a_{nq} \ldots a_{nn} \end{vmatrix}$$

местами p-й и q-й столбцы, мы получим определитель

$$D_2 = \begin{vmatrix} a_{11} & a_{12} \ldots a_{1q} \ldots a_{1p} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2q} \ldots a_{2p} \ldots a_{2n} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} \ldots a_{nq} \ldots a_{np} \ldots a_{nn} \end{vmatrix}.$$

Каждый член определителя D_1 будет членом и определителя D_2, так как его множители расположены и в D_2 в разных строках и разных столбцах, и обратно. Возьмем какой-нибудь член определителя D_1:

$$a_{i_11}a_{i_22}\ldots a_{i_pp}\ldots a_{i_qq}\ldots a_{i_nn}.$$

Так как его множители расположены в порядке следования столбцов в D_1, то он входит в определитель D_1 со знаком $(-1)^{[i_1, i_2, \ldots, i_p, \ldots, i_q, \ldots, i_n]}$. Для того чтобы найти знак этого члена в определителе D_2, расположим его множители в порядке следования столбцов в D_2:

$$a_{i_11}a_{i_22}\ldots a_{i_qq}\ldots a_{i_pp}\ldots a_{i_nn}$$

(элемент $a_{i_q q}$ содержится в p-м столбце определителя D_2, а элемент $a_{i_p p}$ — в q-м). Первые индексы в определителе D_2, так же как и в определителе D_1, указывают номера строк; поэтому в определитель D_2 рассматриваемое произведение войдет со знаком $(-1)^{[i_1, i_2, \ldots, i_q, \ldots, i_p, \ldots, i_n]}$. Но перестановка $i_1, i_2, \ldots, i_q, \ldots, i_p, \ldots, i_n$ получается из перестановки $i_1, i_2, \ldots, i_p, \ldots, i_q, \ldots, i_n$ посредством одной транспозиции, а значит, числа

$$[i_1, i_2, \ldots, i_p, \ldots, i_q, \ldots, i_n] \text{ и } [i_1, i_2, \ldots, i_q, \ldots, i_p, \ldots, i_n]$$

разной четности. Таким образом, каждый член определителя D_1 в определитель D_2 входит с противоположным знаком, и значит, $D_2 = -D_1$.

Для того чтобы доказать соответствующее предложение для строк, перейдем к транспонированным определителям D_1' (полученному из определителя D_1) и D_2' (полученному из D_2). Если определитель D_2 получается из D_1 перестановкой p-й и q-й строк, то D_2' получается из D_1' перестановкой p-го и q-го столбцов, и значит, $D_2' = -D_1'$. Но по свойству 1 $D_1' = D_1$ и $D_2' = D_2$, а поэтому $D_2 = -D_1$.

Следствие. *Определитель с двумя одинаковыми строками или с двумя одинаковыми столбцами равен нулю.*

Для доказательства поменяем местами одинаковые строки (или столбцы) определителя D; от этого он, конечно, не изменится. А так как, по свойству 2, он должен при этом изменить знак, то $D = -D$, откуда $D = 0$.

Свойство 3. *Если все элементы строки или столбца определителя умножить на одно и то же число, то определитель умножится на то же число.*

Доказательство проведем, например, для столбцов. Умножив все элементы k-го столбца определителя

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

на c, мы получим определитель

$$D_1=\begin{vmatrix} a_{11} & a_{12} \dots ca_{1k} \dots a_{1n} \\ a_{21} & a_{22} \dots ca_{2k} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots ca_{nk} \dots a_{nn} \end{vmatrix},$$

равный

$$\sum(-1)^{[i_1,i_2,\dots,i_n]}a_{i_11}a_{i_22}\dots(ca_{i_kk})\dots a_{i_nn}=$$
$$=c\sum(-1)^{[i_1,i_2,\dots,i_n]}a_{i_11}a_{i_22}\dots a_{i_kk}\dots a_{i_nn}=cD.$$

Соответствующее свойство для строк легко доказывается переходом к транспонированным определителям.

Таким образом, *общий множитель всех элементов строки или столбца определителя можно выносить за знак определителя.*

Следствие. *Определитель с двумя пропорциональными строками или столбцами равен нулю.*

В самом деле, вынося «множитель пропорциональности» строки (столбца) за знак определителя, придем к определителю с двумя одинаковыми строками (столбцами), который равен нулю ввиду следствия из свойства 2.

Свойство 4. *Если каждый элемент k-го столбца определителя представлен в виде суммы двух слагаемых: $a_{ik}=b_{ik}+c_{ik}$, т. е. если*

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

то D можно следующим образом представить в виде суммы двух определителей:

$$\begin{vmatrix} a_{11} & a_{12} \dots b_{1k} \dots a_{1n} \\ a_{21} & a_{22} \dots b_{2k} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots b_{nk} \dots a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} \dots c_{1k} \dots a_{1n} \\ a_{21} & a_{22} \dots c_{2k} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots c_{nk} \dots a_{nn} \end{vmatrix}=D_1+D_2.$$

Аналогичное утверждение справедливо и для строк.

Доказательство вытекает из равенства

$$D=\sum(-1)^{[i_1,i_2,\dots,i_n]}a_{i_11}a_{i_22}\dots(b_{i_kk}+c_{i_kk})\dots a_{i_nn}=$$
$$=\sum(-1)^{[i_1,i_2,\dots,i_n]}a_{i_11}a_{i_22}\dots b_{i_kk}\dots a_{i_nn}+$$
$$+\sum(-1)^{[i_1,i_2,\dots,i_n]}a_{i_11}a_{i_22}\dots c_{i_kk}\dots a_{i_nn}=D_1+D_2.$$

Замечание. Легко видеть, что справедливо и следующее, более общее утверждение: *Если каждый элемент k-го столбца определителя D представлен в виде суммы p слагаемых:* $a_{ik} = \overset{1}{a}_{ik} + \overset{2}{a}_{ik} + \ldots + \overset{p}{a}_{ik}$, *то определитель D можно представить в виде суммы p определителей:*

$$D = \sum_{j=1}^{p} \begin{vmatrix} a_{11} & a_{12} & \ldots & \overset{j}{a}_{1k} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & \overset{j}{a}_{2k} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & \overset{j}{a}_{nk} & \ldots & a_{nn} \end{vmatrix}.$$

Следствие. *Определитель не изменится, если ко всем элементам какой-либо его строки или какого-либо столбца прибавить соответствующие элементы параллельного ряда, умноженные на одно и то же число.*

Действительно, пусть дан определитель

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1p} & \ldots & a_{1q} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2p} & \ldots & a_{2q} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{np} & \ldots & a_{nq} & \ldots & a_{nn} \end{vmatrix}.$$

Прибавив ко всем элементам его p-го столбца соответствующие элементы q-го столбца, умноженные на одно и то же число c, мы получим определитель

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1p} + ca_{1q} & \ldots & a_{1q} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2p} + ca_{2q} & \ldots & a_{2q} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{np} + ca_{nq} & \ldots & a_{nq} & \ldots & a_{nn} \end{vmatrix}.$$

Ввиду свойства 4 определитель D_1 равен

$$\begin{vmatrix} a_{11} & \ldots & a_{1p} & \ldots & a_{1q} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2p} & \ldots & a_{2q} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{np} & \ldots & a_{nq} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \ldots & ca_{1q} & \ldots & a_{1q} & \ldots & a_{1n} \\ a_{21} & \ldots & ca_{2q} & \ldots & a_{2q} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & ca_{nq} & \ldots & a_{nq} & \ldots & a_{nn} \end{vmatrix},$$

т. е. равен D (второе слагаемое равно нулю как определитель с двумя пропорциональными столбцами).

§ 4. Миноры и алгебраические дополнения

Минором M_{ik} элемента a_{ik} определителя D n-го порядка называется определитель $(n-1)$-го порядка, получающийся из D вычеркиванием i-й строки и k-го столбца.

Алгебраическим дополнением A_{ik} элемента a_{ik} называется его минор, взятый со знаком $(-1)^{i+k}$:

$$A_{ik} = (-1)^{i+k} M_{ik}.$$

Теорема 2. *Если все элементы k-го столбца (строки) определителя D, кроме, быть может, одного, a_{ik}, равны нулю, то определитель D равен произведению a_{ik} на алгебраическое дополнение этого элемента:*

$$D = a_{ik} A_{ik}.$$

Доказательство. Рассмотрим сначала частный случай, когда в определителе D все элементы первого столбца, кроме a_{11}, равны нулю:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

В каждый член определителя D входит в точности по одному элементу из первого столбца; но так как все эти элементы, отличные от a_{11}, равны нулю, то в определителе D все те члены, в которые из первого столбца входит не a_{11}, а какой-либо другой элемент, равны нулю. Следовательно,

$$D = \sum (-1)^{[1, i_2, \dots, i_n]} a_{11} a_{i_2 2} \dots a_{i_n n},$$

где индексы $i_2, \dots, i_n$ принимают значения $2, 3, \dots, n$. Множитель a_{11} является общим для всех слагаемых, поэтому его можно вынести за знак суммы. С другой стороны, так как единица, стоящая на первом месте, не образует ни одной инверсии, то $[1, i_2, \dots, i_n] = [i_2, \dots, i_n]$, и значит,

$$D = a_{11} \sum (-1)^{[i_2, \dots, i_n]} a_{i_2 2} \dots a_{i_n n},$$

где суммирование распространяется на всевозможные перестановки $i_2, i_3, \dots, i_n$ чисел $2, 3, \dots, n$. А так как сумма

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

равна определителю $(n-1)$-го порядка, получающемуся из D вычеркиванием первой строки и первого столбца, т. е. равна M_{11}, и $A_{11} = (-1)^{1+1} M_{11} = M_{11}$, то

$$D = a_{11} M_{11} = a_{11} A_{11}.$$

Рассмотрим теперь общий случай, когда все элементы k-го столбца определителя D, кроме a_{ik}, равны нулю, т. е. когда определитель имеет вид

$$D = \begin{vmatrix} a_{11} & a_{12} \ldots 0 & \ldots a_{1n} \\ a_{21} & a_{22} \ldots 0 & \ldots a_{2n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{i1} & a_{i2} \ldots a_{ik} & \ldots a_{in} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} \ldots 0 & \ldots a_{nn} \end{vmatrix}.$$

Переместим i-ю строку определителя D на первое место, последовательно меняя ее местами с $(i-1)$-й, $(i-2)$-й, и т. д., наконец, с первой строкой. На это потребуется $i-1$ транспозиций строк, при каждой из которых определитель умножается на -1. Затем переместим k-й столбец определителя D на первое место, последовательно меняя его местами с $(k-1)$-м, $(k-2)$-м, и т. д., наконец, с первым столбцом. Для этого потребуется $k-1$ транспозиций столбцов, при каждой из которых определитель тоже умножается на -1. В конечном счете мы получим определитель

$$D_1 = \begin{vmatrix} a_{ik} & a_{i1} \ldots a_{in} \\ 0 & a_{11} \ldots a_{1n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ 0 & a_{n1} \ldots a_{nn} \end{vmatrix},$$

отличающийся от определителя D только знаком $(-1)^{i-1} \cdot (-1)^{k-1} = (-1)^{i+k}$. Но, как мы показали, определитель D_1 равен произведению a_{ik} на определитель $(n-1)$-го порядка, получающийся из D_1 вычеркиванием первого столбца и первой строки, или, что то же самое, получающийся из D вычеркиванием k-го столбца и i-й строки, т. е.

$$D_1 = a_{ik} M_{ik}$$

и, следовательно,

$$D = (-1)^{i+k} D_1 = (-1)^{i+k} a_{ik} M_{ik} = a_{ik} A_{ik}.$$

Доказанная теорема дает возможность, используя еще следствие из свойства 4, вычислить определитель какого угодно порядка.

Пример *Вычислить определитель пятого порядка:*

$$D = \begin{vmatrix} 2 & 0 & 1 & 3 & 1 \\ -1 & 1 & 2 & 2 & 3 \\ 1 & 4 & 0 & -1 & 5 \\ 2 & 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 3 & 1 \end{vmatrix}.$$

Решение. Вычитая из первого столбца определителя D удвоенный третий (иными словами, прибавляя к первому столбцу третий, умноженный на -2), из четвертого вычитая утроенный третий и из пятого — третий столбец, получим

$$D = \begin{vmatrix} 0 & 0 & 1 & 0 & 0 \\ -5 & 1 & 2 & -4 & 1 \\ 1 & 4 & 0 & -1 & 5 \\ -4 & 1 & 3 & -8 & -1 \\ 3 & 2 & -1 & 6 & 2 \end{vmatrix} = a_{13}A_{13} = 1 \cdot (-1)^{1+3} \begin{vmatrix} -5 & 1 & -4 & 1 \\ 1 & 4 & -1 & 5 \\ -4 & 1 & -8 & -1 \\ 3 & 2 & 6 & 2 \end{vmatrix}.$$

В полученном определителе четвертого порядка будем таким же образом «делать нули»: прибавим к первому столбцу четвертый, умноженный на 5, от второго отнимем четвертый и к третьему прибавим четвертый, умноженный на 4:

$$D = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 26 & -1 & 19 & 5 \\ -9 & 2 & -12 & -1 \\ 13 & 0 & 14 & 2 \end{vmatrix} = 1 \cdot (-1)^{1+4} \begin{vmatrix} 26 & -1 & 19 \\ -9 & 2 & -12 \\ 13 & 0 & 14 \end{vmatrix}.$$

Мы пришли к определителю третьего порядка, который уже можно вычислить либо непосредственно, либо сведя его к определителю второго порядка: прибавив ко второй строке удвоенную первую, получим

$$\begin{vmatrix} 26 & -1 & 19 \\ 43 & 0 & 26 \\ 13 & 0 & 14 \end{vmatrix} = (-1)(-1)^{1+2} \begin{vmatrix} 43 & 26 \\ 13 & 14 \end{vmatrix} = 2 \begin{vmatrix} 43 & 13 \\ 13 & 7 \end{vmatrix} = 2(301-169)=264;$$

значит, определитель $D = -264$.

§ 5 Разложение определителя по элементам строки или столбца

Теорема 3. *Каждый определитель равен сумме произведений элементов любой его строки (столбца) на их алгебраические дополнения.*

Мы докажем, что при всех $i, k = 1, 2, \ldots, n$

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$$

(разложение по элементам i-й строки) и

$$D = a_{1k}A_{1k} + a_{2k}A_{2k} + \ldots + a_{nk}A_{nk}$$

(разложение по элементам k-го столбца).

Для доказательства заметим прежде всего, что если два определителя отличаются друг от друга только элементами одного столбца (строки), то алгебраические дополнения элементов этих столбцов (строк) в обоих определителях одинаковы, так как при вычислении этих дополнений столбцы (строки), которыми отличаются определители, вычеркиваются.

Докажем теперь для определителя D справедливость, например, разложения по k-му столбцу. Для этого представим его в следующем виде:

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} + 0 + \ldots + 0 & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & 0 + a_{2k} + \ldots + 0 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & 0 + 0 + \ldots + a_{nk} & \ldots & a_{nn} \end{vmatrix}$$

(здесь каждый элемент k-го столбца представлен в виде суммы n слагаемых, $n-1$ из которых равны нулю). По свойству 4 (см. замечание на стр. 26) имеем

$$D = D_1 + D_2 + \ldots + D_n,$$

где

$$D_1 = \begin{vmatrix} a_{11} & \ldots & a_{1k} & \ldots & a_{1n} \\ a_{21} & \ldots & 0 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{vmatrix},$$

$$D_2 = \begin{vmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2k} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{vmatrix}, \ldots, \quad D_n = \begin{vmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ a_{21} & \ldots & 0 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nk} & \ldots & a_{nn} \end{vmatrix}.$$

Определитель D_1 равен произведению элемента a_{1k} на его алгебраическое дополнение *в этом определителе*. Однако так как определитель D_1 лишь k-м столбцом отличается от определителя D, то это алгебраическое дополнение совпадает с алгебраическим дополнением A_{1k} элемента a_{1k} *в определителе* D:

$$D_1 = a_{1k}A_{1k}.$$

Аналогично,

$$D_2 = a_{2k}A_{2k}, \ \ldots, \ D_n = a_{nk}A_{nk}.$$

Мы доказали, что

$$D = a_{1k}A_{1k} + a_{2k}A_{2k} + \ldots + a_{nk}A_{nk}.$$

Соответствующее равенство для строк легко получается переходом к транспонированному определителю.

Пример. *Вычислить определитель четвертого порядка*

$$D = \begin{vmatrix} -5 & 1 & -4 & 1 \\ 1 & 4 & -1 & 5 \\ -4 & 1 & -8 & -1 \\ 3 & 2 & 6 & 2 \end{vmatrix}.$$

Решение. Разложим определитель, например, по элементам первой строки

$$D = (-5)(-1)^{1+1}\begin{vmatrix} 4 & -1 & 5 \\ 1 & -8 & -1 \\ 2 & 6 & 2 \end{vmatrix} +$$

$$+ 1(-1)^{1+2}\begin{vmatrix} 1 & -1 & 5 \\ -4 & -8 & -1 \\ 3 & 6 & 2 \end{vmatrix} + (-4)(-1)^{1+3}\begin{vmatrix} 1 & 4 & 5 \\ -4 & 1 & -1 \\ 3 & 2 & 2 \end{vmatrix} +$$

$$+ 1(-1)^{1+4}\begin{vmatrix} 1 & 4 & -1 \\ -4 & 1 & -8 \\ 3 & 2 & 6 \end{vmatrix} = -5\cdot 74 - (-15) - 4(-31) - 33 =$$

$$= -370 + 15 + 124 - 33 = -264.$$

Теорема 4. *Сумма произведений элементов любой строки (или столбца) определителя на алгебраические дополнения соответствующих элементов параллельного ряда равна нулю.*

Доказательство. Пусть дан определитель

$$D = \begin{vmatrix} a_{11} & a_{12} \ldots a_{1i} \ldots a_{1k} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2i} \ldots a_{2k} \ldots a_{2n} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1} & a_{n2} \ldots a_{ni} \ldots a_{nk} \ldots a_{nn} \end{vmatrix}. \tag{12}$$

Рассмотрим другой определитель, D_1, отличающийся от D лишь тем, что в k-м его столбце повторен i-й столбец:

$$D_1 = \begin{vmatrix} a_{11} & a_{12} \dots a_{1i} \dots a_{1i} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2i} \dots a_{2i} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{ni} \dots a_{ni} \dots a_{nn} \end{vmatrix}.$$

Определитель D_1 равен нулю, как определитель с двумя одинаковыми столбцами (следствие из свойства 2). Разложив его по элементам k-го столбца, получим

$$D_1 = a_{1i}A_{1k} + a_{2i}A_{2k} + \dots + a_{ni}A_{nk},$$

где A_{jk} — алгебраические дополнения элементов k-го столбца определителя D_1; но так как определитель D_1 лишь k-м столбцом отличается от D, то они будут и алгебраическими дополнениями элементов k-го столбца определителя D. Таким образом, при всех i и $k \neq i$

$$a_{1i}A_{1k} + a_{2i}A_{2k} + \dots + a_{ni}A_{nk} = 0.$$

Аналогично, при всех i и $k \neq i$

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \dots + a_{in}A_{kn} = 0.$$

§ 6. Системы n линейных уравнений с n неизвестными

Рассмотрим *систему n линейных уравнений с n неизвестными*:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \dots \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{cases} \tag{13}$$

Решением системы (13) называется любая совокупность значений неизвестных $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_n = \alpha_n$, при подстановке которых все уравнения системы обращаются в тождества. Предположим, что определитель, составленный из коэффициентов при неизвестных системы (13), отличен от нуля:

$$D=\begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix} \neq 0.$$

Умножим первое уравнение системы на A_{11}, второе — на A_{21}, и т. д., последнее на A_{n1} и сложим их все. Мы получим уравнение

$$\begin{aligned} x_1(a_{11}A_{11}+a_{21}A_{21}+\dots+a_{n1}A_{n1})+ \\ +x_2(a_{12}A_{11}+a_{22}A_{21}+\dots+a_{n2}A_{n1})+\dots \\ \dots+x_n(a_{1n}A_{11}+a_{2n}A_{21}+\dots+a_{nn}A_{n1})= \\ =b_1A_{11}+b_2A_{21}+\dots+b_nA_{n1}, \end{aligned} \tag{14}$$

или

$$x_1D=b_1A_{11}+b_2A_{21}+\dots+b_nA_{n1}, \tag{15}$$

так как заключенные в скобки коэффициенты при неизвестных x_2, x_3, ..., x_n в уравнении (14) по теореме 4 равны нулю, а коэффициент при x_1, ввиду теоремы 3, равен D. При этом правая часть

$$b_1A_{11}+b_2A_{21}+\dots+b_nA_{n1}=D_1,$$

где D_1 — определитель, получающийся из D при замене первого столбца столбцом свободных членов. (В правых частях равенств (14) и (15) стоит разложение определителя D_1 по первому столбцу.) Аналогично уравнению (15), получаем

$$x_2D=D_2, \dots, x_nD=D_n, \tag{15a}$$

где D_i есть определитель, получающийся из D заменой i-го столбца столбцом свободных членов.

Система (15)—(15a) является *следствием* системы (13). Таким образом, мы доказали, что если система (13) имеет решение, то оно будет решением и системы (15)—(15a), и значит,

$$x_1=\frac{D_1}{D}, \quad x_2=\frac{D_2}{D}, \dots, x_n=\frac{D_n}{D}. \tag{16}$$

Формулы (16) называются *формулами Крамера*.

Непосредственной подстановкой этих значений неизвестных во все уравнения системы (13) можно убеиться, что они действительно образуют ее решение, В самом деле, подставляя значения (16) в i-е уравнение системы (13), будем иметь

$$a_{i1}\frac{D_1}{D}+a_{i2}\frac{D_2}{D}+\ldots+a_{in}\frac{D_n}{D}=\frac{1}{D}[a_{i1}(b_1A_{11}+b_2A_{21}+\ldots$$
$$\ldots+b_nA_{n1})+a_{i2}(b_1A_{12}+b_2A_{22}+\ldots+b_nA_{n2})+\ldots$$
$$\ldots+a_{in}(b_1A_{1n}+b_2A_{2n}+\ldots+b_nA_{nn})]=$$
$$=\frac{1}{D}[b_1(a_{i1}A_{11}+a_{i2}A_{12}+\ldots+a_{in}A_{1n})+b_2(a_{i1}A_{21}+$$
$$+a_{i2}A_{22}+\ldots+a_{in}A_{2n})+\ldots+b_n(a_{i1}A_{n1}+a_{i2}A_{n2}+\ldots$$
$$\ldots+a_{in}A_{nn})]=\frac{1}{D}b_iD=b_i.$$

Здесь скобки при всех b_k, кроме b_i, равны нулю по теореме 4, а сумма

$$a_{i1}A_{i1}+a_{i2}A_{i2}+\ldots+a_{in}A_{in}$$

равна D по теореме 3.

Этим доказана следующая

Теорема 5. *В случае, когда* $D\neq 0$, *система* (13) *имеет единственное решение, определяемое по формулам Крамера* (16).

§ 7. Ранг матрицы

Снова будем рассматривать таблицы чисел (матрицы), не требуя теперь, чтобы число строк матрицы совпадало с числом ее столбцов. Для таких (вообще говоря, *прямоугольных*) *матриц* мы введем важное понятие ранга.

Рассмотрим прямоугольную матрицу, состоящую из m строк и n столбцов ($[m\times n]$-матрицу). Пусть $k\leqslant m$ и $k\leqslant n$. Выделим в этой матрице какие-нибудь k строк и k столбцов. Из элементов, стоящих на пересечении выделенных строк и столбцов, составим определитель k-го порядка. Все такие определители называются *минорами* нашей матрицы. Ясно, что из $[m\times n]$-матрицы можно составить $C_m^k\cdot C_n^k$ миноров k-го порядка. Так, например, из матрицы

$$A=\begin{bmatrix}3&2&4&2\\2&0&1&1\\0&4&5&1\end{bmatrix}$$

можно составить $C_4^1 \cdot C_3^1 = 12$ миноров первого порядка — это сами элементы матрицы A, $C_4^2 \cdot C_3^2 = 6 \cdot 3 = 18$ миноров второго порядка:

$$\begin{vmatrix} 3 & 2 \\ 2 & 0 \end{vmatrix}, \quad \begin{vmatrix} 3 & 4 \\ 2 & 1 \end{vmatrix}, \quad \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 4 \\ 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 4 & 2 \\ 1 & 1 \end{vmatrix},$$

$$\begin{vmatrix} 3 & 2 \\ 0 & 4 \end{vmatrix}, \quad \begin{vmatrix} 3 & 4 \\ 0 & 5 \end{vmatrix}, \quad \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 4 \\ 4 & 5 \end{vmatrix}, \quad \begin{vmatrix} 2 & 2 \\ 4 & 1 \end{vmatrix}, \quad \begin{vmatrix} 4 & 2 \\ 5 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 0 \\ 0 & 4 \end{vmatrix},$$

$$\begin{vmatrix} 2 & 1 \\ 0 & 5 \end{vmatrix}, \quad \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 0 & 1 \\ 4 & 5 \end{vmatrix}, \quad \begin{vmatrix} 0 & 1 \\ 4 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 \\ 5 & 1 \end{vmatrix}$$

и $C_4^3 \cdot C_3^3 = 4$ минора третьего порядка:

$$\begin{vmatrix} 3 & 2 & 4 \\ 2 & 0 & 1 \\ 0 & 4 & 5 \end{vmatrix}, \quad \begin{vmatrix} 3 & 2 & 2 \\ 2 & 0 & 1 \\ 0 & 4 & 1 \end{vmatrix}, \quad \begin{vmatrix} 3 & 4 & 2 \\ 2 & 1 & 1 \\ 0 & 5 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 4 & 2 \\ 0 & 1 & 1 \\ 4 & 5 & 1 \end{vmatrix}.$$

Нетрудно проверить, что все миноры третьего порядка матрицы A равны нулю, а миноры второго порядка во всяком случае не все равны нулю (отличен от нуля уже первый из выписанных выше миноров второго порядка). Поэтому мы будем говорить, что р а н г м а т р и ц ы A р а в е н 2.

Рангом матрицы называется *наивысший порядок отличного от нуля минора этой матрицы.*

Таким образом, если ранг матрицы равен r, то среди миноров этой матрицы есть по крайней мере один минор -го порядка, отличный от нуля, в то время как все ее иноры порядка $r+1$ и выше равны нулю. Ранг матрицы A мы будем обозначать через $r(A)$.

Для вычисления ранга матрицы ее сначала приводят возможно более простому виду с помощью так называемых элементарных преобразований.

Элементарными преобразованиями матрицы называтся следующие ее преобразования:

1. Транспонирование, т. е. замена каждой строки толбцом с тем же номером.

2. Перестановка двух строк или двух столбцов.

3. Умножение всех элементов строки или столбца на юбое число c, отличное от нуля.

4. Прибавление ко всем элементам строки или столба соответствующих элементов параллельного ряда, множенных на одно и то же число.

Теорема 6 (об элементарных преобразованиях). *При элементарных преобразованиях матрицы ее ранг не меняется.*

Доказательство. Рассмотрим каждое преобразование отдельно. В первых трех случаях наше утверждение почти очевидно:

1. По свойству 1 определителей каждый минор транспонированной матрицы равен некоторому минору данной матрицы, и обратно.

2. После перестановки двух строк или двух столбцов матрицы A приходим к новой матрице, каждый минор которой либо равен некоторому минору матрицы A, либо отличается от некоторого минора матрицы A только знаком.

3. При умножении всех элементов строки или столбца матрицы на число c одни ее миноры не меняются, а другие умножаются на c; но так как $c \neq 0$, то наивысший порядок отличного от нуля минора этой матрицы не изменится.

4. Рассмотрим матрицу B, получающуюся из матрицы A прибавлением ко всем элементам ее i-го столбца соответствующих элементов k-го столбца, умноженных на одно и то же число c:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2k} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mk} & \dots & a_{mn} \end{bmatrix},$$

$$B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} + ca_{1k} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} + ca_{2k} & \dots & a_{2k} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mi} + ca_{mk} & \dots & a_{mk} & \dots & a_{mn} \end{bmatrix}.$$

Пусть ранг $r(A)$ матрицы A равен r. Покажем, что ранг матрицы B не больше чем r. Для этого достаточно показать, что каждый минор матрицы B порядка выше r равен 0. Пусть D будет минор порядка выше r матрицы B. Если D не содержит i-го столбца матрицы B, то он в точности равен соответствующему минору матрицы A, и, значит, равен 0 как минор порядка выше r, составленный из матрицы ранга r.

Если D содержит и i-й и k-й столбцы матрицы B, то по свойству 4 он тоже равен соответствующему минору матрицы A, и значит, равен 0, Наконец, если определитель D содержит i-й, но не содержит k-го столбца матрицы B, то по свойству 4 его можно представить в виде суммы двух определителей; $D = D_1 + D_2$, один из которых равен соответствующему минору матрицы A, а другой отличается от некоторого минора матрицы A множителем $\pm c$. (Знак минус здесь получается из-за того, что столбец с элементами a_{ik} может оказаться «не на своем месте». Так, например,

$$D = \begin{vmatrix} a_{21} & a_{22} + ca_{24} & a_{23} \\ a_{41} & a_{42} + ca_{44} & a_{43} \\ a_{51} & a_{52} + ca_{54} & a_{53} \end{vmatrix} = \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{vmatrix} + c \begin{vmatrix} a_{21} & a_{24} & a_{23} \\ a_{41} & a_{44} & a_{43} \\ a_{51} & a_{54} & a_{53} \end{vmatrix}.)$$

Следовательно, каждый из определителей D_1 и D_2 равен 0 и $D = 0$. Таким образом, каждый минор матрицы B порядка выше чем r равен нулю, и значит, $r(B) \leqslant r(A)$.

Но матрица A, в свою очередь, получается из матрицы B с помощью элементарного преобразования четвертого типа: чтобы получить матрицу A, надо к i-му столбцу матрицы B прибавить ее k-й столбец, умноженный на $-c$. По доказанному, ранг матрицы при этом не увеличивается, т. е. $r(A) \leqslant r(B)$. Следовательно,

$$r(A) = r(B).$$

Пример. *С помощью элементарных преобразований вычислить ранг матрицы*

$$A = \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ 0 & 4 & 5 & 1 \end{bmatrix}.$$

Решение. Вычитая из третьей строки удвоенную первую, сокращая второй столбец на 2 и вычитая после этого из первого столбца утроенный второй, из третьего — второй и из четвертого — удвоенный второй, последовательно получаем

$$A \sim \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{bmatrix},$$

где знак $\sim$ указывает, что соединяемые им матрицы получаются одна из другой элементарными преобразованиями и, значит, имеют один и тот же ранг.

Прибавляя далее к третьей строке утроенную вторую, сокращая первый столбец на 2, прибавляя его к третьему и вычитая из четвертого и поменяв, наконец, местами первые два столбца, будем иметь

$$A \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Мы видим, что ранг матрицы A равен 2.

§ 8. Понятие о линейной зависимости

Если обозначить строки матрицы A (см. § 7) через

$$e_1 = (3, 2, 1, 2), \; e_2 = (2, 0, -1, 1), \; e_3 = (0, 4, 5, 1),$$

то очевидно, что имеет место равенство

$$e_3 = 2e_1 - 3e_2,$$

понимаемое в смысле *поэлементного сложения*: каждый элемент строки e_3 равен соответствующему элементу строки e_1, умноженному на 2, без соответствующего элемента строки e_2, умноженного на 3.

Вообще, если $e_1, e_2, \ldots, e_m$ — строки какой-то матрицы A и, например,

$$e_m = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_{m-1} e_{m-1}, \tag{17}$$

где $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ — какие-то числа, мы будем говорить, что m-я строка этой матрицы *линейно выражается* через первые $m-1$ ее строк, или что e_m является *линейной комбинацией* строк $e_1, e_2, \ldots, e_{m-1}$. Из равенства (17) вытекает, что

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_{m-1} e_{m-1} - e_m = 0,$$

где нуль в правой части понимается как *нулевая строка* (т. е. как строка, состоящая из n нулей).

Мы будем говорить, что строки $e_1, e_2, \ldots, e_m$ матрицы A *линейно зависимы*, если можно подобрать такие числа $\gamma_1, \gamma_2, \ldots, \gamma_m$, не равные нулю одновременно, что

$$\gamma_1 e_1 + \gamma_2 e_2 + \ldots + \gamma_m e_m = 0. \tag{18}$$

Если таких чисел γ_i не существует, т. е. если равенство (18) имеет место только в том случае, когда все $\gamma_i = 0$, то говорят, что строки $e_1, e_2, \ldots, e_m$ *линейно независимы*.

Ясно, что если одна из строк матрицы линейно выражается через остальные, то строки этой матрицы между собой линейно зависимы. Обратно, пусть между строками матрицы A имеется линейная зависимость (18). Так как хотя бы одно из чисел γ_i, например γ_m, отлично от нуля, то

$$e_m = -\frac{\gamma_1}{\gamma_m} e_1 - \frac{\gamma_2}{\gamma_m} e_2 - \ldots - \frac{\gamma_{m-1}}{\gamma_m} e_{m-1},$$

т. е. в этом случае по крайней мере одна из строк матрицы линейно выражается через остальные.

Аналогичное понятие линейной зависимости можно ввести и для столбцов матрицы.

Теорема 7 (о ранге матрицы). *Если ранг матрицы равен r, то в этой матрице можно найти r линейно независимых строк (столбцов), через которые линейно выражаются все остальные ее строки (столбцы).*

Доказательство. Пусть дана $[m \times n]$-матрица A ранга r. Предположим, для определенности, что отличный от нуля минор r-го порядка (так называемый базисный минор) этой матрицы расположен в левом верхнем углу, т. е. что

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \qquad D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix} \neq 0.$$

Докажем, что в таком случае *первые r строк матрицы A будут линейно независимы.* (Если отличен от нуля не этот, а какой-нибудь другой минор r-го порядка матрицы A, то линейно независимыми будут именно те строки, которые образуют этот, базисный минор.) Предположим, что, наоборот, эти строки линейно зависимы; тогда одна из них, пусть, для определенности, e_r, линейно выражается через остальные:

$$e_r = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_{r-1} e_{r-1}.$$

Вычтем из r-й строки матрицы A первую строку, умноженную на α_1, вторую, умноженную на α_2, и т. д., наконец, $(r-1)$-ю, умноженную на α_{r-1}. После таких преобразований r-я строка матрицы A окажется состоящей из одних нулей. При этом определитель D, который, ввиду следствия из свойства 4, не должен был бы меняться, станет равным нулю. Полученное противоречие и доказывает линейную независимость первых r строк матрицы A.

Докажем теперь вторую часть теоремы — о том, что *все остальные строки матрицы A линейно выражаются через первые ее r строк*. Пусть $r<k\leqslant m$ и $1\leqslant l\leqslant n$; рассмотрим определитель $(r+1)$-го порядка:

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2l} \\ \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rl} \\ a_{k1} & a_{k2} & \dots & a_{kr} & a_{kl} \end{vmatrix}.$$

Он равен нулю при всех k и l: если $l\leqslant r$, то у него два одинаковых столбца, если же $l>r$, то это — минор $(r+1)$-го порядка матрицы ранга r.

Разложим определитель Δ по элементам последнего столбца

$$\Delta=a_{1l}A_1+a_{2l}A_2+\ldots+a_{rl}A_r+a_{kl}A_{r+1}=0. \qquad (19)$$

Алгебраические дополнения $A_1, A_2, \ldots, A_r, A_{r+1}$ элементов последнего столбца зависят от k, но не зависят от l, так как при их вычислении последний столбец вычеркивается. Кроме того, $A_{r+1}=D\neq 0$, и значит, равенство (19) можно разделить на A_{r+1}; это дает

$$a_{kl}=\alpha_1 a_{1l}+\alpha_2 a_{2l}+\ldots+\alpha_r a_{rl},$$

где коэффициенты $\alpha_i=-\frac{A_i}{D}$ не зависят от l. Подставляя $l=1, 2, \ldots, n$, будем иметь

$$a_{k1}=\alpha_1 a_{11}+\alpha_2 a_{21}+\ldots+\alpha_r a_{r1},$$
$$a_{k2}=\alpha_1 a_{12}+\alpha_2 a_{22}+\ldots+\alpha_r a_{r2},$$
$$\dots\dots\dots\dots\dots\dots$$
$$a_{kn}=\alpha_1 a_{1n}+\alpha_2 a_{2n}+\ldots+\alpha_r a_{rn}.$$

Но это означает, что k-я строка матрицы A линейно выражается через первые r ее строк:

$$e_k=\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_r e_r.$$

Следствие 1. *Максимальное число линейно независимых столбцов матрицы равно максимальному числу линейно независимых строк*, так как при транспонировании матрицы ее строки становятся столбцами, а ранг матрицы не меняется.

Следствие 2. *Для того чтобы определитель был равен нулю, необходимо и достаточно, чтобы его строки (столбцы) были линейно зависимы.*

Действительно, если определитель n-го порядка D равен нулю, то ранг соответствующей матрицы меньше n, и значит, ее строки (столбцы) линейно зависимы. Обратно, если строки (столбцы) определителя линейно зависимы, то ранг соответствующей матрицы меньше n, и этот определитель (n-го порядка) равен нулю.

§ 9. Произвольные системы линейных уравнений

Рассмотрим теперь *систему m линейных уравнений с n неизвестными*:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{20}$$

где число уравнений не предполагается равным числу неизвестных.

Решением системы (20) называется совокупность n значений неизвестных $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_n = \alpha_n$, при подстановке которых все уравнения системы обращаются в тождества. Система, имеющая хотя бы одно решение, называется *совместной*; система, не имеющая ни одного решения,— *несовместной*. Система, имеющая единственное решение, называется *определенной*; система, имеющая более одного решения,— *неопределенной*.

Рассмотрим две матрицы: матрицу A, *составленную из коэффициентов при неизвестных системы* (20), и матрицу

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

получаемую из A добавлением столбца свободных членов и называемую *расширенной матрицей*. Ясно, что

$r(B) \geqslant r(A)$, так как каждый минор матрицы A будет минором и матрицы B, но не наоборот.

Теорема 8 (критерий совместности системы линейных уравнений). *Для совместности системы* (20) *необходимо и достаточно, чтобы ранг расширенной матрицы B был равен рангу матрицы коэффициентов A.*

Доказательство необходимости. Предположим, что система (20) совместна, т. е. что существуют такие числа $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_n = \alpha_n$, что

$$
\begin{gathered}
a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n = b_1, \\
a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2n}\alpha_n = b_2, \\
\cdots\cdots\cdots\cdots \\
c_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n = b_m.
\end{gathered}
$$

Вычитая из последнего столбца матрицы B первый ее столбец, умноженный на α_1, второй, умноженный на α_2, и т. д., наконец, n-й, умноженный на α_n, мы получим матрицу

$$
C = \begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} & 0 \\
a_{21} & a_{22} & \ldots & a_{2n} & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
a_{m1} & a_{m2} & \ldots & a_{mn} & 0
\end{bmatrix},
$$

ранг которой по теореме об элементарных преобразованиях, равен рангу матрицы $B : r(C) = r(B)$.

Но ясно также, что $r(C) = r(A)$, так как все ненулевые миноры матрицы C равны соответствующим минорам матрицы A, и обратно. Следовательно, $r(B) = r(A)$.

Доказательство достаточности. Пусть

$$r(B) = r(A) = r,$$

и предположим, для определенности, что отличный от нуля определитель r-го порядка матрицы A расположен в левом верхнем ее углу:

$$
D = \begin{vmatrix}
a_{11} & a_{12} & \ldots & a_{1r} \\
a_{21} & a_{22} & \ldots & a_{2r} \\
\cdot & \cdot & \cdot & \cdot \\
a_{r1} & a_{r2} & \ldots & a_{rr}
\end{vmatrix} \neq 0.
$$

Тогда первые r строк матрицы B линейно независимы, а *так как ранг ее в точности равен r, то осталь-*

ные строки матрицы B линейно выражаются через первые r ее строк. Но это означает, что *первые r уравнений системы* (20) *независимы, а остальные $m - r$ ее уравнений* «являются их линейными комбинациями», т. е. просто *являются их следствиями.* В этом случае система на самом деле состоит лишь из r независимых уравнений. Нам достаточно поэтому решить первые r уравнений системы; их решения автоматически будут удовлетворять и остальным $m - r$ уравнениям.

Далее возможны два случая:

1. $r = n$. Тогда систему, состоящую из первых r уравнений системы (20)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = b_r, \end{cases}$$

можно решить, например, по формулам Крамера. В этом случае система имеет е д и н с т в е н н о е решение. Она — *совместная* и *определенная.*

2. $r<n$. Возьмем первые r уравнений системы и, оставив в левых частях первые r неизвестных, остальные перенесем в правые части:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1r}x_r=b_1-a_{1,r+1}x_{r+1}-\ldots-a_{1n}x_n, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2r}x_r=b_2-a_{2,r+1}x_{r+1}-\ldots-a_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+a_{r2}x_2+\ldots+a_{rr}x_r=b_r-a_{r,r+1}x_{r+1}-\ldots-a_{rn}x_n. \end{cases} \tag{21}$$

«Свободным неизвестным» $x_{r+1}, x_{r+2}, \ldots, x_n$ можно придавать к а к и е у г о д н о значения, получая при этом соответствующие значения неизвестных $x_1, x_2, \ldots, x_r$ из системы (21). Это — случай *совместной*, но *не определенной* системы. Общие формулы решения можно получить, если решить систему (21) относительно x_1, $x_2, \ldots, x_r$, например по формулам Крамера.

Этим и завершается доказательство теоремы 8: если $r(B) = r(A)$, то система (20) — совместная (определенная или неопределенная), если $r(B) > r(A)$, то система (20) — несовместна.

Теперь мы в состоянии ответить на вопрос, остававшийся пока открытым: что можно сказать о системе n линейных уравнений с n неизвестными, определитель которой равен нулю. Для такой системы ранг матрицы коэффициентов $r<n$, так как единственный минор n-го порядка этой матрицы, по условию, равен нулю. Если ранг расширенной матрицы B такой системы тоже равен r, то система будет *совместной*, но, поскольку $r<n$, *неопределенной*; если же ранг матрицы B больше r, то система *несовместна*.

Пример. *Решить следующие системы уравнений:*

1. $\begin{cases} x_1+2x_2+3x_3=2, \\ x_1-x_2+x_3=0, \\ x_1+3x_2-x_3=-2, \\ 3x_1+4x_2+3x_3=0. \end{cases}$ 2. $\begin{cases} x_1+2x_2+3x_3-x_4=0, \\ x_1-x_2+x_3+2x_4=4, \\ x_1+5x_2+5x_3-4x_4=-4, \\ x_1+8x_2+7x_3-7x_4=-8. \end{cases}$

3. $\begin{cases} x_1+2x_2+3x_3-x_4=0, \\ x_1-x_2+x_3+2x_4=4, \\ x_1+5x_2+5x_3-4x_4=-4, \\ x_1+8x_2+7x_3-7x_4=6. \end{cases}$

Решение.

1. Здесь $r(A)=3$, $r(B)=3$; система *совместная, определенная.* Так как

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 1 & 3 & -1 \end{vmatrix} = 14 \neq 0,$$

то из первых трех уравнений системы, например по формулам Крамера, находим

$$x_1=-1,\ x_2=0,\ x_3=1.$$

2. Здесь $r(A)=2$, $r(B)=2$; система *совместная*, но *не определенная*. Определитель

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

и из первых двух уравнений системы

$$\begin{cases} x_1+2x_2=-3x_3+x_4, \\ x_1-x_2=4-x_3-2x_4 \end{cases}$$

находим

$$x_1 = \frac{8}{3} - \frac{5}{3}x_3 - x_4, \quad x_2 = -\frac{4}{3} - \frac{2}{3}x_3 + x_4,$$

где неизвестным x_3 и x_4 можно придавать любые значения.

3. Здесь $r(A) = 2$, $r(B) = 3$, и система *несовместна.*

§ 10. Однородные системы

Однородные линейные уравнения — это уравнения, свободные члены которых равны нулю:

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

Система (22) однородных уравнений (или *однородная система* линейных уравнений) всегда *совместна*, так как имеет, например, нулевое решение:

$$x_1 = 0, x_2 = 0, \ldots, x_n = 0$$

(т. е. решение, в котором значения всех неизвестных равны нулю). Это следует также из теоремы 8, так как в этом случае, разумеется, $r(B) = r(A)$.

Важно выяснить, при каком условии однородная система (22) является *неопределенной*, а значит,— что бывает особенно важно — имеет и ненулевые решения. Ответ на этот вопрос дает следующая

Теорема 9. *Для того чтобы система* (22) *имела ненулевые решения, необходимо и достаточно, чтобы ранг r ее матрицы коэффициентов был меньше n.*

Действительно, если $r = n$, то, как видно из доказательства теоремы 8, система (22) имеет единственное и, значит, только нулевое решение:

$$x_1 = x_2 = \ldots = x_n = 0.$$

Если же $r < n$, то система (22) является неопределенной (ведь несовместной она быть не может), и значит, она имеет бесчисленное множество решений, в том числе и бесчисленное множество ненулевых решений.

Из доказанной теоремы непосредственно вытекает

Теорема 10. *Для того чтобы однородная система n линейных уравнений с n неизвестными обладала ненулевыми решениями, необходимо и достаточно, чтобы ее определитель D был равен нулю.*

Доказательство. Условие

$$D = 0$$

здесь необходимо, так как если $D \neq 0$, то система имеет единственное и, значит, только нулевое решение. Это условие также и достаточно, так как если $D = 0$, то ранг матрицы коэффициентов системы $r<n$, и система имеет бесчисленное множество (ненулевых) решений.

Пусть

$$x_1 = \alpha_1, \; x_2 = \alpha_2, \; \ldots, \; x_n = \alpha_n$$

— какое-нибудь ненулевое решение однородной системы (22). Это решение можно рассматривать как строку $e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, состоящую из n элементов. Тогда строка

$$ce_1 = (c\alpha_1, c\alpha_2, \ldots, c\alpha_n)$$

тоже, очевидно, будет решением системы (22). Далее, если

$$e_2 = (\beta_1, \beta_2, \ldots, \beta_n)$$

— какое-то другое решение системы (22), то при любых c_1 и c_2 линейная комбинация

$$c_1e_1 + c_2e_2 = (c_1\alpha_1 + c_2\beta_1, \; c_1\alpha_2 + c_2\beta_2, \; \ldots, \; c_1\alpha_n + c_2\beta_n)$$

этих решений тоже будет решением системы, так как если

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n = 0,$$
$$a_{i1}\beta_1 + a_{i2}\beta_2 + \ldots + a_{in}\beta_n = 0,$$

то и

$$a_{i1}(c_1\alpha_1 + c_2\beta_1) + a_{i2}(c_1\alpha_2 + c_2\beta_2) + \ldots + a_{in}(c_1\alpha_n + c_2\beta_n) = 0.$$

Итак, *любая линейная комбинация решений однородной системы* (22) *тоже будет ее решением.* Интересно поэтому найти такие линейно независимые решения системы (22), через которые линейно выражались бы все остальные ее решения.

Линейно независимая система решений $e_1, e_2, \ldots, e_k$ уравнений (22) *называется фундаментальной,* *если каждое решение системы* (22) *является линейной комбинацией решений* $e_1, e_2, \ldots, e_k$.

Теорема 11 (о существовании фундаментальных систем решений). *Если ранг* r *матрицы коэффициентов системы уравнений* (22) *меньше* n, *то эта система обладает фундаментальными системами решений.*

Доказательство. Пусть ранг r матрицы A коэффициентов системы (22) меньше n, и пусть, для определенности, минор D r-го порядка, стоящий в левом верхнем углу матрицы A, отличен от нуля:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix} \neq 0.$$

Перенеся свободные неизвестные $x_{r+1}, \ldots, x_n$ первых r уравнений системы (22) в правые части, получим систему

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = -a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = -a_{2,r+1}x_{r+1} - \ldots - a_{2n}x_n, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = -a_{r,r+1}x_{r+1} - \ldots - a_{rn}x_n. \end{array} \right. \tag{23}$$

Придавая свободным неизвестным значения

$$x_{r+1} = 1,\ x_{r+2} = 0,\ \ldots,\ x_n = 0,$$

получим соответствующие значения $x_1 = \alpha_1$, $x_2 = \alpha_2, \ldots$ $\ldots, x_r = \alpha_r$ первых r неизвестных. Это дает нам строку — решение системы (22)

$$(\alpha_1, \alpha_2, \ldots, \alpha_r, 1, 0, \ldots, 0).$$

Аналогично, придавая свободным неизвестным значения

$$x_{r+1} = 0,\ x_{r+2} = 1,\ \ldots,\ x_n = 0$$

и вычисляя соответствующие значения неизвестных $x_1 = \beta_1$, $x_2 = \beta_2, \ldots, x_r = \beta_r$, получим строку

$$(\beta_1, \beta_2, \ldots, \beta_r, 0, 1, \ldots, 0),$$

и т. д. Так мы найдем всего $k = n - r$ решений системы (22):

$$e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_r, 1, 0, \ldots, 0),$$

$$\begin{aligned} e_2 &= (\beta_1, \beta_2, \ldots, \beta_r, 0, 1, \ldots, 0), \\ &\cdots\cdots\cdots\cdots \\ e_k &= (\xi_1, \xi_2, \ldots, \xi_r, 0, 0, \ldots, 1). \end{aligned} \tag{24}$$

Эти k строк между собой линейно независимы, ибо ранг образованной ими матрицы

$$\begin{bmatrix} \alpha_1 & \alpha_2 \ldots \alpha_r & 1 & 0 \ldots 0 \\ \beta_1 & \beta_2 \ldots \beta_r & 0 & 1 \ldots 0 \\ \cdot & \cdots & \cdot & \cdots \\ \xi_1 & \xi_2 \ldots \xi_r & 0 & 0 \ldots 1 \end{bmatrix} \tag{25}$$

в точности равен k. (В этой матрице есть отличный от нуля минор k-го порядка, например, содержащий последние k столбцов.)

Покажем теперь, что решения $e_1, e_2, \ldots, e_k$ (24) действительно образуют ф у н д а м е н т а л ь н у ю систему. Для этого остается показать, что *каждое решение системы* (22) *линейно выражается через* $e_1, e_2, \ldots, e_k$. Итак, пусть

$$e = (\vartheta_1, \vartheta_2, \ldots, \vartheta_r, \vartheta_{r+1}, \ldots, \vartheta_n)$$

— произвольное решение системы (22). Рассмотрим строку

$$e_0 = e - \vartheta_{r+1}e_1 - \vartheta_{r+2}e_2 - \ldots - \vartheta_n e_k.$$

Легко видет, что все элементы, состоящие на последних k местах этой строки, равны нулю, т. е. что

$$e_0 = (\rho_1, \rho_2, \ldots, \rho_r, 0, 0, \ldots, 0).$$

Будучи линейной комбинацией решений, строка e_0 сама будет решением системы (22). А так как значения всех свободных неизвестных в e_0 равны нулю, то из о д н о р о д н о й в этом случае системы (23), определитель которой отличен от нуля, получаем, что и значения всех остальных неизвестных в e_0 должны быть равны нулю, т. е. что e_0 есть нулевая строка:

$$e_0 = e - \vartheta_{r+1} e_1 - \vartheta_{r+2} e_2 - \ldots - \vartheta_n e_k = (0, 0, \ldots, 0),$$

и

$$e = \vartheta_{r+1} e_1 + \vartheta_{r+2} e_2 + \ldots + \vartheta_n e_k,$$

что и требовалось доказать.

Заметим, что для того чтобы получить фундаментальную систему решений, мы могли бы придавать свободным неизвестным и какие угодно другие значения, лишь бы соответствующий определитель k-го порядка был отличен от нуля. Так можно найти сколько угодно фундаментальных систем решений, каждая из которых состоит из $k = n - r$ строк. Из результатов следующей главы будет видно, что любая фундаментальная система решений уравнений (22) состоит в точности из $n - r$ решений.

Таким образом, можно сказать, что *общее решение* системы (22) линейных однородных уравнений имеет вид

$$c_1e_1 + c_2e_2 + \ldots + c_ke_k, \tag{26}$$

где $e_1, e_2, \ldots, e_k$ какая-то (какая угодно!) фундаментальная система решений, а $c_1, c_2, \ldots, c_k$ — произвольные числа.

Сделаем еще одно, важное для дальнейшего,

Замечание. Рассмотрим систему уравнений

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}\right. \tag{27}$$

и соответствующую ей систему однородных уравнений

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{array}\right. \tag{28}$$

Пусть $e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ — какое-то фиксированное решение системы (27) и $e_2 = (\beta_1, \beta_2, \ldots, \beta_n)$ — любое другое ее решение. Тогда разность

$$e_1 - e_2 = (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \ldots, \alpha_n - \beta_n)$$

удет решением системы (28): если $a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots$ $\ldots + a_{in}\alpha_n = b_i$ и $a_{i1}\beta_1 + a_{i2}\beta_2 + \ldots + a_{in}\beta_n = b_i$, то $a_{i1}(\alpha_1-\beta_1) + a_{i2}(\alpha_2-\beta_2) + \ldots + a_{in}(\alpha_n-\beta_n) = b_i - b_i = 0.$

Далее, если $e_3=(\gamma_1, \gamma_2, \ldots, \gamma_n)$ — произвольное решение однородной системы (28), то строка $e_1+e_3=(\alpha_1+\gamma_1, \alpha_2+\gamma_2, \ldots, \alpha_n+\gamma_n)$ будет удовлетворять системе (27): если $a_{i1}\alpha_1+a_{i2}\alpha_2+\ldots+a_{in}\alpha_n=b_i$ и $a_{i1}\gamma_1+a_{i2}\gamma_2+\ldots+a_{in}\gamma_n=0$, то

$$a_{i1}(\alpha_1+\gamma_1)+a_{i2}(\alpha_2+\gamma_2)+\ldots+a_{in}(\alpha_n+\gamma_n)=b_i+0=b_i.$$

Отсюда следует, что все решения системы (27) можно получить, прибавляя к одному какому-нибудь ее решению всевозможные решения однородной системы (28). Иными словами, *общее решение системы* (27) *равно сумме общего решения однородной системы* (28) *и произвольного, но фиксированного решения системы* (27): если $e_1, e_2, \ldots, e_k$ — фундаментальная система решений однородной системы (28) и e_0 — произвольное фиксированное решение системы (27), то общее решение системы (27) имеет вид

$$e_0+c_1e_1+c_2e_2+\ldots+c_ke_k,$$

где $c_1, c_2, \ldots, c_k$ — произвольные числа.

§ 11. Метод Гаусса

Формулы Крамера, представляющие большой теоретический интерес, серьезного практического значения, однако, не имеют, так как их применение приводит к слишком громоздким вычислениям. Практически для решения систем линейных уравнений чаще всего применяется метод Гаусса, состоящий в последовательном исключении неизвестных по следующей схеме. Для того чтобы решить систему уравнений

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m, \end{array}\right. \qquad (29)$$

выписывают расширенную матрицу этой системы:

$$B=\left[\begin{array}{ccc|c} a_{11} & a_{12} & \ldots a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots a_{mn} & b_m \end{array}\right],$$

где чертой отделен столбец свободных членов; затем над строками матрицы B производят элементарные преобразования: разрешается изменять порядок строк (что соответствует изменению порядка уравнений), умножать строки на любые отличные от нуля числа (что отвечает умножению соответствующих уравнений на эти числа) и прибавлять к любой строке матрицы B любую другую ее строку, умноженную на любое число (что соответствует прибавлению к одному из уравнений системы другого уравнения, умноженного на это число). С помощью таких преобразований каждый раз получается расширенная матрица новой системы, равносильной исходной. При этом стараются привести матрицу B к возможно более простому виду, из которого решение системы видно непосредственно.

Рассмотрим подробнее метод Гаусса на трех конкретных примерах.

$$1.\ \begin{cases} x_1 + 2x_2 + 3x_3 = 2, \\ x_1 - x_2 + x_3 = 0, \\ x_1 + 3x_2 - x_3 = -2, \\ 3x_1 + 4x_2 + 3x_3 = 0. \end{cases} \tag{30}$$

Расширенная матрица этой системы имеет вид

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 1 & -1 & 1 & 0 \\ 1 & 3 & -1 & -2 \\ 3 & 4 & 3 & 0 \end{array}\right].$$

Вычитая первую строку из второй и из третьей и утроенную первую из четвертой, получим матрицу

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & -3 & -2 & -2 \\ 0 & 1 & -4 & -4 \\ 0 & -2 & -6 & -6 \end{array}\right].$$

Эта матрица — расширенная матрица системы

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 2, \\ \qquad -3x_2 - 2x_3 = -2, \\ \qquad\quad x_2 - 4x_3 = -4, \\ \qquad -2x_2 - 6x_3 = -6, \end{cases} \tag{31}$$

которая получается из заданной системы (30), если первое уравнение вычесть из второго и третьего, а утроенное первое вычесть из четвертого. Поэтому система (31) является следствием системы (30) — каждое решение системы (30) будет удовлетворять и системе (31). Но и обратно, система (30) может быть получена из системы (31) посредством аналогичных преобразований: первое уравнение прибавляется ко второму и третьему, а утроенное первое — к четвертому. Поэтому система (30) будет, в свою очередь, следствием системы (31), и значит, обе системы равносильны — они имеют одни и те же решения.

Далее, прибавив утроенную третью строку ко второй и удвоенную третью к четвертой, получим

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 2 \\ 0 & 0 & -14 & -14 \\ 0 & 1 & -4 & -4 \\ 0 & 0 & -14 & -14 \end{array}\right].$$

Вычитая вторую строку из четвертой и сокращая ее на — 14, будем иметь

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -4 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Но это — расширенная матрица системы

$$\left\{\begin{aligned} x_1 + 2x_2 + 3x_3 &= 2, \\ x_3 &= 1, \\ x_2 - 4x_3 &= -4, \end{aligned}\right.$$

равносильной заданной системе (30), и значит, решением системы (30) будет

$$x_3 = 1,\quad x_2 = -4 + 4x_3 = 0,\quad x_1 = 2 - 2x_2 - 3x_3 = 2 - 3 = -1.$$

В этом случае ранг расширенной матрицы равен рангу матрицы коэффициентов и равен, очевидно, трем.

$$2.\quad \left\{\begin{aligned} x_1 + 2x_2 + 3x_3 - x_4 &= 0, \\ x_1 - x_2 + x_3 + 2x_4 &= 4, \\ x_1 + 5x_2 + 5x_3 - 4x_4 &= -4, \\ x_1 + 8x_2 + 7x_3 - 7x_4 &= -8. \end{aligned}\right.$$

Выписав расширенную матрицу этой системы, после очевидных преобразований получим

$$\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 1 & -1 & 1 & 2 & 4 \\ 1 & 5 & 5 & -4 & -4 \\ 1 & 8 & 7 & -7 & -8 \end{array}\Bigg] \sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 3 & 2 & -3 & -4 \\ 0 & 6 & 4 & -6 & -8 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

ткуда $\begin{cases} x_1+2x_2+3x_3-x_4=0, \\ -3x_2-2x_3+3x_4=4, \end{cases}$ и значит,

$$x_2 = -\frac{4}{3} - \frac{2}{3}x_3 + x_4,$$

$$x_1 = -2x_2 - 3x_3 + x_4 = \frac{8}{3} - \frac{5}{3}x_3 - x_4.$$

десь ранг расширенной матрицы равен рангу матрицы оэффициентов и равен, очевидно, двум.

3. $\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 0, \\ x_1 - x_2 + x_3 + 2x_4 = 4, \\ x_1 + 5x_2 + 5x_3 - 4x_4 = -4, \\ x_1 + 8x_2 + 7x_3 - 7x_4 = 6. \end{cases}$

меем, очевидно,

$$\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 1 & -1 & 1 & 2 & 4 \\ 1 & 5 & 5 & -4 & -4 \\ 1 & 8 & 7 & -7 & 6 \end{array}\Bigg] \sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 3 & 2 & -3 & -4 \\ 0 & 6 & 4 & -6 & 6 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 14 \end{array}\right],$$

значит, система несовместна, так как равносильная ей истема содержит уравнение

$0\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4=14$ (последняя строка).

десь ранг матрицы коэффициентов равен, как легко идеть, двум, а ранг расширенной матрицы равен трем.

Пример. *Методом Гаусса решить однородную систему уравнений*

$$\begin{cases} x_1+2x_2+3x_3+4x_4+5x_5=0,\\ 2x_1+3x_2+4x_3+5x_4+x_5=0,\\ 3x_1+4x_2+5x_3+x_4+2x_5=0,\\ x_1+3x_2+5x_3+12x_4+9x_5=0,\\ 4x_1+5x_2+6x_3-3x_4+3x_5=0. \end{cases}$$

и найти ее фундаментальную систему решений.

Решение. Выпишем расширенную матрицу системы (при этом нулевой столбец можно, конечно, не писать). После понятных преобразований будем иметь

$$\begin{bmatrix} 1&2&3&4&5\\ 2&3&4&5&1\\ 3&4&5&1&2\\ 1&3&5&12&9\\ 4&5&6&-3&3 \end{bmatrix} \sim \begin{bmatrix} 1&2&3&4&5\\ 0&-1&-2&-3&-9\\ 0&-2&-4&-11&-13\\ 0&1&2&8&4\\ 0&-3&-6&-19&-17 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} 1&2&3&4&5\\ 0&-1&-2&-3&-9\\ 0&0&0&-5&5\\ 0&0&0&5&-5\\ 0&0&0&-10&10 \end{bmatrix} \sim \begin{bmatrix} 1&2&3&4&5\\ 0&1&2&3&9\\ 0&0&0&-1&1\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{bmatrix},$$

т. е. заданная система равносильна следующей:

$$\begin{cases} x_1+2x_2+3x_3+4x_4+5x_5=0,\\ \qquad x_2+2x_3+3x_4+9x_5=0,\\ \qquad\qquad -x_4+x_5=0. \end{cases}$$

Здесь $r=3$, и три неизвестных можно выразить через остальные, например, так:

$$x_4=x_5,$$
$$x_2=-2x_3-3x_4-9x_5=-2x_3-12x_5,$$
$$x_1=-2x_2-3x_3-4x_4-5x_5=x_3+15x_5.$$

Фундаментальную систему можно получить, если свободным неизвестным x_3, x_5 придавать значения $x_3=1$, $x_5=0$ (тогда $x_1=1$, $x_2=-2$, $x_4=0$) и значения $x_3=0$, $x_5=1$ (тогда $x_1=15$, $x_2=-12$, $x_4=1$). Это дает фундаментальную систему решений:

$$e_1=(1,\,-2,\,1,\,0,\,0),\qquad e_2=(15,\,-12,\,0,\,1,\,1).$$

Общее решение системы имеет вид

$$e=\alpha_1e_1+\alpha_2e_2=(\alpha_1+15\alpha_2,\,-2\alpha_1-12\alpha_2,\,\alpha_1,\,\alpha_2,\,\alpha_2),$$

где α_1 и α_2 — произвольные числа.

ГЛАВА II

n-МЕРНОЕ ПРОСТРАНСТВО

§ 1. Что такое поле

В первой главе мы рассматривали системы линейных уравнений, коэффициентами которых являются числа. Мы намеренно не уточняли, какие именно числа; читатель мог считать эти коэффициенты произвольными вещественными числами — тогда и решения системы будут вещественными. Однако с тем же успехом он мог считать, что это — комплексные числа; тогда и решения системы были бы образованы комплексными числами, но все теоремы из главы I остались бы справедливыми и для этого случая. С другой стороны, можно было бы ограничиться рассмотрением систем уравнений, скажем, с рациональными коэффициентами. Их решения будут образованы тоже рациональными числами, но все предложения первой главы останутся справедливыми.

Здесь все дело в том, что вещественные числа (а также комплексные или одни только рациональные числа) можно складывать и перемножать по известным правилам арифметики, получая при этом такие же числа. Это выражают словами: вещественные числа (а также комплексные, рациональные числа) образуют поле.

Полем называется множество F элементов), для которых определены две алгебраические операции — сложение и умножение* (так что сумма $a+b$ и произведение ab любых двух элементов a и b из F принадлежат F), *причем выполнены следующие условия* (аксиомы поля):

1. $a+b=b+a$ *для всех* a, b *из* F (сложение коммутативно).

*) Предполагается, что множество F состоит более чем из одного элемента.

2. $(a+b)+c=a+(b+c)$ *для всех* a, b, c *из* F (сложение ассоциативно).

3. *В множестве* F *имеется нуль, т.е. такой элемент* 0, *что для каждого* a *из* F *сумма* $a+0=a$.

4. *Для каждого* a *из* F *существует такой (противоположный* a) *элемент* $-a$, что $a+(-a)=0$.

5. $ab=ba$ *для всех* a, b *из* F (умножение коммутативно).

6. $(ab)c=a(bc)$ *для всех* a, b, c *из* F (умножение ассоциативно).

7. *В множестве* F *имеется единица — такой элемент* 1, *что для всякого* a *из* F *имеем* $a\cdot 1=a$.

8. *Для каждого отличного от нуля элемента* a *из* F *имеется такой (обратный* a) *элемент* a^{-1}, *что* $aa^{-1}=1$.

9. $(a+b)c=ac+bc$ *для всех* a, b, c *из* F (умножение дистрибутивно относительно сложения).

Ясно, что если коэффициенты системы линейных уравнений с n неизвестными принадлежат полю F, то и решение ее (если оно существует) следует искать среди наборов из n элементов поля F. Поля, которые на практике встречаются чаще всего,— это поле вещественных чисел, поле комплексных чисел и (реже) поле рациональных чисел. Поле рациональных чисел является, очевидно, частью (или, как говорят, *подполем*) поля вещественных чисел; последнее же содержится в качестве подполя в поле комплексных чисел.

В дальнейшем, говоря о числах, мы всегда будем иметь в виду элементы некоторого фиксированного числового поля F — обычно это будет либо поле вещественных чисел, либо поле комплексных чисел.

Полагая, что с полем вещественных чисел читатель достаточно знаком, мы изложим кратко необходимые для дальнейшего сведения о комплексных числах.

§ 2. Поле комплексных чисел

Комплексным числом называется выражение вида $a+bi$ (или, что то же самое, $a+ib$), где a и b — любые вещественные числа, а i — некоторый новый символ. По определению, $a+bi=c+di$ в том и только в том случае, если $a = c$ и $b = d$. По определению же,

$$a - bi = a + (-b)i,$$

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

и

$$(a + bi)(c + di) = (ac - bd) + (bc + ad)i.$$

Легко видеть, что так определенные сложение и умножение комплексных чисел коммутативны и ассоциативны (проверьте это!). Комплексное число $0+0i$ можно обозначить просто через 0: для любого комплексного числа $\alpha = a + bi$ имеем

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

Комплексное число $-\alpha = (-a) + (-b)i$ *будет* **противоположным** α, так как

$$\alpha + (-\alpha) = (a + bi) + [(-a) + (-b)i] = 0.$$

Комплексное число $1 + 0i$ обозначим просто 1: для любого комплексного числа $\alpha = a + bi$ произведение $\alpha \cdot 1$ равно α, так как

$$\alpha \cdot 1 = (a + bi)(1 + 0i) = (a \cdot 1 - b \cdot 0) + \\ + (b \cdot 1 + a \cdot 0)i = \alpha.$$

Далее, для каждого комплексного числа $\alpha = a + bi \neq 0$ существует обратное ему число $\alpha^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i$, произведение которого на α равно 1:

$$\alpha\alpha^{-1} = (a + bi)\left(\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i\right) = \frac{a^2 + b^2}{a^2 + b^2} + \frac{ba - ab}{a^2 + b^2}i = 1.$$

Наконец, сложение и умножение комплексных чисел связаны дистрибутивным законом — это легко проверяется непосредственно.

В множестве комплексных чисел рассмотрим числа вида $a + 0i$ — такое число можно обозначить просто через a (выше мы уже сделали это для 0 и 1). Сумма таких чисел $a = a + 0i$ и $b = b + 0i$, равная

$$(a + 0i) + (b + 0i) = (a + b) + 0i = a + b,$$

и их произведение

$$(a+0i)(b+0i)=ab+0i=ab$$

имеют такой же вид. Таким образом, числа вида $a+0i=a$ в поле комплексных чисел образуют подполе, которое можно отождествить с полем вещественных чисел.

Заметим, что

$$c\cdot(a+bi)=(c+0i)(a+bi)=ca+cbi.$$

Далее имеем

$$(0+i)(0+i)=-1+0i=-1.$$

Число $0+i$ можно обозначить просто i. При этом $i^2=i\cdot i=-1$. Далее, любое комплексное число

$$a+bi=(a+0i)+(0+bi)=a+b(0+i)$$

можно рассматривать как сумму вещественного числа a и произведения (вещественного) числа b на i; число a называется «вещественной частью», а bi — «мнимой частью» комплексного числа $a+bi$.

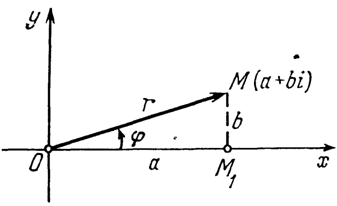


Рис. 2.

Число $\bar{\alpha}=a-bi$ называется *комплексно-сопряженным* к $\alpha=a+bi$. Легко видеть, что $\overline{\alpha+\beta}=\bar{\alpha}+\bar{\beta}$, $\overline{\alpha\beta}=\bar{\alpha}\bar{\beta}$ и что $\bar{\alpha}=\alpha$ *в том и только в том случае, когда* α *вещественно* (проверьте это!). Заметим, что сумма $\alpha+\bar{\alpha}=(a+bi)+(a-bi)=2a$ и произведение $\alpha\bar{\alpha}=(a+bi)(a-bi)=a^2+b^2$ комплексно-сопряженных чисел вещественны.

Комплексные числа удобно изображать точками, или, лучше — векторами плоскости. Комплексному числу $a+bi$ отвечает вектор $\overline{OM}$, где M — точка с координатами (a, b) в прямоугольной декартовой системе координат (рис. 2).

Пусть вектор $\overline{OM}$ представляет комплексное число $\alpha=a+bi$. Тогда угол $\varphi=xOM$ называется *аргументом* числа α (этот угол определен с точностью до кратного 2π), а $r = |\overline{OM}|$ — его *модулем* (заметим, что $r \geqslant 0$). Таким образом, φ и r — это просто полярные координаты точки M. Так как $a = r\cos\varphi$, $b = r\sin\varphi$, то

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

— это так называемая *тригонометрическая форма* комплексного числа. При этом, очевидно, $r = \sqrt{a^2 + b^2}$, $\cos\varphi = \frac{a}{r}$, $\sin\varphi = \frac{b}{r}$.

Вычислим произведение двух комплексных чисел, заданных в тригонометрической форме. Пусть $\alpha = r_1(\cos\varphi_1 + i\sin\varphi_1)$, $\beta = r_2(\cos\varphi_2 + i\sin\varphi_2)$. Тогда

$$\alpha\beta = r_1r_2[(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + \\ + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)] = \\ = r_1r_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)].$$

Таким образом, *при перемножении комплексных чисел их модули перемножаются, а аргументы складываются.* Отсюда легко выводится, что *при делении комплексных чисел их модули делятся, а аргументы вычитаются.*

Если $\alpha = a + bi = r(\cos\varphi + i\sin\varphi)$, то комплексно-сопряженное число

$$\bar{\alpha} = r(\cos\varphi - i\sin\varphi) = r[\cos(-\varphi) + i\sin(-\varphi)]$$

имеет тот же модуль r и противоположный аргумент $-\varphi$. Произведение $\alpha\bar{\alpha} = r^2$, и значит, модуль r комплексного числа α равен $\sqrt{\alpha\bar{\alpha}}$.

Заметим, что если модуль α равен 1 ($r = 1$), то $\alpha\bar{\alpha} = 1$, и значит, $\alpha^{-1} = \bar{\alpha}$.

Далее, при любом целом положительном n имеем

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

— это так называемая формула Муавра.

Рассмотрим еще операцию извлечения корня из комплексного числа. При этом мы ограничимся корнями из единицы. Общий случай читателю предлагается рассмотреть самостоятельно.

Мы имеем, очевидно, $1 = \cos 0 + i\sin 0$. Однако аргумент комплексного числа определен не однозначно, а с точностью до кратного 2π,— и здесь нам будет важно, что число 1 можно представить так:

$$1 = \cos 2\pi k + i \sin 2\pi k,$$

где k — любое целое число: $k = 0, \pm 1, \pm 2, \ldots$

Пусть $\sqrt[n]{1} = r\ (\cos\varphi + i\sin\varphi)$. Тогда, по формуле Муавра, $1 = r^n(\cos n\varphi + i\sin n\varphi)$ и, значит, $1 = r^n$ и $2\pi k = n\varphi$, откуда $r = 1$ (ибо модуль комплексного числа — вещественное положительное число) и $\varphi = \frac{2\pi k}{n}$, т. е.

$$\sqrt[n]{1} = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}, \quad k = 0, \pm 1, \pm 2, \ldots$$

Положим $\varepsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}$. При $k=0, 1, 2, \ldots, n-1$ получаем $\varepsilon_0 = 1$, $\varepsilon_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$, $\varepsilon_2 = \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n}, \ldots, \varepsilon_{n-1} = \cos\frac{2(n-1)\pi}{n} + i\sin\frac{2(n-1)\pi}{n}$. Если изображать комплексные числа точками плоскости (см. рис. 2 выше), то полученные при $k = 0, 1, 2, \ldots, n-1$ значения ε_k будут *расположены в вершинах правильного n-угольника, вписанного в единичную окружность; одна из вершин этого n-угольника находится в точке* $(1, 0)$.

При остальных значениях k мы не получим новых значений корня из 1, так как, например,

$$\varepsilon_n = \cos\frac{2\pi}{n}n + i\sin\frac{2\pi}{n}n = 1 = \varepsilon_0,$$

$$\varepsilon_{n+1} = \cos\frac{2\pi}{n}(n+1) + i\sin\frac{2\pi}{n}(n+1) = \varepsilon_1,$$

и т. д. — эти значения будут *периодически повторяться.* Аналогично,

$$\varepsilon_{-1} = \cos\left(-\frac{2\pi}{n}\right) + i\sin\left(-\frac{2\pi}{n}\right) = \varepsilon_{n-1}, \quad \varepsilon_{-2} = \varepsilon_{n-2},$$

и т. д. Таким образом, *корень n-й степени из* 1 *имеет ровно n различных значений.*

Произведение двух корней n-й степени из 1 *тоже есть корень n-й степени из* 1 (легко проверить, что $\varepsilon_p\varepsilon_q =$

$=\varepsilon_{p+q}$). *Любая* (целая) *степень корня n-й степени из* 1 *тоже будет корнем n-й степени из* 1.

Рассмотрим несколько примеров. При $n=2$ имеются 2 (изображенные на рис. 3) корня из 1:

$$\varepsilon_0 = 1 \quad \text{и} \quad \varepsilon_1 = \cos\frac{2\pi}{2} + i\sin\frac{2\pi}{2} = -1;$$

при $n=3$ — три корня

$$\varepsilon_0 = 1, \quad \varepsilon_1 = \frac{-1+i\sqrt{3}}{2}, \quad \varepsilon_2 = \frac{-1-i\sqrt{3}}{2},$$

расположенных в вершинах правильного треугольника.

При $n=4$ корни из 1 — это $\varepsilon_0 = 1$, $\varepsilon_1 = i$, $\varepsilon_2 = -1$, $\varepsilon_3 = -i$ (они расположены в вершинах квадрата). При $n=6$ корни из 1 образуют правильный шестиугольник с вершинами

$$\varepsilon_0 = 1, \quad \varepsilon_1 = \frac{1+i\sqrt{3}}{2}, \quad \varepsilon_2 = \frac{-1+i\sqrt{3}}{2},$$

$$\varepsilon_3 = -1, \quad \varepsilon_4 = \frac{-1-i\sqrt{3}}{2}, \quad \varepsilon_5 = \frac{1-i\sqrt{3}}{2}.$$

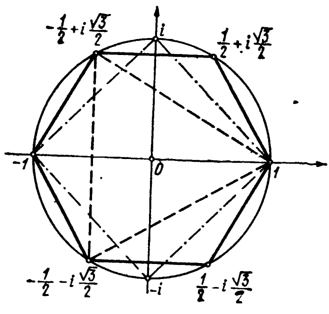


Рис. 3.

Пусть $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}$ — корни n-й степени из 1. Если возвести ε_1 в различные (целые, положительные) степени $k = 1, 2, 3, \ldots, n$, мы получим по одному разу все корни n-й степени из 1, так как, очевидно,

$$\varepsilon_1^1 = \varepsilon_1, \quad \varepsilon_1^2 = \varepsilon_2, \ldots, \varepsilon_1^n = 1 = \varepsilon_0.$$

Аналогичным свойством могут обладать и другие корни из 1. Так, при $n=4$, возводя в степени 1, 2, 3, 4 корень $\varepsilon_3 = -i$, мы получим $\varepsilon_3^1 = \varepsilon_3$, $\varepsilon_3^2 = -1 = \varepsilon_2$, $\varepsilon_3^3 = i = \varepsilon_1$, $\varepsilon_3^4 = 1 = \varepsilon_0$ — тоже все корни четвертой степени из 1. *Корень n-й степени из* 1, *при возведении которого в степени* $k = 1, 2, 3, \ldots, n$ *получаются по одному разу все* *корни n-степени из* 1, *называется* ***первообразным.*** Так, при $n = 4$ корни $\varepsilon_1 = i$ и $\varepsilon_3 = -i$ являются первообразными, корень же $\varepsilon_2 = -1$ первообразным не является, так как $\varepsilon_2^1 = \varepsilon_2$, $\varepsilon_2^2 = \varepsilon_0$, $\varepsilon_2^3 = \varepsilon_2$, $\varepsilon_2^4 = \varepsilon_0$.

§ 3. Определение векторного пространства

Мы начнем с примера, хорошо известного читателю. В геометрии важную роль играет понятие **вектора**, или направленного отрезка. Векторы можно складывать между собой и умножать на числа. Сумма $\overline{OC}$ векторов

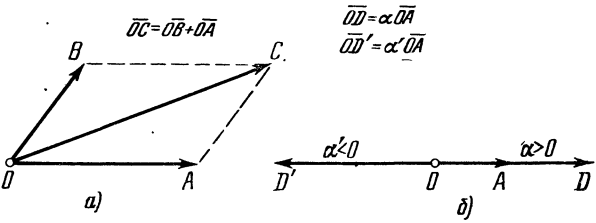


Рис. 4.

$\overline{OA}$ и $\overline{OB}$ определяется как диагональ параллелограмма $OACB$ (рис. 4, *а*; это определение можно распространить и на тот случай, когда прямые OA и OB совпадают), а произведение $\overline{OD}$ вектора $\overline{OA}$ на число α определяется из условий: $OD = |\alpha| \cdot OA$ и векторы $\overline{OD}$ и $\overline{OA}$ направлены в одну сторону, если $\alpha > 0$, и в противоположные стороны, если $\alpha < 0$ (рис. 4, *б*).

Но совокупность всех плоских или всех пространственных векторов — это только примеры (хотя и очень важные примеры) векторных пространств.

В главе I мы видели, что если имеются два **решения** $e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $e_2 = (\beta_1, \beta_2, \ldots, \beta_n)$ некоторой системы линейных однородных уравнений, то их сумма $e_1 + e_2 = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$ и произведение любого из них, например e_1 на произвольное число c (которое естественно считать принадлежащим тому же числовому полю F, что и коэффициенты уравнений): $ce_1 = (c\alpha_1, c\alpha_2, \ldots, c\alpha_n)$ тоже будут решениями той же системы. Аналогичная ситуация, когда имеется множество каких-то элементов, которые можно *склады-*

вать между собой и *умножать на числа,* получая в результате элементы того же самого множества, встречается в математике очень часто. Так, например, складывать между собой и умножать на числа можно многочлены от t с вещественными или комплексными коэффициентами — в результате получаются такие же многочлены. Если складываются и умножаются на числа многочлены, степени которых не превосходят данного числа n, то и полученные при этом многочлены будут степени не выше n. Складывать между собой и умножать на числа можно и произвольные функции от t — в результате снова получаются функции от t. Если функции, к которым применяются эти операции, непрерывны на каком-то отрезке $[a, b]$ (или, скажем, на всей числовой прямой), то и полученные в результате функции обладают тем же свойством.

Наконец, разумеется, и просто числа, образующие некоторое поле F, можно складывать между собой и умножать на числа из F; более того, вместо одного числа можно рассматривать пары, тройки и вообще упорядоченные наборы (строки), состоящие из n чисел:

$$(x_1, x_2, \ldots, x_n)$$

(такие строки выше служили решениями данной системы линейных уравнений, теперь же от них не требуется ничего!). Строки можно складывать:

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = \\ = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

и умножать на числа:

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n),$$

получая всякий раз такую же строку.

Все это — различные примеры векторных пространств (причем последний пример особенно важен для дальнейшего). Для того чтобы охватить все эти и другие возможные случаи, введем такое

Определение 1. *Множество R элементов $x, y, z, \ldots$ называется* ***векторным,*** *или* ***линейным, пространством,*** *если для любых двух его элементов x, y опреде-*

лена сумма $x+y \in R$ *) *и для каждого элемента* $x \in R$ *и каждого числа* α *(взятого из фиксированного числового поля* F*) определено произведение* $\alpha x \in R$*, причем выполнены следующие условия:*

1. $x+y=y+x$ *для всех* $x, y \in R$.
2. $(x+y)+z=x+(y+z)$ *для всех* $x, y, z \in R$.
3. *Существует такой* (нулевой) *элемент* $0 \in R$, что $x+0=x$ *для всех элементов* $x \in R$.
4. *Для каждого элемента* $x \in R$ *существует такой элемент* $-x$ (называемый противоположным к x), *что* $x+(-x)=0$.
5. $1 \cdot x = x$ *для всех* $x \in R$.
6. $\alpha(\beta x)=(\alpha\beta)x$ *для всех* $\alpha, \beta \in F$ *и* $x \in R$ **).
7. $(\alpha+\beta)x=\alpha x+\beta x$ *для всех* $\alpha, \beta \in F$ *и* $x \in R$.
8. $\alpha(x+y)=\alpha x+\alpha y$ *для всех* $\alpha \in F$ *и* $x, y \in R$.

Элементы векторного пространства называются ***векторами.***

Поле F во всем дальнейшем мы будем считать либо полем вещественных, либо полем комплексных чисел и, в соответствии с этим, будем говорить о *вещественном* или о *комплексном пространстве* R. Иногда же, не уточняя, о каком именно поле идет речь, мы будем говорить о векторном пространстве R *над полем* F.

Примеры. Можно говорить о векторном пространстве P_n многочленов степени не выше n с вещественными или комплексными коэффициентами, о векторном пространстве C функций, непрерывных на данном отрезке $[a, b]$, о векторном пространстве решений данной системы линейных однородных уравнений, наконец, просто о векторном пространстве строк, состоящих из n (вещественных или комплексных) чисел.

Из определения 1 непосредственно вытекают следующие

---

*) Символы $\in$, $\subseteq$, $\subset$ называются *знаками включения*. Запись $a \in A$ означает, что a является элементом множества A. Запись $A \subseteq B$ означает, что множество A является частью множества B (т. е. что каждый элемент a из A принадлежит также и B); запись $A \subset B$ означает, что множество A является правильной частью множества B, т. е. что A содержится в B, с ним не совпадая.

**) Чтобы не путать векторы с числами, мы в тех случаях, где может возникнуть недоразумение, условимся обозначать числа греческими, а векторы — латинскими буквами.

Простейшие свойства векторного пространства.

1. *Единственность нуля.* Предположим, что в пространстве R имеются два нулевых элемента, 0_1 и 0_2. Тогда, так как для любого x из R имеем $x+0_1=x$ и $x+0_2=x$, то, в частности, $0_2+0_1=0_2$ и $0_1+0_2=0_1$, откуда, ввиду $0_1+0_2=0_2+0_1$, получаем $0_1=0_2$.

2. *Единственность противоположного элемента.* Предположим, что у элемента x имеются два противоположных элемента, y и z; тогда $x+y=0$ и $x+z=0$. Следовательно, $y+x+z=y+(x+z)=y+0=y$ и $y+x+z=(y+x)+z=0+z=z$, откуда $y=z$.

3. *Для каждого элемента $x \in R$ произведение $0x=0$**). В самом деле, для каждого x имеем $0x=(0+0)x=0x+0x$. Прибавляя к левой и правой частям последнего равенства $-0x$, получим $0=0x$.

4. *Для любого $\alpha \in F$ и $0 \in R$ произведение $\alpha 0=0$.* Действительно, $\alpha 0=\alpha(0+0)=\alpha 0+\alpha 0$. Прибавляя к левой и правой частям равенства $-\alpha 0$, получим $0=\alpha 0$.

5. *Если произведение $\alpha x=0$, то либо $\alpha=0$, либо $x=0$.* В самом деле, пусть $\alpha \neq 0$; тогда

$$x=1\cdot x=\left(\frac{1}{\alpha}\alpha\right)x=\frac{1}{\alpha}(\alpha x)=\frac{1}{\alpha}0=0.$$

6. *Для каждого x элемент $(-1)x$ является противоположным к x.* Действительно, $x+(-1)x=1\cdot x+(-1)x=[1+(-1)]x=0\cdot x=0$, и значит,

$$(-1)x=-x.$$

Ввиду условия 2 определения 1, можно говорить о сумме трех $x+y+z=(x+y)+z$ (или, что то же самое, $x+(y+z)$) (и бо́льшего числа) элементов из R.

Разностью $x-y$ векторов x и y называется вектор $z=x+(-y)$.

§ 4. Размерность и базис

Определение 2. *Векторы $a_1, a_2, \ldots, a_k$ векторного пространства R называются* ***линейно зависимыми,*** *если существуют такие числа $\alpha_1, \alpha_2, \ldots, \alpha_k$, не равные*

*) Один и тот же символ 0 употреблен здесь и как число (слева) и как вектор (в правой части). Здесь и дальше из контекста всегда будет ясно, что означает символ 0 — число нуль или нулевой вектор.

одновременно нулю, что

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0.$$

Векторы, не являющиеся линейно зависимыми, называются ***линейно независимыми.***

Если векторы $a_1, a_2, \ldots, a_k$ линейно зависимы:

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0,$$

и, например, $\alpha_k \neq 0$, то

$$a_k = -\frac{\alpha_1}{\alpha_k} a_1 - \frac{\alpha_2}{\alpha_k} a_2 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} a_{k-1},$$

или

$$a_k = \xi_1 a_1 + \xi_2 a_2 + \ldots + \xi_{k-1} a_{k-1}, \tag{1}$$

где $\xi_i = -\frac{\alpha_i}{\alpha_k}$. Если имеет место равенство (1), то говорят, что вектор a_k является *линейной комбинацией* векторов $a_1, a_2, \ldots, a_{k-1}$, а также, что вектор a_k *линейно выражается* через $a_1, a_2, \ldots, a_{k-1}$. Таким образом, *если векторы $a_1, a_2, \ldots, a_k$ линейно зависимы, то по крайней мере один из них линейно выражается через остальные.* Ясно, что верно и обратное, т. е. что если один из векторов линейно выражается через остальные, то все эти векторы в совокупности линейно зависимы.

Примеры. На плоскости можно найти сколько угодно пар линейно независимых векторов — линейно независимы любые два неколлинеарных, т. е. не параллельных одной прямой, вектора. Но любые три вектора плоскости линейно зависимы.

В обычном (трехмерном) пространстве любые три некомпланарных (т. е. не параллельных одной плоскости) вектора a, b, c линейно независимы (так как если $\alpha a + \beta b + \gamma c = 0$ и, например $\gamma \neq 0$, то $c = -\frac{\alpha}{\gamma} a - \frac{\beta}{\gamma} b$, и вектор c компланарен векторам a, b). Однако любые четыре пространственных вектора a, b, c, d будут линейно зависимыми. (Докажите это.)

Определение 3. *Векторное пространство R называется* ***n-мерным,*** *если в нем можно найти n линейно независимых векторов, но больше чем n линейно независимых векторов оно не содержит.*

***Размерность** пространства — это максимальное число содержащихся в нем линейно независимых векторов.*

Так, размерность множества всех плоских векторов равна 2, размерность множества пространственных векторов равна 3; понятно, что размерность n-мерного пространства, по определению, равна n. Размерность пространства R условимся обозначать через $d(R)$.

Пространство, имеющее конечную размерность, называется конечномерным. Пространство, в котором можно найти сколь угодно много линейно независимых векторов, называется бесконечномерным. Примером бесконечномерного пространства может служить множество P всевозможных многочленов от t или множество C всех функций от t, непрерывных на данном отрезке $[a, b]$ (или непрерывных на всей числовой прямой), и т. д.

Определение 4. *Совокупность n линейно независимых векторов n-мерного векторного пространства R называется его **базисом**.*

Теорема 1. *Каждый вектор x линейного n-мерного пространства R можно представить, и притом единственным способом, в виде линейной комбинации векторов базиса.*

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ — произвольный базис n-мерного пространства R и $x \in R$. Так как каждые $n+1$ векторов (n-мерного!) пространства R линейно зависимы, то зависимы, в частности, и векторы $e_1, e_2, \ldots, e_n, x$, т. е. существуют такие не равные одновременно нулю числа $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha$, что

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n + \alpha x = 0.$$

При этом $\alpha \neq 0$, ибо если $\alpha = 0$, то хоть одно из чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$ было бы отлично от нуля, и векторы $e_1, e_2, \ldots, e_n$ были бы линейно зависимы. Следовательно,

$$x = -\frac{\alpha_1}{\alpha} e_1 - \frac{\alpha_2}{\alpha} e_2 - \ldots - \frac{\alpha_n}{\alpha} e_n.$$

Полагая $-\frac{\alpha_i}{\alpha} = x_i$, будем иметь

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Это представление x через $e_1, e_2, \ldots, e_n$ единственно, так как если $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$ и

$x = y_1e_1 + y_2e_2 + \ldots + y_ne_n$, то

$$(y_1 - x_1)e_1 + (y_2 - x_2)e_2 + \ldots + (y_n - x_n)e_n = 0$$

и ввиду линейной независимости векторов $e_1, e_2, \ldots, e_n$,

$$y_1 - x_1 = y_2 - x_2 = \ldots = y_n - x_n = 0,$$

откуда

$$y_1 = x_1, \quad y_2 = x_2, \quad \ldots, \quad y_n = x_n.$$

Числа $x_1, x_2, \ldots, x_n$ называются ***координатами*** *вектора x в базисе* $e_1, e_2, \ldots, e_n$. Таким образом, теорема 1 утверждает, что если задан базис n-мерного векторного пространства R, то каждый вектор из R имеет (единственным образом определенные) координаты в этом базисе. При этом ясно, что если координаты двух векторов x и y совпадают, то эти векторы одинаковы, так как тогда $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n = y$. Поэтому задавать вектор можно, просто указывая его координаты $x_1, x_2, \ldots, x_n$. При этом так и пишут: вектор $x = (x_1, x_2, \ldots, x_n)$.

Пусть мы имеем два вектора, заданные своими координатами в некотором базисе. Тогда *при сложении этих векторов их соответственные координаты складываются*: если

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n \text{ и } y = y_1e_1 + y_2e_2 + \ldots + y_ne_n,$$

то

$$x + y = (x_1 + y_1)e_1 + (x_2 + y_2)e_2 + \ldots + (x_n + y_n)e_n.$$

При умножении вектора на число все его координаты умножаются на это число: если

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$

то

$$\alpha x = (\alpha x_1)e_1 + (\alpha x_2)e_2 + \ldots + (\alpha x_n)e_n.$$

У нулевого вектора все координаты равны нулю, так как из равенства $\alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ne_n = 0$, ввиду линейной независимости векторов $e_1, e_2, \ldots, e_n$, вытекает, что $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Вектор, противоположный $x = (x_1, x_2, \ldots, x_n)$, равен, очевидно, $(-x_1, -x_2, \ldots, -x_n)$.

Теорема 2. *Если $e_1, e_2, \ldots, e_n$ — линейно независимые векторы пространства R и каждый вектор $x \in R$ линейно выражается через $e_1, e_2, \ldots, e_n$, то эти векторы образуют базис R.*

Доказательство. Векторы $e_1, e_2, \ldots, e_n$, по условию, линейно независимы. Остается доказать, что в пространстве R нет более чем n линейно независимых векторов. Возьмем произвольные $m > n$ векторов из R: $a_1, a_2, \ldots, a_m$. По условию, каждый из них можно линейно выразить через $e_1, e_2, \ldots, e_n$:

$$\begin{aligned} a_1 &= \alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n, \\ a_2 &= \alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_m &= \alpha_{1m}e_1 + \alpha_{2m}e_2 + \ldots + \alpha_{nm}e_n. \end{aligned}$$

Рассмотрим матрицу

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{bmatrix}.$$

Так как число строк этой матрицы равно n, то ее ранг не больше чем n, и значит, среди ее столбцов имеется не более чем n линейно независимых. Но так как $m > n$, то m столбцов этой матрицы между собой линейно зависимы. А это значит, что линейно зависимы и векторы $a_1, a_2, \ldots, a_m$. Мы нашли, что пространство R n-мерно и $e_1, e_2, \ldots, e_n$ — его базис.

Из теоремы 2 вытекает, что пространство R^n упорядоченных строк из n чисел n-мерно. Действительно, n строк $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, ..., $e_n = (0, 0, \ldots, 1)$ линейно независимы, так как из равенства

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = (\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, 0, \ldots, 0)$$

вытекало бы, что $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. С другой стороны, каждая строка $e = (\xi_1, \xi_2, \ldots, \xi_n)$ линейно выражается через $e_1, e_2, \ldots, e_n$:

$$e = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n.$$

Строки $e_1, e_2, \ldots, e_n$ образуют, следовательно, базис пространства R^n.

Пространство P_n многочленов степени не выше n имеет размерность $n + 1$. В самом деле, многочлены

$$1, t, t^2, \ldots, t^n$$

между собой линейно независимы, и каждый многочлен от t степени не выше n через них выражается очевидным образом.

Теорема 3. *В конечномерном векторном пространстве каждое множество линейно независимых векторов можно включить в некоторый базис.*

Доказательство. Пусть векторы $e_1, e_2, \ldots, e_k$ пространства R линейно независимы. Если каждый из остальных векторов из R линейно выражается через $e_1, e_2, \ldots, e_k$, то, по теореме 2, это уже базис. Если же найдется вектор e_{k+1}, линейно не выражающийся через $e_1, e_2, \ldots, e_k$, то $k+1$ векторов $e_1, e_2, \ldots, e_k, e_{k+1}$ линейно независимы. Действительно, если бы имело место равенство

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k + \alpha e_{k+1} = 0,$$

то $\alpha \neq 0$, ввиду линейной независимости векторов $e_1, e_2, \ldots, e_k$, и вектор e_{k+1} линейно выражался бы через $e_1, e_2, \ldots, e_k$.

Присоединим вектор e_{k+1} к $e_1, e_2, \ldots, e_k$. Если все векторы пространства R линейно выражаются через $e_1, e_2, \ldots, e_k, e_{k+1}$, то это уже базис. Если же найдется вектор e_{k+2}, не выражающийся линейно через $e_1, e_2, \ldots, e_k, e_{k+1}$, присоединим его к ним; новая система векторов $e_1, e_2, \ldots, e_k, e_{k+1}, e_{k+2}$ будет линейно независимой, и т. д.

Этот процесс не может продолжаться до бесконечности, так как пространство R, по условию, конечномерно, и, следовательно, в нем не может быть бесконечного множества $e_1, e_2, e_3, \ldots$ линейно независимых векторов. Поэтому, в конце концов, мы получим такую линейно независимую систему векторов $e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_n$, через которую уже будут линейно выражаться все остальные векторы из R. Ввиду теоремы 2 это и будет базис пространства R, содержащий заданные векторы $e_1, e_2, \ldots, e_k$.

§ 5. Изоморфизм векторных пространств

Пусть R — n-мерное векторное пространство и $e_1, e_2, \ldots, e_n$ — некоторый его базис. По теореме 1, каждый вектор $x \in R$ однозначно представляется в виде линейной комбинации

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

векторов $e_1, e_2, \ldots, e_n$.

Если вектору x поставить в соответствие строку $(x_1, x_2, \ldots, x_n)$, то, как мы видели в § 4, при сложении векторов соответствующие им строки тоже складываются, а при умножении вектора на число соответствующая ему строка умножается на то же число.

Таким образом, отправляясь от самого общего определения n-мерного векторного пространства, мы пришли к тому, что это пространство «устроено» в некотором смысле так же, как пространство всевозможных строк из n чисел. Значит, все n-мерные векторные пространства над одним и тем же полем F устроены одинаково; они, как принято говорить, изоморфны между собой. Точный смысл этого термина содержится в следующем определении.

Определение 5. *Векторные пространства R и R' над одним и тем же полем F* (в частности, два вещественных или два комплексных векторных пространства) *называются изоморфными, если между их элементами можно установить взаимно однозначное соответствие*) такое, что если $x \leftrightarrow x'$ (x соответствует x') и $y \leftrightarrow y'$, где $x, y \in R$, $x', y' \in R'$, то*

$$x + y \leftrightarrow x' + y'$$

и при любом $\alpha \in F$

$$\alpha x \leftrightarrow \alpha x'$$

(или, короче, $(x+y)' = x' + y'$ и $(\alpha x)' = \alpha x'$).

Из этого определения сразу видно, что два векторных пространства, изоморфных третьему, изоморфны между собой.

Имеет место следующая

Теорема 4. *Для того чтобы два векторных пространства (определенных над одним и тем же полем F)*

*) Говорят, что задано *взаимно однозначное отображение* множества M на множество N (в частности, M на M), если каждому элементу $a \in M$ поставлен в соответствие определенный элемент $b \in N$, причем каждый элемент $b \in N$ поставлен в соответствие одному определенному элементу $a \in M$. В этом случае говорят также, что между множествами M и N установлено *взаимно однозначное соответствие*.

были изоморфны, необходимо и достаточно, чтобы они имели одинаковые размерности.

Доказательство достаточности. Пусть даны два n-мерных векторных пространства R и R' над полем F. Выберем в каждом из них по базису:

$$e_1, e_2, \ldots, e_n \quad \text{в} \quad R \quad \text{и} \quad e'_1, e'_2, \ldots, e'_n \quad \text{в} \quad R'.$$

Вектору x, имеющему в базисе $e_1, e_2, \ldots, e_n$ координаты $x_1, x_2, \ldots, x_n$, поставим в соответствие вектор x' из R', имеющий те же самые координаты в базисе $e'_1, e'_2, \ldots, e'_n$. Тогда, поскольку при сложении векторов их соответственные координаты складываются, а при умножении на число — умножаются на то же число, будем иметь: если

$$x \leftrightarrow x', \ y \leftrightarrow y',$$

то

$$x + y \leftrightarrow x' + y'$$

и для любого $\alpha \in F$

$$\alpha x \leftrightarrow \alpha x'.$$

Следовательно, R' изоморфно R.

Доказательство необходимости. Для того чтобы доказать, что векторные пространства R и R' разных размерностей не изоморфны между собой, заметим прежде всего, что *при «изоморфном» соответствии между двумя пространствами нулевому вектору одного пространства соответствует нулевой вектор другого.* Действительно, пусть 0 — нулевой вектор из R и $0'$ — соответствующий ему вектор из R', x' — произвольный вектор из R' и $x \leftrightarrow x'$, где $x \in R$. Тогда, по определению,

$$0 + x \leftrightarrow 0' + x'.$$

Но $0+x=x$, а так как $x \leftrightarrow x'$ и соответствие между R и R' — взаимно однозначное, то

$$0' + x' = x',$$

т. е. $0'$ — нулевой вектор пространства R'.

Если пространства R и R' изоморфны и векторам $a_1, a_2, \ldots, a_k$ из R соответствуют векторы $a'_1, a'_2, \ldots, a'_k$ пространства R', то *из линейной зависимости векторов $a_1, a_2, \ldots, a_k$ вытекает, что и векторы $a'_1, a'_2, \ldots, a'_k$ тоже линейно* *зависимы, и обратно.* Действительно, пусть, например, $\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0$. Тогда вектору $\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k$ пространства R, равному $0 \in R$, в пространстве R' соответствует вектор $\alpha_1 a_1' + \alpha_2 a_2' + \ldots + \alpha_k a_k'$ и, значит,

$$\alpha_1 a_1' + \alpha_2 a_2' + \ldots + \alpha_k a_k' = 0'.$$

Следовательно максимальное число линейно независимых векторов в изоморфных пространствах должно быть одинаковым, а значит, размерности этих пространств — равные. (В частности, бесконечномерное пространство не изоморфно никакому пространству конечной размерности.)

В силу теоремы 4 *единственной характеристикой конечномерного векторного пространства, определенного над данным полем F, является его размерность.* По своей алгебраической структуре все n-мерные векторные пространства над полем F одинаковы. Можно, следовательно, сказать, что n-мерное векторное пространство — это пространство всевозможных строк из n чисел.

Поскольку мы уже условились, что основное поле F — фиксированное числовое поле, то n-мерное векторное пространство можно обозначать просто через R^n: одно обозначение для всех n-мерных векторных пространств над одним и тем же полем F законно, потому что все n-мерные векторные пространства над полем F одинаковы (изоморфны).

§ 6. Переход к новому базису

Пусть в пространстве R^n имеются два базиса:

$$e_1, e_2, \ldots, e_n \quad \text{и} \quad e_1', e_2', \ldots, e_n'.$$

Первый условимся называть старым базисом, второй — новым. Каждый из элементов нового базиса, по теореме 1, можно линейно выразить через векторы старого базиса

$$\begin{aligned} e_1' &= a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n, \\ e_2' &= a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ e_n' &= a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n. \end{aligned} \tag{2}$$

Можно сказать, что новые базисные векторы *получаются из старых с помощью матрицы*

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

(причем коэффициенты их разложений по старым базисным векторам образуют столбцы этой матрицы). Матрица A называется *матрицей перехода* от базиса e_1, e_2, ..., e_n к базису $e'_1, e'_2, \dots, e'_n$.

Определитель матрицы A не равен нулю, так как в противном случае ее столбцы, а следовательно, и векторы $e'_1, e'_2, \dots, e'_n$ были бы линейно зависимы.

Обратно, если определитель матрицы A отличен от нуля, то столбцы ее линейно независимы, и значит, векторы $e'_1, e'_2, \dots, e'_n$, получающиеся из базисных векторов e_1, e_2, ..., e_n с помощью матрицы A, линейно независимы, т. е. образуют некоторый базис. Значит, матрицей перехода может служить любая квадратная матрица порядка n с отличным от нуля определителем.

Посмотрим теперь, как связаны между собой координаты одного и того же вектора в старом и новом базисах. Пусть $x = x_1e_1 + x_2e_2 + \dots + x_ne_n$ — в старом базисе и в то же время $x = x'_1e'_1 + x'_2e'_2 + \dots + x'_ne'_n$ — в новом.

Подставляя в последнее равенство вместо $e'_1, e'_2, \dots, e'_n$ их выражения (2) через e_1, e_2, ..., e_n, получим $x = x'_1(a_{11}e_1 + a_{21}e_2 + \dots + a_{n1}e_n) + x'_2(a_{12}e_1 + a_{22}e_2 + \dots + a_{n2}e_n) + \dots + x'_n(a_{1n}e_1 + a_{2n}e_2 + \dots + a_{nn}e_n) = (a_{11}x'_1 + a_{12}x'_2 + \dots + a_{1n}x'_n)e_1 + (a_{21}x'_1 + a_{22}x'_2 + \dots + a_{2n}x'_n)e_2 + \dots + (a_{n1}x'_1 + a_{n2}x'_2 + \dots + a_{nn}x'_n)e_n$. Ввиду единственности разложения вектора x по базису e_1, e_2, ..., e_n отсюда следует, что

$$\begin{aligned} x_1 &= a_{11}x'_1 + a_{12}x'_2 + \dots + a_{1n}x'_n, \\ x_2 &= a_{21}x'_1 + a_{22}x'_2 + \dots + a_{2n}x'_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= a_{n1}x'_1 + a_{n2}x'_2 + \dots + a_{nn}x'_n. \end{aligned}$$

Таким образом, старые координаты вектора x получаются из новых его координат с помощью той же матрицы A, только коэффициенты соответствующих разложений образуют строки этой матрицы.

Пример. Пусть e_1, e_2 — единичные векторы, направленные по осям прямоугольной декартовой системы координат. Повернем оси координат на угол φ против часовой стрелки, и пусть e_1', e_2' — новые базисные векторы. Углы, образуемые вектором e_1' с векторами e_1 и

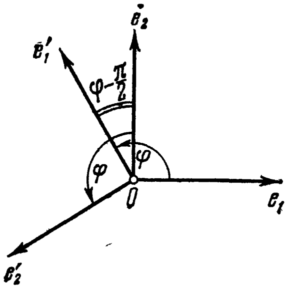


Рис. 5.

e_2, равны соответственно φ и $\varphi - \frac{\pi}{2}$ (рис. 5). Поэтому координаты этого вектора в базисе e_1, e_2 равны $\cos\varphi$ и $\cos\left(\varphi - \frac{\pi}{2}\right) = \sin\varphi$, значит, $e_1' = \cos\varphi \cdot e_1 + \sin\varphi \cdot e_2$. Аналогично, углы вектора e_2' с векторами e_1 и e_2 равны соответственно $\frac{\pi}{2} + \varphi$ и φ; координаты его в базисе e_1, e_2 равны $\cos\left(\frac{\pi}{2} + \varphi\right) = -\sin\varphi$ и $\cos\varphi$, и значит, $e_2' = -\sin\varphi \cdot e_1 + \cos\varphi \cdot e_2$.

Таким образом, матрицей перехода здесь будет

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix},$$

а выражения старых координат через новые имеют вид

$$x_1 = \cos\varphi \cdot x_1' - \sin\varphi \cdot x_2',$$
$$x_2 = \sin\varphi \cdot x_1' + \cos\varphi \cdot x_2',$$

§ 7. Подпространства векторного пространства

Определение 6. ***Подпространство** векторного пространства R — это множество R_1 его элементов, само являющееся векторным пространством относительно введенных в R операций сложения и умножения на число.*

Для того чтобы убедиться в том, что множество R_1 элементов векторного пространства R является его подпространством, необходимо проверить, что для любых двух векторов x и y из R_1 их сумма $x+y$ тоже принадлежит R_1 и что для каждого вектора x из R_1 и произвольного $\alpha \in F$ произведение αx тоже принадлежит R_1.

Покажем, что этого и достаточно. Действительно, аксиомы 1, 2 и 5—8 векторного пространства, справедливые в R, будут выполняться, в частности, и для элементов из R_1. Далее, если какой-то вектор $x \in R_1$, то и произведения $0 \cdot x = 0$ и $(-1)x = -x$ тоже принадлежат R_1. Следовательно, нулевой вектор принадлежит R_1 и для каждого x из R_1 *вектор* $-x$ тоже принадлежит R_1.

Размерность любого подпространства векторного пространства не превосходит размерности самого пространства: ведь линейно независимые векторы подпространства R_1 будут линейно независимыми и во всем пространстве, а значит, максимальное число линейно независимых векторов подпространства не превосходит размерности всего пространства.

Примеры. В обычном трехмерном пространстве (рассматриваемом как множество принадлежащих ему векторов) подпространствами будут все плоскости и все прямые, проходящие через начало координат. Подпространствами любого пространства будут само пространство R и множество, состоящее из одного нуля. В пространстве P_n многочленов степени не выше n подпространствами будут, например, все P_k при $k<n$ — ведь складывая и умножая на числа многочлены степени не выше k, мы будем получать снова такие же многочлены. С другой стороны, каждое из пространств P_n содержится в качестве подпространства в пространстве P всех многочленов с вещественными коэффициентами, а это последнее является подпространством пространства C непрерывных функций.

Рассмотрим систему линейных однородных уравнений, ранг матрицы коэффициентов которой равен r:

$$\left\{\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{aligned}\right. \tag{3}$$

и пусть в векторном пространстве R^n зафиксирован какой-то базис. Если каждое решение $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ системы (3) рассматривать как вектор пространства R^n, то из результатов § 10 главы I вытекает, что *совокупность всех решений системы* (3) *является k-мерным подпространством* (*где $k = n - r$*) *в R^n, базисом которого служит любая фундаментальная система решений*.

Покажем, что и, обратно, *каждое подпространство векторного пространства в любом базисе определяется некоторой системой линейных однородных уравнений.*

Действительно, пусть R_1 — k-мерное подпространство в R^n и $e_1', e_2', \ldots, e_k'$ — базис R_1. Дополним эту линейно независимую систему векторов R^n до базиса $e_1', e_2', \ldots, e_k', e_{k+1}', \ldots, e_n'$ всего пространства R^n. Легко видеть, что если $x_1', x_2', \ldots, x_n'$ — координаты относительно этого («старого») базиса, то подпространство R_1 в этом базисе определяется системой уравнений:

$$x_{k+1}' = 0, \quad x_{k+2}' = 0, \ \ldots, \ x_n' = 0.$$

Далее, если $x_1, x_2, \ldots, x_n$ — координаты относительно любого другого («нового») базиса $e_1, e_2, \ldots, e_n$, то, как показано в § 6, имеют место, в частности, равенства

$$\begin{aligned} x_{k+1}' &= a_{k+1,1}x_1 + a_{k+1,2}x_2 + \ldots + a_{k+1,n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n, \end{aligned}$$

где a_{ij} — некоторые числа, и значит, подпространство R_1 в базисе $e_1, e_2, \ldots, e_n$ определяется системой линейных однородных уравнений

$$\left\{\begin{aligned} &a_{k+1,1}x_1 + a_{k+1,2}x_2 + \ldots + a_{k+1,n}x_n = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0. \end{aligned}\right.$$

§ 8. Линейные многообразия

Пусть дано векторное пространство R^n, в котором выбран некоторый базис.

Рассмотрим (совместную) систему линейных, вообще говоря, неоднородных уравнений:

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{array}\right. \tag{4}$$

ранг матрицы коэффициентов которой равен r, и пусть $k = n - r$.

Определение 7. *Совокупность векторов пространства R^n, координаты которых удовлетворяют системе*

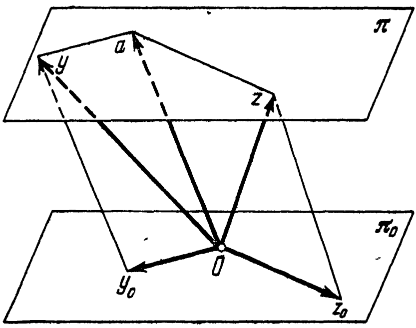


Рис. 6.

линейных уравнений (4), *называется линейным многообразием.*

Согласно замечанию, сделанному в конце § 10 главы I, общее решение x системы (4) равно сумме общего решения x_0 соответствующей (т. е. с теми же коэффициентами при неизвестных) однородной системы (3) и произвольного, но фиксированного решения $a = (a_1, a_2, \ldots, a_n)$ системы (4). Таким образом, *линейное многообразие решений системы* (4) *получается, если к каждому вектору из подпространства решений соответствующей однородной системы* (3) *прибавить один и тот же вектор a* (см. рис. 6, где концы векторов, образующих линейное многообразие, принадлежат плоскости π, получающейся из подпространства π_0 параллельным переносом на вектор a).

Покажем, что и, обратно, *если к каждому вектору подпространства $R_1 \subset R^n$ прибавить один и тот же вектор a, то получится линейное многообразие.* Пусть подпространство R_1 определяется системой линейных однородных уравнений (3) и $a = (a_1, a_2, \ldots, a_n)$. Положим

$$a_{i1}a_1 + a_{i2}a_2 + \ldots + a_{in}a_n = b_i, \qquad i = 1, 2, \ldots, n, \quad (5)$$

и рассмотрим систему уравнений (4). Ввиду условий (5), вектор a является одним из решений этой (вообще говоря, неоднородной) системы. Следовательно, линейное многообразие, определяемое системой (4), совпадает с заданным множеством $R_1 + a$ векторов.

Линейное многообразие (4) называется *k-мерным,* если k-мерно соответствующее ему подпространство (3).

§ 9. Пересечение и сумма подпространств

Определение 8. Пусть в векторном пространстве R имеются два подпространства R_1 *и* R_2. Их *пересечением* $R_3 = R_1 \cap R_2$ *называется множество всевозможных векторов из R, принадлежащих одновременно и R_1, и R_2.*

Легко видеть, что *пересечение двух подпространств R_1 и R_2 является подпространством* (содержащимся и в R_1, и в R_2).

Определение 9. Если R_1 и R_2 — подпространства линейного пространства R, то их *суммой* $R_4 = R_1 + R_2$ *называется множество всех векторов вида $u + v$, где $u \in R_1$ и $v \in R_2$.*

Сумма двух подпространств является подпространством (возможно, совпадающим с R). Действительно, если $x, y \in R_4$, то $x = u_1 + v_1$, $y = u_2 + v_2$, где $u_1, u_2 \in R_1$ и $v_1, v_2 \in R_2$, и тогда $x + y = (u_1 + u_2) + (v_1 + v_2)$, где $u_1 + u_2 \in R_1$ и $v_1 + v_2 \in R_2$, поэтому $x + y \in R_4$.

Далее, если $\alpha \in F$, то $\alpha x = \alpha u_1 + \alpha v_1$, где $\alpha u_1 \in R_1$, $\alpha v_1 \in R_2$ и, следовательно, $\alpha x \in R_4$.

Подпространство R_1 (так же, как и R_2) *содержится в R_4,* ибо каждый элемент $x \in R_1$ можно представить в виде суммы $x + 0$, где $x \in R_1$, а $0 \in R_2$.

Теорема 5. *Если* R_1 *и* R_2 — *подпространства векторного пространства* R *и* $R_3 = R_1 \cap R_2$, *а* $R_4 = R_1 + R_2$, *то*

$$d(R_1) + d(R_2) = d(R_3) + d(R_4). \qquad (6)$$

Доказательство. В подпространстве R_3 выберем какой-нибудь базис

$$e_1,\ e_2,\ e_3,\ \ldots,\ e_k. \qquad (7)$$

Дополним множество (7) векторов, принадлежащих одновременно и R_1, и R_2, до базиса R_1:

$$e_1,\ e_2,\ \ldots,\ e_k,\ f_{k+1},\ \ldots,\ f_p, \qquad (8)$$

с одной стороны, и до базиса R_2:

$$e_1,\ e_2,\ \ldots,\ e_k,\ g_{k+1},\ \ldots,\ g_q \qquad (9)$$

— с другой (теорема 3). Покажем, что векторы

$$e_1,\ e_2,\ \ldots,\ e_k,\ f_{k+1},\ \ldots,\ f_p,\ g_{k+1},\ \ldots,\ g_q \qquad (10)$$

линейно независимы. Тогда, по теореме 2, они образуют базис в R_4, ибо если вектор $z \in R_4$, то $z = x + y$, где $x \in R_1$, $y \in R_2$, и значит, x линейно выражается через векторы (8), а y — через векторы (9). Но тогда вектор z линейно выражается через векторы (10).

Допустим, что векторы (10) линейно зависимы:

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k + \beta_{k+1} f_{k+1} + \ldots + \beta_p f_p + \\ + \gamma_{k+1} g_{k+1} + \ldots + \gamma_q g_q = 0. \qquad (11)$$

Тогда вектор

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k + \beta_{k+1} f_{k+1} + \ldots + \beta_p f_p,$$

равный $-(\gamma_{k+1} g_{k+1} + \ldots + \gamma_q g_q)$, принадлежит одновременно и R_1, и R_2, а значит, и их пересечению R_3. Но в таком случае он должен линейно выражаться через базисные векторы (7) подпространства R_3; пусть

$$a = \sigma_1 e_1 + \sigma_2 e_2 + \ldots + \sigma_k e_k.$$

Отсюда, ввиду единственности разложения вектора a по базису пространства R_1,

$$\sigma_i = \alpha_i,\ i = 1,\ 2,\ \ldots,\ k,\ \text{и}\ \beta_{k+1} = \beta_{k+2} = \ldots = \beta_p = 0.$$

а тогда из равенства (11) следует, что

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k+\gamma_{k+1} g_{k+1}+\ldots+\gamma_q g_q=0,$$

и, ввиду линейной независимости векторов (9),

$$\alpha_i=0,\ i=1,\ 2,\ \ldots,\ k, \text{ и } \gamma_{k+1}=\gamma_{k+2}=\ldots=\gamma_q=0.$$

Таким образом, векторы (10) образуют базис пространства R_4, и значит, его размерность равна числу этих векторов:

$$k+(p-k)+(q-k)=p+q-k.$$

Но $d(R_1)=p$, $d(R_2)=q$ и $d(R_3)=k$. Мы доказали, что *сумма размерностей двух подпространств равна размерности их суммы, сложенной с размерностью пересечения.*

Так, в четырехмерном пространстве R^4 два двумерных подпространства R_1 и R_2 могут пересекаться по нулевому вектору, и тогда их сумма совпадает со всем пространством; в этом случае равенство (6) превращается в $2+2=0+4$. Они могут пересекаться по прямой (одномерному подпространству), и тогда их сумма трехмерна; это соответствует равенству $2+2=1+3$. Наконец, R_1 и R_2 могут совпадать, тогда их пересечение и сумма тоже двумерны, и равенство (6) дает $2+2=2+2$.

Два трехмерных подпространства в R^4 либо пересекаются по плоскости (двумерному подпространству), и тогда $3+3=2+4$, либо совпадают: $3+3=3+3$. (Другие случаи здесь невозможны, так как сумма этих подпространств не более чем четырехмерна).

Если R_1 — двумерное, а R_2 — трехмерное подпространства в R^4, то либо они пересекаются по прямой: $2+3=1+4$, либо R_1 содержится в R_2: $2+3=2+3$.

Определение 10. *Если пространство R является суммой своих подпространств R_1 и R_2, пересечение R_3 которых состоит лишь из нулевого вектора, то говорят, что R_4 есть* **прямая сумма** *подпространств R_1 и R_2, и пишут*

$$R=R_1\oplus R_2.$$

Если $R=R_1\oplus R_2$, то очевидно, что

$$d(R)=d(R_1)+d(R_2).$$

Так, обычное трехмерное пространство R^3 будет, очевидно, прямой суммой любой (проходящей через начало координат) плоскости π и любой не лежащей в этой плоскости (но проходящей через начало) прямой l. Пространство R^3 распадается также и на сумму любых двух своих несовпадающих (проходящих через начало) плоскостей, но эта сумма не будет прямой.

Теорема 6. *Если $R = R_1 \oplus R_2$, то каждый вектор из R единственным способом представляется в виде $u + v$, где $u \in R_1$, $v \in R_2$.*

Доказательство. Каждый вектор из R, по определению суммы подпространств, представляется в виде $u + v$, где $u \in R_1$, $v \in R_2$. Предположим, что какой-то вектор x из R разложен в такую сумму двумя способами:

$$x = u_1 + v_1 = u_2 + v_2.$$

Тогда вектор $u_1 - u_2 = v_2 - v_1$ принадлежит одновременно и R_1, и R_2, т. е. он принадлежит R_3 и, значит, равен нулю, откуда $u_1 = u_2$ и $v_2 = v_1$.

Пусть R — какое-то векторное пространство и $a_1, a_2, \ldots, a_k \in R$. Совокупность всевозможных линейных комбинаций этих векторов

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k$$

является, очевидно, подпространством в R. Мы будем говорить, что это подпространство *порождается векторами* $a_1, a_2, \ldots, a_k$. Его называют также *линейной оболочкой* векторов $a_1, a_2, \ldots, a_k$. Нетрудно видеть, что *линейная оболочка векторов $a_1, a_2, \ldots, a_k$ совпадает с пересечением всех подпространств, содержащих эти векторы.*

§ 10. Определение аффинного пространства

Выше мы неоднократно иллюстрировали общее понятие векторного пространства на примере (обычной) плоскости или (обычного трехмерного) пространства. Однако эти иллюстрации были, строго говоря, не совсем точными — ведь основным понятием той геометрии, которая изучается в средней школе, является *точка*, и все геометрические образы можно понимать как *множества точек*, в то время как в определении векторного пространства точки вообще не фигурируют.

В школьном курсе геометрии понятие *вектора* появляется позже понятия точки: вектором там называют упорядоченную пару точек (направленный отрезок) $\overline{AB}$, определяя далее условия равенства векторов и правила их сложения и умножения на число.

Нам сейчас придется поступить иначе. Располагая уже определением векторного пространства, мы дополним его, введя в рассмотрение еще и *точки*. Полученное таким образом множество (векторов и точек) — его называют точечно-векторным, или аффинным, пространством,— будет уже ближе к тому пространству, которое изучается в курсе элементарной геометрии, хотя и не будет еще полностью с ним совпадать. Дело в том, что само понятие «аффинного» пространства предполагает, что это пространство лишено *метрики*, т. е. способа измерения длин и углов. Оно станет вполне идентичным (во всяком случае для двух- и трехмерного случаев) обычному пространству лишь после введения в нем соответствующей метрики (см. ниже, главу IV).

Определение 11. *Пусть имеются векторное пространство* R (элементы его по-прежнему обозначаются строчными латинскими буквами) *и, кроме того, множество элементов, которые мы будем называть* ***точками*** (и обозначать прописными латинскими буквами), *причем каждой упорядоченной паре точек* M, N *поставлен в соответствие один и только один вектор* x *из* R (хотя разным парам точек может быть поставлен в соответствие один и тот же вектор); мы будем писать в этом случае $\overline{MN}=x$. Будем предполагать, что *это соответствие между точками и векторами обладает следующими свойствами:*

1. *Для каждой точки* M *и каждого вектора* x *найдется одна и только одна такая точка* N, *что* $\overline{MN}=x$.

2. *Для любых трех точек* M, N, P

$$\overline{MN}+\overline{NP}=\overline{MP}.$$

Все точки и все векторы вместе образуют тогда ***аффинное пространство.***

Аффинное пространство называется *n-мерным*, если n-мерно соответствующее ему векторное пространство R.

Итак, аффинное пространство A — это множество элементов двух родов: точек и векторов, связь между которыми задается с помощью операции *откладывания векторов*. Произвольный вектор x можно отложить от любой точки M, получив при этом определенную точку N, и тогда $\overline{MN}=x$. Точка M *называется* *началом*, а точка N — концом вектора $\overline{MN}$.

Далее, почти очевидны следующие предложения.

1. *Если* $\overline{MN}=\overline{QP}$, *то* $\overline{MQ}=\overline{NP}$. (Это вытекает из равенства $\overline{MN}+\overline{NP}=\overline{MP}=\overline{MQ}+\overline{QP}$.)

В частности, так как $\overline{MN}+\overline{NN}=\overline{MN}=\overline{MM}+\overline{MN}$, то $\overline{MM}=\overline{NN}$, т. е. *все векторы, у которых начало и конец совпадают, равны между собой. Такой вектор* $\overline{NN}$ *является нулевым*, так как $\overline{MN}+\overline{NN}=\overline{MN}$.

2. *Вектор* $\overline{NM}$ *является противоположным* $\overline{MN}$, так как $\overline{MN}+\overline{NM}=\overline{MM}=0$.

§ 11. Введение координат в аффинном пространстве

В n-мерном аффинном пространстве A координаты точек можно ввести следующим образом. Выберем какую-нибудь точку O в качестве *начала координат*. Тогда для каждого вектора x, ввиду условия 1 определения 11, найдется и притом только одна такая точка X, что $\overline{OX}=x$. Так будет установлено взаимно однозначное соответствие между всеми точками и всеми векторами из A: точке X ставится в соответствие вектор $x=\overline{OX}$, концом которого она является (в обычном трехмерном пространстве это — откладывание всех векторов пространства от начала координат).

Далее, в соответствующем A векторном пространстве R выберем какой-нибудь базис $e_1, e_2, \ldots, e_n$. Тогда каждый вектор x из A будет определяться строкой своих координат: $x=(x_1, x_2, \ldots, x_n)$. Эти же координаты мы отнесем и соответствующей вектору x точке X; будем писать в этом случае: $X(x_1, x_2, \ldots, x_n)$.

Таким образом, если в n-мерном аффинном пространстве A выбрана система координат (т. е. точка O как начало отсчета и базис $e_1, e_2, \ldots, e_n$ в соответствующем A векторном пространстве R), то каждой точке из A будет однозначно сопоставлена строка из n чисел — ее координат. У точки O все координаты будут равны нулю, так как ей соответствует, очевидно, нулевой вектор 0.

Если $X(x_1, x_2, \ldots, x_n)$ и $Y(y_1, y_2, \ldots, y_n)$ — две точки аффинного пространства A, то, ввиду равенства, $\overline{OX}+\overline{XY}=\overline{OY}$ имеем $\overline{XY}=\overline{OY}-\overline{OX}$, т. е. *координа-*

ты вектора $\overline{XY}$ *равны разностям* $y_i - x_i$ *координат его конца и начала.*

Можно показать, что, подобно векторным пространствам, все аффинные пространства одной и той же размерности тоже «устроены одинаково», так что (если зафиксировано основное поле F) размерность аффинного пространства является его единственной характеристикой. Поэтому n-мерное аффинное пространство мы можем обозначать далее просто через A^n.

§ 12. Переход к новой системе координат

Посмотрим, как преобразуются координаты точки аффинного пространства A^n при переходе к новой системе координат.

Пусть сперва изменяется только начало координат. Предположим, что новое начало помещено в точку O', координаты которой в старой системе $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Для любой точки X из A^n имеем

$$\overline{OO'} + \overline{O'X} = \overline{OX}. \tag{12}$$

Координаты вектора $\overline{OX} = (x_1, x_2, \ldots, x_n)$ — это координаты точки X в старой системе координат; координаты вектора $\overline{O'X} = (x_1', x_2', \ldots, x_n')$ — координаты точки X в новой системе; координаты вектора OO' — это координаты точки O' в старой системе, т. е. $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Из равенства (12) получаем $\overline{O'X} = \overline{OX} - \overline{OO'}$, или, в координатах,

$$(x_1', x_2', \ldots, x_n') = (x_1, x_2, \ldots, x_n) - (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

и $x_i' = x_i - \alpha_i$, $i = 1, 2, \ldots, n$, *т. е. новые координаты точки получаются, если из старых ее координат вычесть координаты нового начала в старой системе координат.*

Пусть теперь начало координат не меняется, но в векторном пространстве R^n, соответствующем A^n, выбирается новый базис с матрицей перехода

$$\begin{bmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} \ldots a_{nn} \end{bmatrix}, \tag{13}$$

т. е. старый базис, образованный векторами $e_1, e_2, \ldots, e_n$, заменяется новым, составленным из векторов $e'_1, e'_2, \ldots, e'_n$, где

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n, \quad i = 1, 2, \ldots, n.$$

Так как координаты точки X — это, по определению, координаты вектора $x = \overline{OX}$, то, как следует из § 6, старые координаты точки будут выражаться через новые ее координаты по формулам

$$x_i = a_{i1}x'_1 + a_{i2}x'_2 + \ldots + a_{in}x'_n, \qquad i = 1, 2, \ldots, n.$$

В общем случае, когда и начало координат O переносится в точку $O'(\alpha_1, \alpha_2, \ldots, \alpha_n)$, и базис $e_1, e_2, \ldots, e_n$ с помощью матрицы перехода (13) заменяется новым, старые координаты $x_1, x_2, \ldots, x_n$ произвольной точки X и новые ее координаты $x'_1, x'_2, \ldots, x'_n$ связаны соотношениями

$$x_i = a_{i1}x'_1 + a_{i2}x'_2 + \ldots + a_{in}x'_n + \alpha_i, \qquad i = 1, 2, \ldots, n.$$

§ 13. k-мерные плоскости в аффинном пространстве

Пусть в n-мерном аффинном пространстве A^n установлена система координат. Рассмотрим снова (совместную) систему уравнений (4), ранг матрицы коэффициентов которой равен r, и пусть $k = n - r$.

Определение 12. *Множество всех точек из A^n, координаты которых удовлетворяют системе уравнений* (4), *называется **k-мерной плоскостью**; одномерные плоскости называются также **прямыми**, а $(n-1)$-мерные плоскости — **гиперплоскостями**.*

Понятно, что каждую гиперплоскость (для которой $r = 1$) можно задать всего одним линейным уравнением

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b.$$

В обычном трехмерном пространстве гиперплоскости — это обычные плоскости, а на обычной плоскости — это прямые.

Можно показать, что при переходе к новой системе координат в A^n точки, удовлетворяющие системе уравнений (4), будут удовлетворять некоторой новой системе уравнений, ранг матрицы коэффициентов которой тоже равен r.

Пусть π будет k-мерная плоскость, определяемая системой уравнений (4). Соответствующая система (3) линейных однородных уравнений определяет некоторую k-мерную плоскость π_0, «проходящую через начало координат». Если все векторы отложены от начала координат, то те векторы, концы которых принадлежат π_0, образуют подпространство, а векторы, концы которых принадлежат π, образуют k-мерное линейное многообразие. Это многообразие получается, если ко всем векторам подпространства π_0 прибавить один и тот же вектор a. Можно сказать поэтому, что *k-мерная плоскость π получается из π_0 параллельным переносом на вектор a.* Это позволяет дать следующее

Определение 13. *Две k-мерные плоскости параллельны, если определяющие их системы таковы, что соответствующие однородные системы равносильны* (имеют одни и те же решения). *k-мерная плоскость π_1 и l-мерная плоскость π_2 параллельны* (при $l>k$), *если π_1 параллельна какой-нибудь k-мерной плоскости, содержащейся в π_2* (в этом случае определяющие π_1 и π_2 системы таковы, что однородная система, соответствующая π_2, является следствием однородной системы, соответствующей π_1).

Пусть снова π будет k-мерная плоскость, определяемая системой уравнений (4). *Общее решение* системы (4) в векторной форме имеет вид

$$x=\alpha_1 c_1+\alpha_2 c_2+\ldots+\alpha_k c_k+a; \tag{14}$$

где

$$x_0=\alpha_1 c_1+\alpha_2 c_2+\ldots+\alpha_k c_k$$

— общее решение соответствующей однородной системы (3) и a — некоторый фиксированный вектор (одно из решений системы (4)). Если $c_i=(c_{i1}, c_{i2}, \ldots, c_{in})$ при $i=1, 2, \ldots, k$ и $a=(a_1, a_2, \ldots, a_n)$, то записывая равенство (14) в координатах, получим *параметрические* *уравнения k-мерной плоскости:*

$$x_j=\alpha_1 c_{1j}+\alpha_2 c_{2j}+\ldots+\alpha_k c_{kj}+a_j, \qquad j=1, 2, \ldots, n.$$

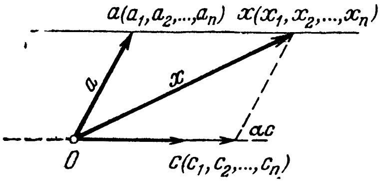


Рис. 7.

Если ранг матрицы коэффициентов системы (4) равен $n-1$, соответствующую (одномерную) плоскость выше мы назвали *прямой*. В этом случае общее решение системы (4) в векторной форме имеет вид

$$x=\alpha c+a, \qquad (15)$$

где $x_0=\alpha c$ — общее решение соответствующей однородной системы и $a=(a_1, a_2, \ldots, a_n)$ — некоторый фиксированный вектор (рис 7). Если $c=(c_1, c_2, \ldots, c_n)$ и $x=(x_1, x_2, \ldots, x_n)$, то, записывая уравнение (15) в координатах, получим *параметрические уравнения прямой*:

$$\begin{aligned} x_1 &= \alpha c_1 + a_1, \\ x_2 &= \alpha c_2 + a_2, \\ &\ldots\ldots \\ x_n &= \alpha c_n + a_n, \end{aligned}$$

которые, исключая параметр α, можно переписать в виде

$$\frac{x_1-a_1}{c_1}=\frac{x_2-a_2}{c_2}=\ldots=\frac{x_n-a_n}{c_n}.$$

(Это — *канонические уравнения прямой.* Они имеют смысл и в том случае, если некоторые из знаменателей обращаются в нуль — тогда равны нулю и соответствующие числители). Если в пространстве A^n даны две точки $A(a_1, a_2, \ldots, a_n)$ и $B(b_1, b_2, \ldots, b_n)$, то проходящая через них *прямая AB* определяется, очевидно, уравнениями

$$\frac{x_1-a_1}{b_1-a_1}=\frac{x_2-a_2}{b_2-a_2}=\ldots=\frac{x_n-a_n}{b_n-a_n}. \qquad (16)$$

Обозначив равные отношения (16) через β, получим *параметрические уравнения прямой AB*:

$$x_i = a_i + \beta(b_i - a_i),$$

или

$$x_i = (1-\beta)a_i + \beta b_i, \qquad i = 1, 2, \ldots, n$$

Полагая $1-\beta=\alpha$, будем иметь

$$x_i = \alpha a_i + \beta b_i, \text{ где } \alpha+\beta=1. \tag{17}$$

Если α и β — вещественные, причем $0 \leqslant \alpha \leqslant 1$ и, значит $0 \leqslant \beta \leqslant 1$, то говорят, что соответствующая точка x припринадлежит *отрезку AB*.

Если ранг матрицы коэффициентов системы (4) равен $n-2$, то определяемая ею плоскость двумерна, и общее решение системы в векторной форме имеет вид

$$x = \alpha_1 c_1 + \alpha_2 c_2 + a,$$

где $x_0 = \alpha_1 c_1 + \alpha_2 c_2$ — общее решение соответствующей однородной системы (3), а $a = (a_1, a_2, \ldots, a_n)$ — некоторый фиксированный вектор. Если $c_i = (c_{i1}, c_{i2}, \ldots, c_{in})$, $i = 1, 2$, то координаты точек $X(x_1, x_2, \ldots, x_n)$ этой плоскости определяются формулами

$$x_k = \alpha_1 c_{1k} + \alpha_2 c_{2k} + a_k, \qquad k = 1, 2, \ldots, n$$

(*параметрические уравнения двумерной плоскости*).

Пусть в аффинном пространстве A^n даны две плоскости: k-мерная плоскость π, определяемая системой уравнений (4), и l-мерная плоскость π', определяемая системой уравнений

$$\left\{\begin{array}{l} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n = d_1, \\ c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n = d_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_{s1}x_1 + c_{s2}x_2 + \ldots + c_{sn}x_n = d_s. \end{array}\right. \tag{18}$$

Тогда их *пересечение* (т. е. множество точек, принадлежащих одновременно и π, и π') будет определяться системой, состоящей из всех уравнений системы (4) и всех уравнений системы (18), и значит тоже будет некоторой плоскостью (которая, в частности, может состоять из одной точки или даже вообще не содержать ни одной точки, если объединенная система окажется несовместной). Легко видеть, что *каждая k-мерная плоскость в A^n является пересечением некоторых $r=n-k$ гиперплоскостей.*

§ 14. Выпуклые множества в аффинном пространстве

Определение 14. *Множество точек вещественного аффинного пространства называется выпуклым, если вместе с каждыми двумя своими точками A и B оно содержит и все точки отрезка AB.*

Легко видеть, что *пересечение любого числа выпуклых множеств выпукло.*

Определение 15. *Множество точек аффинного пространства называется ограниченным, если координаты всех его точек в некоторой системе координат в совокупности ограничены* (легко видеть, что тогда они будут ограничены и во всех системах координат).

Пусть в (вещественном) аффинном пространстве A^n задана гиперплоскость

$$a_1x_1+a_2x_2+\ldots+a_nx_n=b. \tag{19}$$

Этой гиперплоскостью все точки из A^n разбиваются на два *полупространства*: A_1 — множество точек, для которых $a_1x_1+a_2x_2+\ldots+a_nx_n\geqslant b$, и A_2 — множество точек, для которых $a_1x_1+a_2x_2+\ldots+a_nx_n\leqslant b$. Полупространства A_1 и A_2 пересекаются по самой гиперплоскости (19).

Теорема 7. *Каждое полупространство аффинного пространства A^n является выпуклым множеством.*

Доказательство. Пусть точки $P(p_1, p_2, \ldots, p_n)$ и $Q=(q_1, q_2, \ldots, q_n)$ из A^n принадлежат, например, полупространству A_1; тогда

$$a_1p_1+a_2p_2+\ldots+a_np_n\geqslant b, \quad a_1q_1+a_2q_2+\ldots+a_nq_n\geqslant b.$$

Если $X(x_1, x_2, \ldots, x_n)$ — произвольная точка отрезка PQ, то по формулам (17) $x_i=\alpha p_i+\beta q_i$, $i=1, 2, \ldots, n$, где $\alpha, \beta\geqslant 0$ и $\alpha+\beta=1$. Для этой точки X имеем

$$\begin{aligned}a_1x_1&+a_2x_2+\ldots+a_nx_n=\\&=a_1(\alpha p_1+\beta q_1)+a_2(\alpha p_2+\beta q_2)+\ldots+a_n(\alpha p_n+\beta q_n)=\\&=\alpha(a_1p_1+a_2p_2+\ldots+a_np_n)+\beta(a_1q_1+a_2q_2+\ldots+a_nq_n)\geqslant\\&\geqslant\alpha b+\beta b=(\alpha+\beta)b=b,\end{aligned}$$

т. е. произвольная точка X отрезка PQ принадлежит A_1.

Гиперплоскость (19), как пересечение выпуклых множеств A_1 и A_2, *является выпуклым множеством. Каждая k-мерная плоскость в A^n*, как пересечение нескольких гиперплоскостей, *выпукла.*

Пусть в A^n даны m полупространств, определяемых неравенствами

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \geqslant b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \geqslant b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \geqslant b_m. \end{cases} \qquad (20)$$

(Все знаки неравенств здесь одного смысла — этого всегда можно достичь, умножая в случае необходимости обе части неравенства на -1.) Пересечение этих полупространств, называемое *выпуклой многогранной областью*, определяет *множество решений системы линейных неравенств* (20). Если это пересечение ограничено, оно называется (*выпуклым*) *многогранником* n-мерного пространства A^n.

ГЛАВА III

ЛИНЕЙНЫЕ ОПЕРАТОРЫ

§ 1. Определение и примеры

Определение 1. Говорят, что *в векторном пространстве R задан* ***оператор, или преобразование,*** $\mathcal{A}$*), *если каждому вектору $x \in R$ поставлен в соответствие определенный вектор $\mathcal{A}(x)$* или; как мы чаще будем писать, $\mathcal{A}x \in R$.

Оператор (преобразование) A называется ***линейным,*** *если для любых двух векторов x и y из R и произвольного числа $\alpha \in F$*

1) $\mathcal{A}(x+y) = \mathcal{A}x + \mathcal{A}y$,
2) $\mathcal{A}(\alpha x) = \alpha\mathcal{A}x$.

Вектор Ax называется образом вектора x, а вектор x — прообразом вектора $\mathcal{A}x$ при преобразовании $\mathcal{A}$.

Выберем в пространстве R базис $e_1, e_2, \ldots, e_n$. Тогда если $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, то в силу линейности оператора $\mathcal{A}$ имеем

$$\mathcal{A}e_i = x_1\mathcal{A}e_1 + x_2\mathcal{A}e_2 + \ldots + x_n\mathcal{A}e_n, \ i = 1, 2, \ldots, n.$$

Но так как $\mathcal{A}e_i$ (где $i = 1, 2, 3, \ldots, n$) — это тоже вектор из R, то $\mathcal{A}e_i$ можно разложить по базису $e_1, e_2, \ldots$ $\ldots, e_n$; пусть

$$\mathcal{A}e_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n, \ i = 1, 2, \ldots, n.$$

Тогда

$$\mathcal{A}x = x_1(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n) + \\ + x_2(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n) + \ldots \\ \ldots + x_n(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n) =$$

*) На плоскости и в обычном трехмерном пространстве чаще говорят о *преобразовании*. Мы будем пользоваться и тем, и другим терминами, но чаще — первым из них (*оператор*).

$$= (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)e_1 + (a_{21}x_1 + a_{22}x_2 + \ldots$$
$$\ldots + a_{2n}x_n)e_2 + \ldots + (a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n)e_n.$$

Если $x'_1, x'_2, \ldots, x'_n$ — координаты вектора $\mathcal{A}x$ в том же базисе $e_1, e_2, \ldots, e_n$, т. е. если

$$\mathcal{A}x = x'_1e_1 + x'_2e_2 + \ldots + x'_ne_n,$$

то, ввиду единственности разложения вектора по базису, имеем

$$\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x'_n &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n. \end{aligned} \tag{1}$$

Таким образом, каждому линейному оператору $\mathcal{A}$ в данном базисе $e_1, e_2, \ldots, e_n$ отвечает матрица

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

i-й столбец которой образован коэффициентами разложения вектора $\mathcal{A}e_i$ по базису $e_1, e_2, \ldots, e_n$; при этом коэффициенты разложений (1) координат вектора $\mathcal{A}x$ по координатам вектора x образуют строки матрицы A.

Если в n-мерном векторном пространстве R задан базис, то не только каждому линейному оператору $\mathcal{A}$ отвечает определенная матрица A, но и, обратно, *каждая квадратная матрица A порядка n может рассматриваться как матрица некоторого линейного оператора.* Действительно, пусть $e_1, e_2, \ldots, e_n$ — базис пространства R, и пусть дана матрица (2) n-го порядка (мы будем писать короче: $A = [a_{ik}]$). Обозначим через $\mathcal{A}$ оператор, переводящий произвольный вектор $x = x_1e_1 + x_2e_2 + \ldots$ $\ldots + x_ne_n$ в вектор $\mathcal{A}x = x'_1e_1 + x'_2e_2 + \ldots + x'_ne_n$, где $x'_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$ при $i = 1, 2, \ldots, n$.

Покажем, что этот оператор — линейный. В самом деле, произвольный другой вектор $y = y_1e_1 + y_2e_2 + \ldots$ $\ldots + y_ne_n$ он переводит в $\mathcal{A}y = y'_1e_1 + y'_2e_2 + \ldots + y'_ne_n$, где $y'_i = a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n$, вектор $x + y =$

$=(x_1+y_1)e_1+(x_2+y_2)e_2+\ldots+(x_n+y_n)e_n$ — в вектор $\mathcal{A}(x+y)=z_1e_1+z_2e_2+\ldots+z_ne_n$, где $z_i=$ $=a_{i1}(x_1+y_1)+a_{i2}(x_2+y_2)+\ldots+a_{in}(x_n+y_n)=$ $=x_i'+y_i'$. Поэтому

$$\mathcal{A}(x+y)=\mathcal{A}x+\mathcal{A}y.$$

Далее, для любого $\alpha\in F$ имеем $\alpha x=(\alpha x_1)e_1+$ $+(\alpha x_2)e_2+\ldots+(\alpha x_n)e_n$ и $\mathcal{A}(\alpha x)=t_1e_1+t_2e_2+\ldots$ $\ldots+t_ne_n$, где $t_i=a_{i1}(\alpha x_1)+a_{i2}(\alpha x_2)+\ldots+a_{in}(\alpha x_n)=\alpha x_i'$. Следовательно,

$$\mathcal{A}(\alpha x)=\alpha\mathcal{A}x,$$

и оператор $\mathcal{A}$ — линейный.

Таким образом, *если в векторном пространстве R^n задан базис, то каждому линейному оператору отвечает определенная квадратная матрица порядка n и, обратно, каждой такой матрице отвечает определенный линейный оператор*. Поэтому линейный оператор и соответствующую ему (в данном базисе) матрицу мы будем обозначать одной и той же буквой: $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, ... — линейные операторы, A, B, C, ... — соответствующие им матрицы. Матрица A называется *матрицей линейного оператора* $\mathcal{A}$.

Легко видеть, что для всякого линейного оператора $\mathcal{A}$

$$\mathcal{A}0=0.$$

При этом, если $\mathcal{A}x=0$ *только при* $x=0$, *то оператор называется невырожденным; если же найдется такой вектор $x\neq 0$, что $\mathcal{A}x=0$, то оператор $\mathcal{A}$ — вырожденный.*

Пусть $A=[a_{ik}]$ — матрица линейного оператора $\mathcal{A}$. Рассмотрим систему линейных однородных уравнений

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

Ввиду теоремы 10 из главы I, для существования ненулевого решения этой системы (и значит, для существования ненулевого вектора $x=x_1e_1+x_2e_2+\ldots+x_ne_n$ такого, что $\mathcal{A}x=0$) необходимо и достаточно, чтобы определитель матрицы A (обозначим его через $|A|$) был равен нулю. Следовательно, *для того чтобы оператор $\mathcal{A}$ был невырожденным, необходимо и достаточно, чтобы оп-*

ределитель матрицы A этого оператора (в любом базисе) *был отличен от нуля.* Матрица, определитель которой отличен от нуля, называется *невырожденной матрицей.*

Рассмотрим несколько примеров.

1. Пусть $\mathscr{A}$ — *поворот* всех векторов обычной *плоскости xOy* (короче — поворот плоскости xOy) вокруг начала координат на угол φ против часовой стрелки. Это преобразование линейно, ибо безразлично, сначала ли сложить векторы a и b, а потом повернуть их сумму на угол φ, или сначала повернуть векторы, а потом их сложить (рис. 8); так же безразлично, умножить ли сначала

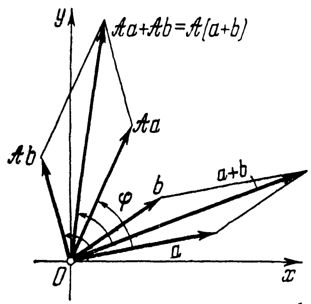


Рис. 8.

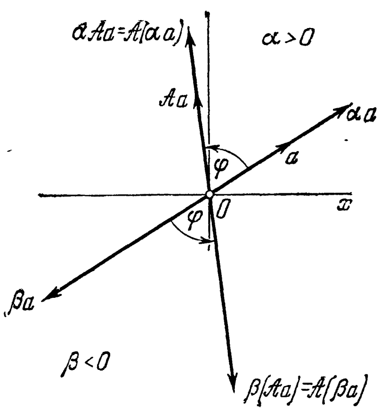


Рис. 9.

вектор a на число α, а затем повернуть его на угол φ или сделать это в обратном порядке (рис. 9).

Предположим, что базисные векторы — единичные и взаимно ортогональные. Вектор $\mathscr{A}e$ — единичный вектор, образующий угол φ с e_1 и угол $\varphi - \frac{\pi}{2}$ с e_2.

Следовательно,

$$\mathscr{A}e_1 = \cos\varphi \cdot e_1 + \sin\varphi \cdot e_2.$$

Единичный вектор $\mathscr{A}e_2$ образует с e_1 угол $\frac{\pi}{2} + \varphi$, а с e_2 — угол φ.

Следовательно,

$$\mathscr{A}e_2 = -\sin\varphi \cdot e_1 + \cos\varphi \cdot e_2.$$

Таким образом,

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$

2. Пусть $\mathscr{A}$ — *поворот* обычного *трехмерного пространства* на угол φ вокруг оси Oz. Если e_1, e_2, e_3 — единичные векторы прямоугольной декартовой системы координат, то

$$\mathscr{A}e_1 = \cos\varphi \cdot e_1 + \sin\varphi \cdot e_2,$$
$$\mathscr{A}e_2 = -\sin\varphi \cdot e_1 + \cos\varphi \cdot e_2,$$
$$\mathscr{A}e_3 = e_3,$$

и значит, матрица этого преобразования

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

3. В обычном трехмерном пространстве пусть $\mathscr{A}a$ будет *ортогональной проекцией* вектора *a на плоскость xOy*. Линейность этого преобразования вытекает из того, что проекция суммы векторов равна сумме проекций слагаемых и что проекция произведения вектора на число равна произведению проекции вектора на это число. Если базис выбран так, как в примере 2, то очевидно, что

$$\mathscr{A}e_1 = e_1, \quad \mathscr{A}e_2 = e_2, \quad \mathscr{A}e_3 = 0,$$

и следовательно,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

4. В обычном трехмерном пространстве пусть $\mathscr{A}a$ будет вектор, *симметричный с вектором a относительно плоскости xOy*. Линейность этого преобразования очевидна. При этом

$$\mathscr{A}e_1 = e_1, \quad \mathscr{A}e_2 = e_2, \quad \mathscr{A}e_3 = -e_3,$$

и матрица преобразования имеет вид:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

5. В пространстве P_n многочленов от t степени не выше n положим

$$\mathscr{A}(x(t)) = x'(t).$$

Линейность этого «оператора дифференцирования» вытекает из основных правил дифференциального исчисления. Чтобы найти его матрицу, выберем в качестве базиса, например, векторы

$$e_0 = 1,\ e_1 = t,\ e_2 = \frac{t^2}{2!},\ \ldots,\ e_n = \frac{t^n}{n!}.$$

Тогда

$$\mathscr{A}e_0 = 0, \quad \mathscr{A}e_1 = e_0, \quad \mathscr{A}e_2 = e_1,\ \ldots,\ \mathscr{A}e_n = e_{n-1}$$

и

$$A = \begin{bmatrix} 0 & 1 & 0 \ldots 0 \\ 0 & 0 & 1 \ldots 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ 0 & 0 & 0 \ldots 1 \\ 0 & 0 & 0 \ldots 0 \end{bmatrix}.$$

6. Обозначим через $\mathscr{E}$ так называемый **тождественный оператор**, определяемый равенством: $\mathscr{E}x = x$ для любого $x \in R$. Тогда $\mathscr{E}e_i = e_i$ для всех $i = 1, 2, \ldots, n$, и следовательно, матрица оператора $\mathscr{E}$ в *любом* базисе имеет вид

$$E = \begin{bmatrix} 1 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 \ldots 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ 0 & 0 & 0 \ldots 1 \end{bmatrix}.$$

7. Обозначим через $\mathscr{O}$ так называемый **нулевой оператор**, определяемый равенством $\mathscr{O}x = 0$ для всех $x \in R$. Матрица этого оператора состоит из одних нулей.

Ясно, что операторы 1, 2, 4 и 6 — невырожденные, а операторы 3, 5 и 7 — вырожденные.

Теорема 1. *При линейном преобразовании векторного пространства каждое подпространство переходит в подпространство.*

Доказательство. Пусть R_1 — подпространство векторного пространства R^n. Обозначим через $\mathcal{A}R_1$ множество всех векторов, являющихся образами векторов из R_1 при линейном преобразовании $\mathcal{A}$. Нам надо доказать, что $\mathcal{A}R_1$ — подпространство. Пусть векторы x и y принадлежат $\mathcal{A}R_1$. Это значит, что $x = \mathcal{A}x'$ и $y = \mathcal{A}y'$, где $x' \in R_1$ и $y' \in R_1$. Но тогда

$$x + y = \mathcal{A}x' + \mathcal{A}y' = \mathcal{A}(x' + y') \in \mathcal{A}R_1,$$

так как $x' + y' \in R_1$, и при любом $\alpha \in F$

$$\alpha x = \alpha\mathcal{A}x' = \mathcal{A}(\alpha x') \in \mathcal{A}R_1,$$

так как $\alpha x' \in R_1$. Таким образом, $\mathcal{A}R_1$ — подпространство.

(Легко понять, что размерность $\mathcal{A}R_1$ не превышает размерности R_1.)

Теорема 2. *При линейном преобразовании векторного пространства каждое линейное многообразие переходит в линейное многообразие.*

Доказательство. Пусть M — линейное многообразие в R^n. Тогда существует такое подпространство R_1 и такой вектор a, что $M = R_1 + a$ (см. выше стр. 78). Если $\mathcal{A}$ — линейный оператор, то $\mathcal{A}M = \mathcal{A}R_1 + \mathcal{A}a$. Ввиду теоремы 1, $\mathcal{A}R_1$ является линейным подпространством и, значит, AM — линейное многообразие (см. стр. 79).

Пусть A^n — n-мерное аффинное пространство и R^n — соответствующее ему векторное пространство, в котором задан линейный оператор $\mathcal{A}$. Этот оператор можно следующим образом распространить и на точки из A^n. Предположим, что в A^n выбрана система координат. Тогда, если вектор $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$ при преобразовании $\mathcal{A}$ переходит в $\mathcal{A}x = x_1'e_1 + x_2'e_2 + \ldots + x_n'e_n$, то, по определению, точка $X\ (x_1, x_2, \ldots, x_n)$ (конец вектора $\overline{OX} = x$) переходит в $X'\ (x_1', x_2', \ldots, x_n')$ (конец вектора $\overline{OX'} = \mathcal{A}x$).

Из теоремы 2 непосредственно вытекает, что *при линейном преобразовании аффинного пространства k-мерная плоскость переходит в плоскость* (не большей размерности). В частности, прямые переходят в прямые или в точки.

§ 2. Действия над линейными операторами

А. Сложение линейных операторов.

Если $\mathscr{A}$ и $\mathscr{B}$ — *два линейных оператора в векторном пространстве* R, *то их* *суммой* $\mathscr{A}+\mathscr{B}$ *называется оператор* $\mathscr{C}$, *определяемый равенством*

$$\mathscr{C}x=\mathscr{A}x+\mathscr{B}x$$

для любого $x \in R$.

Легко видеть, что сумма линейных операторов тоже будет линейным оператором. Если линейные операторы $\mathscr{A}$ и $\mathscr{B}$ имеют (в некотором базисе) соответственно матрицы $A=[a_{ik}]$ и $B=[b_{ik}]$, то матрицей оператора $\mathscr{C}=\mathscr{A}+\mathscr{B}$ будет $C=[c_{ik}]$, где $c_{ik}=a_{ik}+b_{ik}$. Матрица C называется *суммой матриц* A и B. Таким образом, по определению,

$$[a_{ik}]+[b_{ik}]=[a_{ik}+b_{ik}].$$

(Разумеется, складывать можно лишь матрицы одного и того же порядка.)

Сложение линейных операторов (и сложение матриц) обладает, очевидно, следующими свойствами:

1. $\mathscr{A}+\mathscr{B}=\mathscr{B}+\mathscr{A}$.
2. $(\mathscr{A}+\mathscr{B})+\mathscr{C}=\mathscr{A}+(\mathscr{B}+\mathscr{C})$.
3. $\mathscr{A}+\mathscr{O}=\mathscr{A}$ для любого $\mathscr{A}$.
4. Если через $-\mathscr{A}$ обозначить оператор, определяемый тем, что $(-\mathscr{A})x=-\mathscr{A}x$ для всех $x \in R$, то $-\mathscr{A}$ будет линейным оператором и

$$(-\mathscr{A})+\mathscr{A}=\mathscr{O}.$$

Матрицу оператора $-\mathscr{A}$ обозначим через $-A$; тогда ясно, что если $A=[a_{ik}]$, то $-A=[-a_{ik}]$.

Б. Умножение линейного оператора на число.

Если $\mathscr{A}$ — *линейный оператор в пространстве* R *и* $\alpha \in F$, *то произведением* $\mathscr{A}$ *на* α *называется оператор* $\alpha\mathscr{A}$, *определяемый следующим образом:*

$$(\alpha\mathscr{A})x=\alpha(\mathscr{A}x)$$

для каждого вектора x *из* R.

Ясно, что $\alpha\mathscr{A}$ — тоже линейный оператор и что его матрица αA получается из матрицы A оператора $\mathscr{A}$ умножением каждого ее элемента на α:

$$\alpha[a_{ik}]=[\alpha a_{ik}].$$

Матрица αA называется *произведением матрицы A на число α.*

Для умножения линейного оператора на число справедливы, очевидно, следующие тождества:

1. $1\cdot\mathscr{A}=\mathscr{A}$;
$0\cdot\mathscr{A}=\mathscr{O}$;
$(-1)\mathscr{A}=-\mathscr{A}$
2. $\alpha(\beta A)=(\alpha\beta)\mathscr{A}$.
3. $(\alpha+\beta)\mathscr{A}=\alpha\mathscr{A}+\beta\mathscr{A}$.
4. $\alpha(\mathscr{A}+\mathscr{B})=\alpha\mathscr{A}+\alpha\mathscr{B}$.

Аналогичные тождества справедливы и для умножения матрицы на число.

В. Умножение линейных операторов.

***Произведением** $\mathscr{A}\mathscr{B}$ операторов $\mathscr{A}$ и $\mathscr{B}$ называется оператор $\mathscr{C}$, определяемый следующим образом:*

$$\mathscr{C}x=\mathscr{A}(\mathscr{B}x)$$

для каждого вектора x из R.

Таким образом, перемножение операторов состоит в *последовательном их применении* одного за другим; при этом сначала производится преобразование $\mathscr{B}$, а затем уже полученный вектор $\mathscr{B}x$ подвергается преобразованию $\mathscr{A}$.

Так, если $\mathscr{A}$ есть поворот плоскости против часовой стрелки на угол φ, а $\mathscr{B}$ — поворот (в том же направлении) на угол ψ, то $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$ будет поворотом на угол $\varphi+\psi$. Если $\mathscr{A}$ — симметрия плоскости xOy относительно оси Ox, а $\mathscr{B}$ — симметрия относительно Oy, то $\mathscr{A}^2=\mathscr{B}^2=\mathscr{E}$ — тождественное преобразование, а $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$ — симметрия относительно начала координат. Если $\mathscr{A}$ — ортогональное проектирование обычного трехмерного пространства на плоскость или на прямую, то $\mathscr{A}^2=\mathscr{A}$. Если $\mathscr{A}$ — дифференцирование в пространстве многочленов, то оператор $\mathscr{A}^2$ — это взятие второй производной.

Произведение линейных операторов тоже будет линейным оператором.

Действительно,

$$1)\quad (\mathscr{A}\mathscr{B})(x+y)=\mathscr{A}(\mathscr{B}(x+y))=\mathscr{A}(\mathscr{B}x+\mathscr{B}y)=$$
$$=\mathscr{A}(\mathscr{B}x)+\mathscr{A}(\mathscr{B}y)=(\mathscr{A}\mathscr{B})x+(\mathscr{A}\mathscr{B})y,$$

$$2)\ (\mathcal{A}\mathcal{B})(\alpha x) = \mathcal{A}(\mathcal{B}(\alpha x)) = \mathcal{A}(\alpha\mathcal{B}x) = \\ = \alpha\mathcal{A}(\mathcal{B}x) = \alpha(\mathcal{A}\mathcal{B})x.$$

Найдем, как выражается матрица C линейного оператора $\mathcal{C} = \mathcal{A}\mathcal{B}$ через матрицы $A = [a_{ik}]$ и $B = [b_{ik}]$ линейных операторов $\mathcal{A}$ и $\mathcal{B}$. Имеем

$$\mathcal{C}e_k = \mathcal{A}(\mathcal{B}e_k) = \mathcal{A}(b_{1k}e_1 + b_{2k}e_2 + \ldots + b_{nk}e_n) =$$
$$= b_{1k}\mathcal{A}e_1 + b_{2k}\mathcal{A}e_2 + \ldots + b_{nk}\mathcal{A}e_n = b_{1k}(a_{11}e_1 + a_{21}e_2 + \ldots$$
$$\ldots + a_{n1}e_n) + b_{2k}(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n) + \ldots$$
$$\ldots + b_{nk}(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n) =$$
$$= (a_{11}b_{1k} + a_{12}b_{2k} + \ldots + a_{1n}b_{nk})e_1 +$$
$$+ (a_{21}b_{1k} + a_{22}b_{2k} + \ldots + a_{2n}b_{nk})e_2 + \ldots$$
$$\ldots + (a_{n1}b_{1k} + a_{n2}b_{2k} + \ldots + a_{nn}b_{nk})e_n;$$

значит, если $\mathcal{C}e_k = c_{1k}e_1 + c_{2k}e_2 + \ldots + c_{nk}e_n$, то

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{in}b_{nk},$$

где $i, k = 1, 2, \ldots, n$.

Мы видим, что для того *чтобы получить элемент матрицы C, стоящий в пересечении ее i-й строки и k-го столбца, надо каждый элемент i-й строки матрицы A умножить на соответствующий элемент k-го столбца матрицы B и все полученные произведения сложить.* (Говорят и короче: элемент c_{ik} равен произведению «i-й строки матрицы A на k-й столбец матрицы B».) Матрица C называется *произведением матриц* A и B.

Пример.

$$\begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2\cdot4+1\cdot1 & 2\cdot3+1\cdot2 \\ 3\cdot4+5\cdot1 & 3\cdot3+5\cdot2 \end{bmatrix} = \begin{bmatrix} 9 & 8 \\ 17 & 19 \end{bmatrix}.$$

Произведение тех же матриц в обратном порядке равно

$$\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 4\cdot2+3\cdot3 & 4\cdot1+3\cdot5 \\ 1\cdot2+2\cdot3 & 1\cdot1+2\cdot5 \end{bmatrix} = \begin{bmatrix} 17 & 19 \\ 8 & 11 \end{bmatrix}.$$

Мы видим, что *умножение матриц* (вообще говоря) *не коммутативно*.

Рассмотрим свойства умножения линейных операторов и умножения матриц.

1. *Если $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$ — линейные операторы, то*

$$(\mathscr{A}\mathscr{B})\mathscr{C} = \mathscr{A}(\mathscr{B}\mathscr{C}).$$

Действительно, для любого вектора $x \in R$ имеем

$$[(\mathscr{A}\mathscr{B})\mathscr{C}]x = (\mathscr{A}\mathscr{B})(\mathscr{C}x) = \mathscr{A}(\mathscr{B}(\mathscr{C}x))$$

и

$$[\mathscr{A}(\mathscr{B}\mathscr{C})]x = \mathscr{A}[(\mathscr{B}\mathscr{C})x] = \mathscr{A}(\mathscr{B}(\mathscr{C}x));$$

таким образом, *умножение линейных операторов* (а, следовательно, и матриц) *ассоциативно.*

Произведение $(\mathscr{A}\mathscr{B})\mathscr{C} = \mathscr{A}(\mathscr{B}\mathscr{C})$ линейных операторов, состоящее в последовательном их выполнении: сначала $\mathscr{C}$, затем $\mathscr{B}$ и, наконец, $\mathscr{A}$,— обозначается обычно просто через $\mathscr{A}\mathscr{B}\mathscr{C}$ — без скобок.

2. *Для любого линейного оператора* $\mathscr{A}$

$$\mathscr{A}\mathscr{E} = \mathscr{E}\mathscr{A} = \mathscr{A}.$$

Матрица E тождественного оператора $\mathscr{E}$ (см. выше, стр. 97) называется *единичной матрицей.* Для любой матрицы A (того же порядка, что и E)

$$AE = EA = A.$$

3. *Умножение и сложение линейных операторов связаны дистрибутивными законами:*

$$(\mathscr{A} + \mathscr{B})\mathscr{C} = \mathscr{A}\mathscr{C} + \mathscr{B}\mathscr{C} \quad \text{и} \quad \mathscr{C}(\mathscr{A} + \mathscr{B}) = \mathscr{C}\mathscr{A} + \mathscr{C}\mathscr{B},$$

так как для любого вектора $x \in R$

$$((\mathscr{A} + \mathscr{B})\mathscr{C})x = (\mathscr{A} + \mathscr{B})(\mathscr{C}x) = \mathscr{A}(\mathscr{C}x) + \mathscr{B}(\mathscr{C}x) = \\ = (\mathscr{A}\mathscr{C})x + (\mathscr{B}\mathscr{C})x = (\mathscr{A}\mathscr{C} + \mathscr{B}\mathscr{C})x$$

и

$$(\mathscr{C}(\mathscr{A} + \mathscr{B}))x = \mathscr{C}((\mathscr{A} + \mathscr{B})x) = \mathscr{C}(\mathscr{A}x + \mathscr{B}x) = \\ = \mathscr{C}(\mathscr{A}x) + \mathscr{C}(\mathscr{B}x) = (\mathscr{C}\mathscr{A})x + (\mathscr{C}\mathscr{B})x = (\mathscr{C}\mathscr{A} + \mathscr{C}\mathscr{B})x.$$

Аналогичные тождества справедливы и для матриц.

Вспомним теперь основные законы сложения и умножения чисел, сформулированные на стр. 55—56 (аксиомы поля). Для сложения и умножения матриц мы доказали справедливость всех этих законов, кроме пятого и восьмого. Пример на стр. 101 показывает, что умножение матриц, а значит, и умножение линейных операторов, вообще говоря, н е к о м м у т а т и в н о.

Что же касается существования линейного оператора, обратного к данному, то справедливо следующее предложение:

Для каждого *невырожденного* *линейного оператора* $\mathscr{A}$ *существует такой — обратный к* $\mathscr{A}$ *линейный оператор* $\mathscr{A}^{-1}$, *что*

$$\mathscr{A}\mathscr{A}^{-1} = \mathscr{A}^{-1}\mathscr{A} = \mathscr{E}$$

(и соответственно *для каждой матрицы* A, *определитель которой отличен от нуля, существует такая обратная к* A *матрица* A^{-1}, *что* $AA^{-1} = A^{-1}A = E$).

Докажем это. Пусть $\mathscr{A}$ — невырожденный линейный оператор, имеющий в некотором базисе $e_1, e_2, \ldots, e_n$ матрицу A. Мы докажем сначала существование обратной к A матрицы, т. е. такой матрицы, A^{-1}, что

$$AA^{-1} = A^{-1}A = E;$$

тогда линейный оператор $\mathscr{A}^{-1}$, имеющий в том же базисе $e_1, e_2, \ldots, e_n$ матрицу A^{-1}, будет обратным к $\mathscr{A}$: ведь последовательное применение операторов $\mathscr{A}^{-1}$ и $\mathscr{A}$ одного за другим будет линейным оператором с единичной матрицей, т. е. тождественным оператором.

Итак, пусть дана матрица $A = [a_{ij}]$, определитель которой отличен от нуля. Рассмотрим матрицу, составленную из алгебраических дополнений соответствующих элементов матрицы A:

$$\begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{bmatrix}.$$

При транспонировании ее получается матрица $\widetilde{A}$, называемая присоединенной к матрице A:

$$\widetilde{A} = \begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}.$$

Перемножая данную матрицу A и матрицу $\widetilde{A}$, получим

$$A\widetilde{A} = \widetilde{A}A = \begin{bmatrix} |A| & 0 & \ldots & 0 \\ 0 & |A| & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & |A| \end{bmatrix}$$

(теоремы 3 и 4 главы I). А следовательно, матрица

$$A^{-1} = \frac{\tilde{A}}{|A|}$$

будет обратной к A.

Пример. *Найти матрицу, обратную к*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Решение. Определитель матрицы A равен 4. Алгебраические дополнения ее элементов: $A_{11} = 2$, $A_{12} = -2$, $A_{13} = 2$, $A_{21} = -2$, $A_{22} = 4$, $A_{23} = -2$, $A_{31} = -8$, $A_{32} = 10$, $A_{33} = -4$ и, значит,

$$A^{-1} = \frac{1}{4}\begin{bmatrix} 2 & -2 & -8 \\ -2 & 4 & 10 \\ 2 & -2 & -4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 & -4 \\ -1 & 2 & 5 \\ 1 & -1 & -2 \end{bmatrix}.$$

Заметим, что *если операторы $\mathcal{A}$ и $\mathcal{B}$ невырожденные, то таким же будет и их произведение* (так как из равенства $(\mathcal{A}\mathcal{B})x = \mathcal{A}(\mathcal{B}x) = 0$ вытекает, что $\mathcal{B}x = 0$ и, значит, $x = 0$), причем

$$(\mathcal{A}\mathcal{B})^{-1} = \mathcal{B}^{-1}\mathcal{A}^{-1}$$

(а для матриц $(AB)^{-1} = B^{-1}A^{-1}$), так как

$$(\mathcal{A}\mathcal{B})(\mathcal{B}^{-1}\mathcal{A}^{-1}) = \mathcal{A}[\mathcal{B}(\mathcal{B}^{-1}\mathcal{A}^{-1})] = \mathcal{A}[(\mathcal{B}\mathcal{B}^{-1})\mathcal{A}^{-1}] = \\ = \mathcal{A}(\mathcal{E}\mathcal{A}^{-1}) = \mathcal{A}\mathcal{A}^{-1} = \mathcal{E}.$$

Теорема 3. *Определитель произведения двух матриц равен произведению определителей сомножителей:* если $AB = C$, то

$$|C| = |A|\,|B|.$$

Доказательство. Пусть

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

Тогда, как известно,

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}.$$

Определитель матрицы C равен

$$|C| = \begin{vmatrix} \sum\limits_{j_1=1}^{n} a_{1j_1}b_{j_1 1} & \sum\limits_{j_2=1}^{n} a_{1j_2}b_{j_2 2} & \ldots & \sum\limits_{j_n=1}^{n} a_{1j_n}b_{j_n n} \\ \sum\limits_{j_1=1}^{n} a_{2j_1}b_{j_1 1} & \sum\limits_{j_2=1}^{n} a_{2j_2}b_{j_2 2} & \ldots & \sum\limits_{j_n=1}^{n} a_{2j_n}b_{j_n n} \\ \ldots & \ldots & \ldots & \ldots \\ \sum\limits_{j_1=1}^{n} a_{nj_1}b_{j_1 1} & \sum\limits_{j_2=1}^{n} a_{nj_2}b_{j_2 2} & \ldots & \sum\limits_{j_n=1}^{n} a_{nj_n}b_{j_n n} \end{vmatrix}.$$

По свойству 4 определителей его можно представить в виде суммы

$$\sum_{j_1, j_2, \ldots, j_n} \begin{vmatrix} a_{1j_1}b_{j_1 1} & a_{1j_2}b_{j_2 2} & \ldots & a_{1j_n}b_{j_n n} \\ a_{2j_1}b_{j_1 1} & a_{2j_2}b_{j_2 2} & \ldots & a_{2j_n}b_{j_n n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{nj_1}b_{j_1 1} & a_{nj_2}b_{j_2 2} & \ldots & a_{nj_n}b_{j_n n} \end{vmatrix},$$

где индексы $j_1, j_2, \ldots, j_n$ независимо друг от друга пробегают все значения $1, 2, 3, \ldots, n$ (всего в этой сумме n^n слагаемых). Однако можно считать, что в определителе, стоящем под знаком суммы, все индексы $j_1, j_2, \ldots, j_n$ различны, так как те определители, у которых имеются одинаковые индексы j_k, равны нулю как определители с пропорциональными столбцами. Таким образом, в этой сумме остаются только $n!$ слагаемых, отвечающих разным наборам $j_1, j_2, \ldots, j_n$. Вынося теперь за знак определителя общий множитель элементов каждого столбца, получим

$$|C| = \sum_{j_1, j_2, \ldots, j_n} b_{j_1 1}b_{j_2 2} \ldots b_{j_n n} \begin{vmatrix} a_{1j_1} & a_{1j_2} & \ldots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \ldots & a_{2j_n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{nj_1} & a_{nj_2} & \ldots & a_{nj_n} \end{vmatrix},$$

где суммирование ведется по всевозможным перестановкам $j_1, j_2, \ldots, j_n$ чисел $1, 2, \ldots, n$.

В определителе, стоящем под знаком (последней) суммы, переставим столбцы так, чтобы вторые индексы их элементов расположились в порядке возрастания. Это можно сделать посредством нескольких транспозиций столбцов. Так как при переходе от одной перестановки к другой той же четности требуется четное число транспозиций, а при переходе к перестановке другой четности — нечетное число транспозиций (а перестановка $1, 2, \ldots, n$ — четная), то определитель в правой части последнего равенства равен $(-1)^{[j_1, j_2, \ldots, j_n]}|A|$. Таким образом, получаем

$$|C| = \sum_{j_1, j_2, \ldots, j_n} (-1)^{[j_1, j_2, \ldots, j_n]} b_{j_1 1} b_{j_2 2} \ldots b_{j_n n} |A| = |A||B|.$$

Следствие. $|A^{-1}| = |A|^{-1}$, что вытекает из равенства

$$1 = |E| = |AA^{-1}| = |A|\,|A^{-1}|.$$

§ 3. Прямоугольные матрицы

Матрица, состоящая из m строк и n столбцов, называется $[m \times n]$-матрицей. Можно определить *сложение* $[m \times n]$-*матриц*, полагая

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \ldots & b_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \ldots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \ldots & a_{2n}+b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \ldots & a_{mn}+b_{mn} \end{bmatrix},$$

и *умножение* $[m \times n]$-*матрицы на число* α — равенством

$$\alpha \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \ldots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \ldots & \alpha a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha a_{m1} & \alpha a_{m2} & \ldots & \alpha a_{mn} \end{bmatrix}.$$

Легко видеть, что относительно этих операций сложения и умножения на число $[m \times n]$-матрицы (в частности, квадратные матрицы порядка n) с элементами из поля F сами образуют векторное пространство над полем F. Обозначим $[m \times n]$-матрицу, у которой элемент i-й строки и k-го столбца равен 1, а все остальные элементы равны нулю, через e_{ik}. Тогда ясно, что эти матрицы e_{ik}, где $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots, n$, линейно независимы и что каждая $[m \times n]$-матрица является их линейной комбинацией. Следовательно, размерность пространства $[m \times n]$-матриц равна mn. В частности, пространство всех квадратных матриц порядка n имеет размерность n^2.

Прямоугольную матрицу можно рассматривать как матрицу *линейного оператора, отображающего* одно векторное пространство в другое. А именно, пусть имеются два векторных пространства R^n и R^m, вообще говоря, разных размерностей n и m, но над одним и тем же числовым полем F, и предположим, что каждому вектору $x \in R^n$ поставлен в соответствие вектор $\mathcal{A}x \in R^m$ так, что выполнены следующие условия:

1. $\mathcal{A}(x+y) = \mathcal{A}x + \mathcal{A}y$,
2. $\mathcal{A}(\alpha x) = \alpha\mathcal{A}x$

для всех $x, y \in R^n$ и $\alpha \in F$.

Мы говорим тогда, что $\mathcal{A}$ есть *линейный оператор, отображающий пространство R^n в R^m, или линейное отображение R^n в R^m.*

Выберем в пространстве R^n базис $e_1, e_2, \ldots, e_n$, а в пространстве R^m базис $f_1, f_2, \ldots, f_m$. Вектор Ae_i, где $i = 1, 2, \ldots, n$, принадлежит R^m, и следовательно, его можно разложить по базису $f_1, f_2, \ldots, f_m$; пусть

$$\begin{aligned} \mathcal{A}e_1 &= a_{11}f_1 + a_{21}f_2 + \ldots + a_{m1}f_m, \\ \mathcal{A}e_2 &= a_{12}f_1 + a_{22}f_2 + \ldots + a_{m2}f_m, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathcal{A}e_n &= a_{1n}f_1 + a_{2n}f_2 + \ldots + a_{mn}f_m. \end{aligned}$$

Таким образом, линейному оператору $\mathcal{A}$, отображающему пространство R^n в R^m, соответствует *прямоугольная матрица*

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

столбцы которой образованы коэффициентами разложений векторов e_i по векторам f_j.

По аналогии с умножением квадратных частиц можно определить и умножение прямоугольных матриц. Такое умножение выполнимо только в том случае, если длина строки левого множителя равна длине столбца правого, т. е. когда число столбцов левого множителя равно числу строк правого. Произведение $[m \times n]$-матрицы на $[n \times p]$-матрицу будет, очевидно, $[m \times p]$-матрицей. В частности, произведение $[m \times n]$-матрицы на $[n \times 1]$-матрицу, т. е. на столбец, будет $[m \times 1]$-матрицей, т. е. столбцом, а произведение $[1 \times m]$-матрицы, т. е. строки, на $[m \times n]$-матрицу будет $[1 \times n]$-матрицей, т. е. строкой.

Примеры

1. $$\begin{bmatrix} -1 & 3 & 2 & 0 \\ 1 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & -3 & 3 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 7 \\ 2 & -5 & 4 \end{bmatrix}.$$

2. $$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

3. $$[1 \quad 2 \quad 3] \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} = [14 \quad 20].$$

Нетрудно понять «геометрический смысл» операции умножения прямоугольных матриц. Пусть имеются три векторных пространства, вообще говоря, разных размерностей: R^n, R^m, R^p, и пусть даны два линейных оператора: $\mathscr{A}$, отображающий R^m в R^p, и $\mathscr{B}$, отображающий R^n в R^m. Оператор $\mathscr{A}\mathscr{B}$, ставящий в соответствие каждому вектору $x \in R^n$ вектор $\mathscr{A}(\mathscr{B}x)$ пространства R^p, называется *произведением* операторов $\mathscr{A}$ и $\mathscr{B}$. Легко видеть, что $\mathscr{A}\mathscr{B}$ является линейным оператором, отображающим R^n в R^p, и что если оператору $\mathscr{A}$ отвечает $[p \times m]$-матрица A, а оператору $\mathscr{B}$ — $[m \times n]$-матрица B, то матрицей оператора $\mathscr{A}\mathscr{B}$ будет $[p \times n]$-матрица AB.

Как и для квадратных матриц, умножение прямоугольных матриц ассоциативно: $A(BC) = (AB)C$, и дистрибутивно относительно сложения: $(A+B)C = AC + BC$, $C(A+B) = CA + CB$,— разумеется, если матрицы таковы, что все эти действия над ними выполнимы. Кроме того, если A — произвольная

$[m\times n]$-матрица, то $AE_n=A$ и $E_mA=A$, где E_k — единичная матрица порядка k.

Рассмотрим снова систему линейных уравнений

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \cdots\cdots\cdots\cdots&\cdots\cdots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=b_m. \end{aligned}\right\} \quad (3)$$

Обозначим через A матрицу из коэффициентов при неизвестных этой системы, через X — столбец, составленный из неизвестных, и через B — столбец, составленный из правых частей. Тогда систему (3) можно записать в виде одного матричного уравнения

$$AX=B.$$

Если матрица A квадратная и ее определитель отличен от нуля, то существует обратная к ней матрица A^{-1}. Умножая обе части последнего равенства слева на A^{-1}, получим

$$A^{-1}(AX)=A^{-1}B, \quad \text{откуда} \quad X=A^{-1}B.$$

В более подробной записи это — *формулы Крамера* (ср. выше, стр. 33).

Если $\mathcal{A}$ — линейный оператор в пространстве R и A — его матрица в некотором базисе $e_1, e_2, \ldots, e_n$, в котором $x=(x_1, x_2, \ldots, x_n)$ и $\mathcal{A}x=(x_1', x_2', \ldots, x_n')$, то формулы (1) из § 1 можно записать в виде одного матричного уравнения

$$Y=AX,$$

где X — столбец из координат вектора x, а Y — столбец из координат вектора $\mathcal{A}x$.

Наконец, если C — матрица перехода от базиса $e_1, e_2, \ldots, e_n$ к (новому) базису $e_1', e_2', \ldots, e_n'$, то

$$X_{\text{ст}}=CX_{\text{нов}},$$

где $X_{\text{ст}}$ — столбец старых, а $X_{\text{нов}}$ — столбец новых координат вектора x (см. формулы в § 6 главы II). Из последней формулы непосредственно вытекает равенство: $X_{\text{нов}}=C^{-1}X_{\text{ст}}$, т. е. что новые координаты получаются из старых с помощью матрицы, обратной матрице перехода, что впрочем вполне очевидно и так.

Укажем здесь еще один, практически более удобный, чем изложенный выше, способ вычисления матрицы A^{-1}, обратной данной невырожденной матрице A.

Выпишем рядом матрицу A и единичную матрицу E и над строками их будем одновременно производить элементарные преобразования до тех пор, пока матрица A не превратится в единичную. При этом исходная единичная матрица превратится в A^{-1}.

Рассмотрим пример. Пусть нам надо найти матрицу A^{-1}, обратную матрице $A=\begin{bmatrix}2 & 1\\ 4 & 3\end{bmatrix}$. Выпишем рядом с A единичную матрицу E и над строками полученной «объединенной (прямоугольной) матрицы» будем производить элементарные преобразования: сначала отнимем от второй строки утроенную первую, затем разделим вторую строку на —2, вычтем удвоенную вторую строку из первой и, наконец, переставим строки. Так мы получим последовательно:

$$\left[\begin{array}{cc|cc}2 & 1 & 1 & 0\\ 4 & 3 & 0 & 1\end{array}\right]\sim\left[\begin{array}{cc|cc}2 & 1 & 1 & 0\\ -2 & 0 & -3 & 1\end{array}\right]\sim\left[\begin{array}{cc|cc}2 & 1 & 1 & 0\\ 1 & 0 & \frac{3}{2} & -\frac{1}{2}\end{array}\right]\sim$$

$$\sim\left[\begin{array}{cc|cc}0 & 1 & -2 & 1\\ 1 & 0 & \frac{3}{2} & -\frac{1}{2}\end{array}\right]\sim\left[\begin{array}{cc|cc}1 & 0 & \frac{3}{2} & -\frac{1}{2}\\ 0 & 1 & -2 & 1\end{array}\right].$$

В последнем «блоке» левая матрица — единичная, а правая равна A^{-1}.

Можно было бы производить элементарные преобразования не над строками, а над столбцами — одновременно матриц A и E,— но всегда *либо только над строками, либо только над столбцами*. Во втором случае (при элементарных преобразованиях столбцов) удобнее располагать матрицы «столбиком»: единичную матрицу E помещать *под* матрицей A. Так, в нашем примере мы будем иметь последовательно:

$$\left[\begin{array}{cc}2 & 1\\ 4 & 3\\ \hline 1 & 0\\ 0 & 1\end{array}\right]\sim\left[\begin{array}{cc}0 & 1\\ -2 & 3\\ \hline 1 & 0\\ -2 & 1\end{array}\right]\sim\left[\begin{array}{cc}0 & 1\\ 1 & 3\\ \hline -\frac{1}{2} & 0\\ 1 & 1\end{array}\right]\sim$$

$$\sim\left[\begin{array}{cc}0 & 1\\ 1 & 0\\ \hline -\frac{1}{2} & \frac{3}{2}\\ 1 & -2\end{array}\right]\sim\left[\begin{array}{cc}1 & 0\\ 0 & 1\\ \hline \frac{3}{2} & -\frac{1}{2}\\ -2 & 1\end{array}\right].$$

Здесь удвоенный второй столбец вычитается из первого, первый столбец делится на —2, утроенный первый столбец вычитается из второго и, наконец, столбцы меняются местами. После того как верхняя матрица превратилась в единичную, нижняя будет равна A^{-1}.

Для того чтобы обосновать эти действия, заметим следующее.

1. Умножение произвольной $[n \times m]$-матрицы A слева на матрицу B порядка n, получающуюся из единичной матрицы умножением ее i-й строки на число c, равносильно умножению на c i-й строки самой матрицы A. В то же время умножение матрицы A на аналогичную матрицу B_1 поряка m справа равносильно умножению на c i-го столбца матрицы A. (Проверьте это сами.)

2. Умножение $[n \times m]$-матрицы A слева на матрицу C порядка n, получающуюся из единичной матрицы перестановкой ее i-й и k-й строк, равносильно перестановке i-й и k-й строк самой матрицы A, а умножение матрицы A на аналогичную матрицу C_1 порядка m справа равносильно перестановке i-го и k-го столбцов матрицы A. (Проверьте и это.)

3. Умножение $[n \times m]$-матрицы A слева на матрицу D, получающуюся из единичной матрицы порядка n прибавлением к ее i-й строке k-й строки, умноженной на c, равносильно аналогичной операции над строками самой матрицы A. Так, например,

$$\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}=\begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}.$$

В то же время умножение матрицы A справа на матрицу D_1 порядка m, получающуюся из единичной матрицы прибавлением к ее i-му столбцу k-го столбца, умноженного на число c, равносильно аналогичной операции над столбцами самой матрицы A. (Докажите все это сами.) Так, например,

$$\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}.$$

Пусть теперь A^{-1} — матрица, обратная A, тогда $AA^{-1} = E$. Элементарные преобразования над строками матрицы A равносильны умножению ее *слева* на некоторые специальным образом подобранные матрицы. На те же матрицы одновременно умножается и матрица E. В нашем примере элементарные преобразования над строками матрицы $\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$ отвечают таким действиям:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}AA^{-1} =$$
$$=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}E.$$

Но произведение

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}A$$

равно E, и значит, правая часть последнего равенства равна A^{-1}.

Элементарные преобразования столбцов отвечают умножению равенства $A^{-1}A = E$ (обратите внимание на то, что теперь мы написали матрицу A с п р а в а) *справа* на определенным образом подобранные матрицы. Так, в том же примере мы имеем

$$A^{-1}A\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}=$$

$$=E\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

А так как произведение

$$A\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

равно E, то правая часть последнего равенства равна A^{-1}.

Заметим, что аналогичный прием можно применить и при решении матричного уравнения, скажем, вида $AX = B$, где A — квадратная матрица порядка n, X — искомая и B — данная $[n \times m]$-матрицы: производим элементарные преобразования с т р о к одновременно матриц A и B до тех пор, пока матрица A не превратится в единичную; при этом матрица B превратится в $A^{-1}B = X$. (Сравните это с методом Гаусса на стр. 50, которым, в сущности, решается матричное уравнение $AX = B$, где A — матрица из коэффициентов при неизвестных, X — столбец неизвестных и B — столбец правых частей.)

§ 4. Изменение матрицы линейного оператора при переходе к новому базису

Пусть линейный оператор $\mathscr{A}$, действующий в пространстве R, в базисе $e_1, e_2, \ldots, e_n$ имеет матрицу $A = [a_{ik}]$, а в базисе $e'_1, e'_2, \ldots, e'_n$, вообще говоря, другую матрицу $A_1 = [a'_{ik}]$. Найдем, как связаны между собой матрицы A и A_1.

Обозначим через $C = [c_{ik}]$ матрицу перехода от базиса $e_1, e_2, \ldots, e_n$ к базису $e'_1, e'_2, \ldots, e'_n$. Тогда

$$e'_i = c_{1i}e_1 + c_{2i}e_2 + \ldots + c_{ni}e_n, \quad \text{где} \quad i = 1, 2, \ldots, n.$$

Будем матрицу C рассматривать как матрицу линейного оператора $\mathscr{C}$ в базисе $e_1, e_2, \ldots, e_n$. Тогда очевидно, что

$$\mathscr{C}e_i = c_{1i}e_1 + c_{2i}e_2 + \ldots + c_{ni}e_n = e'_i,$$

и значит, линейный оператор $\mathscr{C}$ переводит векторы $e_1, e_2, \ldots, e_n$ соответственно в векторы $e'_1, e'_2, \ldots, e'_n$.

Определитель матрицы C отличен от нуля (§ 6 главы II), а значит, для $\mathscr{C}$ существует обратный оператор $\mathscr{C}^{-1}$ такой, что $\mathscr{C}^{-1}e_1' = e_1, \mathscr{C}^{-1}e_2' = e_2, \ldots, \mathscr{C}^{-1}e_n' = e_n$. По условию,

$$\mathscr{A}e_i' = a_{1i}'e_1' + a_{2i}'e_2' + \ldots + a_{ni}'e_n'.$$

Применяя к обеим частям этого равенства оператор $\mathscr{C}^{-1}$, получим

$$\mathscr{C}^{-1}\mathscr{A}e_i' = a_{1i}'e_1 + a_{2i}'e_2 + \ldots + a_{ni}'e_n.$$

Подставляя в левую часть последнего равенства $e_i' = \mathscr{C}e_i$, будем иметь

$$\mathscr{C}^{-1}\mathscr{A}\mathscr{C}e_i = a_{1i}'e_1 + a_{2i}'e_2 + \ldots + a_{ni}'e_n, \tag{4}$$

т. е. матрицей оператора $\mathscr{C}^{-1}\mathscr{A}\mathscr{C}$ в базисе $e_1, e_2, \ldots, e_n$ является матрица A_1. Но, с другой стороны, матрица этого оператора равна произведению матриц операторов $\mathscr{C}^{-1}$, $\mathscr{A}$ и $\mathscr{C}$ в базисе $e_1, e_2, \ldots, e_n$, т. е.

$$A_1 = C^{-1}AC. \tag{5}$$

Отсюда, в частности, следует, что *определитель матрицы линейного оператора не зависит от базиса:*

$$|A_1| = |C^{-1}AC| = |C^{-1}||A||C| = |C|^{-1}|A||C| = |A|$$

(см. следствие на стр. 106).

Пример. *В базисе e_1, e_2 преобразование $\mathscr{A}$ имеет матрицу $A = \begin{bmatrix} 6 & -2 \\ 6 & -1 \end{bmatrix}$. Написать матрицу этого преобразования в базисе*

$$e_1' = e_1 + 2e_2, \quad e_2' = 2e_1 + 3e_2.$$

Решение. Матрица перехода здесь $C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, а обратная к ней матрица $C^{-1} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$. Следовательно,

$$A_1 = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 6 & -2 \\ 6 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 6 \\ 4 & 9 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Формулу (5) можно получить еще и следующим образом. Как показано в § 3, имеем

$$X_{\text{ст}} = CX_{\text{нов}} \quad \text{и} \quad Y_{\text{ст}} = CY_{\text{нов}},$$

$$Y_{\text{ст}} = AX_{\text{ст}} \quad \text{и} \quad Y_{\text{нов}} = A_1X_{\text{нов}}.$$

Следовательно, $CY_{\text{нов}} = Y_{\text{ст}} = AX_{\text{ст}} = ACX_{\text{нов}}$, откуда $Y_{\text{нов}} = C^{-1}ACX_{\text{нов}}$. Но $Y_{\text{нов}} = A_1X_{\text{нов}}$, и значит, $A_1 = C^{-1}AC$. (Легко видеть, что из матричного равенства $BX = B_1X$, справедливого при всех X — если B и B_1 одного строения, вытекает, что $B = B_1$.)

§ 5. Ранг и дефект линейного оператора

Определение 2. *Пусть $\mathscr{A}$ — линейный оператор, действующий в пространстве R. Совокупность $\mathscr{A}R$ всевозможных векторов вида $\mathscr{A}x$, где $x \in R$, называется* ***областью значений*** *оператора $\mathscr{A}$, или* ***образом пространства*** *R при преобразовании $\mathscr{A}$, а множество N всевозможных векторов x, для которых $\mathscr{A}x = 0$, —* ***ядром*** *оператора $\mathscr{A}$.*

Покажем, что *область значений и ядро линейного оператора $\mathscr{A}$ являются подпространствами в R.*

Действительно, для области значений это вытекает из теоремы 1, если рассматриваемое в ней подпространство R_1 совпадает со всем пространством R.

С другой стороны, если $x, y \in N$, т. е. если $\mathscr{A}x = 0$ и $\mathscr{A}y = 0$, то и $\mathscr{A}(x + y) = \mathscr{A}x + \mathscr{A}y = 0$ и $\mathscr{A}(\alpha x) = \alpha\mathscr{A}x = 0$, т. е. $x + y \in N$ и $\alpha x \in N$, и значит, N — подпространство.

Размерность области значений оператора $\mathscr{A}$ совпадает с рангом матрицы A (и называется *рангом оператора $\mathscr{A}$*). Действительно, подпространство $\mathscr{A}R$ порождается векторами

$$\mathscr{A}e_1, \mathscr{A}e_2, \ldots, \mathscr{A}e_n, \tag{6}$$

где $e_1, e_2, \ldots, e_n$ — любой базис пространства R, и значит, размерность $\mathscr{A}R$ равна максимальному числу линейно независимых векторов в системе (6), т. е. равна максимальному числу линейно независимых столбцов матрицы A.

Размерность ядра N называется ***дефектом*** *линейного оператора $\mathscr{A}$.*

Теорема 4. *Сумма ранга и дефекта линейного оператора равна размерности n пространства.*

Доказательство. Если ранг линейного оператора $\mathscr{A}$ равен r, то среди векторов $\mathscr{A}e_1, \mathscr{A}e_2, \ldots, \mathscr{A}e_n$ найдется r линейно независимых, через которые линейно выражаются все остальные. Пусть, для определенности, это будут $\mathscr{A}e_1, \mathscr{A}e_2, \ldots, \mathscr{A}e_r$.

Обозначим через L подпространство, порожденное в R векторами $e_1, e_2, \ldots, e_r$, и покажем, что (r-мерное) *подпространство L и ядро N пересекаются только по нулевому вектору.* Действительно, если $x \in L \cap N$, то $x \in L$, т. е. $x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r$, и $x \in N$, т. е. $\mathcal{A}x = \alpha_1 \mathcal{A}e_1 + \alpha_2 \mathcal{A}e_2 + \ldots + \alpha_r \mathcal{A}e_r = 0$. Но так как векторы $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_r$ линейно независимы, то $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$ и $x = 0$.

Покажем теперь, что *подпространства L и N порождают все R* (т. е. что их сумма совпадает с R). Пусть x — произвольный вектор из R. Тогда $\mathcal{A}x \in \mathcal{A}R$ и, следовательно, $\mathcal{A}x = \beta_1 \mathcal{A}e_1 + \beta_2 \mathcal{A}e_2 + \ldots + \beta_r \mathcal{A}e_r$. Вектор $y = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_r e_r$ принадлежит, очевидно, L, а разность $z = x - y \in N$, так как $\mathcal{A}z = \mathcal{A}(x - y) = \mathcal{A}x - \mathcal{A}y = 0$. Мы нашли, что $x = y + z$, где $y \in L$, а $z \in N$.

Таким образом, пространство R равно прямой сумме подпространств L и N, а значит, его размерность n равна сумме размерностей этих подпространств.

В дальнейшем нам понадобится еще такое

Определение 2'. *Пусть $\mathcal{A}$ — линейный оператор, отображающий пространство R_1 в пространство R_2 (вообще говоря, другой размерности). Тогда множество $\mathcal{A}R_1 \subseteq R_2$ всех векторов y из R_2 вида $y = \mathcal{A}x$, где $x \in R_1$, называется областью значений оператора $\mathcal{A}$ (или образом пространства R при отображении $\mathcal{A}$), а множество $N \subseteq R_1$ всех векторов x из R_1 таких, что $\mathcal{A}x = 0$, — его ядром.*

Нетрудно видеть, что область *значений оператора $\mathcal{A}$ является подпространством в R_2, а его ядро — подпространством в R_1* (докажите это).

§ 6. Невырожденный линейный оператор

Линейный оператор $\mathcal{A}$ мы назвали невырожденным, если из равенства $\mathcal{A}x = 0$ вытекает, что $x = 0$ (§ 1). Далее, в § 1 было показано, что матрица невырожденного линейного оператора в любом базисе имеет отличный от нуля определитель, а в § 2, — что *для всякого невырожденного линейного оператора $\mathcal{A}$ существует обратный линейный оператор $\mathcal{A}^{-1}$*. Наоборот, *если для линейного оператора $\mathcal{A}$ существует обратный оператор*

$\mathscr{A}^{-1}$, *то этот оператор — невырожденный*, так как из равенства $\mathscr{A}x = 0$, применяя к обеим его частям оператор $\mathscr{A}^{-1}$, получаем $\mathscr{A}^{-1}(\mathscr{A}x) = \mathscr{A}^{-1}0 = 0$; но $\mathscr{A}^{-1}(\mathscr{A}x) = (\mathscr{A}^{-1}\mathscr{A})x = \mathscr{E}x = x$, и значит, $x = 0$.

Ранг невырожденного линейного оператора в пространстве $R^n = R$ равен n, так как определитель его матрицы отличен от нуля; *дефект невырожденного линейного оператора равен нулю*. Обратно, *всякий линейный оператор ранга n будет*, очевидно, *невырожденным*. Область значений невырожденного линейного оператора n-мерна и, значит, совпадает со всем R: невырожденный линейный оператор отображает R на все R. Ядро невырожденного линейного оператора состоит лишь из нулевого вектора. Невырожденный линейный оператор *взаимно однозначно отображает пространство R на себя*, так как из равенства $\mathscr{A}x = \mathscr{A}y$ вытекает, что $\mathscr{A}(x - y) = 0$ и, значит, $x - y = 0$, т. е. $x = y$.

Под действием *невырожденного линейного оператора линейно независимые векторы переходят в линейно независимые*. Действительно, если векторы $e_1, e_2, \ldots, e_k$ линейно независимы и

$$\alpha_1\mathscr{A}e_1 + \alpha_2\mathscr{A}e_2 + \ldots + \alpha_k\mathscr{A}e_k = \\ = \mathscr{A}(\alpha_1e_1 + \alpha_2e_2 + \ldots \alpha_ke_k) = 0,$$

то

$$\alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ke_k = 0 \text{ и } \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0.$$

Следовательно, если $\mathscr{A}$ — невырожденный линейный оператор и подпространство $R_1 \subseteq R$ r-мерно, то и подпространство $\mathscr{A}R_1$ имеет ту же размерность r.

Покажем, что *если $\mathscr{A}$ — линейный оператор ранга r, а $\mathscr{B}$ — невырожденный линейный оператор, то оба оператора $\mathscr{A}\mathscr{B}$ и $\mathscr{B}\mathscr{A}$ будут ранга r*.

Действительно, область значений (невырожденного) линейного оператора $\mathscr{B}$ совпадает со всем пространством: $\mathscr{B}R = R$, и следовательно, область значений оператора $\mathscr{A}\mathscr{B}$ r-мерна, т. е. ранг оператора $\mathscr{A}\mathscr{B}$ равен r.

С другой стороны, область значений $\mathscr{A}R$ оператора $\mathscr{A}$ r-мерна, а так как оператор $\mathscr{B}$-невырожденный, то он переводит r-мерное подпространство $\mathscr{A}R$ в r-мерное же подпространство $\mathscr{B}\mathscr{A}R$, и значит, область значений оператора $\mathscr{B}\mathscr{A}$ тоже r-мерна, т. е. ранг оператора $\mathscr{B}\mathscr{A}$ равен r.

§ 7. Инвариантные подпространства

Пусть R_1 — подпространство векторного пространства R и $\mathscr{A}$ — действующий в R линейный оператор. Образ $\mathscr{A}x$ вектора x из R_1, вообще говоря, не обязан принадлежать R_1. Особый интерес представляют такие подпространства, векторы которых действием оператора $\mathscr{A}$ не выводятся из этих подпространств.

Определение 3. *Подпространство R_1 пространства R называется **инвариантным** относительно линейного оператора $\mathscr{A}$, если образ $\mathscr{A}x$ каждого вектора x из R_1 принадлежит R_1* (иными словами, если $\mathscr{A}R_1 \subseteq R_1$).

Примеры. 1. Пусть $\mathscr{A}$ — поворот вокруг оси Oz обычного трехмерного пространства. Инвариантными подпространствами будут, например, плоскость xOy и ось Oz.

2. Если $\mathscr{A}$ — ортогональное проектирование того же пространства R^3 на плоскость xOy, то инвариантными подпространствами будут: плоскость xOy, все плоскости, проходящие через ось Oz, сама ось Oz и все прямые, содержащиеся в плоскости xOy (и проходящие через начало координат).

3. В пространстве P_n многочленов степени не выше n подпространства P_k при всех k, $0 \leqslant k \leqslant n$, инвариантны относительно оператора дифференцирования.

4. В любом пространстве каждое подпространство инвариантно относительно тождественного и нулевого операторов.

5. В любом пространстве само пространство R и его подпространство, состоящее из одного нулевого вектора, инвариантны относительно любого линейного оператора.

Покажем, что *пересечение и сумма подпространств, инвариантных относительно линейного оператора $\mathscr{A}$, инвариантны относительно $\mathscr{A}$.*

Действительно, если подпространства R_1 и R_2 инвариантны относительно $\mathscr{A}$ и $x \in R_1 \cap R_2$, то $x \in R_1$ и $x \in R_2$, а значит, $Ax \in R_1$ и $Ax \in R_2$, т. е. $Ax \in R_1 \cap R_2$.

С другой стороны, если $x \in R_1 + R_2$, то $x = u + v$, где $u \in R_1$, $v \in R_2$. Но тогда $\mathscr{A}u \in R_1$, $Av \in R_2$ и $\mathscr{A}x = \mathscr{A}u + \mathscr{A}v \in R_1 + R_2$.

Теорема 5. *Если $\mathscr{A}$ — невырожденный линейный оператор и R_1 — подпространство, инвариантное относительно $\mathscr{A}$, то R_1 инвариантно и относительно $\mathscr{A}^{-1}$.*

Доказательство. Пусть $e_1, e_2, \ldots, e_r$ — базис подпространства R_1. Тогда векторы $\mathscr{A}e_1, \mathscr{A}e_2, \ldots, \mathscr{A}e_r$, тоже принадлежащие R_1 (ввиду инвариантности R_1), линейно независимы (см. § 6), и значит, они тоже образуют базис R_1, т. е. произвольный вектор $x \in R_1$ представляется в виде

$$x = \alpha_1 \mathscr{A}e_1 + \alpha_2 \mathscr{A}e_2 + \ldots + \alpha_r \mathscr{A}e_r.$$

Но тогда и

$$\mathscr{A}^{-1}x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r$$

принадлежит R_1.

Сделаем еще одно, полезное для дальнейшего

Замечание. Пусть $\mathscr{A}$ — произвольный линейный оператор, действующий в n-мерном пространстве R; предположим, что R распадается в прямую сумму $R = R_1 \oplus R_2$ своих подпространств R_1 и R_2, *инвариантных относительно* $\mathscr{A}$; $e_1, e_2, \ldots, e_r$ — базис R_1 и $e_{r+1}, e_{r+2}, \ldots, e_n$ — базис R_2. Ввиду инвариантности подпространств R_1 и R_2, имеют место равенства

$$\mathscr{A}e_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ri}e_r \text{ при } i = 1, 2, \ldots, r$$

и

$$\mathscr{A}e_k = a_{r+1,k}e_{r+1} + a_{r+2,k}e_{r+2} + \ldots + a_{nk}e_n \quad \text{при } k = r+1, r+2, \ldots, n$$

(так как $\mathscr{A}e_i \in R_1$ при $i = 1, 2, \ldots, r$ и $\mathscr{A}e_k \in R_2$ при $k = r+1, r+2, \ldots, n$). Тогда матрица оператора $\mathscr{A}$ в базисе $e_1, e_2, \ldots, e_n$ всего пространства имеет вид

$$A = \begin{bmatrix} a_{11} & a_{12} \ldots a_{1r} & 0 & 0 & \ldots 0 \\ a_{21} & a_{22} \ldots a_{2r} & 0 & 0 & \ldots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} \ldots a_{rr} & 0 & 0 & \ldots 0 \\ 0 & 0 \ \ldots 0 & a_{r+1,r+1} & a_{r+1,r+2} & \ldots a_{r+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 \ \ldots 0 & a_{n,r+1} & a_{n,r+2} & \ldots a_{nn} \end{bmatrix}. \qquad (7)$$

Можно сказать, что матрица A «распадается на клетки»:

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

где A_1 — матрица оператора $\mathscr{A}$ в подпространстве R_1, A_2 — матрица оператора $\mathscr{A}$ в подпространстве R_2, а прямоугольные матрицы в левом нижнем и правом верхнем углах матрицы A состоят из одних нулей.

Таким образом, *знaя матрицы* A_1 и A_2 *оператора* $\mathscr{A}$ *в подпространствах* R_1 *и* R_2, *мы можем составить из них матрицу оператора* $\mathscr{A}$ *во всем пространстве* R.

Верно и обратное утверждение: если матрица оператора $\mathscr{A}$ в некотором базисе имеет «клеточный» вид (7), то пространство R очевидным образом распадается в прямую сумму инвариантных относительно $\mathscr{A}$ подпространств R_1 и R_2.

§ 8. Собственные векторы и собственные значения линейного оператора

В предыдущем параграфе мы познакомились с определением подпространства, инвариантного относительно данного линейного оператора. При этом особый интерес представляют одномерные инвариантные подпространства. Пусть R_1 — такое подпространство и $x \in R_1$ (где $x \neq 0$); тогда $\mathscr{A}x \in R_1$, и значит, $\mathscr{A}x = \lambda_0 x$, где λ_0 — число. Если y — любой другой вектор из R_1, то $y = \alpha x$ и

$$\mathscr{A}y = \mathscr{A}(\alpha x) = \alpha \mathscr{A}x = \alpha(\lambda_0 x) = \lambda_0(\alpha x) = \lambda_0 y.$$

Определение 4. *Вектор* $x \neq 0$ *называется* ***собственным вектором*** *линейного оператора* $\mathscr{A}$, *если найдется такое число* λ_0, *что* $\mathscr{A}x = \lambda_0 x$; *это* λ_0 *называется соотвествующим вектору* x ***собственным значением*** *оператора* $\mathscr{A}$ (*матрицы* A).

Как мы только что видели, если R_1 — одномерное инвариантное относительно оператора $\mathscr{A}$ подпространство R, то каждый ненулевой вектор из R_1 является собственным вектором оператора $\mathscr{A}$ и притом с одним и тем же собственным значением. Обратно, если x — собственный вектор оператора $\mathscr{A}$, то порожденное им одномерное подпространство R_1 (состоящее из всех векторов вида αx) будет, очевидно, инвариантным относительно $\mathscr{A}$.

Как найти собственные значения и собственные векторы линейного оператора? Предположим, что x — собственный вектор, а λ_0 — соответствующее ему собственное значение линейного оператора $\mathscr{A}$. Тогда $\mathscr{A}x = \lambda_0 x$, Выберем в пространстве R какой-нибудь базис $e_1, e_2, \ldots$ $\ldots, e_n$, и пусть $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, а матрица оператора $\mathcal{A}$ в этом базисе $A = [a_{ik}]$. Тогда (см. § 1)

$$\mathcal{A}x = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)e_1 + \\ + (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n)e_2 + \ldots \\ \ldots + (a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n)e_n = \\ = \lambda_0(x_1e_1 + x_2e_2 + \ldots + x_ne_n),$$

откуда, ввиду единственности разложения вектора $\mathcal{A}x$ по базису $e_1, e_2, \ldots, e_n$,

$$\left\{\begin{array}{l} (a_{11} - \lambda_0)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + (a_{22} - \lambda_0)x_2 + \ldots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda_0)x_n = 0. \end{array}\right. \tag{8}$$

Для существования ненулевого решения этой (однородной) системы необходимо и достаточно, чтобы ее определитель был равен нулю:

$$\begin{vmatrix} a_{11} - \lambda_0 & a_{12} & \ldots a_{1n} \\ a_{21} & a_{22} - \lambda_0 & \ldots a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots a_{nn} - \lambda_0 \end{vmatrix} = 0$$

(теорема 10 из главы I). Левая часть последнего равенства совпадает со значением при $\lambda = \lambda_0$ определителя $|A - \lambda E|$ матрицы $A - \lambda E$, который является многочленом относительно λ степени n. Коэффициенты этого многочлена $\varphi(\lambda)$, называемого **характеристическим многочленом матрицы** A, принадлежат, очевидно, основному полю F. Ниже (теорема 6) будет показано, что многочлен $\varphi(\lambda)$ на самом деле **не зависит от выбора базиса**, и поэтому его можно назвать *характеристическим многочленом* **оператора** $\mathcal{A}$.

Мы доказали, что *каждое собственное значение линейного оператора $\mathcal{A}$ является корнем его характеристического многочлена.* Обратно, *каждый корень λ_0 характеристического многочлена оператора $\mathcal{A}$ будет его собственным значением* — соответствующие λ_0 собственные векторы находятся из системы уравнений (8), которая в этом случае обязательно имеет ненулевые решения, так как ее определитель равен нулю.

Теорема 6. *Характеристический многочлен линейного оператора не зависит от выбора базиса.*

Доказательство. Пусть $\varphi_e(\lambda) = |A - \lambda E|$ — характеристический многочлен оператора $\mathcal{A}$ в базисе $e_1, e_2, \ldots, e_n$. Предположим, что новый базис $e'_1, e'_2, \ldots, e'_n$ получается из старого с помощью матрицы C. Тогда характеристический многочлен оператора $\mathcal{A}$ в базисе $e'_1, e'_2, \ldots, e'_n$

$$\varphi_{e'}(\lambda) = |C^{-1}AC - \lambda E| = |C^{-1}AC - C^{-1}\lambda EC| = \\ = |C^{-1}(A - \lambda E)C| = |C^{-1}||A - \lambda E||C| = \\ = |C|^{-1}|A - \lambda E||C| = |A - \lambda E| = \varphi_e(\lambda).$$

Пусть

$$\varphi(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\alpha_1\lambda^{n-1} + \ldots + \alpha_n$$

— характеристический многочлен оператора $\mathcal{A}$. Легко видеть, что α_1 равно сумме $a_{11} + a_{22} + \ldots + a_{nn}$ диагональных элементов матрицы A (эта сумма называется *следом* матрицы A и обозначается символом $\operatorname{tr} A$ *)). С другой стороны, $\alpha_n = \varphi(0)$ есть *определитель* матрицы A; поэтому *для того чтобы оператор $\mathcal{A}$ был невырожденным, необходимо и достаточно, чтобы $\varphi(0)$ было отлично от нуля, т. е. чтобы оператор $\mathcal{A}$ не имел нулевых собственных значений* (что, впрочем, ясно и непосредственно).

Для *тождественного* оператора все ненулевые векторы пространства являются, очевидно, собственными (с собственным значением, равным *единице*).

Для *нулевого* оператора все ненулевые векторы пространства являются собственными (с собственным значением, равным нулю).

Найдем собственные значения и собственные векторы преобразования 1 *из* § 1.

Характеристический многочлен

$$\varphi(\lambda) = \begin{vmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{vmatrix} = \lambda^2 - 2\cos\varphi\cdot\lambda + 1.$$

*) От английского слова trace — след; иногда также употребляют обозначение Sp A от немецкого слова Spur — след.

Его корни $\lambda_{1,2} = \cos\varphi \pm i\sin\varphi$ комплексны. Значит, в вещественной плоскости, и если φ не кратно π, это преобразование не имеет собственных значений.

Если $\varphi = 2\pi k$, преобразование является *тождественным*, и каждый вектор плоскости — собственный (причем $\lambda = 1$).

Если $\varphi = (2k+1)\pi$, преобразование является *центральной симметрией*, и каждый вектор плоскости будет собственным с собственным значением, равным -1.

В комплексном случае система (8) приводится к уравнению $ix_1 + x_2 = 0$ для собственного значения $\lambda_1 = \cos\varphi + i\sin\varphi$ и к уравнению $ix_1 - x_2 = 0$ — для корня $\lambda_2 = \cos\varphi - i\sin\varphi$. Это дает два линейно независимых собственных вектора $(1, -i)$ и $(1, i)$.

Рассмотрим еще один пример.

Пример. *Найти собственные значения и собственные векторы линейного преобразования $\mathscr{A}$ с матрицей*

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}.$$

Решение. Характеристический многочлен преобразования $\mathscr{A}$:

$$\varphi(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{vmatrix} = \lambda^2 - 5\lambda - 6.$$

Его корни $\lambda_1 = 6$, $\lambda_2 = -1$. Собственные векторы находятся из двух систем уравнений:

$$\begin{cases} (1-\lambda_i)\,x_1 + 2x_2 = 0, \\ 5x_1 + (4-\lambda_i)\,x_2 = 0, \end{cases}$$

$i = 1, 2$, каждая из которых, поскольку ее определитель равен нулю, сводится к одному уравнению.

При $\lambda = 6$ это — уравнение $5x_1 - 2x_2 = 0$, из которого находим: $x_1 : x_2 = 2 : 5$, и в качестве собственного вектора, соответствующего $\lambda = 6$, можно взять $a_1 = (2, 5)$ (или любой вектор, кратный a_1). При $\lambda = -1$ имеем уравнение $x_1 + x_2 = 0$, из которого $x_1 : x_2 = -1$, и соответствующий собственный вектор $a_2 = (1, -1)$ (или любой вектор, кратный ему).

Особенно простой вид принимает матрица линейного оператора, имеющего n линейно независимых собственных векторов. В самом деле, пусть линейный оператор $\mathscr{A}$ имеет n линейно независимых собственных векторов $e_1, e_2, \ldots, e_n$ с собственными значениями, соответственно равными $\lambda_1, \lambda_2, \ldots, \lambda_n$. Векторы $e_1, e_2, \ldots, e_n$ примем за базисные, тогда, ввиду равенств

$$\mathscr{A}e_i = \lambda_i e_i \quad \text{при} \quad i = 1, 2, \ldots, n$$

матрица оператора $\mathcal{A}$ будет иметь вид

$$A=\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}$$

(такая матрица называется *диагональной*). Верно и обратное: *если матрица A оператора $\mathcal{A}$ в некотором базисе является диагональной, то все векторы этого базиса являются собственными векторами оператора $\mathcal{A}$.*

Однако далеко не каждый линейный оператор в n-мерном векторном пространстве имеет n-линейно независимых собственных векторов. Один из случаев, когда можно утверждать, что базис из собственных векторов («собственный базис») существует, подсказывается следующей теоремой:

Теорема 7. *Собственные векторы линейного оператора, отвечающие попарно различным собственным значениям, линейно независимы.*

Доказательство проведем индукцией по числу рассматриваемых собственных векторов. Для одного вектора x это ясно, так как, по определению собственного вектора, он отличен от нуля (и значит, из равенства $\alpha x=0$ вытекает, что $\alpha=0$).

Пусть наше утверждение справедливо для $k-1$ векторов $x_1, x_2, \dots, x_{k-1}$, и предположим, что k собственных векторов

$$x_1, x_2, \dots, x_k,$$

отвечающих попарно различным собственным значениям $\lambda_1, \lambda_2, \dots, \lambda_k$, линейно зависимы:

$$\alpha_1 x_1+\alpha_2 x_2+\dots+\alpha_k x_k=0. \tag{9}$$

Применяя к обеим частям этого равенства оператор $\mathcal{A}$, получим

$$\alpha_1\mathcal{A}x_1+\alpha_2\mathcal{A}x_2+\dots+\alpha_k\mathcal{A}x_k=0,$$

или

$$\alpha_1\lambda_1 x_1+\alpha_2\lambda_2 x_2+\dots+\alpha_k\lambda_k x_k=0. \tag{10}$$

С другой стороны, умножая равенство (9) на λ_k, будем иметь

$$\alpha_1\lambda_k x_1+\alpha_2\lambda_k x_2+\dots+\alpha_k\lambda_k x_k=0. \tag{11}$$

Вычитая равенство (11) из равенства (10), получим

$$\alpha_1(\lambda_1-\lambda_k)x_1+\alpha_2(\lambda_2-\lambda_k)x_2+\ldots$$
$$\ldots+\alpha_{k-1}(\lambda_{k-1}-\lambda_k)x_{k-1}=0,$$

а так как, по условию, все λ_i различны и в силу предположения индукции векторы $x_1, x_2, \ldots, x_{k-1}$ линейно независимы, то $\alpha_1=\alpha_2=\ldots=\alpha_{k-1}=0$, а тогда из равенства (9) имеем $\alpha_k x_k=0$ и $\alpha_k=0$. Теорема доказана.

Таким образом, если *линейный оператор* $\mathscr{A}$ *имеет* n *попарно различных собственных значений, то отвечающие им собственные векторы линейно независимы, и матрица этого оператора в соответствующем базисе имеет диагональный вид.*

Поскольку многочлен с вещественными коэффициентами не обязательно имеет хотя бы один вещественный корень, то в вещественном пространстве не для всякого линейного оператора найдется хотя бы одно одномерное инвариантное подпространство. Однако имеет место следующая

Теорема 8. *Для всякого линейного оператора, действующего в вещественном пространстве размерности* $n>2$, *существует одномерное или двумерное инвариантное подпространство.*

Доказательство. Если характеристический многочлен оператора $\mathscr{A}$ имеет хотя бы один вещественный корень, то этот оператор имеет собственный вектор, и значит, в R существует одномерное инвариантное относительно $\mathscr{A}$ подпространство.

Если характеристический многочлен не имеет вещественных корней, мы сошлемся на так называемую основную теорему алгебры комплексных чисел:

Каждый многочлен с комплексными (в частности, с вещественными) *коэффициентами имеет хотя бы один* (комплексный) *корень.*

В силу этой теоремы (которую мы здесь не доказываем) характеристический многочлен, не имеющий вещественного корня, будет иметь хотя бы один комплексный корень $\lambda=\alpha+i\beta$, где $\beta\neq 0$.

Решая для этого λ систему уравнений (8), мы найдем соответствующие (комплексные) решения: $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, ..., $z_n = x_n + iy_n$, и значит, будут справедливы равенства

$$\begin{cases} a_{11}(x_1 + iy_1) + a_{12}(x_2 + iy_2) + \ldots + a_{1n}(x_n + iy_n) = \\ \qquad\qquad = (\alpha + i\beta)(x_1 + iy_1), \\ a_{21}(x_1 + iy_1) + a_{22}(x_2 + iy_2) + \ldots + a_{2n}(x_n + iy_n) = \\ \qquad\qquad = (\alpha + i\beta)(x_2 + iy_2), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}(x_1 + iy_1) + a_{n2}(x_2 + iy_2) + \ldots + a_{nn}(x_n + iy_n) = \\ \qquad\qquad = (\alpha + i\beta)(x_n + iy_n). \end{cases}$$

Приравнивая действительные и мнимые части, получим две системы равенств:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = \alpha x_1 - \beta y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = \alpha x_2 - \beta y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = \alpha x_n - \beta y_n \end{cases} \tag{12}$$

и

$$\begin{cases} a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n = \alpha y_1 + \beta x_1, \\ a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n = \alpha y_2 + \beta x_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n = \alpha y_n + \beta x_n. \end{cases} \tag{13}$$

Рассмотрим два (вещественных) вектора

$$u = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

и

$$v = y_1e_1 + y_2e_2 + \ldots + y_ne_n.$$

Равенства (12) показывают, что $\mathcal{A}u = \alpha u - \beta v$, а равенства (13) — что $\mathcal{A}v = \alpha v + \beta u$. Но тогда подпространство R_1, порожденное векторами u и v, инвариантно относительно $\mathcal{A}$, так как если $x \in R_1$, т. е. $x = \xi u + \eta v$, то и

$$\mathcal{A}x = \xi\mathcal{A}u + \eta\mathcal{A}v = \xi(\alpha u - \beta v) + \eta(\alpha v + \beta u) = \\ = (\alpha\xi + \beta\eta)u + (\alpha\eta - \beta\xi)v$$

принадлежит R_1. Это подпространство двумерно, так как если бы векторы u и v были линейно зависимыми:

$v = \gamma u$, то мы имели бы

$$\mathcal{A}u = \alpha u - \beta v = (\alpha - \beta\gamma) u,$$

и вектор u был бы собственным вектором оператора $\mathcal{A}$ с вещественным собственным значением $\alpha - \beta\gamma$.

§ 9. Спектр линейного оператора

Из сформулированной в предыдущем параграфе «основной теоремы алгебры» непосредственно следует, что *в комплексном векторном пространстве R каждый линейный оператор $\mathcal{A}$ имеет хотя бы один собственный вектор* и, значит, *в R существует одномерное инвариантное относительно $\mathcal{A}$ подпространство.*

Далее, из этой основной теоремы вытекает, что многочлен n-й степени (с комплексными коэффициентами) имеет в точности n (комплексных) корней, среди которых, впрочем, могут быть и равные. Действительно, пусть $f(t)$ — многочлен степени n и t_1 — его корень; тогда $f(t)$ делится на $t - t_1$, т. е. $f(t) = (t - t_1) f_1(t)$, где $f_1(t)$ — многочлен $(n-1)$-й степени, тоже с комплексными коэффициентами. Но $f_1(t)$ тоже имеет хотя бы один корень t_2, и тогда $f_1(t) = (t - t_2) f_2(t)$, откуда $f(t) = (t - t_1)(t - t_2) f_2(t)$, и т. д. Через n шагов мы получим равенство

$$f(t) = (t - t_1)(t - t_2) \ldots (t - t_n) c, \tag{14}$$

где $t_1, t_2, \ldots, t_n$ — корни многочлена $f(t)$, а c — число. Если множитель $t - t_m$ входит в разложение (14) k раз, то соответствующий корень t_m называется *корнем кратности k*, или *k-кратным корнем.*

Покажем теперь, что многочлен $f(t)$ не может иметь корней, отличных от $t_1, t_2, \ldots, t_n$, в частности, он *не может иметь более чем n корней.* Действительно, если t_0 — корень многочлена $f(t)$, то

$$f(t_0) = (t_0 - t_1)(t_0 - t_2) \ldots (t_0 - t_n) c = 0,$$

и значит, одна из разностей $t_0 - t_m = 0$, откуда $t_0 = t_m$, где $m = 1, 2, \ldots$ или n.

Пусть теперь $\varphi(\lambda)$ — характеристический многочлен линейного оператора $\mathcal{A}$ и $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все его корни (собственные значения оператора $\mathcal{A}$), причем *каждый* *из них взят столько раз, какова его кратность.* Мы видели выше (см. стр. 121), что

$$\varphi(\lambda) = (-1)^n[\lambda^n - (a_{11} + a_{22} + \ldots a_{nn})\lambda^{n-1} + \ldots]. \qquad (15)$$

С другой стороны,

$$\varphi(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n) = \\ = (-1)^n[\lambda^n - (\lambda_1 + \lambda_2 + \ldots + \lambda_n)\lambda^{n-1} + \ldots].$$

Следовательно, *сумма всех собственных значений* $\lambda_1 + \lambda_2 + \ldots + \lambda_n$ *оператора* $\mathcal{A}$ *равна следу*

$$a_{11} + a_{22} + \ldots + a_{nn} = \operatorname{tr} A$$

его матрицы. Но так как след $\operatorname{tr} A$ — это один из коэффициентов характеристического многочлена (см. стр. 121), то он не зависит от базиса и поэтому может быть назван *следом* самого *оператора* $\mathcal{A}$.

Легко видеть, что для любых двух линейных операторов $\mathcal{A}$ и $\mathcal{B}$

$$\operatorname{tr}(\mathcal{A} + \mathcal{B}) = \operatorname{tr}\mathcal{A} + \operatorname{tr}\mathcal{B}.$$

Покажем еще, что

$$\operatorname{tr}(\mathcal{A}\mathcal{B}) = \operatorname{tr}(\mathcal{B}\mathcal{A}).$$

Действительно, если $A = [a_{ik}]$, $B = [b_{ik}]$, то

$$\operatorname{tr}(\mathcal{A}\mathcal{B}) = \sum_{k=1}^{n} a_{1k}b_{k1} + \sum_{k=1}^{n} a_{2k}b_{k2} + \ldots + \sum_{k=1}^{n} a_{nk}b_{kn} = \\ = \sum_{l=1}^{n}\left(\sum_{k=1}^{n} a_{lk}b_{kl}\right) = \sum_{k=1}^{n}\left(\sum_{l=1}^{n}(b_{kl}a_{lk})\right) = \\ = \sum_{l=1}^{n} b_{1l}a_{l1} + \sum_{l=1}^{n} b_{2l}a_{l2} + \ldots + \sum_{l=1}^{n} b_{nl}a_{ln} = \operatorname{tr}(BA).$$

Из доказанного равенства, в частности, вытекает, что для любых $\mathcal{A}$, $\mathcal{B}$ (где $\mathcal{B}$ — невырожденный оператор) имеем

$$\operatorname{tr}(\mathcal{B}^{-1}\cdot\mathcal{A}\mathcal{B}) = \operatorname{tr}(\mathcal{A}\mathcal{B}\cdot\mathcal{B}^{-1}) = \operatorname{tr}(\mathcal{A}\cdot\mathcal{B}\mathcal{B}^{-1}) = \operatorname{tr}\mathcal{A}.$$

Отметим еще несколько свойств собственных значений.

Пусть $\mathcal{A}$ — линейный оператор, λ_1 — его собственное значение и x — соответствующий собственный вектор, Тогда $\mathcal{A}x = \lambda x$. Применяя к обеим частям этого равенства оператор $\mathcal{A}$, получим $\mathcal{A}(\mathcal{A}x) = \mathcal{A}(\lambda x)$, или $\mathcal{A}^2x = \lambda\mathcal{A}x = \lambda^2 x$, т. е. λ^2 — собственное значение оператора $\mathcal{A}^2$ для того же собственного вектора x. Аналогично показывается, что при любом натуральном k число λ^k есть собственное значение оператора $\mathcal{A}^k$, и для любого многочлена $f(t)$ число $f(\lambda)$ — собственное значение оператора $f(\mathcal{A})$, отвечающее тому же собственному вектору x. Можно доказать и такую более общую теорему: *если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все собственные значения оператора $\mathcal{A}$, взятые с учетом их кратностей, и $f(t)$ — произвольный многочлен, то $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ — это все собственные значения оператора $f(\mathcal{A})$, причем $f(\lambda_i)$ взято столько раз, какова кратность λ_i.*

Далее, если оператор $\mathcal{A}$ — невырожденный, то, применяя к обеим частям равенства $\mathcal{A}x = \lambda x$ оператор $\mathcal{A}^{-1}$, получим $\mathcal{A}^{-1}(\mathcal{A}x) = \mathcal{A}^{-1}(\lambda x)$, или $x = \lambda\mathcal{A}^{-1}x$, откуда $\mathcal{A}^{-1}x = \lambda^{-1}x$, т. е. λ^{-1} является собственным значением оператора $\mathcal{A}^{-1}$ с тем же собственным вектором x ($\lambda \neq 0$, так как оператор $\mathcal{A}$ — невырожденный). Мы видим, что действиям над линейными операторами отвечают соответствующие действия над их собственными значениями. Поэтому набор этих чисел — собственных значений оператора $\mathcal{A}$ в каком-то смысле определяет этот оператор. *Множество всех собственных значений линейного оператора $\mathcal{A}$ называется его* ***спектром***.

§ 10. Жорданова нормальная форма

Этот параграф, несколько более трудный, чем остальные, дальше не используется и при первом чтении может быть пропущен.

Пусть $\mathcal{A}$ — линейный оператор, действующий в векторном пространстве R размерности n. Как было показано в § 8, если в R найдется n линейно независимых собственных векторов оператора $\mathcal{A}$ то в базисе, состоящем из этих векторов, матрица оператора $\mathcal{A}$ приводится к наиболее простому — диагональному виду

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots\ldots & 0 \\ 0 & \lambda_2 & \ldots\ldots & 0 \\ \cdot & \cdot & \cdot\;\cdot\;\cdot\;\cdot & \cdot \\ 0 & 0 & \ldots\ldots & \lambda_n \end{bmatrix},$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения $\mathcal{A}$. Так будет, в частности, в том случае, если характеристический многочлен оператора $\mathcal{A}$ имеет *n попарно различных* корней (см. стр. 123); так будет, как мы увидим ниже, и в случае любого так называемого самосопряженного оператора (как в комплексном, так и в вещественном евклидовом пространстве; см. главу V, стр. 170), и в случае любого унитарного оператора в комплексном евклидовом векторном пространстве (стр. 182).

Однако, как уже тоже было сказано выше (§ 8), к такому простому, диагональному виду приводится матрица далеко не всякого линейного оператора. Рассмотрим, например, линейный оператор с матрицей $A = \begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix}$ в некотором базисе e_1, e_2; характеристический многочлен его $\varphi(\lambda) = (2-\lambda)^2$ имеет два одинаковых корня $\lambda_1 = \lambda_2 = 2$. Собственные векторы этого оператора определяются уравнением $0x_1 + 1x_2 = 0$ (где x_1 и x_2 — координаты вектора), или $x_2 = 0$ — это только векторы, коллинеарные e_1 (ср. стр. 338). Поэтому не существует базиса, образованного собственными векторами оператора $\mathcal{A}$, и, значит, его матрица ни в каком базисе не приводится к диагональному виду.

Поэтому возникает вопрос о каком-то другом, достаточно простом виде, к которому можно привести матрицу всякого линейного оператора. В комплексном пространстве таким «простейшим», каноническим видом принято считать так называемую жорданову форму матрицы.

Определение 5. *Жордановой клеткой называется квадратная матрица вида*

$$I_{\lambda_0} = \begin{bmatrix} \lambda_0 & 1 & 0 & \ldots\ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots\ldots & 0 \\ 0 & 0 & & \ldots\ldots & 0 \\ \cdot & \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot\cdot & \cdot \\ 0 & 0 & 0 & \ldots\ldots & \lambda_0 \end{bmatrix}, \tag{16}$$

в которой на главной диагонали стоит одно и то же число λ_0, над главной диагональю — всюду число 1, а все остальные элементы матрицы — нулевые.

Порядок жордановой клетки может быть каким угодно. В частности, он может быть равен и 1; в этом случае клетка имеет простейший вид: $I_{\lambda_0} = [\lambda_0]$. Легко видеть, что характеристический многочлен оператора $\mathcal{I}_{\lambda_0}$, матрицей которого служит жорданова клетка (16) порядка n, равен $(\lambda_0 - \lambda)^n$; он имеет одно собственное значение λ_0 кратности n, и все его собственные векторы коллинеарны e_1. Матрица оператора $\mathcal{I}_{\lambda_0}$ при $n > 1$ ни в каком базисе не приводится к диагональному виду (ср. с примером на стр. 129).

Определение 6. *Жордановой матрицей называется матрица вида*

$$I = \left[\begin{array}{c:c:c:c} I_{\lambda_1} & & & \\ \hdashline & I_{\lambda_2} & & \\ \hdashline & & & \\ \hdashline & & & I_{\lambda_s} \end{array}\right],$$

где I_{λ_k} $(k = 1, 2, \ldots, s)$ — *жордановы клетки (вообще говоря, разных порядков), а все остальные клетки этой матрицы — нулевые* (т. е. состоят из одних нулей).

Легко видеть, что числа $\lambda_1, \lambda_2, \ldots, \lambda_s$ являются собственными значениями оператора $\mathcal{I}$ с матрицей I. Конечно, эти значения не обязательно должны быть разными, некоторые из них могут и совпадать.

Рассмотрим пример жордановой матрицы пятого порядка:

$$I = \begin{bmatrix} \alpha & 1 & 0 & 0 & 0 \\ 0 & \alpha & 1 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & \beta & 1 \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix}.$$

Эта матрица состоит из двух жордановых клеток — третьего и второго порядков. Числа α и β являются собственными значениями оператора $\mathcal{I}$, характеристический многочлен которого равен $(\alpha - \lambda)^3 (\beta - \lambda)^2$. Если e_1, e_2, e_3, e_4, e_5 — базис, соответствующий матрице I, то имеем, очевидно,

$$\mathcal{A}e_1 = \alpha e_1, \quad \mathcal{A}e_2 = \alpha e_2 + e_1, \quad \mathcal{A}e_3 = \alpha e_3 + e_2,$$
$$\mathcal{A}e_4 = \beta e_4, \quad \mathcal{A}e_5 = \beta e_5 + e_4.$$

Базисные векторы e_1 и e_4 являются собственными векторами оператора $\mathcal{J}$ с собственными значениями α и β соответственно. Их можно назвать векторами нулевого слоя. Векторы e_2 и e_5 являются «собственными с точностью до векторов нулевого слоя» e_1 и e_4. Это значит, что, скажем, $\mathcal{A}e_2$ отличается от αe_2 лишь на вектор e_1, при пренебрежении которым вектор e_2 можно считать собственным. Эти векторы e_2 и e_5 можно назвать поэтому векторами первого слоя. Аналогично, вектор e_3 является «собственным с точностью до вектора e_2», т. е. с точностью до вектора первого слоя. Можно сказать поэтому, что это — вектор второго слоя.

Нашей ближайшей целью будет доказательство следующей важной теоремы.

Теорема 9. *Матрицу всякого линейного оператора, действующего в комплексном векторном пространстве, можно привести к жордановой форме.*

Это значит, что базис векторного пространства, в котором действует рассматриваемый оператор, можно выбрать так, что матрица оператора в этом базисе будет жордановой матрицей.

Для доказательства теоремы 9 нам понадобится следующая

Лемма. *Пусть $\mathcal{B}$ — линейный оператор, действующий в векторном пространстве R размерности n, N — его ядро и $\mathcal{B}R$ — область значений* (см. § 5). *Обозначим через M пересечение этих подпространств, т. е. пусть $M = N \cap \mathcal{B}R$. Выберем в M базис $x_1, x_2, \ldots, x_m$, дополним его до базиса $x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_k$ ядра N, с одной стороны, и до базиса $x_1, x_2, \ldots, x_m, x'_{m+1}, \ldots, x'_r$ подпространства $\mathcal{B}R$ — с другой. Пусть далее $y_1, y_2, \ldots, y_m$ — прообразы векторов $x_1, x_2, \ldots, x_m$ и $z_{m+1}, z_{m+2}, \ldots, z_r$ — прообразы векторов $x'_{m+1}, x'_{m+2}, \ldots, x'_r$ при преобразовании $\mathcal{B}$.* (Это значит, что $\mathcal{B}y_i = x_i$ при $i = 1, 2, \ldots, m$ и $\mathcal{B}z_i = x'_i$ при $i = m+1, m+2, \ldots, r$. Схематически это можно изобразить так:

$$
\begin{array}{ccccl}
 & & & \diagup & x_{m+1}, x_{m+2}, \ldots, x_k \text{ базис } N, \\
x_1, & x_2, & \ldots, & x_m & \\
 & & & \diagdown & x'_{m+1}, x'_{m+2}, \ldots, x'_r \text{ базис } \mathcal{B}R.) \\
\uparrow & \uparrow & & \uparrow & \uparrow \qquad\quad \uparrow \qquad\qquad \uparrow \\
y_1, & y_2, & \ldots, & y_m & z_{m+1}, z_{m+2}, \ldots, z_r
\end{array}
$$

Тогда векторы

$$x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_k, y_1, y_2, \ldots, y_m, z_{m+1}, z_{m+2}, \ldots, z_r \quad (17)$$

образуют базис пространства R.

Доказательство. Число векторов системы (17) равно $k+r$; при этом размерность r подпространства $\mathcal{B}R$ равна рангу матрицы B (см. § 5), а размерность k ядра N (дефект оператора $\mathcal{B}$) равна $n-r$ (теорема 4). Значит, число $k+r$ векторов системы (17) равно размерности n пространства R, — и нам остается доказать, что эти векторы линейно независимы. Предположим, что они линейно зависимы, т. е. что имеет место равенство

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k + \beta_1 y_1 + \ldots \\ \ldots + \beta_m y_m + \beta_{m+1} z_{m+1} + \ldots + \beta_r z_r = 0. \quad (18)$$

Применим к обеим частям этого равенства оператор $\mathcal{B}$:

$$\alpha_1 \mathcal{B} x_1 + \alpha_2 \mathcal{B} x_2 + \ldots + \alpha_k \mathcal{B} x_k + \\ + \beta_1 \mathcal{B} y_1 + \ldots + \beta_m \mathcal{B} y_m + \beta_{m+1} \mathcal{B} z_{m+1} + \ldots + \beta_r \mathcal{B} z_r = 0.$$

Но так как $\mathcal{B} x_i = 0$ при $i = 1, 2, \ldots, k$, $\mathcal{B} y_i = x_i$ при $i = 1, 2, \ldots, m$ и $\mathcal{B} z_i = x_i'$ при $i = m+1, \ldots, r$, мы имеем

$$\beta_1 x_1 + \ldots + \beta_m x_m + \beta_{m+1} x'_{m+1} + \ldots + \beta_r x'_r = 0.$$

Однако векторы $x_1, x_2, \ldots, x_m, x'_{m+1}, \ldots, x'_r$ образуют базис $\mathcal{B}R$ и, значит, линейно независимы, т. е. $\beta_i = 0$ при $i = 1, 2, \ldots, r$. Поэтому из равенства (18) следует, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0,$$

и ввиду линейной независимости векторов $x_1, x_2, \ldots, x_k$ также и $\alpha_i = 0$ при $i = 1, 2, \ldots, k$.

Так как векторы системы (17) линейно независимы, а число их равно размерности пространства R, то они образуют базис R.

Теперь мы можем перейти к доказательству теоремы 9. Доказательство это мы разобьем на несколько шагов.

I. Сначала мы рассмотрим частный случай, когда линейный оператор $\mathcal{A}$, представляемый в некотором базисе матрицей $A=[a_{ik}]$, имеет только *одно* собственное значение α, т. е. когда его характеристический многочлен имеет вид $(\alpha-\lambda)^n$.

1. Введем некоторые обозначения.

Пусть x — собственный вектор оператора $\mathcal{A}$, тогда $\mathcal{A}x=\alpha x$, или $(\mathcal{A}-\alpha\mathcal{E})x=0$. Обозначим оператор $\mathcal{A}-\alpha\mathcal{E}$ через $\mathcal{A}_\alpha$. Тогда $\mathcal{A}_\alpha x=(\mathcal{A}-\alpha\mathcal{E})x=0$. Таким образом, оператор $\mathcal{A}_\alpha$ переводит в нуль, «аннулирует» каждый собственный вектор (и, не считая нуля,— только собственные векторы) оператора $\mathcal{A}$. Обозначим через N_1 ядро оператора $\mathcal{A}_\alpha$ — оно состоит из всех собственных векторов оператора $\mathcal{A}$ с добавлением нуля.

Далее, обозначим R_1 область значений $\mathcal{A}_\alpha R$ оператора $\mathcal{A}_\alpha$. Подпространство R_1 инвариантно относительно $\mathcal{A}$, так как $\mathcal{A}R_1=\mathcal{A}(\mathcal{A}_\alpha R)=\mathcal{A}_\alpha(\mathcal{A}R)\subseteq\subseteq\mathcal{A}_\alpha R=R_1$ (легко видеть, что $\mathcal{A}\mathcal{A}_\alpha=\mathcal{A}_\alpha\mathcal{A}$).

Так как R — векторное пространство над полем **комплексных** чисел, то в (инвариантном) подпространстве R_1 (если $R_1\neq 0$) найдется собственный вектор оператора $\mathcal{A}$. Этот вектор, поскольку все собственные значения оператора $\mathcal{A}$ равны α, аннулируется оператором $\mathcal{A}_\alpha$ — и значит, он принадлежит ядру N_2 оператора $\mathcal{A}_\alpha$, рассматриваемого в R_1. Но если ядро N_2 оператора $\mathcal{A}_\alpha$ в R_1 имеет ненулевую размерность, то его область значений $R_2=\mathcal{A}_\alpha R_1=\mathcal{A}_\alpha^2 R$ имеет размерность, меньшую размерности R_1, и значит, включение $R_2==\mathcal{A}_\alpha^2 R\subset\mathcal{A}_\alpha R=R_1$ — строгое.

Заметим, что N_2 совпадает с пересечением $N_1\cap R_1$, ибо N_2 состоит из всех тех и только тех векторов подпространства R_1, которые аннулируются оператором $\mathcal{A}_\alpha$; но из тех же векторов состоит и пересечение $N_1\cap R_1$.

Продолжим это построение. Если подпространства R_p и N_p, где $p=1, 2, \ldots$, уже определены, обозначим через N_{p+1} ядро и через R_{p+1} область значений оператора $\mathcal{A}_\alpha$ в подпространстве R_p. Если $R_p\neq 0$, то ядро N_{p+1} имеет ненулевую размерность (ибо наше пространство — комплексное!), и значит, $R_{p+1}\subset R_p$. Пересечение же $N_p\cap R_p=N_{p+1}$.

Так мы получим (строго) убывающую цепочку подпространств

$$R\supset R_1=\mathcal{A}_\alpha R\supset R_2=\mathcal{A}_\alpha^2 R\supset R_3=\mathcal{A}_\alpha^3 R\supset\ldots,$$

которая, поскольку размерности этих подпространств убывают, должна закончиться нулем. Пусть $R_{k+1} = \mathscr{A}_\alpha^{k+1}R = 0$, но $R_k = \mathscr{A}_\alpha^k R \neq 0$.

Поскольку $R_{k+1} = 0$, то ядро N_{k+1} оператора $\mathscr{A}_\alpha$ в R_k совпадает с R_k. Но из того, что $N_k \cap R_k = N_{k+1} = R_k$, вытекает, что $R_k \subseteq N_k$.

2. Теперь мы будем строить искомый базис пространства R, начиная с базиса подпространства R_k, и, переходя последовательно от R_k к R_{k-1}, от R_{k-1} к R_{k-2}, и т. д., постепенно дополним его до базиса всего пространства R.

Итак, выберем в подпространстве R_k базис $x_1, x_2, \ldots, x_{p_1}$. Перейдем к подпространству R_{k-1}; так как $R_k \subseteq N_k$, то мы можем дополнить базис $x_1, x_2, \ldots, x_{p_1}$ подпространства R_k до базиса $x_1, x_2, \ldots, x_{p_1}, x_{p_1+1}, \ldots, x_{p_2}$ ядра N_k. Пусть $y_1, y_2, \ldots, y_{p_1}$ — те векторы из R_{k-1}, которые оператором $\mathscr{A}_\alpha$ переводятся соответственно в векторы $x_1, x_2, \ldots, x_{p_1}$. Другими словами, $y_1, \ldots, y_{p_1}$ — это прообразы векторов $x_1, x_2, \ldots, x_{p_1}$: $\mathscr{A}_\alpha y_i = x_i$ при $i = 1, 2, \ldots, p_1$. Схематически:

$$\begin{array}{cccccccc} x_1, & x_2, & \ldots, & x_{p_1}, & x_{p_1+1}, & \ldots, & x_{p_2}. \\ \uparrow & \uparrow & & \uparrow & & & \\ y_1 & y_2 & & y_{p_1} & & & \end{array}$$

Как видно из леммы (стр. 131), векторы

$$x_1, x_2, \ldots, x_{p_1}, x_{p_1+1}, \ldots, x_{p_2}, y_1, y_2, \ldots, y_{p_1}$$

образуют базис подпространства R_{k-1} (размерность которого равна, следовательно, $p_1 + p_2$).

Перейдем далее к подпространству R_{k-2}. В пересечении $N_k = N_{k-1} \cap R_{k-1}$ уже построен базис $x_1, x_2, \ldots, x_{p_1}, x_{p_1+1}, \ldots, x_{p_2}$. Дополним эту систему векторов до базиса

$$x_1, x_2, \ldots, x_{p_1}, x_{p_1+1}, \ldots, x_{p_2}, x_{p_2+1}, \ldots, x_{p_3}$$

ядра N_{k-1}. С другой стороны, векторы

$$x_1, x_2, \ldots, x_{p_1}, x_{p_1+1}, \ldots, x_{p_2}, y_1, y_2, \ldots, y_{p_1}$$

образуют б а з и с подпространства R_{k-1}. Пусть $y_{p_1+1}, \ldots, y_{p_2}$ — те векторы (из R_{k-1}), которые оператором $\mathscr{A}_\alpha$

переводятся соответственно в векторы $x_{p_1+1}, \ldots, x_{p_2}$, а $z_1, z_2, \ldots, z_{p_1}$ — векторы (из R_{k-2}), которые переводятся в векторы $y_1, y_2, \ldots, y_{p_1}$. Схематически это можно изобразить так:

$$\begin{array}{llll}
 & & \nearrow x_{p_2+1}, \ldots, x_{p_3} & \text{базис } N_{k-1}, \\
x_1, x_2, \ldots, x_{p_1}, & x_{p_1+1}, \ldots, x_{p_2} & & \\
 & \uparrow \qquad\qquad \uparrow & \searrow y_1, \ldots, y_{p_1} & \text{базис } R_{k-1}. \\
 & y_{p_1+1}, \ldots, y_{p_2} & \quad \uparrow \qquad\quad \uparrow & \\
 & & \quad z_1, \ldots, z_{p_1} &
\end{array}$$

По лемме векторы

$$x_1, x_2, \ldots, x_{p_1}, x_{p_1+1}, \ldots, x_{p_2}, x_{p_2+1}, \ldots, x_{p_3},$$

$$y_1, y_2, \ldots, y_{p_1}, y_{p_1+1}, \ldots, y_{p_2}, z_1, z_2, \ldots, z_{p_1}$$

образуют б а з и с подпространства R_{k-2}. При этом очевидно, что

$\mathcal{A}_\alpha x_i = 0$ при $i = 1, 2, \ldots, p_1, p_1 + 1, \ldots, p_2, p_2 + 1, \ldots, p_3$;
$\mathcal{A}_\alpha y_i = x_i$ при $i = 1, 2, \ldots, p_1, p_1 + 1, \ldots, p_2$;
$\mathcal{A}_\alpha z_i = y_i$ при $i = 1, 2, \ldots, p_1$.

Это построение мы продолжаем до тех пор, пока не получим базис всего пространства R. В нем все векторы x_i будут векторами нулевого слоя (это — собственные векторы оператора $\mathcal{A}$), все y_i — векторами первого слоя, z_i — векторами второго слоя, и т. д.

Для ясности рассмотрим подробнее частный случай, когда $k = 3$, $p_1 = 2$, $p_2 = 2$, $p_3 = 5$, $p_4 = 7$. Здесь базис пространства R мы получим в виде

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7, y_1, y_2, y_3, y_4, y_5, z_1, z_2, u_1, u_2.$$

При этом $\mathcal{A}_\alpha x_i = 0$ при $i = 1, 2, \ldots, 7$; $\mathcal{A}_\alpha y_i = x_i$ при $i = 1, 2, 3, 4, 5$; $\mathcal{A}_\alpha z_i = y_i$ при $i = 1, 2$; $\mathcal{A}_\alpha u_i = z_i$ при $i = 1, 2$, или $\mathcal{A} x_i = \alpha x_i$ при $i = 1, 2, \ldots, 7$; $\mathcal{A} y_i = \alpha y_i + x_i$ при $i = 1, 2, 3, 4, 5$, $\mathcal{A} z_i = \alpha z_i + y_i$ при $i = 1, 2$; $\mathcal{A} u_i = \alpha u_i + z_i$ при $i = 1, 2$.

Расположим теперь эти базисные векторы в порядке

$$x_1, y_1, z_1, u_1, x_2, y_2, z_2, u_2, x_3, y_3, x_4, y_4, x_5, y_5, x_6, x_7.$$

Легко видеть, что в этом базисе матрица оператора $\mathcal{A}$

приведется к жордановой форме:

$$
I=\left[\begin{array}{cccc|cccc|cc|cc|cc|c|c}
\alpha & 1 & 0 & 0 & & & & & & & & & & & & \\
0 & \alpha & 1 & 0 & & & & & & & & & & & & \\
0 & 0 & \alpha & 1 & & & & & & & & & & & & \\
0 & 0 & 0 & \alpha & & & & & & & & & & & & \\
\hline
 & & & & \alpha & 1 & 0 & 0 & & & & & & & & \\
 & & & & 0 & \alpha & 1 & 0 & & & & & & & & \\
 & & & & 0 & 0 & \alpha & 1 & & & & & & & & \\
 & & & & 0 & 0 & 0 & \alpha & & & & & & & & \\
\hline
 & & & & & & & & \alpha & 1 & & & & & & \\
 & & & & & & & & 0 & \alpha & & & & & & \\
\hline
 & & & & & & & & & & \alpha & 1 & & & & \\
 & & & & & & & & & & 0 & \alpha & & & & \\
\hline
 & & & & & & & & & & & & \alpha & 1 & & \\
 & & & & & & & & & & & & 0 & \alpha & & \\
\hline
 & & & & & & & & & & & & & & \alpha & \\
\hline
 & & & & & & & & & & & & & & & \alpha
\end{array}\right].
$$

Все невыписанные элементы матрицы I равны нулю.

В общем случае доказательство завершается аналогично.

Легко видеть, что в матрице I будет p_1 клеток порядка $k+1$, p_2-p_1 клеток порядка k, p_3-p_2 клеток порядка $k-1$, ... и, наконец, $p_{k+1}-p_k$ клеток порядка 1. (Конечно, не исключено, что для некоторых i $p_{i+1}=p_i$, и тогда клетки соответствующих порядков будут отсутствовать.) Общее же число жордановых клеток равно

$$p_1+(p_2-p_1)+(p_3-p_2)+\ldots+(p_{k+1}-p_k)=p_{k+1},$$

а размерность n всего пространства R равна

$$(k+1)p_1+k(p_2-p_1)+(k-1)(p_3-p_2)+\ldots$$
$$\ldots+1(p_{k+1}-p_k)=p_1+p_2+\ldots+p_{k+1}.$$

Числа p_1, p_1+p_2, $p_1+p_2+p_3$, ..., $p_1+p_2+\ldots+p_k$ — это размерности подпространств $R_k=\mathscr{A}_\alpha^k R$, $R_{k-1}=\mathscr{A}_\alpha^{k-1}R, \ldots, R_1=\mathscr{A}_\alpha R$ — они равны соответственно рангам матриц $\mathscr{A}_\alpha^k$, $\mathscr{A}_\alpha^{k-1}$, ..., $\mathscr{A}_\alpha$. Обозначив ранг матрицы $\mathscr{A}_\alpha^i$ через r_i (и через r_0 ранг единичной матрицы порядка n, т. е. полагая $r_0 = n$), будем иметь

$$p_1 = r_k,\ p_1 + p_2 = r_{k-1},$$
$$p_1 + p_2 + p_3 = r_{k-2}, \ldots, p_1 + p_2 + \ldots + p_k = r_1,$$

откуда $p_k = r_1 - r_2$, $p_{k-1} = r_2 - r_3$, ..., $p_2 = r_{k-1} - r_k$, $p_1 = r_k (= r_k - r_{k+1}$, так как $r_{k+1} = 0)$. Ранги матриц $\mathcal{A}_\alpha^k, \mathcal{A}_\alpha^{k-1}, \ldots, \mathcal{A}_\alpha$ можно найти непосредственно, по ним определяются числа $p_1, p_2, \ldots, p_k$,— а значит и вид искомой жордановой матрицы I.

Прежде чем перейти к доказательству теоремы в общем случае, рассмотрим пример. Пусть оператор $\mathcal{A}$ в некотором базисе e_1, e_2, e_3 имеет матрицу

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Его характеристический многочлен

$$\varphi(\lambda) = \begin{vmatrix} 2-\lambda & 1 & -1 \\ 0 & 3-\lambda & -1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} 3-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix} = (2-\lambda)^3.$$

Собственное значение здесь одно, равное 2, кратности 3. Матрица оператора $\mathcal{A}_2 = \mathcal{A} - 2\mathcal{E}$:

$$A_2 = \begin{vmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{vmatrix}.$$

Ясно, что $A_2^2 = 0$, и значит, здесь $k = 1$, $r_0 = 3$, $r_1 = 1$, $r_2 = 0$. Так как $p_1 = r_1 = 1$ и $p_2 = r_0 - r_1 = 2$, жорданова форма нашей матрицы будет содержать одну жорданову клетку порядка 2 и одну — порядка 1, т. е. это будет матрица

$$I = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Далее, если нам надо найти и новый базис, заметим, что $\mathcal{A}_2 e_1 = 0$, $\mathcal{A}_2 e_2 = (1, 1, 1)$, $\mathcal{A}_2 e_3 = (-1, -1, -1)$. Следовательно, образ $R_1 = \mathcal{A}_2 R$ пространства R — это одномерное подпространство с (1, 1, 1). Ядро N_1 оператора $\mathcal{A}_2$ определяется уравнением $\xi_2 - \xi_3 = 0$ (здесь ξ_2 и ξ_3 — координаты соответствующего вектора). Оно двумерно. Его базис образуют, например, векторы (1, 0, 0) и (1, 1, 1). При этом $R_1 \subset N_1$.

В качестве «жорданова базиса» можно взять, следовательно, векторы $x_1 = (1, 1, 1)$, $x_2 = (1, 0, 0)$ (вектор x_2 дополняет базис R_1 до базиса N_1) и $y_1 = (0, 1, 0)$ (y_1 — это прообраз вектора x_1 при преобразовании $\mathcal{A}_2$).

Для контроля проделаем следующую выкладку. В нашем случае матрица C перехода от старого базиса e_1, e_2, e_3 к новому x_1, y_1, x_2 такова:

$$C=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}, \quad \text{и значит,} \quad C^{-1}=\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & -1\\ 1 & 0 & -1\end{bmatrix}.$$

Вычислим произведение $C^{-1}AC$:

$$\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & -1\\ 1 & 0 & -1\end{bmatrix}\begin{bmatrix}2 & 1 & -1\\ 0 & 3 & -1\\ 0 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}; \quad \text{оно равно} \quad \begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix}$$

— найденной выше жордановой матрице.

II. Перейдем теперь к доказательству теоремы 9 в общем случае.

1. Пусть $\mathcal{A}$ — произвольный линейный оператор, действующий в пространстве R размерности n над полем комплексных чисел, α — одно из его собственных значений и $e_1, e_2, \ldots, e_k$ — базис ядра N_1 оператора $\mathcal{A}_\alpha$. Дополним эту систему векторов до базиса

$$e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_n \tag{19}$$

всего пространства R. В базисе (19) матрица оператора $\mathcal{A}$ имеет вид

$$A=\begin{bmatrix}\alpha & 0 & \ldots & 0 & a_{1,k+1} & \ldots & a_{1n}\\ 0 & \alpha & \ldots & 0 & a_{2,k+1} & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & \alpha & a_{k,k+1} & \ldots & a_{k,n}\\ 0 & 0 & \ldots & 0 & a_{k+1,k+1} & \ldots & a_{k+1,n}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & 0 & a_{n,k+1} & \ldots & a_{nn}\end{bmatrix}.$$

Характеристический многочлен $\varphi(\lambda)$ оператора $\mathcal{A}$, очевидно, таков:

$$\varphi(\lambda)=(\alpha-\lambda)^k\begin{vmatrix}a_{k+1,k+1}-\lambda & a_{k+1,k+2} & \ldots & a_{k+1,n}\\ a_{k+2,k+1} & a_{k+2,k+2}-\lambda & \ldots & a_{k+2,n}\\ \cdot & \cdot & \cdot & \cdot\\ a_{n,k+1} & a_{n,k+2} & \ldots & a_{nn}-\lambda\end{vmatrix}=$$

$$=(\alpha-\lambda)^k\psi(\lambda), \quad \text{где} \quad \psi(\lambda)=\begin{vmatrix}a_{k+1,k+1}-\lambda & \ldots & a_{k+1,n}\\ a_{k+2,k+1} & \ldots & a_{k+2,n}\\ \cdot & \cdot & \cdot\\ a_{n,k+1} & \ldots & a_{nn}-\lambda\end{vmatrix}.$$

Рассмотрим образ $R_1 = \mathscr{A}_\alpha R$ пространства R при преобразовании $\mathscr{A}_\alpha$. Так как ядро N_1 оператора $\mathscr{A}_\alpha$ в R имеет размерность k, то размерность r подпространства R_1 равна $n-k$. За базис R_1 можно принять любые r линейно независимых образов элементов исходного базиса $e_1, e_2, \ldots, e_n$ при отображении $\mathscr{A}_\alpha$, т. е. любые r линейно независимых векторов из $\mathscr{A}_\alpha e_1$, $\mathscr{A}_\alpha e_2$, $\ldots$, $\mathscr{A}_\alpha e_n$ (см. стр. 114), или, что то же самое, любые r линейно независимых столбцов матрицы $\mathscr{A}_\alpha$. Но матрица эта имеет вид

$$A_\alpha = \begin{bmatrix} 0 \ldots\ldots 0 & a_{1,k+1} & \ldots\ldots\ldots a_{1n} \\ 0 \ldots\ldots 0 & a_{2,k+1} & \ldots\ldots\ldots a_{2n} \\ \cdots & \cdots & \cdots \\ 0 \ldots\ldots 0 & a_{k,k+1} & \ldots\ldots\ldots a_{kn} \\ 0 \ldots\ldots 0 & a_{k+1,k+1}-\alpha & \ldots\ldots\ldots a_{k+1,n} \\ \cdots & \cdots & \cdots \\ 0 \ldots\ldots 0 & a_{n,k+1} & \ldots\ldots\ldots a_{nn}-\alpha \end{bmatrix}.$$

Так как первые k столбцов этой матрицы — нулевые, то последние $n-k=r$ столбцов ее линейно независимы; следовательно, они и образуют искомый базис R_1. Обозначим векторы этого базиса через $g_{k+1}, g_{k+2}, \ldots, g_n$ (таким образом, $g_i = \mathscr{A}_\alpha e_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{i-1,i}e_{i-1} + (a_{ii}-\alpha)e_i + a_{i+1,i}e_{i+1} + \ldots + a_{ni}e_n$ при $i = k+1, k+2, \ldots, n$).

Найдем матрицу оператора $\mathscr{A}$ в базисе $g_{k+1}, g_{k+2}, \ldots, g_n$ пространства R_1. Для этого надо найти образы $\mathscr{A}g_i$ базисных векторов g_i при действии оператора $\mathscr{A}$. Но $\mathscr{A}g_i = \mathscr{A}(\mathscr{A}_\alpha e_i) = \mathscr{A}_\alpha \mathscr{A} e_i = \mathscr{A}_\alpha (a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n) = a_{1i}\mathscr{A}_\alpha e_1 + a_{2i}\mathscr{A}_\alpha e_2 + \ldots + a_{ni}\mathscr{A}_\alpha e_n$. А так как $\mathscr{A}_\alpha e_i = 0$ при $i = 1, 2, \ldots, k$ и $\mathscr{A}_\alpha e_i = g_i$ при $i = k+1, k+2, \ldots, n$, то

$$\mathscr{A}g_i = a_{k+1,i}\, g_{k+1} + a_{k+2,i}\, g_{k+2} + \ldots + a_{ni} g_n,$$

и значит, матрицей оператора $\mathscr{A}$ в подпространстве R_1 размерности $r = n-k$ (в базисе $g_{k+1}, g_{k+2}, \ldots, g_n$) является клетка

$$\begin{bmatrix} a_{k+1,k+1} & a_{k+1,k+2} & \ldots & a_{k+1,n} \\ a_{k+2,k+1} & a_{k+2,k+2} & \ldots & a_{k+2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,k+1} & a_{n,k+2} & \ldots & a_{nn} \end{bmatrix}.$$

Как видно из п. 1, характеристический многочлен $\psi(\lambda)$ оператора $\mathscr{A}$ в R_1 равен $\frac{\varphi(\lambda)}{(\alpha-\lambda)^k}$, где $\varphi(\lambda)$ — характеристический многочлен оператора $\mathscr{A}$ в R. Он может все еще иметь корень, равный α, но кратность этого корня будет на k единиц меньше, чем кратность того же корня для оператора $\mathscr{A}$ в пространстве R. В то же время ясно, что при переходе к подпространству R_1 все остальные собственные значения оператора $\mathscr{A}$ не меняются и не изменяют своих кратностей.

Если оператор $\mathscr{A}$ в R_1 имеет собственное значение, равное α, то точно так же, как выше, переходя к подпространству $R_2 = \mathscr{A}_\alpha^2 R$, мы можем еще понизить кратность корня α, не меняя кратностей остальных собственных значений. Продолжая это построение, мы придем, в конце концов, к подпространству $R_s = \mathscr{A}_\alpha^s R$, в котором оператор $\mathscr{A}$ совсем не имеет собственных значений, равных α. В этом случае дефект оператора $\mathscr{A}_\alpha$ в R_s равен 0, и значит, ранг его равен размерности этого подпространства, т. е. $\mathscr{A}_\alpha R_s = R_s$ — оператор $\mathscr{A}_\alpha$, рассматриваемый в подпространстве R_s, является невырожденным. В этом случае ранг матрицы $\mathscr{A}_\alpha^{s+1}$ совпадает с рангом матрицы $\mathscr{A}_\alpha^s$ (см. стр. 116).

2. Предположим, что оператор $\mathscr{A}$ имеет собственные значения $\alpha_1, \alpha_2, \ldots, \alpha_p$ с кратностями, соответственно равными $k_1, k_2, \ldots, k_p$. Применяя надлежащее число раз описанный в п. 1 прием, мы можем построить подпространство

$$\mathscr{A}_{\alpha_1}^{s_1}\mathscr{A}_{\alpha_2}^{s_2}\ldots\mathscr{A}_{\alpha_{i-1}}^{s_{i-1}}\mathscr{A}_{\alpha_{i+1}}^{s_{i+1}}\ldots\mathscr{A}_{\alpha_p}^{s_p}R = \mathscr{B}_i R$$

$\left(\text{где } \mathscr{B}_i = \mathscr{A}_{\alpha_1}^{s_1}\mathscr{A}_{\alpha_2}^{s_2}\ldots\mathscr{A}_{\alpha_{i-1}}^{s_{i-1}}\mathscr{A}_{\alpha_{i+1}}^{s_{i+1}}\ldots\mathscr{A}_{\alpha_p}^{s_p}\right)$, в котором оператор $\mathscr{A}$ совсем не имеет собственных значений, равных $\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_p$. В подпространстве $\mathscr{B}_i R$ у оператора $\mathscr{A}$ будет лишь одно собственное значение, равное α_i, причем оно будет той же кратности k_i, что и у оператора $\mathscr{A}$ в пространстве R. Конечно, подпространство $\mathscr{B}_i R$ инвариантно относительно $\mathscr{A}$, так как $\mathscr{A}\mathscr{B}_i R = \mathscr{B}_i\mathscr{A}R \subseteq$

$\subseteq \mathcal{B}_i R$. Оно инвариантно также и относительно каждого из операторов $\mathcal{A}_{\alpha i}$, где $i=1, 2, \ldots, p$.

Размерность подпространства $\mathcal{B}_i R$, как видно из построения, равна кратности k_i собственного значения α_i.

В подпространстве $\mathcal{B}_i R$ каждый из операторов $\mathcal{A}_{\alpha_1}, \mathcal{A}_{\alpha_2}, \ldots, \mathcal{A}_{\alpha_{i-1}}, \mathcal{A}_{\alpha_{i+1}}, \ldots, \mathcal{A}_{\alpha_p}$ является невырожденным и, следовательно, оператор $\mathcal{B}_i$ как произведение невырожденных операторов тоже будет невырожденным, т. е. $\mathcal{B}_i(\mathcal{B}_i R) = \mathcal{B}_i R$ при всех $i=1, 2, \ldots, p$. В то же время очевидно, что $\mathcal{B}_i(\mathcal{B}_j R)=0$ при $i \neq j$.

3. Покажем, что если $e_1^i, e_2^i, \ldots, e_{k_i}^i$ — базис подпространства $\mathcal{B}_i R$, то $k_1+k_2+\ldots+k_p=n$ векторов e_j^i (где $i=1, 2, \ldots, p$ и для каждого i отвечающие ему j пробегают значения $1, 2, \ldots, k_i$) образуют базис пространства R. Так как число этих векторов равно размерности R, то нам достаточно доказать их линейную независимость.

Предположим, что какая-то линейная комбинация векторов e_j^i обращается в нуль. Обозначая сумму всех тех из этих векторов, которые принадлежат $\mathcal{B}_i R$, через a_i, получим равенство

$$a_1+a_2+\ldots+a_p=0,$$

где, конечно, некоторые из слагаемых могут и обращаться в нуль. Применим к обеим частям оператор $\mathcal{B}_j$; учитывая, что $\mathcal{B}_j a_k=0$ при $j \neq k$, получим $\mathcal{B}_j a_j=0$. Но так как оператор $\mathcal{B}_j$, действующий в подпространстве $\mathcal{B}_j R$, — невырожденный, то $a_j=0$. Таким образом, базисные векторы всех инвариантных относительно $\mathcal{A}$ подпространств $\mathcal{B}_1 R, \mathcal{B}_2 R, \ldots, \mathcal{B}_p R$ линейно независимы, и значит, они образуют базис всего пространства R. В этом базисе матрица оператора $\mathcal{A}$ разобьется на клетки:

$$A=\left[\begin{array}{c:c:c:c} A_1 & & & \\ \hdashline & A_2 & & \\ \hdashline & & & \\ \hdashline & & & A_p \end{array}\right],$$

где A_j — это матрица оператора $\mathcal{A}_j$ в подпространстве $\mathcal{B}_j R$.

Так как оператор $\mathcal{A}$ в подпространстве $\mathcal{B}_j R$ имеет лишь одно собственное значение α_j, то, как показано в п. 1, соответствующим выбором базиса в $\mathcal{B}_j R$ клетку $\mathcal{A}_j$ можно привести к жордановой форме. Тем самым приведется к жордановой форме и матрица оператора $\mathcal{A}$ во всем пространстве R (см. замечание на стр. 118).

Рассмотрим пример. Пусть оператор $\mathcal{A}$ в некотором базисе имеет матрицу

$$A = \begin{bmatrix} 0 & -2 & 1 & 2 \\ 1 & 3 & -1 & -2 \\ 0 & 2 & 1 & -2 \\ 0 & -1 & 0 & 2 \end{bmatrix}.$$

Его характеристический многочлен

$$\varphi(\lambda) = \begin{vmatrix} -\lambda & -2 & 1 & 2 \\ 1 & 3-\lambda & -1 & -2 \\ 0 & 2 & 1-\lambda & -2 \\ 0 & -1 & 0 & 2-\lambda \end{vmatrix} = \lambda^4 - 6\lambda^3 + 13\lambda^2 - 12\lambda + 4 =$$

$$= (\lambda - 1)^2 (\lambda - 2)^2.$$

Собственные значения $\lambda_1 = 1$ и $\lambda_2 = 2$, оба кратности 2. При этом

$$A_1 = A - E = \begin{bmatrix} -1 & -2 & 1 & 2 \\ 1 & 2 & -1 & -2 \\ 0 & 2 & 0 & -2 \\ 0 & -1 & 0 & 1 \end{bmatrix},$$

$$A_1^2 = \begin{bmatrix} -1 & -2 & 1 & 2 \\ 1 & 2 & -1 & -2 \\ 2 & 6 & -2 & -6 \\ -1 & -3 & 1 & 3 \end{bmatrix}.$$

Ранг матрицы A_1 равен 2; значит, существуют два линейно независимых собственных вектора, отвечающих собственному значению, равному 1. Легко видеть, что здесь и ранг матрицы A_1^2 равен 2.

Далее, имеем

$$A_2 = A - 2E = \begin{bmatrix} -2 & -2 & 1 & 2 \\ 1 & 1 & -1 & -2 \\ 0 & 2 & -1 & -2 \\ 0 & -1 & 0 & 0 \end{bmatrix},$$

$$A_2^2=\begin{bmatrix}2 & 2 & -1 & -2\\ -1 & -1 & 1 & 2\\ 2 & 2 & -1 & -2\\ -1 & -1 & 1 & 2\end{bmatrix},\quad A_2^3=\begin{bmatrix}-2 & -2 & 1 & 2\\ 1 & 1 & -1 & -2\\ -2 & -2 & 1 & 2\\ 1 & 1 & -1 & -2\end{bmatrix}.$$

Ранг матрицы A_2 равен 3, значит, существует лишь одномерное подпространство, отвечающее собственному значению, равному 2. Отсюда уже ясно, что искомой жордановой формой матрицы A будет

$$I=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 2\end{bmatrix}.$$

Заметим, что ранги матриц A_2^2 и A_2^3 одинаковы и равны 2.

ГЛАВА IV

ЕВКЛИДОВО ПРОСТРАНСТВО

§ 1. Скалярное произведение

Мы определили векторное пространство, в котором можно складывать векторы и умножать их на числа, ввели понятия размерности, базиса, линейного оператора, а теперь в этом пространстве мы введем *метрику*, т. е. *способ измерять длины и углы*. Метрику в векторном пространстве удобнее всего ввести, используя понятие скалярного произведения.

В обычном трехмерном пространстве скалярным произведением двух векторов называется *произведение их длин, умноженное на косинус угла между ними*. Это скалярное умножение коммутативно: $(x, y) = (y, x)$, ассоциативно относительно умножения вектора на число: $(\alpha x, y) = \alpha(x, y)$ и дистрибутивно относительно сложения векторов: $(x + y, z) = (x, z) + (y, z)$; кроме того, скалярный квадрат (x, x) любого ненулевого вектора x положителен.

В случае n-мерного векторного пространства у нас нет понятия длины и угла, и мы введем скалярное произведение аксиоматически. Его определение мы дадим для случая, когда основное поле F есть поле комплексных чисел. Читатель, собирающийся изучать вещественное евклидово пространство, должен всюду, где над числом α из поля F стоит черточка, просто ее опустить: ведь в том случае, когда число α вещественно (и, кстати сказать, только в этом случае) $\overline{\alpha} = \alpha$.

Определение 1. *Говорят, что в векторном пространстве R задано* ***скалярное произведение****, если каждой паре векторов x, y из R поставлено в соответствие число $(x, y) \in F$ так, что выполнены следующие условия:*

1. *Для любых двух векторов x и y*

$$(x, y) = \overline{(y, x)}.$$

[(В случае вещественного пространства $(x, y) = (y, x)$.)]

2. *Для каждого вектора x и любого $\alpha \in F$*

$$(\alpha x, y) = \alpha (x, y).$$

3. *Для любых трех векторов x, y, z*

$$(x + y, z) = (x, z) + (y, z).$$

(Эти условия называются аксиомами скалярного умножения.) Пространство R называется в этом случае *пространством со скалярным произведением.*

Из условия 1 непосредственно вытекает, что $(x, x) = \overline{(x, x)}$, т. е. что скалярный квадрат любого вектора x является вещественным числом.

Пространство со скалярным произведением, удовлетворяющее кроме условий 1—3 *еще и условию*

4. *Для любого вектора x скалярный квадрат $(x, x) \geqslant 0$, и из равенства $(x, x) = 0$ вытекает, что $x = 0$, называется* ***евклидовым*** *векторным пространством* *).

Из равенств 1—3 легко получаются следующие соотношения:

$$2.' \quad (x, \alpha y) = \overline{(\alpha y, x)} = \overline{\alpha (y, x)} = \bar{\alpha}\overline{(y, x)} = \bar{\alpha}(x, y).$$

(В вещественном случае $(x, \alpha y) = \alpha (x, y)$.)

$$3.' \quad (z, x + y) = \overline{(x + y, z)} = \overline{(x, z) + (y, z)} =$$
$$= \overline{(x, z)} + \overline{(y, z)} = (z, x) + (z, y).$$

Примеры 1. Пусть в n-мерном векторном пространстве R зафиксирован определенный базис. Тогда скалярное произведение векторов $x = (x_1, x_2, \ldots, x_n)$ и $y = (y_1, y_2, \ldots, y_n)$ можно определить равенством

$$(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n \tag{1}$$

(в вещественном пространстве $(x, y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$). Справедливость условий 1—4 проверяется непосредственно.

2. В пространстве P многочленов от t с вещественными коэффициентами и в пространстве C функций, непрерывных на отрезке

*) Удовлетворяющее условиям 1—3 «произведение» векторов комплексного векторного пространства, сопоставляющее каждым двум векторам $x, y \in R$ комплексное число (x, y), часто называют также *эрмитовым скалярным произведением*, а то пространство, которое мы назвали евклидовым,— *эрмитовым* (или *унитарным*) комплексным векторным пространством. Мы, однако, предпочтем использовать здесь более простые и привычные термины «скалярное произведение» и «евклидово пространство».

$[a, b]$, скалярное произведение можно определить равенством

$$(x, y) = \int_a^b x(t)\, y(t)\, dt. \tag{2}$$

Справедливость условий 1—3 очевидна, а 4-е следует из того, что непрерывная неотрицательная функция, интеграл от которой равен нулю, тождественно равна нулю.

Определение 2. *Длиной*, или *модулем*, или *нормой*, *вектора x в евклидовом пространстве называется корень квадратный из его скалярного квадрата*

$$|x| = \sqrt{(x, x)}.$$

Векторы x и *y, скалярное произведение (x, y) которых равно нулю* (а значит, равно нулю и произведение $(y, x) = \overline{(x, y)}$), *называются ортогональными.* В этом случае мы будем также писать $x \perp y$.

В любом пространстве со скалярным произведением справедлива

«Теорема Пифагора». *Если векторы x и y ортогональны, то*

$$|x + y|^2 = |x|^2 + |y|^2.$$

Действительно, если $(x, y) = 0$ то, ввиду условий 1—3,

$$|x + y|^2 = (x + y, x + y) = \\ = (x, x) + (x, y) + (y, x) + (y, y) = |x|^2 + |y|^2.$$

В любом евклидовом пространстве справедливо

Неравенство Коши—Буняковского: для любых векторов x, y из R

$$|(x, y)| \leqslant |x|\,|y|.$$

Доказательство проведем отдельно для вещественного и комплексного случаев.

А. Пространство R — вещественное.

Если $\alpha \in F$, то для вектора $x - \alpha y$ по условию 4 имеем неравенство

$$(x - \alpha y, x - \alpha y) \geqslant 0,$$

из которого, ввиду условий 1 — 3, получаем

$$(x, x) - 2\alpha(x, y) + \alpha^2(y, y) \geqslant 0,$$

или

$$|x|^2 - 2\alpha(x, y) + \alpha^2|y|^2 \geqslant 0.$$

Это — квадратный трехчлен относительно α. Так как он должен быть неотрицательным при всех значениях α, то он не может иметь двух различных вещественных корней и, значит, его дискриминант неположителен:

$$(x, y)^2 - |x|^2|y|^2 \leqslant 0,$$

откуда

$$|(x, y)| \leqslant |x|\,|y|,$$

что и требовалось доказать.

Б. Пространство R — комплексное.

И в этом случае для любых двух векторов x, y из R и любого (комплексного) числа α

$$(x - \alpha y, x - \alpha y) \geqslant 0,$$

откуда, в силу условий 1 — 3, получаем

$$(x, x) - \alpha(y, x) - \bar{\alpha}(x, y) + \alpha\bar{\alpha}(y, y) \geqslant 0.$$

Полагая $\alpha = \beta \frac{(x, y)}{|(x, y)|}$, где β — произвольное вещественное число, и учитывая, что $(x, y)\overline{(x, y)} = |(x, y)|^2$, будем иметь

$$|x|^2 - 2\beta|(x, y)| + \beta^2|y|^2 \geqslant 0.$$

Мы получили квадратный трехчлен относительно β с вещественными коэффициентами. Так как он неотрицателен при всех β, то его дискриминант неположителен и, значит,

$$|(x, y)|^2 - |x|^2|y|^2 \leqslant 0,$$

что и требовалось доказать.

Легко видеть, что равенство $|(x, y)| = |x|\,|y|$ будет иметь место в том и только в том случае, если для некоторого $\alpha \in F$ имеем $x - \alpha y = 0$, т. е. если векторы x и y пропорциональны: $x = \alpha y$.

В вещественном евклидовом пространстве можно определить угол φ между ненулевыми векторами x

и y. По определению,

$$\cos\varphi = \frac{(x, y)}{|x|\,|y|}. \tag{3}$$

Легко видеть, что, ввиду неравенства Коши — Буняковского, $|\cos\varphi| \leqslant 1$.

Из неравенства Коши — Буняковского, если применить его к пространству P со скалярным произведением (1), получается «неравенство Коши»: для любых чисел a_i, b_i

$$|a_1\bar{b}_1 + a_2\bar{b}_2 + \ldots + a_n\bar{b}_n|^2 \leqslant$$
$$\leqslant (|a_1|^2 + |a_2|^2 + \ldots + |a_n|^2)(|b_1|^2 + |b_2|^2 + \ldots + |b_n|^2)$$

— для комплексного пространства и

$$(a_1b_1 + a_2b_2 + \ldots + a_nb_n)^2 \leqslant$$
$$\leqslant \left(a_1^2 + a_2^2 + \ldots + a_n^2\right)\left(b_1^2 + b_2^2 + \ldots + b_n^2\right)$$

— в вещественном случае, а для пространства C со скалярным произведением (2) — «неравенство Буняковского»

$$\left[\int_a^b x(t)\,y(t)\,dt\right]^2 \leqslant \int_a^b [x(t)]^2\,dt \cdot \int_a^b [y(t)]^2\,dt,$$

справедливое для любых двух непрерывных функций $x(t)$ и $y(t)$.

В евклидовом пространстве справедливо так называемое *неравенство треугольника*: *для любых двух векторов x, $y \in R$*

$$|x + y| \leqslant |x| + |y|.$$

Доказательство в вещественном случае очень просто. Пользуясь неравенством Коши—Буняковского, получаем

$$|x+y|^2 = (x+y,\ x+y) = |x|^2 + 2(x, y) + |y|^2 \leqslant$$
$$\leqslant |x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2,$$

откуда

$$|x+y| \leqslant |x| + |y|.$$

Пусть теперь пространство R — комплексное. Имеем, очевидно,

$$|x \pm y|^2 = (x + y,\ x + y) = (x, x) \pm (x, y) + (y, x) + (y, y).$$

Скалярное произведение (x, y) есть, вообще говоря, комплексное число, пусть $(x, y) = a + bi$. Тогда $(y, x) = \overline{(x, y)} = a - bi$ и

$$(x, y) + (y, x) = 2a \leqslant 2\sqrt{a^2 + b^2} = 2\,|(x, y)|.$$

Следовательно,

$$|x + y|^2 \leqslant |x|^2 + 2\,|(x, y)| + |y|^2,$$

а это, в силу неравенства Коши—Буняковского, не превосходит

$$|x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2.$$

Таким образом, $|x + y|^2 \leqslant (|x| + |y|)^2$ и, значит,

$$|x + y| \leqslant |x| + |y|.$$

Равенство $|x + y| = |x| + |y|$ будет выполняться, если, во-первых, $y = \alpha x$ (и тогда $|(x, y)| = |x|\,|y|$) и если, во-вторых, $a = \sqrt{a^2 + b^2}$, т. е. если $b = 0$ и скалярное произведение (x, y) вещественно и положительно. А тогда $(y, x) = (\alpha x, x) = \alpha(x, x) > 0$ и $\alpha > 0$.

§ 2. Ортонормированный базис

Определение 3. *Базис $e_1, e_2, \ldots, e_n$ евклидова пространства R называется* ***ортогональным***, *если*

$$(e_i, e_k) = 0 \quad \text{при } i \neq k.$$

Если, кроме того,

$$|e_i| = 1 \quad \textit{при} \quad i = 1, 2, \ldots, n,$$

то базис называется ***ортонормированным***.

Лемма. *Попарно ортогональные и отличные от нуля векторы линейно независимы.*

Доказательство. Пусть векторы $x_1, x_2, \ldots, x_m$ попарно ортогональны: $(x_i, x_k) = 0$ при $i \neq k$, и отличны от нуля. Предположим, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m = 0.$$

Умножая обе части этого равенства скалярно на x_i, $i = 1, 2, \ldots, m$, будем иметь

$$\alpha_1 (x_1, x_i) + \alpha_2 (x_2, x_i) + \ldots + \alpha_m (x_m, x_i) = 0,$$

откуда, поскольку $(x_i, x_k) = 0$ при $i = k$ и $(x_i, x_i) \neq 0$ при всех $i = 1, 2, \ldots, m$, вытекает, что $\alpha_i = 0$ при всех $i = 1, 2, \ldots, m$.

Теорема 1. *Во всяком евклидовом пространстве R имеются ортонормированные базисы.*

Доказательство. Пусть $g_1, g_2, \ldots, g_n$ — произвольный базис пространства R. Положим $f_1 = g_1$ и $f_2 = g_2 + \alpha f_1$, причем α подберем так, чтобы векторы f_1 и f_2 были ортогональны:

$$(g_2 + \alpha f_1, f_1) = (g_2, f_1) + \alpha (f_1, f_1) = 0,$$

откуда

$$\alpha = -\frac{(g_2, f_1)}{(f_1, f_1)}.$$

Так как $f_1 \neq 0$, то знаменатель (f_1, f_1) последней дроби отличен от нуля. Ввиду линейной независимости векторов g_1 и g_2 полученный вектор f_2 — ненулевой.

Допустим теперь, что попарно ортогональные и отличные от нуля векторы $f_1, f_2, \ldots, f_{k-1}$ уже найдены. Положим

$$f_k = g_k + \lambda_1 f_1 + \lambda_2 f_2 + \ldots \lambda_{k-1} f_{k-1},$$

и подберем числа $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ так, чтобы вектор f_k был ортогонален к $f_1, f_2, \ldots, f_{k-1}$. Для этого нужно, чтобы выполнялись равенства

$$(f_k, f_i) = (g_k, f_i) + \lambda_i (f_i, f_i) = 0 \quad \text{при} \quad i = 1, 2, \ldots, k-1,$$

откуда

$$\lambda_i = -\frac{(g_k, f_i)}{(f_i, f_i)}.$$

Знаменатель (f_i, f_i) здесь отличен от нуля, так как все векторы f_i при $i = 1, 2, \ldots, k-1$, по предположению, — ненулевые. Так как векторы $g_1, g_2, \ldots, g_k$ линейно независимы, то и полученный вектор f_k тоже будет ненулевым.

Это построение мы будем продолжать до тех пор, пока не найдем последнего (ненулевого) вектора

$$f_n = g_n + \xi_1 f_1 + \xi_2 f_2 + \ldots + \xi_{n-1} f_{n-1},$$

ортогонального всем предыдущим векторам $f_1, f_2, \ldots, f_{n-1}$. В силу последней леммы векторы $f_1, f_2, \ldots, f_n$

линейно независимы и, значит, образуют (ортогональный) базис. Если теперь каждый из векторов f_i поделить на его модуль, то получится **ортонормированный** базис, образованный векторами

$$e_1 = \frac{f_1}{|f_1|},\ e_2 = \frac{f_2}{|f_2|},\ \ldots,\ e_n = \frac{f_n}{|f_n|}.$$

Легко видеть, что *если первые k векторов $g_1, g_2, \ldots, g_k$ были попарно ортогональными, то $f_1 = g_1$, $f_2 = g_2, \ldots, f_k = g_k$, а если они были, кроме того, единичными, то $e_1 = g_1$, $e_2 = g_2, \ldots, e_k = g_k$.*

Примененный здесь способ получения ортонормированной системы векторов из заданной линейно независимой системы носит название **процесса ортогонализации**.

Замечание. *Если R_1 — подпространство R и $e_1, e_2, \ldots, e_k$ — ортонормированный базис R_1, то векторы $e_1, e_2, \ldots, e_k$ можно включить в ортонормированный базис всего пространства.*

Для доказательства достаточно дополнить $e_1, e_2, \ldots, e_k$ до базиса пространства R и произвести ортогонализацию полученного множества векторов, начиная с $e_1, e_2, \ldots, e_k$.

Пример. *Найти ортогональный базис в пространстве многочленов степени не выше 4, определенных на отрезке* $[-1, 1]$.

Решение. В качестве исходного базиса возьмем

$$g_0 = 1,\quad g_1 = t,\quad g_2 = t^2,\quad g_3 = t^3,\quad g_4 = t^4.$$

Положим

$$f_0 = g_0 = 1$$

и $f_1 = g_1 + \alpha f_0$. Так как $(g_1, f_0) = \int\limits_{-1}^{1} t\,dt = 0$, то $\alpha = 0$ и

$$f_1 = g_1 = t.$$

Далее, положим $f_2 = g_2 + \beta f_0 + \gamma f_1$. Имеем $(g_2, f_0) = \int\limits_{-1}^{1} t^2 dt = \frac{2}{3}$, $(f_0, f_0) = \int\limits_{-1}^{1} dt = 2$, откуда $\beta = -\frac{1}{3}$, $(g_2, f_1) = \int\limits_{-1}^{1} t^3 dt = 0$, и значит, $\gamma = 0$. Следовательно,

$$f_2 = t^2 - \frac{1}{3}.$$

Пусть $f_3 = g_3 + \lambda f_0 + \mu f_1 + \nu f_2$. Имеем $(g_3, f_0) = \int\limits_{-1}^{1} t^3 dt = 0$, откуда

$\lambda = 0$; $(g_3, f_1) = \int\limits_{-1}^{1} t^4 dt = \frac{2}{5}$, $(f_1, f_1) = \int\limits_{-1}^{1} t^2 dt = \frac{2}{3}$, значит, $\mu = -\frac{3}{5}$,

и $(g_3, f_2) = \int\limits_{-1}^{1} \left(t^5 - \frac{1}{3}t^3\right) dt = 0$, т. е. $\nu = 0$. Следовательно,

$$f_3 = t^3 - \frac{3}{5}t.$$

Положим, наконец, $f_4 = g_4 + \xi f_0 + \eta f_1 + \zeta f_2 + \rho f_3$. Тогда, поскольку $(g_4, f_0) = \int\limits_{-1}^{1} t^4 dt = \frac{2}{5}$, а $(f_0, f_0) = 2$, то $\xi = -\frac{1}{5}$; далее

$(g_4, f_1) = \int\limits_{-1}^{1} t^5 dt = 0$, значит, $\eta = 0$. Затем имеем

$$(g_4, f_2) = \int\limits_{-1}^{1} t^4 \left(t^2 - \frac{1}{3}\right) dt = \frac{2}{7} - \frac{2}{15} = \frac{16}{105}$$

и

$$(f_2, f_2) = \int\limits_{-1}^{1} \left(t^4 - \frac{2}{3}t^2 + \frac{1}{9}\right) dt = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45},$$

откуда $\zeta = -\frac{6}{7}$; наконец, $(g_4, f_3) = \int\limits_{-1}^{1} t^4 \left(t^3 - \frac{3}{5}t\right) dt = 0$, т. е. $\rho = 0$. Следовательно,

$$f_4 = t^4 - \frac{1}{5} - \frac{6}{7}\left(t^2 - \frac{1}{3}\right) = t^4 - \frac{6}{7}t^2 + \frac{3}{35}.$$

Полученные многочлены

$$f_0, f_1, f_2, f_3, f_4$$

— это (с точностью до множителей) первые пять из так называемых многочленов Лежандра, играющих важную роль в математической физике.

Найдем *выражение скалярного произведения векторов в координатах*. Пусть $e_1, e_2, \ldots, e_n$ — произвольный базис пространства R со скалярным произведением и

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$
$$y = y_1e_1 + y_2e_2 + \ldots + y_ne_n.$$

Тогда

$$(x, y) = \left(\sum_{i=1}^{n} x_ie_i, \sum_{k=1}^{n} y_ke_k\right) = \sum_{i,k=1}^{n} (x_ie_i, y_ke_k) =$$
$$= \sum_{i,k=1}^{n} x_i\bar{y}_k(e_i, e_k) = \sum_{i,k=1}^{n} g_{ik}x_i\bar{y}_k,$$

где i и k независимо друг от друга пробегают значения $1, 2, \ldots, n$, а $g_{ik} = (e_i, e_k)$. Если пространство R евклидово, а $e_1, e_2, \ldots, e_n$ — ортонормированный базис, то $(e_i, e_k) = 0$ при $i \neq k$, $(e_i, e_i) = 1$ при всех $i = 1, 2, \ldots, n$ и, значит,

$$(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n.$$

Легко видеть, что, и обратно, если в базисе e_1 $e_2, \ldots, e_n$ скалярное произведение векторов

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

и

$$y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$$

равно

$$x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n,$$

то этот базис ортонормированный, так как в этом случае $(e_i, e_i) = 1$ и $(e_i, e_k) = 0$ при $i \neq k$.

Пусть $e_1, e_2, \ldots, e_n$ — ортонормированный базис в евклидовом пространстве R и $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$. Умножив обе части последнего равенства скалярно на e_i, получим $(x, e_i) = x_i$, т. е. i-я координата вектора x в ортонормированном базисе равна скалярному произведению x на единичный вектор e_i. Это скалярное произведение можно назвать (ортогональной) проекцией вектора x на вектор e_i. Таким образом, *координаты вектора в ортонормированном базисе — это его проекции на базисные векторы.*

Пусть R и R' — два n-мерных евклидовых пространства. Если в каждом из них выбрать ортонормированный базис ($e_1, e_2, \ldots, e_n$ в R и $e'_1, e'_2, \ldots, e'_n$ в R')

и поставить в соответствие каждому вектору x из R вектор x' из R' с теми же координатами, то, как известно (см. § 5 главы II), сумме элементов из R будет отвечать сумма соответствующих элементов из R' и произведению элемента из R на число — произведение соответствующего элемента из R' на то же число. При этом, если

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n, \quad y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$$

(и значит,

$$x' = x_1e_1' + x_2e_2' + \ldots + x_ne_n', \quad y' = y_1e_1' + y_2e_2' + \ldots + y_ne_n'),$$

то скалярное произведение

$$(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n = (x', y').$$

Таким образом, пространства R и R' устроены одинаково: соответствующие векторы их имеют одинаковые длины $\left(|x| = \sqrt{(x, x)} = \sqrt{(x', x')} = |x'|\right)$, а в случае вещественного пространства и углы между парами соответствующих друг другу векторов равны между собой:

$$\left(\cos(\widehat{x, y}) = \frac{(x, y)}{|x|\,|y|} = \frac{(x', y')}{|x'|\,|y'|} = \cos(\widehat{x', y'})\right).$$

Таким образом, все евклидовы векторные пространства над одним и тем же полем изоморфны и, как говорят, «изометричны» между собой, т. е. обладают в некотором смысле одинаковыми метриками; следовательно, единственной характеристикой евклидова пространства над данным полем F является его р а з м е р н о с т ь.

§ 3. Ортогональное дополнение

О п р е д е л е н и е 4. *Два подпространства R_1 и R_2 евклидова пространства R называются* в з а и м н о о р т о г о н а л ь н ы м и, *если каждый вектор из R_1 ортогонален каждому вектору из R_2* (мы будем писать в этом случае $R_1 \perp R_2$).

Так, в обычном трехмерном пространстве проходящая через начало координат плоскость π (понимаемая как множество принадлежащих π векторов) и перпендикулярная к ней (и тоже проходящая через начало) прямая l ортогональны (рис. 10, a). Наоборот, две взаимно перпендикулярные в смысле элементарной геометрии плоскости π_1 и π_2 (рис. 10, *б*) не будут ортогональными подпространствами в смысле этого определения: ведь из того, что $a_1 \in \pi_1$, а $a_2 \in \pi_2$, совсем не следует, что $a_1 \perp a_2$.

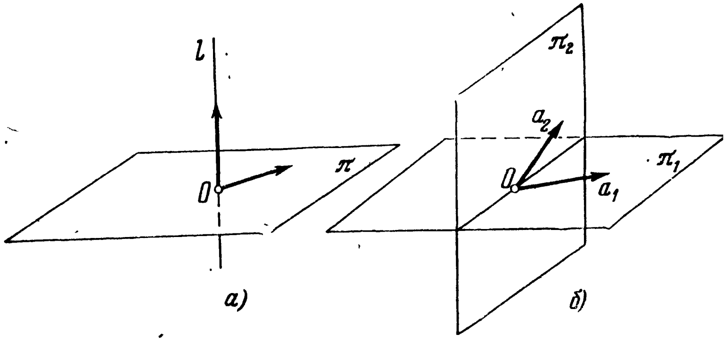


Рис. 10.

Для того чтобы подпространства R_1 и R_2 были взаимно ортогональными, необходимо и достаточно, чтобы все базисные векторы одного были ортогональны всем базисным векторам другого. Необходимость следует из определения 4, для доказательства достаточности предположим, что $e_1, e_2, \ldots, e_k$ — базис R_1 и $f_1, f_2, \ldots, f_m$ — базис R_2, причем $(e_i, f_j) = 0$ для всех $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, m$, тогда для каждого $x = x_1e_1 + x_2e_2 + \ldots + x_ke_k$ и каждого $y = y_1f_1 + y_2f_2 + \ldots + y_mf_m$ скалярное произведение

$$(x, y) = \sum_{i,j=1}^{n} x_i\bar{y}_j (e_i, f_j) = 0,$$

и значит, эти векторы ортогональны.

Покажем, что *два взаимно ортогональных подпространства пересекаются только по нулевому вектору.*

Действительно, пусть R_1 и R_2 — взаимно ортогональные подпространства R. Если вектор $x \in R_1 \cup R_2$, то $x \in R_1$ и $x \in R_2$; но тогда $(x, x) = 0$ и, значит, $x = 0$.

Пусть R_1 — произвольное подпространство евклидова пространства R. Выберем в R_1 ортонормированный базис $e_1, e_2, \ldots, e_r$ и дополним его до ортонормированного базиса $e_1, e_2, \ldots, e_r, e_{r+1}, \ldots, e_n$ всего пространства. Векторы $e_{r+1}, \ldots, e_n$ порождают $(n-r)$-мерное подпространство R_2, очевидно, ортогональное R_1.

Покажем, что *каждый вектор x из R, ортогональный R_1, принадлежит R_2*. Действительно, если вектор

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

ортогонален R_1, то

$$(x, e_i) = x_i = 0 \text{ при } i = 1, 2, \ldots, r,$$

и значит,

$$x = x_{r+1} e_{r+1} + \ldots + x_n e_n \in R_2.$$

Определение 5. *Подпространство R_2, образованное всевозможными векторами из R, ортогональными ко всем векторам из R_1, называется **ортогональным дополнением** R_1*; это подпространство R_2 мы будем обозначать через $R_1^{\perp}$.

Легко видеть, что ортогональное дополнение r-мерного подпространства $(n-r)$-мерно и что ортогональное дополнение к $R_1^{\perp}$ совпадает с R_1, т. е. что

$$(R_1^{\perp})^{\perp} = R_1.$$

Подпространства R_1 и $R_1^{\perp}$ порождают все R и пересекаются по нулевому вектору. Следовательно, *евклидово пространство R является прямой суммой любого своего подпространства и его ортогонального дополнения:*

$$R = R_1 \oplus R_1^{\perp}.$$

Поэтому каждый вектор x из R однозначно представляется в виде суммы $x = y + z$, где $y \in R_1$, $z \in R_1^{\perp}$ (теорема 6 главы II). Вектор y можно назвать *ортогональной проекцией вектора x на подпространство R_1*. В случае вещественного пространства можно определить и **угол между вектором** x и **подпространством** R_1 — его естественно считать равным углу между вектором x и проекцией y вектора x на R_1, а значит, косинус этого угла равен

$$\frac{(x, y)}{|x||y|} = \frac{(y+z, y)}{|x||y|} = \frac{(y, y)}{|x||y|} = \frac{|y|^2}{|x||y|} = \frac{|y|}{|x|}.$$

Рассмотрим снова систему линейных однородных уравнений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \qquad (4)$$

Этой системе можно дать следующую геометрическую интерпретацию. В евклидовом пространстве R^n (в ортонормированном базисе) *задано* m *векторов* $a_i = (\bar{a}_{i1}, \bar{a}_{i2}, \ldots, \bar{a}_{in})$, $i = 1, 2, \ldots, m$. Задача состоит в том, чтобы *найти все векторы* $x = (x_1, x_2, \ldots, x_n)$, *ортогональные каждому из векторов* $a_1, a_2, \ldots, a_m$.

Пусть ранг матрицы $A = [a_{ik}]$ равен r. Если вектор x ортогонален ко всем векторам a_i, то он ортогонален и к порождаемому ими r-мерному подпространству R_1. Таким образом, векторы-решения x образуют ортогональное дополнение $R_1^{\perp}$ подпространства R_1. Размерность $R_1^{\perp}$ (т. е. максимальное число линейно независимых решений системы (4)) равна, как мы видели, $n - r$. Каждая фундаментальная система решений уравнений (4) — это базис подпространства $R_1^{\perp}$.

§ 4. Евклидово (точечно-векторное) пространство

Пусть A^n — вещественное n-мерное аффинное пространство и R^n — соответствующее ему векторное пространство, в котором введена евклидова метрика (т. е. задано скалярное произведение, удовлетворяющее условиям 1—4 из § 1). В пространстве A^n можно определить расстояние между любыми двумя его точками, M и N, полагая его равным модулю вектора $\overline{MN}$, а в случае вещественного пространства — и угол, MPN, считая его равным углу между векторами $\overline{PM}$ и $\overline{PN}$.

Пространство A^n с введенной в нем таким образом метрикой называется просто **евклидовым пространством** (в отличие от введенного выше евклидова векторного пространства *)).

Таким образом, n-мерное евклидово пространство может быть определено с помощью следующих пяти групп аксиом:

I. Аксиомы сложения векторов (1—4 на стр. 64).

*) Которое, впрочем, для краткости (и если нет опасности смешения этих двух понятий) также называют короче — евклидовым пространством.

II. Аксиомы умножения вектора на число (5—8 на стр. 64).

III. Аксиома размерности: *существуют n линейно независимых векторов, но нет больше чем n линейно независимых векторов* (ср. стр. 66).

IV. Аксиомы, связывающие векторы и точки (1—2 на стр. 83).

V. Аксиомы скалярного умножения (1—4 на стр. 144—145).

Можно показать, что все n-мерные евклидовы пространства над одним и тем же полем тоже «устроены одинаково» (изоморфны и изометричны между собой). В частности, при $n = 2$ это — обычная плоскость, при $n = 3$ — обычное трехмерное пространство.

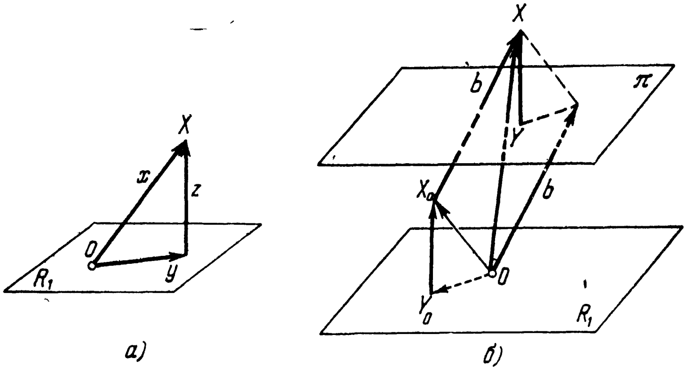


Рис. 11.

Пусть в вещественном пространстве A^n заданы k-мерная плоскость R_1, проходящая через начало координат:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{cases} \tag{5}$$

(k-мерное подпространство) и точка $X(\xi_1, \xi_2, \ldots, \xi_n)$. Тогда вектор $x = \overline{OX}$ можно представить в виде $x = y + z$, где $y \in R_1$ и $z \in R_1^{\perp}$ (рис. 11, a). Длина вектора

r называется **расстоянием точки X от подпространства R_1.**

Пусть теперь в A^n заданы **произвольная** k-мерная плоскость π:

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

и точка X $(\xi_1, \xi_2, \ldots, \xi_n)$. Плоскость π получается из соответствующего ей **подпространства** $R_1 = \pi_0$, определяемого системой уравнений (5), параллельным переносом на некоторый вектор b. При этом точка X получается переносом на тот же вектор b из некоторой точки X_0 (и, значит, $\overline{OX} = \overline{OX_0} + b$; см. рис. 11, *б*). **Расстояние точки X от k-мерной плоскости π** естественно считать равным расстоянию точки X_0 от подпространства $\pi_0 = R_1$. (Можно показать, что расстояние точки x от k-мерной плоскости π — это **наименьшее** из расстояний точки x от всех точек плоскости π).

k-мерная плоскость π_1 и l-мерная плоскость π_2 *ортогональны*, если ортогональны соответствующие им подпространства π_1^0 и π_2^0. В этом случае каждый вектор $\overline{MN}$, где точки M, $N \in \pi_1$, ортогонален каждому вектору $\overline{PQ}$, где P, $Q \in \pi_2$ (рис. 12).

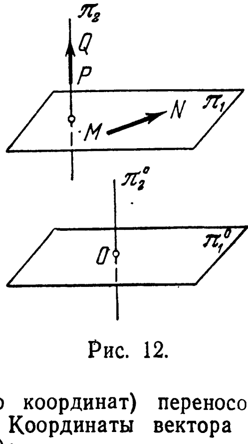


Рис. 12.

Пусть в **вещественном** пространстве A^n выбрана **ортонормированная** система координат. Рассмотрим линейное уравнение

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = \alpha, \qquad (6)$$

коэффициенты $a_1, a_2, \ldots, a_n$ левой части которого не равны нулю одновременно. Это уравнение определяет некоторую **гиперплоскость** π (см. стр. 86), которая получается из подпространства π_0:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0 \qquad (7)$$

(гиперплоскости, проходящей через начало координат) переносом на некоторый вектор $b = (b_1, b_2, \ldots, b_n)$. Координаты вектора b удовлетворяют, как мы знаем, уравнению (6):

$$a_1b_1 + a_2b_2 + \ldots + a_nb_n = \alpha$$

(см. стр. 79). Последнее равенство, положив $a = (a_1, a_2, \ldots, a_n)$, можно переписать в виде

$$ab = \alpha. \tag{8}$$

Вектор a ортогонален подпространству π_0, так как для каждого вектора $x = (x_1, x_2, \ldots, x_n) \in \pi_0$ скалярное произведение $ax = a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$.

Положим $M = \pm|a| = \pm\sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}$, причем знак здесь выберем так, чтобы $\alpha/M = p$ было неотрицательно. (Если $\alpha = 0$, то знак может быть выбран произвольно.) Уравнение

$$\frac{a_1}{M}x_1 + \frac{a_2}{M}x_2 + \ldots + \frac{a_n}{M}x_n - \frac{\alpha}{M} = 0$$

называется *нормальным уравнением гиперплоскости* π. Вектор $m = a/M$ является, очевидно, единичным вектором (т. е. вектором длины 1), коллинеарным a, и значит, ортогональным π_0.

Пусть теперь нам дана точка $X(\xi_1, \xi_2, \ldots, \xi_n)$ и надо найти *расстояние точки X от гиперплоскости* π.

Точка X получается сдвигом на вектор b из некоторой точки X_0. Это значит, что $\overline{OX} = \overline{OX}_0 + b$ (см. тот же рис. 11, *б*). Теперь нам остается найти расстояние от точки X_0 до подпространства π_0. Представим вектор $\overline{OX}_0$ в виде $\overline{OX}_0 = y + z$, где $y \in \pi_0$, а $z \perp \pi_0$. Тогда искомое расстояние будет равно длине вектора z. Но вектор z, как и вектор m, ортогональный π_0, коллинеарен m и, значит, найдется такое число λ, что $z = \lambda m$. Так как вектор m — единичный, то искомое расстояние, равное $|z|$, равно $|\lambda|$.

Итак, мы имеем равенство

$$\overline{OX} = y + \lambda m + b.$$

Умножим его скалярно на m:

$$\overline{OX} \cdot m = ym + \lambda m^2 + bm.$$

Но $ym = 0$, так как $y \in \pi_0$, а $m \perp \pi_0$; $m^2 = 1$ и $bm = b \cdot \frac{a}{M} = \frac{\alpha}{M}$ (см. (8)). Следовательно, $\overline{OX} \cdot m = \lambda + \frac{\alpha}{M}$, откуда $\lambda = \overline{OX} \cdot m - \frac{\alpha}{M}$. Наконец, скалярное произведение $\overline{OX} \cdot m$ в координатах равно

$$\xi_1\frac{a_1}{M} + \xi_2\frac{a_2}{M} + \ldots + \xi_n\frac{a_n}{M}$$

и значит

$$|\lambda| = \frac{|a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n - \alpha|}{\sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}}.$$

Таким образом, *для того чтобы найти расстояние от точки до гиперплоскости, надо подставить координаты этой точки в левую часть нормального уравнения гиперплоскости и взять полученную величину по модулю.* (Вспомните формулу расстояния от точки до прямой на плоскости и от точки до плоскости в пространстве!)

Гиперсферой в евклидовом пространстве A^n называется совокупность всех точек, отстоящих на одно и то же расстояние r (радиус гиперсферы) от некоторой фиксированной точки Q (центра). Уравнение гиперсферы радиуса r с центром в точке $Q(\alpha_1, \alpha_2, \ldots, \alpha_n)$ в ортонормированной системе координат, как легко видеть, имеет вид

$$(x_1-\alpha_1)^2+(x_2-\alpha_2)^2+\ldots+(x_n-\alpha_n)^2=r^2.$$

Следовательно, гиперсфера является частным случаем поверхности второго порядка (ср. главу VII). Гиперсфера S *касается* гиперплоскости π, если она имеет с этой гиперплоскостью единственную общую точку.

Задачи. (Задачи 1—7 относятся к четырехмерному пространству; система координат везде ортонормированная.)

1. Найдите расстояние точки $(-1, 3, 5, 1)$ от начала координат, от координатных осей, от координатных (двумерных) плоскостей и от координатных гиперплоскостей.

2. Найдите точки пересечения прямых

а) $\dfrac{x-1}{1}=\dfrac{y+3}{3}=\dfrac{z-1}{4}=\dfrac{t}{5}$;

б) $\dfrac{x-1}{2}=\dfrac{y+3}{3}=\dfrac{z-1}{4}=\dfrac{t-17}{5}$;

в) $\dfrac{x-1}{2}=\dfrac{y+3}{3}=\dfrac{z-1}{4}=\dfrac{t}{5}$

с гиперплоскостью $x+3y-4z+t=5$.

3. Найдите условия, при которых прямая

$$\frac{x-x_0}{b_1}=\frac{y-y_0}{b_2}=\frac{z-z_0}{b_3}=\frac{t-t_0}{b_4}$$

принадлежит гиперплоскости $a_1x+a_2y+a_3z+a_4t=\alpha$.

4. Докажите, что гиперплоскость, касающаяся гиперсферы, ортогональна радиусу, проведенному в точку касания.

5. Напишите уравнение гиперсферы, имеющей центр в точке $(5, -1, 4, 0)$ и касающейся гиперплоскости $x-3y+z+5t=6$.

6. Пересечение гиперсферы

$$(x-5)^2+y^2+(z+5)^2+(t-2)^2=25$$

и гиперплоскости

$$7x-5y+z+5t=20$$

есть некоторая сфера трехмерного пространства. Найдите ее центр и радиус.

7. Напишите уравнение гиперплоскости, проходящей через двумерную плоскость

$$\begin{cases} x-2y+3z+5t=2,\\ 3x-y-z+2t=3 \end{cases}$$

и

а) проходящую через точку $(2, 5, -3, 0)$,

б) ортогональную гиперплоскости $2x+y+4z-t=5$, Пусть в n-мерном пространстве A^n даны n попарно ортогональных векторов одинаковой длины $a_1, a_2, \ldots, a_n$. Тогда натянутым на них (n-мерным) *кубом* называется совокупность всевозможных векторов вида

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n,$$

где

$$0 \leqslant \alpha_i \leqslant 1, \quad i = 1, 2, \ldots, n.$$

k-мерная грань куба — это множество таких его точек, для которых $n-k$ из коэффициентов α_i, принимают постоянные значения, равные 0 или 1.

8. Найдите число k-мерных граней n-мерного куба.

9. Найдите угол между диагональю n-мерного куба (т. е. вектором $a_1 + a_2 + \ldots + a_n$) и его ребром a_j (заметьте, кстати, что этот угол не зависит от j).

10 Найдите угол между диагональю n-мерного куба и его k-мерной гранью.

ГЛАВА V

ЛИНЕЙНЫЕ ОПЕРАТОРЫ В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

§ 1. Линейный функционал

Определение 1. *Линейный оператор f, отображающий векторное пространство R в числовое поле F, называется линейным функционалом,* или *линейной функцией.*

Таким образом, если f — линейный функционал, то для каждого вектора $x \in R$ определено число $f(x)$ из основного поля F так, что выполнены следующие условия:

1) $f(x+y) = f(x) + f(y)$,

2) $f(\alpha x) = \alpha f(x)$,

где x и y — произвольные векторы из R, а $\alpha \in F$.

Для того чтобы найти выражение линейного функционала в координатах, выберем в пространстве R базис $e_1, e_2, \ldots, e_n$. Если $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$ — произвольный вектор из R, то

$$f(x) = f(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = \\ = x_1f(e_1) + x_2f(e_2) + \ldots + x_nf(e_n).$$

Обозначив $f(e_i) = a_i$, где $i = 1, 2, \ldots, n$, получим

$$f(x) = x_1a_1 + x_2a_2 + x_na_n.$$

Таким образом, при фиксированном базисе линейный функционал представляется линейной формой*), т. е. выражением вида

$$f(x) = a_1x_1 + a_2x_2 + \ldots + a_nx_n.$$

Если пространство R евклидово, а базис $e_1, e_2, \ldots, e_n$ — ортонормированный, то $f(x) = (x, a)$ — скаляр-

*) Слово «форма» означает «однородный многочлен», т. е. многочлен, являющийся суммой одночленов одной и той же степени.

ному произведению вектора x и некоторого (зависящего только от f, но не от x) вектора $a = (\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n)$.

Легко видеть, что верно и обратное: если в евклидовом векторном пространстве R задан вектор $a = (\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n)$, то $f(x) = (x, a)$ — скалярное произведение вектора x и вектора a — является **линейным функционалом**, так как $f(x+y) = (x+y, a) = (x, a) + (y, a) = f(x) + f(y)$ и $f(\alpha x) = (\alpha x, a) = \alpha(x, a) = \alpha f(x)$.

§ 2. Оператор, сопряженный данному

Лемма. *Если в евклидовом пространстве R $(x, u) = (x, v)$ для всех векторов x, то $u = v$.*

Доказательство. Из равенства $(x, u) = (x, v)$ вытекает, что $(x, u-v) = 0$ при всех x. Подставив $x = u - v$, получим $(u-v, u-v) = 0$. Но так как пространство R евклидово, то $u - v = 0$ и $u = v$.

Пусть R — евклидово пространство и $\mathscr{A}$ — линейный оператор в нем. Покажем, что *при фиксированном y скалярное произведение $f_y(x) = (\mathscr{A}x, y)$ является линейным функционалом относительно x. Действительно,*

$$f_y(x_1 + x_2) = (\mathscr{A}(x_1 + x_2), y) = (\mathscr{A}x_1 + \mathscr{A}x_2, y) = \\ = (\mathscr{A}x_1, y) + (\mathscr{A}x_2, y) = f_y(x_1) + f_y(x_2)$$

и

$$f_y(\alpha x) = (\mathscr{A}(\alpha x), y) = (\alpha\mathscr{A}x, y) = \alpha(\mathscr{A}x, y) = \alpha f_y(x).$$

Как показано в § 1, найдется такой вектор y' из R, что при всех x $f_y(x) = (\mathscr{A}x, y) = (x, y')$. Этот вектор y' зависит только от y (не от x!) и можно положить поэтому $y' = \mathscr{A}^*y$. Вектор $\mathscr{A}^*y$ определяется вектором y, т. е. $\mathscr{A}^*$ — **оператор**, переводящий вектор y в новый вектор y' (который мы и обозначаем $\mathscr{A}^*y$). Покажем, что этот оператор — линейный. Действительно, при всех $x, y, z \in R$ мы имеем

$$(\mathscr{A}x, y+z) = (x, \mathscr{A}^*(y+z))$$

и

$$(\mathscr{A}x, y+z) = (\mathscr{A}x, y) + (\mathscr{A}x, z) = \\ = (x, \mathscr{A}^*y) + (x, \mathscr{A}^*z) = (x, \mathscr{A}^*y + \mathscr{A}^*z),$$

откуда $(x, \mathcal{A}^*(y+z)) = (x, \mathcal{A}^*y + \mathcal{A}^*z)$ и, ввиду леммы,

$$\mathcal{A}^*(y+z) = \mathcal{A}^*(y) + \mathcal{A}^*(z).$$

Аналогично, если $\alpha \in F$, то для любых $x, y \in R$ имеем

$$(\mathcal{A}x, \alpha y) = (x, \mathcal{A}^*(\alpha y))$$

и

$$(\mathcal{A}x, \alpha y) = \bar{\alpha}(\mathcal{A}x, y) = \bar{\alpha}(x, \mathcal{A}^*y) = (x, \alpha\mathcal{A}^*y),$$

откуда, по той же лемме,

$$\mathcal{A}^*(\alpha y) = \alpha\mathcal{A}^*y.$$

Определение 2. *Линейный оператор $\mathcal{A}^*$ такой, что при всех $x, y \in R$*

$$(\mathcal{A}x, y) = (x, \mathcal{A}^*y),$$

называется сопряженным $\mathcal{A}$.

Легко видеть, что оператор, сопряженный $\mathcal{A}$,— единственный, так как из равенства $(x, \mathcal{B}y) = (x, \mathcal{C}y)$, справедливого при всех $x, y \in R$, вытекает (по той же лемме), что $\mathcal{B}y = \mathcal{C}y$ при всех y и, значит, $\mathcal{B} = \mathcal{C}$.

Пусть $A = [a_{ik}]$ — матрица линейного оператора $\mathcal{A}$ в ортонормированном базисе $e_1, e_2, \ldots, e_n$, $A' = [a_{ki}]$ — матрица, транспонированная к A, $A^* = \bar{A}' = [\bar{a}_{ki}]$ — матрица, элементы которой комплексно-сопряжены элементам матрицы A'. Обозначим через $\mathcal{A}_1$ линейный оператор, имеющий в том же базисе матрицу A^*, и покажем, что $\mathcal{A}_1 = \mathcal{A}^*$, т. е. что $\mathcal{A}_1$ и есть оператор, сопряженный $\mathcal{A}$. Мы имеем, очевидно,

$$(\mathcal{A}e_i, e_k) = (a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n, e_k) = a_{ki}$$

и

$$(e_i, \mathcal{A}_1e_k) = (e_i, \bar{a}_{k1}e_1 + \bar{a}_{k2}e_2 + \ldots + \bar{a}_{kn}e_n) = a_{ki},$$

т. е. $(\mathcal{A}e_i, e_k) = (e_i, \mathcal{A}_1e_k)$ при всех i, k. А тогда, если

$$x = \sum_{i=1}^{n} x_ie_i \text{ и } y = \sum_{k=1}^{n} y_ke_k,$$

то

$$(\mathscr{A}x, y) = \left(\mathscr{A}\sum_{i=1}^{n} x_i e_i, \sum_{k=1}^{n} y_k e_k\right) = \sum_{i,k=1}^{n} x_i \bar{y}_k (\mathscr{A}e_i, e_k)$$

и

$$(x, \mathscr{A}_1 y) = \left(\sum_{i=1}^{n} x_i e_i, \mathscr{A}_1 \sum_{k=1}^{n} y_k e_k\right) =$$

$$= \sum_{i,k=1}^{n} x_i \bar{y}_k (e_i, \mathscr{A}_1 e_k) = \sum_{i,k=1}^{n} x_i \bar{y}_k (\mathscr{A}e_i, e_k) = (\mathscr{A}x, y),$$

т. е. для всех x, y $(\mathscr{A}x, y) = (x, \mathscr{A}_1 y)$ и оператор $\mathscr{A}_1$ является сопряженным $\mathscr{A}$, т. е. $\mathscr{A}_1 = \mathscr{A}^*$.

Таким образом, *для каждого линейного оператора* $\mathscr{A}$ *в евклидовом пространстве существует и притом только один сопряженный ему линейный оператор* $\mathscr{A}^*$, *матрица которого в любом ортонормированном базисе является транспонированной и комплексно-сопряженной матрице оператора* $\mathscr{A}$.

Покажем, что $\mathscr{A}^{**} = \mathscr{A}$. Действительно, имеем

$$(x, \mathscr{A}^{**}y) = (\mathscr{A}^*x, y) = \overline{(y, \mathscr{A}^*x)} = \overline{(\mathscr{A}y, x)} = (x, \mathscr{A}y)$$

при всех x, y, и значит, опять по той же лемме, $\mathscr{A}^{**}y = \mathscr{A}y$ при всех y, т. е. $\mathscr{A}^{**} = \mathscr{A}$.

Свойства оператора, сопряженного данному.

1. $\mathscr{E}^* = \mathscr{E}$, так как

$$(x, \mathscr{E}^*y) = (\mathscr{E}x, y) = (x, y) = (x, \mathscr{E}y),$$

и, согласно лемме, $\mathscr{E}^* = \mathscr{E}$.

2. $(\mathscr{A} + \mathscr{B})^* = \mathscr{A}^* + \mathscr{B}^*$, так как

$$(x, (\mathscr{A}+\mathscr{B})^*y) = ((\mathscr{A}+\mathscr{B})x, y) = (\mathscr{A}x + \mathscr{B}x, y) =$$
$$= (\mathscr{A}x, y) + (\mathscr{B}x, y) = (x, \mathscr{A}^*y) + (x, \mathscr{B}^*y) =$$
$$= (x, \mathscr{A}^*y + \mathscr{B}^*y) = (x, (\mathscr{A}^* + \mathscr{B}^*)y),$$

и, по лемме, $(\mathscr{A} + \mathscr{B})^* = \mathscr{A}^* + \mathscr{B}^*$.

3. $(\mathscr{A}\mathscr{B})^* = \mathscr{B}^*\mathscr{A}^*$, так как

$$(x, (\mathscr{A}\mathscr{B})^*y) = ((\mathscr{A}\mathscr{B})x, y) = (\mathscr{A}(\mathscr{B}x), y) =$$
$$= (\mathscr{B}x, \mathscr{A}^*y) = (x, \mathscr{B}^*\mathscr{A}^*y);$$

по лемме, $(\mathscr{A}\mathscr{B})^* = \mathscr{B}^*\mathscr{A}^*$.

4. *Если $\mathcal{A}^{-1}$ существует, то* $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$, так как из равенства $(\mathcal{A}\mathcal{A}^{-1}) = \mathcal{E}$ и пп. 3 и 1 вытекает, что $(\mathcal{A}\mathcal{A}^{-1})^* = \mathcal{E}^*$, или $(\mathcal{A}^{-1})^*\mathcal{A}^* = \mathcal{E}$, т. е. что $(\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$.

5. *Если α — число, то* $(\alpha\mathcal{A})^* = \bar{\alpha}\mathcal{A}^*$, так как

$$(x, (\alpha\mathcal{A})^*y) = (\alpha\mathcal{A}x, y) = \alpha(\mathcal{A}x, y) =$$
$$= \alpha(x, \mathcal{A}^*y) = (x, \bar{\alpha}\mathcal{A}^*y),$$

и, по лемме, $(\alpha\mathcal{A})^* = \bar{\alpha}\mathcal{A}^*$.

Теорема 1. *Если подпространство R_1 инвариантно относительно линейного оператора $\mathcal{A}$, то его ортогональное дополнение $R_1^{\perp}$ инвариантно относительно сопряженного $\mathcal{A}$ оператора $\mathcal{A}^*$.*

Доказательство. Пусть x — произвольный вектор из $R_1^{\perp}$, y — произвольный вектор из R_1. Тогда

$$(\mathcal{A}^*x, y) = (x, \mathcal{A}y) = 0,$$

так как $\mathcal{A}y \in R_1$ и, значит, $x \perp \mathcal{A}y$. Следовательно, вектор $\mathcal{A}^*x \in R_1^{\perp}$, и $R_1^{\perp}$ инвариантно относительно $\mathcal{A}^*$.

Пусть $f(t)$ — произвольный многочлен с комплексными, вообще говоря, коэффициентами. Обозначим через $\bar{f}(t)$ многочлен, все коэффициенты которого являются комплексно-сопряженными к соответствующим коэффициентам многочлена $f(t)$. Так, если $f(t) = (1+i)t^2 + (2-i)t + 1$, то $\bar{f}(t) = (1-i)t^2 + (2+i)t + 1$; если $f(t) = 2t^2 + 3t - 5$, то $\bar{f}(t) = f(t)$, и т. д.

Теорема 2. *Если $\varphi(\lambda)$ — характеристический многочлен линейного оператора $\mathcal{A}$, то характеристическим многочленом сопряженного $\mathcal{A}$ оператора $\mathcal{A}^*$ будет $\bar{\varphi}(\lambda)$.*

Доказательство. Пусть

$$\varphi(\lambda) = |A - \lambda E| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots a_{1n} \\ a_{21} & a_{22}-\lambda & \dots a_{2n} \\ \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots a_{nn}-\lambda \end{vmatrix}.$$

Тогда характеристический многочлен оператора $\mathcal{A}^*$ равен

$$|A^* - \lambda E| = |\bar{A}' - \lambda E| = |\bar{A}' - \lambda E'| = |(\bar{A} - \lambda E)'| =$$
$$= |\bar{A} - \lambda E| = \begin{bmatrix} \bar{a}_{11}-\lambda & \bar{a}_{12} & \dots \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22}-\lambda & \dots \bar{a}_{2n} \\ \dots & \dots & \dots \\ \bar{a}_{n1} & \bar{a}_{n2} & \dots \bar{a}_{nn}-\lambda \end{bmatrix} = \bar{\varphi}(\lambda).$$

Следствие. *Если λ_1 — собственное значение оператора $\mathscr{A}$ кратности k,* то $\bar{\lambda}_1$ — *собственное значение оператора $\mathscr{A}^*$ той же кратности k.*

Действительно, если

$$\varphi(\lambda) = (-1)^n (\lambda - \lambda_1)^k (\lambda - \lambda_2)^{\cdot} \ldots (\lambda - \lambda_m)^{\cdot},$$

где $\lambda_2, \lambda_3, \ldots, \lambda_m \neq \lambda_1$, то

$$\bar{\varphi}(\lambda) = (-1)^n (\lambda - \bar{\lambda}_1)^k (\lambda - \bar{\lambda}_2)^{\cdot} \ldots (\lambda - \bar{\lambda}_m)^{\cdot},$$

где тоже

$$\bar{\lambda}_2, \bar{\lambda}_3, \ldots, \bar{\lambda}_m \neq \bar{\lambda}_1.$$

В частности, в вещественном пространстве R характеристический многочлен линейного оператора $\mathscr{A}^*$ равен характеристическому многочлену оператора $\mathscr{A}$ — и все их собственные значения одинаковы (т. е. спектры их тождественны).

§ 3. Самосопряженный оператор

Определение 3. *Линейный оператор $\mathscr{A}$, совпадающий со своим сопряженным, т. е. такой, что $\mathscr{A}^* = \mathscr{A}$, называется самосопряженным.* В вещественном пространстве самосопряженный оператор называют также *симметрическим*, а в комплексном пространстве — *эрмитовым.*

Таким образом, если $\mathscr{A}$ — самосопряженный оператор, то тождественно при всех x и y из R

$$(\mathscr{A}x, y) = (x, \mathscr{A}y).$$

Свойства самосопряженных операторов.

1. *Тождественный оператор является самосопряженным*, так как $\mathscr{E}^* = \mathscr{E}$.

2 *Сумма самосопряженных операторов является самосопряженным оператором*, так как если $\mathscr{A}^* = \mathscr{A}$ и $\mathscr{B}^* = \mathscr{B}$, то $(\mathscr{A} + \mathscr{B})^* = \mathscr{A}^* + \mathscr{B}^* = \mathscr{A} + \mathscr{B}$.

3. *Для того чтобы произведение самосопряженных операторов было самосопряженным оператором, необходимо и достаточно, чтобы эти операторы были перестановочны между собой*, т. е. чтобы имело место равенство $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$. Действительно, если $\mathscr{A}^* = \mathscr{A}$ и $\mathscr{B}^* = \mathscr{B}$, то $(\mathscr{A}\mathscr{B})^* = \mathscr{B}^*\mathscr{A}^* = \mathscr{B}\mathscr{A}$, что равно $\mathscr{A}\mathscr{B}$

в том и только в том случае, если операторы $\mathcal{A}$ и $\mathcal{B}$ перестановочны.

4. *Оператор, обратный к невырожденному самосопряженному оператору, является самосопряженным*, так как если $\mathcal{A}^* = \mathcal{A}$, то $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1} = \mathcal{A}^{-1}$.

5. *Если $\mathcal{A}$ — самосопряженный оператор, то для того, чтобы произведение $\alpha\mathcal{A}$ было самосопряженным, необходимо и достаточно, чтобы число α было вещественным, так как в этом случае $(\alpha\mathcal{A})^* = \bar{\alpha}\mathcal{A}^* = \alpha\mathcal{A}$.*

Теорема 3. *Если $\mathcal{A}$ — самосопряженный оператор и R_1 — подпространство, инвариантное относительно $\mathcal{A}$, то и $R_1^{\perp}$ инвариантно относительно $\mathcal{A}$.*

Доказательство. По теореме 1 $R_1^{\perp}$ инвариантно относительно $\mathcal{A}^*$, но $\mathcal{A}^* = \mathcal{A}$, следовательно, $R_1^{\perp}$ инвариантно относительно $\mathcal{A}$.

Далее рассмотрим отдельно самосопряженные операторы в вещественном и в комплексном векторных пространствах.

А. Пространство R вещественно.

Пусть $\mathcal{A}$ — самосопряженный (симметрический) оператор в вещественном векторном пространстве и $A = [a_{ik}]$ — его матрица в ортонормированном базисе. Тогда матрицей оператора $\mathcal{A}^*$ в том же базисе будет транспонированная к A матрица $A' = [a_{ki}]$ (см. § 2), и так как $\mathcal{A}^* = \mathcal{A}$, то $A' = A$, т. е. $a_{ik} = a_{ki}$ при всех i, k. Обладающая этим свойством матрица A называется симметрической (она «симметрична относительно главной диагонали»). Пример симметрической матрицы:

$$\begin{bmatrix} 1 & 1+\sqrt{2} \\ 1+\sqrt{2} & \pi \end{bmatrix}.$$

Обратно, *линейный оператор, имеющий в ортонормированном базисе симметрическую матрицу, будет*, очевидно, *самосопряженным.*

Теорема 4. *Все корни характеристического многочлена самосопряженного оператора $\mathcal{A}$ вещественны.*

Доказательство. Пусть $\lambda = \alpha + i\beta$ - комплексный корень характеристического многочлена самосопряженного оператора A. Тогда, как видно из доказательства теоремы 8 главы III, в пространстве R имеется двумерное подпространство, порожденное векторами u и v такими, что

$$\begin{cases} \mathcal{A}u = \alpha u - \beta v, \\ \mathcal{A}v = \beta u + \alpha v, \end{cases} \tag{1}$$

где $\beta \neq 0$ и векторы u и v не равны нулю. (Если само пространство R двумерно и в нем нет собственных векторов, то это подпространство совпадает с R.) Умножая скалярно первое из равенств (1) на v, второе — на u, получим

$$(\mathcal{A}u, v) = \alpha(u, v) - \beta(v, v)$$

и

$$(u, \mathcal{A}v) = \beta(u, u) + \alpha(u, v).$$

Но так как $(\mathcal{A}u, v) = (u, \mathcal{A}v)$, то $\beta(|u|^2 + |v|^2) = 0$, и либо $|u|^2 + |v|^2 = 0$, либо $\beta = 0$, что противоречит предположению.

Теорема 5. *Матрица самосопряженного оператора в некотором ортонормированном базисе приводится к диагональному виду.*

Доказательство. Пусть λ_1 — один из корней характеристического многочлена самосопряженного оператора $\mathcal{A}$. По теореме 4, λ_1 вещественно. Соответствующий λ_1 собственный вектор обозначим через e_1; тогда $\mathcal{A}e_1 = \lambda_1 e_1$. Вектор e_1 можно считать единичным, так как в противном случае его можно было бы заменить вектором $\frac{e_1}{|e_1|}$ — единичным собственным вектором с тем же собственным значением λ_1.

Обозначим через R_1 одномерное (инвариантное) подпространство, порожденное вектором e_1. Его ортогональное дополнение $R_1^{\perp}$ будет инвариантно относительно $\mathcal{A}$, и в нем оператор $\mathcal{A}$ остается, конечно, самосопряженным. Пусть λ_2 (вещественное) собственное значение оператора $\mathcal{A}$ в подпространстве $R_1^{\perp}$; соответствующий (единичный) собственный вектор обозначим через e_2; тогда

$$\mathcal{A}e_2 = \lambda_2 e_2.$$

Обозначим через R_2 (инвариантное) подпространство, порожденное векторами e_1 и e_2; тогда подпространство

$R_2^\perp$ тоже инвариантно относительно $\mathcal{A}$. Продолжая это построение, мы найдем n попарно ортогональных (и значит, линейно независимых) единичных собственных векторов оператора $\mathcal{A}$. В базисе, состоящем из этих векторов, матрица оператора $\mathcal{A}$ приведется к диагональному виду

$$\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Геометрический смысл самосопряженного преобразования виден из последней теоремы: если

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

— произвольный вектор из R, то

$$Ax = x_1 \lambda_1 e_1 + x_2 \lambda_2 e_2 + \ldots + x_n \lambda_n e_n.$$

Таким образом, при соответствующем $\mathcal{A}$ преобразовании точек точка $X(x_1, x_2, \ldots, x_n)$ переходит в точку

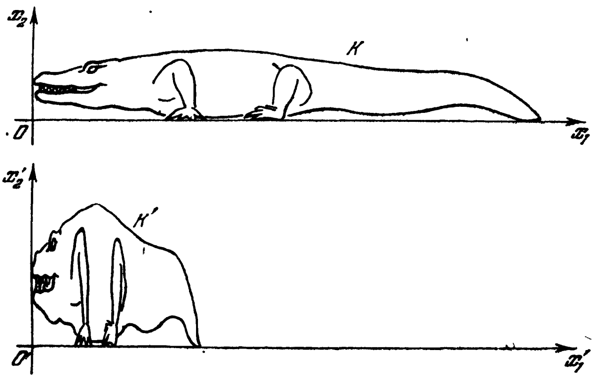


Рис. 13.

$X'(\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n)$ и, значит, в базисе, состоящем из собственных векторов оператора $\mathcal{A}$, оно сводится к n растяжениям вдоль координатных осей с коэффициентами, соответственно равными $\lambda_1, \lambda_2, \ldots, \lambda_n$ (см. рис. 13, на котором изображено действие на фигуру K евклидовой плоскости самосопряженного преобразования с собственными значениями $\lambda_1 = \frac{1}{3}$ и $\lambda_2 = 2$).

Б. Пространство R — комплексное.

Пусть $\mathscr{A}$ — самосопряженный (эрмитов) оператор в комплексном векторном пространстве и $A = [a_{ik}]$ — его матрица в ортонормированном базисе. Тогда $\bar{A}' = A$, т. е. $\bar{a}_{ik} = a_{ki}$ при всех i, k. Такая матрица называется эрмитовой. Таким образом, если A — эрмитова матрица, то ее элементы, симметричные относительно главной диагонали, являются комплексно-сопряженными. В частности, элементы главной диагонали вещественны, так как $a_{ii} = \bar{a}_{ii}$ при всех i. Пример эрмитовой матрицы:

$$\begin{bmatrix} 2 & 1+i & i \\ 1-i & 0 & 3+2i \\ -i & 3-2i & 1 \end{bmatrix}.$$

Итак, *матрица эрмитова оператора в любом ортонормированном базисе является эрмитовой.* Очевидно, что и, обратно, *линейный оператор, матрица которого в каком-то ортонормированном базисе является эрмитовой,— эрмитов.*

Теорема 4′ *Собственные значения самосопряженного (эрмитова) оператора вещественны.*

Доказательство этой теоремы для комплексного пространства совсем просто. В самом деле, пусть x — собственный вектор и λ — соответствующее ему собственное значение эрмитова оператора $\mathscr{A}$, тогда

$$(\mathscr{A}x, x) = (x, \mathscr{A}x),$$

или

$$(\lambda x, x) = (x, \lambda x),$$

откуда

$$\lambda(x, x) = \bar{\lambda}(x, x),$$

и так как $(x, x) \neq 0$, то $\lambda = \bar{\lambda}$, т. е. λ — вещественно.

Таким образом, *спектр самосопряженного оператора* (и в вещественном, и в комплексном пространствах) *расположен на вещественной оси.* Далее так же, как для вещественного пространства, в комплексном случае доказывается.

Теорема 5'. *Матрица самосопряженного (эрмитова) оператора в некотором ортонормированном базисе приводится к диагональному виду* (где все диагональные элементы вещественны).

§ 4. Ортогональный оператор

В этом параграфе евклидово пространство R предполагается вещественным.

Определение 4. *Линейный оператор $\mathscr{A}$ в вещественном евклидовом пространстве R называется ортогональным, если*

$$(\mathscr{A}x, \mathscr{A}y) = (x, y)$$

для всех x, y из R.

Таким образом, ортогональный оператор сохраняет скалярное произведение, и значит, он сохраняет длины векторов и углы между ними.

Если $\mathscr{A}$ — ортогональный оператор и $\mathscr{A}^*$ — сопряженный ему оператор, то

$$(x, y) = (\mathscr{A}x, \mathscr{A}y) - (x, \mathscr{A}^*\mathscr{A}y)$$

для всех x, $y \in R$. Следовательно, $\mathscr{A}^*\mathscr{A} = \mathscr{E}$ — по лемме из § 2, или

$$\mathscr{A}^* = \mathscr{A}^{-1}.$$

При этом равенство $\mathscr{A}^* = \mathscr{A}^{-1}$ является необходимым и достаточным условием для того, чтобы линейный оператор $\mathscr{A}$ был ортогональным. Отсюда, в частности, видно, что ортогональный оператор всегда невырожденный.

Свойства ортогональных операторов.

1. *Тождественный оператор $\mathscr{E}$ является ортогональным,* так как $\mathscr{E}x = x$ для всякого x, и значит,

$$(\mathscr{E}x, \mathscr{E}y) = (x, y).$$

2. *Произведение ортогональных операторов является ортогональным,* так как если операторы $\mathscr{A}$ и $\mathscr{B}$ ортогональны, то

$$(\mathscr{A}\mathscr{B}x, \mathscr{A}\mathscr{B}y) = (\mathscr{B}x, \mathscr{B}y) = (x, y),$$

3. *Оператор, обратный ортогональному оператору, тоже является ортогональным, так как если* $\mathcal{A}^* = \mathcal{A}^{-1}$, то $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^{-1}$ (см. п. 4 из § 2).

4. *Если* $\mathcal{A}$ — *ортогональный оператор, то произведение* $\alpha\mathcal{A}$ *будет ортогональным в том и только в том случае, если* $\alpha = \pm 1$, — это видно из равенства $(\alpha\mathcal{A}x, \alpha\mathcal{A}y) = \alpha^2(\mathcal{A}x, \mathcal{A}y) = \alpha^2(x, y)$.

Ясно, что *ортогональный оператор переводит любой ортонормированный базис в ортонормированный.* Покажем, что верно и обратное: *линейный оператор* $\mathcal{A}$, *переводящий хотя бы один ортонормированный базис в ортонормированный,* является *ортогональным.* Действительно, пусть ортонормированный базис $e_1, e_2, \ldots, e_n$ оператором $\mathcal{A}$ переводится в ортонормированный базис $e_1', e_2', \ldots, e_n'$. Тогда, если

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$

$$y = y_1e_1 + y_2e_2 + \ldots + y_ne_n,$$

то

$$\mathcal{A}x = x_1e_1' + x_2e_2' + \ldots + x_ne_n',$$

$$\mathcal{A}y = y_1e_1' + y_2e_2' + \ldots + y_ne_n'$$

и

$$(\mathcal{A}x, \mathcal{A}y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n = (x, y).$$

Пусть $A = [a_{ik}]$ — матрица ортогонального оператора в ортонормированном базисе $e_1, e_2, \ldots, e_n$. Так как под действием ортогонального оператора ортонормированный базис переходит в ортонормированный, то образы $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n$ базисных векторов $e_1, e_2, \ldots, e_n$ *сами образуют ортонормированный базис.* А значит, $(\mathcal{A}e_i, \mathcal{A}e_k) = 0$ при $i \neq k$ и $(\mathcal{A}e_i, \mathcal{A}e_i) = 1$ при всех i, т. е.

$$a_{1i}a_{1k} + a_{2i}a_{2k} + \ldots + a_{ni}a_{nk} = 0 \text{ при } i \neq k$$

и (2)

$$a_{1i}^2 + a_{2i}^2 + \ldots + a_{ni}^2 = 1 \quad \text{при всех} \quad i.$$

Таким образом, столбцы матрицы A, рассматриваемые как векторы, сами образуют ортонормированную систему. Это же верно и для строк. Действительно, если $\mathcal{A}$ — ортогональный оператор, то $\mathcal{A}^* = \mathcal{A}^{-1}$ — оператор тоже ортогональный, и значит, столбцы матрицы A^*, т. е. строки матрицы A, тоже образуют ортонормированную систему:

$$a_{i1}a_{k1} + a_{i2}a_{k2} + \ldots + a_{in}a_{kn} = 0 \text{ при } i \neq k$$

и (3)

$$a_{i1}^2 + a_{i2}^2 + \ldots + a_{in}^2 = 1 \quad \text{при всех} \quad i.$$

Матрица A, для которой

$$A' = A^{-1},$$

называется *ортогональной матрицей;* она характеризуется соотношениями (2) и (равносильными им) соотношениями (3). Мы показали, что *матрица ортогонального оператора в любом ортонормированном базисе является ортогональной;* обратно, если в каком-то ортонормированном базисе матрица оператора $\mathscr{A}$ ортогональна: $A^{-1} = A'$, то $\mathscr{A}^{-1} = \mathscr{A}^*$, и оператор $\mathscr{A}$ является ортогональным.

Теорема 6. *Если подпространство R_1 инвариантно относительно ортогонального оператора $\mathscr{A}$, то его ортогональное дополнение $R_1^\perp$ тоже инвариантно относительно $\mathscr{A}$.*

Доказательство. Так как $\mathscr{A}$ — ортогональный оператор, то $\mathscr{A}^{-1} = \mathscr{A}^*$. По теореме 1, подпространство $R_1^\perp$ инвариантно относительно оператора $\mathscr{A}^* = \mathscr{A}^{-1}$; но в таком случае в силу теоремы 5 главы III оно инвариантно и относительно $(\mathscr{A}^{-1})^{-1} = \mathscr{A}$.

Теорема 7. *Собственные значения ортогонального оператора равны ± 1.*

Доказательство. Пусть x — собственный вектор и λ — соответствующее ему собственное значение ортогонального оператора $\mathscr{A}$. Тогда

$$(x, x) = (\mathscr{A}x, \mathscr{A}x) = (\lambda x, \lambda x) = \lambda^2 (x, x),$$

откуда получаем (поскольку $(x, x) \neq 0$)

$$\lambda^2 = 1 \quad \text{и} \quad \lambda = \pm 1.$$

Теорема 8. *Определитель ортогональной матрицы равен ± 1.*

Доказательство. Из равенства $AA' = E$ следует, что $|A|\,|A'| = |AA'| = |E| = 1$. Но так как $|A'| = |A|$, то $|A|^2 = 1$ и $|A| = \pm 1$.

Выясним, что собой представляет произвольный ортогональный оператор, действующий в (вещественном) n-мерном евклидовом пространстве. Пусть сначала $\mathcal{A}$ — ортогональное преобразование прямой R^1 и $e \in R^1$. Тогда $\mathcal{A}e \in R^1$ и, значит, $\mathcal{A}e = \lambda e$, где $\lambda = \pm 1$, т. е. $\mathcal{A}e = \pm e$, и $\mathcal{A}$ — либо *тождественное преобразование*, либо *центральная симметрия*.

Пусть теперь $\mathcal{A}$ — ортогональное преобразование плоскости R^2 и

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

— его матрица в некотором ортонормированном базисе. Тогда, как мы знаем,

$$a_{11}^2 + a_{21}^2 = 1, \quad a_{12}^2 + a_{22}^2 = 1, \quad a_{11}a_{12} + a_{21}a_{22} = 0.$$

В силу первых двух равенств найдутся такие φ и ψ, что $a_{11} = \cos\varphi$, $a_{21} = \sin\varphi$, и $a_{12} = \cos\psi$, $a_{22} = \sin\psi$. Но тогда третье равенство дает

$$\cos\varphi \cdot \cos\psi + \sin\varphi \cdot \sin\psi = \cos(\psi - \varphi) = 0,$$

откуда следует, что

$$\psi - \varphi = \frac{\pi}{2} \quad \text{или} \quad \frac{3\pi}{2}.$$

В первом случае $a_{12} = \cos\psi = -\sin\varphi$, $a_{22} = \sin\psi = \cos\varphi$, и мы имеем

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}, \tag{4a}$$

т. е. преобразование $\mathcal{A}$ — *это поворот на угол φ вокруг начала координат*. (В частности, при $\varphi = 0$ это *тождественное преобразование*, а при $\varphi = \pi$ — *симметрия относительно начала координат*.)

Во втором случае $a_{12} = \sin\varphi$, $a_{22} = -\cos\varphi$ и

$$A = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}.$$

Эта матрица — симметрическая, значит, ортогональное преобразование $\mathcal{A}$ является и самосопряженным; т. е. в некотором (вообще говоря, новом) ортонормированном базисе e_1, e_2 его матрица приводится к диагональному виду. Но так как собственные значения здесь находятся непосредственно:

$$\varphi(\lambda) = \begin{bmatrix} \cos\varphi - \lambda & \sin\varphi \\ \sin\varphi & -\cos\varphi - \lambda \end{bmatrix} = \lambda^2 - 1,$$

откуда $\lambda = \pm 1$, то матрица преобразования $\mathcal{A}$ приведется к виду

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \qquad (4б)$$

Произвольный вектор x, в новом базисе равный $x_1e_1 + x_2e_2$, преобразуется в $x' = x_1e_1 - x_2e_2$. Это — *симметрия относительно прямой,* определяемой вектором e_1 — первым базисным вектором нового базиса (рис. 14).

Таким образом, *ортогональное преобразование плоскости — это либо п о в о р о т вокруг начала координат на некоторый угол* φ (в частности, т о ж д е с т в е н н о е п р е о б р а з о в а н и е или ц е н т р а л ь н а я с и м м е т р и я; определитель такого преобразования равен $+1$), *либо — о с е в а я с и м м е т р и я* (с определителем, равным -1).

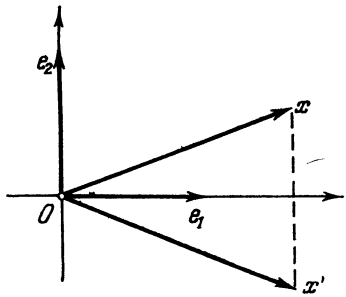


Рис. 14.

Из доказанного, в частности, вытекают две теоремы (плоской) элементарной геометрии:

1. *Произведение двух осевых симметрий является поворотом вокруг точки пересечения осей симметрии* (так как это — ортогональное преобразование с определителем, равным $+1$).

2. *Произведение поворота и симметрии, ось которой проходит через центр поворота, является симметрией относительно некоторой новой оси, проходящей через ту же точку* (так как это — ортогональное преобразование с определителем, равным -1).

Перейдем теперь к общему случаю ортогонального оператора, действующего в n-мерном евклидовом пространстве.

Теорема 9. *Матрица ортогонального оператора в некотором ортонормированном базисе приводится к виду*

$$\begin{bmatrix} 1 &&&&&&&&&&&& \\ & 1 &&&&&&&&&&& \\ && \ddots &&&&&&&&&& \\ &&& 1 &&&&&&&&& \\ &&&& -1 &&&&&&&& \\ &&&&& \ddots &&&&&&& \\ &&&&&& -1 &&&&&& \\ &&&&&&& \cos\varphi_1 & -\sin\varphi_1 &&&& \\ &&&&&&& \sin\varphi_1 & \cos\varphi_1 &&&& \\ &&&&&&&&& \cos\varphi_2 & -\sin\varphi_2 && \\ &&&&&&&&& \sin\varphi_2 & \cos\varphi_2 && \\ &&&&&&&&&&& \ddots & \\ &&&&&&&&&&&& \cos\varphi_k & -\sin\varphi_k \\ &&&&&&&&&&&& \sin\varphi_k & \cos\varphi_k \end{bmatrix}. \quad (5)$$

(Все остальные элементы этой матрицы равны нулю.)

Доказательство проведем методом математической индукции. Мы уже установили справедливость этой теоремы при $n = 1$ и $n = 2$. Предположим, что теорема верна для всех пространств, размерность которых меньше n, и пусть R — n-мерное евклидово пространство и $\mathcal{A}$ — действующий в нем ортогональный оператор. Возможны два случая.

1. *Оператор $\mathcal{A}$ имеет вещественное собственное значение* (это обязательно будет так, если n нечетно): $\lambda = \pm 1$. Пусть e_1 — соответствующий (единичный) собственный вектор (тогда $\mathcal{A}e_1 = \pm e_1$) и R_1 — порожденное вектором e_1 одномерное подпространство. В силу теоремы 6, $(n-1)$-мерное подпространство $R_1^{\perp}$ инвариантно относительно $\mathcal{A}$. Ясно, что и в нем $\mathcal{A}$ будет ортогональным оператором. По предположению индукции, в $R_1^{\perp}$ можно найти ортонормированный базис e_2, $e_3, \ldots, e_n$, в котором матрица оператора $\mathcal{A}$ приведется к виду (5). Учитывая замечание, сделанное в § 7 главы III, получаем, что (возможно, после соответствующего изменения нумерации базисных векторов) матрица оператора $\mathcal{A}$ во всем пространстве R в некотором ортонормированном базисе приведется к виду (5).

2. *Оператор $\mathcal{A}$ не имеет вещественных собственных значений.* По теореме 8 из главы III, в R найдется двумерное инвариантное подпространство R_1. По доказанному выше, в плоскости R_1 можно найти ортонормиро-

ванный базис e_1, e_2, в котором матрица оператора $\mathscr{A}$ приведется к виду (4а). (Другой случай, (4б), здесь невозможен, так как, по предположению, оператор $\mathscr{A}$ не имеет вещественных собственных значений.)

Подпространство $R_1^{\perp}$ инвариантно относительно $\mathscr{A}$. По предположению индукции, в $R_1^{\perp}$ можно найти такой ортонормированный базис $e_3, e_4, \ldots, e_n$, в котором матрица оператора $\mathscr{A}$ приведется к виду (5). (В этом случае n обязательно четно, и на главной диагонали этой матрицы совсем не будет чисел 1 и -1.) Ввиду замечания из § 7 главы III матрица оператора $\mathscr{A}$ всего пространства R в ортонормированном базисе $e_1, e_2, \ldots, e_n$ приведется к виду (5).

Геометрический смысл ортогонального преобразования виден из последней теоремы. Так как каждая матрица вида (5) является произведением нескольких матриц вида

$$\begin{bmatrix} 1 & & & & & & & & & \\ & 1 & & & & & & & & \\ & & \cdot & & & & & & & \\ & & & \cdot & & & & & & \\ & & & & 1 & & & & & \\ & & & & & -1 & & & & \\ & & & & & & 1 & & & \\ & & & & & & & \cdot & & \\ & & & & & & & & \cdot & \\ & & & & & & & & & 1 \end{bmatrix} \tag{6}$$

и нескольких матриц вида

$$\begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cdot\ 1 & & & & & \\ & & & & \cos\varphi & -\sin\varphi & & & \\ & & & & \sin\varphi & \cos\varphi & & & \\ & & & & & & 1 & & \\ & & & & & & & \cdot & \\ & & & & & & & & 1 \end{bmatrix}, \tag{7}$$

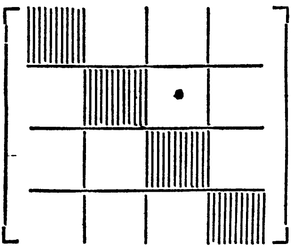
то ортогональное преобразование $\mathscr{A}$ можно осуществить, произведя последовательно несколько *симметрий* *относительно «координатных гиперплоскостей»* (матрица каждого такого преобразования имеет вид (6)) и несколько *поворотов вокруг «$(n-2)$-мерных осей»* (каждый из которых имеет матрицу вида (7))—это преобразование представляет собой одинаковый поворот, осуществляемый одновременно во всех двумерных плоскостях, перпендикулярных к $(n-2)$-мерной «оси» поворота.

Объединяя в матрице (5) два соседних элемента $+1$ или -1 в «клетки»

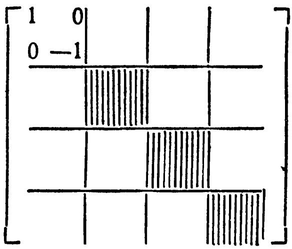
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos 0 & -\sin 0 \\ \sin 0 & \cos 0 \end{bmatrix} \text{ и } \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{bmatrix},$$

мы (возможно, после изменения нумерации базисных векторов) получим четыре типа ортогональных матриц (заштрихованы «клетки» вида $\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$, где φ, в частности, может равняться нулю или π, а в пустых клетках все элементы равны нулю):

Для четного n:

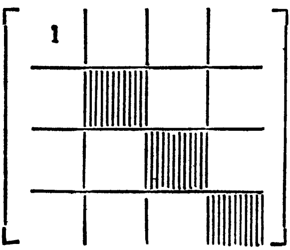


и

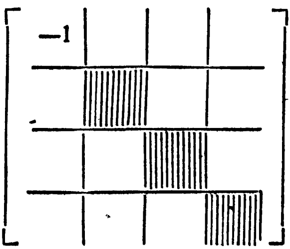


если $|A|=1$, если $|A|=-1$.

Для нечетного n:



и



если $|A|=1$, если $|A|=-1$.

§ 5. Унитарный оператор

В этом параграфе евклидово пространство R предполагается комплексным.

Определение 4'. *Линейный оператор $\mathscr{A}$, действующий в комплексном евклидовом пространстве, называется унитарным, если*

$$(\mathscr{A}x, \mathscr{A}y) = (x, y)$$

для всех x, y из R.

Таким образом, унитарный оператор является аналогом ортогонального оператора. Так же как и ортогональный оператор (в вещественном пространстве), он сохраняет длины векторов и ортогональные векторы переводит в ортогональные. В частности, *любой ортонормированный базис унитарный оператор переводит в ортонормированный базис.* Верно и обратное: *линейный оператор, преобразующий хотя бы один ортонормированный базис в ортонормированный, является унитарным.* Легко видеть, что если оператор $\mathscr{A}$ — унитарный, то $\mathscr{A}^* = \mathscr{A}^{-1}$, и обратно.

Свойства 1—3 ортогональных операторов (см. стр. 173) переносятся на унитарные операторы без изменений. Фактически сохраняется и свойство 4:

4. Если $\mathscr{A}$ — унитарный оператор, то для того, *чтобы оператор $\alpha\mathscr{A}$ был унитарным, необходимо и достаточно, чтобы α по модулю было равно* 1, ибо

$$(\alpha\mathscr{A}x, \alpha\mathscr{A}y) = \alpha\bar{\alpha}(\mathscr{A}x, \mathscr{A}y) = |\alpha|^2 (x, y).$$

Пусть A — матрица унитарного оператора $\mathscr{A}$ в ортонормированном базисе $e_1, e_2, \ldots, e_n$. Тогда образы $\mathscr{A}e_1, \mathscr{A}e_2, \ldots, \mathscr{A}e_n$ базисных векторов $e_1, e_2, \ldots, e_n$ сами образуют ортонормированный базис: $(\mathscr{A}e_i, \mathscr{A}e_k) = 0$ при $i \neq k$ и $(\mathscr{A}e_i, \mathscr{A}e_i) = 1$, т. е.

$$a_{1i}\bar{a}_{1k} + a_{2i}\bar{a}_{2k} + \ldots + a_{ni}\bar{a}_{nk} = 0 \text{ при } i \neq k,$$

$$a_{1i}\bar{a}_{1i} + a_{2i}\bar{a}_{2i} + \ldots + a_{ni}\bar{a}_{ni} = |a_{1i}|^2 + |a_{2i}|^2 + \ldots \qquad (2')$$
$$\ldots + |a_{ni}|^2 = 1 \text{ при всех } i.$$

Далее, если $\mathscr{A}$ — унитарный оператор, то оператор $\mathscr{A}^* = \mathscr{A}^{-1}$ — тоже унитарный, и значит, столбцы *матри-*

цы A^*, т. е. строки матрицы $\overline{A}$, тоже образуют ортонормированную систему:

$$\overline{a}_{k1}a_{i1}+\overline{a}_{k2}a_{i2}+\ldots+\overline{a}_{kn}a_{in}=0 \quad \text{при} \quad i\neq k,$$
$$\overline{a}_{i1}a_{i1}+\overline{a}_{i2}a_{i2}+\ldots+\overline{a}_{in}a_{in}= \tag{3'}$$
$$=|a_{i1}|^2+|a_{i2}|^2+\ldots+|a_{in}|^2=1$$

при $i=1, 2, \ldots, n$.

Матрица A, для которой $A^*=A^{-1}$, т. е. матрица, элементы которой удовлетворяют условиям (2′) (или равносильным им условиям (3′)), называется *унитарной* матрицей. Таким образом, *матрица унитарного оператора в любом ортонормированном базисе является унитарной*. Обратно, *если в каком-то ортонормированном базисе матрица оператора $\mathcal{A}$ унитарна: $A^*=A^{-1}$*, то $\mathcal{A}^*=\mathcal{A}^{-1}$ и *оператор $\mathcal{A}$ является унитарным.*

Теорема 6 переносится на унитарные операторы без изменений: *ортогональное дополнение $R_1^{\perp}$ подпространства R_1, инвариантного относительно унитарного оператора $\mathcal{A}$, инвариантно относительно $\mathcal{A}$*. Теорема 7 принимает такой вид:

Теорема 7′. *Собственные значения унитарного оператора по модулю равны* 1.

Доказательство. Пусть x — собственный вектор и λ — соответствующее собственное значение унитарного оператора $\mathcal{A}$; тогда $\mathcal{A}x=\lambda x$ и $(\mathcal{A}x, \mathcal{A}x)=(x, x)$.

Но $(\mathcal{A}x, \mathcal{A}x)=(\lambda x, \lambda x)=\lambda\overline{\lambda}(x, x)=|\lambda|^2(x, x)$; а так как $(x, x)\neq 0$ (x — собственный вектор, и значит, $x\neq 0$), то $|\lambda|^2=1$, или $|\lambda|=1$.

Таким образом, *спектр унитарного оператора расположен на единичной окружности комплексной плоскости.*

Новой является следующая

Теорема 10. *Матрица унитарного оператора $\mathcal{A}$ комплексного евклидова пространства R в некотором ортонормированном базисе приводится к диагональному виду* (где все элементы главной диагонали по модулю равны 1).

Доказательство. Пусть λ_1 — одно из собственных значений (унитарного) оператора $\mathcal{A}$. По теореме 7′, $|\lambda_1|=1$. Соответствующий λ_1 (единичный) собственный вектор обозначим через e_1. Тогда $\mathcal{A}e_1=\lambda_1e_1$.

Пусть R_1 — одномерное подпространство, порожденное вектором e_1. Его ортогональное дополнение $R_1^{\perp}$ инвариантно относительно $\mathscr{A}$.

Если, далее, λ_2 (где $|\lambda_2| = 1$) — собственное значение оператора $\mathscr{A}$ в подпространстве $R_1^{\perp}$ и e_2 — соответствующий (единичный) собственный вектор, то $\mathscr{A}e_2 = \lambda_2 e_2$. Обозначим через R_2 (инвариантное) подпространство, порожденное векторами e_1 и e_2. Тогда подпространство $R_2^{\perp}$ тоже инвариантно относительно $\mathscr{A}$. Продолжая это построение, мы найдем n попарно ортогональных (и, следовательно, линейно независимых) единичных векторов $e_1, e_2, \ldots, e_n$ — собственных векторов оператора $\mathscr{A}$. В базисе, состоящем из этих векторов, матрица оператора $\mathscr{A}$ имеет диагональный вид

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & \lambda_n \end{bmatrix}.$$

Все элементы, стоящие на главной диагонали этой матрицы, по модулю равны 1.

Отсюда, в частности, видно, что определитель матрицы унитарного оператора в любом базисе (он ведь не зависит от базиса!) по модулю равен 1 (ср. с теоремой 8).

§ 6. Произвольный линейный оператор в евклидовом пространстве

Теорема 11. *Всякий линейный оператор $\mathscr{A}$ в комплексном евклидовом пространстве можно представить в виде $\mathscr{A} = \mathscr{B} + i\mathscr{C}$, где $\mathscr{B}$ и $\mathscr{C}$ — эрмитовы операторы.*

Доказательство. Допустим, что такое представление возможно; тогда

$$\mathscr{A}^* = \mathscr{B}^* + (i\mathscr{C})^* = \mathscr{B}^* - i\mathscr{C}^* = \mathscr{B} - i\mathscr{C},$$

так как $\mathscr{B}^* = \mathscr{B}$ и $\mathscr{C}^* = \mathscr{C}$. Из равенств $\mathscr{A} = \mathscr{B} + i\mathscr{C}$ и $\mathscr{A}^* = \mathscr{B} - i\mathscr{C}$ находим, что

$$\mathscr{B} = \frac{1}{2}(\mathscr{A} + \mathscr{A}^*) \quad \text{и} \quad \mathscr{C} = \frac{i}{2}(\mathscr{A}^* - \mathscr{A}).$$

Легко видеть, что операторы

$$\mathcal{B}=\frac{1}{2}(\mathcal{A}+\mathcal{A}^*) \text{ и } \mathcal{C}=\frac{i}{2}(\mathcal{A}^*-\mathcal{A})$$

действительно являются самосопряженными и что $\mathcal{A}=\mathcal{B}+i\mathcal{C}$.

Представление $\mathcal{A}=\mathcal{B}+i\mathcal{C}$ напоминает разложение комплексного числа на вещественную и мнимую части (ведь самосопряженный оператор имеет вещественный спектр!). Более содержательна, однако, следующая

Теорема 12. *Каждый невырожденный линейный оператор $\mathcal{A}$ в евклидовом пространстве можно представить в виде произведения $\mathcal{A}=\mathcal{U}\mathcal{C}$, где $\mathcal{C}$ — самосопряженный оператор с положительными собственными значениями* (такой оператор называется положительно определенным, или просто положительным), *а $\mathcal{U}$ — унитарный* (а в случае вещественного пространства — *ортогональный*) *оператор* (собственные значения которого, как известно, по модулю равны 1).

[Такое разложение в произведение вида $\mathcal{U}\mathcal{C}$ линейного оператора напоминает тригонометрическую форму комплексного числа: если $\alpha\neq 0$, то $\alpha=r(\cos\varphi+i\sin\varphi)$, где $r>0$, а число $\cos\varphi+i\sin\varphi$ по модулю равно 1.]

Доказательство. Заметим сначала, что если $\mathcal{A}$ — произвольный линейный оператор в евклидовом пространстве, то оператор $\mathcal{B}=\mathcal{A}^*\mathcal{A}$ (так же, как и $\mathcal{A}\mathcal{A}^*$) является самосопряженным, так как

$$\mathcal{B}^*=(\mathcal{A}^*\mathcal{A})^*=\mathcal{A}^*\mathcal{A}^{**}=\mathcal{A}^*\mathcal{A}=\mathcal{B}.$$

Если оператор $\mathcal{A}$ — невырожденный, то при $x\neq 0$ и $\mathcal{A}x\neq 0$, а значит $(\mathcal{A}x, \mathcal{A}x)>0$. Покажем, что в этом случае все собственные значения оператора $\mathcal{B}=\mathcal{A}^*\mathcal{A}$ положительны. Действительно, пусть λ — собственное значение, а x — соответствующий собственный вектор оператора $\mathcal{B}$. Тогда $x\neq 0$ и $\mathcal{B}x=\lambda x$. В этом случае $(\mathcal{B}x, x)=(\mathcal{A}^*\mathcal{A}x, x)=(\mathcal{A}x, \mathcal{A}x)>0$. Но $(\mathcal{B}x, x)=(\lambda x, x)=\lambda(x, x)$. А так как $(x, x)>0$, то и $\lambda>0$.

Докажем теперь само утверждение теоремы. Если оно справедливо, т. е. если оператор $\mathcal{A}$ можно представить в указанном виде $\mathcal{A}=\mathcal{U}\mathcal{C}$, то оператор

$$\mathcal{B}=\mathcal{A}^*\mathcal{A}=(\mathcal{U}\mathcal{C})^*\mathcal{U}\mathcal{C}=\mathcal{C}^*\mathcal{U}^*\mathcal{U}\mathcal{C}=\mathcal{C}^2.$$

Возьмем в качестве базиса пространства R тот (ортонормированный) базис, в котором матрица (самосопряженного) оператора $\mathcal{B}=\mathcal{A}^*\mathcal{A}$ приводится к диагональному виду

$$B=\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}, \tag{8}$$

где по доказанному все $\lambda_i>0$. Обозначим через $\mathcal{C}$ «положительный квадратный корень» из $\mathcal{B}$, т. е. оператор $\mathcal{C}$, матрица которого в том же базисе имеет вид

$$C=\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}.$$

Тогда ясно, что $\mathcal{C}$ — положительно определенный оператор и что $\mathcal{C}^2=\mathcal{B}$. Если теперь положить $\mathcal{A}=\mathcal{U}\mathcal{C}$, то оператор $\mathcal{U}=\mathcal{A}\mathcal{C}^{-1}$, и нам остается только показать, что оператор $\mathcal{U}$ — эрмитов (в вещественном случае — ортогональный). Но это видно из равенства

$$\mathcal{U}^*\mathcal{U}=(\mathcal{A}\mathcal{C}^{-1})^*\mathcal{A}\mathcal{C}^{-1}=(\mathcal{C}^{-1})^*\mathcal{A}^*\mathcal{A}\mathcal{C}^{-1}= \\ =\mathcal{C}^{-1}\mathcal{B}\mathcal{C}^{-1}=\mathcal{C}^{-1}\mathcal{C}^2\mathcal{C}^{-1}=\mathcal{E}$$

— и теорема доказана.

Аналогично можно доказать, что всякий невырожденный линейный оператор $\mathcal{A}$ можно представить и в виде $\mathcal{A}=\mathcal{C}_1\mathcal{U}_1$, где $\mathcal{C}_1$ — положительно определенный, а $\mathcal{U}_1$ — унитарный (ортогональный) операторы. Можно доказать, что указанное в теореме разложение единственно.

В случае вещественного пространства можно сказать, таким образом, что каждое невырожденное линейное преобразование сводится к нескольким симметриям относительно гиперплоскостей, нескольким поворотам около $(n-2)$-мерных «осей» и нескольким растяжениям вдоль попарно ортогональных прямых.

Пример. Пусть оператор $\mathcal{A}$ в базисе (e_1, e_2) имеет невырожденную $(|A|=4\neq 0)$ матрицу

$$A=\begin{bmatrix}-\sqrt{2} & -3\\ 2 & \sqrt{2}\end{bmatrix}.$$

Произведение

$$A^*A=\begin{bmatrix}-\sqrt{2} & 2\\ -3 & \sqrt{2}\end{bmatrix}\begin{bmatrix}-\sqrt{2} & -3\\ 2 & \sqrt{2}\end{bmatrix}=\begin{bmatrix}6 & 5\sqrt{2}\\ 5\sqrt{2} & 11\end{bmatrix}=B$$

— симметрическая матрица, из которой надо «извлечь квадратный корень». Собственные значения B — это $\lambda_1=16$ и $\lambda_2=1$. Соответствующие собственные векторы $e_1'=(1, \sqrt{2})$ и $e_2=(-\sqrt{2}, 1)$. В базисе (e_1', e_2') матрица оператора $\mathcal{B}$ приводится к виду $B_1=\begin{bmatrix}16 & 0\\ 0 & 1\end{bmatrix}$. Значит, «положительный квадратный корень» из нее — это $C_1=\begin{bmatrix}4 & 0\\ 0 & 1\end{bmatrix}$. Матрицей перехода от базиса (e_1, e_2) к базису (e_1', e_2') будет $\Delta=\begin{bmatrix}1 & -\sqrt{2}\\ \sqrt{2} & 1\end{bmatrix}$, а обратной к ней — $\Delta^{-1}=\frac{1}{3}\begin{bmatrix}1 & \sqrt{2}\\ -\sqrt{2} & 1\end{bmatrix}$. Следовательно, в старом базисе (e_1, e_2) матрица оператора $\mathcal{C}$ — это

$$C=\Delta C_1\Delta^{-1}=\begin{bmatrix}1 & -\sqrt{2}\\ \sqrt{2} & 1\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & 1\end{bmatrix}\cdot\frac{1}{3}\begin{bmatrix}1 & \sqrt{2}\\ -\sqrt{2} & 1\end{bmatrix}=\begin{bmatrix}2 & \sqrt{2}\\ \sqrt{2} & 3\end{bmatrix}$$

и тогда

$$U=AC^{-1}=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$$

(при этом $\mathcal{A}=\mathcal{U}\mathcal{C}$ — оператор $\mathcal{A}$ представлен в виде произведения положительно определенного и унитарного операторов).

ГЛАВА VI

БИЛИНЕЙНЫЕ И КВАДРАТИЧНЫЕ ФОРМЫ

Результаты первых пяти параграфов этой главы относятся к вещественному пространству. В последнем, шестом, параграфе они обобщаются на комплексный случай.

§ 1. Билинейный функционал. Билинейная и квадратичная формы

Определение 1. *Заданная в (вещественном) векторном пространстве R функция двух переменных $A(x, y)$*, относящая каждой паре x, y векторов число $A(x, y)$, *называется билинейной функцией,* или *билинейным функционалом, если*

$$A(x+y, z) = A(x, z) + A(y, z),$$
$$A(\alpha x, y) = \alpha A(x, y),$$
$$A(z, x+y) = A(z, x) + A(z, y),$$
$$A(x, \alpha y) = \alpha A(x, y),$$

где x, y, z — произвольные векторы из R и α — любое (вещественное) число.

Таким образом, $A(x, y)$ есть линейный функционал по x при фиксированном y и линейный функционал по y при фиксированном x.

Примером билинейного функционала может служить скалярное произведение (x, y) векторов (вещественного) евклидова пространства.

Найдем *выражение билинейного функционала в координатах.* Пусть в пространстве R задан базис $e_1, e_2, \ldots$ $\ldots, e_n$, и пусть

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n,$$
$$y = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n.$$

Тогда

$$A(x, y) =$$
$$= A(x_1e_1 + x_2e_2 + \ldots + x_ne_n,\ y_1e_1 + y_2e_2 + \ldots + y_ne_n) =$$
$$= \sum_{i,k=1}^{n} x_iy_kA(e_i, e_k) = \sum_{i,k=1}^{n} a_{ik}x_iy_k,$$

где коэффициенты $a_{ik} = A(e_i, e_k)$ зависят от базиса и не зависят от x и y. Таким образом, *в заданном базисе билинейный функционал представляется* ***б и л и н е й н о й ф о р м о й***, т. е. выражением вида $\sum_{i,k=1}^{n} a_{ik}x_iy_k$ *). Матрица $A = [a_{ik}]$ называется **м а т р и ц е й** этой **б и л и н е й н о й ф о р м ы**. В частности, скалярное произведение (x, y) представляется билинейной формой

$$\sum_{i,k=1}^{n} g_{ik}x_iy_k, \quad \text{где} \quad g_{ik} = (e_i, e_k).$$

Билинейную форму $A(x, y)$ можно рассматривать как матричное произведение

$$X'AY,$$

где X — столбец (и значит, транспонированная к X матрица X' — строка) из координат вектора x, Y — столбец из координат вектора y и A — матрица билинейной формы.

Найдем, *как изменяется матрица билинейной формы при переходе к новому базису*. Пусть в базисе $e_1, e_2, \ldots, e_n$

$$A(x, y) = \sum_{i,k=1}^{n} a_{ik}x_iy_k, \quad \text{где} \quad a_{ik} = A(e_i, e_k),$$

и пусть $e'_1, e'_2, \ldots, e'_n$ — новый базис, в котором

$$A(x, y) = \sum_{p,q=1}^{n} b_{pq}x'_py'_q, \quad \text{где} \quad b_{pq} = A(e'_p, e'_q).$$

Положим $A = [a_{ik}]$, $B = [b_{ik}]$ и обозначим через $C = [c_{ik}]$ матрицу перехода от старого базиса к новому;

---

*) Сам билинейный функционал $f(x, y)$ часто тоже называют *билинейной формой*.

тогда

$$b_{pq} = A(e'_p, e'_q) =$$
$$= A(c_{1p}e_1 + c_{2p}e_2 + \ldots + c_{np}e_n,\ c_{1q}e_1 + c_{2q}e_2 + \ldots + c_{nq}e_n) =$$
$$= \sum_{i,k=1}^{n} c_{ip}c_{kq}A(e_i, e_k) = \sum_{i,k=1}^{n} c_{ip}c_{kq}a_{ik} = \sum_{i,k=1}^{n} c_{ip}a_{ik}c_{kq}.$$

Обозначив c_{ip} через d_{pi}, получим

$$b_{pq} = \sum_{i,k=1}^{n} d_{pi}a_{ik}c_{kq}.$$

Матрица $[d_{pi}] = C'$ является транспонированной к матрице $C = [c_{ip}]$. Далее, так как $\sum_{k=1}^{n} a_{ik}c_{kq}$ есть элемент, стоящий в i-строке и q-м столбце матрицы AC, то

$$\sum_{i,k=1}^{n} d_{pi}a_{ik}c_{kq} = \sum_{i=1}^{n} d_{pi}\left(\sum_{k=1}^{n} a_{ik}c_{kq}\right)$$

— это элемент, стоящий в p-й строке и q-м столбце матрицы $C'AC$. Таким образом,

$$B = C'AC. \qquad (1)$$

Заметим, что так как матрица перехода C (а значит, и C') является невырожденной (т. е. имеет ранг n), то ранг матрицы B равен рангу матрицы A (см. § 6 главы III). Таким образом, *ранг матрицы билинейной формы не зависит от выбора базиса* и может быть назван поэтому *рангом* самой *билинейной формы* (*билинейного функционала*).

Приведем еще другой вывод формулы (1). В обозначениях § 3 главы III имеем $X_{\text{ст}} = CX_{\text{нов}}$ и $Y_{\text{ст}} = CY_{\text{нов}}$. Далее, из равенства $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$ (§ 2 главы V) для матриц вытекает равенство $(AB)' = B'A'$ — оно справедливо, впрочем, не только для квадратных матриц, и, значит, $X'_{\text{ст}} = X'_{\text{нов}}C'$. Следовательно,

$$A(x, y) = X'_{\text{ст}}A_{\text{ст}}Y_{\text{ст}} = X'_{\text{нов}}C'A_{\text{ст}}CY_{\text{нов}}.$$

Но $A(x, y) = X'_{\text{нов}}A_{\text{нов}}Y_{\text{нов}}$ и, значит, $A_{\text{нов}} = C'A_{\text{ст}}C$ (легко видеть, что из равенства $X'B_1Y = X'B_2Y$, справедливого для любой строки X' и любого столбца Y, вытекает, что $B_1 = B_2$).

Билинейный функционал $A(x, y)$ называется **симметрическим**, если для всех x и y из R

$$A(x, y) = A(y, x).$$

В этом случае $a_{ik} = A(e_i, e_k) = A(e_k, e_i) = a_{ki}$, т. е. матрица $[a_{ik}]$ соответствующей билинейной формы (в любом базисе) будет **симметрической**; обратно, если матрица билинейной формы (в каком-то базисе) — симметрическая, то и соответствующий билинейный функционал будет симметрическим (почему?). Примером симметрического билинейного функционала может служить скалярное произведение векторов пространства со скалярным произведением. Последний пример является вполне общим, так как и, обратно, каждый симметрический билинейный функционал $A(x, y)$ удовлетворяет, очевидно, условиям 1—3 из § 1 главы IV и, значит, может быть принят за скалярное произведение.

Если в *симметрической* билинейной форме $A(x, y)$ положить $y = x$, то получится *квадратичная форма* $A(x, x)$. При этом матрица A квадратичной формы $A(x, x)$ — это, по определению, **симметрическая** матрица A отвечающей $A(x, x)$ билинейной формы $A(x, y)$. Заметим, что по квадратичной форме породившая ее симметрическая билинейная форма определяется однозначно. Действительно, пусть $A(y, x) = A(x, y)$ при всех x и y. Тогда

$$A(x+y, x+y) = A(x, x) + 2A(x, y) + A(y, y),$$

откуда

$$A(x, y) = \frac{1}{2}[A(x+y, x+y) - A(x, x) - A(y, y)].$$

Билинейная функция $A(x, y)$ называется *кососимметрической, если*

$$A(x, y) = -A(y, x)$$

при всех $x, y \in R$. В заданном базисе кососимметрическая функция представляется *кососимметрической формой*

$$A(x, y) = \sum_{i,k=1}^{n} a_{ik} x_i y_k,$$

где $a_{ik} = A(e_i, e_k) = -A(e_k, e_i) = -a_{ki}$ при всех i, k и, в частности, $a_{ii}=0$ при всех i. Так, в трехмерном пространстве кососимметрическая форма имеет вид

$$a_{12}(x_1y_2-x_2y_1)+a_{13}(x_1y_3-x_3y_1)+a_{23}(x_2y_3-x_3y_2).$$

Пусть $A(x, y)$ — произвольный билинейный функционал. Тогда $B(x, y)=\frac{1}{2}[A(x, y)+A(y, x)]$ является, очевидно, симметрическим, а $C(x, y)=\frac{1}{2}[A(x, y)-A(y, x)]$ — кососимметрическим функционалами. Но

$$A(x, y)=B(x, y)+C(x, y);$$

следовательно, *каждый билинейный функционал может быть представлен в виде суммы симметрического и кососимметрического функционалов.*

§ 2. Приведение квадратичной формы к сумме квадратов

Теорема 1. *Пусть $A(x, x)$ — произвольная квадратичная форма в n-мерном векторном пространстве. Тогда найдется такой базис, в котором эта форма приводится к сумме квадратов* (т. е. в котором все коэффициенты при попарных произведениях координат вектора x равны нулю).

Доказательство проведем индукцией, по числу входящих в форму переменных. Если в $A(x, x)$ входит лишь одна координата, скажем,

$$A(x, x)=a_{11}x_1^2,$$

то наше утверждение очевидно. Предположим, что оно справедливо для всех квадратичных форм, зависящих от $m-1$ координат, и рассмотрим квадратичную форму, зависящую от m переменных:

$$A(x, x)=a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2+\ldots+a_{mm}x_m^2.$$

Если здесь есть хотя бы один квадрат с отличным от нуля коэффициентом, например, если $a_{mm}\neq 0$, то соберем все члены, содержащие x_m:

$$2a_{1m}x_1x_m+2a_{2m}x_2x_m+\ldots+2a_{m-1,m}x_{m-1}x_m+a_{mm}x_m^2$$

и «выделим полный квадрат»:

$$2a_{1m}x_1x_m + 2a_{2m}x_2x_m + \ldots + 2a_{m-1,m}x_{m-1}x_m + a_{mm}x_m^2 =$$
$$= \frac{1}{a_{mm}}(a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{m-1,m}x_{m-1} + a_{mm}x_m)^2 -$$
$$- \frac{1}{a_{mm}}(a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{m-1,m}x_{m-1})^2.$$

Тогда

$$A(x, x) = \frac{1}{a_{mm}}(a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{mm}x_m)^2 + B(x, x),$$

где квадратичная форма $B(x, x)$ зависит уже только от $m-1$ координат: $x_1, x_2, \ldots, x_{m-1}$. Положим

$$y_1 = x_1,\ y_2 = x_2,\ \ldots,\ y_{m-1} = x_{m-1},$$
$$y_m = a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{mm}x_m,$$
$$y_{m+1} = x_{m+1},\ \ldots,\ y_n = x_n.$$

Так как определитель

$$D = \begin{vmatrix} 1 & 0 & 0 & \ldots 0 & \ldots 0 \\ 0 & 1 & 0 & \ldots 0 & \ldots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{1m} & a_{2m} & a_{3m} & \ldots a_{mm} & \ldots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots 0 & \ldots 1 \end{vmatrix} = a_{mm} \neq 0,$$

то этот переход к новым координатам вызывается переходом к некоторому новому базису — с матрицей перехода, обратной матрице определителя D (см. § 6 главы II).

По предположению индукции, форму $B(x, x)$, зависящую от $m-1$ переменных $x_1, x_2, \ldots, x_{m-1}$, посредством перехода к новому базису можно привести к сумме квадратов. При этом окончательно приведется к сумме квадратов и форма $A(x, x)$.

Мы предполагали, что хотя бы один из квадратов входит в форму $A(x, x)$ с ненулевым коэффициентом. Если это не так, т. е. если все $a_{ii} = 0$, то допустим, что, например, $a_{12} \neq 0$, и положим

$$x_1 = y_1 + y_2,\ x_2 = y_1 - y_2,\ x_3 = y_3,\ \ldots,\ x_n = y_n$$

— это соответствует переходу к новому базису

$$e_1' = e_1 + e_2, \quad e_2' = e_1 - e_2, \quad e_3' = e_3, \ldots, e_n' = e_n$$

с матрицей перехода

$$\begin{bmatrix} 1 & 1 & 0 & \ldots & 0 \\ 1 & -1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

(определитель этой матрицы равен $-2 \neq 0$). При этом произведение $x_1 x_2$ обратится в $y_1^2 - y_2^2$, и мы придем к первому случаю.

Мы доказали, что если в n-мерном векторном пространстве R задана произвольная квадратичная форма, то в R можно найти такой базис, в котором эта форма приведется к сумме квадратов:

$$A(x, x) = a_1 x_1'^2 + a_2 x_2'^2 + \ldots + a_n x_n'^2, \tag{2}$$

где $x_1', x_2', \ldots, x_n'$ — координаты вектора x в новом базисе. Коэффициенты a_i могут быть и положительными и отрицательными; некоторые из них могут быть равными нулю. Сделав еще одну подстановку $\sqrt{|a_i|}\, x_i' = z_i$, если $a_i \neq 0$ и $x_j' = z_j$, если $a_j = 0$, приведем квадратичную форму $A(x, x)$ к виду

$$A(x, x) = \pm z_1^2 \pm z_2^2 \pm \ldots \pm z_m^2,$$

где коэффициент перед каждым неизвестным $z_1, z_2, \ldots$ $\ldots, z_m$ равен $+1$, или -1, или, после изменения нумерации базисных векторов,— к виду

$$A(x, x) = z_1^2 + z_2^2 + \ldots + z_p^2 - z_{p+1}^2 - \ldots - z_{p+q}^2.$$

Пример. *Квадратичную форму* $x_1^2 + 2x_1x_2 + 3x_2^2 + 4x_1x_3 + 6x_2x_3 + 3x_3^2$ *привести к сумме квадратов.*

Решение.

$$A(x, x) = (x_1 + x_2 + 2x_3)^2 + 2x_2^2 + 2x_2x_3 - x_3^2 =$$
$$= (x_1 + x_2 + 2x_3)^2 - (x_2 - x_3)^2 + 3x_2^2 = z_1^2 + z_2^2 - z_3^2,$$

где $z_1 = x_1 + x_2 + 2x_3, \quad z_2 = \sqrt{3}x_2, \; z_3 = x_2 - x_3.$

§ 3. Закон инерции квадратичных форм

Приводя квадратичную форму $A(x, x)$ к сумме квадратов разными способами, мы будем получать в формуле (2), вообще говоря, разные коэффициенты. Однако имеет место следующее важное обстоятельство:

Теорема 2 (закон инерции квадратичных форм). *Если квадратичная форма приводится к сумме квадратов в двух разных базисах, то число членов с положительными коэффициентами, так же как и число членов с отрицательными коэффициентами, в обоих случаях одно и то же.*

Доказательство (от противного). Предположим, что в базисе $e_1, e_2, \ldots, e_n$ квадратичная форма $A(x, x)$ имеет вид

$$A(x, x) = x_1^2 + x_2^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2, \tag{3}$$

где x_i — координаты вектора x в этом базисе; и пусть в другом базисе $e_1', e_2', \ldots, e_n'$:

$$A(x, x) = x_1'^2 + x_2'^2 + \ldots + x_k'^2 - x_{k+1}'^2 - \ldots - x_{k+m}'^2, \tag{4}$$

где x_i' — координаты вектора x в новом базисе. Пусть, например, $p > k$. Рассмотрим в пространстве R подпространство R_1, порожденное векторами $e_1, e_2, \ldots, e_p$, и подпространство R_2, порожденное векторами $e_{k+1}', e_{k+2}', \ldots$ $\ldots, e_n'$. Так как сумма их размерностей, равная $p + (n-k)$, больше n, то их пересечение имеет ненулевую размерность (теорема 5 из § 9 главы II), т. е. существует вектор $x \neq 0$, принадлежащий $R_1 \cap R_2$. Этот вектор можно представить как в виде

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_p e_p,$$

так и в виде

$$x = \beta_{k+1} e_{k+1}' + \beta_{k+2} e_{k+2}' + \ldots + \beta_n e_n'.$$

Для вектора x по формуле (3)

$$A(x, x) = \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_p^2 > 0,$$

так как хотя бы одно из $\alpha_i \neq 0$; в то же время по формуле (4)

$$A(x, x) = -\beta_{k+1}^2 - \beta_{k+2}^2 - \ldots - \beta_{k+m}^2 \leqslant 0$$

(последнее неравенство — нестрогое, потому что возможно, что $k+m<n$).

Мы пришли к противоречию, откуда и следует, что $p \leqslant k$. Аналогично получаем и неравенство $p \geqslant k$. Следовательно, $p=k$. Так же доказывается, что $q=m$.

Легко видеть, что *сумма* $p+q$ *равна рангу* r *квадратичной формы* $A(x, x)$. *Разность* $p-q$ *называется* *сигнатурой формы* $A(x, x)$.

§ 4. Определенные формы

Определение 2. *Квадратичная форма* $A(x, x)$ *называется положительно (отрицательно) определенной, если* $A(x, x)>0$ $(A(x, x)<0)$ *при всех* $x \neq \theta$, *и положительно (отрицательно) полуопределенной, если* $A(x, x) \geqslant 0$ $(A(x, x) \leqslant 0)$ *при всех* x.

Так, если $A(x, y)=(x, y)$ — скалярное произведение в евклидовом пространстве, то соответствующая квадратичная форма $A(x, x)=(x, x)$ (скалярный квадрат вектора x) является положительно определенной.

Ясно, что положительно определенная квадратичная форма приводится к сумме квадратов с положительными коэффициентами, а положительно полуопределенная форма — с неотрицательными коэффициентами (некоторые из которых могут равняться нулю). Важным условием положительной определенности формы является следующая

Теорема 3 (критерий Сильвестра). *Для того чтобы квадратичная форма* $A(x, x)$ *была положительно определенной, необходимо и достаточно, чтобы были положительны все «угловые миноры» матрицы* $A=[a_{ik}]$, *т. е. чтобы имели место неравенства*

$$\Delta_1=a_{11}>0,$$

$$\Delta_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}>0,\ \Delta_3=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}>0,\ \ldots,\ \Delta_n=|A|>0.$$

Доказательство проведем индукцией по числу входящих в форму переменных.

Для квадратичной формы, зависящей от одной переменной, $A(x, x)=a_{11}x_1^2$, и наше утверждение очевидно. Предположим, что оно справедливо для всех квадратичных форм, зависящих от $n-1$ переменных, и рассмотрим квадратичную форму $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_i x_k$, зависящую от переменных $x_1, x_2, \ldots, x_n$.

1. Доказательство необходимости. Если представить положительно определенную форму $A(x, x)$ в виде

$$A(x, x) = \sum_{i,k=1}^{n-1} a_{ik}x_i x_k + 2\sum_{i=1}^{n-1} a_{in}x_i x_n + a_{nn}x_n^2,$$

то квадратичная форма $B(x', x') = \sum_{i,k=1}^{n-1} a_{ik}x_i x_k$, зависящая от $n-1$ переменных $x_1, x_2, \ldots, x_{n-1}$ (и рассматриваемая, конечно, в $(n-1)$ мерном пространстве), будет положительно определенной, так как если $B(x', x') \leqslant 0$ при $x' = (x_1, x_2, \ldots, x_{n-1})$, то при $x = (x_1, x_2, \ldots, x_{n-1}, 0)$ мы имели бы $A(x, x) \leqslant 0$. По предположению индукции, все угловые миноры матрицы квадратичной формы $B(x', x')$ положительны, т. е.

$$\Delta_1 = a_{11} > 0, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0, \ldots$$

$$\ldots, \Delta_{n-1} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1,n-1} \\ a_{12} & a_{22} & \ldots & a_{2,n-1} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1,n-1} & a_{2,n-1} & \ldots & a_{n-1,n-1} \end{vmatrix} > 0.$$

Остается доказать, что и $\Delta_n = |A| > 0$.

Мы знаем, что положительно определенная квадратичная форма $A(x, x)$ в некотором базисе $e'_1, e'_2, \ldots, e'_n$ приводится к сумме квадратов

$$A(x, x) = x'^2_1 + x'^2_2 + \ldots + x'^2_n.$$

В этом новом базисе определитель ее матрицы равен 1 и, значит, он больше нуля. Однако при переходе к новому базису матрица билинейной формы преобразуется по формуле (стр. 189)

$$B = C'AC,$$

где A — ее матрица в старом базисе, B — в новом и C — матрица перехода от старого базиса к новому. Следовательно,

$$|B| = |C'|\,|A|\,|C| = |A|\,|C|^2. \tag{5}$$

Но так как $|C| \neq 0$ и $|B| > 0$, то и

$$|A| = \Delta_n > 0.$$

2. Доказательство достаточности. Предположим, что все угловые миноры матрицы квадратичной формы $A(x, x)$

положительны:

$$\Delta_1 > 0,\ \Delta_2 > 0,\ \ldots,\ \Delta_{n-1} > 0,\ \Delta_n = |A| > 0,$$

и докажем, что квадратичная форма $A(x, x)$ положительно определенная. Из предположения индукции вытекает, прежде всего, положительная определенность квадратичной формы $B(x', x') = \sum_{i,k=1}^{n-1} a_{ik}x_i x_k$ от $n-1$ переменных (в $(n-1)$-мерном пространстве). Следовательно, $B(x', x')$ в некотором новом базисе приводится к сумме квадратов:

$$B(x', x') = x_1'^2 + x_2'^2 + \ldots + x_{n-1}'^2.$$

Сделав соответствующую замену переменных $x_1, x_2, \ldots, x_{n-1}$ и положив, кроме того, $x_n = x_n'$, мы получим

$$A(x, x) = x_1'^2 + x_2'^2 + \ldots + x_{n-1}'^2 + \\ + 2\left(b_{1n}x_1'x_n' + b_{2n}x_2'x_n' + \ldots + b_{n-1,n}x_{n-1}'x_n'\right) + a_{nn}x_n'^2,$$

где b_{in} — какие-то новые коэффициенты. Далее имеем

$$A(x, x) = \left(x_1' + b_{1n}x_n'\right)^2 + \\ + \left(x_2' + b_{2n}x_n'\right)^2 + \ldots + \left(x_{n-1}' + b_{n-1,n}x_n'\right)^2 + bx_n'^2,$$

где, очевидно, $b = a_{nn} - b_{1n}^2 - b_{2n}^2 - \ldots - b_{n-1,n}^2$, и, полагая

$$x_i' + b_{in}x_n' = y_i, \quad i = 1, 2, \ldots, n-1, \quad x_n' = y_n$$

(что соответствует переходу к новому базису, с матрицей, определитель которой равен единице), получим

$$A(x, x) = y_1^2 + y_2^2 + \ldots + y_{n-1}^2 + by_n^2.$$

Определитель матрицы этой квадратичной формы равен b, а так как знак его, как видно из формулы (5), совпадает со знаком Δ_n, то $b > 0$, и значит, квадратичная форма $A(x, x)$ — положительно определенная. Теорема доказана.

Теперь нетрудно найти и условия *отрицательной определенности квадратичной формы* $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_i x_k$.

Для того чтобы квадратичная форма $A(x, x)$ была отрицательно определенной, необходимо и достаточно, чтобы квадратичная форма

$$-A(x, x) = \sum_{i,k=1}^{n} (-a_{ik})\, x_i x_k$$

была положительно определенной, а значит — чтобы все угловые миноры матрицы

$$\begin{bmatrix} -a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{12} & -a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{1n} & -a_{2n} & \dots & -a_{nn} \end{bmatrix},$$

т. е.

$$-a_{11}, \quad \begin{vmatrix} -a_{11} & -a_{12} \\ -a_{12} & -a_{22} \end{vmatrix}, \quad \dots, \quad \begin{vmatrix} -a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{12} & -a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{1n} & -a_{2n} & \dots & -a_{nn} \end{vmatrix},$$

были положительны. Но это означает, что

$$\Delta_1 = a_{11} < 0, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} < 0, \dots,$$

т. е. что *знаки угловых миноров матрицы A чередуются, начиная со знака минус.*

Пример. При исследовании на экстремум функции

$$F(x, y, z) = 2x^2 + y^2 + 11z^2 - 2xy + 4xz - 6yz - 2y + 8z,$$

находим, что ее частные производные обращаются в нули при

$$x = 1, \ y = 2, \ z = 0.$$

Второй дифференциал функции $F(x, y, z)$ имеет вид

$$d^2F = 2(2dx^2 - 2dx\,dy + dy^2 + 4dx\,dz - 6dy\,dz + 11dz^2).$$

В скобках — квадратичная форма относительно дифференциалов независимых переменных dx, dy, dz. Угловые миноры ее матрицы

$$\Delta_1 = 2, \quad \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 1, \quad \Delta_3 = \begin{vmatrix} 2 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 11 \end{vmatrix} = 1$$

положительны Следовательно, эта квадратичная форма положительно определенная, и заданная функция имеет в точке (1, 2, 0) минимум.

§ 5. Билинейные и квадратичные формы в евклидовом пространстве

Лемма. *Пусть R (вещественное) евклидово пространство и $C = [c_{ik}]$ — матрица перехода от одного ортонормированного базиса $e_1, e_2, \ldots, e_n$ к другому, тоже ортонормированному базису $e'_1, e'_2, \ldots, e'_n$. Тогда C — ортогональная матрица.*

Доказательство. По условию,

$$e'_i = c_{1i}e_1 + c_{2i}e_2 + \ldots + c_{ni}e_n, \qquad i = 1, 2, \ldots, n.$$

Рассмотрим линейный оператор $\mathscr{C}$ с матрицей C в базисе $e_1, e_2, \ldots, e_n$. Мы имеем

$$\mathscr{C}e_i = c_{1i}e_1 + c_{2i}e_2 + \ldots + c_{ni}e_n = e'_i, \qquad i = 1, 2, \ldots, n.$$

Но оператор $\mathscr{C}$, переводящий хотя бы один ортонормированный базис в ортонормированный же,— ортогональный (см. § 4 главы V). Следовательно, C — ортогональная матрица.

Пусть теперь в евклидовом пространстве R выбран ортонормированный базис $e_1, e_2, \ldots, e_n$, и пусть дан билинейный функционал $A(x, y)$, который в этом базисе представляется билинейной формой

$$A(x, y) = \sum_{i,k=1}^{n} a_{ik}x_iy_k,$$

где $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, $y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$. Рассмотрим линейный оператор $\mathscr{A}$ с той же матрицей A в том же базисе $e_1, e_2, \ldots, e_n$. При переходе к новому базису $e'_1, e'_2, \ldots, e'_n$ с матрицей перехода C матрица A билинейной формы перейдет в

$$C'AC,$$

а матрица линейного оператора $\mathscr{A}$ — в

$$C^{-1}AC,$$

т. е., вообще говоря, эти матрицы преобразуются не одинаково. Однако если новый базис $e'_1, e'_2, \ldots, e'_n$ — тоже ортонормированный, то матрица перехода C ортогональна и $C' = C^{-1}$. В этом случае матрица билинейной формы $A(x, y)$ и матрица линейного оператора $\mathscr{A}$ преобразуются одинаково. Таким образом, в евклидовом пространстве каждому билинейному функционалу соответствует вполне определенный линейный оператор (имеющий ту же матрицу в любом ортонормированном базисе).

Если $A(x, y)$ — *симметрический* билинейный функционал, то соответствующий линейный оператор $\mathscr{A}$ будет *самосопряженным*. Но матрица самосопряженного оператора в некотором ортонормированном же базисе приводится к диагональному виду с собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_n$ на главной диагонали. При этом, если

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{j=1}^{n} y_j e_j,$$

то билинейная форма $A(x, y) = \sum_{i=1}^{n} \lambda_i x_i y_i$, а соответствующая квадратичная форма $A(x, x)$ приводится к сумме квадратов:

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2.$$

Пример. *Квадратичную форму*

$$A(x, x) = 66x^2 - 24xy + 59y^2$$

в евклидовом пространстве R^2 переходом к новому ортонормированному базису привести к сумме квадратов.

Решение. Характеристический многочлен матрицы A этой формы

$$\varphi(\lambda) = \begin{vmatrix} 66-\lambda & -12 \\ -12 & 59-\lambda \end{vmatrix} = \lambda^2 - 125\lambda + 3750.$$

Его корни $\lambda_1 = 75$, $\lambda_2 = 50$.

В новом базисе (состоящем из собственных векторов оператора A, соответствующих собственным значениям λ_1 и λ_2)

$$A(x, x) = 75x_1'^2 + 50y_1'^2.$$

Легко видеть, что квадратичная форма $A(x, x)$ тождественно равна скалярному произведению $(\mathscr{A}x, x)$. Выше мы назвали положительно определенным самосопряженный оператор с положительными собственными значениями. Ясно, что *для того чтобы самосопряженный оператор $\mathscr{A}$ был положительно определенным, необходимо и достаточно, чтобы была положительно* *определенной соответствующая квадратичная форма* $A(x, x)$, т. е. чтобы при всех $x \neq 0$ выполнялось неравенство $(\mathcal{A}x, x) = A(x, x) > 0$.

Методами математического анализа можно показать, что наименьшее из собственных значений $\lambda_1, \lambda_2, \ldots$ $\ldots, \lambda_n$ самосопряженного оператора $\mathcal{A}$ равно минимуму, а наибольшее — максимуму квадратичной формы $A(x, x) = (\mathcal{A}x, x)$ на «единичной сфере». $x_1^2 + x_2^2 + \ldots$ $\ldots + x_n^2 = 1$. (Для того чтобы доказать это, надо найти экстремумы функции $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2$ при условии, что $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$.)

§ 6. Билинейный функционал в комплексном векторном пространстве

Билинейный функционал в комплексном пространстве иногда задают определением 1; однако чаще используется следующее

Определение 1′. *Функция $A(x, y)$ двух переменных, заданная в комплексном векторном пространстве R, называется билинейным функционалом, если для всех x, y, z из R и любого (комплексного) числа α*

$$A(x+y, z) = A(x, z) + A(y, z),$$
$$A(\alpha x, y) = \alpha A(x, y),$$
$$A(z, x+y) = A(z, x) + A(z, y),$$
$$A(x, \alpha y) = \bar{\alpha} A(x, y).$$

(В дальнейшем мы будем относить определение 1 только к вещественному пространству, понимая под билинейным функционалом в комплексном пространстве функцию, удовлетворяющую условиям определения 1′ *).)

*) В научной и учебной литературе удовлетворяющую перечисленным условиям функцию $A(x, y)$ двух векторов комплексного линейного пространства чаще называют *билинейной эрмитовой*, или *полуторалинейной*, и говорят, что она является *линейной* по первому переменному x и *антилинейной* по второму переменному y (или является линейной *1-го рода* по переменному x и *линейной* 2-го рода по переменному y). Мы, однако, и здесь сохраним термин «билинейная форма» (или «билинейный функционал»).

Легко видеть, что в комплексном векторном пространстве билинейный функционал в координатах представляется билинейной формой

$$A(x, y) = \sum_{i,k=1}^{n} a_{ik} x_i \bar{y}_k, \quad \text{где } a_{ik} = A(e_i, e_k).$$

В частности, скалярное произведение (x, y) представляется билинейной формой $\sum_{i,k=1}^{n} g_{ik} x_i \bar{y}_k$, где $g_{ik} = (e_i, e_k)$; в ортонормированном базисе $(x, y) = \sum_{i=1}^{n} x_i \bar{y}_i$.

При переходе к новому базису с матрицей перехода $C = [c_{ik}]$ билинейная форма $A(x, y) = \sum_{i,k=1}^{n} a_{ik} x_i y_k$ преобразуется в $\sum_{i,k=1}^{n} b_{pq} x'_p \bar{y}'_q$, где $b_{pq} = A(e'_p, e'_q)$. Если C' — матрица, транспонированная к C, а $\bar{C}$ — матрица, комплексно сопряженная к матрице C (все ее элементы являются комплексно-сопряженными к соответствующим элементам матрицы C), и $B = [b_{pq}]$, то

$$B = C'A\bar{C}.$$

Это можно показать, например, так: в обозначениях § 3 главы III $X_{\text{ст}} = CX_{\text{нов}}$, $Y_{\text{ст}} = CY_{\text{нов}}$ и $A(x, y) = X'_{\text{ст}} A_{\text{ст}} \bar{Y}_{\text{ст}} = X'_{\text{нов}} C' A_{\text{ст}} \bar{C} \bar{Y}_{\text{нов}}$. Но $A(x, y) = X'_{\text{нов}} A_{\text{нов}} \bar{Y}_{\text{нов}}$, и, значит,

$$A_{\text{нов}} = C'A_{\text{ст}}\bar{C}.$$

Билинейный функционал, а также соответствующая ему билинейная форма $A(x, y)$ называются *эрмитовыми, если* $A(x, y) = \overline{A(y, x)}$ при всех $x, y \in R$. В этом случае $a_{ik} = A(e_i, e_k) = \overline{A(e_k, e_i)} = \bar{a}_{ki}$ при всех $i, k = 1, 2, \ldots, n$, т. е. $A = A^*$. Очевидно и обратное: если матрица A билинейной формы равна A^*, т. е. если при всех i, k $a_{ik} = \bar{a}_{ki}$, то соответствующий билинейный функционал — эрмитов. Таким образом, *для того чтобы билинейный функционал был эрмитовым, необходимо и достаточно, чтобы матрица соответствующей билинейной формы* (в любом базисе) *была эрмитовой* (ср. стр. 172).

Пусть $A(x, y)$ — эрмитов билинейный функционал. Положив $y = x$, мы получим *эрмитов квадратичный* *функционал* (форму) $A(x, x)$. В этом случае $A(x, x) = \overline{A(x, x)}$ при всех $x \in R$, и значит, *эрмитова квадратичная форма принимает только вещественные значения.*

В комплексном векторном пространстве *каждая эрмитова квадратичная форма* $A(x, x)$ *в некотором базисе приводится к виду*

$$A(x, x) = \xi_1|x_1|^2 + \xi_2|x_2|^2 + \ldots + \xi_n|x_n|^2, \tag{6}$$

где все ξ_i вещественны (это нетрудно доказать, видоизменив соответствующим образом доказательство теоремы 1, см. стр. 191), *причем если эрмитова квадратичная форма в двух разных базисах приведена к виду* (6), *то число положительных и число отрицательных квадратов в обоих случаях одно и то же.* Наконец, определение положительно определенной эрмитовой формы и критерий Сильвестра без труда переносятся на комплексный случай.

Пусть теперь R — (комплексное) евклидово пространство. Аналогично вещественному случаю (см. лемму на стр. 199), можно показать, что *матрица перехода C от одного ортонормированного базиса к другому, тоже ортонормированному, унитарна.*

Пусть $e_1, e_2, \ldots, e_n$ — ортонормированный базис в R, и пусть $A(x, y)$ — эрмитов билинейный функционал, который в этом базисе представляется билинейной формой

$$A(x, y) = \sum_{i,k=1}^{n} a_{ik} x_i \bar{y}_k,$$

где $a_{ik} = \bar{a}_{ki}$ при всех i, k.

Рассмотрим линейный оператор $\mathscr{A}'$ с матрицей A', транспонированной по отношению к матрице A билинейной формы $A(x, y)$. Так как эта форма эрмитова, то $A^* = A$ и оператор $\mathscr{A}'$ — эрмитов, а следовательно, его матрица, в некотором, тоже ортонормированном, базисе $e'_1, e'_2, \ldots, e'_n$ приводится к диагональному виду

$$B = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

причем все λ_i — вещественны (§ 3 главы V). Если C — матрица перехода к новому базису, то $B = C^{-1}A'C$ (§ 4 главы III). А так как матрица перехода C унитарна, то $C^{-1} = \bar{C}'$, и значит,

$$B = \bar{C}' A C.$$

Далее, так как матрица B диагональная, то она совпадает со своей транспонированной: $B' = B$. Но

$$B' = (\bar{C}' A' C)' = C' A \bar{C}.$$

С другой стороны, при переходе к базису $e'_1, e'_2, \ldots, e'_n$ матрица $\bar{A}$ билинейной формы $A(x, y)$ тоже преобразуется в $C'A\bar{C}$ (см. стр. 202), и значит, в новом базисе она совпадает с матрицей $B' = B$. Следовательно, в базисе $e'_1, e'_2, \ldots, e'_n$ билинейная форма $\overline{A(x, y)}$ имеет вид

$$A(x, y) = \lambda_1 x_1 \bar{y}_1 + \lambda_2 x_2 \bar{y}_2 + \ldots + \lambda_n x_n \bar{y}_n,$$

где λ_i — собственные значения линейного оператора $\mathscr{A}'$ (а значит, и собственные значения оператора $\mathscr{A}$); соответствующая квадратичная форма $A(x, x)$ приведется при этом к сумме квадратов:

$$A(x, x) = \lambda_1 |x_1|^2 + \lambda_2 |x_2|^2 + \ldots + \lambda_n |x_n|^2.$$

ГЛАВА VII

ИССЛЕДОВАНИЕ КРИВЫХ И ПОВЕРХНОСТЕЙ ВТОРОГО ПОРЯДКА

В трех первых параграфах этой главы рассматривается вещественное двумерное пространство (плоскость) с обычной, евклидовой метрикой, а в последнем параграфе — трехмерное вещественное евклидово пространство.

§ 1. Приведение общего уравнения кривой второго порядка к каноническому виду

Установим на плоскости прямоугольную декартову систему координат и рассмотрим общее *уравнение второй степени*

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_1x + 2a_2y + a = 0. \tag{1}$$

Множество точек, координаты которых удовлетворяют уравнению (1), *называется* ***линией*** *(или* ***кривой****)* ***второго порядка.*** Как известно, при некоторых частных значениях коэффициентов уравнение (1) будет уравнением *эллипса* $\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\right)$, *гиперболы* $\left(\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\right)$ или *параболы* ($y^2 = 2px$). Мы докажем, что *уравнение* (1) *всегда является уравнением* одной из этих кривых: *эллипса*, *гиперболы* или *параболы* (не считая случаев вырождения — *пары прямых*, если левая часть уравнения распадается в произведение двух линейных множителей, *точки* или «*пустого множества*», вовсе не содержащего точек).

Обозначим через e_1 и e_2 единичные векторы, направленные по осям выбранной (прямоугольной) системы координат. Группу старших членов

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 \tag{2}$$

уравнения (1) можно рассматривать как квадратичную форму от координат x, y вектора (x, y). Как было показано в § 5 главы VI, эта квадратичная форма в некотором (ортонормированном же) базисе e_1', e_2' приводится к сумме квадратов

$$\lambda_1 x'^2 + \lambda_2 y'^2, \tag{3}$$

где λ_1 и λ_2 — собственные значения матрицы

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \tag{4}$$

а e_1' и e_2' — соответствующие им собственные векторы.

Пусть вектор e_1' получается из вектора e_1 поворотом на угол φ против часовой стрелки. Так как вектор e_2 ортогонален e_1, а e_2' ортогонален e_1', то вектор e_2' получается из вектора e_2 либо поворотом на угол φ, либо поворотом на угол φ и симметрией относительно начала координат. Во втором случае заменим его на вектор $e_2'' = -e_2'$ который тоже будет собственным вектором матрицы (4) с тем же собственным значением λ_2: если $\mathcal{A}e_2' = \lambda_2 e_2'$, то

$$\mathcal{A}e_2'' = \mathcal{A}(-e_2') = -\mathcal{A}e_2' = -\lambda_2 e_2' = \lambda_2 e_2''.$$

Таким образом, можно считать, что новый базис e_1' e_2' получается из старого поворотом на некоторый угол φ против часовой стрелки, т. е. что

$$e_1' = \cos\varphi \cdot e_1 + \sin\varphi \cdot e_2,$$
$$e_2' = -\sin\varphi \cdot e_1 + \cos\varphi \cdot e_2.$$

Но в таком случае старые координаты x, y (вектора, а значит, и соответствующей точки) и новые координаты x', y' связаны соотношениями

$$\begin{aligned} x &= \cos\varphi \cdot x' - \sin\varphi \cdot y', \\ y &= \sin\varphi \cdot x' + \cos\varphi \cdot y'. \end{aligned} \tag{5}$$

Подставив значения (5) в уравнение (1), мы приведем это уравнение к виду

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2b_1 x' + 2b_2 y' + b = 0, \tag{6}$$

где b_1, b_2, b — некоторые новые коэффициенты. Эта операция называется *отнесением линии к главным осям* —

из дальнейшего будет видно, что если линия (1) представляет собой эллипс или гиперболу, то новые оси координат параллельны главным осям кривой.

Коэффициенты λ_1, λ_2 — это собственные значения матрицы (4); их можно найти из уравнения

$$\varphi(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{12} & a_{22}-\lambda \end{vmatrix} = 0. \tag{7}$$

Они вещественны, так как матрица (4) симметрическая (теорема 4 главы V). Произведение $\lambda_1\lambda_2$ собственных значений равно свободному члену $\varphi(0)$ квадратного уравнения (7), т. е. равно определителю

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

Рассмотрим теперь отдельно два случая: $\delta \neq 0$ и $\delta = 0$.

I. $\delta = \lambda_1\lambda_2 \neq 0$. Преобразуем уравнение (6) следующим образом:

$$\lambda_1\left(x' + \frac{b_1}{\lambda_1}\right)^2 + \lambda_2\left(y' + \frac{b_2}{\lambda_2}\right)^2 + c = 0,$$

где $c = b - \frac{b_1^2}{\lambda_1} - \frac{b_2^2}{\lambda_2}$. Сделаем подстановку

$$x'' = x' + \frac{b_1}{\lambda_1}, \quad y'' = y' + \frac{b_2}{\lambda_2}.$$

Эта подстановка отвечает переносу начала координат в точку $\left(-\frac{b_1}{\lambda_1}, -\frac{b_2}{\lambda_2}\right)$ при сохранении направлений осей. Уравнение (6) приведется тогда к виду

$$\lambda_1 x''^2 + \lambda_2 y''^2 + c = 0. \tag{8}$$

Предположим сначала, что $\lambda_1\lambda_2 > 0$ (т. е. что $\delta > 0$). В этом случае геометрическое место точек, координаты которых удовлетворяют уравнению (8), представляет собой *эллипс* (рис. 15, *а*), если знак c противоположен знаку λ_1; оно сводится к *одной точке*, если $c = 0$, и *совсем* *не содержит точек*, если знак c совпадает со знаком λ_1 *).

Пусть теперь $\lambda_1\lambda_2 < 0$ (т. е. $\delta < 0$); тогда (8) будет уравнением *гиперболы*, если $c \neq 0$ (рис. 15, *б*), и *пары пересекающихся прямых*, если $c = 0$.

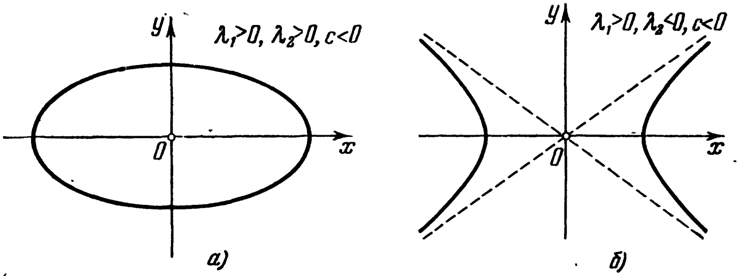


Рис. 15.

В случае I линия представляет собой *центральную кривую второго порядка* (легко видеть, что для содержащей хотя бы одну точку кривой (8) начало координат является центром симметрии).

II. $\delta = \lambda_1\lambda_2 = 0$, и пусть, например, $\lambda_2 \neq 0$. Уравнение (1) приводится к виду

$$\lambda_2 y'^2 + 2b_1 x' + 2b_2 y' + b = 0. \tag{9}$$

Если $b_1 \neq 0$ то, выделив полный квадрат, будем иметь

$$\lambda_2\left(y' + \frac{b_2}{\lambda_2}\right)^2 + 2b_1\left(x' + \frac{b}{2b_1} - \frac{b_2^2}{2\lambda_2 b_1}\right) = 0.$$

После переноса начала координат

$$x'' = x' + \frac{b}{2b_1} - \frac{b_2^2}{2\lambda_2 b_1}, \quad y'' = y' + \frac{b_2}{\lambda_2}$$

уравнение (9) принимает вид

$$\lambda_2 y''^2 + 2b_1 x'' = 0. \tag{10}$$

*) Вместо «точки», определяемой уравнением $\lambda_1 x''^2 + \lambda_2 y''^2 = 0$, говорят также о паре «мнимых прямых» $y'' = \pm i\sqrt{\frac{\lambda_1}{\lambda_2}}x''$, пересекающихся в вещественной (т. е обыкновенной, реально существующей) точке. «Пустое множество точек» $\lambda_1 x''^2 + \lambda_2 y''^2 + c = 0$, где λ_1, λ_2, c — одного знака, называют также «мнимым эллипсом».

Это — каноническое уравнение *параболы* (рис. 16). В случае, когда коэффициент $b_1 = 0$, уравнение (9) приводится к виду

$$\lambda_2\left(y' + \frac{b_2}{\lambda_2}\right)^2 + b - \frac{b_2^2}{\lambda_2} = 0,$$

и после подстановки

$$x'' = x', \quad y'' = y' + \frac{b_2}{\lambda_2}$$

принимает следующий вид:

$$\lambda_2 y''^2 + c = 0, \qquad (11)$$

где $c = b - \frac{b_2^2}{\lambda_2}$.

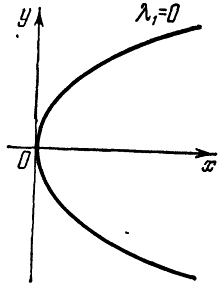


Рис. 16.

Это — *пара параллельных прямых*, если $c\lambda_2 < 0$, *пара совпадающих прямых*, если $c = 0$, и *«пустое множество точек»* (не содержащее ни одной точки) при $c\lambda_2 > 0$ *).

Таким образом, утверждение, сформулированное в начале параграфа, доказано.

§ 2. Инварианты кривой второго порядка

Слово инвариантный значит неизменный. *Инвариантами кривой называются такие выражения, составленные из коэффициентов ее уравнения, которые не меняются при переходе от одной прямоугольной декартовой системы координат к другой такой же системе,* т. е. при поворотах осей координат и при параллельных переносах осей.

Теорема 1. *Для кривой второго порядка* (1) *сумма коэффициентов при квадратах координат*

$$s = a_{11} + a_{22},$$

определитель, составленный из коэффициентов при старших членах:

$$\delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix},$$

*) Уравнение $\lambda_2 y''^2 + c = 0$, где $\lambda_2 c > 0$, определяет, как иногда говорят, «пару мнимых параллельных прямых»: $y'' = \pm i\sqrt{\frac{c}{\lambda_2}}$.

и определитель третьего порядка

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a \end{vmatrix}$$

являются инвариантами.

Доказательство. Рассмотрим отдельно перенос начала координат и поворот координатных осей. Предположим прежде, что *начало координат* (при сохранении направлений осей) *переносится* в точку с координатами (α, β). Тогда

$$x = x' + \alpha,$$
$$y = y' + \beta,$$

где x' и y' — новые координаты. Подставляя эти значения x и y в уравнение (1), получим

$$a_{11}(x'+\alpha)^2 + 2a_{12}(x'+\alpha)(y'+\beta) + a_{22}(y'+\beta)^2 + \\ + 2a_1(x'+\alpha) + 2a_2(y'+\beta) + a = 0,$$

или

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + \\ + 2(a_{11}\alpha + a_{12}\beta + a_1)x' + 2(a_{12}\alpha + a_{22}\beta + a_2)y' + \\ + (a_{11}\alpha^2 + 2a_{12}\alpha\beta + a_{22}\beta^2 + 2a_1\alpha + 2a_2\beta + a) = 0. \qquad (12)$$

Мы видим, что группа старших членов вообще не изменилась, отсюда инвариантность s и δ очевидна. (Заметим, кстати, что коэффициент при x' равен $2(a_{11}\alpha + a_{12}\beta + a_1) = f'_x(\alpha, \beta)$ — частной производной от левой части уравнения (1) по x, взятой при $x = \alpha$, $y = \beta$; коэффициент при y' равен $f'_y(\alpha, \beta)$, а свободный член равен $f(\alpha, \beta)$, так что окончательно преобразованное уравнение принимает вид

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + f'_x(\alpha, \beta)\,x' + \\ + f'_y(\alpha, \beta)\,y' + f(a, \beta) = 0$$

Для уравнения (12) определитель Δ равен

$$\begin{vmatrix} a_{11} & a_{12} & a_{11}\alpha + a_{12}\beta + a_1 \\ a_{12} & a_{22} & a_{12}\alpha + a_{22}\beta + a_2 \\ a_{11}\alpha + a_{12}\beta + a_1 & a_{12}\alpha + a_{22}\beta + a_2 & \begin{matrix} a_{11}\alpha^2 + 2a_{12}\alpha\beta + a_{22}\beta^2 + \\ + 2a_1\alpha + 2a_2\beta + a \end{matrix} \end{vmatrix}.$$

Вычитая из последней строки этого определителя первую, умноженную на α, и вторую, умноженную на β, получим

$$\begin{vmatrix} a_{11} & a_{12} & a_{11}\alpha + a_{12}\beta + a_1 \\ a_{12} & a_{22} & a_{12}\alpha + a_{22}\beta + a_2 \\ a_1 & a_2 & a_1\alpha + a_2\beta + a \end{vmatrix}.$$

А проделав такие же операции над столбцами полученного опеределителя, найдем, что он равен

$$\begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a \end{vmatrix},$$

т. е. равен старому определителю Δ. Таким образом, инвариантность Δ при переносах начала координат тоже доказана.

Далее, при *повороте осей* координат на угол φ мы переходим от одного ортонормированного базиса к другому — такому же; следовательно, матрица квадратичной формы

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2$$

преобразуется так же, как матрица соответствующего линейного преобразования (см. § 5 главы VI). Но для линейного преобразования с матрицей

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$$

коэффициенты его характеристического уравнения

$$\varphi(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{12} & a_{22} - \lambda \end{vmatrix} =$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}^2 = \lambda^2 - s\lambda + \delta$$

вообще не зависят от выбора базиса (теорема 6 § 8 главы III). Этим доказана инвариантность s и δ при поворотах координатных осей. Аналогично, можно доказать и инвариантность определителя Δ.

В самом деле, если мы перейдем к новому базису e_1', e_2', где

$$e_1' = \cos\varphi \cdot e_1 + \sin\varphi \cdot e_2,$$

$$e_2' = -\sin\varphi \cdot e_1 + \cos\varphi \cdot e_2,$$

то координаты преобразуются по формулам

$$\begin{aligned} x &= \cos\varphi\cdot x' - \sin\varphi\cdot y', \\ y &= \sin\varphi\cdot x' + \cos\varphi\cdot y'. \end{aligned} \tag{13}$$

В трехмерном евклидовом пространстве R^3 в ортонормированном базисе e_1, e_2, e_3 рассмотрим квадратичную форму от трех переменных:

$$F(x, y, z) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_1xz + 2a_2yz + az^2,$$

которая при $z = 1$ превращается в $f(x, y)$. При переходе к новому базису с матрицей перехода

$$\begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

координаты в R^3 преобразуются по формулам

$$\begin{aligned} x &= \cos\varphi\cdot x' - \sin\varphi\cdot y', \\ y &= \sin\varphi\cdot x' + \cos\varphi\cdot y', \\ z &= z'. \end{aligned} \tag{14}$$

Если $f(x, y)$ при подстановке (13) переходит в

$$b_{11}x'^2 + 2b_{12}x'y' + b_{22}y'^2 + 2b_1x' + 2b_2y' + a$$

(свободный член при этом, очевидно, не меняется), то ясно, что $F(x, y, z)$ при подстановке (14) перейдет в

$$b_{11}x'^2 + 2b_{12}x'y' + b_{22}y'^2 + 2b_1x'z' + 2b_2y'z' + az'^2.$$

Но при переходе к новому (ортонормированному!) базису определитель матрицы квадратичной формы не меняется, следовательно, для формы $F(x, y, z)$ имеет место равенство

$$\begin{vmatrix} b_{11} & b_{12} & b_1 \\ b_{12} & b_{22} & b_2 \\ b_1 & b_2 & a \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a \end{vmatrix}.$$

Левая часть его есть определитель Δ для $f(x, y)$ в новом базисе e_1', e_2' а правая часть — в старом. Следовательно, и при поворотах координатных осей этот определитель также не меняется. Теорема полностью доказана.

По значению δ можно судить о типе кривой: если $\delta > 0$, перед нами кривая эллиптического типа (эллипс, точка или «пустое множество» — «мнимый эллипс»), если $\delta < 0$ — кривая гиперболического типа (гипербола или пара пересекающихся вещественных прямых), если $\delta = 0$ — кривая параболического типа (парабола или пара параллельных прямых, возможно, совпадающих или даже не существующих — «мнимых»).

Установленная в теореме 1 инвариантность выражеий s, δ и Δ облегчает приведение уравнения кривой к аноническому виду. Так, например, в случае центральой кривой, т. е. при $\delta \neq 0$, уравнение кривой, как мы идели, приводится к виду

$$\lambda_1 x^2 + \lambda_2 y^2 + c = 0,$$

де λ_1, λ_2 — собственные значения линейного оператора матрицей $\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$. Но для последнего уравнения

$$\Delta = \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{vmatrix} = \lambda_1 \lambda_2 c,$$

ткуда $\lambda_1 \lambda_2 c = \Delta$, или $\delta c = \Delta$, и $c = \frac{\Delta}{\delta}$. Таким образом, каноническое», т. е уже упрощенное, уравнение центральной кривой второго порядка, будет иметь вид

$$\lambda_1 x^2 + \lambda_2 y^2 + \frac{\Delta}{\delta} = 0.$$

Если

$$\delta > 0 \text{ и } \Delta \neq 0,$$

о наша кривая — эллипс или «мнимый эллипс». на будет эллипсом (вещественным), если λ_1 и $\frac{\Delta}{\delta}$ азных знаков, т. е. если $\lambda_1 \frac{\Delta}{\delta} < 0$; но так как $\delta > 0$ λ_1 — одного знака с s, то это будет, если $s\Delta < 0$. Криая будет «мнимым эллипсом» в том случае, когда > 0. Если же $\delta > 0$ и $\Delta = 0$, то кривая представляет обой точку.

Если $\delta < 0$, то кривая является гиперболой при $\neq 0$ и распадается на пару пересекающихся рямых при $\Delta = 0$.

Для параболы, уравнение которой приведено к виу (10),

$$\Delta = \begin{vmatrix} 0 & 0 & b_1 \\ 0 & \lambda_2 & 0 \\ b_1 & 0 & 0 \end{vmatrix} = -b_1^2 \lambda_2,$$

откуда

$$b_1 = \pm\sqrt{-\frac{\Delta}{\lambda_2}} = \pm\sqrt{-\frac{\Delta}{s}}.$$

Здесь $b_1 \neq 0$ и, значит, $\Delta \neq 0$.

В случае пары параллельных прямых (различных, совпадающих или «мнимых») уравнение кривой приводится к виду

$$\lambda_2 y^2 + c = 0.$$

В этом случае

$$\Delta = \begin{vmatrix} 0 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{vmatrix} = 0.$$

Соберем результаты двух последних параграфов в следующую таблицу:

$\delta > 0$ Кривая эллиптического типа	$\Delta \neq 0$	$s\Delta < 0$. Эллипс
		$s\Delta > 0$. «Мнимый эллипс»
	$\Delta = 0$	Точка (пара пересекающихся в этой точке «мнимых прямых»)
$\delta < 0$ Кривая гиперболического типа	$\Delta \neq 0$	Гипербола
	$\Delta = 0$	Пара пересекающихся прямых
$\delta = 0$ Кривая параболического типа	$\Delta \neq 0$	Парабола
	$\Delta = 0$	Пара параллельных прямых (различных, совпадающих или «мнимых»)

Из этой таблицы, в частности, видно, что *определитель Δ равен нулю в том и только в том случае, когд кривая распадается на пару* (действительных или «мнимых») *прямых*.

Примеры. *Определить типы следующих кривых и привест их уравнения к каноническому виду.*

1. $3x^2 - 2xy + 3y^2 + 2x - 4y + 1 = 0.$
2. $3x^2 - 2xy + 3y^2 + 2x - 4y + 2 = 0.$
3. $x^2 + y^2 + 2x + 1 = 0.$
4. $x^2 + 2xy - y^2 - 6x + 4y - 3 = 0.$
5. $x^2 + 3xy + 2y^2 + 2x + 5y - 3 = 0.$

6. $x^2-2xy+y^2+4x-6y+1=0.$
7. $x^2+4xy+4y^2-2x-4y-3=0.$
8. $x^2+4xy+4y^2-2x-4y+1=0.$
9. $x^2+4xy+4y^2+2x+4y+2=0.$

Решение.

1. $$\delta=\begin{vmatrix}3 & -1\\ -1 & 3\end{vmatrix}=8>0;$$

это — кривая эллиптического типа. Так как

$$\Delta=\begin{vmatrix}3 & -1 & 1\\ -1 & 3 & -2\\ 1 & -2 & 1\end{vmatrix}=-3\neq 0,$$

о кривая не распадается. Поскольку $s=3+3=6$ и $s\Delta=-18<$ <0, то кривая представляет собой *эллипс*.

Далее,

$$\varphi(\lambda)=\lambda^2-s\lambda+\delta=\lambda^2-6\lambda+8=0,$$
$$\lambda=3\pm 1,\ \lambda_1=4,\ \lambda_2=2.$$

Каноническое уравнение кривой

$$4x'^2+2y'^2-\frac{3}{8}=0,\quad \text{или}\quad \frac{32}{3}x'^2+\frac{16}{3}y'^2=1;$$

полуоси этого эллипса

$$a=\frac{\sqrt{3}}{4\sqrt{2}}\approx 0{,}3;\quad b=\frac{\sqrt{3}}{4}\approx 0{,}4.$$

2. $$\delta=\begin{vmatrix}3 & -1\\ -1 & 3\end{vmatrix}=8>0;$$

кривая эллиптического типа. Здесь

$$\Delta=\begin{vmatrix}3 & -1 & 1\\ -1 & 3 & -2\\ 1 & -2 & 2\end{vmatrix}=5\neq 0;$$

ривая не распадается Так как $s=6$ и $s\Delta=30>0$, то это — «мнимый эллипс» («пустое множество» точек).

3. $$\delta=\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix}=1>0;$$

кривая эллиптического типа.

$$\Delta=\begin{vmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 1\end{vmatrix}=0.$$

Эта кривая, уравнение которой можно записать в виде

$$(x+1)^2+y^2=0,$$

представляет собой *точку*

$$x=-1, \quad y=0$$

(ее можно также понимать как *пару* пересекающихся в этой точке *«мнимых прямых»* $x+iy+1=0$ и $x-iy+1=0$).

4. $$\delta=\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}=-2<0;$$

кривая гиперболического типа. Поскольку

$$\Delta=\begin{vmatrix} 1 & 1 & -3 \\ 1 & -1 & 2 \\ -3 & 2 & -3 \end{vmatrix}=-1\neq 0,$$

то это — *гипербола*. Далее,

$$s=0, \quad \delta=-2 \quad \text{и} \quad \varphi(\lambda)=\lambda^2-2; \quad \lambda_1=\sqrt{2}, \quad \lambda_2=-\sqrt{2}.$$

Каноническое уравнение кривой

$$\sqrt{2}x'^2-\sqrt{2}y'^2+\frac{1}{2}=0, \quad \text{или} \quad 2\sqrt{2}y'^2-2\sqrt{2}x'^2=1;$$

полуоси этой гиперболы $a=b\approx 0{,}6$.

5. $$\delta=\begin{vmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & 2 \end{vmatrix}=-\frac{1}{4}<0;$$

кривая гиперболического типа. Здесь

$$\Delta=\begin{vmatrix} 1 & \frac{3}{2} & 1 \\ \frac{3}{2} & 2 & \frac{5}{2} \\ 1 & \frac{5}{2} & -3 \end{vmatrix}=0$$

и, значит, кривая распадается на *пару пересекающихся прямых*. Следовательно, левая часть уравнения кривой распадается на два линейных множителя. Чтобы найти эти множители, можно поступить, например, следующим образом. Уравнение

$$x^2+3xy+2y^2+2x+5y-3=x^2+(3y+2)x+2y^2+5y-3=0$$

решим относительно x (так как уже известно, что левая часть уравнения распадается на линейные множители, то x будет рационально выражаться через y):

$$x = -\left(\frac{3}{2}y + 1\right) \pm \sqrt{\frac{9}{4}y^2 + 3y + 1 - 2y^2 - 5y + 3} =$$

$$= -\left(\frac{3}{2}y + 1\right) \pm \left(\frac{1}{2}y - 2\right),$$

$$x_1 = -y - 3, \quad x_2 = -2y + 1.$$

Левая часть уравнения распадается, следовательно, на множители: $(x + y + 3)(x + 2y - 1) = 0$, и кривая распадается на пару прямых:

$$x + y + 3 = 0 \quad \text{и} \quad x + 2y - 1 = 0.$$

6.
$$\delta = \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} = 0;$$

кривая параболического типа. Так как

$$\Delta = \begin{vmatrix} 1 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 1 \end{vmatrix} = -1 \neq 0,$$

то это — *парабола*. Далее,

$$s = 2, \quad \delta = 0 \quad \text{и} \quad \varphi(\lambda) = \lambda^2 - 2\lambda, \quad \lambda_1 = 0, \quad \lambda_2 = 2, \quad b_1 = \pm\sqrt{\frac{1}{2}}.$$

Каноническое уравнение кривой:

$$2y'^2 \pm \sqrt{2}x' = 0, \quad \text{или} \quad y'^2 = \pm\frac{1}{\sqrt{2}}x'.$$

7.
$$\delta = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0;$$

кривая параболического типа. Далее

$$\Delta = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & -3 \end{vmatrix} = 0,$$

эта кривая распадается на *пару параллельных прямых:*

$$x + 2y + 1 = 0, \quad x + 2y - 3 = 0.$$

8.
$$\delta = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0;$$

ривая параболического типа. Здесь

$$\Delta = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{vmatrix} = 0,$$

и наша кривая состоит из *двух совпавших прямых:*

$$f(x, y) = (x + 2y - 1)^2 = 0, \quad x + 2y - 1 = 0,$$

9. $$\delta = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0;$$

кривая параболического типа. Здесь

$$\Delta = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 2 \end{vmatrix} = 0;$$

эта кривая представляет собой «пустое множество» точек. (Ее уравнение можно переписать так:

$$f(x, y) = (x + 2y + 1)^2 + 1 = (x + 2y + 1 + i)(x + 2y + 1 - i) = 0;$$

говорят поэтому, что она представляет собой *«пару параллельных мнимых прямых»*).

§ 3. Определение центра и главных осей центральной кривой. Отыскание вершины и оси параболы

В этом параграфе мы будем предполагать, что $\Delta \neq 0$, т. е. что кривая не распадается на пару прямых.

Пусть дано общее уравнение второго порядка (1). Найдем собственные значения λ_1, λ_2 матрицы (4) и соответствующие им собственные векторы e'_1, e'_2. Мы знаем, что в базисе, образованном этими векторами, квадратичная форма $a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ приводится к сумме квадратов $\lambda_1 x'^2 + \lambda_2 y'^2$, а уравнение (1) — к виду (6). Собственные векторы e'_1 и e'_2 матрицы (4) находятся, как известно, из систем уравнений

$$\begin{cases} (a_{11} - \lambda_1) x_1 + a_{12} y_1 = 0, \\ a_{12} x_1 + (a_{22} - \lambda_1) y_1 = 0 \end{cases} \quad \text{и} \quad \begin{cases} (a_{11} - \lambda_2) x_2 + a_{12} y_2 = 0, \\ a_{12} x_2 + (a_{22} - \lambda_2) y_2 = 0, \end{cases}$$

каждая из которых, поскольку ее определитель

$$\begin{vmatrix} a_{11} - \lambda_i & a_{12} \\ a_{12} & a_{22} - \lambda_i \end{vmatrix}, \qquad i = 1, 2,$$

равен нулю, сводится к одному уравнению, например,

$$(a_{11} - \lambda_1) x_1 + a_{12} y_1 = 0 \quad \text{для первой системы;}$$

$$(a_{11} - \lambda_2) x_2 + a_{12} y_2 = 0 \quad \text{для второй системы.}$$

Следовательно, для $e_1' = (x_1, y_1)$ имеем

$$\frac{y_1}{x_1} = \frac{\lambda_1 - a_{11}}{a_{12}},$$

а для $e_2' = (x_2, y_2)$ —

$$\frac{y_2}{x_2} = \frac{\lambda_2 - a_{11}}{a_{12}}.$$

Таким образом, угловые коэффициенты новых осей координат в старой системе равны

$$k_1 = \frac{\lambda_1 - a_{11}}{a_{12}} \quad \text{(для новой оси } x\text{, соответствующей } \lambda_1\text{)}$$

и

$$k_2 = \frac{\lambda_2 - a_{11}}{a_{12}} \quad \text{(для новой оси } y\text{, соответствующей } \lambda_2\text{)}.$$

В дальнейшем достаточно, как мы видели, лишь переноса начала координат для того, чтобы уравнение кривой привелось к каноническому виду; следовательно, k_1 и k_2 определяют направления главных осей кривой (1).

Предположим, что мы рассматриваем *центральную* кривую второго порядка, т. е. что $\delta \neq 0$. Для того чтобы найти центр кривой, т. е. начало новой системы координат, воспользуемся следующим элементарным соображением. Мы уже видели, что если, не меняя направлений осей, перенести начало координат в точку (α, β), т. е. если положить

$$x = x' + \alpha, \quad y = y' + \beta,$$

то уравнение (1) приведется к виду

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + f_x'(\alpha, \beta)\,x' + f_y'(\alpha, \beta)\,y' + f(\alpha, \beta) = 0.$$

Рассмотрим систему уравнений

$$\left\{\begin{aligned} \frac{1}{2} f_x'(x, y) &= a_{11}x + a_{12}y + a_1 = 0, \\ \frac{1}{2} f_y'(x, y) &= a_{12}x + a_{22}y + a_2 = 0. \end{aligned}\right. \qquad (15)$$

Так как ее определитель δ, по предположению, не равен нулю, то она имеет (единственное) решение α, β.

Если перенести начало координат в точку (α, β), то в уравнении кривой исчезнут члены с первыми степенями x', y', и значит, новое начало координат будет *центром* кривой. Таким образом, *центр центральной кривой второго порядка* (эллипса и гиперболы) *определяется из системы уравнений* (15).

Рассмотрим теперь нецентральную кривую второго порядка (при $\delta = 0$). Так как мы условились, что $\Delta \neq 0$, то это — *парабола.* Собственные значения матрицы (4) пусть будут $\lambda_1 = 0$ и λ_2; направления новых осей определяются по-прежнему:

$$k_1 = \frac{\lambda_1 - a_{11}}{a_{12}} = -\frac{a_{11}}{a_{12}} \text{ (для оси } Ox', \text{ соответствующей } \lambda_1 = 0)$$

и

$$k_2 = \frac{\lambda_2 - a_{11}}{a_{12}} \text{ (для оси } Oy', \text{ соответствующей } \lambda_2).$$

Новое начало координат, т. е. *вершину* (α, β) параболы, можно найти следующим образом. Для параболы, заданной каноническим уравнением $y^2 = 2px$, ось Oy служит касательной в вершине. Новая ось Oy в старых координатных осях имеет угловой коэффициент $k_2 = \frac{\lambda_2 - a_{11}}{a_{12}}$. Так как она служит касательной к параболе в ее вершине (α, β), то k_2 должно равняться производной y'_x в этой точке. Чтобы найти y'_x, продифференцируем уравнение (1) по x, считая y функцией от x; мы получим

$$f'_x(x, y) + f'_y(x, y)\, y'_x = 0,$$

или, подробнее,

$$(a_{11}x + a_{12}y + a_1) + (a_{12}x + a_{22}y + a_2)\, y'_x = 0,$$

откуда

$$y'_x = -\frac{a_{11}x + a_{12}y + a_1}{a_{12}x + a_{22}y + a_2}.$$

Следовательно, в вершине (α, β) параболы

$$k_2 = -\frac{a_{11}\alpha + a_{12}\beta + a_1}{a_{12}\alpha + a_{22}\beta + a_2},$$

откуда

$$(a_{11}\alpha + a_{12}\beta + a_1) + k_2(a_{12}\alpha + a_{22}\beta + a_2) = 0,$$

или, короче,

$$f'_x(\alpha, \beta) + k_2 f'_y(\alpha, \beta) = 0.$$

Таким образом, *координаты вершины* (α, β) *параболы можно найти, решив систему уравнений, состоящую из уравнения*

$$f'_x(x, y) + k_2 f'_y(x, y) = 0 \tag{16}$$

и уравнения (1)

Выясним геометрический смысл уравнения (16), в более подробной записи имеющего вид

$$(a_{11}x + a_{12}y + a_1) + k_2(a_{12}x + a_{22}y + a_2) = 0.$$

Это — прямая, принадлежащая пучку, который определяется прямыми

$$a_{11}x + a_{12}y + a_1 = 0 \text{ и } a_{12}x + a_{22}y + a_2 = 0.$$

Угловые коэффициенты $-\frac{a_{11}}{a_{12}}$ и $-\frac{a_{12}}{a_{22}}$ этих прямых равны между собой, так как $\delta = 0$, и равны k_1; следовательно, эти прямые параллельны новой оси Ox. Значит, и принадлежащая определяемому ими пучку прямая (16) тоже параллельна новой оси Ox. Но так как она проходит через вершину, то это — *ось симметрии параболы, ее главный диаметр.*

§ 4. Исследование общего уравнения поверхности второго порядка

В этом параграфе мы будем заниматься только приведением общего уравнения ***поверхности второго порядка*** к каноническому виду.

Пусть в прямоугольной декартовой системе координат в пространстве задано уравнение

$$f(x, y, z) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}xz + 2a_{23}yz + a_{33}z^2 + 2a_1x + \\ + 2a_2y + 2a_3z + a = 0. \tag{17}$$

Рассмотрим квадратичную форму от трех переменных:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}xz + 2a_{23}yz + a_{33}z^2.$$

В некотором, тоже ортонормированном базисе она приводится к сумме квадратов:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2.$$

При этом уравнение (17) приводится к виду

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2b_1 x' + 2b_2 y' + 2b_3 z' + b = 0$$

Здесь возможны три случая:

I. Все λ_i отличны от нуля.

II. Одно из λ_i равно нулю.

III. Два из λ_i равны нулю.

Рассмотрим каждый из этих случаев отдельно.

I. $\lambda_1\lambda_2\lambda_3 \neq 0$. Точно таким же образом, как и в случае кривой второго порядка, можно избавиться от членов первой степени:

$$\lambda_1\left(x' + \frac{b_1}{\lambda_1}\right)^2 + \lambda_2\left(y' + \frac{b_2}{\lambda_2}\right)^2 + \lambda_3\left(z' + \frac{b_3}{\lambda_3}\right)^2 + c = 0.$$

Сделав подстановку

$$x'' = x' + \frac{b_1}{\lambda_1}, \qquad y'' = y' + \frac{b_2}{\lambda_2}, \qquad z'' = z' + \frac{b_3}{\lambda_3},$$

т. е. выполнив некоторый параллельный перенос осей координат, мы получим уравнение

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 + c = 0.$$

Это — уравнение *центральной поверхности* второго порядка (новое начало координат является ее ц е н т р о м).

Будем считать, что $c \leqslant 0$ (в противном случае умножим уравнение на -1). При $c < 0$ возможны следующие случаи:

1. $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ — *эллипсоид.*
2. $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 < 0$ — *однополостный гиперболоид.*
3. $\lambda_1 > 0$, $\lambda_2 < 0$, $\lambda_3 < 0$ — *двуполостный гиперболоид.*
4. $\lambda_1 < 0$, $\lambda_2 < 0$, $\lambda_3 < 0$ — *«пустое множество»* точек (его называют также *«мнимым эллипсоидом»*).

Если $c = 0$ и все λ_i одного знака, получается *точка* («мнимый конус»); при $c = 0$ и λ_i разных знаков — *конус.*

II. Один из коэффициентов λ_i равен нулю; пусть, например, $\lambda_3 = 0$. Тогда соответствующим переносом начала координат уравнение поверхности можно привести к виду

$$\lambda_1 x''^2 + \lambda_2 y''^2 + 2b_3 z'' + b = 0. \tag{18}$$

Здесь возможны случаи $b_3 = 0$ и $b_3 \neq 0$.

При $b_3 = 0$ уравнение (18) имеет вид

$$\lambda_1 x''^2 + \lambda_2 y''^2 + b = 0.$$

Это — уравнение *цилиндрической поверхности*, вид которой определяется ее направляющей $\lambda_1 x''^2 + \lambda_2 y''^2 + b = 0$ в плоскости $x''O'y''$ (*эллиптический цилиндр, гиперболический цилиндр, пара пересекающихся плоскостей, одна прямая*, или *пара «мнимых плоскостей»*, пересекающихся по вещественной прямой, *«пустое множество» точек*, или *«мнимый эллиптический цилиндр»*).

При $b_3 \neq 0$ уравнение (18) приводится к виду

$$\lambda_1 x''^2 + \lambda_2 y''^2 + 2b_3 z'' = 0.$$

Если $\lambda_1\lambda_2 > 0$, это — *эллиптический параболоид*, при $\lambda_1\lambda_2 < 0$ — *гиперболический параболоид*.

III. Среди чисел λ_i два равны нулю, пусть, например, $\lambda_2 = 0$ и $\lambda_3 = 0$. Уравнение (17) приводится к виду

$$\lambda_1 x'^2 + 2b_2 y' + 2b_3 z' + b = 0. \qquad (19)$$

Если $b_2 = 0$ и $b_3 = 0$ — это *пара параллельных плоскостей*, *различных* при $\lambda_1 b < 0$, *совпадающих* при $b = 0$ и *«мнимых»* при $\lambda_1 b > 0$.

Наконец, если хотя бы один из коэффициентов b_2, b_3 уравнения (19) отличен от нуля, положим

$$x' = x'', \quad y' = \frac{b_2 y'' + b_3 z''}{\sqrt{b_2^2 + b_3^2}}, \quad z' = \frac{b_3 y'' - b_2 z''}{\sqrt{b_2^2 + b_3^2}},$$

что, как легко видеть, отвечает переходу к новому (тоже ортонормированному) базису с матрицей перехода

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & \dfrac{b_2}{\sqrt{b_2^2 + b_3^2}} & \dfrac{b_3}{\sqrt{b_2^2 + b_3^2}} \\ 0 & \dfrac{b_3}{\sqrt{b_2^2 + b_3^2}} & \dfrac{-b_2}{\sqrt{b_2^2 + b_3^2}} \end{Vmatrix}.$$

При этом уравнение (19) преобразуется в

$$\lambda_1 x''^2 + 2\sqrt{b_2^2 + b_3^2}\, y'' + b = 0,$$

а это последнее уравнение, так как $\sqrt{b_2^2 + b_3^2} \neq 0$, посредством переноса начала координат преобразуется в

$$\lambda_1 x'''^2 + 2\sqrt{b_2^2 + b_3^2}\, y''' = 0.$$

Это — *параболический цилиндр*.

Заметим без доказательства, что, как и в случае кривой второго порядка, при преобразовании уравнения поверхности второго порядка можно использовать и н в а р и а н т ы. Здесь это будут

$$s_1 = a_{11} + a_{22} + a_{33},$$

$$s_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix}, \quad \delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$$

(с точностью до знаков — это коэффициенты характеристического многочлена матрицы $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}\Bigg)$ и определитель

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_1 \\ a_{12} & a_{22} & a_{23} & a_2 \\ a_{13} & a_{23} & a_{33} & a_3 \\ a_1 & a_2 & a_3 & a \end{vmatrix}.$$

Уравнение центральной поверхности приводится к виду

$$\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2 + \frac{\Delta}{\delta} = 0.$$

Определитель Δ обращается в нуль в том и только в том случае, если поверхность является конической или цилиндрической (в частности, распадается на пару плоскостей — различных, совпадающих или «мнимых»).

ГЛАВА VIII

ПОНЯТИЕ О ТЕНЗОРАХ

В этой главе пространство R предполагается вещественным.

§ 1. Примеры тензоров

Прежде чем дать общее определение тензора, рассмотрим несколько примеров.

1. *Линейный функционал.* Пусть $f(x)$ — линейный функционал (§ 1 главы V) в n-мерном векторном пространстве R. Выберем в R базис $e_1, e_2, \ldots, e_n$ и пусть

$$x = x^1 e_1 + x^2 e_2 + \ldots + x^n e_n$$

— произвольный вектор из R. (Номера координат мы условимся теперь писать сверху; целесообразность этого будет видна из дальнейшего.) Тогда

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

где $a_i = f(e_i)$.

Перейдем к новому базису $e'_1, e'_2, \ldots, e'_n$ и пусть новые базисные векторы получаются из старых по формулам

$$e'_i = c_i^1 e_1 + c_i^2 e_2 + \ldots + c_i^n e_n = \sum_{k=1}^{n} c_i^k e_k. \tag{2}$$

В матрице перехода

$$C = \begin{bmatrix} c_1^1 & c_2^1 & \ldots & c_n^1 \\ c_1^2 & c_2^2 & \ldots & c_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ c_1^n & c_2^n & \ldots & c_n^n \end{bmatrix} \tag{3}$$

мы условимся теперь обозначать номер строки — верхним индексом, а номер столбца — нижним. Пусть в новом базисе $x = \sum_{i=1}^{n} x'^i e'_i$; тогда

$$f(x) = \sum_{i=1}^{n} a'_i x'^i,$$

где

$$a'_i = f(e'_i) = f\left(\sum_{k=1}^{n} c_i^k e_k\right) = \sum_{k=1}^{n} c_i^k f(e_k) = \sum_{k=1}^{n} c_i^k a_k. \quad (4)$$

Таким образом, *линейный функционал $f(x)$ в каждом базисе определяется строкой из n чисел $a_1, a_2, \ldots, a_n$, причем при переходе к новому базису эти числа преобразуются по формулам* (4), *т. е. точно так же, как базисные векторы* (2).

Примем теперь для сокращения записей следующее соглашение (правило Эйнштейна): *если в каком-нибудь выражении один и тот же индекс, скажем i, встречается дважды, один раз наверху и один раз внизу, то имеется в виду, что по этому индексу производится суммирование* (в пределах $i = 1, 2, \ldots, n$), *а знак суммы $\sum_{i=1}^{n}$ в этом случае опускается.* Так, например, по определению,

$$c_i^k e_k = \sum_{k=1}^{n} c_i^k e_k, \quad c_i^k x'^i = \sum_{i=1}^{n} c_i^k x'^i, \quad b_i^{pq} a_r^i = \sum_{i=1}^{n} b_i^{pq} a_r^i, \text{ и т. п.}$$

В этих обозначениях равенство (2) можно переписать так:

$$e'_i = c_i^k e_k,$$

равенство (1) — так:

$$f(x) = a_i x^i,$$

а равенство (4) — так:

$$a'_i = c_i^k a_k.$$

Аналогично, если в одном и том же выражении имеются по две или более пар одинаковых индексов (каждый из которых стоит один раз наверху и один раз внизу), то мы также всегда будем считать, что по этим индексам производится суммирование, причем все эти индексы независимо друг от друга пробегают значения 1, 2, ..., n. Так, например,

$$a_m^{ik}b_{ip}^m = \sum_{i,m=1}^{n} a_m^{ik}b_{ip}^m, \quad a_{ipj}^{ijk} = \sum_{i,j=1}^{n} a_{ipj}^{ijk}, \text{ и т. п.}$$

2. *Вектор*. В заданном базисе $e_1, e_2, \ldots, e_n$ каждый вектор x представляется строкой из n чисел $(x^1, x^2, \ldots, x^n)$ — его координат.

В новом базисе $e'_1, e'_2, \ldots, e'_n$ тот же вектор представляется другой строкой $(x'^1, x'^2, \ldots, x'^n)$, причем если (3) — матрица перехода от первого базиса ко второму, то, как было показано в § 6 главы II,

$$x^k = c_j^k x'^j. \tag{5}$$

Это — выражение старых координат через новые. Выразим отсюда новые координаты x'^j через старые x^i. Пусть $C^{-1} = [b_k^j]$ — матрица, обратная матрице перехода C. Тогда равенство

$$CC^{-1} = C^{-1}C = E$$

равносильно тому, что

$$c_k^i b_j^k = b_k^i c_j^k = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Положим

$$\delta_j^i = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j \end{cases}$$

(так называемый *символ Кронекера*). Тогда

$$c_k^i b_j^k = b_k^i c_j^k = \delta_j^i.$$

Умножив обе части равенства (5) на b_k^i (и, естественно, суммируя по k), мы получим

$$b_k^i x^k = b_k^i c_j^k x'^j = \delta_j^i x'^j = x'^i$$

(так как $\delta_j^i = 0$ при $i \neq j$ и $\delta_i^i = 1$), или

$$x'^i = b_k^i x^k.$$

Таким образом, *новые координаты x'^i вектора x получаются из старых его координат x^i с помощью матрицы C^{-1}, обратной матрице перехода C, причем* *коэффициенты разложений x'^i по x^i образуют* с т р о к и *матрицы* C^{-1}.

В двух рассмотренных примерах (линейный функционал, вектор) есть нечто общее, позволяющее заключить их в рамки общего определения. И линейный функционал, и вектор в каждом базисе определяются n числами, соответственно $a_1, a_2, \ldots, a_n$ и $x^1, x^2, \ldots, x^n$, причем *при переходе к новому базису эти числа преобразуются линейно* — с матрицей C, т. е. т а к ж е, как базисные векторы, в случае линейного функционала, и с матрицей C^{-1}, о б р а т н о й матрице C,— в случае вектора. Коэффициенты линейной формы (так же, как координаты вектора) представляют собой пример *тензора,* если назвать тензором заданную в каждом базисе систему чисел, линейно преобразующихся при переходе от одного базиса к другому. Точное определение этого понятия будет дано ниже; пока же мы только еще добавим, что оба рассмотренных тензора являются о д н о в а л е н т н ы м и, так как определяются системами чисел $a_1, a_2, \ldots, a_n$ или $x^1, x^2, \ldots, x^n$, зависящими от одного индекса. Коэффициенты линейной формы при переходе к новому базису, преобразующиеся так же, как базисные векторы, образуют тензор к о в а р и а н т н ы й, т. е. «сопреобразующийся»— преобразующийся одинаково с базисными векторами. Координаты вектора — пример к о н т р а в а р и а н т н о г о, т. е. «противопреобразующегося» тензора.

Рассмотрим еще три примера.

3. *Билинейный функционал.* Пусть в n-мерном векторном пространстве R задан билинейный функционал $A(x, y)$ (§ 1 главы VI). Тогда, если $x = x^i e_i$ и $y = y^k e_k$ — произвольные векторы из R, то

$$A(x, y) = A(x^i e_i, y^k e_k) = x^i y^k A(e_i, e_k) = a_{ik} x^i y^k,$$

где $a_{ik} = A(e_i, e_k)$: в заданном базисе $e_1, e_2, \ldots, e_n$ билинейный функционал $A(x, y)$ представляется билинейной формой $a_{ik} x^i y^k$ (по i и по k суммирование!) от координат векторов x и y с коэффициентами a_{ik} (ср. стр. 188)

Перейдем к новому базису $e'_1, e'_2, \ldots, e'_n$ с матрицей перехода (3). Тогда, если $x = x'^i e'_i$ и $y = y'^k e'_k$, то

$$A(x, y) = A(x'^i e'_i, y'^k e'_k) = x'^i y'^k A(e'_i, e'_k) = a'_{ik} x'^i y'^k, \quad (6)$$

где

$$a'_{ik} = A(e'_i, e'_k) = A(c^j_i e_j, c^h_k e_h) = c^j_i c^h_k A(e_j, e_h) = c^j_i c^h_k a_{jh} \quad (7)$$

(ср. стр. 189).

Таким образом, билинейный функционал $A(x, y)$ в каждом базисе определяется системой из n^2 чисел a_{ik}, зависящих от двух индексов, причем при переходе к новому базису эти числа преобразуются по закону (7), т. е. по каждому из этих двух индексов так же, как базисные векторы. Это — пример тензора валентности два (зависящего от двух индексов), ковариантного по обоим индексам (дважды ковариантного).

4. *Линейный оператор.* Каждый линейный оператор $\mathcal{A}$ в n-мерном векторном пространстве R в заданном базисе $e_1, e_2, \ldots, e_n$ представляется матрицей $A = [a^i_k]$ (здесь опять верхний индекс — номер строки, нижний — номер столбца). При переходе к новому базису $e'_1, e'_2, \ldots, e'_n$ с матрицей перехода C эта матрица A преобразуется в $C^{-1}AC$ (§ 4 главы III). Вспомним, как выражаются элементы a'^i_k матрицы $C^{-1}AC$ через элементы a^i_k матрицы A. В матрице AC элемент p-й строки и k-го столбца равен $a^p_j c^j_k$. В матрице $C^{-1}AC$ элемент i-й строки и k-го столбца — это $b^i_p a^p_j c^j_k$, т. е.

$$a'^i_k = b^i_p a^p_j c^j_k = c^j_k b^i_p a^p_j. \quad (8)$$

Таким образом, линейный оператор $\mathcal{A}$ в каждом базисе определяется системой из n^2 чисел a^i_k, занумерованных двумя индексами, нижним и верхним, причем при переходе к новому базису эти числа преобразуются по формуле (8) — по нижнему индексу, так же как базисные векторы, а по верхнему — с обратной матрицей, «контравариантно» базисным векторам. Это — еще один пример тензора валентности два (зависящего от двух индексов), в этом случае один раз ковариантного и один раз контравариантного (смешанный двухвалентный тензор).

5. *Символ Кронекера.* Рассмотрим смешанный двухвалентный тензор, координаты которого в некотором

фиксированном базисе $e_1, e_2, \ldots, e_n$ определяются равенствами

$$\delta_j^i = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j \end{cases}$$

(см. стр. 227).

В новом базисе $e'_1, e'_2, \ldots, e'_n$ имеем

$$\delta'^i_j = c_j^p b_q^i \delta_p^q = c_j^p b_p^i = \delta_j^i.$$

Таким образом, *координаты тензора* δ_j^i *одинаковы во всех системах координат.* (Это можно объяснить тем, что в первоначальном базисе $e_1, e_2, \ldots, e_n$ элементы δ_i^i составляют единичную матрицу, и значит, соответствующий тензор определяет *тождественное* преобразование, матрица которого — одна и та же во всех базисах).

§ 2. Определение и простейшие свойства тензоров

Пусть в n-мерном векторном пространстве R в каждом базисе задана система из n^{p+q} чисел $a_{k_1 k_2 \ldots k_p}^{i_1 i_2 \ldots i_q}$ (занумерованных p нижними и q верхними индексами, которые независимо друг от друга пробегают значения $1, 2, \ldots, n$); *предположим, что при переходе к новому базису с матрицей перехода* (3) *эти числа преобразуются по закону*

$$a'^{i_1 i_2 \ldots i_q}_{k_1 k_2 \ldots k_p} = c_{k_1}^{j_1} c_{k_2}^{j_2} \ldots c_{k_p}^{j_p} b_{h_1}^{i_1} b_{h_2}^{i_2} \ldots b_{h_q}^{i_q} a_{j_1 j_2 \ldots j_p}^{h_1 h_2 \ldots h_q}. \tag{9}$$

Тогда мы говорим, что имеем $(p+q)$-валентный ***тензор,*** *p раз ковариантный и q раз контравариантный. Числа* $a_{k_1 k_2 \ldots k_p}^{i_1 i_2 \ldots i_q}$ называются *координатами тензора.*

Скаляр, т. е. величину, во всех системах координат имеющую одно и то же значение, можно рассматривать как тензор нулевой валентности.

Ясно, что *если координаты двух тензоров одинакового строения* (т. е. таких, у которых одинаковы числа ко- и контравариантных индексов) *совпадают в одном каком-нибудь базисе, то они совпадают и во всех остальных* (и значит, эти тензоры равны), так как при переходе к новому базису координаты обоих тензоров преобразуются одинаково. Поэтому *для того, чтобы задать тензор данного строения, достаточно задать его координаты в какой-нибудь одной системе координат.* А это можно сделать без каких-либо ограничений: *в качестве координат тензора в данном базисе можно выбрать совершенно произвольные числа.* Действительно, пусть в базисе $e_1, e_2, \ldots, e_n$ произвольно заданы n^{p+q} чисел $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$. Тогда координаты соответствующего тензора в любом другом базисе найдутся по формуле (9), и нам остается только проверить, что *при переходе от любого базиса $e'_1, e'_2, \ldots, e'_n$ к любому другому базису $e''_1, e''_2, \ldots, e''_n$ координаты полученного тензора тоже преобразуются по формуле* (9).

Покажем это на примере трехвалентного тензора a_k^{ij}. Пусть при переходе от базиса $e_1, e_2, \ldots, e_n$ к базису $e'_1, e'_2, \ldots, e'_n$ имеем

$$e'_j = c_j^m e_m, \tag{10}$$

а при переходе от базиса $e_1, e_2, \ldots, e_n$ к базису $e''_1, e''_2, \ldots, e''_n$ —

$$e''_i = \tilde{c}_i^k e_k. \tag{11}$$

Из равенства (10) получаем

$$b_k^j e'_j = b_k^j c_j^m e_m = \delta_k^m e_m = e_k, \tag{12}$$

а из равенства (11) —

$$\tilde{b}_m^i e''_i = \tilde{b}_m^i \tilde{c}_i^k e_k = \delta_m^k e_k = e_m. \tag{13}$$

Здесь матрица $[b_k^i]$ — обратная к матрице $[c_k^i]$, а матрица $[\tilde{b}_k^i]$ — обратная к $[\tilde{c}_k^i]$. Из равенств (11) и (12) следует, что

$$e''_i = \tilde{c}_i^k e_k = \tilde{c}_i^k b_k^j e'_j = d_i^j e'_j, \qquad \text{где} \qquad d_i^j = \tilde{c}_i^k b_k^j,$$

а из равенств (10) и (13) — что

$$e'_j = c_j^m e_m = c_j^m \tilde{b}_m^i e''_i = f_j^i e''_i, \qquad \text{где} \qquad f_j^i = c_j^m \tilde{b}_m^i.$$

Таким образом, матрицей перехода от базиса $e'_1, e'_2, \ldots, e'_n$ к базису $e''_1, e''_2, \ldots, e''_n$ будет матрица

$$[d_i^j] = [\tilde{c}_i^k b_k^j],$$

а обратный к ней — матрица

$$[f^i_j] = [c^m_j \tilde{b}^i_m].$$

Мы имеем

$$a'^{pq}_r = c^u_r b^p_s b^q_t a^{st}_u \tag{14}$$

и

$$a''^{ij}_k = \tilde{c}^h_k \tilde{b}^i_m \tilde{b}^j_l a^{ml}_h. \tag{15}$$

Из равенства (14) получаем

$$b^r_h c^m_p c^l_q a'^{pq}_r = b^r_h c^m_p c^l_q c^u_r b^p_s b^q_t a^{st}_u =$$
$$= (b^r_h c^u_r)(c^m_p b^p_s)(c^l_q b^q_t) a^{st}_u = \delta^u_h \delta^m_s \delta^l_t a^{st}_u = a^{ml}_h.$$

Подставляя это значение a^{ml}_h в равенство (15), будем иметь

$$a''^{ij}_k = \tilde{c}^h_k \tilde{b}^i_m \tilde{b}^j_l b^r_h c^m_p c^l_q a'^{pq}_r = (\tilde{c}^h_k b^r_h)(\tilde{b}^i_m c^m_p)(\tilde{b}^j_l c^l_q) a'^{pq}_r = d^r_k f^i_p f^j_q a'^{pq}_r;$$

т. е. формулы преобразования координат тензора при переходе от базиса $e'_1, e'_2, \ldots, e'_n$ к базису $e''_1, e''_2, \ldots, e''_n$ имеют в точности такое строение, какое требуется.

В общем случае доказательство аналогично.

Из сказанного вытекает, что если, например, x^i — произвольный одновалентный контравариантный тензор, то его можно рассматривать как совокупность координат некоторого вектора. Действительно, если в одном каком-нибудь базисе $e_1, e_2, \ldots, e_n$ взять вектор с координатами $x^1, x^2, \ldots, x^n$, то и во всех остальных базисах координаты этого вектора и заданного тензора совпадут. Точно так же каждый дважды ковариантный тензор a_{ij} можно рассматривать, как совокупность коэффициентов некоторой билинейной формы, а каждый смешанный двухвалентный тензор a^i_j — как совокупность элементов матрицы некоторого линейного оператора, и т. д.

Тензор, координаты которого не меняют своего значения при транспозиции любых двух индексов из данного множества индексов $i, j, \ldots, m$ (причем все эти индексы — только верхние или только нижние), *называется симметрическим по этим индексам.* Примером симметрического тензора может служить совокупность коэффициентов симметрической билинейной формы, *Свойство тензора быть симметрическим не зависит от выбора базиса* Рассмотрим, например, трехвалентный тензор a_{ij}^k, и пусть в базисе $e_1, e_2, \ldots, e_n$

$$a_{ji}^k = a_{ij}^k.$$

Тогда в базисе $e'_1, e'_2, \ldots, e'_n$

$$a'^k_{ij} = c_i^s c_j^t b_r^k a_{st}^r \quad \text{и} \quad a'^k_{ji} = c_j^s c_i^t b_r^k a_{st}^r.$$

Заменяя во втором равенстве a_{st}^r на a_{ts}^r, найдем, что

$$a'^k_{ji} = c_j^s c_i^t b_r^k a_{ts}^r.$$

Но сумма не зависит от обозначения индекса, по которому производится суммирование; поэтому, заменяя s на t, а t на s, получим

$$a'^k_{ji} = c_j^t c_i^s b_r^k a_{st}^r = a'^k_{ij}.$$

Кососимметрическим по данным индексам $i, j, \ldots, m$ (только верхним или только нижним) *называется тензор, координаты которого* $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$ *изменяют знак при любой транспозиции индексов из выделенной группы* (не меняясь при этом по абсолютной величине). *Свойство тензора быть кососимметрическим по данной группе индексов* тоже *не зависит от выбора базиса.*

Примером кососимметрического тензора может служить кососимметрическая билинейная форма.

§ 3. Операции над тензорами

1. *Сложение.* Пусть даны два тензора одинакового строения $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$ и $\bar{a}_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$. *Сумма* их в каждой системе координат определяется равенством

$$d_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} = a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} + \bar{a}_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}.$$

Легко видеть, что *сумма двух тензоров будет тензором такого же строения.*

2. *Умножение.* Пусть даны два тензора

$$a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} \quad \text{и} \quad \bar{a}_{j_1j_2\ldots j_s}^{m_1m_2\ldots m_r}$$

какого угодно строения. *Произведение* их в каждом базисе определяется как совокупность $n^{p+q+s+r}$ чисел

$$f^{i_1i_2\ldots i_q m_1m_2\ldots m_r}_{k_1k_2\ldots k_p j_1j_2\ldots j_s} = a^{i_1i_2\ldots i_q}_{k_1k_2\ldots k_p} \bar{a}^{m_1m_2\ldots m_r}_{j_1j_2\ldots j_s}.$$

Покажем, что *произведение двух тензоров — тоже тензор* (в нашем случае валентности $p+q+r+s$, $p+s$ раз ковариантный и $q+r$ раз контравариантный). Действительно, в новом базисе

$$a'^{i_1i_2\ldots i_q}_{k_1k_2\ldots k_p} = c^{t_1}_{k_1}c^{t_2}_{k_2}\ldots c^{t_p}_{k_p} b^{i_1}_{l_1}b^{i_2}_{l_2}\ldots b^{i_q}_{l_q} a^{l_1l_2\ldots l_q}_{t_1t_2\ldots t_p}$$

и

$$\bar{a}'^{m_1m_2\ldots m_r}_{j_1j_2\ldots j_s} = c^{h_1}_{j_1}c^{h_2}_{j_2}\ldots c^{h_s}_{j_s} b^{m_1}_{g_1}b^{m_2}_{g_2}\ldots b^{m_r}_{g_r} \bar{a}^{g_1g_2\ldots g_r}_{h_1h_2\ldots h_s},$$

поэтому,

$$f'^{i_1i_2\ldots i_q m_1m_2\ldots m_r}_{k_1k_2\ldots k_p j_1j_2\ldots j_s} = a'^{i_1i_2\ldots i_q}_{k_1k_2\ldots k_p} \bar{a}'^{m_1m_2\ldots m_r}_{j_1j_2\ldots j_s} =$$

$$= c^{t_1}_{k_1}c^{t_2}_{k_2}\ldots c^{t_p}_{k_p} c^{h_1}_{j_1}c^{h_2}_{j_2}\ldots c^{h_s}_{j_s} b^{i_1}_{l_1}b^{i_2}_{l_2}\ldots b^{i_q}_{l_q} b^{m_1}_{g_1}b^{m_2}_{g_2}\ldots b^{m_r}_{g_r} \times$$

$$\times a^{l_1l_2\ldots l_q}_{t_1t_2\ldots t_p} \bar{a}^{g_1g_2\ldots g_r}_{h_1h_2\ldots h_s} =$$

$$= c^{t_1}_{k_1}\ldots c^{t_p}_{k_p} c^{h_1}_{j_1}\ldots c^{h_s}_{j_s} b^{i_1}_{l_1}\ldots b^{i_q}_{l_q} b^{m_1}_{g_1}\ldots b^{m_r}_{g_r} f^{l_1l_2\ldots l_q g_1g_2\ldots g_r}_{t_1t_2\ldots t_p h_1h_2\ldots h_s}.$$

Умножение тензоров не коммутативно. Рассмотрим, например, произведение двух одновалентных ковариантных тензоров при $n = 2$. Координаты одного: a_1 и a_2, координаты другого: b_1 и b_2. Произведение $c_{ik} = a_i b_k$ — дважды ковариантный тензор, координаты которого

$$c_{11} = a_1b_1,\ c_{12} = a_1b_2,\ c_{21} = a_2b_1,\ c_{22} = a_2b_2.$$

Произведение тех же тензоров в обратном порядке — дважды ковариантный тензор с координатами $d_{ik} = b_i a_k$:

$$d_{11} = b_1a_1 = a_1b_1 (= c_{11}), \quad d_{12} = b_1a_2 = a_2b_1 (= c_{21}),$$

$$d_{21} = b_2a_1 = a_1b_2 (= c_{12}), \quad d_{22} = b_2a_2 = a_2b_2 (= c_{22}).$$

Он, вообще говоря, отличен от первого.

Так как *скаляр*, т. е. величина, во всех системах координат имеющая одно и то же значение, является тензором нулевой валентности, то *при умножении тензора на скаляр* (т. е. при умножении всех координат тензора на этот скаляр) *получается тензор того же строения.*

Вычитание тензоров одинакового строения сводится к умножению вычитаемого на -1 и сложению (при этом получается, очевидно, тензор того же строения).

3. *Свертывание тензоров.* Эта специфическая для тензоров операция определяется следующим образом. Пусть дан, например, тензор a^{ij}_{mpq}. Выделим в нем два какие-нибудь индекса, например, j и p (один наверху, другой внизу), отберем среди всех координат тензора те, у которых эти индексы одинаковы, и сложим их все. Мы получим

$$a^{ij}_{mjq} = a^{i1}_{m1q} + a^{i2}_{m2q} + \ldots + a^{in}_{mnq} = b^{i}_{mq}$$

— *свертку тензора* a^{ij}_{mpq} *по индексам* j *и* p.

Так, например, тензор a^{i}_{jk} при $n=2$ имеет восемь координат $a^1_{11}, a^1_{12}, a^1_{21}, a^1_{22}, a^2_{11}, a^2_{12}, a^2_{21}, a^2_{22}$. Свертывая этот тензор по индексам i и j, будем иметь

$$a^1_{1k} + a^2_{2k} = b_k,$$

или, подробнее: $b_1 = a^1_{11} + a^2_{21}$, $b_2 = a^1_{12} + a^2_{22}$.

Свертывая тот же тензор по индексам i и k, получим $a^i_{ji} = c_j$, т. е.

$$c_1 = a^1_{11} + a^2_{12}, \quad c_2 = a^1_{21} + a^2_{22}.$$

Покажем, что *в результате свертывания тензора получается снова* т е н з о р, *имеющий на один нижний и на один верхний индекс меньше, чем исходный тензор.* Произведем, например, свертывание тензора a^{ij}_{mpq} по индексам j и p. Пусть $a^{ij}_{mjq} = b^i_{mq}$. В новом базисе координаты исходного тензора имеют вид

$$a'^{ij}_{mpq} = c^r_m c^s_p c^t_q b^i_g b^j_h a^{gh}_{rst}.$$

Выбрав координаты, у которых $p = j$, и просуммировав по $j = 1, 2, \ldots, n$, получим

$$b'^{i}_{mq} = a'^{ij}_{mjq} = c^r_m c^s_j c^t_q b^i_g b^j_h a^{gh}_{rst}.$$

Но $c^s_j b^j_h = \delta^s_h$, а $\delta^s_h a^{gh}_{rst} = a^{gh}_{rht} = b^g_{rt}$. Следовательно,

$$b'^{i}_{mq} = c^r_m c^t_q b^i_g b^g_{rt}.$$

В общем случае доказательство аналогично.

Операция свертывания может быть применена к тензору и несколько раз. Так, например, свертывая тензор a^{ijk}_{pqr} при $n=2$ по индексам i и q и по индексам k и p, получим тензор $\alpha^j_r = a^{ijk}_{kir}$, или, подробнее:

$$\alpha^1_1 = a^{111}_{111} + a^{211}_{121} + a^{112}_{211} + a^{212}_{221}, \qquad \alpha^1_2 = a^{111}_{112} + a^{211}_{122} + a^{112}_{212} + a^{212}_{222},$$

$$\alpha^2_1 = a^{121}_{111} + a^{221}_{121} + a^{122}_{211} + a^{222}_{221}, \qquad \alpha^2_2 = a^{121}_{112} + a^{221}_{122} + a^{122}_{212} + a^{222}_{222}.$$

При p-кратном свертывании тензора p раз ко- и p раз контравариантного получается, очевидно, *инвариант*, или *скаляр*, — величина, не зависящая от выбора базиса. Это — один из способов получения численных инвариантов. Так, при свертывании тензора a^i_j, определяющего линейный оператор $\mathcal{A}$, получаем инвариант a^i_i, называемый *следом* оператора (след a^i_i — это сумма элементов главной диагонали матрицы A; его инвариантность мы уже установили в § 8 главы III: a^i_i — это коэффициент при λ^{n-1} характеристического многочлена оператора $\mathcal{A}$).

Особенно часто операция свертывания применяется по отношению к произведению двух тензоров — по индексам, взятым в разных сомножителях. Если произведение тензоров α^i_j и β^{hm}_k свертывается по индексам j и h, мы будем говорить короче, что *тензоры α^i_j и β^{hm}_k свертываются по индексам j и h*. Так, например, при свертывании тензора a_i (определяющего линейный функционал $f(x)$) с вектором $x=(x^1, x^2, \ldots, x^n)$ получается скаляр $a_i x^i$, равный, очевидно, $f(x)$.

При двукратном свертывании тензора a_{ij}, определяющего билинейный функционал $A(x, y)$ с парой векторов $x=(x^1, x^2, \ldots, x^n)$ и $y=(y^1, y^2, \ldots, y^n)$, получается скаляр $a_{ij}x^i y^j$, равный значению функционала $A(x, y)$ для данных векторов x и y.

При свертывании тензора a^i_j, определяющего линейный оператор $\mathcal{A}$ с вектором $x=(x^1, x^2, \ldots, x^n)$, получается контравариантный тензор $y^i = a^i_j x^j$. Как следует из § 1 главы III, это — заданный своими координатами y^i преобразованный вектор $\mathcal{A}x$.

Пусть даны два тензора a^i_j и b^h_k, определяющие соответственно линейные операторы $\mathcal{A}$ и $\mathcal{B}$. Свертка их по индексам j и h — смешанный двухвалентный тензор $a^i_j b^j_k$, тоже определяющий, следовательно, некоторый линейный оператор $\mathcal{D}$. Легко видеть, что оператор $\mathcal{D}$ равен произведению $\mathcal{AB}$ операторов $\mathcal{A}$ и $\mathcal{B}$ (в смысле § 2 главы III). Свертка $a^i_j b^h_i$ тех же тензоров по индексам i и k соответствует произведению $\mathcal{BA}$ тех же операторов в обратном порядке.

4. *Симметрирование и альтернирование тензора.* Пусть $A_{i_1 i_2 \dots i_k}$ — произвольный тензор, у которого выделены какие-то k индексов $i_1, i_2, \dots, i_k$ — все только верхние или только нижние. Тогда тензор

$$A_{(i_1 i_2 \dots i_k)} = \frac{1}{k!} \sum_{(i_1, i_2, \dots, i_k)} A_{i_1 i_2 \dots i_k},$$

где, суммирование распространено по всевозможным перестановкам $i_1, i_2, \dots, i_k$ выделенных индексов, будет, очевидно, симметрическим, а тензор

$$A_{[i_1 i_2 \dots i_k]} = \frac{1}{k!} \sum_{(i_1, i_2, \dots, i_k)} (-1)^{[i_1, i_2, \dots, i_k]} A_{i_1, i_2, \dots i_k}$$

— кососимметрическим. Операции получения тензоров $A_{(i_1 i_2 \dots i_k)}$ и $A_{[i_1 i_2 \dots i_k]}$ из данного тензора $A_{i_1 i_2 \dots i_k}$ называются соответственно *симметрированием* и *альтернированием* тензора $A_{i_1 i_2 \dots i_k}$ по индексам $i_1, i_2, \dots, i_k$.

Так, на стр. 191 билинейный функционал $B(x, y)$ был получен симметрированием, а $C(x, y)$ — альтернированием билинейного функционала $A(x, y)$.

§ 4. Тензоры в евклидовом пространстве

Пусть теперь R — n-мерное евклидово пространство. Конечно, все, что говорилось о тензорах в произвольном векторном пространстве, распространяется и на этот случай. Но тензоры в евклидовом пространстве обладают еще и некоторыми специфическими свойствами.

В евклидовом пространстве R для любых двух векторов x, y определено их скалярное произведение (x, y), являющееся симметрическим билинейным функционалом. В заданном базисе $e_1, e_2, \dots, e_n$ оно представляется симметрической билинейной формой

$$(x, y) = g_{ik}x^i y^k,$$

где $g_{ik} = (e_i, e_k)$. Взятые во всех системах координат величины g_{ik} образуют, как мы видели в § 1, дважды ковариантный тензор, который называется (ковариантным) *метрическим тензором пространства R*. Свертка метрического тензора g_{ik} с вектором $x = (x^1, x^2, \ldots, x^k)$

$$x_i = g_{ik}x^k \tag{16}$$

является одновалентным ковариантным тензором. Числа x_i также определяют вектор x, т. е. в некотором смысле тоже являются его координатами; их можно назвать ковариантными координатами вектора x, в отличие от его контравариантных координат x^i. Выясним геометрический смысл ковариантных координат. Так как

$$x_i = g_{ik}x^k = (e_i, e_k)x^k = (e_i, e_kx^k) = (e_i, x),$$

то *ковариантные координаты x_i — это проекции вектора x на базисные векторы*. (Напомним, что контравариантные координаты вектора x — это коэффициенты его разложения

$$x = x^i e_i$$

по базису $e_1, e_2, \ldots, e_n$.)

В ортонормированном базисе

$$g_{ik} = (e_i, e_k) = \begin{cases} 1, & \text{если } i = k, \\ 0, & \text{если } i \neq k, \end{cases}$$

и значит, $x_i = x^i$, т. е. ко- и контравариантные координаты вектора совпадают.

Двойная свертка $g_{ik}x^iy^k$ метрического тензора g_{ik} с векторами $x = (x^1, x^2, \ldots, x^n)$ и $y = (y^1, y^2, \ldots, y^n)$ — это скалярное произведение (x, y); двойная свертка $g_{ik}x^ix^k$ его с вектором x — скалярный квадрат (x, x) вектора x.

Определитель $|g_{ik}|$ матрицы $[g_{ik}]$ отличен от нуля. Действительно, при переходе к новому базису ранг матрицы билинейной формы, в том числе и матрицы $[g_{ik}]$, не меняется. Но в ортонормированном базисе матрица $[g_{ik}]$ — единичная, и ее определитель равен 1; следовательно, и во всех других базисах определитель матрицы $[g_{ik}]$ отличен от нуля.

Пусть $[g^{ik}]$ — матрица, обратная матрице $[g_{ik}]$ в каком-то фиксированном базисе $e_1, e_2, \ldots, e_n$. Тогда $g^{ik}g_{kj} = \delta^i_j$ при всех $i, j = 1, 2, \ldots, n$. Построим дважды контравариантный тензор, координаты которого в базисе $e_1, e_2, \ldots, e_n$ равны g^{ik}; тогда координаты этого тензора во всех остальных базисах определятся по формуле (9). В каждом новом базисе $e'_1, e'_2, \ldots, e'_n$, ввиду тензорного характера операций умножения и свертывания, будем иметь

$$g'^{ik}g'_{kj} = \delta'^i_j = \delta^i_j,$$

и значит, *координаты тензора* g^{ik} *во всех системах координат образуют матрицу, обратную матрице* $[g_{ik}]$. Тензор g^{ik} называется *контравариантным метрическим тензором.*

Переход от контравариантных координат вектора к ковариантным его координатам по формуле (16) можно назвать *опусканием индекса.* Чтобы *поднять индекс,* т. е. перейти от ковариантных координат вектора к его контравариантным координатам, умножим обе части равенства (16) на g^{ji} (и, конечно, просуммируем по i); мы получим $g^{ji}x_i = g^{ji}g_{ik}x^k = \delta^j_k x^k = x^j$.

Операцию опускания или поднятия индекса в евклидовом пространстве (эта операция носит выразительное название *жонглирования индексами*) можно применить к тензору любого строения. Пусть дан, например, трехвалентный тензор a^{ij}_k, один раз ковариантный и два раза контравариантный. Свертка его с метрическим тензором

$$g_{ih}a^{hj}_k = b^j_{ik} \tag{17}$$

будет дважды ковариантным и один раз контравариантным тензором. Свертка

$$g_{pj}b^p_{ik} = c_{jik} \tag{18}$$

— трижды ковариантным, а наоборот, свертка

$$g^{qk}a^{ij}_q = d^{kij} \tag{19}$$

— трижды контравариантным тензором. Если оба базиса $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ — ортонормированные, то

$$g_{ik} = g^{ik} = \begin{cases} 1, & \text{если } i = k, \\ 0, & \text{если } i \neq k, \end{cases}$$

и равенства (17)—(19) последовательно дают $a_k^{ij} = b_{ik}^j$, $b_{ik}^j = c_{jik}$, $a_k^{ij} = d^{kij}$. Таким образом,

$$a_k^{ij} = b_{ik}^j = c_{jik} = d^{kij}.$$

В этом случае ко- и контравариантные индексы при переходе к новому базису ведут себя одинаково, и *закон преобразования тензора определяется исключительно его валентностью*

Последнее можно объяснить еще и следующим образом. Если оба базиса $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ ортонормированные, то матрица перехода C от первого базиса ко второму ортогональна, т. е. $C^{-1} = C'$. Но тогда $b_k^i = c_i^k$ — и в формуле (9) для преобразования, например, тензора a_k^{ij}: $a'^{ij}_k = c_k^p b_q^i b_r^j a_p^{qr}$, можно заменить одно или оба b на c, или, наоборот, c заменить на b. Так, заменяя b_q^i на c_i^q, получим

$$a'^{ij}_k = \sum_{q=1}^{n} c_k^p c_i^q b_r^j a_p^{qr}.$$

Здесь пришлось использовать знак суммы $\sum\limits_{q=1}^{n}$, так как индекс q, по которому происходит суммирование, оба раза стоит наверху. Полагая $a_p^{qr} = \alpha_{qp}^r$, получим для α_{ik}^j закон преобразования в виде

$$\alpha'^j_{ik} = c_k^p c_i^q b_r^j \alpha_{qp}^r = c_i^q c_k^p b_r^j \alpha_{pq}^r,$$

т. е. тензор a_{ik}^j является дважды ковариантным и один раз контравариантным. Но его координаты в обоих базисах равны соответствующим координатам тензора a_k^{ij} один раз ковариантного и дважды контравариантного.

Так, выше (§ 5 главы VI) мы уже видели, что при переходе от одного ортонормированного базиса к другому, тоже ортонормированному, матрица билинейной формы (дважды ковариантный тензор) и матрица линейного преобразования (один раз ковариантный и один раз контравариантный тензор) преобразуются одинаково.

ГЛАВА IX

ОСНОВНЫЕ ПОНЯТИЯ СПЕЦИАЛЬНОЙ ТЕОРИИ ОТНОСИТЕЛЬНОСТИ

Содержание настоящей главы — это всего лишь некоторая интерпретация физических законов. Ясно, что сами эти законы нельзя вывести из линейной алгебры. Впрочем, переход от чисто алгебраических рассмотрений (в первых параграфах главы) к физическим эффектам будет довольно плавным.

§ 1. Двумерные пространства со скалярным произведением

Пусть R — вещественное векторное пространство, в котором определено *скалярное произведение*, т. е. каждой паре векторов x, y из R поставлено в соответствие (вещественное) число (x, y), так что:

1) $(x, y) = (y, x)$,
2) $(\alpha x, y) = \alpha(x, y)$,
3) $(x + y, z) = (x, z) + (y, z)$ при всех x, y, z из R всех вещественных α. Заметьте, что мы не требуем выполнения условия 4 (стр. 145).

Длиной, или *модулем*, вектора называется корень квадратный из его скалярного квадрата. При этом, вообще говоря, ненулевой вектор может иметь нулевую и аже мнимую длину. (Если $(x, x) = -a^2 < 0$, то, по определению, $|x| = ai$, где $a > 0$, а $i = \sqrt{-1}$.)

Если в пространстве R выбран базис, то скалярное роизведение представляется симметрической билинейой формой

$$(x, y) = \sum_{i,k=1}^{n} g_{ik} x_i y_k$$

т координат векторов x и y. Соответствующая квадраичная форма в некотором (вообще говоря, другом) базисе приводится к «сумме квадратов»

$$(x, x) = x_1^2 + x_2^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$$

(§ 2 главы VI). При этом число p положительных и число q отрицательных квадратов являются инвариантами пространства R (закон инерции квадратичных форм, § 3 главы VI) и определяют его тип.

Так, для двумерного пространства (плоскости) R возможны такие значения p и q:

1) $p = 2$, $q = 0$;
1′) $p = 0$, $q = 2$;
2) $p = 1$, $q = 0$;
2′) $p = 0$, $q = 1$;
3) $p = 1$, $q = 1$.

В случае 1) в некотором (ортонормированном) базисе скалярный квадрат произвольного вектора $x = x_1e_1 + x_2e_2$ равен $x_1^2 + x_2^2$, и это пространство *евклидово*.

В случае 1′) $(x, x) = -x_1^2 - x_2^2$, и пространство несущественно отличается от евклидова.

В случае 2) (или 2′), что почти то же самое) квадратичная форма (x, x) содержит только один квадрат, и в некотором базисе $(x, x) = x_2^2$ (соответственно $-x_2^2$). Такая плоскость называется *полуевклидовой*.

Наконец, в случае 3) квадратичная форма (x, x) в некотором базисе приводится к разности квадратов $x_1^2 - x_2^2$; такая плоскость называется *псевдоевклидовой*.

§ 2. Полуевклидова плоскость

Пусть R — двумерное *векторное пространство* с полуевклидовой метрикой и e_1, e_2 — такой его базис, в котором скалярный квадрат (x, x) произвольного вектора $x = x_1e_1 + x_2e_2$ равен x_2^2. Тогда, в частности,

$$(e_1, e_1) = 0, \quad (e_2, e_2) = 1$$

и

$$(e_1 + e_2, e_1 + e_2) = 1 = (e_1, e_1) + 2(e_1, e_2) + (e_2, e_2) = 2(e_1, e_2) + 1,$$

откуда

$$(e_1, e_2) = 0$$

(вектор e_1 — нулевой длины, вектор e_2 — единичный, e_1 и e_2 ортогональны). Такой базис условимся называть *каноническим*.

Пусть $x = x_1e_1 + x_2e_2$ и $y = y_1e_1 + y_2e_2$ — произвольные векторы из R; тогда их скалярное произведение (x, y) равно

$$x_1y_1(e_1, e_1) + (x_1y_2 + x_2y_1)(e_1, e_2) + x_2y_2(e_2, e_2) = x_2y_2,$$

а модуль вектора x равен

$$|x| = \sqrt{x_2^2} = |x_2|.$$

Предположим, что e_1', e_2' — другой, тоже канонический базис в пространстве R и

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

— матрица перехода от первого базиса ко второму, т. е. что

$$e_1' = a_{11}e_1 + a_{21}e_2 \quad \text{и} \quad e_2' = a_{12}e_1 + a_{22}e_2.$$

Тогда

$$(e_1', e_1') = a_{21}^2 = (e_1, e_1) = 0,$$

откуда $a_{21} = 0$ и

$$(e_2', e_2') = a_{22}^2 = (e_2, e_2) = 1,$$

т. е. $a_{22} = \pm 1$. Таким образом, матрица перехода от одного канонического базиса к другому имеет вид

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & \pm 1 \end{bmatrix}. \tag{1}$$

Зафиксируем теперь какой-то канонический базис e_1, e_2 и *угол* между векторами $x = x_1e_1 + x_2e_2$ и $y = y_1e_1 + y_2e_2$, по определению, положим равным

$$\left|\frac{y_1}{y_2} - \frac{x_1}{x_2}\right|. \tag{2}$$

Так определенный угол, вообще говоря, не инвариантен относительно перехода к новому (даже каноническому) базису. Посмотрим, какие еще ограничения надо наложить на матрицу перехода для того, чтобы угол (2) не зависел от системы (канонических) координат. При переходе к новому (каноническому) базису с матрицей перехода (1) координаты векторов x и y соответственно преобразуются в

$$x_1' = a_{11}x_1 + a_{12}x_2, \quad x_2' = \pm x_2$$

и

$$y_1' = a_{11}y_1 + a_{12}y_2, \quad y_2' = \pm y_2,$$

причем знаки у x_2' и y_2' одинаковы. Тогда угол между векторами x и y в новом базисе в силу определения (2) должен быть равен

$$\left|\frac{y_1'}{y_2'} - \frac{x_1'}{x_2'}\right| = \left|\frac{a_{11}y_1 + a_{12}y_2}{\pm y_2} - \frac{a_{11}x_1 + a_{12}x_2}{\pm x_2}\right| = |a_{11}|\left|\frac{y_1}{y_2} - \frac{x_1}{x_2}\right|;$$

он будет иметь прежнее значение в том и только в том случае, если $a_{11} = \pm 1$. Поэтому, зафиксировав один какой-то канонический базис, мы дальше *будем допускать только такие базисы* e_1', e_2', *матрицы перехода к которым от базиса* e_1, e_2 *имеют вид*

$$\begin{bmatrix} \pm 1 & v \\ 0 & \pm 1 \end{bmatrix} \qquad (3)$$

(мы положили $a_{12} = v$).

Легко видеть, что если матрица перехода от базиса e_1, e_2 к базису e_1', e_2', имеет вид (3) и матрица перехода от базиса e_1', e_2' к базису e_1'', e_2'' — тоже вида (3), то и матрица перехода от базиса e_1, e_2 к базису e_1'', e_2'' будет такого же вида.

Обозначим через A_0 матрицу

$$A_0 = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix}.$$

Тогда, очевидно,

$$A_1 = \begin{bmatrix} -1 & v \\ 0 & 1 \end{bmatrix} = A_0\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & v \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} A_0,$$

$$A_3 = \begin{bmatrix} -1 & v \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} A_0 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Рассмотрим теперь двумерное *точечно-векторное пространство,* в котором *расстояние* между точками $X(x_1, x_2)$ и $Y(y_1, y_2)$ считается равным модулю вектора $\overline{XY} = (y_1 - x_1,\ y_2 - x_2)$ в полуевклидовой метрике, т. е. равным

$$|y_2 - x_2|.$$

Если представить точки полуевклидовой плоскости точками обычной (евклидовой) плоскости с теми же координатами, то $|y_2 - x_2|$ — это евклидова длина проекции вектора на ось ординат. (В частности, длина любого отрезка, параллельного e_1, будет равна нулю.) Две точки, расстояние между которыми равно нулю, назовем *параллельными*, подобно тому как в евклидовой геометрии параллельными называются прямые, угол между которыми равен нулю; тогда параллельные точки — это точки, принадлежащие одной прямой, параллельной вектору e_1.

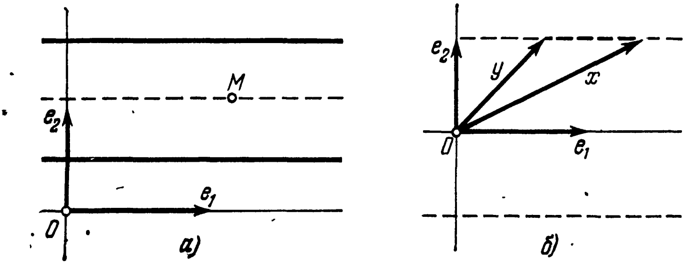


Рис. 17.

Окружность радиуса r с центром в данной точке $M(\alpha_1, \alpha_2)$, т. е. совокупность всех точек, отстоящих от точки M на одно и то же полуевклидово расстояние r, — это пара прямых, параллельных оси абсцисс и отстоящих от данной точки M на (евклидово) расстояние r (рис. 17, *а*). Центром такой окружности будет также любая точка прямой, проходящей через M и параллельной тем же прямым. Уравнение окружности радиуса r с центром в точке $M(\alpha_1, \alpha_2)$ имеет вид

$$(x_2 - \alpha_2)^2 = r^2.$$

В частности, уравнение «единичной окружности» (окружности радиуса единица) с центром в начале координат имеет вид

$$x_2^2 = 1.$$

Углом между прямыми называется угол между параллельными им векторами. Если $x=(\xi,\ 1)$ и $y=(\eta,\ 1)$ — два вектора единичной длины, то угол между ними равен

$$\left|\frac{\eta}{1}-\frac{\xi}{1}\right|=|\eta-\xi|;$$

он измеряется той «дугой», которую эти векторы высекают на «единичной окружности» (рис. 17, *б*). Заметим, что в полуевклидовой метрике *смежные углы равны* между собой. Действительно, угол между (единичными) векторами $(x,\ 1)$ и $(y,\ 1)$ равен $|y-x|$; смежный угол — между векторами $(-x,\ -1)$ и $(y,\ 1)$ равен

$$\left|\frac{y}{1}-\frac{-x}{-1}\right|=|y-x|.$$

Приведем несколько примеров теорем «элементарной полуевклидовой геометрии». Будем называть треугольником фигуру, образованную тремя точками, никакие две из которых не параллельны.

Теорема 1. *Бо́льшая сторона треугольника равна сумме двух других его сторон.*

Действительно, так как $AB=A'B'$, $AC=A'C'$, $BC=B'C'$ (рис. 18, *а*) и $A'B'=A'C'+B'C'$, то $AB=AC+BC$, или

$$c=a+b.$$

Эту теорему можно считать аналогом теоремы косинусов в евклидовой геометрии.

Теорема 2. *Бо́льший угол треугольника равен сумме двух других его углов.*

Для доказательства проведем прямую $CE \parallel BA$ (см. рис. 18, *б*). Тогда, очевидно, $\angle ACE=A$, а $\angle ECD=B$. Но $\angle ACE+\angle ECD=\angle ACD=C$ и, значит,

$$C=A+B.$$

Теорема 3. *Стороны треугольника пропорциональны противолежащим углам.*

Для доказательства проведем $CD \parallel e_1$ (см. рис. 18, *в*). Тогда $A=\frac{CD}{b}$ (где CD равно модулю разности абсцисс точек D и C — евклидовой длине отрезка CD), $B=\frac{CD}{a}$, значит, $A\cdot b=B\cdot a$, откуда

$$\frac{A}{a}=\frac{B}{b}.$$

(Эту теорему можно считать аналогом теоремы синусов евклидовой геометрии.)

Из трех последних теорем видна определенная «двойственность» теорем полуевклидовой плоскости, выражающаяся в равноправии сторон и углов треугольника. Если в формулировках этих теорем заменить слово «сторона» словом «угол», и наоборот, то, из теоремы 1 получится теорема 2, а из теоремы 2 — теорема 1; они двойственны друг другу. Теорема 3 двойственна сама себе.

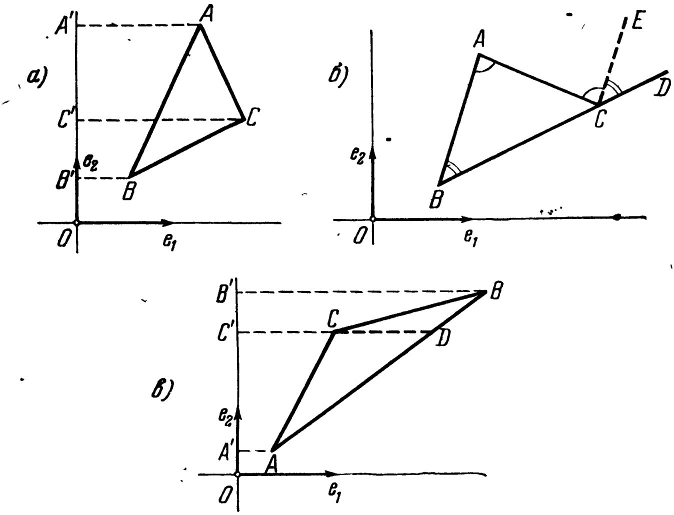


Рис. 18.

Такой двойственности нет на обычной, евклидовой плоскости, на которой имеются параллельные прямые (угол между которыми равен нулю), но нет «параллельных точек» (расстояние между которыми равно нулю). Эта «несправедливость» устранена в полуевклидовой геометрии, где, наряду с параллельными прямыми, имеются и «параллельные точки».

Задачи. Докажите, что в полуевклидовой плоскости

1. Отрезок, соединяющий середины боковых сторон (средняя линия) треугольника, параллелен основанию и равен его половине.

1′. Точка пересечения биссектрис двух углов треугольника параллельна противолежащей вершине, причем угол между этими биссектрисами равен половине третьего угла треугольника.

2. В равнобедренном треугольнике середина основания параллельна вершине.

2′. В равнобедренном треугольнике биссектриса угла при вершине параллельна основанию

3. Биссектриса треугольника делит противолежащую сторону на части, пропорциональные двум другим сторонам.

3′. Медиана треугольника делит соответствующий угол на части, пропорциональные двум другим углам.

4. Медианы треугольника пересекаются в одной точке и делятся в ней в отношении 2 : 1, считая от вершины.

4′ Биссектрисы треугольника пересекают противоположные стороны в трех точках, лежащих на одной прямой. Эта прямая делит каждый угол между стороной и биссектрисой противолежащего угла в отношении 2 ·1, считая от стороны.

(Легко видеть, что утверждения 1 и 1′, 2 и 2′, и т. д. двойственны друг другу. Им можно дать и двойственные одно другому доказательства.)

5. Сформулируйте свойства параллелограмма и докажите их.

5′. Дайте определение фигуры, двойственной параллелограмму («антипараллелограмм»). Сформулируйте и докажите ее свойства.

6. Сформулируйте и докажите признаки равенства треугольников.

7. Данный угол циркулем и линейкой разделите на n равных частей.

8. Дайте определение центрального угла, вписанного угла. Покажите, что вписанный угол равен половине центрального угла, опирающегося на ту же дугу, и что угол с вершиной вне круга измеряется полуразностью, а угол с вершиной внутри круга — полусуммой дуг, заключенных между его сторонами.

Интересна еще одна фигура, родственная окружности евклидовой плоскости и называемая *циклом*. Она определяется как геометрическое место точек, из которых данный отрезок виден под данным углом (в евклидовой геометрии, как известно, это определение снова приводит к окружности). Можно показать, что цикл полуевклидовой плоскости в евклидовой плоскости изображается **параболой**, однако подробное обсуждение этих вопросов завело бы нас слишком далеко. Вместо этого мы отсылаем читателя к книге И. М. Яглома «Принцип относительности Галилея и неевклидова геометрия».

§ 3. Псевдоевклидова плоскость

Пусть R — двумерное векторное пространство с псевдоевклидовой метрикой и e_1, e_2 — тот его базис, в котором скалярный квадрат произвольного вектора $x = x_1e_1 + x_2e_2$ равен $x_1^2 - x_2^2$. Тогда, в частности,

$$(e_1, e_1) = 1, \quad (e_2, e_2) = -1$$

и

$$(e_1 + e_2, e_1 + e_2) = 1 - 1 = 0 = (e_1, e_1) + 2(e_1, e_2) + (e_2, e_2) = 2(e_1, e_2),$$

откуда

$$(e_1, e_2) = 0$$

(т. е. вектор e_1 — единичный, вектор e_2 — «мнимо-единичный», e_1 и e_2 ортогональны). Такой базис будем называть *ортонормированным*. В ортонормированном базисе скалярное произведение векторов

$$x = x_1e_1 + x_2e_2 \text{ и } y = y_1e_1 + y_2e_2$$

равно

$$(x, y) = x_1y_1(e_1, e_1) + (x_1y_2 + x_2y_1)(e_1, e_2) + \\ + x_2y_2(e_2, e_2) = x_1y_1 - x_2y_2,$$

а модуль вектора x равен

$$|x| = \sqrt{x_1^2 - x_2^2}.$$

Рассмотрим теперь двумерное точечно-векторное пространство, в котором расстояние между точками $X(x_1, x_2)$ и $Y(y_1, y_2)$ считается равным модулю вектора $\overline{XY} = (y_1 - x_1, y_2 - x_2)$ в псевдоевклидовой метрике, т. е. равным

$$\sqrt{(y_1 - x_1)^2 - (y_2 - x_2)^2}.$$

Окружность радиуса r с центром в точке $M(\alpha_1, \alpha_2)$ — это совокупность всех точек, удаленных на одно и то же (псевдоевклидово) расстояние r от точки M. Значит, уравнение окружности радиуса r с центром в точке $M(\alpha_1, \alpha_2)$ будет иметь вид

$$(x_1 - \alpha_1)^2 - (x_2 - \alpha_2)^2 = r^2.$$

Таким образом, если точки псевдоевклидовой плоскости представить точками евклидовой плоскости с теми же координатами, то окружность представится гиперболой, если ее радиус $r \neq 0$, и парой пересекающихся прямых при $r = 0$ (рис. 19). Радиус такой окружности может быть *положительным*, *нулевым* или даже *«чисто мнимым»*. Так, уравнение окружности положительного радиуса $r = a$ с центром в начале координат будет иметь вид

$$x_1^2 - x_2^2 = a^2$$

(гипербола с горизонтальной вещественной осью). Окружность мнимого радиуса $r = ai$ (с тем же центром) имеет уравнение

$$x_2^2 - x_1^2 = a^2$$

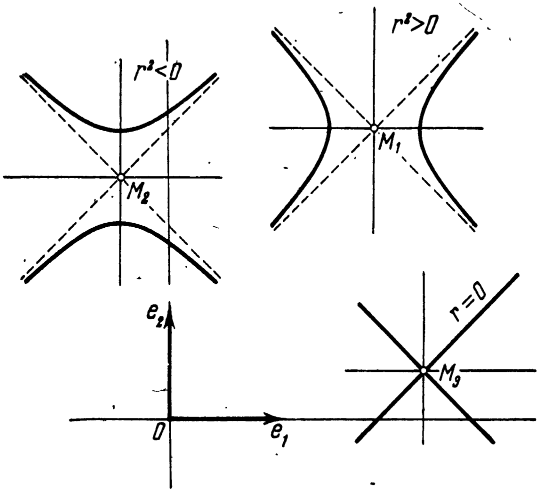


Рис. 19.

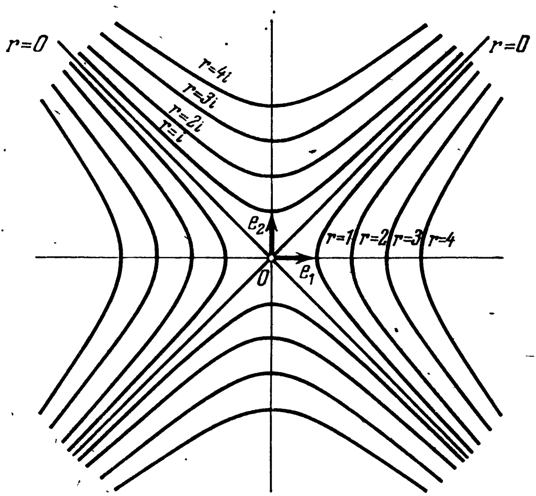


Рис. 20.

(гипербола с вертикальной вещественной осью). Эти два семейства окружностей разделяются окружностью нулевого радиуса

$$x_1^2 - x_2^2 = 0$$

(пара прямых — общие асимптоты обоих семейств гипербол; см. рис. 20).

Если векторы x и y ортогональны, т. е. если их скалярное произведение равно нулю:

$$(x, y) = x_1 y_1 - x_2 y_2 = 0,$$

то

$$\frac{x_1}{x_2} = \frac{y_2}{y_1}$$

— угловые коэффициенты этих векторов, рассматриваемых в евклидовой метрике, взаимно обратны и, значит, векторы, ортогональные в псевдоевклидовой метрике, при изображении на евклидовой плоскости по направлению симметричны друг другу относительно биссектрисы I—III координатных углов (см. рис. 21, на котором $e_1 \perp e_2$, $a_1 \perp a_2$, $b_1 \perp b_2$).

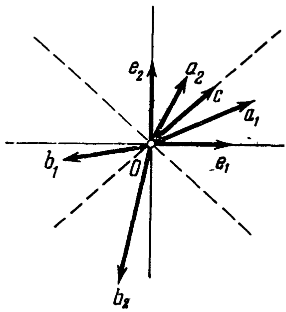


Рис. 21.

Каждый вектор, у которого $|x_1| = |x_2|$, *ортогонален самому себе* и имеет *нулевую длину*. Для векторов с вещественными длинами $|x_1| \geqslant |x_2|$, а для векторов мнимых длин $|x_1| < |x_2|$ (см. тот же рис. 21, на котором векторы e_1, a_1, b_1 имеют вещественные длины, векторы e_2, a_2, b_2 — мнимые длины, а вектор c ортогонален самому себе и $|c| = 0$).

Задачи. Докажите, что в псевдоевклидовой плоскости

1. Диагонали прямоугольника равны между собой.
2. Диагонали ромба взаимно перпендикулярны.
3. Сумма квадратов диагоналей параллелограмма равна сумме квадратов его сторон.
4. Средняя линия треугольника параллельна основанию и равна его половине.
5. Медианы треугольника пересекаются в одной точке и делятся в ней в отношении 2 : 1.
6. Серединные перпендикуляры к сторонам треугольника пересекаются в одной точке.

7. Высоты треугольника пересекаются в одной точке.

8. В равнобедренном треугольнике медиана является и биссектрисой и высотой.

Сообщим также без доказательства, что в псевдоевклидовой геометрии можно ввести понятие угла так, что для треугольника будут справедливы следующие соотношения:

$$a^2 = b^2 + c^2 - 2bc \operatorname{ch} A$$

(теорема косинусов — но косинус тут гиперболический!) и

$$\frac{a}{\operatorname{sh} A} = \frac{b}{\operatorname{sh} B}$$

(теорема синусов, где синусы — гиперболические).

За подробностями отсылаем читателя к той же книге И. М. Яглома (см. выше стр. 248).

§ 4. Псевдоортогональный оператор

Линейный оператор $\mathscr{A}$ *псевдоевклидова пространства называется* ***псевдоортогональным***, *если он сохраняет скалярное произведение*, т. е. если для всех $x, y \in R$

$$(\mathscr{A}x, \mathscr{A}y) = (x, y).$$

Пусть $\mathscr{A}$ — псевдоортогональный оператор в псевдоевклидовой плоскости R и

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

— его матрица в ортонормированном базисе e_1, e_2. Мы имеем

$$\mathscr{A}e_1 = e_{11}e_1 + a_{21}e_2,$$

$$\mathscr{A}e_2 = a_{12}e_1 + a_{22}e_2.$$

По определению,

$$(\mathscr{A}e_1, \mathscr{A}e_1) = (e_1, e_1) = 1,$$

$$(\mathscr{A}e_2, \mathscr{A}e_2) = (e_2, e_2) = -1$$

и

$$(\mathscr{A}e_1, \mathscr{A}e_2) = (e_1, e_2) = 0,$$

т. е.

$$\begin{aligned} a_{11}^2 - a_{21}^2 &= 1, \\ a_{12}^2 - a_{22}^2 &= -1, \end{aligned} \tag{4a}$$

и

$$a_{11} \cdot a_{12} - a_{21} \cdot a_{22} = 0. \tag{4б}$$

Из равенств (4а) видно, что $a_{11} \neq 0$ и $a_{22} \neq 0$. Из равенства (4б) следует, что

$$\frac{a_{21}}{a_{11}} = \frac{a_{12}}{a_{22}}. \tag{5}$$

Обозначив равные отношения (5) через β, получим

$$\begin{aligned} a_{21} &= \beta a_{11}, \\ a_{12} &= \beta a_{22}. \end{aligned} \tag{6}$$

Подставляя эти значения в равенства (4а), найдем, что

$$a_{11}^2 - \beta^2 a_{11}^2 = 1, \quad \text{откуда} \quad a_{11} = \frac{1}{\pm\sqrt{1-\beta^2}},$$

и

$$\beta^2 a_{22}^2 - a_{22}^2 = -1, \quad \text{откуда} \quad a_{22} = \frac{1}{\pm\sqrt{1-\beta^2}}.$$

Таким образом, матрица оператора $\mathcal{A}$ имеет вид

$$A = \begin{bmatrix} \frac{1}{\pm\sqrt{1-\beta^2}} & \frac{\beta}{\pm\sqrt{1-\beta^2}} \\ \frac{\beta}{\pm\sqrt{1-\beta^2}} & \frac{1}{\pm\sqrt{1-\beta^2}} \end{bmatrix}, \tag{7}$$

причем, как видно из равенств (6), оба элемента первого столбца, так же как и оба элемента второго столбца, берутся с одним и тем же знаком. Матрицу такого вида будем называть *псевдоортогональной*.

Если обозначить через A_0 матрицу

$$A_0 = \begin{bmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} \\ \frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{bmatrix},$$

то, как легко видеть,

$$A_1 = \begin{bmatrix} -\frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} \\ -\frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{bmatrix} = A_0 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} \\ \frac{\beta}{\sqrt{1-\beta^2}} & -\frac{1}{\sqrt{1-\beta^2}} \end{bmatrix} = A_0 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} -\frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} \\ -\frac{\beta}{\sqrt{1-\beta^2}} & -\frac{1}{\sqrt{1-\beta^2}} \end{bmatrix} = A_0 \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

(преобразования $\mathscr{A}_1$ и $\mathscr{A}_2$ отличаются от $\mathscr{A}_0$ осевой, а $\mathscr{A}_3$ — центральной симметрией).

Определители $|A_0| = |A_3| = 1$, $|A_1| = |A_2| = -1$.

Заметим, что, поскольку $\frac{1}{\sqrt{1-\beta^2}} \geqslant 1$, то найдется такое φ, что $\frac{1}{\sqrt{1-\beta^2}} = \operatorname{ch}\varphi$, $\frac{\beta}{\sqrt{1-\beta^2}} = \operatorname{sh}\varphi$, и тогда

$$A_0 = \begin{bmatrix} \operatorname{ch}\varphi & \operatorname{sh}\varphi \\ \operatorname{sh}\varphi & \operatorname{ch}\varphi \end{bmatrix}.$$

Это преобразование называется *гиперболическим поворотом.*

Пусть в псевдоевклидовой плоскости R имеются два ортонормированных базиса, e_1, e_2 и e_1', e_2' и

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

— матрица перехода от первого ко второму. Рассмотрим линейный оператор $\mathscr{A}$ с матрицей A в базисе e_1, e_2 и покажем, что он — псевдоортогональный. Действительно, по условию,

$$\mathscr{A}e_1 = a_{11}e_1 + a_{21}e_2 = e_1' \text{ и } \mathscr{A}e_2 = a_{12}e_1 + a_{22}e_2 = e_2'.$$

Если $x = x_1e_1 + x_2e_2$ и $y = y_1e_1 + y_2e_2$ — произвольные векторы из R, то

$$\mathscr{A}x = x_1\mathscr{A}e_1 + x_2\mathscr{A}e_2 = x_1e_1' + x_2e_2'$$

и

$$\mathscr{A}y = y_1\mathscr{A}e_1 + y_2\mathscr{A}e_2 = y_1e_1' + y_2e_2'.$$

А так как оба базиса e_1, e_2 и e_1', e_2' — ортонормированные, то скалярное произведение

$$(\mathcal{A}x, \mathcal{A}y) = x_1y_1 - x_2y_2 = (x, y).$$

Значит, оператор $\mathcal{A}$ — псевдоортогональный, и его матрица имеет вид (7).

§ 5. Пространство событий. Принцип относительности Галилея

Предположим, что точка M движется вдоль прямой линии l, на которой установлена *система отсчета* S. Это значит, что на этой прямой расположена шкала с соответствующими делениями и в каждой точке прямой имеются синхронизированные между собой часы.

Пусть в момент времени t координата точки M равна x. Это обстоятельство, или, как мы будем говорить, «событие», можно отметить на некоторой (двумерной) плоскости P точкой с координатами (x, t). Плоскость P называется *пространством событий*.

С течением времени координаты точки в пространстве событий меняются, даже если точка M не меняет своего положения на прямой l — за счет изменения времени t. Таким образом, существование точки в пространстве и времени будет отмечено некоторой *линией* в плоскости P. Прямой эта линия будет в том и только в том случае, если точка M движется по прямой l с постоянной скоростью, и тогда ее положение в плоскости P будет определяться уравнением

$$x = ut + b,$$

где $b = x(0)$ — положение точки в момент $t = 0$. Если точка M неподвижна на прямой l («движется с нулевой скоростью»), то соответствующая ей в плоскости P прямая параллельна оси t.

Предположим, что вдоль прямой l равномерно со скоростью v движется другая система отсчета, S', причем в начальный момент времени начала координат обеих систем совпадают: $x = x' = 0$ при $t = 0$. Тогда координата x точки M в системе S и координата ее x' в системе S' будут связаны соотношением

$$x = x' + vt,$$

При этом считается, что время t в системе S и время t' в системе S' одно и то же: для одного и того же события $t = t'$.

Преобразования

$$\begin{aligned} x &= x' + vt, \\ t &= t', \end{aligned} \tag{8}$$

или, что то же самое,

$$\begin{aligned} x' &= x - vt, \\ t' &= t \end{aligned}$$

называются *преобразованиями Галилея*. Из них дифференцированием по t получаем

$$\frac{dx}{dt} = \frac{dx'}{dt} + v,$$

или

$$u = u' + v, \tag{9}$$

где u — скорость точки в системе S, а u' — скорость ее в системе S'. Это — закон сложения скоростей в классической механике: скорость u точки в старой системе отсчета равна ее скорости u' в новой системе, сложенной с «переносной» скоростью v (скоростью движения новой системы отсчета относительно старой). Дифференцируя по t еще раз, получаем

$$\frac{d^2x}{dt^2} = \frac{d^2x'}{dt^2}.$$

Таким образом, *ускорения точки M в системе S и в системе S' одинаковы,* откуда делается вывод, что одинаковые силы вызывают в обеих системах одинаковые следствия (описываемые вторым законом Ньютона: вызванное силой F ускорение прямо пропорционально этой силе). Другими словами это выражают, говоря, что *законы механики инварианты относительно преобразований Галилея* (принцип относительности Галилея).

Вернемся к формулам (8). Они показывают, что при переходе от системы S к системе S' координаты точек пространства событий подвергаются линейному преобразованию с матрицей

$$\begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix}. \tag{10}$$

Это обстоятельство наводит на мысль ввести в пространстве событий *полуевклидову метрику.* Тогда расстояние между событиями $A(x_1, t_1)$ и $B(x_2, t_2)$ будет иметь определенный физический смысл: оно будет равно $|t_2 - t_1|$ — *временно́му интервалу, протекшему между событиями A и B.*

Далее, так как переход от одной системы координат к другой задается матрицей (10), то инвариантным окажется и введенное в § 2 понятие угла. Чтобы выяснить его физический смысл, рассмотрим две равномерно движущиеся по прямой l точки M_1 и M_2. Скорости их обозначим соответственно через u_1 и u_2. В плоскости P их движения определяются прямыми m_1 и m_2. Пусть $A_0(x_0, t_0)$ — точка пересечения этих прямых (это значит, что при $t = t_0$ обе точки, M_1 и M_2, находились в одном и том же месте прямой l — имели абсциссу x_0). Предположим, что при $t = t_1$ точка M_1 имеет абсциссу x_1, а при $t = t_2$ точка M_2 — абсциссу x_2. Тогда угол между прямыми m_1 и m_2 (в полуевклидовой метрике) равен углу между векторами $\overline{A_0A_1}$ и $\overline{A_0A_2}$, где $A_1(x_1, t_1)$, $A_2(x_2, t_2)$ (рис. 22), и значит, он равен

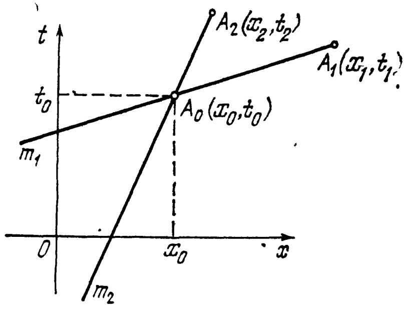


Рис. 22.

$$\left| \frac{x_2 - x_0}{t_2 - t_0} - \frac{x_1 - x_0}{t_1 - t_0} \right| = |u_2 - u_1|$$

— *относительной скорости* движения этих точек.

При такой интерпретации расстояния и угла теоремы 1, 2 и 3 на стр. 246 получают определенный физический смысл, установить который предоставляется читателю.

§ 6. Принцип относительности Эйнштейна

Из закона сложения скоростей (9) естественно сделать следующий вывод: если система отсчета S' равномерно движется относительно S со скоростью v и если свет в системе S распространяется со скоростью c, то в системе S' его скорость должна быть равна $c-v$ в направлении движения системы S' и $c+v$ — в противоположном направлении. Однако в 1881 г. американским физиком А. Майкельсоном было экспериментально установлено, что *на движущейся Земле солнечный свет распространяется с одинаковой скоростью во всех направлениях.*

После попыток многих ученых как-то согласовать результаты опытов Майкельсона с теорией, в 1905 г. была опубликована фундаментальная работа А. Эйнштейна, в которой излагалась новая теория пространства и времени — так называемая *специальная теория относительности.* Мы рассмотрим здесь только самые основные, простейшие понятия этой теории.

В основу теории Эйнштейна был положен *закон постоянства скорости света* во всех инерциальных*) системах отсчета.

Таким образом, *принцип относительности Галилея состоит в невозможности установить равномерное движение одной механической системы относительно другой с помощью каких-либо механических экспериментов внутри этой системы. Принцип относительности Эйнштейна утверждает, что это невозможно сделать, исходя не только из механических, но также и из каких-либо оптических явлений* (связанных, как известно, с электромагнетизмом).

Но приняв закон постоянства скорости света, Эйнштейн вынужден был *отказаться от предположения о существовании абсолютного времени,* годного для измерения временных интервалов сразу во всех инерциальных системах отсчета.

То, что эта относительность времени с необходимостью вытекает из закона постоянства скорости света,

*) В физике *инерциальной* называют такую систему отсчета, в которой тело без действия на него внешних сил движется равномерно и прямолинейно.

можно видеть на следующем простом примере *). Представим себе очень большой по линейным размерам поезд, скорость которого сравнима со скоростью света («поезд Эйнштейна»). Пусть в этом поезде у окна находится наблюдатель, который в некоторый момент времени зажигает фонарик, испускающий луч света в потолок. На потолке имеется зеркало, отразившись от которого, луч возвращается к наблюдателю. Путь луча света с точки зрения этого наблюдателя — дважды проходимый отрезок AB (рис. 23, *а*). Для наблюдателя же, находящегося вне поезда, путь луча света представится в виде ломаной линии, состоящей из боковых сторон равнобедренного треугольника A_1BA_2, высота которого равна AB (рис. 23, *б*). Следовательно, путь, проходимый светом, с точки зрения наблюдателя вне поезда, больше, чем для пассажира поезда. А так как скорость

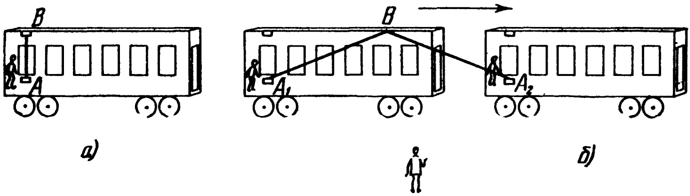


Рис. 23.

света постоянна, то время, которое потребуется свету на этот путь по часам наблюдателя вне поезда, будет больше, чем для пассажира поезда: часы внутри поезда отстают по сравнению с часами на станции.

Закон постоянства скорости света делает относительным и понятие одновременности, что хорошо видно на другом примере. Предположим, что в центре вагона того же поезда Эйнштейна находится наблюдатель, который в некоторый момент времени зажигает фонарик. В дверях вагона имеется механизм, благодаря которому двери открываются, как только до них доходит свет. Наблюдатель в центре вагона увидит, что задняя и передняя двери открываются одновременно. С точки же зрения наблюдателя вне поезда передняя дверь вагона уходит

*) Этот и следующий примеры заимствованы из брошюры Л. Д. Ландау и Ю. Б. Румера [18].

от светового луча, тогда как задняя идет к нему навстречу. Ввиду постоянства скорости света, с точки зрения наблюдателя вне поезда свет достигает задней двери вагона раньше, чем передней, и она откроется раньше.

Более того, даже последовательность событий может быть разной и для этих двух наблюдателей. Так, если (например, из-за неисправности механизма дверей) задняя дверь откроется несколько позже, чем на нее попадет свет, то, если эта разница во времени достаточно мала, наблюдатель вне поезда все-таки увидит заднюю дверь открывающейся раньше, чем передняя, хотя для наблюдателя в центре вагона последовательность этих событий будет обратной.

§ 7. Преобразования Лоренца

Итак, мы вынуждены отказаться от предположения, что время — одно и то же во всех равномерно движущихся друг относительно друга системах отсчета. Мы уже не можем считать, что для одного и того же события $t' = t$. Как же связаны между собой координаты x, t точки в системе S и координаты x', t' ее в системе S', движущейся относительно S равномерно со скоростью v? В классической механике эта связь линейна (преобразования Галилея). *Мы сохраним это предположение о линейной зависимости x', t' от x, t* — тогда переходу от S к S' будет отвечать переход к новому базису в пространстве событий. Какова же метрика этого пространства?

Пусть в некоторый момент времени (начальный для обеих систем S и S') их начала координат совпадают: $x = x' = 0$ при $t = t' = 0$. Предположим, далее, что при $t = t' = 0$ из общего начала координат обеих систем пущен световой сигнал, принятый в системе S в точке x в момент t, а в системе S' — в точке x' в момент времени t'. Ввиду постоянства скорости света c

$$\left|\frac{x}{t}\right| = \left|\frac{x'}{t'}\right| = c,$$

откуда $x^2 - c^2t^2 = 0$ и $x'^2 - c^2t'_2 = 0$. Таким образом, *если выражение*

$$x^2 - c^2t^2 \qquad (11)$$

равно нулю в одной инерциальной системе отсчета, то оно обращается в нуль и во всех остальных. Мы сделаем теперь еще одно дополнительное предположение — о том, что *выражение* (11) *вообще является инвариантом, т. е. что оно одинаково во всех инерциальных системах отсчета.*

Положив $x = x_1$ и $ct = x_2$ (и соответственно $x' = x_1'$ и $ct' = x_2'$), мы можем наше пространство событий рассматривать как *псевдоевклидову плоскость*, в которой выражение (11), равное

$$x_1^2 - x_2^2,$$

является квадратом расстояния точки (x_1, x_2) от начала координат, или, что то же самое, квадратом длины соответствующего вектора. Но базис, в котором квадрат длины вектора имеет такой вид, является *ортонормированным* (см. начало § 3). Ортонормированным будет по той же причине и соответствующий базис системы S', а значит, матрица A перехода от базиса системы S к базису S' псевдоортогональна:

$$A = \begin{bmatrix} \dfrac{1}{\pm\sqrt{1-\beta^2}} & \dfrac{\beta}{\pm\sqrt{1-\beta^2}} \\ \dfrac{\beta}{\pm\sqrt{1-\beta^2}} & \dfrac{1}{\pm\sqrt{1-\beta^2}} \end{bmatrix}$$

(причем в каждом из столбцов стоит какой-то один знак).

Следовательно,

$$e_1' = \frac{e_1 + \beta e_2}{\pm\sqrt{1-\beta^2}}, \quad e_2' = \frac{\beta e_1 + e_2}{\pm\sqrt{1-\beta^2}}.$$

Рассмотрим сначала случай, когда оба знаменателя положительны, и матрица A имеет вид

$$A_0 = \begin{bmatrix} \dfrac{1}{\sqrt{1-\beta^2}} & \dfrac{\beta}{\sqrt{1-\beta^2}} \\ \dfrac{\beta}{\sqrt{1-\beta^2}} & \dfrac{1}{\sqrt{1-\beta^2}} \end{bmatrix}.$$

Тогда координаты x_1, x_2 и x'_1, x'_2 связаны соотношениями

$$x_1 = \frac{x'_1 + \beta x'_2}{\sqrt{1-\beta^2}}, \quad x_2 = \frac{\beta x'_1 + x'_2}{\sqrt{1-\beta^2}},$$

или, в старых обозначениях,

$$x = \frac{x' + \beta c t'}{\sqrt{1-\beta^2}}, \quad t = \frac{\frac{\beta}{c} x' + t'}{\sqrt{1-\beta^2}}. \tag{12}$$

Выражая отсюда x' и t' через x и t, получим формулы

$$x' = \frac{x - \beta c t}{\sqrt{1-\beta^2}}, \quad t' = \frac{-\frac{\beta}{c} x + t}{\sqrt{1-\beta^2}}. \tag{13}$$

Каков физический смысл параметра β? Предположим, что в системе S' покоится точка M; пусть, например, это будет начало координат $x' = 0$. По первой из формул (13) для этой точки имеем

$$x - \beta c t = 0, \text{ или } \frac{x}{t} = \beta c.$$

Но $\frac{x}{t}$ есть скорость точки M в системе S, равная, очевидно, скорости v системы S' относительно S. Следовательно,

$$v = \beta c, \text{ и } \beta = \frac{v}{c}.$$

Подставив это значение β в формулы (12) и (13), получим

$$x = \frac{x' + v t'}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t = \frac{\frac{v}{c^2} x' + t'}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{14}$$

и

$$x' = \frac{x - v t}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t' = \frac{-\frac{v}{c^2} x + t}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{15}$$

Преобразования (14) и (15) называются *преобразованиями Лоренца.* Заметим, что формулы (15) получаются из формул (14) простым изменением знака у v.

Мы предполагали, что в матрице перехода от базиса e_1, e_2 к базису e_1', e_2' все знаменатели положительны. Покажем, как исключить остальные случаи. Если бы во втором столбце матрицы перехода стояли знаки минус (а в первом какие угодно), то мы получили бы формулы

$$x = \frac{\pm x' - vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t = \frac{\pm \frac{v}{c^2} x' - t'}{\sqrt{1 - \frac{v^2}{c^2}}},$$

и, например, при $x' = 0$, т. е. в начале координат системы S', увеличению t' соответствовало бы уменьшение t, что невозможно, так как при этом последовательность **всех** событий в точке x' системы S' была бы обратной последовательности тех же событий в системе S. Если же знаки минус стоят в первом столбце матрицы перехода (а во втором столбце стоят знаки плюс), то получаются формулы

$$x = \frac{-x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t = \frac{-\frac{v}{c^2} x' + t'}{\sqrt{1 - \frac{v^2}{c^2}}},$$

от которых к формулам (14) можно перейти, изменив знак у x', т. е. изменив на противоположное направление оси Ox'.

Таким образом, мы можем ограничиться исследованием преобразований Лоренца (14) и (15). Формулы Лоренца имеют смысл лишь при $\left|\frac{v}{c}\right| < 1$, откуда следует, что $|v| < c$, т. е. *что движение со скоростью, превышающей скорость света, невозможно.*

Если v мало по сравнению с c, то $\sqrt{1 - \frac{v^2}{c^2}} \approx 1$,

а тогда

$$x' \approx x - vt, \; t' \approx t.$$

Таким образом, *при малом v* (по сравнению с c) *преобразования Лоренца переходят в преобразования Галилея классической механики.*

Пусть Ox и Ot — координатные оси пространства событий системы S, Ox' и Ot' — оси системы S' (рис. 24). Как мы знаем, оси Ox' и Ot', если изображать их на евклидовой плоскости, симметричны, друг другу относительно биссектрис MM' и NN' координатных углов первой системы. Ось Ot' можно рассматривать как график движения начала координат системы S' относительно S: для всех ее точек $x'=0$. Наоборот, ось Ot — это график движения начала координат системы S относительно S'. Тангенс угла наклона оси Ot' к Ox по абсолютной величине равен

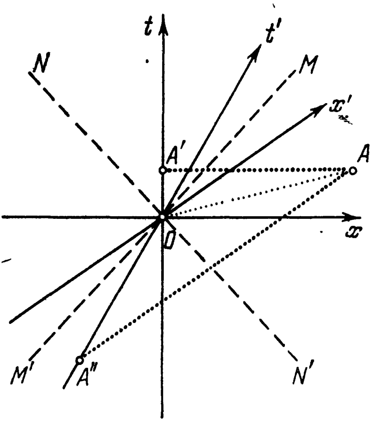


Рис. 24.

$$\left|\frac{ct}{x}\right| = \frac{c}{|v|},$$

где $\frac{x}{t} = v$ — скорость движения системы S' относительно S. А так как $|v| < c$, то тангенс этот по модулю больше единицы, и значит, *все временные оси Ot лежат внутри угла MON, а следовательно, все пространственные оси Ox — внутри угла MON'.*

Для прямых MM' и NN' имеем $\left|\frac{x}{ct}\right| = 1$, т. е. $|v| = c$; во всех системах отсчета это — график движения со скоростью света.

§ 8. Некоторые следствия из формул Лоренца

1. *Правило сложения скоростей.* Из равенства (15) получаем

$$\frac{dx'}{dt'} = \frac{dx'}{dt} : \frac{dt'}{dt} =$$

$$= \frac{\frac{dx}{dt} - v}{\sqrt{1 - \frac{v^2}{c^2}}} : \frac{-\frac{v}{c^2}\frac{dx}{dt} + 1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\frac{dx}{dt} - v}{-\frac{v}{c^2}\frac{dx}{dt} + 1} = \frac{u - v}{-\frac{uv}{c^2} + 1},$$

или

$$u' = \frac{u - v}{-\frac{uv}{c^2} + 1}, \tag{16}$$

откуда

$$u = \frac{u' + v}{\frac{u'v}{c^2} + 1}.$$

Это — новая формула сложения скоростей. Если u и v малы по сравнению с c, то $u' \approx u - v$. Если $u = c$, то из формулы (16) получаем

$$u' = \frac{c - v}{-\frac{v}{c} + 1} = c,$$

и обратно, если $u' = c$, то и

$$u = \frac{c + v}{\frac{v}{c} + 1} = c$$

(и, значит, из формул Лоренца вытекает закон постоянства скорости света).

2. *Относительность одновременности.* Предположим, что события A и B в системе S происходят *в один и тот же момент времени* t в точках с разными абсциссами x_1 и x_2. Тогда в системе S' по второй из формул (15) эти события будут происходить в моменты времени

$$t_1' = \frac{-\frac{v}{c^2}x_1 + t}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{и} \quad t_2' = \frac{-\frac{v}{c^2}x_2 + t}{\sqrt{1 - \frac{v^2}{c^2}}},$$

откуда

$$t_2' - t_1' = \frac{\frac{v}{c^2}(x_1 - x_2)}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

т. е. *события, одновременные в одной системе отсчета, не будут одновременными в другой.* При этом разность $t_2'-t_1'$ может быть как положительной, так и отрицательной, в зависимости от знака разности x_1-x_2. (Это хорошо видно на чертеже: если события A и B одновременны в системе S, то отрезок AB должен быть параллелен оси Ox, а если они одновременны в системе S', то он должен быть параллелен оси Ox'.)

Больше того, даже последовательность событий может быть в системах S и S' не одинаковой. Так, на рис. 24 события A и A' одновременны в системе S ($AA' \parallel Ox$), причем A' происходит, очевидно, позднее O, а следовательно, и A в системе S происходит после O. В системе S' одновременны события A и A'' ($AA'' \parallel Ox'$), и значит, событие A (вместе с A'') предшествует O.

Здесь, естественно, возникает такой вопрос: не может ли случиться так, что, например, событие O, в системе S послужившее причиной события A, в системе S' окажется происходящим после A, что противоречило бы принципу причинности. Покажем, что на самом деле этого быть не может.

Точки, отвечающие событиям, которые в системе S происходят после события O,— это все те и только те точки, которые лежат выше оси Ox; точки, отвечающие событиям, которые происходят после события O в системе S',— это точки, лежащие выше оси Ox'. Так как все пространственные оси проходят внутри угла MON' (см. конец § 7), и, очевидно, каждая такая прямая служит пространственной осью некоторой системы отсчета, то *пересечение всех полуплоскостей, лежащих выше какой-либо из пространственных осей,— это угол MON, заполненный всеми теми и только теми событиями, которые следуют за O во всех системах отсчета* (его можно назвать *«областью будущего»*). *Аналогично, угол $M'ON'$ представляет собой множество всех тех событий, которые во всех системах отсчета происходят до события O* (*«область прошедшего»*).

Точки же, лежащие в углах MON' и NOM', отвечают событиям, которые в одних системах отсчета предшествуют O, а других — следуют за O. Однако ни одно из этих событий не может иметь своей причиной событие O. Действительно, если событие O послужило причиной события $A(x, t)$ (см. тот же рис. 24), то какое-то возмущение должно успеть дойти от O до A. Однако это невозможно, так как длина вектора $\overline{OA}$ вещественна и, значит, для него во всех системах отсчета

$$x^2 - c^2t^2 > 0, \quad \text{т. е.} \quad x_2 > c^2t^2,$$

откуда

$$\left|\frac{x}{t}\right| > c,$$

т. е. скорость $u = \frac{x}{t}$ такого возмущения должна была бы быть больше скорости света, что невозможно.

Аналогично показывается, что для любых двух событий A и B закон причинности не нарушается: если A может служить причиной B, т. е. если существует сигнал, распространяющийся (в данной системе отсчета S) от A к B со скоростью $v < c$, то A предшествует B во всех инерциальных системах отсчета.

3. *Сокращение длин.* Пусть в системе S покоится стержень длины l; координаты концов его обозначим x_1 и x_2, где $x_1 < x_2$; тогда

$$l = x_2 - x_1.$$

Для того чтобы измерить длину l' стержня в системе S', надо отметить координаты его концов в какой-то (один и тот же!) момент времени t'. Если эти координаты x_1' и x_2', то по первой из формул (14)

$$x_1 = \frac{x_1' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad x_2 = \frac{x_2' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}},$$

откуда имеем

$$l = x_2 - x_1 = \frac{x_2' - x_1'}{\sqrt{1 - \frac{v^2}{c^2}}},$$

или, так как длина l' стержня в системе S' равна $x_2' - x_1'$,

$$l' = l\sqrt{1 - \frac{v^2}{c^2}} < l.$$

Таким образом, *длина l' стержня в той системе отсчета, относительно которой этот стержень движется, меньше, чем длина его l в той системе отсчета, относительно которой он находится в покое.*

Поясним этот результат на чертеже. Пусть точка A лежит на пересечении гиперболы

$$x^2 - c^2t^2 = l^2$$

с осью Ox (рис. 25); тогда в системе S ее расстояние от начала координат равно l. Если $AA' \| Ot$, то точки A и A' в системе S находятся на одном и том же расстоянии l от начала координат. (Это — одна и та же точка, покоящаяся в системе S, в разные моменты времени.) Но в системе S' расстояние точки A' от начала координат равно OA'; оно меньше OB, равного l.

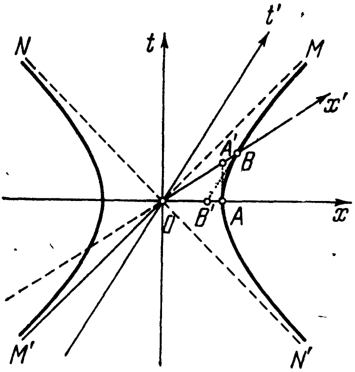


Рис. 25.

Наоборот, точка B, лежащая на пересечении гиперболы $x^2 - c^2t^2 = l^2$ с осью Ox', в системе S' находится на расстоянии l от начала координат. Если $BB' \| Ot'$, то точка B' в системе S' находится от начала координат на том же расстоянии l; однако в системе S расстояние точки B' от точки O равно $OB' < OA = l$ — *релятивистское*, т. е. связанное с теорией относительности *сокращение длин взаимно.*

Если v мало́ по сравнению со скоростью света, то указанное сокращение длин в движущейся системе отсчета настолько мало́, что практически обнаружить его невозможно. $\left(\text{Поскольку } l' = l\sqrt{1 - \frac{v^2}{c^2}} \approx l\left(1 - \frac{v^2}{2c^2}\right),\right.$ то разность $l - l' \approx \frac{lv^2}{2c^2} = \frac{l}{2}\left(\frac{v}{c}\right)^2$ — второго порядка относительно $\left.\frac{v}{c}.\right)$ Так, с космической ракеты (при скорости 12 *км/сек*) диаметр Земли (12 000 *км*) покажется укороченным всего на 1 *см*.

4. *Замедление времени.* Пусть в системе S на неподвижных в ней часах протекает время T от t_1 до t_2:

$$T = t_2 - t_1.$$

Найдем значения t'_1, соответствующее t_1, и t'_2, соответствующее t_2, *в одной и той же точке* с абсциссой x' в системе S'. По второй из формул (14)

$$t_1 = \frac{\frac{v}{c^2}x' + t'_1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t_2 = \frac{\frac{v}{c^2}x' + t'_2}{\sqrt{1 - \frac{v^2}{c^2}}},$$

откуда

$$T = t_2 - t_1 = \frac{t'_2 - t'_1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{T'}{\sqrt{1 - \frac{v^2}{c^2}}},$$

где $T' = t'_2 - t'_1$ — отрезок времени, протекающий в системе S', когда в системе S протекает время T от t_1 до t_2; значит,

$$T' = T\sqrt{1 - \frac{v^2}{c^2}} < T.$$

Таким образом, *временно́й интервал T' между двумя событиями, происходящими в одной и той же точке системы S', которая движется относительно S, меньше, чем время T, протекшее между теми же событиями в системе S.*

Поясним этот результат на чертеже. Рассмотрим гиперболу

$$x_2 - c^2t^2 = -c^2T^2$$

(рис. 26), и пусть точка A лежит на пересечении этой гиперболы с осью Ot; тогда ее временно́е расстояние от точки O, т. е. время, протекшее от события O до A, в системе S равно T. Если $AA' \parallel Ox$, то события A и A' одновременны в системе S. Но в системе S' время, протекшее от O до A', равно OA'; оно меньше OB, равного T.

Наоборот, точка B в системе S' по времени удалена от точки O на интервал T. Если $BB' \parallel Ox'$, то события B

и B' одновременны в системе S'; однако в системе S временнóе расстояние точки B' от O равно OB' — оно меньше OA, равного T — *лоренцово замедление времени взаимно.*

Если скорость v мала по сравнению со скоростью света, то это замедление времени в движущейся системе отсчета составляет всего $\frac{T}{2}\left(\frac{v}{c}\right)^2$ — оно второго порядка относительно $\frac{v}{c}$, и обнаружить его практически невозможно. Так, земные сутки покажутся космонавту сократившимися меньше чем на $\frac{1}{10\,000}$ *сек.*

5. *Увеличение массы движущегося тела.* Мы не будем рассматривать дальнейших выводов теории относительности; упомянем еще только один феномен — увеличение массы движущегося тела.

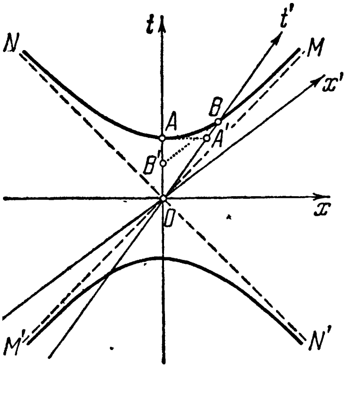


Рис. 26.

Если на тело действует постоянная сила, то скорость движения его в обычных условиях возрастает пропорционально времени действия силы. Однако, ввиду существования предельной скорости, эта пропорциональность не может сохраняться и при больших скоростях. При скоростях, сравнимых со скоростью света, дальнейшее нарастание скорости замедляется — тело как бы оказывает большее сопротивление действующей на него силе. Можно сказать, что *масса тела увеличивается.* При этом оказывается, что

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}},$$

где m — масса движущегося тела, v — скорость его движения и m_0 — масса покоя, т. е. масса тела в той системе отсчета, относительно которой оно покоится. Так, в современных ускорителях электрон разгоняется до скорости, отличающейся от скорости света всего на десятки метров в секунду, при этом его масса увеличивается в тысячи раз. (Действительно, если скорость электрона v отличается от скорости света, скажем на 30 *м/сек*, то масса этого электрона

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{m_0}{\sqrt{\frac{c+v}{c}\frac{c-v}{c}}} \approx \frac{m_0}{\sqrt{2\frac{c-v}{c}}} =$$

$$= \frac{m_0}{\sqrt{2\cdot\frac{30}{30\cdot 10^7}}} = 1000\sqrt{5}m_0 > 2000m_0$$

— увеличивается более чем в 2000 раз.)

Мы рассмотрели движение точки по прямой линии. В общем случае, когда одна пространственная система отсчета движется относительно другой равномерно и прямолинейно, направление этого движения можно принять за направление оси Ox, и тогда в классической механике

$$x' = x - vt, \qquad y' = y, \qquad z' = z, \qquad t' = t,$$

а в теории относительности

$$x' = \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}}, \quad y' = y, \quad z' = z, \quad t' = \frac{-\frac{v}{c^2}x + t}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Пространство событий в этом случае четырехмерно. Сокращение длин (только в направлении движения) и замедление времени в движущейся системе отсчета происходят в том же отношении $\sqrt{1-\frac{v^2}{c^2}}$.

ГЛАВА X

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ГРУПП

§ 1. Примеры групп. Определение группы

Рассмотрим *множество всех целых чисел.* При сложении двух целых чисел получается снова целое число. Если одно из слагаемых равно (целому) числу 0, то сумма равна другому слагаемому: $a+0=a$; для каждого целого числа a противоположное к нему число $-a$ (сумма которого с данным числом a равна 0) тоже является целым. Операция сложения (в частности, целых) чисел коммутативна ($a+b=b+a$ для любых двух чисел a и b) и ассоциативна ($(a+b)+c=a+(b+c)$ для любых трех чисел a, b, c).

Далее, если из множества всех целых чисел выделить подмножество чисел, *делящихся на данное число* k, то и оно обладает такими же свойствами. Это множество тоже «замкнуто» относительно «операции сложения» — сумма любых двух чисел, делящихся на k, делится на k; это множество содержит 0 (нуль делится на любое число); и, наконец, если a делится на k, то и $-a$ делится на k.

Аналогичными свойствами обладают и *множество всех рациональных чисел, множество всех вещественных чисел* или *всех комплексных чисел* — каждое из них замкнуто относительно операции сложения; нуль является одновременно числом рациональным, вещественным и комплексным; для каждого (комплексного) числа a имеется противоположное к нему число $-a$ такое, что $a+(-a)=0$, причем $-a$ при вещественном a будет вещественным, а при рациональном a — рациональным. Операция сложения в множестве комплексных чисел (а значит, и подавно, в множестве вещественных и в множестве рациональных чисел) коммутативна и ассоциативна. Все это — примеры «групп по сложению».

Рассмотрим теперь множество всех *отличных от нуля вещественных чисел* и «операцию умножения» в нем. Произведение любых двух таких чисел — снова отличное от нуля вещественное число; произведение любого числа a на (вещественное, отличное от нуля) число 1 равно a, и для каждого (отличного от нуля!) вещественного числа a имеется обратное ему (и тоже отличное от нуля) вещественное число a^{-1}, произведение которого на a равно 1.

Аналогичными свойствами обладает и множество всех *отличных от нуля рациональных чисел,* множество всех *положительных вещественных чисел* или всех *положительных рациональных чисел,* а также множество всех *отличных от нуля комплексных чисел* или множество *комплексных чисел, по модулю равных* 1. Каждое из них замкнуто относительно операции умножения, все они содержат единицу и у каждого из их элементов имеется обратный элемент, принадлежащий тому же множеству. Умножение комплексных (а значит, и вещественных, и рациональных) чисел коммутативно ($ab = ba$ для всех a и b) и ассоциативно ($(ab)c = a(bc)$ для всех a, b, c).

Это — примеры «групп по умножению». Можно привести и более неожиданный пример: группу по умножению образует, например, *пара чисел,* 1 и -1. Впрочем, множество, состоящее *из одного числа* 1 (или 0), тоже образует группу по умножению (соответственно по сложению). Комплексные числа 1, i, -1, $-i$ также образуют, очевидно, группу по умножению.

Складывать можно не только числа, но, например, *векторы* линейного пространства R, причем это сложение подчиняется тем же законам, что и сложение чисел: оно коммутативно и ассоциативно, в R имеется нулевой элемент 0 такой, что $x+0=x$ для любого $x \in R$, и для всякого вектора $x \in R$ имеется противоположный ему вектор $-x$, такой, что $x+(-x)=0$.

Складывать можно *матрицы одного и того же строения* (т. е. $[m \times n]$-матрицы, где m и n — какие-то заранее заданные целые положительные числа). Это сложение ассоциативно и коммутативно, имеется нулевая матрица, прибавление которой не меняет второго слагаемого — это матрица, состоящая из одних нулей, и для каждой матрицы $[a_{ik}]$ имеется противоположная к ней матрица $[-a_{ik}]$ — такая, что $[a_{ik}]+[-a_{ik}]$ есть нулевая матрица. Если рассматривать только так называемые *целочисленные матрицы* (т. е. матрицы с целыми элементами a_{ik}), то и суммой двух таких матриц будет матрица такого же строения, нулевая матрица является целочисленной, и для каждой целочисленной матрицы, противоположной к ней, будет тоже целочисленная матрица. Все это — тоже примеры групп по сложению.

С другой стороны, и перемножать можно не только числа, но, например, *невырожденные квадратные матрицы одного и того же порядка n* с вещественными (или только с рациональными или, наоборот, с комплексными) элементами. Произведение двух таких матриц тоже будет невырожденной матрицей (теорема 3 главы III) с вещественными (соответственно с рациональными, комплексными) элементами; единичная матрица является невырожденной, и у каждой невырожденной матрицы имеется обратная (тоже невырожденная и тоже с вещественными или соответственно рациональными, комплексными элементами). Умножение матриц ассоциативно, однако оно не коммутативно. Множество всех невырожденных матриц порядка n с вещественными (рациональными, комплексными) элементами представляет собой пример некоммутативной группы по умножению.

Дадим теперь общее определение группы.

Определение 1. ***Группой** называется множество G элементов a, b, ..., для которых определена некоторая алгебраическая операция (обычно называемая умножением или сложением), ставящая в соответствие каждой упорядоченной паре a, b элементов из G третий элемент $c = a \circ b$, причем так, что выполнены следующие условия:*

1. *Эта операция ассоциативна: для любых трех элементов a, b, c из G*

$$(a \circ b) \circ c = a \circ (b \circ c).$$

2. *В G существует «нейтральный» элемент e такой,* что

$$a \circ e = e \circ a = a$$

для каждого $a \in G$.

3. *Для каждого элемента a из G существует «обратный» ему элемент a^{-1} такой, что*

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

Группа, в которой дополнительно выполняется коммутативный закон:

4. *Для любых двух элементов $a, b \in G$*

$$a \circ b = b \circ a,$$

называется коммутативной, или абелевой.

Группа, состоящая из конечного числа элементов, называется конечной группой. Число элементов конечной группы называется ее порядком. Группа, не являющаяся конечной, называется бесконечной.

В том случае, когда «групповая операция» $a \circ b$ называется сложением и обозначается знаком $+$, группа G называется *группой по сложению*, или *аддитивной группой*. В этом случае «нейтральный элемент» e обычно обозначается символом 0 и называется *нулём*, а элемент, обратный к a, обозначается через $-a$ и называется *противоположным* к a. В том случае, когда групповая операция называется умножением, $a \circ b$ обозначается через ab, группа называется *группой по умножению*, или *мультипликативной группой*, а нейтральный элемент называется *единицей* и часто обозначается символом 1.

Пользуясь ассоциативным законом, можно определить произведение (сумму) трех и большего числа элементов группы. Так как $(ab)c = a(bc)$, то имеет смысл говорить просто о произведении abc трех элементов, равном, по определению, $(ab)c = a(bc)$. Так же как для линейных пространств, можно доказать *единственность единичного (нулевого) элемента группы* и *единственность элемента, обратного (противоположного) данному.*

Заметим, что для каждого элемента a группы $(a^k)^{-1} = (a^{-1})^k$, так как $a^k (a^{-1})^k = \underbrace{a \ldots a}_{k \text{ раз}} \underbrace{a^{-1} \ldots a^{-1}}_{k \text{ раз}} = e$; вместо $(a^{-1})^k$ мы будем также писать a^{-k}. Далее, в каждой (например, мультипликативной) группе однозначно разрешимы уравнения $ax = b$ (решением которого, очевидно, будет $x = a^{-1}b$) и $ya = b$ (для которого $y =$

$=ba^{-1}$); ясно, что если группа коммутативна, то эти уравнения не различаются и $x = y$).

Еще одним важным примером группы может служить группа вращений правильного многоугольника. Пусть дан правильный n-угольник $A_1A_2\ldots A_n$, и пусть O — его центр (сделайте чертеж). Рассмотрим всевозможные повороты плоскости вокруг точки O, при которых этот n-угольник совмещается сам с собой. Таких поворотов, очевидно, n:

a_0 — поворот на $\angle 0$ (тождественное преобразование),

a_1 — поворот на $\angle \frac{2\pi}{n}$,

a_2 — поворот на $\angle \frac{2\pi}{n}\cdot 2$,

. . . .

a_{n-1} — поворот на $\angle \frac{2\pi}{n}(n-1)$.

По определению, *умножение* поворотов — это их последовательное выполнение одного за другим:

$$a_k \circ a_l = a_{k+l};$$

при этом естественно считать, что $a_{k+n} = a_k$ для любого k, в частности, $a_n = a_0$. Это умножение, очевидно, ассоциативно (и коммутативно). Поворот a_0 является единичным элементов группы и $a_k^{-1} = a_{n-k}$ для всех $k = 0, 1, \ldots, n-1$.

Если положить $a_1 = a$, мы будем иметь $a_2 = a^2$, $a_3 = a^3$, ..., $a_{n-1} = a^{n-1}$ и, наконец, $a_n = a^n = a_0$. Можно сказать, что эта группа *образована степенями одного из своих элементов* (или что она *«порождается»* одним из своих элементов), а именно, элемента $a = a_1$. Такие группы называются **циклическими**. Группа вращений правильного n-угольника является *циклической группой порядка n*; обозначается эта группа символом C_n.

Группа целых чисел (по сложению) тоже является циклической — она порождается одним из своих элементов: ведь $2 = 1 + 1$, $3 = (1 + 1) + 1$, -1 есть элемент, противоположный к 1, и т. д. Эта группа является *бесконечной циклической группой*; *обозначается она символом* C_∞.

Рассмотрим еще один пример — группу V самосовмещений, или *группу симметрии, ромба* (она называется еще *клейновской группой* четвертого порядка). Пусть дан ромб $ABCD$ (рис. 27). Он переходит в себя при следующих преобразованиях:

b_1 — тождественное преобразование,
b_2 — симметрия относительно AC,
b_3 — симметрия относительно BD,
b_4 — симметрия относительно центра O.

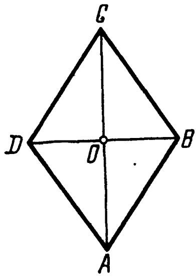


Рис. 27.

Произведение (т. е. результат последовательного выполнения одного за другим) любых двух из этих преобразований — снова одно из них. Эти преобразования образуют группу, которую можно представить следующей «таблицей умножения»:

	b_1	b_2	b_3	b_4
b_1	b_1	b_2	b_3	b_4
b_2	b_2	b_1	b_4	b_3
b_3	b_3	b_4	b_1	b_2
b_4	b_4	b_3	b_2	b_1

(Ассоциативность этого умножения будет вытекать из результатов § 3.)

Аналогичную таблицу умножения, где слева стоят левые множители b_i, сверху — правые b_k, а на пересечении соответствующих строки и столбца — их произведение b_ib_k, можно написать для каждой *конечной* группы. Таблицы такого рода называются **таблицами Кэли**. Легко видеть, что в *каждой строке и в каждом столбце таблицы Кэли все элементы стоят по одному разу* (так как из равенства $b_ib_j = b_ib_k$ умножением слева на b_i^{-1} получаем $b_j = b_k$, и из равенства $b_jb_i = b_kb_i$ тоже следует, что $b_j = b_k$). Ясно также, что *если группа коммутативна, то ее таблица Кэли симметрична относительно главной диагонали* (т. е. при всех i и k элемент, стоящий в пересечении i-й строки и k-го столбца, равен элементу, стоящему в пересечении k-й строки и i-го столбца).

§ 2. Подгруппа

Определение 2. *Подгруппой группы G называется совокупность G_1 элементов группы G, сама являющаяся группой относительно заданной в G операции.*

Так, в аддитивной группе вещественных чисел содержится подгруппа целых чисел, а в ней при любом k — подгруппа чисел, кратных k. Сама группа вещественных чисел содержится в качестве подгруппы в группе комплексных чисел.

В мультипликативной группе отличных от нуля комплексных чисел содержится подгруппа вещественных чисел, а в ней — подгруппа рациональных чисел, подгруппа положительных вещественных чисел.

Много интересных подгрупп содержит мультипликативная группа невырожденных матриц порядка n (*полная линейная группа*), например, с вещественными элементами. Отметим, в частности, подгруппу ортогональных матриц и подгруппу унимодулярных матриц (т. е. матриц с определителем, равным 1). Подгруппами полной линейной группы являются также группа матриц с определителем, равным ± 1, группа матриц с положительным определителем, группа диагональных матриц, группа скалярных матриц, т. е. матриц вида cE, где $c \neq 0$ — любое число, а E — единичная матрица, группа треугольных матриц, т. е. матриц, у которых все элементы снизу (сверху) от главной диагонали равны нулю.

Для того чтобы убедиться в том, что подмножество G_1 группы G является ее подгруппой, надо прежде всего проверить, что *произведение* (*сумма*) *любых двух элементов из G_1 принадлежит G_1* и что *если $a \in G_1$, то и $a^{-1} \in G_1$*. Но этого и достаточно, так как ассоциативный закон, справедливый во всей группе G, будет выполняться и для элементов из G_1, а элемент e (или 0) как произведение aa^{-1} (как сумма $a+(-a)$) тоже будет принадлежать G_1.

Пусть дана группа G и $a \in G$. Рассмотрим всевозможные степени (положительные и отрицательные)

$$\ldots, a^{-2}, a^{-1}, e = a^0, a, a^2, a^3, \ldots$$

элемента a. Они образуют, очевидно, подгруппу — *циклическую подгруппу, порожденную элементом* a. При этом возможны два случая: либо все эти степени элемента a различны, либо среди них имеются одинаковые. Последнее наверняка будет, например, в любой конечной группе. Пусть, скажем, $a^m = a^l$, где $m > l$. Тогда $a^{m-l} = e$. Обозначим через k наименьшую положительную степень, такую что $a^k = e$. Тогда, для того чтобы имело место равенство $a^n = e$, необходимо и достаточно, чтобы n делилось на k. Действительно, если $n = ks$, то $a^n = (a^k)^s = e$. С другой стороны, если $a^n = e$ и $n = kp + q$, где $0 \leqslant q < k$, то, так как $a^n = a^{kp} \cdot a^q = a^q = e$, и k — наименьшая положительная степень, в которой $a^k = e$, то $q = 0$; и n делится на k. В этом случае элемент a называется *элементом k-го порядка*. Если все степени элемента a различны, то он называется *элементом бесконечного порядка*. (Таким будет, например, любой отличный от 0 элемент аддитивной группы целых чисел.)

Для того чтобы убедиться в том, что данное множество G_1 элементов конечной группы образует в ней подгруппу, достаточно проверить, что произведение (сумма) любых двух элементов множества G_1 принадлежит G_1. Действительно, в конечной группе каждый элемент a — конечного порядка, и если $a \in G_1$ и $a^k = e$ (откуда уже следует, что $e = \underbrace{a \cdot a \cdot \ldots \cdot a}_{k \text{ раз}}$ принадлежит множеству G_1), то $a^{k-1} \cdot a = a \cdot a^{k-1} = e$, и элемент $a^{k-1} \in G_1$ является обратным к a.

Легко видеть, что *пересечение двух подгрупп группы G* само *является подгруппой в G*. Каждая группа имеет подгруппу, состоящую из одной единицы (нуля), и каждая группа сама является своей подгруппой (эти подгруппы называются несобственными). Ясно, *что подгруппа коммутативной группы всегда будет коммутативной*, в то время как подгруппа некоммутативной группы может быть и некоммутативной, и коммутативной; так, (некоммутативная) полная линейная группа содержит коммутативную подгруппу скалярных матриц.

§ 3. Группы преобразований. Симметрическая группа *n*-й степени

Важный класс групп составляют так называемые группы преобразований. Пусть M — произвольное множество. *Преобразованием множества M мы теперь будем называть любое взаимно однозначное отображение P этого множества на себя.* Это значит, что для каждого элемента x из M однозначно определен его образ $Px = x' \in M$, причем каждый элемент x' из M служит образом единственного элемента x, называемого его *прообразом.*

Умножением преобразований называется последовательное выполнение их одного за другим: по определению,

$$(PQ)x = P(Qx).$$

Умножение преобразований ассоциативно — это можно доказать точно так же, как доказывалось выше для линейных операторов (§ 2 главы III), но, вообще говоря, не коммутативно (не коммутативно уже умножение линейных операторов). Роль единицы в этом умножении играет *тождественное преобразование* E, ставящее в соответствие каждому элементу x из M его самого. Для каждого преобразования P множества M существует *обратное преобразование* P^{-1}, ставящее в соответствие каждому элементу x' из M его (единственный по условию) прообраз x; при этом, очевидно,

$$PP^{-1} = P^{-1}P = E.$$

Если множество M конечно и состоит из n элементов, то всевозможные взаимно однозначные отображения этого множества на себя называются *подстановками*, а соответствующая группа преобразований обозначается через S_n и называется *группой подстановок* из n элементов, или *симметрической группой n-й степени.*

Рассмотрим *симметрическую группу третьей степени* S_3 — группу всех взаимно однозначных отображений множества, состоящего из трех элементов a, b, c, — например, это могут быть числа 1, 2, 3, на себя. Так как из трех элементов можно составить всего шесть различных перестановок:

$$123, \quad 132, \quad 321, \quad 213, \quad 231, \quad 312,$$

то и число различных подстановок для них равно шести. Обозначать их удобно следующим образом:

$$P_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

$$P_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad P_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

где, например, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ — это такое отображение множества 1, 2, 3 на себя, при котором $1 \to 2$ (1 отображается в 2), $2 \to 3$ и $3 \to 1$. Подстановки, отличающиеся только порядком следования столбцов, например,

$$\begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix} \text{ и } \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

не считаются различными. Умножение подстановок — это их последовательное выполнение (сначала правого множителя, а затем — левого), поэтому, например,

$$P_6 P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = P_3,$$

ибо в правом множителе $1 \to 1$, в левом $1 \to 3$, следовательно, в произведении $1 \to 3$, и т. д. Единицей при этом умножении служит тождественная подстановка $P_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ и для каждой подстановки имеется обратная ей:

$$P_2^{-1} = P_2, \quad P_3^{-1} = P_3, \quad P_4^{-1} = P_4, \quad P_5^{-1} = P_6, \quad P_6^{-1} = P_5.$$

Для того чтобы получить подстановку, обратную данной, надо лишь поменять местами ее строки:

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

Группу S_3 можно представить такой таблицей Кэли:

S_3	P_1	P_2	P_3	P_4	P_5	P_6
P_1	P_1	P_2	P_3	P_4	P_5	P_6
P_2	P_2	P_1	P_5	P_6	P_3	P_4
P_3	P_3	P_6	P_1	P_5	P_4	P_2
P_4	P_4	P_5	P_6	P_1	P_2	P_3
P_5	P_5	P_4	P_2	P_3	P_6	P_1
P_6	P_6	P_3	P_4	P_2	P_1	P_5

Группа S_3 некоммутативна, так как, например,

$$P_4P_5 = P_2, \quad \text{а} \quad P_5P_4 = P_3$$

(таблица Кэли этой группы не симметрична относительно главной диагонали).

Мы подробно рассмотрели группу подстановок из трех элементов; обратимся теперь к общему случаю. Подстановку из n элементов — например, чисел $1, 2, \ldots$ $\ldots, n$ — можно обозначить символом

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \end{pmatrix},$$

показывающим, что 1 переходит в α_1, 2 — в α_2, и т. д.; здесь $\alpha_1, \alpha_2, \ldots, \alpha_n$ — это те же числа $1, 2, 3, \ldots, n$, но расположенные, вообще говоря, в каком-то другом порядке. Расположение столбцов в этой записи не играет роли и, например,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 4 & 2 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \quad \text{и т. д.}$$

Число подстановок из n элементов равно, очевидно, $n!$, Перемножаются подстановки в общем случае так же, как подстановки из трех элементов. Так, например,

$$\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 3 & 4 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 3 & 1 & 2\end{pmatrix}=\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 4 & 2 & 3\end{pmatrix}.$$

(Сначала выполняется правая подстановка, а потом левая: здесь $1\to 4$, а затем $4\to 1$; далее, $2\to 3$, а затем $3\to 4$, и т. д.) Умножение подстановок *ассоциативно*, но, вообще говоря, *не коммутативно*. Подстановка

$$\begin{pmatrix}1 & 2 & 3\dots n\\ 1 & 2 & 3\dots n\end{pmatrix}$$

играет роль единицы и называется *тождественной подстановкой*. У каждой подстановки имеется *обратная*:

$$\begin{pmatrix}1 & 2 & \dots n\\ \alpha_1 & \alpha_2 & \dots \alpha_n\end{pmatrix}^{-1}=\begin{pmatrix}\alpha_1 & \alpha_2 & \dots \alpha_n\\ 1 & 2 & \dots n\end{pmatrix}.$$

Группа подстановок из n элементов (*симметрическая группа n-й степени*) *имеет*, очевидно, *порядок* $n!$.

Подстановки бывают двух типов: четные и нечетные. Подстановка называется *четной*, если обе составляющие ее перестановки (т. е. верхняя строка и нижняя) — одинаковой четности, и *нечетной* — в противном случае. Это определение не зависит от способа записи подстановки: если поменять местами ее столбцы, то в обеих составляющих ее перестановках произойдет по одной транспозиции, отчего четность каждой из них изменится, а значит, четность самой подстановки не изменится.

Теорема 1. *Произведение двух подстановок одинаковой четности является четной подстановкой, а произведение двух подстановок разной четности — нечетной подстановкой.*

Доказательство. Рассмотрим произведение двух подстановок

$$AB=\begin{pmatrix}\beta_1 & \beta_2 & \dots \beta_n\\ \gamma_1 & \gamma_2 & \dots \gamma_n\end{pmatrix}\begin{pmatrix}\alpha_1 & \alpha_2 & \dots \alpha_n\\ \beta_1 & \beta_2 & \dots \beta_n\end{pmatrix}=\begin{pmatrix}\alpha_1 & \alpha_2 & \dots \alpha_n\\ \gamma_1 & \gamma_2 & \dots \gamma_n\end{pmatrix}.$$

Если подстановки A и B одинаковой четности, то либо они обе четны, либо обе нечетны. В первом случае перестановки $\alpha_1\alpha_2\ldots\alpha_n$ и $\beta_1\beta_2\ldots\beta_n$, а также $\beta_1\beta_2\ldots\beta_n$ и $\gamma_1\gamma_2\ldots\gamma_n$ — одинаковой четности, и значит, перестановки $\alpha_1\alpha_2\ldots\alpha_n$ и $\gamma_1\gamma_2\ldots\gamma_n$ — одинаковой четности. Во втором случае перестановки $\alpha_1\alpha_2\ldots\alpha_n$ и $\beta_1\beta_2\ldots\beta_n$ — разной четности, но и перестановки $\beta_1\beta_2\ldots\beta_n$ и $\gamma_1\gamma_2\ldots\gamma_n$ — тоже разной четности, а значит, перестановки $\alpha_1\alpha_2\ldots\alpha_n$ и $\gamma_1\gamma_2\ldots\gamma_n$ — одинаковой четности. В обоих случаях подстановка AB — четна.

Если подстановки A и B — разной четности, то либо подстановка A четна, а B нечетна, либо — наоборот. В обоих случаях перестановки $\alpha_1\alpha_2\ldots\alpha_n$ и $\gamma_1\gamma_2\ldots\gamma_n$ разной четности, и значит, подстановка AB — нечетна.

Следствие. *Все четные подстановки симметрической группы S_n образуют в ней подгруппу.* Порядок этой подгруппы равен, очевидно, $\frac{n!}{2}$. Она называется *знакопеременной* подгруппой симметрической группы и обозначается символом A_n.

§ 4. Изоморфизм групп

В симметрической группе третьей степени S_3 имеются три подгруппы второго порядка: $\{P_1, P_2\}$, $\{P_1, P_3\}$, $\{P_1, P_4\}$ с таблицами Кэли:

	P_1	P_2
P_1	P_1	P_2
P_2	P_2	P_1

	P_1	P_3
P_1	P_1	P_3
P_3	P_3	P_1

	P_1	P_4
P_1	P_1	P_4
P_4	P_4	P_1

Если рассматривать их независимо от группы S_3, они отличаются друг от друга только обозначениями элементов. В группе S_3 имеется еще подгруппа A третьего порядка $\{P_1, P_5, P_6\}$ с таблицей Кэли:

	P_1	P_5	P_6
P_1	P_1	P_5	P_6
P_5	P_5	P_6	P_1
P_6	P_6	P_1	P_5

Сравним ее с группой вращений правильного треугольника:

	a_0	a_1	a_2
a_0	a_0	a_1	a_2
a_1	a_1	a_2	a_0
a_2	a_2	a_0	a_1

Эти группы тоже отличаются только обозначениями элементов. Такие группы называются изоморфными; их можно считать одинаковыми, поскольку с точки зрения их «групповых» свойств (т. е. тех свойств, которые единственно изучаются в теории групп) они не различаются между собой.

Дадим теперь определение изоморфизма групп.

Определение 3. *Группы G_1 и G_2 называются изоморфными, если между их элементами можно установить взаимно однозначное соответствие, сохраняющее групповую операцию, т. е. такое, что если*

$$x_1,\ y_1 \in G_1,\ x_2,\ y_2 \in G_2 \text{ и } x_1 \leftrightarrow x_2,\ y_1 \leftrightarrow y_2,$$

то

$$x_1 \circ y_1 \leftrightarrow x_2 \circ y_2.$$

Соответствие $\leftrightarrow$ можно рассматривать как такое (взаимно однозначное) отображение f (скажем, мультипликативной) группы G_1 на (например, мультипликативную же) группу G_2, что для всех $x, y \in G_1$

$$f(xy) = f(x)f(y).$$

Заметим, что если f — изоморфное отображение группы G_1 на группу G_2, то $f(e_1) = e_2$, где e_i, $i = 1, 2$,— единица группы G_i и $[f(x_1^{-1})] = [f(x_1)]^{-1}$ для каждого $x_1 \in G_1$. Действительно, пусть x_2 — произвольный элемент группы G_2 и x_1 — такой элемент группы G_1, что $f(x_1) = x_2$. Тогда

$$x_2 = f(x_1) = f(x_1e_1) = f(x_1)f(e_1) = x_2f(e_1)$$

и

$$x_2 = f(x_1) = f(e_1x_1) = f(e_1)f(x_1) = f(e_1)x_2,$$

а значит, $f(e_1) = e_2$ — единица группы G_2 (ср. стр. 72). Далее,

$$f(x_1)f(x_1^{-1}) = f(x_1x_1^{-1}) = f(e_1) = e_2$$

и

$$f(x_1^{-1})f(x_1) = f(x_1^{-1}x_1) = f(e_1) = e_2,$$

и значит,

$$[f(x_1)]^{-1} = f(x_1^{-1}).$$

Легко проверить, что *все группы второго порядка* (*а также все группы третьего порядка*) *между собой изоморфны*. Но для порядка четыре существуют уже две неизоморфные между собой группы: *группа вращений квадрата* C_4 и *группа симметрии ромба* V.

Выше мы назвали циклической группой группу, образованную степенями одного из своих элементов. Можно сказать, что *циклическая группа порядка n — это группа, изоморфная группе вращений правильного n-угольника* (легко видеть, что все циклические группы одного и того же порядка изоморфны между собой!), а *бесконечная циклическая группа — это группа, изоморфная аддитивной группе целых чисел*.

Заметим еще, что операции в изоморфных группах могут обозначаться по-разному. Так, мультипликативная группа положительных чисел изоморфна аддитивной группе вещественных чисел. Изоморфное соответствие между ними устанавливается отображением

$$f(a) = \log_c a,$$

где $c \neq 1$ — произвольно выбранное фиксированное положительное число.

§ 5. Разложение группы по подгруппе

Рассмотрим сначала следующий пример. Пусть G будет *аддитивная группа целых чисел* и A — ее *подгруппа*, состоящая из всех чисел, *кратных* k. Разобьем группу G на классы, относя к одному классу числа, *дающие при делении на* k *одинаковые остатки.* Тогда для того, чтобы два числа x и y попали в один и тот же класс, необходимо и достаточно, чтобы их разность делилась на k и, значит, принадлежала подгруппе A:

$$x-y = kn \in A,$$

откуда $x = y + a$, где $a \in A$.

Так мы получим, очевидно, k классов, считая одним из классов и подгруппу A.

Схематически это разложение группы целых чисел по подгруппе чисел, кратных k, при $k = 5$ можно представить следующим образом:

A	...	—20,	—15,	—10,	—5,	0,	5,	10	15,	20,	...
$1+A$	...	—19,	—14,	—9,	—4,	1,	6,	11,	16,	21,	...
$2+A$	...	—18,	—13,	—8	—3,	2,	7,	12,	17,	22,	...
$3+A$	...	—17,	—12,	—7,	—2,	3,	8,	13,	18,	23,	...
$4+A$	...	—16,	—11,	—6,	—1,	4,	9,	14,	19,	24,	...

Введем теперь *операцию сложения в множестве самих классов.* Пусть даны два класса B и C. Выберем в каждом из них по одному элементу (по представителю), скажем, $b \in B$ и $c \in C$, и сложим их, а суммой классов $B+C$ условимся считать класс, содержащий сумму $b+c$. Такое определение сложения классов будет иметь смысл, если тот класс, в котором содержится сумма $b+c$, *не зависит от выбора представителей b и c в классах B и C*; проверим, что это действительно так. Если b' тоже принадлежит B, а $c' \in C$, то

$$b-b' = kn_1 \quad \text{и} \quad c-c' = kn_2,$$

а тогда

$$(b+c)-(b'+c') = k(n_1+n_2)$$

делится на k и, значит, суммы $b+c$ и $b'+c'$ принадлежат одному и тому же классу.

Так определенное сложение классов ассоциативно и коммутативно, ибо этими свойствами обладает сложение в самой группе G. Класс, совпадающий с подгруппой A, играет роль нуля, так как в качестве представителя из A можно взять нуль, а $g+0=g$ при всех $g \in G$. Наконец, для каждого класса B имеется противоположный ему: если $b \in B$, то класс, содержащий $-b$, будет противоположным к B, так как $b+(-b)=0 \in A$. Поэтому совокупность построенных классов сама образует группу относительно определенного нами сложения классов.

Полученная группа (группа классов) называется *фактор-группой группы целых чисел по подгруппе чисел, кратных k*. Она является, очевидно, циклической группой порядка k.

Аналогичная конструкция применима и в общем случае. Пусть G — произвольная, на этот раз мультипликативная группа, и A — некоторая ее подгруппа. Обозначим через xA множество всех элементов вида xa, где $a \in A$; xA называется левым смежным классом *группы G по подгруппе A*.

Каждый элемент y, принадлежащий классу xA, назовем эквивалентным x (будем писать $y \sim x$). Отметим следующие свойства этого понятия:

1. *Каждый элемент x эквивалентен самому себе*: $x \sim x$ (рефлексивность отношения $\sim$), так как $x = xe \in xA$.

2. *Если $y \sim x$, то $x \sim y$ (симметричность отношения $\sim$)*.

Действительно, если $y \sim x$, т. е. $y \in xA$, то $y = xa$, где $a \in A$, а тогда $x = ya^{-1} \in yA$, и значит, $x \sim y$.

3. *Если* $x \sim y$ *и* $y \sim z$, *то* $x \sim z$ (транзитивность отношения $\sim$). По условию, $x \sim y$, т. е. $x = ya_1$, и $y \sim z$, т. е. $y = za_2$, где a_1, $a_2 \in A$. Но тогда $x = (za_2)a_1 = z(a_2a_1) \in zA$, и значит, $x \sim z$.

Сделаем теперь отступление общего характера. Предположим, что для элементов некоторого множества M задано отношение $\sim$ (запись $x \sim y$ читается: «x эквивалентно y»), обладающее свойствами рефлексивности (всегда $x \sim x$), симметричности (если $x \sim y$, то $y \sim x$) и транзитивности (если $x \sim y$ и $y \sim z$, то $x \sim z$); тогда говорят, что в этом множестве задано *отношение эквивалентности*. Примерами отношения эквивалентности могут служить равночисленность конечных наборов предметов, параллельность прямых, подобие треугольников, и т. д.

Теорема 2. *Если в множестве M задано отношение эквивалентности, то это множество разбивается на непересекающиеся классы эквивалентных между собой элементов.*

Доказательство. Обозначим через $S(x)$ множество всех элементов, эквивалентных x (элемент $x' \in S(x)$ в том и только в том случае, если $x' \sim x$). Покажем сначала, *что если элементы x и y эквивалентны, то соответствующие классы $S(x)$ и $S(y)$ совпадают.* Действительно, если $x' \in S(x)$, то $x' \sim x$. Но так как $x \sim y$, то $x' \sim y$ и, значит, $x' \in S(y)$. Мы видим, что каждый элемент x' из класса $S(x)$ принадлежит $S(y)$. Аналогично показывается, что каждый элемент y' из класса $S(y)$ принадлежит $S(x)$. Следовательно, $S(x) = S(y)$.

Покажем, далее, что *если элементы x и y не эквивалентны, то классы $S(x)$ и $S(y)$ не пересекаются.* Действительно, если $z \in S(x) \cap S(y)$, то $z \sim x$ и $z \sim y$, а тогда $x \sim y$. Теорема доказана.

Вернемся к нашей группе G и введенному в ней выше отношению $\sim$. По теореме 2 группа G разбивается на (непересекающиеся) классы эквивалентных между собой элементов. Эти классы называются *левыми смеж-*

ными классами группы G по *подгруппе* A. Одним из этих классов будет, очевидно, сама подгруппа A.

Если G — конечная группа, то все ее смежные классы по данной подгруппе A состоят из одного и того же числа элементов (элементы смежного класса xA взаимно однозначно соответствуют элементам подгруппы A, так как если бы xa_1 равнялось xa_2, то a_1 было бы равно a_2). Отсюда вытекает важная

Теорема Лагранжа. *Порядок подгруппы конечной группы является делителем порядка группы.*

Доказательство. Пусть G — конечная группа порядка n и A — ее подгруппа порядка k. Разложим группу G на левые смежные классы по подгруппе A. Если j — число полученных классов, то, поскольку каждый класс состоит из k элементов, общее число элементов группы

$$n = kj,$$

и значит, n делится на k. Число j (тоже являющееся, очевидно, делителем n) называется **индексом** подгруппы A в группе G.

Каждый элемент g группы G порождает в ней циклическую подгруппу $\{g\}$, состоящую из всех степеней этого элемента. Порядок подгруппы $\{g\}$ совпадает с порядком элемента g в группе G. Ввиду теоремы Лагранжа *порядок каждого элемента конечной группы является делителем порядка группы.*

Всякая конечная группа, порядок которой — простое число, является циклической, так как циклическая подгруппа, порожденная в ней любым из ее элементов (кроме e), должна совпадать со всей группой.

Аналогично левостороннему разложению, можно построить **правостороннее** разложение группы G по подгруппе A (на классы Ax). В коммутативном случае оба разложения совпадают (состоят из одних и тех же классов).

В некоммутативной группе левостороннее и правостороннее разложения могут оказаться различными. Рассмотрим, например, разложение симметрической группы S_3 по ее подгруппе $B = \{P_1, P_2\}$. Левостороннее разложение состоит из классов

$$B, \quad P_5B = P_4B = \{P_4, P_5\}, \quad P_3B = P_6B = \{P_3, P_6\},$$

правостороннее разложение — из классов

$$B, \quad BP_6 = BP_4 = \{P_4, P_6\}, \quad BP_5 = BP_3 = \{P_3, P_5\}.$$

В то же время левостороннее и правостороннее разложения группы S_3 по ее подгруппе третьего порядка $A = \{P_1, P_5, P_6\}$ совпадают: каждое из них состоит из двух классов

$$A = \{P_1, P_5, P_6\} \text{ и } AP_2 = P_2A = \{P_2, P_3, P_4\}.$$

§ 6. Нормальная подгруппа

Обобщим теперь конструкцию, которая в начале § 5 привела нас к понятию группы классов (фактор-группы) аддитивной группы целых чисел. Пусть A — подгруппа произвольной группы G. Образуем всевозможные левые смежные классы группы G по подгруппе A и попытаемся определить умножение этих классов следующим образом: если даны два класса B и C, выберем из них по представителю: $b \in B$, $c \in C$, перемножим этих представителей и в качестве произведения BC возьмем тот класс, в котором содержится bc. Необходимо только проверить, *не зависит ли это определение произведения классов от выбора представителей в них.* Итак, пусть $b' \sim b$, $c' \sim c$; можем ли мы утверждать, что $b'c' \sim bc$? По условию, $b' = ba_1$ и $c' = ca_2$, откуда $b'c' = ba_1ca_2$. Если группа G коммутативна, то

$$a_1c = ca_1 \tag{1}$$

и

$$b'c' = bc(a_1a_2), \quad \text{т. е.} \quad b'c' \sim bc.$$

В некоммутативной группе равенство (1), вообще говоря, места не имеет. Однако для нашей цели достаточно следующего, более слабого, чем коммутативность, условия: *достаточно, чтобы произведение a_1c можно было представить в виде ca_3*, где $a_3 \in A$, причем a_3, вообще говоря, отлично от a_1. Если это так, то $b'c' = bc(a_3a_2)$, т. е. $b'c' \sim bc$, и произведение классов *не зависит от выбора представителей в них.*

Итак, мы будем теперь предполагать, что подгруппа A обладает следующим свойством: *для каждого элемента $a \in A$ и произвольного элемента $g \in G$ найдется элемент $\tilde{a} \in A$ такой, что $\tilde{a}g = ga$.* Это значит, что *для* *любого* $a \in A$ *и произвольного* $g \in G$ *произведение* $g^{-1}ag = \tilde{a} \in A$. Обозначив через $g^{-1}Ag$ множество всевозможных элементов вида $g^{-1}ag$, где $a \in A$, дадим следующее

Определение 4. *Подгруппа A группы G называется ее* ***нормальной подгруппой***, или ***нормальным делителем***, или еще ***инвариантной подгруппой***, *если для любого элемента $g \in G$*

$$g^{-1}Ag \subseteq A.$$

Теорема 3. *Пересечение двух нормальных подгрупп группы G само является нормальной подгруппой G.*

Доказательство. Пусть A_1 и A_2 — нормальные подгруппы группы G и $A = A_1 \cap A_2$. Мы знаем, что A — подгруппа в G (§ 2). Далее, так как $A \subseteq A_1$ и $A \subseteq A_2$, то для любого элемента $g \in G$

$$g^{-1}Ag \subseteq g^{-1}A_1g \subseteq A_1$$

и

$$g^{-1}Ag \subseteq g^{-1}A_2g \subseteq A_2,$$

а значит,

$$g^{-1}Ag \subseteq A_1 \cap A_2 = A,$$

т. е. A — нормальная подгруппа группы G.

Теорема 4. *Для того чтобы подгруппа A группы G была ее нормальной подгруппой, необходимо и достаточно, чтобы для любого элемента $g \in G$ имело место равенство*

$$g^{-1}Ag = A. \tag{2}$$

Доказательство. Достаточность условия (2) следует из определения 4. Для того чтобы доказать его необходимость, предположим, что A — нормальная подгруппа группы G; тогда для любого элемента $g \in G$ $g^{-1}Ag \subseteq A$, а значит, и $gAg^{-1} = (g^{-1})^{-1}Ag^{-1} \subseteq A$, откуда, в свою очередь, следует, что $A \subseteq g^{-1}Ag$. Но если $g^{-1}Ag \subseteq A$ и $A \subseteq g^{-1}Ag$, то $g^{-1}Ag = A$.

Теорема 5. *Для того чтобы подгруппа A группы G была ее нормальной подгруппой, необходимо и достаточно, чтобы левые и правые смежные классы группы G по подгруппе A совпадали.*

Доказательство. Из равенства

$$g^{-1}Ag = A$$

вытекает, что

$$Ag = gA,$$

т. е. что для любого $g \in G$ левый и правый смежные классы, содержащие этот элемент, совпадают.

Обратно, если для любого $g \in G$

$$Ag = gA,$$

то

$$g^{-1}Ag = A,$$

и A — нормальная подгруппа.

Так, в симметрической группе S_3 подгруппа $A = \{P_1, P_5, P_6\}$ будет нормальной подгруппой, подгруппы же $\{P_1, P_2\}$, $\{P_1, P_3\}$ и $\{P_1, P_4\}$ нормальными не являются.

Легко видеть, что при любом *n знакопеременная подгруппа A_n является нормальной подгруппой симметрической группы S_n*, так как разложение группы S_n по подгруппе A_n (и левостороннее, и правостороннее) состоит из двух классов — самой подгруппы A_n и множества B всех остальных элементов (т. е. множества всех нечетных подстановок). Совершенно аналогично этому *в любой группе всякая подгруппа индекса 2 является нормальной подгруппой.*

В коммутативной группе любая подгруппа является, очевидно, нормальной.

§ 7. Фактор-группа

Пусть G — произвольная группа, A — ее нормальная подгруппа и S — множество всевозможных смежных класов группы G по подгруппе A (напоминаем, что левые и правые смежные классы в этом случае совпадают). *В множестве классов S введем операцию умножения, полагая*

$$xA \cdot yA = xyA.$$

Так как подгруппа A является нормальной, то *произведение $xA \cdot yA$ не зависит от выбора представителей x и y в перемножаемых классах.*

Проверим, что у нас получилась группа.

1. Ассоциативность умножения классов вытекает из ассоциативности умножения в группе G:

$$(xA \cdot yA) \cdot zA = (xy)A \cdot zA = (xy)zA = \\ = x(yz)A = xA \cdot (yz)A = xA \cdot (yA \cdot zA).$$

2. Единичным элементом служит сама подгруппа A:

$$A \cdot xA = eA \cdot xA = exA = xA,$$

$$xA \cdot A = xA \cdot eA = xeA = xA.$$

3. Обратным к классу xA будет класс $x^{-1}A$, так как

$$xA \cdot x^{-1}A = xx^{-1}A = eA = A,$$

$$x^{-1}A \cdot xA = x^{-1}xA = eA = A.$$

Полученная группа обозначается через G/A и называется ***фактор-группой*** *группы G по нормальной подгруппе A.*

Фактор-группа коммутативной группы коммутативна, так как в этом случае для любых двух классов

$$xA \cdot yA = (xy)A = (yx)A = yA \cdot xA.$$

Порядок фактор-группы конечной группы равен индексу нормальной подгруппы A в группе G, и значит, *является делителем порядка n группы G.*

Фактор-группа симметрической группы S_n по ее подгруппе A_n состоит из двух элементов и является, следовательно, циклической группой второго порядка.

Пример. *Покажем, что в группе G всех невырожденных матриц порядка n (например, с вещественными элементами) подгруппа A унимодулярных матриц* (т. е. матриц с определителем, равным 1) *является нормальной подгруппой, и найдем фактор-группу G/A.*

Унимодулярные матрицы образуют подгруппу в G, так как произведение двух унимодулярных матриц и матрица, обратная унимодулярной, являются унимодулярными (теорема 3 главы III).

Далее, подгруппа A унимодулярных матриц является нормальной, так как если матрица $a \in A$ и, значит, $|a| = 1$, то для любой матрицы $g \in G$

$$|g^{-1}ag| = |g^{-1}\|a\|g| = |g|^{-1}|a\|g| = |a| = 1$$

и

$$g^{-1}ag \in A.$$

Найдем фактор-группу G/A. Покажем прежде всего, что для того, чтобы две матрицы b и c принадлежали одному и тому же смежному классу группы G по подгруппе A, необходимо и достаточно, чтобы они имели равные определители. Действительно, если $b \sim c$, т. е. $b = ca$, где $a \in A$ и, значит, $|a| = 1$, то $|b| = |c||a| = |c|$.

Обратно, если $|b| = |c|$, то $b = c(c^{-1}b)$, где

$$|c^{-1}b| = |c^{-1}||b| = |c|^{-1}|b| = 1$$

и, значит, $c^{-1}b \in A$, т. е. $b \in cA$ и $b \sim c$.

Таким образом, каждый смежный класс G по A вполне характеризуется определителем входящих в него матриц. Перемножению классов отвечает перемножение произвольных представителей из них, и, значит, произведение классов B (матриц с определителем β) и C (матриц с определителем γ) есть класс BC — матриц с определителем $\beta\gamma$. Следовательно, *фактор-группа G/A изоморфна мультипликативной группе отличных от нуля вещественных чисел.*

§ 8. Прямое произведение групп

Определение 5. *Пусть даны группа G и две ее подгруппы G_1 и G_2, причем выполнены следующие условия:*

1) *G_1 и G_2 являются нормальными подгруппами группы G,*

2) *пересечение $G_1 \cap G_2$ состоит только из единицы e,*

3) *каждый элемент группы G может быть представлен в виде произведения a_1a_2, где $a_1 \in G_1$, $a_2 \in G_2$.*

Тогда группа G называется **прямым произведением** *своих подгрупп G_1 и G_2.* (Это записывается так: $G = G_1 \times G_2$.)

Теорема 6. I. *Каждый элемент группы $G = G_1 \times G_2$ однозначно представляется в виде произведения a_1a_2, где $a_1 \in G_1$ и $a_2 \in G_2$.*

II. *Каждый элемент $a_1 \in G_1$ коммутирует с каждым элементом $a_2 \in G_2$* (т. е. $a_1a_2 = a_2a_1$).

Доказательство. I. Предположим, что какой-то элемент группы G двумя способами представлен в виде произведения элементов подгрупп G_1 и G_2:

$$a_1a_2 = b_1b_2 \text{ (где } a_1, b_1 \in G_1,\ a_2, b_2 \in G_2).$$

Умножая обе части последнего равенства слева на b_1^{-1}, а справа — на a_2^{-1}, получим

$$b_1^{-1}a_1 = b_2a_2^{-1}. \tag{3}$$

Но $b_1^{-1}a_1 \in G_1$, а $b_2a_2^{-1} \in G_2$, и, значит, элемент (3) принадлежит пересечению $G_1 \cap G_2$, т. е. он равен e:

$$b_1^{-1}a_1 = e = b_2 a_2^{-1},$$

откуда $b_1 = a_1$ и $b_2 = a_2$.

II. Пусть $a_1 \in G_1$ и $a_2 \in G_2$. Рассмотрим так называемый коммутатор этих элементов:

$$a_1^{-1}a_2^{-1}a_1a_2. \tag{4}$$

Произведение $a_1^{-1}a_2^{-1}a_1 \in G_2$, так как G_2 — нормальная подгруппа, и значит, произведение $(a_1^{-1}a_2^{-1}a_1)\,a_2$ принадлежит G_2. С другой стороны, произведение $a_2^{-1}a_1a_2$ принадлежит G_1, так как G_1 — нормальная подгруппа, и значит, произведение $a_1^{-1}(a_2^{-1}a_1a_2)$ принадлежит G_1. Таким образом, коммутатор (4) принадлежит пересечению $G_1 \cap G_2$, и потому он равен e: $a_1^{-1}a_2^{-1}a_1a_2 = e$. Умножая последнее равенство слева на a_2a_1, получим

$$a_1a_2 = a_2a_1,$$

т. е. любой элемент из G_1 коммутирует с любым элементом из G_2.

Аналогично можно определить *прямое произведение* $G = G_1 \times G_2 \times \ldots \times G_n$ *n множителей*. Здесь все подгруппы G_i являются нормальными подгруппами G, пересечение каждой из подгрупп G_i с подгруппой, порожденной в G всеми остальными множителями $G_1, G_2, \ldots, G_{i-1}, G_{i+1}, \ldots, G_n$, состоит только из единицы, и каждый элемент группы G можно представить в виде произведения $a_1a_2 \ldots a_n$, где $a_i \in G_i$.

Легко видеть, что порядок прямого произведения конечных групп равен произведению порядков сомножителей.

Теорема 7. *Пусть даны две группы A и B; тогда существует такая группа G, которая является прямым произведением своих подгрупп G_1 и G_2, соответственно изоморфных данным группам A и B.*

Доказательство. Будем обозначать элементы группы A буквами $a, a', \ldots$, элементы группы B — буквами $b, b' \ldots$. Рассмотрим множество всевозможных пар элементов (a, b), где $a \in A$, $b \in B$. Произведение двух таких пар, по определению, положим равным

$$(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2).$$

Легко видеть, что множество G пар (a, b) с так определенным умножением является группой, единицей которой служит пара (a_0, b_0), где a_0 — единица группы A, а b_0 — единица группы B. Множество G_1

пар вида (a, b_0) образует в G подгруппу, изоморфную, очевидно, группе A, а множество G_2 пар вида (a_0, b) — подгруппу, изоморфную B.

Покажем, что группа G является прямым произведением своих подгрупп G_1 и G_2. Действительно, пересечение подгрупп G_1 и G_2 состоит только из единицы — пары (a_0, b_0). Каждый элемент (a, b) из G является произведением элемента (a, b_0) из G_1 и элемента (a_0, b) из G_2. Наконец, каждая из подгрупп G_1 и G_2 является в G нормальной подгруппой. Покажем это, например, для G_1. Рассмотрим произведение $g^{-1}(a', b_0)g$, где $g = (a, b)$ — произвольный элемент из G, а (a', b_0) принадлежит G_1. Мы имеем, очевидно,

$$g^{-1}(a', b_0)g = (a, b)^{-1}(a', b_0)(a, b) =$$
$$= (a^{-1}, b^{-1})(a', b_0)(a, b) = (a^{-1}a'a, b_0) \in G_1,$$

и значит, подгруппа G_1 — нормальная.

Так, например, прямое произведение $A \times B$ двух циклических групп $A = \{e_1, a\}$ и $B = \{e_2, b\}$ второго порядка состоит из четырех элементов

$$e = (e_1, e_2) \quad a = (a, e_2), \quad b = (e_1, b), \quad ab = (a, b).$$

Легко видеть, что эта группа изоморфна клейновской группе V четвертого порядка.

Прямое произведение циклической группы $a = \{e_1, a\}$ второго порядка и циклической группы $B = \{e_2, b, b^2\}$ третьего порядка состоит из элементов

$$e = (e_1, e_2), \quad a = (a, e_2), \quad b = (e_1, b),$$
$$ab = (a, b), \quad b^2 = (e_1, b^2), \quad ab^2 = (a, b^2)$$

и является циклической группой шестого порядка, так как, например,

$$(ab)^2 = b^2, \quad (ab)^3 = a, \quad (ab)^4 = b, \quad (ab)^5 = ab^2, \quad (ab)^6 = e.$$

§ 9. Классы сопряженных элементов группы

Определение 6. *Пусть G — произвольная* (для определенности, мультипликативная) *группа и a — один из ее элементов. Каждый элемент b вида $g^{-1}ag$, где $g \in G$, называется сопряженным с a.* (Условимся писать в этом случае $b \approx a$.) Говорят еще, что элемент b получается трансформированием элемента a с помощью элемента g.

Отметим следующие свойства отношения сопряженности $\approx$:

1. Каждый элемент сопряжен самому себе, $a \approx a$ (рефлексивность отношения $\approx$),— так как $a = e^{-1}ae$.

2. Если элемент $b \approx a$, то $a \approx b$ (симметричность отношения $\approx$) — так как из равенства $b = g^{-1}ag$ вытекает, что $a = gbg^{-1} = (g^{-1})^{-1}bg^{-1}$.

3. Если $a \approx b$ и $b \approx c$, то $a \approx c$ (транзитивность отношения $\approx$). Действительно, из равенств $a = g_1^{-1}bg_1$, $b = g_2^{-1}cg_2$ вытекает, что $a = g_1^{-1}g_2^{-1}cg_2g_1 = (g_2g_1)^{-1}c(g_2g_1)$.

Таким образом, отношение сопряженности $\approx$ рефлексивно, симметрично и транзитивно, а значит, оно является отношением эквивалентности (см. стр. 289) и определяет *разбиение группы G на непересекающиеся классы сопряженных между собой элементов.*

Множество элементов, сопряженных с данным элементом a (т. е. элементов вида $g^{-1}ag$, где $g \in G$), мы обозначим через $K(a)$. Очевидно, что $a \in K(a)$.

Классы сопряженных элементов состоят, вообще говоря, не из одного и того же числа элементов. Единица всегда образует отдельный класс, так как $g^{-1}eg = e$ при любом g. Вообще, *каждый элемент, перестановочный со всеми остальными элементами группы, образует отдельный класс.* В коммутативной группе каждый элемент образует отдельный класс, и значит, в *коммутативной группе число классов сопряженных элементов равно порядку группы.* В некоммутативной группе число классов меньше порядка группы.

Порядки сопряженных между собой элементов одинаковы. Действительно, если $a^k = e$ и $b = g^{-1}ag$, то

$$b^k = g^{-1}ag \cdot g^{-1}ag \cdot \ldots \cdot g^{-1}ag = g^{-1}a^kg = e,$$

Обратно, если $b^m = e$, то и $a^m = e$, и значит, те наименьшие степени, в которых элементы a и b равны единице, одинаковы.

Пример. В симметрической группе S_3 элемент P_1 первого порядка, элементы P_2, P_3, P_4 — порядка 2, элементы P_5 и P_6 — порядка 3. Единичный элемент P_1 сопряжен только сам с собой. Три элемента порядка 2 сопряжены между собой, так как, например,

$$P_5^{-1}P_2P_5 = P_4 \quad \text{и} \quad P_6^{-1}P_2P_6 = P_3.$$

Элементы P_5 и P_6 тоже сопряжены между собой, так как $P_2^{-1}P_5P_2 = P_6$. Но элементы P_2, P_3, P_4 второго порядка не могут быть сопряжены с элементами P_5 и P_6 третьего порядка. Таким образом, группа S_3 состоит из трех классов сопряженных элементов:

$$\{P_1\},\ \{P_2, P_3, P_4\},\ \{P_5, P_6\}.$$

Мы видим, что число элементов в каждом классе делит порядок группы.

Теорема 8. *Число элементов в каждом классе сопряженных между собой элементов конечной группы является делителем порядка группы.*

Доказательство. Пусть G — произвольная конечная группа и $a \in G$. Обозначим через $N(a)$ множество всех элементов группы, перестановочных с a, $N(a)$ называется *нормализатором* элемента a.

Проверим, что $N(a)$ является подгруппой группы G. Действительно, если $b \in N(a)$ и $c \in N(a)$, то $ab = ba$ и $ac = ca$, а тогда и $a(bc) = (ab)c = (ba)c = b(ac) = b(ca) = (bc)a$, т. е. и $bc \in N(a)$.

Рассмотрим разложение группы G на правые смежные классы по подгруппе $N = N(a)$ и докажем, что между этими классами и элементами, сопряженными с a, существует взаимно однозначное соответствие. Для этого покажем, что если два элемента x и y принадлежат одному и тому же смежному классу G по N, то при трансформировании ими элемента a получается один и тот же элемент b (сопряженный с a), и обратно. Пусть элементы x и y принадлежат одному и тому же смежному классу G по N; тогда $y = hx$, где $h \in N$. Если $x^{-1}ax = b$, то и

$$y^{-1}ay = x^{-1}h^{-1}ahx = x^{-1}ax = b.$$

Обратно, пусть $p^{-1}ap = b$ и $q^{-1}aq = b$. Тогда $p = (pq^{-1})q$, и нам надо показать, что $pq^{-1} \in N$. Но произведение

$$(pq^{-1})^{-1}a(pq^{-1}) = q(p^{-1}ap)q^{-1} = qbq^{-1} = a,$$

и значит, $a(pq^{-1}) = (pq^{-1})a$, откуда $pq^{-1} \in N$ и $p \in Nq$, т. е. элементы p и q принадлежат одному и тому же смежному классу G по N. Так получается взаимно однозначное соответствие между правыми смежными классами группы G по подгруппе $N = N(a)$ и элементами, сопряженными с a (элементу $b = g^{-1}ag$, сопряженному с a, соответствует правый смежный класс, состоящий из всех тех элементов g группы G, при трансформировании которыми элемента a получается b). Следовательно, число элементов, сопряженных с a, равно числу классов в разложении группы G по подгруппе $N(a)$, т. е. равно индексу *нормализатора элемента a в группе* G и, значит, является делителем порядка группы (см. доказательство теоремы Лагранжа).

Теорема 9. *Для того чтобы подгруппа G_1 группы G была нормальной подгруппой, необходимо и достаточно, чтобы она содержала вместе с каждым своим элементом a и весь класс сопряженных с ним элементов $K(a)$.*

Доказательство этой теоремы непосредственно вытекает из определения нормальной подгруппы.

Пересечение Z нормализаторов всех элементов группы G является подгруппой в G (как и пересечение любого множества подгрупп).

Оно состоит из всех тех элементов группы, каждый из которых коммутирует со всеми элементами группы, и называется *центром* группы. *Центр группы является*, очевидно, *ее нормальной подгруппой.*

§ 10. Классы сопряженных элементов прямого произведения групп

Теорема 10. *Пусть группа G равна прямому произведению $G_1 \times G_2$ своих подгрупп G_1 и G_2. Тогда, если A_1 есть класс сопряженных элементов группы G_1, а A_2 — класс сопряженных элементов группы G_2, то всевозможные произведения вида a_1a_2, где $a_1 \in A_1$, $a_2 \in A_2$, образуют класс сопряженных элементов самой группы G, и обратно, каждый класс сопряженных элементов группы G получается таким образом.*

Доказательство. Нам надо показать, что если $a_1 \approx b_1$ в группе G_1 и $a_2 \approx b_2$ — в G_2, то $a_1a_2 \approx b_1b_2$ *в группе* $G = G_1 \times G_2$, и наоборот.

Пусть $b_1 = g_1^{-1}a_1g_1$ и $b_2 = g_2^{-1}a_2g_2$ (где $g_1 \in G_1$ и $g_2 \in G_2$). Тогда, так как элементы из подгрупп G_1 и G_2 коммутируют между собой, имеем

$$b_1b_2 = (g_1^{-1}a_1g_1)(g_2^{-1}a_2g_2) = g_1^{-1}g_2^{-1}a_1a_2g_1g_2 = \\ = (g_2g_1)^{-1}a_1a_2(g_1g_2) = (g_1g_2)^{-1}a_1a_2(g_1g_2),$$

и значит, $b_1b_2 \approx a_1a_2$ в группе G.

Обратно, предположим, что элементы a_1a_2 и b_1b_2 прямого произведения $G = G_1 \times G_2$ сопряжены в G, т. е. что $b_1b_2 = g^{-1}a_1a_2g$, где $g \in G$ и, значит, $g = g_1g_2$, $g_1 \in G_1$, $g_2 \in G_2$. Тогда

$$b_1b_2 = (g_1g_2)^{-1}a_1a_2(g_1g_2) = g_2^{-1}g_1^{-1}a_1a_2g_1g_2 = \\ = (g_1^{-1}a_1g_1)(g_2^{-1}a_2g_2),$$

откуда, ввиду единственности разложения элементов прямого произведения на компоненты из разных сомножителей, получаем $b_1 = g_1^{-1}a_1g_1$ и $b_2 = g_2^{-1}a_2g_2$, т. е. $b_1 \approx a_1$, в группе G_1 и $b_2 \approx a_2$ — в группе G_2.

Следствие. *Если группа G_1 содержит p классов, а группа G_2 — q классов сопряженных элементов, то число классов сопряженных элементов группы $G = G_1 \times G_2$ равно pq.*

§ 11 Гомоморфизм групп

Определение 7. Говорят, что *группа G гомоморфна группе G'*, или что *имеется гомоморфное отображение f группы G на группу G', если каждому элементу x группы G поставлен в соответствие определенный элемент $f(x)$ группы G' (причем каждый элемент группы G' поставлен в соответствие хотя бы одному элементу группы G) так, что для всех элементов $x, y \in G$*

$$f(xy) = f(x)f(y).$$

Как и изоморфизм, это соответствие «сохраняет групповую операцию». Чем же тогда гомоморфизм отличается от изоморфизма? Тем, что здесь отображение группы G на группу G' не предполагается взаимно однозначным: каждому элементу x группы G отвечает один определенный элемент $f(x)$ из G', но разным элементам из G может быть поставлен в соответствии один и тот же элемент из G'. Таким образом, *изоморфизм является частным случаем гомоморфизма.*

Рассмотрим несколько примеров. *Группа симметрии ромба V* с элементами e, a, b, ab (и «определяющими соотношениями» $a^2 = b^2 = e$, $ab = ba$) *гомоморфна циклической группе C_2 второго порядка* с элементами E, A ($A^2 = E$): можно положить, например,

$$f(e) = f(a) = E \quad f(b) = f(ab) = A.$$

Легко видеть, что произведению любых двух элементов группы G отвечает произведение соответствующих элементов группы G'.

Гомоморфное отображение группы V на группу C_2 можно установить и иначе:

$$\varphi(e) = \varphi(b) = E, \quad \varphi(a) = \varphi(ab) = A,$$

или еще так:

$$\psi(e) = \psi(ab) = E, \quad \psi(a) = \psi(b) = A.$$

Другой пример. *Циклическая группа C_6 шестого порядка* с элементами e, a, a^2, a^3, a^4, a^5 *гомоморфна циклической группе C_2 второго порядка* $\{E, A\}$:

$$f(e) = f(a^2) = f(a^4) = E, \quad f(a) = f(a^3) = f(a^5) = A,$$

и *циклической группе C_3 третьего порядка* $\{E, B, B^2\}$:

$$\varphi(e) = \varphi(a^3) = E, \quad \varphi(a) = \varphi(a^4) = B, \quad \varphi(a^2) = \varphi(a^5) = B^2.$$

Каждая группа гомоморфна себе самой (ибо можно положить $f(x) = x$ для всех элементов $x \in G$). Каждая группа гомоморфна единичной группе (состоящей из одного единичного элемента E): для доказательства достаточно положить $f(x) = E$ для всех элементов $x \in G$.

Легко видеть, что если f — гомоморфное отображение группы G на группу G', то $f(e) = e'$, где e — единица группы G, а e' — единица группы G' и $[f(x)]^{-1} = f(x^{-1})$. Это доказывается так же, как соответствующие утверждения для изоморфного отображения (см. § 4).

Каким вообще группам может быть гомоморфна данная группа G? На этот вопрос полностью отвечает следующая

Теорема о гомоморфизмах. *Каждая группа гомоморфна любой своей фактор-группе. Обратно, если группа G гомоморфна группе G', то G' изоморфна фактор-группе группы G по некоторой нормальной подгруппе H.*

Таким образом, группы, которым гомоморфна данная группа G это, с точностью до изоморфизма,— все ее фактор-группы и только они. Следовательно,— в случае конечной группы G — *порядок каждой такой группы является делителем порядка группы.*

Доказательство. I. Пусть H — нормальная подгруппа группы G. Рассмотрим фактор-группу G/H группы G по H. Поставим в соответствие каждому элементу x группы G тот смежный класс, в котором содержится этот элемент, т. е. положим $f(x) = xH$. Тогда $f(y) = yH$ и $f(xy) = xyH$. Но $xyH = xH \cdot yH = f(x)f(y)$, и значит, $f(xy) = f(x)f(y)$, т. е. группа G гомоморфна своей фактор-группе G/H.

II. Пусть группа G гомоморфна группе G'. Рассмотрим множество H всех тех элементов группы G, которые отображаются в единицу e' группы G', т. е. таких, что $f(x) = e'$. Покажем, что H — подгруппа в G. Действительно, если $h_1 \in H$ и $h_2 \in H$, то $f(h_1) = e'$ и $f(h_2) = e'$, а тогда и $f(h_1h_2) = f(h_1)f(h_2) = e'$, т. е. и $h_1h_2 \in H$. Далее, если $h \in H$, то $f(h) = e'$ и $f(h^{-1}) = [f(h)]^{-1} = e'$, т. е. и $h^{-1} \in H$. Эта подгруппа H, называемая *ядром гомоморфизма* f, является нормальной, так как если $h \in H$, т. е. $f(h) = e'$, то для любого элемента $g \in G$

$$f(g^{-1}hg) = f(g^{-1})f(h)f(g) = [f(g)]^{-1}e'f(g) = e',$$

и значит,

$$g^{-1}hg \in H.$$

Покажем теперь, что фактор-группа G/H изоморфна G'. Фактор-группа G/H образована смежными классами группы G по подгруппе H. Все элементы ядра и только они, отображаются в единицу группы G'. Покажем, что все элементы одного и того же смежного класса G по H отображаются в один и тот же элемент группы

G'. Действительно, если элементы x и y группы G принадлежат одному и тому же смежному классу по H, то $x = yh$, где $h \in H$, а тогда

$$f(x) = f(y)f(h) = f(y)e' = f(y),$$

и значит, элементы x и y отображаются в один и тот же элемент группы G'. Обратно, если $f(u) = f(v)$, то

$$f(u^{-1}v) = f(u^{-1})f(v) = [f(u)]^{-1}[f(v)] = e'$$

и значит, $u^{-1}v \in H$, т. е. $v \in uH$, и элементы u и v принадлежат одному и тому же смежному классу.

Мы установили взаимно однозначное соответствие между элементами группы G' и смежными классами группы G по H. Покажем, что это соответствие является изоморфным.

Пусть x и y — элементы группы G, xH и yH — содержащие их классы, $x_1 = f(x)$ и $y_1 = f(y)$ — соответствующие им элементы группы G'. Классу xH фактор-группы G/H поставим в соответствие элемент $x_1 = f(x)$ группы G'. Тогда имеем

$$\begin{aligned} xH &\leftrightarrow x_1 = f(x), \\ yH &\leftrightarrow y_1 = f(y), \\ xyH &\leftrightarrow f(xy). \end{aligned}$$

Но $xyH = xH \cdot yH$, а $f(xy) = f(x)f(y)$, и значит,

$$xH \cdot yH \leftrightarrow f(x)f(y),$$

т. е. взаимно однозначное соответствие $\leftrightarrow$ между фактор-группой G/H и G' «сохраняет групповую операцию», и значит, группы G/H и G' **изоморфны**.

ГЛАВА XI

ГРУППЫ СИММЕТРИИ ГЕОМЕТРИЧЕСКИХ ФИГУР

В этой главе рассматриваются преобразования обычного трехмерного (евклидова) пространства R (определение преобразования см. в § 3 главы X).

§ 1. Группа движений вещественного евклидова пространства и ее подгруппы

Определение. *Движением называется такое преобразование вещественного евклидова пространства R, при котором расстояния между точками не меняются*: если точка P переходит в P', а точка Q — в Q', то расстояние $P'Q'$ равно PQ.

Такие преобразования образуют, очевидно, группу, называемую *группой движений* евклидова пространства, или *евклидовой группой.*

В группе движений пространства R выделим множество тех движений, при которых некоторая фиксированная точка O (начало координат) остается неподвижной, т. е. переходит сама в себя. Такие движения тоже, очевидно, образуют группу — подгруппу группы движений, называемую *центроевклидовой*, или *полной ортогональной группой.*

Каждому движению с неподвижной точкой O отвечает определенное преобразование соответствующего векторного пространства: если точка P переходит в P', то вектор $\overline{OP}$ переходит в вектор $\overline{OP'}$. При этом преобразовании длины векторов не меняются; легко видеть, что не меняются также и углы между векторами, т. е. рассматриваемое преобразование *сохраняет скалярное произведение векторов.*

Покажем, что *преобразование $\mathcal{A}$ евклидова векторного пространства, сохраняющее скалярное произведение*, т. е. такое, что $(\mathcal{A}x,\ \mathcal{A}y) = (x,\ y)$ при всех x,

$y \in R$, *является линейным* (и следовательно, ортогональным) *преобразованием*, т. е. что $\mathscr{A}(x+y) = \mathscr{A}x + \mathscr{A}y$ и $\mathscr{A}(\alpha x) = \alpha\mathscr{A}x$ при всех $x, y \in R$ и произвольном α. Действительно, имеем

$$\begin{aligned}(\mathscr{A}(x+y) - \mathscr{A}x - \mathscr{A}y,\ \mathscr{A}(x+y) - \mathscr{A}x - \mathscr{A}y) = \\ = (\mathscr{A}(x+y),\ \mathscr{A}(x+y)) - (\mathscr{A}x,\ \mathscr{A}(x+y)) - \\ - (\mathscr{A}y,\ \mathscr{A}(x+y)) - (\mathscr{A}(x+y),\ \mathscr{A}x) + (\mathscr{A}x,\ \mathscr{A}x) + \\ + (\mathscr{A}y,\ \mathscr{A}x) - (\mathscr{A}(x+y),\ \mathscr{A}y) + (\mathscr{A}x,\ \mathscr{A}y) + (\mathscr{A}y,\ \mathscr{A}y) = \\ = (x+y,\ x+y) - (x,\ x+y) - (y,\ x+y) - (x+y,\ x) + \\ + (x,\ x) + (y,\ x) - (x+y,\ y) + (x,\ y) + (y,\ y) = \\ = ((x+y) - x - y,\ (x+y) - x - y) = 0,\end{aligned}$$

откуда (поскольку пространство — евклидово)

$$\mathscr{A}(x+y) - \mathscr{A}x - \mathscr{A}y = 0 \quad \text{и} \quad \mathscr{A}(x+y) = \mathscr{A}x + \mathscr{A}y$$

Далее имеем

$$\begin{aligned}(\mathscr{A}(\alpha x) - \alpha\mathscr{A}x,\ \mathscr{A}(\alpha x) - \alpha\mathscr{A}x) = (\mathscr{A}(\alpha x),\ \mathscr{A}(\alpha x)) - \\ - \alpha(\mathscr{A}x,\ \mathscr{A}(\alpha x)) - \alpha(\mathscr{A}(\alpha x),\ \mathscr{A}x) + \alpha^2(\mathscr{A}x,\ \mathscr{A}x) = \\ = (\alpha x,\ \alpha x) - \alpha(x,\ \alpha x) - \alpha(\alpha x,\ x) + \alpha^2(x,\ x) = 0,\end{aligned}$$

и значит,

$$\mathscr{A}(\alpha x) - \alpha\mathscr{A}x = 0, \text{ т. е. } \mathscr{A}(\alpha x) = \alpha\mathscr{A}x$$

Таким образом, полная ортогональная группа — это группа всех (линейных) ортогональных преобразований векторного пространства.

Пусть теперь R — т р е х м е р н о е евклидово пространство. Напомним, что матрица ортогонального преобразования трехмерного пространства в некотором ортонормированном базисе приводится либо к виду

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} \tag{1}$$

(и тогда преобразование представляет собой п о в о р о т вокруг некоторой оси — «новой» оси x), либо к виду

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} \tag{2}$$

(в этом случае преобразование состоит из поворота вокруг новой оси x и симметрии относительно плоскости, перпендикулярной к этой оси; ясно, что если φ кратно $2\pi k$, то «поворот» представляет собой тождественное преобразование).

Определитель преобразования (1) первого рода (поворота вокруг какой-то оси, проходящей через начало координат), равен $+1$, а определитель преобразования (2) второго рода равен -1. Поэтому все преобразования первого рода (вращения) образуют подгруппу полной ортогональной группы, называемую *группой вращений* (трехмерного) пространства.

На плоскости матрица ортогонального преобразования приводится либо к виду

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}, \quad \text{либо к} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

В первом случае (преобразование первого рода) это — поворот вокруг начала координат на угол φ; определитель его матрицы равен $+1$. Во втором случае (преобразование второго рода) это — симметрия относительно некоторой прямой — «новой» оси x, определитель его матрицы равен -1. Все повороты плоскости образуют подгруппу ортогональной группы — группу вращений плоскости. Ясно, что все рассматриваемые нами пока группы (евклидова, ортогональная, группа вращений) — бесконечны.

Дальше нас будут интересовать различные конечные подгруппы полной ортогональной группы и группы вращений, которые могут быть получены следующим образом. Возьмем какую-нибудь (в определенном смысле «симметричную») фигуру плоскости или пространства и рассмотрим всевозможные ортогональные преобразования (или всевозможные вращения), переводящие эту фигуру в себя. Все такие преобразования, очевидно, образуют группу. Мы будем называть ее *группой симметрии* (соответственно *группой вращений*) рассматриваемой фигуры (в случае пространства иногда говорят не «фигура», а «тело»).

Понятие группы симметрии фигуры является математическим эквивалентом общежитейского представления о «симметричных» и «несимметричных» фигурах, которое само по себе точного смысла не имеет (поэтому выше мы заключили слово «симметричное» в кавычки): «степень симметричности» фигуры определяется «богатством» ее группы симметрии. Так, например, ясно, что ромб «менее симметричен», чем квадрат,— и группа симметрии ромба содержит всего 4 элемента, тогда как группа симметрии квадрата — 8 элементов. (Найдите сами все элементы группы симметрии квадрата; полное описание этой группы, входящей в серию диэдральных групп и обозначаемой через D_4, дается ниже, в § 4.)

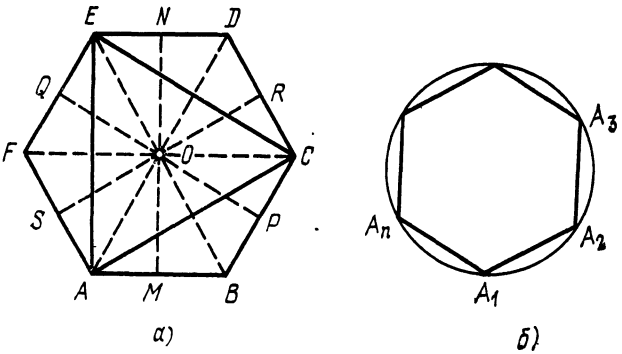


Рис. 28.

Далее, можно сказать, что правильный шестиугольник «более симметричен», чем правильный треугольник. Ведь правильный шестиугольник $ABCDEF$ переходит в себя при всех тех преобразованиях, при которых переходит в себя треугольник ACE (рис. 28, *а*) — при поворотах (вокруг центра) на углы 120° и 240° и симметриях относительно прямых AD, BE и CF, но, кроме того, шестиугольник переходит в себя еще и при поворотах на углы 60°, 180° и 300°, а также при симметриях относительно прямых MN, PQ и RS. Группа симметрии правильного треугольника D_3 содержит 6, а группа симметрии правильного шестиугольника D_6 — 12 элементов (см. ниже, § 4).

«Особенно симметричен» круг, он «гораздо более симметричен», чем любой (даже правильный) многоугольник. Действительно, круг переходит в себя при всех преобразованиях, при которых переходит в себя любой (вписанный в этот круг) правильный n-угольник, (рис. 28, *б*), поэтому группа симметрии круга бесконечна. Легко видеть, что *группа вращений круга изоморфна группе комплексных чисел, по модулю равных* 1. Действительно, круг переходит в себя при повороте (вокруг его центра) на любой угол α, где α достаточно брать в пределах от 0 до 2π, $0 \leqslant \alpha < 2\pi$. Повороту φ на угол α поставим в соответствие комплексное число (по модулю равное 1) $r_\varphi = \cos\alpha + i\sin\alpha$; тогда повороту ψ на угол β будет соответствовать число $r_\psi = \cos\beta + i\sin\beta$, а повороту $\varphi\psi$ на угол $\alpha+\beta$ — число $r_{\varphi\psi} = \cos(\alpha+\beta) + i\sin(\alpha+\beta)$, равное, очевидно, произведению $r_\varphi \cdot r_\psi$.

§ 2. Сопряженные элементы в группе вращений трехмерного пространства

Как известно, *осью* называют прямую, на которой задано определенное направление, или направленную прямую. Ось, проходящую через начало координат, можно задать некоторым вектором. Говоря о повороте вокруг оси l на угол φ, имеют в виду поворот в положительном направлении, т. е. против часовой стрелки, если смотреть с положительного направления оси l.

Пусть сначала G — произвольная группа линейных преобразований трехмерного векторного пространства и $\mathscr{A} \in G$. *Элемент $\mathscr{B}$ сопряжен с $\mathscr{A}$ в группе G, если найдется такое $\mathscr{C} \in G$, что $\mathscr{B} = \mathscr{C}^{-1}\mathscr{A}\mathscr{C}$* (и значит, $\mathscr{A} = \mathscr{C}\mathscr{B}\mathscr{C}^{-1}$). Выясним геометрический смысл понятия сопряженности. Для этого вернемся к равенству (4) на стр. 113. При $n = 3$ оно принимает вид

$$\mathscr{C}^{-1}\mathscr{A}\mathscr{C}e_i = a'_{1i}e_1 + a'_{2i}e_2 + a'_{3i}e_3, \quad i = 1, 2, 3,$$

и его можно интерпретировать следующим образом: преобразование $\mathscr{C}^{-1}\mathscr{A}\mathscr{C}$ в базисе e_1, e_2, e_3 имеет ту же матрицу, что преобразование $\mathscr{A}$ в базисе

$$e'_1 = \mathscr{C}e_1, \quad e'_2 = \mathscr{C}e_2, \quad e'_3 = \mathscr{C}e_3,$$

получающемся из базиса e_1, e_2, e_3 посредством преобразования $\mathscr{C}$.

Пусть теперь G — группа вращений (трехмерного) пространства. Тогда если $\mathscr{B}$ — поворот вокруг оси

l, определяемой вектором x (короче, поворот вокруг оси x) на угол φ, то сопряженный поворот $\mathscr{A} = \mathscr{C}\mathscr{B}\mathscr{C}^{-1}$ — это поворот вокруг оси $\mathscr{C}x$ на тот же самый угол φ. Оси x и $\mathscr{C}x$, где $\mathscr{C} \in G$, называются эквивалентными (относительно группы G).

Таким, образом, *сопряженные повороты в группе вращений трехмерного евклидова пространства — это повороты вокруг эквивалентных осей на один и тот же угол.*

Обратно, *если $\mathscr{B}$ есть поворот вокруг некоторой оси x,* а *$\mathscr{A}$ — поворот вокруг эквивалентной оси $\mathscr{C}x$ на один и тот же угол φ, то повороты $\mathscr{A}$ и $\mathscr{B}$ сопряжены между собой, так как в этом случае*

$$\mathscr{B} = \mathscr{C}^{-1}\mathscr{A}\mathscr{C}.$$

§ 3. Группа вращений правильного n-угольника C_n

Эта группа уже рассматривалась на стр. 279. *Правильный n-угольник* (рис. 29, *а*) совмещается сам с собой при n поворотах вокруг его центра на углы $0, \frac{2\pi}{n}, \frac{2\pi}{n}\cdot 2, \ldots, \frac{2\pi}{n}(n-1)$. Обозначим теперь эти повороты соответственно через $e, r, r^2, \ldots, r^{n-1}$. Образованная ими

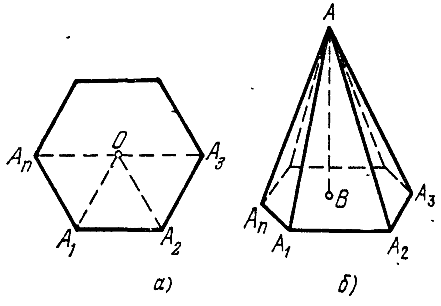


Рис. 29.

группа является циклической группой порядка n и обозначается символом C_n.

Легко видеть, что такова же и группа вращений правильной n-угольной пирамиды: такая пирамида переходит в себя при поворотах вокруг оси AB на те же углы $0, \frac{2\pi}{n}, \frac{2\pi}{n}\cdot 2, \ldots, \frac{2\pi}{n}(n-1)$ (рис. 29, *б*).

Группа C_n, коммутативна, и значит, число классов сопряженных элементов этой группы равно ее порядку.

§ 4. Диэдральные группы D_n

Диэдральную группу тоже можно интерпретировать по-разному. Например, можно определить ее как *группу симметрии* того же *правильного n-угольника*. Тогда к n поворотам вокруг центра на углы $0, \frac{2\pi}{n}, \frac{2\pi}{n}2, \ldots$ $\ldots, \frac{2\pi}{n}(n-1)$ добавятся n отражений относительно осей симметрии многоугольника. Покажем, что никаких других элементов эта группа не содержит, т. е. что она состоит в точности из $2n$ элементов.

Группа симметрии n-угольника D_n содержит, очевидно, подгруппу C_n, состоящую из всех поворотов вокруг центра. Пусть $e, r, r^2, \ldots, r^{n-1}$ — все элементы этой подгруппы (напомним, что r — поворот вокруг центра на угол $\frac{2\pi}{n}$). Обозначим через s отражение относительно одной какой-нибудь фиксированной оси симметрии многоугольника. Тогда $s^2 = e$. Пусть g — произвольный элемент группы D_n. Тогда g есть ортогональное преобразование, определитель которого равен, следовательно, $+1$ или -1. Если $|g| = +1$, то это — вращение, и значит, $g = r^k$, где $k = 0, 1, \ldots, n-1$. Если определитель $|g| = -1$, то определитель произведения sg равен $+1$ (так как определитель s равен -1), и значит, $sg \in C_n$, т. е. $sg = r^k$, откуда следует, что $g = sr^k$. Таким образом, *каждый элемент группы D_n может быть представлен либо в виде r^k, либо в виде sr^k*, где $k = 0, 1, \ldots, n-1$. Значит, *порядок группы D_n равен $2n$.*

Диэдральную группу D_n можно интерпретировать и иначе: ее можно рассматривать как группу вращений правильного n-угольника, но не в его плоскости, а в пространстве. Тогда r^k — это поворот вокруг оси, перпендикулярной плоскости многоугольника и проходящей через его центр, на угол $\frac{2\pi}{n}k$, а s — поворот в пространстве вокруг одной из осей симметрии многоугольника на угол π.

Наконец, диэдральную группу можно еще рассматривать как группу вращений правильного диэдра — правильной бипирамиды, состоящей из двух одинаковых правильных пирамид, сложенных своими основаниями (рис. 30). Тогда r^k — это поворот вокруг оси AB диэдра на угол $\frac{2\pi}{n}k$, а s — одно из «опрокидываний» — поворот на угол π вокруг одной из осей симметрии лежащего в основании диэдра многоугольника.

При $n = 1$ диэдр вырождается в отрезок, и группа D_1 изоморфна C_2. При $n = 2$ диэдр вырождается в ромб, и группа D_2 изоморфна группе симметрии ромба V. При $n = 3$ получается группа симметрии треугольника; легко видеть, что она изоморфна S_3 — группе подстановок трех его вершин. При $n \geqslant 3$ диэдральные группы некоммутативны.

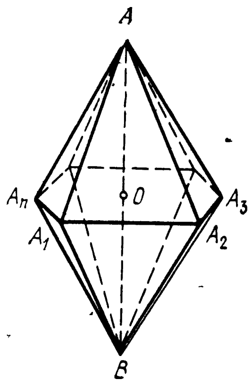


Рис. 30.

Найдем *классы сопряженных элементов* диэдральной группы.

При поворотах вокруг горизонтальных осей (осей симметрии многоугольника) вертикальная ось AB диэдра переходит в BA (опрокидывается); она является, как говорят, двусторонней осью — ось AB эквивалентна BA. Следовательно, поворот вокруг оси AB на угол α сопряжен с поворотом вокруг оси BA на тот же угол α, т. е. с поворотом вокруг оси AB на угол $-\alpha$. Таким образом, повороты r^k и r^{n-k} сопряжены между собой, причем, очевидно, имеет место равенство

$$sr^ks = r^{n-k}.$$

(Проверьте сами это равенство, выписав и перемножив соответствующие матрицы!) Никакой другой оси ось AB не эквивалентна, и поворот r^k сопряжен только с r^{n-k}.

Следовательно, при нечетном n повороты вокруг оси AB разбиваются на классы сопряженных элементов следующим образом:

$$\{e\},\ \{r,\ r^{n-1}\},\ \{r^2,\ r^{n-2}\},\ \ldots,\ \left\{r^{\frac{n-1}{2}},\ r^{\frac{n+1}{2}}\right\}.$$

Число таких классов равно $\frac{n+1}{2}$. При четном n повороты вокруг оси AB разбиваются на классы:

$$\{e\},\ \{r,\ r^{n-1}\},\ \{r^2,\ r^{n-2}\},\ \ldots,\ \left\{r^{\frac{n}{2}-1},\ r^{\frac{n}{2}+1}\right\},\ \left\{r^{\frac{n}{2}}\right\}.$$

Число этих классов равно $\frac{n}{2}+1$.

Далее, при нечетном n все горизонтальные оси вращения эквивалентны между собой (рис. 31, a) — они

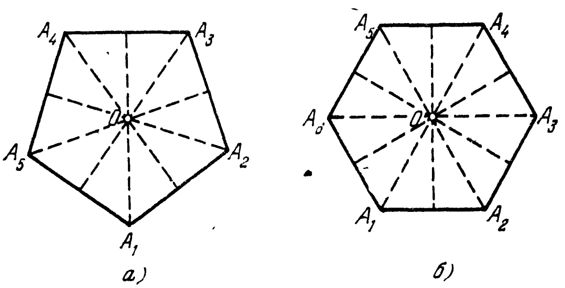


Рис. 31.

переходят друг в друга при поворотах вокруг вертикальной оси AB. А так как углы поворотов вокруг всех этих осей одинаковы — они равны π, то все эти повороты сопряжены между собой и образуют один класс сопряженных элементов.

При четном n многоугольник имеет оси симметрии двух типов: диагонали и прямые, соединяющие середины противоположных сторон (рис. 31, $б$). Все первые оси между собой эквивалентны. Все вторые — тоже, но первые во вторые не переводятся никаким вращением — они не эквивалентны. Значит, при четном n повороты вокруг горизонтальных осей образуют два класса сопряженных элементов: $\{s, sr^2, sr^4, \ldots, sr^{n-2}\}$ и $\{sr, sr^3, \ldots$ $\ldots, sr^{n-1}\}$. (Действительно, имеем: $r^{-1}(sr^k)r = r^{n-1}(sr^k)r = (r^{n-1}s)r^{k+1} = (sr)r^{k+1} = sr^{k+2}$ (см. равенство на стр. 311) — и значит sr^k сопряжено с sr^{k+2} при всех k.)

Таким образом, общее *число классов сопряженных элементов в группе D_n при нечетном n равно* $\frac{n+3}{2}$, *а при четном n оно равно* $\frac{n}{2}+3$. Так, группа D_3 имеет 3 класса сопряженных элементов, группа D_4 имеет 5 классов,

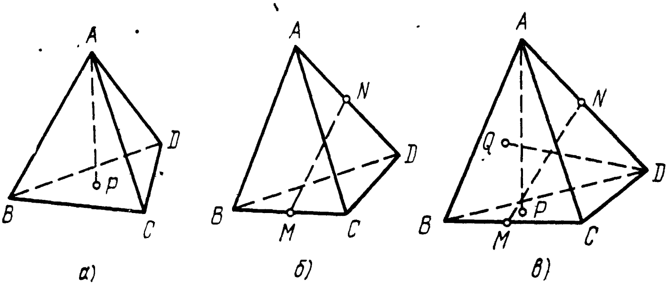


Рис. 32.

группа D_5 имеет 4 класса и группа D_6 имеет 6 классов сопряженных элементов.

Полезно еще заметить, что диэдральная группа D_n порождается двумя «образующими»: элементами r и s со связывающими их «определяющими соотношениями»

$$r^n = e, \quad s^2 = e, \quad srs = r^{n-1}.$$

Все остальные соотношения между элементами этой группы вытекают из «определяющих соотношений». Так, например, $sr^2s = srs^2rs = (srs)(srs) = r^{n-1} \cdot r^{n-1} = r^{n-2}$, и т. д.

§ 5. Группа вращений тетраэдра T

Рассмотрим *правильный тетраэдр $ABCD$*. Он переходит в себя при следующих нетождественных поворотах:

а) При поворотах вокруг каждой из осей типа AP (рис. 32, a), соединяющих вершину тетраэдра с центром противолежащей грани, на углы $\frac{2\pi}{3}$ и $\frac{4\pi}{3}$. Всего таких вращений имеется $4 \times 2 = 8$.

б) При поворотах на угол π вокруг каждой из трех прямых типа MN, соединяющих середины противоположных ребер (рис. 32, *б*). [Так как $MN \perp BC$, $MN \perp AD$, $BM = MC$ и $AN = ND$, то при повороте вокруг прямой MN на угол π точка B перейдет в C, C — в B, A — в D и D — в A.]

Всего, вместе с тождественным поворотом, мы имеем $1 + 8 + 3 = 12$ поворотов, при которых тетраэдр переходит в себя. Им отвечают, очевидно, такие подстановки вершин:

$$\begin{pmatrix} ABCD \\ ABCD \end{pmatrix}, \begin{pmatrix} ABCD \\ ACDB \end{pmatrix}, \begin{pmatrix} ABCD \\ ADBC \end{pmatrix}, \begin{pmatrix} ABCD \\ CBDA \end{pmatrix}, \begin{pmatrix} ABCD \\ DBAC \end{pmatrix}, \begin{pmatrix} ABCD \\ BDCA \end{pmatrix},$$
$$\begin{pmatrix} ABCD \\ DACB \end{pmatrix}, \begin{pmatrix} ABCD \\ BCAD \end{pmatrix}, \begin{pmatrix} ABCD \\ CABD \end{pmatrix}, \begin{pmatrix} ABCD \\ BADC \end{pmatrix}, \begin{pmatrix} ABCD \\ CDAB \end{pmatrix}, \begin{pmatrix} ABCD \\ DCBA \end{pmatrix}.$$

Нетрудно убедиться в том, что все эти подстановки — четные (проверьте это!), и значит, соответствующие им повороты действительно образуют группу T, изоморфную, очевидно, знакопеременной подгруппе A_4 симметрической группы S_4.

Условимся о такой терминологии. Если данная конфигурация переходит в себя при повороте вокруг оси l на угол $\frac{2\pi}{k}$ (причем $\frac{2\pi}{k}$ — это н а и м е н ь ш и й такой ненулевой угол), то ось l будем называть *осью симметрии k-го порядка.*

Поворот вокруг оси l на угол $\frac{2\pi}{k}$ будем обозначать символом c_k (часто и сама эта ось обозначается через c_k), поворот на угол $\frac{2\pi}{k} \cdot 2$ тогда естественно обозначить через c_k^2, и т. д.

Найдем теперь *классы сопряженных элементов* группы T. Каждая из осей симметрии третьего порядка может быть преобразована в любую другую ось третьего порядка при повороте, например, вокруг одной из осей второго порядка. Так, при повороте вокруг оси MN (рис. 32, *в*) точка A переходит в D, B — в C, C — в B и D в A. Плоскость BCD переходит в плоскость BCA, центр P грани BCD — в центр Q грани BCA и ось AP — в ось DQ.

Таким образом, все оси третьего порядка (типа AP) эквивалентны между собой, и все повороты вокруг них на углы $\frac{2\pi}{3}$ между собой сопряжены. Число таких поворотов равно 4, и соответствующий классс сопряженных элементов можно обозначить через $\{4c_3\}$. Точно так же сопряжены между собой и 4 поворота вокруг тех же осей на углы $\frac{4\pi}{3}$; соответствующий класс можно обозначить через $\{4c_3^2\}$. Но повороты c_3 и c_3^2 не сопряжены между собой, так как это — повороты на разные углы.

Далее, каждая из осей второго порядка (типа MN) переходит в любую другую при одном из поворотов вокруг осей третьего порядка; значит, все оси второго порядка между собой эквивалентны, и три поворота вокруг этих осей на угол π между собой сопряжены. Этот класс можно обозначить через $\{3c_2\}$.

Учитывая, что тождественное преобразование составляет отдельный класс, мы получим в группе T *четыре класса сопряженных элементов*, состоящих из одного $\{e\}$, четырех $\{4c_3\}$, четырех $\{4c_3^2\}$ и трех $\{3c_2\}$ элементов.

§ 6. Группа вращений куба O

Легко видеть, что *куб* переходит в себя при следующих нетождественных вращениях:

а) При трех поворотах на углы $\pi/2$, π и $3\pi/2$ вокруг каждой из трех прямых типа MN (рис. 33, *а*), соединяющих центры противоположных граней (оси симметрии четвертого порядка). Всего таких поворотов $3 \times 3 = 9$.

б) При двух поворотах вокруг каждой из четырех диагоналей (осей симметрии третьего порядка, рис. 33, *б*) на углы $2\pi/3$ и $4\pi/3$ (правильный треугольник ACD' при этом переходит в себя). Всего таких поворотов $2 \times 4 = 8$.

в) При шести поворотах на угол π — вокруг каждой из прямых типа PQ (рис. 33, *в*), соединяющих середины противоположных ребер (оси симметрии второго порядка).

Всего, вместе с тождественным преобразованием, мы нашли $1+9+8+6=24$ поворота, при которых куб переходит в себя. Из доказываемой ниже теоремы будет вытекать, что это — все вращения, при которых куб переходит в себя.

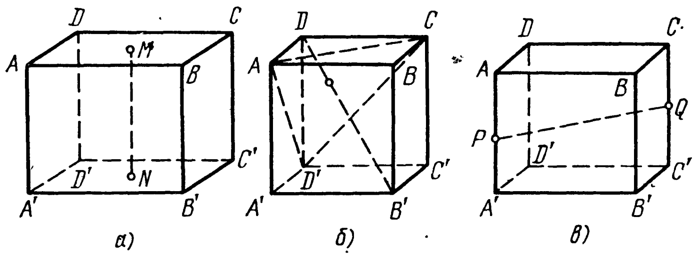


Рис. 33.

Теорема. *Группа O вращений куба изоморфна симметрической группе* S_4 (и значит, порядок этой группы равен 24).

Доказательство. При каждом повороте, при котором куб переходит в себя, каждая его диагональ переходит в одну из диагоналей. У куба 4 диагонали, поэтому каждому вращению куба отвечает определенная подстановка его диагоналей, а произведению вращений — произведение соответствующих подстановок.

Остается доказать, что *разным вращениям куба отвечают разные подстановки диагоналей.* Действительно, если два разных вращения α и β куба приводят к одной и той же подстановке диагоналей, то при (нетождественном) повороте $\alpha\beta^{-1}$ каждая диагональ куба переходит в себя (хотя, возможно, меняются местами концы некоторых диагоналей). Покажем, что такое вращение, при котором каждая диагональ куба переходит в себя, является тождественным.

Предположим, что при повороте γ все диагонали куба перешли в себя. В частности, перейдут в себя диагонали DB' и BD' (рис. 34), а тогда перейдет в себя и содержащая их плоскость $DBB'D'$. Значит, ось этого вращения либо лежит в плоскости $DBB'D'$, причем поворот этот — на угол π, либо к ней перпендикулярна. Но в первом случае переходят в себя только прямые, направленные по оси вращения, и прямые, перпендикулярные к оси. Однако прямоугольник $DBB'D'$ — не квадрат, и значит, его диагонали не перпендикулярны друг другу. Во втором случае, т. е. если ось вращения перпендикулярна плоскости $DBB'D'$, она совпадает с прямой PQ, а тогда не переходят в себя (переставляются) диагонали AC' и $A'C$.

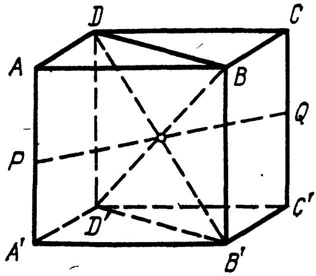


Рис. 34.

Таким образом, группа вращений куба изоморфна симметрической группе S_4.

Найдем теперь, как элементы группы O разбиваются на *классы сопряженных элементов*.

Три оси симметрии четвертого порядка, очевидно, эквивалентны, и значит, повороты вокруг них на углы $\pi/2$ сопряжены между собой. Далее, эти оси являются д в у с т о р о н н и м и (опрокидываются при поворотах на угол π вокруг других осей четвертого порядка), и значит, повороты вокруг них на углы $3\pi/2$ тоже сопряжены не только между собой, но и с поворотами на углы $\pi/2$. Поворот на угол $\pi/2$ можно обозначить через c_4, поворот на угол $3\pi/2$ — через c_4^3. Мы нашли класс, состоящий из шести элементов, который можно обозначить символом $\{3c_4, 3c_4^3\}$, или даже, короче,— символом $\{6c_4\}$.

Все повороты вокруг тех же осей четвертого порядка на углы $\pi/2 \cdot 2 = \pi$ сопряжены между собой (и только между собой). Число таких поворотов равно 3, соответствующий класс можно обозначить через $\{3c_4^2\}$.

Далее, все оси третьего порядка (диагонали) между собой эквивалентны (переходят друг в друга, например, при поворотах вокруг осей четвертого порядка). При этом каждая диагональ является двусторонней осью (опрокидывается при поворотах вокруг перпендикулярных к ней осей второго порядка. Значит, все 8 поворотов вокруг диагоналей на углы $2\pi/3$ и $4\pi/3$ сопряжены между собой. Соответствующий класс можно обозначить через $\{4c_3, 4c_3^2\}$, или просто через $\{8c_3\}$.

Наконец, шесть осей второго порядка переходят друг в друга, например, при поворотах вокруг осей четвертого порядка, и значит, все шесть поворотов вокруг них на угол π сопряжены между собой. Этот класс можно обозначить $\{6c_2\}$.

Учитывая отдельный класс, образуемый тождественным преобразованием, получаем всего пять классов сопряженных элементов, состоящих из одного $\{e\}$, шести $\{6c_4\}$, трех $\{3c_4^2\}$, восьми $\{8c_3\}$ и шести $\{6c_2\}$ элементов.

§ 7. Группа симметрии тетраэдра T_d

Кроме семи осей симметрии *правильный тетраэдр* имеет шесть плоскостей симметрии. К 12 вращениям, при которых тетраэдр переходит в себя (и которые отвечают, как мы видели, четным подстановкам его вершин), добавим одну из симметрий, например, симметрию относительно плоскости ADM (рис. 35) — ей соответствует (нечетная) подстановка $\begin{pmatrix} ABCD \\ ACBD \end{pmatrix}$ вершин тетраэдра. Если умножить эту симметрию на каждый из 12 поворотов, при которых тетраэдр переходит в себя, мы получим еще 12 преобразований, отвечающих нечетным подстановкам вершин. Среди них будут 6 «чистых» симметрий и 6 произведений поворота и симметрии. Кроме этих 24 преобразований, не существует никаких ортогональных преобразований, при которых тетраэдр $ABCD$ переходит в себя (в частности — никаких вращений, кроме рассмотренных в § 5), так как каждое такое преобразование отвечает определенной подстановке его вершин и, значит, совпадает с одним из уже определенных преобразований. Таким об-

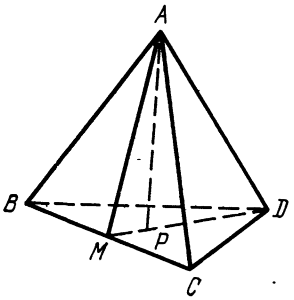


Рис. 35.

разом, *группа симметрии тетраэдра T_d изоморфна симметрической группе S_4* и, значит, она, *изоморфна группе вращений куба O*. Поэтому эта группа тоже состоит из пяти классов сопряженных элементов, содержащих 1, 6, 3, 8 и 6 элементов. Найдем, как распределяются элементы группы T_d по этим классам.

В группе T_d класс из трех элементов образуют, очевидно, повороты вокруг осей второго порядка $\{3c_2\}$. Класс из 8 элементов состоит из всех поворотов вокруг осей третьего порядка: $\{4c_3, 4c_3^2\}$; повороты вокруг осей третьего порядка на углы $2\pi/3$ и $4\pi/3$ в группе T не сопряжены, а в группе T_d они оказываются сопряженными, так как если s — симметрия, скажем, относительно плоскости AMD, а r — поворот относительно оси AP, лежащей в этой плоскости на угол α, то srs есть поворот вокруг той же самой оси AP на угол $-\alpha$. В нашем случае это доказывается следующим равенством:

$$\begin{pmatrix} ABCD \\ ACBD \end{pmatrix}\begin{pmatrix} ABCD \\ ACDB \end{pmatrix}\begin{pmatrix} ABCD \\ ACBD \end{pmatrix} = \begin{pmatrix} ABCD \\ ADBC \end{pmatrix},$$

и значит, $srs = r^{-1}$.

Далее, 6 симметрий относительно плоскостей вида ADM, очевидно, сопряжены между собой (эти плоскости «эквивалентны» — при поворотах переходят друг в друга), они образуют отдельный класс; обозначим его $\{6\sigma\}$. Остальные 6 преобразований — произведения поворота и симметрии тоже, следовательно, образуют отдельный класс $\{6\sigma'\}$.

§ 8. Группа симметрии куба O_h

Кроме 13 осей симметрии *куб* имеет 9 плоскостей симметрии (и центр симметрии): три плоскости симметрии — такие, как $PQMN$ на рис. 36, *а*, и шесть диагональных плоскостей — таких, как $DBB'D'$ на рис. 36, *б*.

Рассмотрим всевозможные ортогональные преобразования пространства, при которых куб переходит в себя. Те из этих преобразований, определитель которых равен 1 (вращения), образуют подгруппу O (изоморфную, как мы видели, симметрической группе S_4). Пусть g — одно из преобразований симметрии куба с определителем, равным — 1. Умножив его на центральную симметрию j (преобразование с определителем, равным — 1), мы получим преобразование a с определителем, равным + 1, т. е. вращение. Из равенства $jg = a$ имеем $g = ja$. Таким образом, все преобразования, при которых куб переходит в себя,— это всевозможные вращения и всевозможные произведения вида ja, где j — центральная симметрия, а a — вращение.

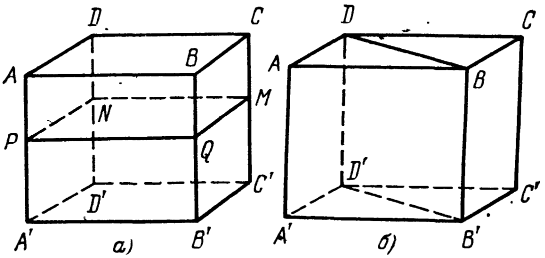


Рис. 36.

Покажем, что *центральная симметрия j перестановочна с любым вращением* и, даже, более того,— *с любым линейным преобразованием* $\mathcal{A}$. Действительно, для любого вектора x имеем $jx = -x$, и, значит,

$$\mathcal{A}(jx) = \mathcal{A}(-x) = -\mathcal{A}x = j(\mathcal{A}x), \text{ т. е. } \mathcal{A}j = j\mathcal{A}.$$

Тождественное преобразование e и центральная симметрия j в группе O_h образуют циклическую подгруппу второго порядка $J = \{e, j\}$ $(j^2 = e)$. Покажем, что *группа O_h равна прямому произведению своих подгрупп O и J.*

1) Подгруппы O и J в группе O_h являются нормальными: O как подгруппа индекса 2 (ср. выше стр. 293), а J — как подгруппа, оба элемента которой коммутируют со всеми элементами группы O_h.

2) Пересечение подгрупп O и J состоит из одного единичного элемента.

3) Каждый элемент группы O_h представляется в виде произведения элемента из O и элемента из J: если $a_1 = e$, a_2, ..., a_{24} — все 24 элемента группы O, то

элементы группы O_h — это

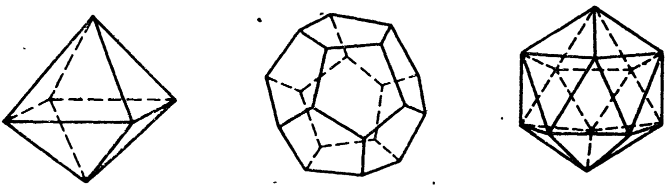
$$ea_1, ea_2, \ldots, ea_{24} \text{ и } ja_1, ja_2, \ldots, ja_{24}.$$

Группа O_h состоит из 48 элементов. Так как в группе O — пять, а в группе J — два класса сопряженных элементов, то число классов сопряженных элементов группы O_h равно 10. Если $\{a_{i_1}, a_{i_2}, \ldots, a_{i_q}\}$ — какой-то класс сопряженных элементов группы O, то в группе O_h ему соответствуют два класса:

$$\{a_{i_1}, a_{i_2}, \ldots, a_{i_q}\} \text{ и } \{ja_{i_1}, ja_{i_2}, \ldots, ja_{i_q}\}.$$

§ 9. Заключение

Правильным многогранником называется такой (выпуклый) многогранник, все грани которого — равные между собой правильные многоугольники и все многогранные углы которого равны между собой. Еще

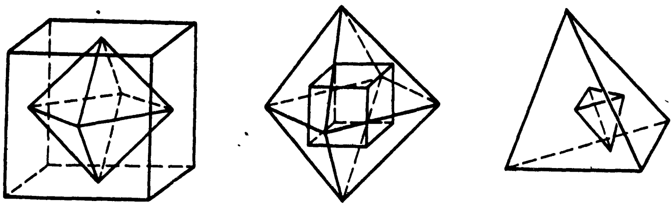


а) б) в)

Рис. 37.

в древности было известно, что существует всего пять правильных многогранников: *правильный тетраэдр*, ограниченный четырьмя правильными треугольными гранями, и имеющий 6 ребер и 4 вершины; *куб* (или *правильный гексаэдр*), ограниченный шестью квадратными гранями и имеющий 12 ребер и 8 вершин; *правильный октаэдр*, ограниченный восемью треугольными гранями и имеющий 12 ребер и 6 вершин (рис. 37, *а*); *правильный додекаэдр*, ограниченный двенадцатью пятиугольными гранями и имеющий 30 ребер и 20 вершин (рис. 37, *б*); и, наконец, *правильный икосаэдр*, ограниченный двадцатью треугольными гранями и имеющий 30 ребер и 12 вершин (рис. 37, *в*).

Куб и правильный октаэдр в определенном смысле *двойственны друг другу*: если соединить центры граней куба, как указано на рис. 38, *а*, то получится правильный октаэдр, и, наоборот, если соединить центры граней



а) *б)* *в)*

Рис. 38.

октаэдра, то получится куб (рис. 38, *б*). Поэтому *группа вращений октаэдра изоморфна группе вращений куба* (и обозначается эта группа буквой O), а группа симметрии октаэдра — группе симметрии куба O_h.

В описанном смысле *правильный тетраэдр двойствен сам себе* (рис. 38, *в*); *правильные же додекаэдр и икосаэдр двойственны друг другу, и их группы вращений изоморфны* — обозначается эта группа буквой I. Она состоит из 60 элементов и изоморфна знакопеременной подгруппе A_5 симметрической группы S_5 пятой степени.

Группа симметрии додекаэдра (икосаэдра) I_h состоит из 120 элементов и является прямым произведением группы I и циклической группы второго порядка.

Кроме перечисленных пяти правильных многогранников, в пространстве существует еще «вырожденный правильный многогранник» — правильный многоугольник, который можно рассматривать как многогранник, состоящий из двух равных правильных n-угольников (двугранник). Циклическая группа C_n есть группа вращений правильного n-угольника, т. е. вырожденного правильного многогранника в содержащей его плоскости, а диэдральная группа D_n — это группа его вращений в пространстве.

Существует такая общая теорема: *Циклические группы* C_n, $n = 1, 2, \ldots$, *диэдральные группы* D_n, $n = 1, 2, \ldots$, *группа вращений тетраэдра* T, *октаэдра* (*куба*) O *и икосаэдра* (*додекаэдра*) I — *это все конечные подгруппы группы вращений трехмерного пространства вокруг неподвижной точки.* «Это,— пишет Г. Вейль,— и есть современный эквивалент того перечня правильных многогранников, который дали древние греки»*).

Полный список всех конечных подгрупп группы ортогональных преобразований содержит, кроме определенных выше групп симметрии многогранников, еще несколько отдельных групп и серий групп; в этой книге они не рассматриваются **).

---

*) Г. Вейль, Симметрия (М.: Наука, 1968), стр. 105.
**) См., например, ту же книгу Г. Вейля.

ГЛАВА XII

ЛИНЕЙНЫЕ ПРЕДСТАВЛЕНИЯ КОНЕЧНЫХ ГРУПП

§ 1. Определения и примеры

Пусть дана группа G, гомоморфная некоторой другой группе G^*. Тогда, по теореме о гомоморфизмах, группа G^* изоморфна фактор-группе группы G по некоторой нормальной подгруппе H. Следовательно, группа G^* «устроена» в некотором смысле проще, чем группа G; в частности, если группа G — конечного порядка, то порядок группы G^* меньше порядка группы G (или равен ему). С другой стороны, группа G^* «подобна» группе G: так, если f — гомоморфизм, отображающий группу G на G^*, то из того, что $ab = c$ (где $a, b, c \in G$), следует, что $f(a)f(b) = f(c)$ (где $f(a), f(b), f(c) \in G^*$).

Говорят, что *группа G^* представляет группу G*, или, точнее, что *гомоморфное отображение f группы G на группу G^* является представлением группы G (группой G^*)*. Существует теорема Кэли, в силу которой *каждая конечная группа порядка n изоморфна* (а изоморфизм есть частный случай гомоморфизма!) *некоторой подгруппе группы подстановок из n элементов.* В этом случае группа G^* устроена в точности так, как группа G, что позволяет назвать это представление группы G группой G^* *точным представлением.* Следовательно, *каждая конечная группа может быть точно представлена некоторой группой подстановок.*

Наиболее интересны для теории и приложений так называемые линейные представления групп.

Говоря о линейном представлении группы G, мы предполагаем, что нам дано (вообще говоря, комплексное) векторное пространство R размерности n, в котором действуют невырожденные линейные операторы. Эти операторы образуют группу G^*, которой гомоморфна наша группа G — группа G^* и представляет группу G. Итак, можно дать следующее

Определение 1. *Гомоморфное отображение* Γ *группы* G *на группу* G^* *невырожденных линейных операторов, действующих в* n*-мерном векторном (комплексном) пространстве* R, *называется* ***линейным представлением группы*** G (группой G^*).

Таким образом, если Γ есть линейное представление группы G группой G^*, то каждому элементу a группы G поставлен в соответствие невырожденный линейный оператор $\Gamma(a) \in G^*$, действующий в пространстве R, так, что для любых $a, b \in G$

$$\Gamma(ab) = \Gamma(a) \cdot \Gamma(b).$$

При этом, как мы знаем, $\Gamma(e) = E$, где e — единица группы G, а E — единица группы G^* (тождественный оператор) и $\Gamma(a^{-1}) = [\Gamma(a)]^{-1}$ для любого $a \in G$ (см. выше стр. 302).

Пространство R, *в котором действуют операторы из группы* G^*, *называется* ***пространством представления*** группы G. Иногда и само это пространство называют *представлением* группы G. Размерность пространства R называется *размерностью*, или, чаще, ***степенью***, рассматриваемого представления.

В приложениях вместо операторов часто рассматривают соответствующие им матрицы. Если в пространстве R выбрать базис, то каждому линейному оператору $\Gamma(a)$ будет отвечать определенная матрица, т. е. каждому элементу a группы G будет поставлена в соответствие (невырожденная) квадратная матрица $\Gamma(a)$*) порядка n так, что

$$\Gamma(ab) = \Gamma(a)\Gamma(b).$$

Если пространство R одномерно, то эти матрицы — первого порядка. В этом случае каждому элементу a группы G поставлено в соответствие (вообще говоря, комплексное) *отличное от нуля число* $\Gamma(a)$, так что

$$\Gamma(ab) = \Gamma(a)\Gamma(b).$$

При этом единичному элементу e группы отвечает число 1.

---

*) Здесь и дальше мы обозначаем оператор $\Gamma(a)$ и соответствующую ему в заданном базисе матрицу одной и той же буквой.

Заметим, что если Γ — одномерное представление группы G и элементы a и b сопряжены в $G: b = c^{-1}ac$, то

$$\Gamma(b) = \Gamma(c^{-1}ac) = \Gamma(c^{-1})\Gamma(a)\Gamma(c) = [\Gamma(c)]^{-1}\Gamma(a)\Gamma(c) = \Gamma(a).$$

Тривиальным, но важным для теории примером может служить одномерное представление (т. е. представление степени 1), при котором *каждому элементу a группы G поставлено в соответствие число* 1, так что $\Gamma(a) = 1$ для каждого $a \in G$. Такое представление называется *единичным* представлением группы G.

Если группа G изоморфна группе G^*, то представление Γ группы G группой G^* называется *точным*; в противном случае представление Γ, по определению, *неточное*.

Если G есть группа линейных операторов, то она сама является одним из своих линейных представлений (причем, очевидно, точным); это представление называют *основным* представлением группы G.

Рассмотрим несколько примеров.

1. Найдем все *одномерные представления циклической группы C_2 второго порядка.* Эта группа состоит из двух элементов e и a, причем $a^2 = e$. Пусть Γ будет одномерное представление группы C_2. Тогда $\Gamma(e) = 1$. Предположим, что $\Gamma(a) = \alpha$; тогда $\Gamma(a^2) = (\Gamma(a))^2 = \alpha^2$. Но так как $a^2 = e$, то $\Gamma(a^2) = \Gamma(e) = 1$, и значит, $\alpha^2 = 1$, т. е. $\alpha = \pm 1$. Это дает два одномерных представления Γ_1 и Γ_2 группы C_2:

C_2	e	a
Γ_1	1	1
Γ_2	1	-1

Здесь первое представление: $\Gamma_1(e) = 1$, $\Gamma_1(a) = 1$ — единичное (неточное); второе: $\Gamma_2(e) = 1$, $\Gamma_2(a) = -1$ — точное.

2. Найдем все *одномерные представления группы C_4 — циклической группы четвертого порядка.* Эта груп-

па состоит из четырех элементов e, a, a^2, a^3, причем $a^4 = e$. Пусть $\Gamma(a) = \alpha$. Тогда $\Gamma(a^2) = \alpha^2$, $\Gamma(a^3) = \alpha^3$, $\Gamma(a^4) = \alpha^4$, и значит, $\alpha^4 = 1$, т. е. $\alpha = \sqrt[4]{1}$. Поэтому α может равняться 1, i, -1, $-i$, что дает четыре одномерных представления:

C_4	e	a	a^2	a^3
Γ_1	1	1	1	1
Γ_2	1	-1	1	-1
Γ_3	1	i	-1	$-i$
Γ_4	1	$-i$	-1	i

первое из которых — единичное; два последних представления являются точными.

3. Пусть V — *группа симметрии ромба*. Она состоит из четырех элементов e, a, b и $ab = ba$, причем $a^2 = b^2 = e$. Если Γ — одномерное представление этой группы и $\Gamma(a) = \alpha$, $\Gamma(b) = \beta$, то $\Gamma(a^2) = \alpha^2 = 1$ и $\Gamma(b^2) = \beta^2 = 1$, т. е. $\alpha = \pm 1$ и $\beta = \pm 1$. Это дает четыре одномерных представления (ни одно из которых не является точным):

V	e	a	b	ab
Γ_1	1	1	1	1
Γ_2	1	1	-1	-1
Γ_3	1	-1	1	-1
Γ_4	1	-1	-1	1

4. Найдем *одномерные представления диэдральной группы* $D_3 \simeq S_3$*). Элементы группы: e, r, r^2, s, sr, sr^2. Пусть Γ — одномерное представление D_3; поскольку r и r^2 сопряжены между собой, если $\Gamma(r) = \alpha$, то $\Gamma(r^2) = \alpha^2 = \Gamma(r) = \alpha$, т. е. $\alpha^2 = \alpha$, — и так как $\alpha \neq 0$, то $\alpha = 1$. Далее, если $\Gamma(s) = \beta$, то $\Gamma(s^2) = \beta^2 = 1$, и $\beta = \pm 1$. Это дает два одномерных представления группы D_3:

D_3	e	r	r^2	s	sr	sr^2
Γ_1	1	1	1	1	1	1
Γ_2	1	1	1	−1	−1	−1

или, короче (поскольку элементы r и r^2, а также элементы s, sr и sr^2 сопряжены между собой):

D_3	e	r, r^2	s, sr, sr^2
Γ_1	1	1	1
Γ_2	1	1	−1

5. Построим *двумерное представление группы* D_3. Эта группа изоморфна группе симметрии правильного треугольника, т. е. D_3 есть группа преобразований плоскости, а значит, она сама является одним из своих представлений. Найдем матрицы этого (основного) представления.

Пусть ABC — правильный треугольник с центром O (рис. 39). Совместим точку O с началом координат, а вершину A треугольника расположим на положительной стороне оси Ox. Обозначим через r поворот вокруг центра треугольника на угол $2\pi/3$, тогда r^2 — поворот на

*) $\simeq$ — знак изоморфизма групп.

угол $4\pi/3$; через s обозначим симметрию относительно оси Ox. Соответствующие матрицы будут иметь вид

$$\Gamma(e) = E, \quad \Gamma(r) = \begin{bmatrix} \cos\frac{2\pi}{3} & -\sin\frac{2\pi}{3} \\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix},$$

$$\Gamma(r^2) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}, \quad \Gamma(s) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\Gamma(sr) = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}, \quad \Gamma(sr^2) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}.$$

[$\Gamma(sr)$ и $\Gamma(sr^2)$ находятся перемножением матриц: $\Gamma(sr) = \Gamma(s)\cdot\Gamma(r)$ и $\Gamma(sr^2) = \Gamma(s)\cdot\Gamma(r^2)$.]

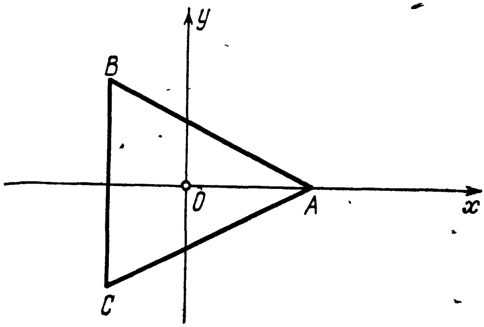


Рис. 39.

6. Диэдральную группу D_3 можно интерпретировать и как *группу движений трехмерного пространства* — как группу вращений диэдра. Если лежащий в основании диэдра треугольник расположить как в примере 5, а ось Oz направить перпендикулярно к плоскости треугольника и если r — это поворот вокруг оси Oz на угол $2\pi/3$, а s — поворот вокруг оси Ox на угол π, то

соответствующие матрицы будут иметь вид

$$\Gamma(e)=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\quad \Gamma(r)=\begin{bmatrix}-\frac{1}{2}&-\frac{1}{2}\sqrt{3}&0\\ \frac{1}{2}\sqrt{3}&-\frac{1}{2}&0\\0&0&1\end{bmatrix},$$

$$\Gamma(r^2)=\begin{bmatrix}-\frac{1}{2}&\frac{1}{2}\sqrt{3}&0\\ -\frac{1}{2}\sqrt{3}&-\frac{1}{2}&0\\0&0&1\end{bmatrix},$$

$$\Gamma(s)=\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix},\quad \Gamma(sr)=\begin{bmatrix}-\frac{1}{2}&-\frac{1}{2}\sqrt{3}&0\\ -\frac{1}{2}\sqrt{3}&\frac{1}{2}&0\\0&0&-1\end{bmatrix},$$

$$\Gamma(sr^2)=\begin{bmatrix}-\frac{1}{2}&\frac{1}{2}\sqrt{3}&0\\ \frac{1}{2}\sqrt{3}&\frac{1}{2}&0\\0&0&-1\end{bmatrix}.$$

Как будет показано дальше (в § 3 гл. XIII), полученное трехмерное представление группы D_3 в некотором смысле «хуже» найденных выше одномерных (пример 1) и двумерного (пример 5) представлений этой группы (оно распадается в «прямую сумму» одномерных и двумерных представлений).

§ 2. Изоморфные представления

Определение 2. *Пусть Γ_1 и Γ_2 — два представления группы G в пространствах R_1 и R_2 соответственно, причем размерности пространств R_1 и R_2 одинаковы,* т. е. пространства R_1 и R_2 изоморфны (см. § 5 главы II). *Представления Γ_1 и Γ_2* называются ***изоморфными*** *(эквивалентными, подобными), если $\mathscr{H}\Gamma_1=\Gamma_2\mathscr{H}$, где $\mathscr{H}$ — изоморфное отображение пространства R_1 на R_2.* (Ясно, что изоморфные представления имеют одинаковые степени.)

Равенство $\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H}$ означает, что $\mathscr{H}\Gamma_1(a) = \Gamma_2(a)\mathscr{H}$ для любого элемента $a \in G$, а это, в свою очередь, означает, что для любого вектора $x \in R_1$

$$\mathscr{H}\Gamma_1(a)x = \Gamma_2(a)\mathscr{H}x. \tag{1}$$

Поясним «геометрический смысл» последнего равенства. Пусть $x \in R_1$ и $y = \mathscr{H}x \in R_2$. Тогда равенство (1) означает, что $\mathscr{H}\Gamma_1(a)x = \Gamma_2(a)y$, т. е. что если

$$x \leftrightarrow y$$

(вектор $x \in R_1$ соответствует вектору $y \in R_2$ при изоморфном отображении $\mathscr{H}$ пространства R_1 на R_2), то $\Gamma_1(a)x \leftrightarrow \Gamma_2(a)y$, т. е. образы $\Gamma_1(a)x$ и $\Gamma_2(a)y$ элементов x и y для любого $a \in G$ тоже соответствуют друг другу при отображении $\mathscr{H}$:

$$\mathscr{H}(\Gamma_1(a)x) = \Gamma_2(a)y.$$

Иными словами, безразлично, отобразить ли сначала вектор $x \in R_1$ посредством $\mathscr{H}$ в пространство R_2, а потом применить к полученному вектору $\mathscr{H}x$ преобразование, соответствующее элементу $a \in G$, или сделать это в обратном порядке; здесь выполняется следующая «коммутативная диаграмма»:

$$\begin{tikzcd}
x \arrow[r, "\mathscr{H}"] \arrow[d, "\Gamma_1(a)"'] & \mathscr{H}(x) \arrow[d, "\Gamma_2(a)"] \\
\Gamma_1(a)x \arrow[r, "\mathscr{H}"] & \mathscr{H}\Gamma_1(a)x = \Gamma_2(a)\mathscr{H}(x)
\end{tikzcd}$$

Изоморфное отображение $\mathscr{H}$ пространства R_1 на R_2 может быть задано квадратной (невырожденной) матрицей H. Тогда для каждого $a \in G$ соответствующие матрицы $\Gamma_1(a)$ и $\Gamma_2(a)$ изоморфных представлений Γ_1 и Γ_2 связаны соотношением $H\Gamma_1(a) = \Gamma_2(a)H$, которое можно переписать в виде

$$\Gamma_1(a) = H^{-1}\Gamma_2(a)H.$$

Это означает, что если пространство R_1 отождествить с (изоморфным ему) пространством R_2 и матрицу H рассматривать как матрицу перехода к новому базису в этом пространстве, то $\Gamma_1(a)$ и $\Gamma_2(a)$ — это матрицы одного и того же оператора, взятые в разных базисах.

Задача. Покажите, что отображения Γ_1 и Γ_2:

$$\Gamma_1(e)=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ \Gamma_1(r)=\begin{bmatrix}-1/2&-1/2\\3/2&-1/2\end{bmatrix},\ \Gamma_1(r^2)=\begin{bmatrix}-1/2&1/2\\-3/2&-1/2\end{bmatrix},$$

$$\Gamma_1(s)=\begin{bmatrix}1&0\\0&-1\end{bmatrix},\ \Gamma_1(sr)=\begin{bmatrix}-1/2&-1/2\\-3/2&1/2\end{bmatrix},\ \Gamma_1(sr^2)=\begin{bmatrix}-1/2&1/2\\3/2&1/2\end{bmatrix},$$

$$\Gamma_2(e)=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ \Gamma_2(r)=\begin{bmatrix}0&1\\-1&-1\end{bmatrix},\ \Gamma_2(r^2)=\begin{bmatrix}-1&-1\\1&0\end{bmatrix},$$

$$\Gamma_2(s)=\begin{bmatrix}0&1\\1&0\end{bmatrix},\ \Gamma_2(sr)=\begin{bmatrix}-1&-1\\0&1\end{bmatrix},\ \Gamma_2(sr^2)=\begin{bmatrix}1&0\\-1&-1\end{bmatrix}$$

являются представлениями диэдральной группы D_3, и проверьте, что эти представления изоморфны.

§ 3. Подпредставление

Определение 3. *Пусть Γ — представление группы G, R — пространство представления и R_1 — подпространство R, инвариантное относительно всех операторов, соответствующих элементам группы G* (в таком случае говорят, что *подпространство R_1* ***инвариантно относительно группы*** G). *Тогда каждому элементу $a\in G$ можно сопоставить оператор* $\Gamma(a)$, действующий в подпространстве R_1. *Эти операторы также образуют представление группы G* (поскольку, если $x\in R_1$, то $\Gamma(a)x\in R_1$ для каждого $a\in G$ и равенство

$$\Gamma(ab)x=\Gamma(a)\Gamma(b)x$$

выполняется для всех векторов $x\in R_1$, так как оно справедливо для любого вектора $x\in R$, а $R_1\subset R$). *Представление Γ_1 группы G в пространстве R_1 называется* ***подпредставлением*** *представления* Γ.

Пример. Рассмотрим двумерное представление циклической группы C_2 второго порядка:

$$\Gamma(e)=\begin{bmatrix}1&0\\0&1\end{bmatrix},\quad \Gamma(a)=\begin{bmatrix}0&1\\1&0\end{bmatrix}.$$

(Легко видеть, что $(\Gamma(a))^2=\Gamma(e)$ и, значит, это действительно — представление.) Пусть e_1, e_2 — тот базис пространства R, в котором взяты эти матрицы. Вектор e_1+e_2 является собственным для обоих преобразований $\Gamma(e)$ и $\Gamma(a)$, так как

$$\Gamma(e)(e_1+e_2) = \Gamma(e)e_1 + \Gamma(e)e_2 = e_1+e_2,$$
$$\Gamma(a)(e_1+e_2) = \Gamma(a)e_1 + \Gamma(a)e_2 = e_2+e_1 = e_1+e_2.$$

Следовательно, порожденное им одномерное подпространство $R_1 = \{e_1+e_2\}$ инвариантно относительно группы C_2. Соответствующее подпредставление Γ_1 является одномерным единичным представлением группы C_2:

$$\Gamma_1(e)(e_1+e_2) = e_1+e_2, \quad \Gamma_1(a)(e_1+e_2) = e_1+e_2,$$

или, в матричной форме:

$$\Gamma_1(e) = 1, \quad \Gamma_1(a) = 1.$$

§ 4. Прямая сумма представлений

Определение 4. *Если пространство представления R группы G является прямой суммой $R = R_1 \oplus R_2$ подпространств R_1* и R_2, *инвариантных относительно группы G, то в каждом из этих подпространств представление Γ определяет по подпредставлению; обозначим эти подпредставления через Γ_1 и Γ_2. Говорят, что представление Γ является **прямой суммой** подпредставлений Γ_1 и Γ_2*, что записывается так:

$$\Gamma = \Gamma_1 \oplus \Gamma_2.$$

Рассмотрим два примера.

1. Пусть Γ — двумерное представление группы C_2, введенное в предыдущем параграфе. Мы видели, что подпространство $R_1 = \{e_1+e_2\}$ инвариантно относительно группы C_2. Подпространство $R_2 = \{e_1-e_2\}$ тоже инвариантно относительно C_2, так как

$$\Gamma(e)(e_1-e_2) = \Gamma(e)e_1 - \Gamma(e)e_2 = e_1-e_2,$$
$$\Gamma(a)(e_1-e_2) = \Gamma(a)e_1 - \Gamma(a)e_2 = e_2-e_1 = -(e_1-e_2).$$

В базисе $f_1 = e_1+e_2$, $f_2 = e_1-e_2$ — матрицы нашего представления имеют вид

$$\Gamma(e) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{и} \quad \Gamma(a) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Таким образом, рассматриваемое представление является прямой суммой двух одномерных представлений: $\Gamma = \Gamma_1 \oplus \Gamma_2$, где

$$\Gamma_1(e) = 1, \quad \Gamma_1(a) = 1$$

и

$$\Gamma_2(e) = 1, \quad \Gamma_2(a) = -1.$$

2. Рассмотрим двумерное представление циклической группы C_4:

$$\Gamma(e)=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ \Gamma(a)=\begin{bmatrix}0&-1\\1&0\end{bmatrix},\ \Gamma(a^2)=\begin{bmatrix}-1&0\\0&-1\end{bmatrix},$$

$$\Gamma(a^3)=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$$

(проверьте сами, что это действительно представление!).

Так как пространство представления двумерно, то инвариантные подпространства (если они существуют) одномерны, и значит, для того чтобы найти их, мы должны найти собственные векторы преобразования $\Gamma(a)$ — они будут собственными и для остальных преобразований, поскольку $\Gamma(a^2)=[\Gamma(a)]^2$ и $\Gamma(a^3)=[\Gamma(a)]^3$.

Собственные значения преобразования $\Gamma(a)$ находятся из уравнения

$$\begin{vmatrix}-\lambda&-1\\1&-\lambda\end{vmatrix}=\lambda^2+1=0,$$

и значит, $\lambda=\pm i$.

При $\lambda=i$ собственные векторы находятся из уравнения $-ix_1-x_2=0$, откуда $x_2=-ix_1$, и собственный вектор $f_1=(1,-i)$.

При $\lambda=-i$ собственные векторы находятся из уравнения $ix_1-x_2=0$, откуда $x_2=ix_1$, т. е. $f_2=(1, i)$.

Далее, имеем

$$\Gamma(a)f_1=if_1,\quad \Gamma(a^2)f_1=-f_1,\quad \Gamma(a^3)f_1=-if_1,$$
$$\Gamma(a)f_2=-if_2,\quad \Gamma(a^2)f_2=-f_2,\quad \Gamma(a^3)f_2=if_2.$$

Таким образом, в базисе f_1, f_2 матрицы представления Γ будут иметь вид

$$\Gamma(e)=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ \Gamma(a)=\begin{bmatrix}i&0\\0&-i\end{bmatrix},\ \Gamma(a^2)=\begin{bmatrix}-1&0\\0&-1\end{bmatrix},$$
$$\Gamma(a^3)=\begin{bmatrix}-i&0\\0&i\end{bmatrix},$$

и представление Γ оказывается прямой суммой двух одномерных представлений:

$$\Gamma_1(e)=1,\quad \Gamma_1(a)=i,\quad \Gamma_1(a^2)=-1,\quad \Gamma_1(a^3)=-i$$

и

$$\Gamma_2(e)=1,\quad \Gamma_2(a)=-i,\quad \Gamma_2(a^2)=-1,\quad \Gamma_2(a^3)=i.$$

Легко видеть, что *из любых двух представлений группы G всегда можно составить представление, являющееся их прямой суммой.*

Заметим также, что *если представление Γ_1 группы G изоморфно Γ_1' и Γ_2 изоморфно Γ_2', то прямая сумма $\Gamma=\Gamma_1\oplus\Gamma_2$ изоморфна $\Gamma'=\Gamma_1'\oplus\Gamma_2'$.*

Действительно, пусть R_1, R_1', R_2, R_2' — соответственно пространства представлений $\Gamma_1, \Gamma_1', \Gamma_2, \Gamma_2'$. Тогда существуют изоморфные отображения $\mathscr{H}_1$ пространства R_1 на R_1' и $\mathscr{H}_2$ — пространства R_2 на R_2' такие, что

$$\mathscr{H}_1\Gamma_1 = \Gamma_1'\mathscr{H}_1 \quad \text{и} \quad \mathscr{H}_2\Gamma_2 = \Gamma_2'\mathscr{H}_2.$$

Пусть, далее, $R = R_1 \oplus R_2$ и $R' = R_1' \oplus R_2'$ — пространства представлений Γ и Γ'. Тогда отображение $\mathscr{H}$ с матрицей

$$H = \begin{bmatrix} H_1 & O \\ O & H_2 \end{bmatrix},$$

где H_i — матрица отображения $\mathscr{H}_i$, $i = 1, 2$, в соответственно выбранных базисах, будет, очевидно, изоморфным отображением пространства R на R'. Далее, ясно, что

$$\Gamma = \begin{bmatrix} \Gamma_1 & O \\ O & \Gamma_2 \end{bmatrix} \quad \text{и} \quad \Gamma' = \begin{bmatrix} \Gamma_1' & O \\ O & \Gamma_2' \end{bmatrix}$$

— матрицы представлений Γ и Γ', и мы имеем, очевидно,

$$H\Gamma = \begin{bmatrix} H_1 & O \\ O & H_2 \end{bmatrix}\begin{bmatrix} \Gamma_1 & O \\ O & \Gamma_2 \end{bmatrix} = \begin{bmatrix} H_1\Gamma_1 & O \\ O & H_2\Gamma_2 \end{bmatrix} = \begin{bmatrix} \Gamma_1'H_1 & O \\ O & \Gamma_2'H_2 \end{bmatrix} =$$

$$= \begin{bmatrix} \Gamma_1' & O \\ O & \Gamma_2' \end{bmatrix}\begin{bmatrix} H_1 & O \\ O & H_2 \end{bmatrix} = \Gamma' H,$$

а значит, представление Γ изоморфно Γ'.

Аналогичное утверждение справедливо, конечно, и для любого числа слагаемых Γ_i.

§ 5. Унитарное представление. Приводимые и неприводимые представления

Определение 5. *Представление* Γ *группы* G *называется* **унитарным**, *если в пространстве представления* R *можно так определить скалярное произведение, что это пространство станет евклидовым, а все операторы* $\Gamma(a)$, *где* $a \in G$, *будут унитарными.*

Лемма. Каждое представление конечной группы является унитарным.

Доказательство. Можно считать, что пространство R — евклидово: мы просто положим, по определению, скалярное произведение векторов $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$ и $y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$ равным

$$(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n.$$

Ясно, что относительно этой метрики пространство R будет евклидовым, а базис $e_1, e_2, \ldots, e_n$ — ортонормированным, так как $(e_i, e_i) = 1$ для всех i и $(e_i, e_j) = 0$ при $i \neq j$.

Если в этой евклидовой метрике все преобразования $\Gamma(a)$ унитарны, то наше утверждение доказано. В противном случае мы изменим скалярное произведение в R, полагая, по определению,

$$\langle x, y\rangle = \sum_{a\in G} (\Gamma(a)x, \Gamma(a)y), \tag{2}$$

где суммирование ведется по всем элементам a группы G. (В формуле (2) $\langle x, y\rangle$ — новое скалярное произведение векторов x и y, а (x, y) — старое скалярное произведение.)

Проверим, что функция $\langle\ ,\ \rangle$ удовлетворяет всем условиям, которые должны выполняться для скалярного произведения в комплексном евклидовом пространстве. Имеем

1. $$\langle y, x\rangle = \sum_{a\in G} (\Gamma(a)y, \Gamma(a)x) = \sum_{a\in G} \overline{(\Gamma(a)x, \Gamma(a)y)} = \overline{\sum_{a\in G} (\Gamma(a)x, \Gamma(a)y)} = \overline{\langle x, y\rangle}.$$

2. $$\langle \alpha x, y\rangle = \sum_{a\in G} (\Gamma(a)(\alpha x), \Gamma(a)y) = \sum_{a\in G} \alpha(\Gamma(a)x, \Gamma(a)y) = \alpha\sum_{a\in G} (\Gamma(a)x, \Gamma(a)y) = \alpha\langle x, y\rangle.$$

3. $$\langle x_1 + x_2, y\rangle = \sum_{a\in G} (\Gamma(a)(x_1 + x_2), \Gamma(a)y) = \sum_{a\in G} (\Gamma(a)x_1, \Gamma(a)y) + \sum_{a\in G} (\Gamma(a)x_2, \Gamma(a)y) = \langle x_1, y\rangle + \langle x_2, y\rangle.$$

4. Если вектор $x \neq 0$, то так как оператор $\Gamma(a)$ — невырожденный, то и $\Gamma(a)x \neq 0$, а тогда $(\Gamma(a)x, \Gamma(a)x) > 0$, и значит, вся сумма $\sum_{a\in G} (\Gamma(a)x, \Gamma(a)x)$ больше нуля С другой стороны, если $x = 0$, то $\langle x, x\rangle$, очевидно, равно 0. Таким образом, скалярный квадрат $\langle x, x\rangle \geqslant 0$ и из равенства $\langle x, x\rangle = 0$ вытекает, что $x = 0$.

Покажем, что *в новой метрике все операторы, соответствующие элементам группы G, унитарны*, т. е. что $\langle\Gamma(b)x, \Gamma(b)y\rangle = \langle x, y\rangle$ для каждого $b \in G$. Действительно,

$$\langle\Gamma(b)x, \Gamma(b)y\rangle = \sum_{a\in G} (\Gamma(a)\Gamma(b)x, \Gamma(a)\Gamma(b)y).$$

Но $\Gamma(a)\Gamma(b) = \Gamma(ab)$, так как Γ — представление группы G, и последняя сумма равна $\sum\limits_{a \in G} (\Gamma(ab)x, \Gamma(ab)y)$. Далее, если a пробегает все элементы (конечной!) группы G:

$$a = a_1, a_2, a_3, \ldots, a_k,$$

а b — один из этих элементов, скажем, $b = a_i$, то произведения

$$ab = a_1a_i, \quad a_2a_i, a_3a_i, \ldots, a_ka_i$$

— это тоже все элементы группы G, но взятые, вообще говоря, в каком-то другом порядке (из равенства $a_ka_i = a_ja_i$ немедленно вытекало бы, что $a_k = a_j$). Следовательно,

$$\sum_{a \in G} (\Gamma(ab)x, \Gamma(ab)y) = \sum_{a \in G} (\Gamma(a)x, \Gamma(a)y) = \langle x, y\rangle$$

и окончательно

$$\langle\Gamma(b)x, \Gamma(b)y\rangle = \langle x, y\rangle.$$

В дальнейшем, говоря о представлении (конечной) группы, мы всегда будем пользоваться тем, что оно унитарно.

Определение 6. *Представление* Γ *группы* G *в пространстве* R *называется* ***приводимым***, *если в* R *имеется нетривиальное* (т. е. отличное от всего пространства R и от «пространства размерности 0», образованного нулевым вектором) *подпространство* R_1, *инвариантное относительно* G (т. е. инвариантное относительно всех преобразований $\Gamma(a)$, где $a \in G$). *Если такого подпространства нет, то представление* Γ ***неприводимо.***

Ясно, что одномерное представление всегда неприводимо.

Теорема 1. *Пусть* Γ — *приводимое представление конечной группы* G *в пространстве* R *и* R — *подпространство* R, *инвариантное относительно группы* G. *Тогда найдется такое, тоже инвариантное относительно* G, *подпространство* R_2, *что*

$$R = R_1 \oplus R_2.$$

(Таким образом, для того чтобы представление конечной группы было разложимо в прямую сумму подпредставлений, необходимо и достаточно, чтобы оно было приводимым.)

Доказательство. По лемме мы можем считать пространство R евклидовым, а все операторы $\Gamma(a)$, соответствующие элементам группы G, — унитарными.

Тогда ортогональное дополнение $R_2 = R_1^{\perp}$ подпространства R_1 тоже инвариантно относительно группы G (см. стр. 182) и $R = R_1 \oplus R_2$ (см. стр. 156).

Для бесконечных групп последняя теорема, вообще говоря, неверна. Рассмотрим бесконечную циклическую группу — аддитивную группу целых чисел. Отображение Г, ставящее в соответствие числу k матрицу

$$\Gamma(k) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix},$$

является (двумерным) представлением этой группы, так как

$$\Gamma(k)\,\Gamma(m) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ k+m & 1 \end{bmatrix} = \Gamma(k+m).$$

Одномерное подпространство $\{e_2\}$ инвариантно относительно всех преобразований $\Gamma(k)$, но для него не найдется инвариантного дополнительного подпространства, так как (двумерное) пространство R представления Г не имеет никаких других подпространств, инвариантных относительно $\Gamma(k)$. Действительно, собственные значения преобразования $\Gamma(k)$ находятся из уравнения

$$\begin{vmatrix} 1-\lambda & 0 \\ k & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = 0,$$

и значит, $\lambda_{1,2} = 1$. Соответствующие собственные векторы удовлетворяют уравнениям

$$\begin{cases} 0x_1 + 0x_2 = 0, \\ kx_1 + 0x_2 = 0, \end{cases}$$

т. е. $x_1 = 0$, а это — только векторы, коллинеарные e_2.

Теорема 2. *Всякое представление конечной группы либо неприводимо, либо является прямой суммой неприводимых представлений.*

Доказательство. Если представление Г группы G в пространстве R приводимо, то, по теореме 1, его можно разложить в прямую сумму представлений меньших размерностей. Если какое-нибудь из слагаемых приводимо, с ним поступим так же, и т. д. Этот процесс не может продолжаться до бесконечности, так как размерности слагаемых уменьшаются, а размерность пространства представления конечна. Окончательно наше представление разложится в прямую сумму неприводимых представлений:

$$\Gamma = \Gamma_1 \oplus \Gamma_2 \oplus \ldots \oplus \Gamma_m.$$

§ 6. Регулярное представление

Для каждой конечной группы можно построить так называемое регулярное представление, играющее важную роль в общей теории представлений групп.

Пусть $a_1, a_2, \ldots, a_k$ — все элементы группы G. Рассмотрим k-мерное векторное пространство R, элементы базиса которого поставим во взаимно однозначное соответствие элементам группы G; короче говоря, мы просто перенумеруем элементы базиса элементами группы G:

$$e_{a_1}, e_{a_2}, \ldots, e_{a_k}.$$

Далее, положим, по определению,

$$\Gamma(a_i)e_{a_j} = e_{a_i a_j}$$

Мы получим k-мерное представление группы G, так как

$$\Gamma(a_i)\Gamma(a_j)e_{a_s} = \Gamma(a_i)e_{a_j a_s} = e_{a_i(a_j a_s)} = e_{(a_i a_j)a_s} = \Gamma(a_i a_j)e_{a_s},$$

т. е. для всех базисных векторов $e_{a_1}, e_{a_2}, \ldots, e_{a_k}$ — а значит, и для всех векторов x пространства R,— имеем

$$\Gamma(a_i)\Gamma(a_j)x = \Gamma(a_i a_j)x;$$

следовательно,

$$\Gamma(a_i)\Gamma(a_j) = \Gamma(a_i a_j).$$

Определенное таким образом представление группы G и называется ее *регулярным представлением.*

Рассмотрим несколько примеров.

1. *Регулярное представление группы C_2.*

Пусть ε и a — элементы группы, причем $a^2 = \varepsilon$ *). Пространство представления будет двумерным с элементами базиса e_ε и e_a. По определению,

$$\Gamma(\varepsilon)e_\varepsilon = e_\varepsilon, \quad \Gamma(\varepsilon)e_a = e_a, \quad \Gamma(a)e_\varepsilon = e_a, \quad \Gamma(a)e_a = e_{a^2} = e_\varepsilon;$$

и соответствующие матрицы имеют вид

$$\Gamma(\varepsilon) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Gamma(a) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

*) Здесь и далее мы часто будем обозначать единицу группы G греческой буквой ε, а не латинской буквой e.

2. *Регулярное представление группы* C_4. Здесь элементы группы ε, a, a^2, a^3 и базис пространства представления образуют четыре вектора

$$e_\varepsilon,\ e_a,\ e_{a^2},\ e_{a^3}.$$

По определению,

$$\Gamma(a)e_\varepsilon = e_a,\quad \Gamma(a)e_a = e_{a^2},\quad \Gamma(a)e_{a^2} = e_{a^3},\quad \Gamma(a)e_{a^3} = e_\varepsilon;$$
$$\Gamma(a^2)e_\varepsilon = e_{a^2},\quad \Gamma(a^2)e_a = e_{a^3},\quad \Gamma(a^2)e_{a^2} = e_\varepsilon,\quad \Gamma(a^2)e_{a^3} = e_a;$$
$$\Gamma(a^3)e_\varepsilon = e_{a^3},\quad \Gamma(a^3)e_a = e_\varepsilon,\quad \Gamma(a^3)e_{a^2} = e_a,\quad \Gamma(a^3)e_{a^3} = e_{a^2};$$

соответствующие матрицы имеют вид

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},\quad \begin{bmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix},\quad \begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix},\quad \begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{bmatrix}.$$

3. *Регулярное представление группы* V. Элементы группы ε, a, b и $ab = ba$, причем $a^2 = b^2 = \varepsilon$; базис пространства представления

$$e_\varepsilon,\ e_a,\ e_b,\ e_{ab}.$$

По определению,

$$\Gamma(a)e_\varepsilon = e_a,\quad \Gamma(a)e_a = e_\varepsilon,\quad \Gamma(a)e_b = e_{ab},\quad \Gamma(a)e_{ab} = e_b;$$
$$\Gamma(b)e_\varepsilon = e_b,\quad \Gamma(b)e_a = e_{ab},\quad \Gamma(b)e_b = e_\varepsilon,\quad \Gamma(b)e_{ab} = e_a;$$
$$\Gamma(ab)e_\varepsilon = e_{ab},\quad \Gamma(ab)e_a = e_b,\quad \Gamma(ab)e_b = e_a,\quad \Gamma(ab)e_{ab} = e_\varepsilon;$$

соответствующие матрицы имеют вид

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},\quad \begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix},\quad \begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix},\quad \begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix}.$$

4. *Регулярное представление диэдральной группы* D_3. Здесь элементы группы ε, r, r^2, s, sr, sr^2, причем $r^3 = s^2 = \varepsilon$, $rs = sr^2$. Базис пространства представления

$$e_\varepsilon,\ e_r,\ e_{r^2},\ e_s,\ e_{sr},\ e_{sr^2}.$$

Проверьте сами, что соответствующие матрицы имеют вид

$$\begin{bmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{bmatrix}, \quad \begin{bmatrix} 0&0&1&0&0&0\\ 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 0&0&0&1&0&0 \end{bmatrix}, \quad \begin{bmatrix} 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 1&0&0&0&0&0\\ 0&0&0&0&0&1\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0 \end{bmatrix},$$

$$\begin{bmatrix} 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0 \end{bmatrix}, \quad \begin{bmatrix} 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 0&0&0&1&0&0\\ 0&0&1&0&0&0\\ 1&0&0&0&0&0\\ 0&1&0&0&0&0 \end{bmatrix}, \quad \begin{bmatrix} 0&0&0&0&0&1\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 1&0&0&0&0&0 \end{bmatrix}.$$

При $k > 1$ регулярное представление приводимо, так как, например, одномерное подпространство, порожденное вектором $f = e_{a_1} + e_{a_2} + \ldots + e_{a_k}$, инвариантно относительно группы G: при всех i

$$\Gamma(a_i) f = \Gamma(a_i)\left(e_{a_1} + e_{a_2} + \ldots + e_{a_k}\right) = \\ = e_{a_i a_1} + e_{a_i a_2} + \ldots + e_{a_i a_k}.$$

Но $a_i a_1, a_i a_2, \ldots, a_i a_k$ — это те же элементы $a_1, a_2, \ldots, a_k$, только, быть может, взятые в каком-то другом порядке. Следовательно, $\Gamma(a_i) f = f$.

Легко видеть, что в каждой строке и в каждом столбце матрицы регулярного представления один из элементов равен 1, а все остальные элементы равны 0 (докажите это сами).

§ 7. Функции, определенные на группе

Пусть G — произвольная (конечная) группа; предположим, что каждому элементу a группы G поставлено в соответствие какое-то, вообще говоря, комплексное число $\varphi(a)$. Мы будем говорить тогда, что *на группе G задана функция* φ.

Если естественным образом определить сложение функций:

$$(\varphi + \psi)(a) = \varphi(a) + \psi(a) \text{ для каждого } a \in G$$

и умножение функции на число:

$$(\alpha\varphi)(a) = \alpha\varphi(a) \text{ для каждого } a \in G \text{ и любого (комплексного) числа } \alpha,$$

то очевидно, что множество всех комплексно-значных (т. е. принимающих комплексные значения) функций, определенных на группе G, станет векторным пространством. Покажем, что *если порядок группы равен k, то это пространство k-мерно.* Действительно, пусть элементы группы будут $a_1, a_2, \ldots, a_k$; рассмотрим k функций $\varphi_i(a)$, где $i = 1, 2, \ldots, k$, определяемых следующим образом:

$$\varphi_i(a_j) = \begin{cases} 1, & \text{если} \quad i = j, \\ 0, & \text{если} \quad i \neq j. \end{cases}$$

Легко видеть, что эти k функций линейно независимы и что *каждая функция, определенная на группе G, является их линейной комбинацией.* В самом деле, если F — произвольная функция на группе G и, скажем, $F(a_m) = \alpha_m$, то, очевидно,

$$F(a) = \sum_{i=1}^{k} \alpha_i \varphi_i(a), \quad \text{т. е.} \quad F = \sum_{i=1}^{k} \alpha_i \varphi_i.$$

Определение 7. *Функция $\varphi(a)$, определенная на группе G, называется центральной, если для любых двух элементов $a, b \in G$*

$$\varphi(ab) = \varphi(ba).$$

Пусть φ — центральная функция на группе G. Тогда для любых $a, b \in G$ имеем

$$\varphi(b^{-1}ab) = \varphi((b^{-1}a)b) = \varphi(b(b^{-1}a)) = \\ = \varphi((bb^{-1})a) = \varphi(a).$$

Обратно, пусть для всех $a, b \in G$ имеет место равенство $\varphi(b^{-1}ab) = \varphi(a)$. Полагая $b^{-1}a = c$ (откуда $a = bc$), получим

$$\varphi(cb) = \varphi(b^{-1}ab) = \varphi(a) = \varphi(bc).$$

Итак, если равенство $\varphi(ab) = \varphi(ba)$ выполняется тождественно для всех элементов группы G, то тождественно выполняется и равенство $\varphi(b^{-1}ab) = \varphi(a)$, и наоборот. Следовательно, *функция φ на группе G в том и только в том случае является центральной, если она принимает равные значения на всех сопряженных между собой элементах группы.* Можно сказать поэтому, что центральная функция определена на классах сопряженных элементов группы.

Множество всех центральных функций является подпространством пространства функций, определенных на группе G так как сумма центральных функций и произведение центральной функции на число тоже являются, очевидно, центральными функциями.

Теорема 3. *Размерность пространства центральных функций, определенных на группе G, равна числу p классов сопряженных элементов этой группы.*

Доказательство. Пусть C_1, C_2, ..., C_p — все классы сопряженных элементов группы G. Если F — произвольная центральная функция на группе G, то она может рассматриваться как функция, определенная на этих классах, т. е. функция, ставящая в соответствие каждому классу C_i определенное число $F(C_i)$.

Рассмотрим p (центральных) функций $\psi_1, \psi_2, \ldots, \psi_p$, где

$$\psi_i(C_j) = \begin{cases} 1 \text{ при } i = j, \\ 0 \text{ при } i \neq j. \end{cases}$$

Функции ψ_i, очевидно, линейно независимы, и каждая центральная функция является их линейной комбинацией: если F - произвольная центральная функция на G и $F(C_j) = \alpha_j$, $j = 1, 2, \ldots, p$, то

$$F = \sum_{i=1}^{p} \alpha_i \psi_i.$$

Значит, размерность пространства центральных функций, определенных на группе G, равна числу классов сопряженных элементов этой группы.

Выше мы видели, что если Γ — одномерное представление группы G, то $\Gamma(b^{-1}ab) = \Gamma(a)$ для любых $a, b \in G$; поэтому одномерные представления любой группы являются определенными на ней центральными функциями. Так, для группы D_3 на стр. 328 мы нашли две определенные на ней центральные функции Γ_1 и Γ_2: Γ_1 тождественно равна 1 и $\Gamma_2(e) = 1$,

$$\Gamma_2(r) = \Gamma_2(r^2) = 1, \qquad \Gamma_2(s) = \Gamma_2(sr) = \Gamma_2(sr^2) = -1,$$

§ 8. Скалярное произведение на группе

В пространстве функций, заданных на группе G порядка k, определим *скалярное произведение*, полагая

$$(\varphi, \psi) = \frac{1}{k} \sum_{a \in G} \varphi(a) \overline{\psi(a)},$$

где суммирование распространяется на все элементы a группы G. Проверим, что при этом будут выполнены все аксиомы скалярного произведения:

1. $(\psi, \varphi) = \frac{1}{k} \sum_{a \in G} \psi(a) \overline{\varphi(a)} = \overline{\frac{1}{k} \sum_{a \in G} \varphi(a) \overline{\psi(a)}} = \overline{(\varphi, \psi)}$.

2. $(\alpha\varphi, \psi) = \frac{1}{k} \sum_{a \in G} \alpha\varphi(a) \overline{\psi(a)} = \alpha(\varphi, \psi)$.

3. $(\varphi_1 + \varphi_2, \psi) =$
$$= \frac{1}{k} \sum_{a \in G} [\varphi_1(a) + \varphi_2(a)] \overline{\psi(a)} = (\varphi_1, \psi) + (\varphi_2, \psi).$$

При этом пространство функций, определенных на группе G, будет евклидовым, так как

4. $(\varphi, \varphi) = \frac{1}{k} \sum_{a \in G} \varphi(a) \overline{\varphi(a)} = \frac{1}{k} \sum_{a \in G} |\varphi(a)|^2 \geqslant 0,$

и если $(\varphi, \varphi) = 0$, то $\varphi(a) = 0$ для каждого $a \in G$, и значит, $\varphi \equiv 0$.

Рассмотрим еще несколько примеров. Выше (в § 1) мы нашли 3 представления диэдральной группы D_3: два одномерных $\Gamma_1 \equiv 1$ и

$$\Gamma_2(e) = \Gamma_2(r) = \Gamma_2(r^2) = 1, \quad \Gamma_2(s) = \Gamma_2(sr) = \Gamma_2(sr^2) = -1,$$

и двумерное, назовем его Γ_3:

$$\Gamma_3(e) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Gamma_3(r) = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix},$$

$$\Gamma_3(r^2) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}; \quad \Gamma_3(s) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\Gamma_3(sr)=\begin{bmatrix}-\frac{1}{2} & -\frac{1}{2}\sqrt{3}\\ -\frac{1}{2}\sqrt{3} & \frac{1}{2}\end{bmatrix},\quad \Gamma_3(sr^2)=\begin{bmatrix}-\frac{1}{2} & \frac{1}{2}\sqrt{3}\\ \frac{1}{2}\sqrt{3} & \frac{1}{2}\end{bmatrix}.$$

Рассмотрим шесть функций на группе D_3: Γ_1, Γ_2 и функции γ_{11}, γ_{12}, γ_{21}, γ_{22}, определенные следующим образом:

$$\gamma_{11}(e)=1,\quad \gamma_{11}(r)=-\frac{1}{2},\quad \gamma_{11}(r^2)=-\frac{1}{2},$$

$$\gamma_{11}(s)=1,\quad \gamma_{11}(sr)=-\frac{1}{2},\quad \gamma_{11}(sr^2)=-\frac{1}{2};$$

$$\gamma_{12}(e)=0,\quad \gamma_{12}(r)=-\frac{1}{2}\sqrt{3},\quad \gamma_{12}(r^2)=\frac{1}{2}\sqrt{3},$$

$$\gamma_{12}(s)=0,\quad \gamma_{12}(sr)=-\frac{1}{2}\sqrt{3},\quad \gamma_{12}(sr^2)=\frac{1}{2}\sqrt{3};$$

$$\gamma_{21}(e)=0,\quad \gamma_{21}(r)=\frac{1}{2}\sqrt{3},\quad \gamma_{21}(r^2)=-\frac{1}{2}\sqrt{3},$$

$$\gamma_{21}(s)=0,\quad \gamma_{21}(sr)=-\frac{1}{2}\sqrt{3},\quad \gamma_{21}(sr^2)=\frac{1}{2}\sqrt{3};$$

$$\gamma_{22}(e)=1,\quad \gamma_{22}(r)=-\frac{1}{2},\quad \gamma_{22}(r^2)=-\frac{1}{2},$$

$$\gamma_{22}(s)=-1,\quad \gamma_{22}(sr)=\frac{1}{2},\quad \gamma_{22}(sr^2)=\frac{1}{2}$$

(таким образом, γ_{ij} — это соответствующие элементы матрицы Γ_3). Вычислим попарные скалярные произведения этих функций:

$$(\Gamma_1,\Gamma_2)=$$
$$=\frac{1}{6}[1\cdot 1+1\cdot 1+1\cdot 1+1(-1)+1(-1)+1(-1)]=0;$$
$$(\Gamma_2,\gamma_{11})=\frac{1}{6}\left[1\cdot 1+1\left(-\frac{1}{2}\right)+1\left(-\frac{1}{2}\right)+\right.$$
$$\left.+(-1)\cdot 1+(-1)\left(-\frac{1}{2}\right)+(-1)\left(-\frac{1}{2}\right)\right]=0. \tag{3}$$

Аналогично получаем

$$(\Gamma_1,\gamma_{11})=(\Gamma_1,\gamma_{12})=(\Gamma_1,\gamma_{21})=(\Gamma_1,\gamma_{22})=(\Gamma_2,\gamma_{12})=$$
$$=(\Gamma_2,\gamma_{21})=(\Gamma_2,\gamma_{22})=0.$$

Далее,

$$(\gamma_{11}, \gamma_{12}) = \frac{1}{6}\left[1\cdot 0 + \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\sqrt{3}\right) + \left(-\frac{1}{2}\right)\cdot\frac{1}{2}\sqrt{3} + \right.$$
$$\left. + 1\cdot 0 + \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\sqrt{3}\right) + \left(-\frac{1}{2}\right)\frac{1}{2}\sqrt{3}\right] = 0. \quad (4)$$

Аналогично,

$$(\gamma_{11}, \gamma_{21}) = (\gamma_{11}, \gamma_{22}) = (\gamma_{12}, \gamma_{21}) =$$
$$= (\gamma_{12}, \gamma_{22}) = (\gamma_{21}, \gamma_{22}) = 0.$$

Все рассматриваемые функции, таким образом, *попарно ортогональны*. Скалярные квадраты этих функций равны

$$(\Gamma_1, \Gamma_1) = \frac{1}{6}(1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2) = 1 = \frac{1}{1}$$

и (5′)

$$(\Gamma_2, \Gamma_2) = \frac{1}{6}(1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2) = 1 = \frac{1}{1}$$

— для одномерных представлений и

$$(\gamma_{11}, \gamma_{11}) = \frac{1}{6}\left(1 + \frac{1}{4} + \frac{1}{4} + 1 + \frac{1}{4} + \frac{1}{4}\right) = \frac{1}{2},$$
$$(\gamma_{12}, \gamma_{12}) = \frac{1}{6}\left(0 + \frac{3}{4} + \frac{3}{4} + 0 + \frac{3}{4} + \frac{3}{4}\right) = \frac{1}{2}, \quad (5'')$$

а также и $(\gamma_{21}, \gamma_{21}) = (\gamma_{22}, \gamma_{22}) = \frac{1}{2}$ — для двумерных представлений. Всюду в знаменателе правой части стоит *степень соответствующего представления*. Ниже (в § 10) мы докажем аналогичные равенства для общего случая.

§ 9. Лемма Шура

Лемма Шура состоит из двух частей с «Прологом» и «Эпилогом».

Пролог леммы Шура*). Пусть Γ_1 и Γ_2 — два линейных представления группы G в пространствах R_1 и R_2 соответственно и $\mathscr{H}$ — линейное отображение прост-

*) Что же касается «Эпилога» леммы Шура, то он появится у нас только в следующей главе (см. стр. 357).

ранства R_1 в R_2. В отличие от отображения, фигурирующего в § 2, $\mathscr{H}$ не предполагается взаимно однозначным; точнее говоря, каждому вектору $x \in R_1$ поставлен в соответствие определенный вектор $\mathscr{H}x \in R_2$, так что

$$\mathscr{H}(x+y) = \mathscr{H}x + \mathscr{H}y \text{ и } \mathscr{H}(\alpha x) = \alpha\mathscr{H}x,$$

но не предполагается, что каждый вектор из R_2 является образом хотя бы одного вектора из R_1 и что из равенства $\mathscr{H}x = \mathscr{H}y$ вытекает $x = y$.

Предположим, далее, что $\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H}$, т. е. что $\mathscr{H}\Gamma_1(a) = \Gamma_2(a)\mathscr{H}$ для каждого элемента $a \in G$ и, значит, $\mathscr{H}\Gamma_1(a)x = \Gamma_2(a)\mathscr{H}x$ для каждого вектора $x \in R_1$. Последнее означает, что безразлично, сначала ли применить к вектору x оператор $\Gamma_1(a)$, соответствующий элементу a группы G, а затем отобразить полученный вектор $\Gamma_1(a)x$ в R_2, или сначала отобразить x в R_2, а потом к полученному вектору $\mathscr{H}x$ применить оператор $\Gamma_2(a)$, соответствующий тому же элементу $a \in G$ (см. «Коммутативную диаграмму» на стр. 331).

Докажем, что в нашем предположении *образ* $\mathscr{H}R_1$ *пространства* R_1 *в* R_2 (т. е. совокупность всевозможных векторов вида $\mathscr{H}x$, где $x \in R_1$) *и ядро* N *отображения* $\mathscr{H}$ (образованное всеми такими векторами $x \in R_1$, что $\mathscr{H}x = 0$), *являются подпространствами, инвариантными относительно группы* G.

То, что $\mathscr{H}R_1$ в R_2 и N в R_1 являются подпространствами, доказывается совсем просто (ср. стр. 115). Докажем инвариантность этих подпространств.

1. *Инвариантность* $\mathscr{H}R_1$. Пусть $x \in \mathscr{H}R_1$, тогда $x = \mathscr{H}x_1$, где $x_1 \in R_1$. Нам надо показать, что вектор $\Gamma_2(a)x$ тоже принадлежит $\mathscr{H}R_1$. Но так как $\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H}$, то

$$\Gamma_2(a)x = \Gamma_2(a)\mathscr{H}x_1 = \mathscr{H}\Gamma_1(a)x_1,$$

и значит, $\Gamma_2(a)x \in \mathscr{H}R_1$.

2. *Инвариантность* N. Пусть $y \in N$, т. е. $\mathscr{H}y = 0$. Нам надо показать, что и $\Gamma_1(a)y \in N$. Но так как $\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H}$, то

$$\mathscr{H}\Gamma_1(a)y = \Gamma_2(a)\mathscr{H}y = 0,$$

и значит, $\Gamma_1(a)y \in N$.

Таким образом, подпространства $\mathscr{H}R_1 \subseteq R_2$ и $N \subseteq R_1$ инвариантны относительно группы G.

Лемма Шура, часть I. *Если в сформулированных выше условиях* (Γ_1 *и* Γ_2 — *представления группы* G *в пространствах* R_1 *и* R_2 *и* $\mathscr{H}$ — *линейное отображение пространства* R_1 *в* R_2 *такое, что* $\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H}$) *представления* Γ_1 *и* Γ_2 *неприводимы, то*

либо $\mathscr{H} = \mathscr{O}$,

либо представления Γ_1 *и* Γ_2 *изоморфны.*

Доказательство. Так как представления Γ_1 и Γ_2 неприводимы, то ни R_1, ни R_2 не могут содержать нетривиальных подпространств, инвариантных относительно группы G. Однако образ $\mathscr{H}R_1$ пространства R_1 является в R_2 инвариантным подпространством; следовательно, либо $\mathscr{H}R_1 = 0$, либо $\mathscr{H}R_1 = R_2$. Но если $\mathscr{H}R_1 = 0$, то $\mathscr{H}$ отображает все R_1 в нуль, и значит, $\mathscr{H} = \mathscr{O}$. Если же $\mathscr{H}R_1 = R_2$, то $\mathscr{H}$ отображает пространство R_1 на все R_2 (т. е. каждый элемент из R_2 будет образом хотя бы одного элемента из R_1).

Далее, ядро N отображения $\mathscr{H}$ является в R_1 инвариантным подпространством, и значит, либо $N = 0$, либо $N = R_1$. Если $N = R_1$, то все пространство R_1 отображается в нулевой вектор, и значит, $\mathscr{H} = \mathscr{O}$. Если же $N = 0$, то отображение $\mathscr{H}$ взаимно однозначно. Действительно, из равенства $\mathscr{H}x = \mathscr{H}y$ вытекает, что $\mathscr{H}(x - y) = 0$, откуда следует, что $x - y \in N$, и значит, $x - y = 0$, т. е. $x = y$.

Итак, если $\mathscr{H} \neq \mathscr{O}$, то $\mathscr{H}$ *взаимно однозначно* отображает пространство R_1 *на все* R_2. Так как, кроме того, это отображение линейное, то оно будет *изоморфным* отображением R_1 на R_2. В силу равенства $\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H}$, представления Γ_1 и Γ_2 *изоморфны.*

Лемма Шура, часть II. *Пусть* Γ — *неприводимое представление группы* G *в пространстве* R *и* $\mathscr{H}$ — *такой линейный оператор в пространстве* R, *что* $\mathscr{H}\Gamma = \Gamma\mathscr{H}$; *тогда* $\mathscr{H}$ *является гомотетией* (*т. е. существует такое число* λ, *что* $\mathscr{H} = \lambda\mathscr{E}$).

Доказательство. Пусть λ — одно из собственных значений *оператора* $\mathscr{H}$, а $x \neq 0$ — соответствующий собственный вектор. Тогда

$$\mathscr{H}x = \lambda x, \text{ или } (\mathscr{H} - \lambda\mathscr{E})x = 0.$$

Из равенства $\mathcal{H}\Gamma = \Gamma\mathcal{H}$ вытекает, что

$$(\mathcal{H} - \lambda\mathcal{E})\Gamma = \Gamma(\mathcal{H} - \lambda\mathcal{E}).$$

Как доказано в «Прологе» (если применить его к оператору $\mathcal{H} - \lambda\mathcal{E}$ при $R_1 = R_2 = R$ и $\Gamma_1 = \Gamma_2 = \Gamma$), ядро N оператора $\mathcal{H} - \lambda\mathcal{E}$ инвариантно относительно группы G, а так как представление Γ неприводимо, то либо $N = 0$, либо $N = R$. Однако равенство $N = 0$ невозможно, поскольку ненулевой вектор x принадлежит N. Следовательно, $N = R$, т. е. все пространство R оператором $\mathcal{H} - \lambda\mathcal{E}$ отображается в нулевой вектор, откуда

$$\mathcal{H} - \lambda\mathcal{E} = \mathcal{O} \text{ и } \mathcal{H} = \lambda\mathcal{E}.$$

§ 10. Следствия из леммы Шура

Вспомогательное предложение. *Пусть Γ_1 и Γ_2 — произвольные представления группы G в пространствах R_1 и R_2 и $\mathcal{H}$ — произвольное линейное отображение R_1 в R_2. Тогда*

$$\mathcal{H}_0 = \sum_{a\in G} \Gamma_2(a^{-1})\,\mathcal{H}\Gamma_1(a),$$

являющееся, как легко видеть, линейным отображением R_1 в R_2, удовлетворяет условию

$$\mathcal{H}_0\Gamma_1 = \Gamma_2\mathcal{H}_0,$$

т. е. для каждого элемента b группы G

$$\mathcal{H}_0\Gamma_1(b) = \Gamma_2(b)\mathcal{H}_0,$$

или, что то же самое,

$$\Gamma_2^{-1}(b)\,\mathcal{H}_0\Gamma_1(b) = \mathcal{H}_0.$$

Доказательство. Мы имеем

$$\Gamma_2^{-1}(b)\,\mathcal{H}_0\Gamma_1(b) = \Gamma_2(b^{-1})\,\mathcal{H}_0\Gamma_1(b) =$$
$$= \sum_{a\in G} \Gamma_2(b^{-1})\,\Gamma_2(a^{-1})\,\mathcal{H}\Gamma_1(a)\,\Gamma_1(b) =$$
$$= \sum_{a\in G} \Gamma_2(b^{-1}a^{-1})\,\mathcal{H}\Gamma_1(ab) = \sum_{a\in G} \Gamma_2((ab)^{-1})\mathcal{H}\Gamma_1(ab).$$

Но если a пробегает все элементы (конечной!) группы

G, а b — фиксированный элемент этой группы, то произведение ab тоже пробегает все элементы группы G только, вообще говоря, в каком-то другом порядке. Следовательно,

$$\sum_{a\in G} \Gamma_2((ab)^{-1})\,\mathscr{H}_0\Gamma_1(ab) = \mathscr{H}_0.$$

Мы доказали, что

$$[\Gamma_2(b)]^{-1}\mathscr{H}_0\Gamma_1(b) = \mathscr{H}_0, \text{ или } \mathscr{H}_0\Gamma_1(b) = \Gamma_2(b)\mathscr{H}_0,$$

где b — любой элемент группы G.

Следствие из I-й части леммы Шура. Пусть Γ_1 и Γ_2 — *неприводимые неизоморфные представления группы* G в пространствах R_1 и R_2 и $\mathscr{H}$ — произвольное (линейное) отображение R_1 в R_2. Предположим, что в пространствах R_1 и R_2 выбраны ортонормированные базисы, и пусть в этих базисах элементы матриц операторов $\Gamma_1(a)$, $\Gamma_2(a)$ и отображения $\mathscr{H}$ будут соответственно $\gamma_{ij}^{(1)}(a)$, $\gamma_{ij}^{(2)}(a)$ и h_{ij}.

Мы видели, что отображение

$$\mathscr{H}_0 = \sum_{a\in G} \Gamma_2(a^{-1})\,\mathscr{H}\Gamma_1(a)$$

пространства R_1 в R_2 удовлетворяет условию

$$\mathscr{H}_0\Gamma_1 = \Gamma_2\mathscr{H}_0.$$

Тогда, по I-й части леммы Шура, поскольку представления Γ_1 и Γ_2 не изоморфны, $\mathscr{H}_0 = 0$.

Элемент $h_{pq}^{(0)}$, стоящий на пересечении p-й строки и q-го столбца матрицы H_0, равен

$$h_{pq}^0 = \sum_{a\in G}\sum_{i,j} \gamma_{pi}^{(2)}(a^{-1})\,h_{ij}\gamma_{jq}^{(1)}(a) = \sum_{i,j} h_{ij} \sum_{a\in G} \gamma_{pi}^{(2)}(a^{-1})\,\gamma_{jq}^{(1)}(a)$$

(здесь суммирование ведется по i от 1 до n_2 и по j от 1 до n_1, где n_l, $l = 1, 2$,— размерность пространства R_l) — и он равен нулю при любом отображении $\mathscr{H}$, т. е. при любых h_{ij}. Следовательно, коэффициент при h_{ij} в правой части последнего равенства равен нулю при всех i, j:

$$\sum_{a\in G} \gamma_{pi}^{(2)}(a^{-1})\,\gamma_{jq}^{(1)}(a) = 0.$$

(Легко видеть, что если сумма $\sum_{i,j} \alpha_{ij} h_{ij}$ равна нулю при подстановке любых значений h_{ij}, то все ее коэффициенты α_{ij} равны нулю. Действительно, положив, например, $h_{11} = 1$, а все остальные $h_{ij} = 0$, мы получим, что $\sum_{i,j} \alpha_{ij} h_{ij} = \alpha_{11}$, и значит, $\alpha_{11} = 0$, и т. д.) Но выше мы условились (см. стр. 335), что все рассматриваемые операторы унитарны. Следовательно, $\Gamma_2(a^{-1}) = [\Gamma_2(a)]^{-1} = [\Gamma_2(a)]^*$. Так как базис в пространстве R_2 — ортонормированный, то элементы матрицы $[\gamma_{ij}^{(2)}(a)]$ оператора $\Gamma^{(2)}(a)$ удовлетворяют условиям: $\gamma_{pi}^{(2)}(a^{-1}) = \bar{\gamma}_{ip}^{(2)}(a)$ при всех i, p. Таким образом, при всех i, j

$$\sum_{a \in G} \bar{\gamma}_{ip}^{(2)}(a) \gamma_{jq}^{(1)}(a) = 0.$$

Если матричные элементы $\gamma_{ip}^{(2)}$, $\gamma_{jq}^{(1)}$ рассматривать как функции, заданные на группе G, то последнее равенство означает, что скалярное произведение любых двух таких функций, взятых для *неизоморфных неприводимых* представлений, равно нулю:

$$(\gamma_{jq}^{(1)}, \gamma_{ip}^{(2)}) = 0 \quad \text{при всех} \quad i,\ j,\ p,\ q,$$

т. е. что эти функции *попарно ортогональны*. Так, в примерах, приведенных в § 8, функции γ_{11}, γ_{12}, γ_{21}, γ_{22} ортогональны функциям Γ_1 и Γ_2, а также Γ_1 ортогональна Γ_2 (см. равенства (3) на стр. 345).

Полезно заметить, что поскольку для каждой группы имеется единичное представление $\Gamma_0(a) = 1$ для всех $a \in G$, то, если $\gamma_{ij}(a)$ — элементы матрицы произвольного (неприводимого) неединичного представления этой группы, то

$$(\gamma_{ij}, \Gamma_0) = \sum_{a \in G} \gamma_{ij}(a) = 0,$$

т. е. *для любых i, j сумма всех значений функции γ_{ij} равна нулю.*

Следствие из II-й части леммы Шура. Пусть Γ — *неприводимое* представление группы G в (n-мерном) пространстве R и $\mathscr{H}$ — произвольный линейный оператор в R. Тогда, как было показано выше (вспомогательное предложение для случая $\Gamma_1 = \Gamma_2 = \Gamma$, $R_1 = R_2 = R$), линейный оператор

$$\mathscr{H}_0 = \sum_{a \in G} \Gamma(a^{-1}) \mathscr{H} \Gamma(a)$$

в пространстве R удовлетворяет условию $\mathscr{H}_0 \Gamma = \Gamma \mathscr{H}_0$, и значит (по лемме Шура, часть II), $\mathscr{H}_0$ — гомотетия:

$$\sum_{a \in G} \Gamma(a^{-1}) \mathscr{H} \Gamma(a) = \lambda \mathscr{E}. \tag{6}$$

Элемент, стоящий на пересечении p-й строки и q-го столбца матрицы $\sum_{a \in G} \Gamma(a^{-1}) \mathscr{H} \Gamma(a)$, равен

$$\sum_{a \in G} \sum_{i,j=1}^{n} \gamma_{pi}(a^{-1}) h_{ij} \gamma_{jq}(a) = \sum_{i,j=1}^{n} h_{ij} \sum_{a \in G} \gamma_{pi}(a^{-1}) \gamma_{jq}(a).$$

Элемент, стоящий на соответствующем месте в матрице $\lambda \mathscr{E}$, равен $\delta_{pq}\lambda$, где

$$\delta_{pq} = \begin{cases} 1, & \text{если } p = q, \\ 0, & \text{если } p \neq q. \end{cases}$$

Найдем значение λ. Для этого вычислим след обеих частей равенства (6). След правой части равен $\operatorname{tr}(\lambda \mathscr{E}) = \lambda n$. След левой части равен

$$\sum_{a \in G} \operatorname{tr}[(\Gamma(a))^{-1} \mathscr{H} \Gamma(a)] = \sum_{a \in G} \operatorname{tr} \mathscr{H} = k \operatorname{tr} \mathscr{H},$$

(см. стр. 127; здесь k — порядок группы G), и так как $\operatorname{tr} \mathscr{H} = \sum_{i,j=1}^{n} \delta_{ij} h_{ij}$, то

$$\lambda n = k \sum_{i,j=1}^{n} \delta_{ij} h_{ij}.$$

Мы нашли, что $\lambda = \frac{k}{n} \sum_{i,j=1}^{n} \delta_{ij} h_{ij}$ и, значит, элемент, стоящий на пересечении p-й строки и q-го столбца матрицы $\lambda \mathscr{E}$, равен $\delta_{pq} \frac{k}{n} \sum_{i,j=1}^{n} \delta_{ij} h_{ij}$. Таким образом, мы имеем

$$\sum_{i,j=1}^{n} h_{ij} \sum_{a\in G} \gamma_{pi}(a^{-1})\,\gamma_{jq}(a) = \delta_{pq}\frac{k}{n}\sum_{i,j=1}^{n} \delta_{ij}h_{ij},$$

причем это равенство справедливо при всех h_{ij}. Следовательно, при всех i, j, p, q

$$\sum_{a\in G} \gamma_{pi}(a^{-1})\,\gamma_{jq}(a) = \frac{k}{n}\delta_{pq}\delta_{ij}. \tag{7}$$

(Если равенство $\sum\limits_{i,j} \alpha_{ij}h_{ij} = \sum\limits_{i,j} \beta_{ij}h_{ij}$ выполняется при подстановке любых значений h_{ij}, то $\sum\limits_{i,j} (\alpha_{ij} - \beta_{ij})\,h_{ij} = 0$ тоже при любых h_{ij}, и значит, $\alpha_{ij} - \beta_{ij} = 0$, т. е. $\alpha_{ij} = \beta_{ij}$ при всех i и j — см. замечание в скобках на стр. 351.) Пользуясь унитарностью оператора $\Gamma(a)$, равенство (7) можно переписать еще и так:

$$\frac{1}{k}\sum_{a\in G} \bar{\gamma}_{ip}(a)\,\gamma_{jq}(a) = \frac{1}{n}\delta_{pq}\delta_{ij},$$

или

$$(\gamma_{jq}, \gamma_{ip}) = \frac{1}{n}\delta_{pq}\delta_{ij},$$

и значит,

$$(\gamma_{jq}, \gamma_{ip}) = \begin{cases} 0, & \text{если } p \neq q \text{ или } i \neq j, \\ \dfrac{1}{n}, & \text{если } p = q \text{ и } i = j, \end{cases}$$

(ср. с равенствами (4), (5′), (5″) на стр. 346).

ГЛАВА XIII

ТЕОРИЯ ХАРАКТЕРОВ

§ 1. Характер представления. Простейшие свойства характеров

Определение 1. *Пусть* Γ — *линейное представление группы* G *в пространстве* R. *Для каждого* $a \in G$ *положим* $\chi(a) = \operatorname{tr}\Gamma(a)$. *Определенная таким образом на группе* G (*комплексно-значная*) *функция* χ *называется* ***характером*** *представления* Γ.

Если Γ — одномерное представление группы G, то для каждого $a \in G$ имеем $\chi(a) = \operatorname{tr}\Gamma(a) = \Gamma(a)$.

Выпишем характеры всех определенных выше представлений группы $S_3 \simeq D_3$. Пусть $\chi_1 = \Gamma_1$ и $\chi_2 = \Gamma_2$ — характеры одномерных представлений Γ_1 и Γ_2 этой группы (см. стр. 328), χ_3 — характер двумерного представления Γ_3, определенного на стр. 328, χ_4 — характер трехмерного представления, указанного на стр. 329 (назовем его Γ_4) и, наконец, χ_5 — характер регулярного представления Γ_5 этой группы (стр. 340). Все эти характеры можно собрать в такую таблицу:

	e	r	r^2	s	sr	sr^2
χ_1	1	1	1	1	1	1
χ_2	1	1	1	−1	−1	−1
χ_3	2	−1	−1	0	0	0
χ_4	3	0	0	−1	−1	−1
χ_5	6	0	0	0	0	0

Характеры играют очень важную роль в теории представлений. Можно сказать, что характер представления определяет это представление, так как дальше будет показано, что *представления с одинаковыми характерами изоморфны.*

Рассмотрим простейшие свойства характеров.

1. *Для любого представления* $\chi(e) = n$, *где* e — *единица группы* G, *а* n — *степень представления.* Действительно, $\Gamma(e)$ есть единичная матрица порядка n, и значит,

$$\chi(e) = \operatorname{tr} \Gamma(e) = n.$$

2. *Характер является центральной функцией на группе*: $\chi(b^{-1}ab) = \chi(a)$. Действительно,

$$\chi(b^{-1}ab) = \operatorname{tr} \Gamma(b^{-1}ab) = \operatorname{tr} \Gamma(a) = \chi(a)$$

(см. стр. 127).

3. *Изоморфные представления имеют одинаковые характеры.* В самом деле, если

$$\mathscr{H}\Gamma_1 = \Gamma_2\mathscr{H},$$

то

$$\Gamma_1 = \mathscr{H}^{-1}\Gamma_2\mathscr{H}$$

и

$$\chi_1(a) = \operatorname{tr} \Gamma_1(a) = \operatorname{tr}(\mathscr{H}^{-1}\Gamma_2(a)\mathscr{H}) = \operatorname{tr} \Gamma_2(a) = \chi_2(a).$$

4. *Если представление* Γ *является прямой суммой представлений* Γ_1 *и* Γ_2, *то характер* χ *представления* Γ *равен сумме характеров* χ_1 *и* χ_2 *представлений* Γ_1 *и* Γ_2:

$$\chi = \chi_1 + \chi_2.$$

По условию, $R = R_1 \oplus R_2$, где R, R_1, R_2 — пространства представлений Γ, Γ_1, Γ_2. Если базис R выбрать так, чтобы первые n_1 (где n_1 — размерность R_1) векторов принадлежали подпространству R_1, а последние n_2 векторов (где n_2 — размерность R_2) — подпространству R_2, то матрица представления $\Gamma(a)$ будет иметь вид

$$\Gamma(a) = \begin{bmatrix} \Gamma_1(a) & O \\ O & \Gamma_2(a) \end{bmatrix},$$

где $\Gamma_1(a)$ — матрица представления Γ_1 в пространстве R_1, а $\Gamma_2(a)$ — матрица представления Γ_2 в пространстве R_2. Но в таком случае, очевидно, для каждого $a \in G$

$$\chi(a) = \operatorname{tr}\Gamma(a) = \operatorname{tr}\Gamma_1(a) + \operatorname{tr}\Gamma_2(a) = \chi_1(a) + \chi_2(a),$$

т. е.

$$\chi = \chi_1 + \chi_2.$$

5. *Для любого* $a \in G$

$$\chi(a^{-1}) = \overline{\chi(a)}.$$

Пусть $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все собственные значения оператора $\Gamma(a)$, причем каждое взято столько раз, какова его кратность. Так как оператор $\Gamma(a)$ — унитарный, то $\Gamma(a^{-1}) = [\Gamma(a)]^{-1} = [\Gamma(a)]^*$, и значит, собственные значения оператора $\Gamma(a^{-1})$ совпадают с собственными значениями $[\Gamma(a)]^*$. Но, ввиду следствия на стр. 168, собственными значениями оператора $[\Gamma(a)]^*$ будут

$$\overline{\lambda}_1, \overline{\lambda}_2, \ldots, \overline{\lambda}_n$$

(где также каждое собственное значение взято столько раз, какова его кратность). Следовательно,

$$\chi(a^{-1}) = \operatorname{tr}\Gamma(a^{-1}) = \operatorname{tr}[\Gamma(a)]^* = \overline{\lambda}_1 + \overline{\lambda}_2 + \ldots + \overline{\lambda}_n = \overline{\operatorname{tr}\Gamma(a)} = \overline{\chi(a)}.$$

6. *Если* χ_0 — *характер регулярного представления* Γ_0 *группы* G *порядка* k, *то*

$$\chi_0(a) = \begin{cases} k, & \text{если } a = \varepsilon, \\ 0, & \text{если } a \neq \varepsilon. \end{cases}$$

Действительно, пусть $a_1 = \varepsilon, a_2, \ldots, a_k$ — все элементы группы G, и базис пространства представления образован векторами $e_{a_1}, e_{a_2}, \ldots, e_{a_k}$. Так как $a_1 = \varepsilon$ — единичный элемент, то

$$\chi_0(a_1) = \operatorname{tr}\Gamma_0(a_1) = k$$

(п. 1). Если $a_i \neq a_1 = \varepsilon$, то

$$\Gamma_0(a_i)\, e_{a_p} = e_{a_i a_p} \neq e_{a_p},$$

так как из равенства $a_i a_p = a_p$ вытекало бы, что $a_i =$

$\neq \varepsilon = a_1$. Следовательно, при $i \neq 1$ базисный вектор e_{a_p} оператором $\Gamma_0(a_i)$ переводится снова в базисный вектор $e_{a_i a_p}$, однако — в вектор, отличный от него самого. Это значит, что для любого p в p-м столбце матрицы $\Gamma_0(a_i)$ единственный отличный от нуля элемент — это единица, стоящая не на главной диагонали, а следовательно, все элементы главной диагонали такой матрицы равны нулю, и значит, след ее равен нулю.

§ 2. Характеры неприводимых представлений

Этот параграф, содержащий дальнейшие следствия леммы Шура, можно назвать ее «Эпилогом».

Теорема 1. *Характеры неприводимых не изоморфных между собой представлений конечной группы образуют ортонормированную систему функций* (отсюда, в частности, будет следовать, что конечная группа имеет конечное число неприводимых представлений).

Как и лемма Шура, эта теорема состоит из двух частей:

I. *Если χ_1 и χ_2 — характеры неприводимых не изоморфных между собой представлений Γ_1 и Γ_2 группы G, то* $(\chi_1, \chi_2) = 0$.

Действительно, пусть k — порядок группы G.

Тогда

$$(\chi_1, \chi_2) = \frac{1}{k}\sum_{a\in G}\chi_1(a)\overline{\chi_2(a)} = \frac{1}{k}\sum_{a\in G}\operatorname{tr}\Gamma_1(a)\,\overline{\operatorname{tr}\Gamma_2(a)} =$$

$$= \frac{1}{k}\sum_{a\in G}\sum_{i}\gamma_{ii}^{(1)}(a)\overline{\sum_{j}\gamma_{jj}^{(2)}(a)} = \frac{1}{k}\sum_{a\in G}\sum_{i,j}\gamma_{ii}^{(1)}(a)\overline{\gamma_{jj}^{(2)}(a)} =$$

$$= \sum_{i,j}\frac{1}{k}\sum_{a\in G}\gamma_{ii}^{(1)}(a)\overline{\gamma_{jj}^{(2)}(a)} = \sum_{i,j}\left(\gamma_{ii}^{(1)}, \gamma_{jj}^{(2)}\right) = 0,$$

так как каждое слагаемое этой суммы равно нулю (см. стр. 351).

II. *Если Γ — неприводимое представление группы G с характером χ, то* $(\chi, \chi) = 1$.

Действительно,

$$(\chi, \chi) = \frac{1}{k} \sum_{a \in G} \chi(a) \overline{\chi(a)} = \frac{1}{k} \sum_{a \in G} \operatorname{tr} \Gamma(a) \overline{\operatorname{tr} \Gamma(a)} =$$

$$= \frac{1}{k} \sum_{a \in G} \sum_{i=1}^{n} \gamma_{ii}(a) \sum_{j=1}^{n} \overline{\gamma_{jj}(a)} = \frac{1}{k} \sum_{a \in G} \sum_{i,j=1}^{n} \gamma_{ii}(a) \overline{\gamma_{jj}(a)} =$$

$$= \sum_{i,j=1}^{n} \frac{1}{k} \sum_{a \in G} \gamma_{ii}(a) \overline{\gamma_{jj}(a)} =$$

$$= \sum_{i,j=1}^{n} (\gamma_{ii}, \gamma_{jj}) = \sum_{i=1}^{n} (\gamma_{ii}, \gamma_{ii}) = \sum_{i=1}^{n} \frac{1}{n} = 1,$$

так как $(\gamma_{ii}, \gamma_{jj}) = 0$ при $i \neq j$, а $(\gamma_{ii}, \gamma_{ii}) = \frac{1}{n}$ (см. стр. 353). Теорема доказана.

Так, для группы $S_3 \simeq D_3$ (см. таблицу на стр. 354) имеем

$$(\chi_1, \chi_2) = \frac{1}{6}(1 + 1 + 1 - 1 - 1 - 1) = 0, \quad (\chi_1, \chi_3) =$$

$$= \frac{1}{6}(2 - 1 - 1) = 0, \quad (\chi_2, \chi_3) = 0, \quad (\chi_1, \chi_1) =$$

$$= \frac{1}{6}(1 + 1 + 1 + 1 + 1 + 1) = 1, \quad (\chi_2, \chi_2) = 1, \quad (\chi_3, \chi_3) = 1.$$

Но, с другой стороны,

$$(\chi_4, \chi_4) = \frac{1}{6}(9 + 1 + 1 + 1) = 2 \text{ и } (\chi_5, \chi_5) = \frac{1}{6} \cdot 36 = 6;$$

значит, представления Γ_4 и Γ_5 этой группы не являются неприводимыми.

Найдем *скалярный квадрат характера* χ_0 *регулярного представления* произвольной группы G порядка k. Выше мы видели, что $\chi_0(a_1) = k$ и $\chi_0(a_i) = 0$, если $i \neq 1$. Следовательно,

$$(\chi_0, \chi_0) = \frac{1}{k} \cdot k \cdot k = k;$$

отсюда снова получаем, что регулярное представление любой (не состоящей из одной единицы) группы приводимо (ср. стр. 341).

§ 3. Дальнейшие свойства характеров

Пусть Γ — линейное представление группы G в пространстве R с характером χ. Разложим Γ в прямую сумму неприводимых представлений Γ_i — пусть

$$\Gamma = \Gamma_1 \oplus \Gamma_2 \oplus \ldots \oplus \Gamma_s.$$

Характер (неприводимого) представления Γ_i (где $i = 1, 2, \ldots, s$) обозначим через χ_i. Тогда характер представления Γ

$$\chi = \chi_1 + \chi_2 + \ldots + \chi_s.$$

Среди представлений Γ_i могут быть и изоморфные между собой, им отвечают равные характеры. Наоборот, характеры неизоморфных между собой неприводимых представлений не могут быть равными, так как если $\chi_1 = \chi_2$, то $(\chi_1, \chi_2) = (\chi_1, \chi_1) = 1$, а не 0, как должно быть по п. 1 предыдущего параграфа. Объединяя слагаемые, отвечающие изоморфным представлениям, последнюю сумму можно переписать так:

$$\chi = m_1\chi_1 + m_2\chi_2 + \ldots + m_q\chi_q, \tag{1}$$

где m_j — «кратность» представления Γ_j (здесь $j = 1, 2, \ldots, q$ и представления $\Gamma_1, \Gamma_2, \ldots, \Gamma_q$ попарно не изоморфны). Пусть теперь Γ' — произвольное *неприводимое* представление группы G с характером χ'. Рассмотрим скалярное произведение

$$(\chi', \chi) = (\chi', m_1\chi_1 + m_2\chi_2 + \ldots + m_q\chi_q) =$$
$$= m_1(\chi', \chi_1) + m_2(\chi', \chi_2) + \ldots + m_q(\chi', \chi_q).$$

Но $(\chi', \chi_i) = 0$, если Γ' не изоморфно Γ_i и $(\chi', \chi_i) = 1$, если представления Γ' и Γ_i изоморфны. Следовательно, *скалярное произведение* (χ', χ) *обязательно является целым числом, которое показывает, сколько раз неприводимое представление* Γ' *содержится в* Γ.

Следствие 1. *Представления* Γ_1 *и* Γ_2, *имеющие одинаковые характеры, изоморфны*, так как каждое неприводимое представление Γ' и в Γ_1, и в Γ_2 содержится одинаковое число раз (см. конец § 4 гл. XII).

Следствие 2. *Разложение представления* Γ *в прямую сумму неприводимых представлений* Γ_i *(с точностью до изоморфизма слагаемых) однозначно.*

Вернемся к (вообще говоря, приводимому) представлению Г с характером χ. Из равенства (1) вытекает, что скалярный квадрат характера χ равен

$$(\chi, \chi) = m_1^2 + m_2^2 + \ldots + m_q^2.$$

Следовательно, *скалярный квадрат характера всегда является целым числом, которое в* *т о м* *и* *т о л ь к о* *в* *т о м* *случае равно* 1, *если представление* Г *н е п р и в о д и м о*.

Можно сказать поэтому, что *для того, чтобы представление было неприводимым, необходимо и достаточно, чтобы скалярный квадрат его характера был равен* 1. (Значит, в частности, представления Γ_1, Γ_2 и Γ_3 группы D_3 неприводимы.) Далее, *для того чтобы два неприводимых представления группы были неизоморфны, необходимо и достаточно, чтобы скалярное произведение их характеров было равно нулю* (необходимость этого была доказана выше, а достаточность вытекает из того, что если два неприводимых представления изоморфны, то их характеры равны: $\chi_1 = \chi_2$ и, значит, скалярное произведение $(\chi_1, \chi_2) = (\chi_1, \chi_1) = 1$).

Пример. Для группы $S_3 \simeq D_3$ (см. стр. 354) имеем

$$(\chi_1, \chi_4) = \frac{1}{6}(3 - 1 - 1 - 1) = 0, \quad (\chi_2, \chi_4) = \frac{1}{6}(3 + 1 + 1 + 1) = 1,$$

$$(\chi_3, \chi_4) = \frac{1}{6}\cdot 6 = 1;$$

и представление Γ_4 равно прямой сумме Γ_2 и Γ_3. Далее,

$$(\chi_1, \chi_5) = \frac{1}{6}\cdot 6 = 1, \quad (\chi_2, \chi_5) = \frac{1}{6}\cdot 6 = 1, \quad (\chi_3, \chi_5) = \frac{1}{6}\cdot 12 = 2,$$

и значит, $\chi_5 = \chi_1 + \chi_2 + 2\chi_3$ — регулярное представление Γ_5 содержит по одному разу каждое из одномерных представлений Γ_1 и Γ_2 и два раза — двумерное представление Γ_3.

§ 4. Основное соотношение

Лемма. Пусть Γ_0 — *р е г у л я р н о е представление группы* G. *Тогда каждое неприводимое представление* Γ_i *этой группы содержится в* Γ_0 *столько раз, какова его степень.*

Доказательство. Пусть χ_0 — характер (регулярного) представления Γ_0 и Γ_i — произвольное неприводимое представление группы G степени n_i с характером χ_i. Тогда

$$(\chi_i, \chi_0) = \frac{1}{k} \sum_{a \in G} \chi_i(a) \overline{\chi_0(a)} = \frac{1}{k} \chi_i(e) \overline{\chi_0(e)} = \chi_i(e) = n_i.$$

Но, как мы уже знаем (см. § 3), скалярное произведение (χ_i, χ_0) показывает, сколько раз (неприводимое) представление Γ_i содержится в (регулярном) представлении Γ_0.

Теорема 2. *Сумма квадратов степеней всех неприводимых (не изоморфных между собой) представлений конечной группы равна порядку группы.* Иными словами, если G — (конечная) группа порядка k и n_1, $n_2, \ldots, n_p$ — степени всех ее неприводимых представлений, то

$$n_1^2 + n_2^2 + \ldots + n_p^2 = k.$$

Доказательство. Пусть Γ_0 — регулярное представление группы G. Если $\Gamma_1, \Gamma_2, \ldots, \Gamma_p$ — все неприводимые представления группы G, $\chi_1, \chi_2, \ldots, \chi_p$ — их характеры и $n_1, n_2, \ldots, n_p$ — их степени, а χ_0 — характер регулярного представления Γ_0, то

$$\chi_0 = n_1\chi_1 + n_2\chi_2 + \ldots + n_p\chi_p.$$

Скалярный квадрат χ_0 равен

$$(\chi_0, \chi_0) = n_1^2 + n_2^2 + \ldots + n_p^2.$$

Но, по доказанному в конце § 2, скалярный квадрат (χ_0, χ_0) регулярного представления равен k. Следовательно,

$$n_1^2 + n_2^2 + \ldots + n_p^2 = k.$$

Заметим, что среди чисел n_i могут быть и равные. Так, для группы D_3 имеем $n_1 = n_2 = 1$ и $n_3 = 2$. Поскольку порядок группы D_3 равен 6 и $1^2 + 1^2 + 2^2 = 6$, то эти представления Γ_1, Γ_2 и Γ_3 — это все неприводимые представления группы D_3.

Так как у каждой группы имеется единичное представление, то среди чисел n_i по крайней мере одно равно 1.

Можно доказать, что *все n_i являются делителями порядка k группы.*

§ 5. Число неприводимых представлений группы

Лемма. *Пусть $f(a)$ — центральная функция, определенная на группе G порядка k и Γ — неприводимое представление группы G в пространстве R размерности n с характером χ. Тогда линейный оператор*

$$\mathscr{H}(f)=\sum_{a\in G}\overline{f(a)}\,\Gamma(a),$$

действующий в пространстве R, является гомотетией с коэффициентом гомотетии, равным $\frac{k}{n}(\chi, f)$.

Доказательство. Для любого элемента b группы G имеем

$$\Gamma^{-1}(b)\,\mathscr{H}(f)\,\Gamma(b)=\sum_{a\in G}\Gamma^{-1}(b)\,\overline{f(a)}\,\Gamma(a)\,\Gamma(b)=$$
$$=\sum_{a\in G}\overline{f(a)}\,\Gamma(b^{-1})\,\Gamma(a)\,\Gamma(b)=\sum_{a\in G}\overline{f(a)}\,\Gamma(b^{-1}ab)=$$
$$=\sum_{a\in G}\overline{f(b^{-1}ab)}\,\Gamma(b^{-1}ab)$$

(так как $f(a)$ — центральная функция, то $f(a)=f(b^{-1}ab)$). Но произведение $b^{-1}ab$, где b — фиксированный элемент группы G, а a пробегает все элементы группы, тоже пробегает (по одному разу) все элементы группы (произвольный элемент x группы равен $x=b^{-1}(bxb^{-1})b$; здесь $a=bxb^{-1}$). Следовательно, последняя сумма равна

$$\sum_{a\in G}\overline{f(a)}\,\Gamma(a)=\mathscr{H}(f),$$

т. е. для любого $b\in G$

$$\Gamma^{-1}(b)\,\mathscr{H}(f)\,\Gamma(b)=\mathscr{H}(f).$$

Из равенства $\Gamma^{-1}\mathscr{H}(f)\Gamma=\mathscr{H}(f)$ вытекает, что $\mathscr{H}(f)\Gamma=\Gamma\mathscr{H}(f)$. Но тогда из второй части леммы Шура следует, что $\mathscr{H}(f)$ является гомотетией, т. е. что для некоторого λ

$$\sum_{a\in G}\overline{f(a)}\,\Gamma(a)=\lambda\mathscr{E}.$$

Для того чтобы найти λ, вычислим след обеих частей этого равенства. В правой части получим $\operatorname{tr}\lambda\mathscr{E}=\lambda n$, След левой части равен

$$\operatorname{tr} \sum_{a \in G} \overline{f(a)}\, \Gamma(a) = \sum_{a \in G} \operatorname{tr} [\overline{f(a)}\, \Gamma(a)] =$$
$$= \sum_{a \in G} \overline{f(a)} \operatorname{tr} \Gamma(a) = \sum_{a \in G} \overline{f(a)}\, \chi(a) = k(\chi, f).$$

Следовательно, $\lambda = \frac{k}{n}(\chi, f)$ и

$$\sum_{a \in G} \overline{f(a)}\, \Gamma(a) = \frac{k}{n}(\chi, f)\, \mathscr{E}.$$

Теорема 3. *Центральная функция f, определенная на группе G и ортогональная ко всем характерам неприводимых представлений Γ_i этой группы, тождественно равна нулю.*

Доказательство. Пусть $\chi_1, \chi_2, \ldots, \chi_p$ — характеры всех неприводимых представлений группы G. Предположим, что $(\chi_i, f) = 0$ для каждого $i = 1, 2, \ldots, p$. Тогда для каждого неприводимого представления Γ_i группы

$$\sum_{a \in G} \overline{f(a)}\, \Gamma_i(a) = \frac{k}{n}(\chi_i, f)\, \mathscr{E}_i = \mathcal{O},$$

где $\mathcal{O}$ есть нулевой оператор. Покажем, что сумма $\sum_{a \in G} \overline{f(a)}\ \Gamma(a) = \mathcal{O}$ и в том случае, когда представление Γ приводимо.

Действительно, пусть Γ является суммой, например, двух (неприводимых) представлений Γ_1 и Γ_2. Тогда в соответственно выбранном базисе матрицы представления Γ имеют вид

$$\Gamma(a) = \begin{bmatrix} \Gamma_1(a) & 0 \\ 0 & \Gamma_2(a) \end{bmatrix},$$

где $\Gamma_i(a)$ — матрица представления Γ_i в пространстве представления R_i, $i = 1, 2$. Но так как

$$\sum_{a \in G} \overline{f(a)}\, \Gamma_1(a) = \mathcal{O} \quad \text{и} \quad \sum_{a \in G} \overline{f(a)}\, \Gamma_2(a) = \mathcal{O},$$

то и

$$\sum_{a\in G}\overline{f(a)}\,\Gamma(a)=\sum_{a\in G}\overline{f(a)}\begin{bmatrix}\Gamma_1(a) & O\\ O & \Gamma_2(a)\end{bmatrix}=$$
$$=\sum_{a\in G}\begin{bmatrix}\overline{f(a)}\,\Gamma_1(a) & O\\ O & \overline{f(a)}\,\Gamma_2(a)\end{bmatrix}=$$
$$=\begin{bmatrix}\sum\limits_{a\in G}\overline{f(a)}\,\Gamma_1(a) & O\\ O & \sum\limits_{a\in G}\overline{f(a)}\,\Gamma_2(a)\end{bmatrix}=O.$$

Аналогичное утверждение верно и для любого числа слагаемых, т. е. *для любого представления* Γ *группы* G *и любой центральной функции* f, *ортогональной ко всем характерам неприводимых представлений этой группы, имеем*

$$\sum_{a\in G}\overline{f(a)}\,\Gamma(a)=\mathcal{O}.$$

Пусть теперь Γ_0 — регулярное представление группы G, $a_1, a_2, \ldots, a_k$ — все ее элементы ($a_1 = \varepsilon$ — единичный элемент), $e_{a_1}, e_{a_2}, \ldots, e_{a_k}$ — базисные векторы пространства представления R. Применим оператор $\sum\limits_{a\in G}\overline{f(a)}\,\Gamma_0(a)$ (являющийся, по доказанному выше, нулевым оператором) к вектору e_{a_1}. Мы получим

$$\sum_{a_i\in G}\overline{f(a_i)}\,\Gamma_0(a_i)\,e_{a_1}=\sum_{a_i\in G}\overline{f(a_i)}\,e_{a_i}=\mathcal{O}.$$

Но так как векторы e_{a_i} линейно независимы, то $\overline{f(a_i)}=0$ для всех $i=1, 2, \ldots, k$, и значит, $f(a_i)=0$ для всех $a_i\in G$, т. е. функция f на группе G тождественно равна нулю.

Теорема 4. *Каждая центральная функция на группе* G *является линейной комбинацией характеров неприводимых представлений этой группы.*

Доказательство. Пусть V — пространство всех центральных функций, определенных на группе G, и $V_1 \rightleftharpoons \{\chi_1, \chi_2, \ldots, \chi_p\}$ — его подпространство, порожденное характерами всех неприводимых представлений. Тогда пространство V равно прямой сумме подпространства V_1 и его ортогонального дополнения $V_1^{\perp}$:

$$V = V_1 \oplus V_1^{\perp}.$$

Однако $V_1^{\perp} = \{0\}$, так как каждая центральная функция, ортогональная ко всем χ_i, равна нулю. Следовательно, $V = V_1$, т. е. *характеры неприводимых представлений группы G образуют в пространстве всех определенных на ней центральных функций (ортонормированный) базис — каждая центральная функция является линейной комбинацией характеров.*

Теорема 5. *Число попарно неизоморфных неприводимых представлений группы G равно числу классов сопряженных элементов этой группы.*

По предыдущей теореме, размерность пространства всех центральных функций, определенных на группе G, равна числу неприводимых и неизоморфных между собой представлений этой группы. В то же время размерность пространства центральных функций равна числу p классов сопряженных элементов группы G (стр. 343), откуда и следует утверждение теоремы.

Окончательно мы имеем соотношение

$$k = n_1^2 + n_2^2 + \ldots + n_p^2,$$

где k — порядок группы G, n_i — степень неприводимого представления Γ_i, p — число классов сопряженных элементов группы. Все n_i являются делителями k и по крайней мере одно из них равно 1.

§ 6. Представления коммутативной группы

Если группа G коммутативна (и только в этом случае), число p ее классов сопряженных элементов равно порядку k группы. Из равенства $k = n_1^2 + n_2^2 + \ldots + n_k^2$ следует, что все (целые положительные) числа n_i равны 1, т. е. что *все неприводимые представления коммутативной группы одномерны*, причем *число их равно порядку группы.* Обратно, если все неприводимые представления группы одномерны, т. е. если все $n_i = 1$, то число p классов сопряженных элементов равно порядку k группы, и группа коммутативна.

Итак, *для того чтобы группа G была коммутативной, необходимо и достаточно, чтобы все ее неприводимые представления были одномерны* (имели степень 1).

Из этого результата, в частности, вытекает, что для групп C_2, C_4, V выше (на стр. 326—327) были найдены в с е их неприводимые представления.

§ 7. Представления циклических групп

Пусть C_n — циклическая группа порядка n, состоящая из элементов $e, a, a^2, \ldots, a^{n-1}$, где $a^n = e$. Так как группа C_n коммутативна, то все ее неприводимые представления о д н о м е р н ы, причем число неизоморфных между собой неприводимых представлений равно n. Пусть $\Gamma(a) = \alpha$; тогда $\Gamma(a^2) = \alpha^2$, $\Gamma(a^3) = \alpha^3, \ldots$, наконец, $\Gamma(a^n) = \alpha^n = \Gamma(e) = 1$. Следовательно, $\alpha^n = 1$ и α есть корень n-й степени из 1. Положим

$$\alpha = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}.$$

Тогда

$$1, \alpha, \alpha^2 = \cos\frac{2\pi}{n}2 + i\sin\frac{2\pi}{n}2, \ldots$$

$$\ldots, \alpha^{n-1} = \cos\frac{2\pi}{n}(n-1) + i\sin\frac{2\pi}{n}(n-1)$$

— все корни n-й степени из 1. Это дает n одномерных представлений группы C_n, совпадающих со своими характерами $\chi_1, \chi_2, \ldots, \chi_n$.

C_n	e	a	a^2		a^{n-1}
χ_1	1	1	1	$, \ldots ,$	1
χ_2	1	α	α^2	$, \ldots ,$	α^{n-1}
χ_3	1	α^2	α^4	$, \ldots ,$	$\alpha^{(n-1)2}$
	...	...	...	$, \ldots ,$	...
	...	...	...	$, \ldots ,$	...
χ_n	1	α^{n-1}	$\alpha^{(n-1)2}$	$, \ldots ,$	$\alpha^{(n-1)(n-1)}$

Полученные представления, будучи одномерными,— неприводимы:

Так как $\chi_1, \chi_2, \ldots, \chi_n$ — различны, то эти n представлений попарно неизоморфны (п. 3, стр. 355).

Таким образом, мы нашли все n неизоморфных между собой неприводимых представлений циклической группы C_n.

§ 8. Представления диэдральных групп

Диэдральная группа D_n при четном n имеет $\frac{n}{2}+3$ классов сопряженных элементов, а при нечетном n число классов сопряженных элементов этой группы равно $\frac{n+3}{2}$.

Для четного n имеем

$$2n = 1^2 \cdot 4 + 2^2\left(\frac{n}{2}-1\right)$$

(число слагаемых в правой части равно $4+\frac{n}{2}-1 = \frac{n}{2}+3$) — и, значит, группа D_n имеет 4 одномерных и $\frac{n}{2}-1$ двумерных неприводимых представлений.

Для нечетного n

$$2n = 1^2 \cdot 2 + 2^2\,\frac{n-1}{2}$$

(число слагаемых здесь равно $2+\frac{n-1}{2}=\frac{n+3}{2}$) — и группа имеет два одномерных и $\frac{n-1}{2}$ двумерных неприводимых представлений.

I. Рассмотрим сначала случай, когда $n = 2m$ четно. Элементы группы $D_n = D_{2m}$ выше мы обозначали так:

$$e,\ r,\ r^2,\ \ldots,\ r^{2m-1},\ s,\ sr,\ sr^2,\ \ldots,\ sr^{2m-1}.$$

Они следующим образом разбиваются на $\frac{n}{2}+3 = m+3$

классов сопряженных элементов:

$$\{e\},\ \{r,\ r^{2m-1}\},\ \{r^2,\ r^{2m-2}\},\ \ldots,\ \{r^{m-1},\ r^{m+1}\},\ \{r^m\}$$
$$\{s,\ sr^2,\ \ldots,\ sr^{2m-2}\},\ \{sr,\ sr^3,\ \ldots,\ sr^{2m-1}\}.$$

Обратим внимание на то, что элементы r^{m-1} и r^{m+1} сопряжены между собой.

Пусть Γ — одномерное представление группы D_n с характером χ и $\Gamma(r)=\chi(r)=\alpha$. Тогда

$$\Gamma(r^{m-1})=\chi(r^{m-1})=\alpha^{m-1} \text{ и } \Gamma(r^{m+1})=\chi(r^{m+1})=\alpha^{m+1}.$$

Но характер на сопряженных элементах принимает одинаковые значения; поэтому $\alpha^{m-1}=\alpha^{m+1}$. Так как $\alpha\neq 0$, то $\alpha^2=1$ и $\alpha=\pm 1$. Если, далее, $\Gamma(s)=\beta$, то $\Gamma(s^2)=\beta^2=1$, т. е. $\beta=\pm 1$. Это дает 4 одномерных представления (совпадающих со своими характерами):

D_{2m}	e	r	r^2	r^3		r^{2m-1}	s	sr	sr^2		sr^{2m-1}
χ_1	1	1	1	1	...	1	1	1	1	...	1
χ_2	1	1	1	1	...	1	−1	−1	−1	...	−1
χ_3	1	−1	1	−1	...	−1	1	−1	1	...	−1
χ_4	1	−1	1	−1	...	−1	−1	1	−1	...	1

;

или короче,

D_{2m}	r^k	sr^k
χ_1	1	1
χ_2	1	−1
χ_3	$(-1)^k$	$(-1)^k$
χ_4	$(-1)^l$	$(-1)^{k+1}$

Легко видеть, что эти представления неприводимы и попарно неизоморфны.

Для того чтобы найти двумерные представления группы D_n, заметим, что так как D_n является группой преобразований плоскости, то она сама и будет одним из своих представлений (основным). Выпишем соответствующие матрицы.

Если r — поворот вокруг начала координат на угол $\frac{2\pi}{n}$, а s — симметрия, скажем, относительно оси Ox, то

$$\Gamma(e)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\quad \Gamma(r^k)=\begin{bmatrix}\cos\frac{2\pi}{n}k & -\sin\frac{2\pi}{n}k\\ \sin\frac{2\pi}{n}k & \cos\frac{2\pi}{k}k\end{bmatrix},$$

$$\Gamma(s)=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},\quad \Gamma(sr^k)=\begin{bmatrix}\cos\frac{2\pi}{n}k & -\sin\frac{2\pi}{n}k\\ -\sin\frac{2\pi}{n}k & -\cos\frac{2\pi}{n}k\end{bmatrix},$$

где $k=0, 1, 2, \ldots, n-1$.

Другие представления этой группы можно получить, отнеся элементу r поворот на угол $\frac{2\pi}{n}\cdot 2$, на угол $\frac{2\pi}{n}\cdot 3, \ldots$, наконец, на угол $\frac{2\pi}{n}(m-1)$. Так мы получим $m-1=\frac{n}{2}-1$ представлений $\Gamma_1, \Gamma_2, \ldots, \Gamma_{m-1}$, где

$$\Gamma_h(r^k)=\begin{bmatrix}\cos\frac{2\pi}{n}hk & -\sin\frac{2\pi}{n}hk\\ \sin\frac{2\pi}{n}hk & \cos\frac{2\pi}{n}hk\end{bmatrix},$$

$$\Gamma_h(sr^k)=\begin{bmatrix}\cos\frac{2\pi}{n}hk & -\sin\frac{2\pi}{n}hk\\ -\sin\frac{2\pi}{n}hk & -\cos\frac{2\pi}{n}hk\end{bmatrix};$$

здесь $h=1, 2, \ldots, m-1$, $k=0, 1, 2, \ldots, n-1$ (при $h=1$ получается выписанное выше представление $\Gamma=\Gamma_1$).

Характер представления Γ_h:

$$\chi_h(r^k)=2\cos\frac{2\pi}{n}hk,\quad \chi_h(sr^k)=0.$$

Кроме найденных $m-1$ представлений $\Gamma_1, \Gamma_2, \ldots, \Gamma_{m-1}$, каких-либо новых представлений мы таким путем не получим, так как если элементу r поставить в соответствие поворот на угол $\frac{2\pi}{n}\cdot m=\pi$, то это представление будет приводимым, ибо при таком преобразовании (центральная симметрия) все векторы пространства являются собственными. При поворотах же на углы $\frac{2\pi}{n}(m+1)$, $\frac{2\pi}{n}(m+2), \ldots$ $\ldots$, и т. д. получатся представления, изоморфные уже найденным представлениям Γ_h.

Итак, мы нашли все $\frac{n}{2}-1$ двумерных представлений группы D_n. Эти представления **неприводимы**, так как одномерных подпространств, инвариантных, например, относительно поворота на угол $\frac{2\pi}{n}h$, где $h = 1, 2, \ldots, m-1$, не существует. Ясно также, что найденные представления Γ_h попарно **не изоморфны** (п. 3, стр. 355).

II. Пусть теперь $n = 2m+1$ **нечетно**. Элементы группы D_n следующим образом разбиваются на классы сопряженных элементов:

$$\{e\},\ \{r, r^{2m}\},\ \{r^2, r^{2m-1}\},\ \ldots,\ \{r^m, r^{m+1}\}, \{s, sr, \ldots, sr^{2m-1}\}.$$

Найдем сначала два **одномерных** представления. Если $\Gamma(r) = \alpha$, то $\Gamma(r^m) = \alpha^m$, $\Gamma(r^{m+1}) = \alpha^{m+1}$, и так как элементы r^m и r^{m+1} сопряжены между собой, то $\alpha^m = \alpha^{m+1}$, откуда $\alpha = 1$. Далее, если $\Gamma(s) = \beta$, то $\Gamma(s^2) = \beta^2 = 1$ и $\beta = \pm 1$. Так мы получаем два одномерных представления (совпадающие со своими характерами):

D_{2m+1}	e	r	r^2		r^{2m}	s	sr		sr^{2m}
χ_1	1	1	1	...	1	1	1	...	1
χ_2	1	1	1	...	1	-1	-1	...	-1

или, короче,

D_{2m+1}	r^k	sr^k
χ_1	1	1
χ_2	1	-1

Двумерные представления этой группы находятся так же, как в случае четного n (см. п. I). А именно, имеем

$$\Gamma_h(r^k) = \begin{bmatrix} \cos\frac{2\pi}{n}hk & -\sin\frac{2\pi}{n}hk \\ \sin\frac{2\pi}{n}hk & \cos\frac{2\pi}{n}hk \end{bmatrix},$$

$$\Gamma_h(sr^k) = \begin{bmatrix} \cos\frac{2\pi}{n}hk & -\sin\frac{2\pi}{n}hk \\ -\sin\frac{2\pi}{n}hk & -\cos\frac{2\pi}{n}hk \end{bmatrix},$$

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

Характер h-го представления

$$\chi_h(r^k) = 2\cos\frac{2\pi}{n}hk, \quad \chi_h(sr^k) = 0.$$

Все эти представления неприводимы и попарно не изоморфны.

Выпишем полностью таблицы характеров для групп D_4, D_5, D_6. При $n = 4$ группа D_4 имеет 4 одномерных и одно двумерное представления.

Их характеры:

D_4	e	r, r^3	r^2	s, sr^2	sr, sr^3
χ_1	1	1	1	1	1
χ_2	1	1	1	−1	−1
χ_3	1	−1	1	1	−1
χ_4	1	−1	1	−1	1
χ_5	2	0	−2	0	0

При $n = 6$ группа имеет 4 одномерных и 2 двумерных представления с характерами:

D_6	e	r, r^5	r^2, r^4	r^3	s, sr^2, sr^4	sr, sr^3, sr^5
χ_1	1	1	1	1	1	1
χ_2	1	1	1	1	-1	-1
χ_3	1	-1	1	-1	1	-1
χ_4	1	-1	1	-1	-1	1
χ_5	2	1	-1	-2	0	0
χ_6	2	-1	-1	2	0	0

Наконец, группа D_5 имеет два одномерных и два двумерных представления. Их характеры:

D_5	e	r, r^4	r^2, r^3	sr^k, $k=0, 1, 2, 3, 4$
χ_1	1	1	1	1
χ_2	1	1	1	-1
χ_3	2	$2\cos\frac{2\pi}{5}$	$2\cos\frac{4\pi}{5}$	0
χ_4	2	$2\cos\frac{4\pi}{5}$	$2\cos\frac{2\pi}{5}$	0

§ 9. Характеры группы вращений тетраэдра

Число классов сопряженных элементов этой группы равно 4, порядок ее равен 12. Так как

$$12 = 1^2 + 1^2 + 1^2 + 3^2,$$

то группа имеет три одномерных и одно трехмерное представления.

Найдем одномерные представления. Пусть $\Gamma(c_3) = \alpha$. Тогда $\Gamma(c_3^2) = \alpha^2$ и $\Gamma(c_3^3) = \alpha^3 = 1$. Следовательно, α — кубический корень из 1. Пусть $\Gamma(c_2) = \beta$; тогда $\Gamma(c_2^2) = \beta^2 = 1$ и $\beta = \pm 1$. Но характер каждого неприводимого неединичного представления должен быть ортогонален характеру единичного представления, т. е.

$$1 + 4\alpha + 4\alpha^2 + 3\beta = 0,$$

или (2)

$$4(1 + \alpha + \alpha^2) + 3(\beta - 1) = 0.$$

А так как $\alpha^3 = 1$, то при $\alpha \neq 1$ имеем $1 + \alpha + \alpha^2 = 0$, и, значит, $\beta = 1$ (Заметим, что для неединичного представления α не может равняться 1, так как в этом случае из равенства (2) вытекало бы, что $\beta = -3$, что невозможно).

Итак, мы нашли характеры трех одномерных представлений группы T:

T	e	$4c_3$	$4c_3^2$	$3c_2$
χ_1	1	1	1	1
χ_2	1	α	α^2	1
χ_3	1	α^2	α	1

где $\alpha = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}$. Эти представления неприводимы и попарно неизоморфны.

Далее, так как группа T является группой преобразований трехмерного пространства, то она и будет одним из своих представлений. Мы не будем находить матрицы этого (основного) представления в каком-нибудь фиксированном базисе, но ограничимся тем, что найдем только **характер** представления. Это приводит к значительным упрощениям, потому что след матрицы линейного преобразования не зависит от базиса, и мы можем для каждого преобразования выбирать базис так, чтобы **его** матрица выглядела возможно проще. Особенно удобно принимать ось вращения за одну из координатных осей.

Так, если принять ось вращения за ось Oz (а оси Ox и Oy, ортогональные оси Oz и между собой, выбрать произвольным образом), то матрицы поворотов на углы $\frac{2\pi}{3}$ и $\frac{4\pi}{3}$ соответственно будут иметь вид

$$\begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} & 0 \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Следы их равны нулю. Матрица поворота вокруг оси Oz на угол π

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

имеет след, равный -1. Это дает характер **трехмерного** представления группы T:

T	e	$4c_3$	$4c_3^2$	$3c_2$
χ_4	3	0	0	-1

Соответствующее представление неприводимо, так как

$$(\chi_4, \chi_4) = \frac{1}{12}(9 + 3) = 1,$$

и не изоморфно, очевидно, ни одному из предыдущих,

§ 10. Характеры группы вращений куба и группы симметрии тетраэдра

Число классов сопряженных элементов группы вращений куба O равно 5, порядок ее 24.

Из равенства

$$24 = 1^2 + 1^2 + 2^2 + 3^2 + 3^2$$

видим, что группа имеет два одномерных, одно двумерное и два трехмерных представления. Найдем их характеры.

Группа O изоморфна группе S_4 подстановок из четырех элементов. Подстановки бывают четные и нечетные. Произведение двух подстановок одинаковой четности — четно, произведение подстановок разной четности — нечетно. Поэтому для группы S_4 (как и для каждой симметрической группы) мы сразу получаем два одномерных представления — единичное и такое, которое всем четным подстановкам ставит в соответствие 1, а всем нечетным подстановкам — 1.

Каким же вращениям куба отвечают четные подстановки? Выше мы нашли, что группа O состоит из 5 классов сопряженных элементов, содержащих один, шесть, три, восемь и шесть элементов. Четные подстановки образуют в группе S_4 нормальную подгруппу 12-го порядка. Но нормальная подгруппа должна содержать целиком несколько классов сопряженных элементов. Кроме того, единичный элемент обязательно должен в нее войти. Следовательно, кроме единицы, в эту нормальную подгруппу войдут элементы, образующие классы из восьми и трех элементов.

Восемь элементов образуют повороты вокруг диагоналей на углы $2\pi/3$ и $4\pi/3$ — класс $\{8c_3\}$. Три элемента — это повороты вокруг осей четвертого порядка на угол π — класс $\{3c_4^2\}$. Остальные элементы группы: класс, состоящий из 6 поворотов вокруг осей четвертого порядка на углы $\pi/2$ и $3\pi/2$,— это $\{6c_4\}$, и класс, состоящий из 6 поворотов вокруг осей второго порядка — $\{6c_2\}$.

Итак, характеры одномерных представлений группы O:

O	e	$8c_3$	$3c_4^2$	$6c_4$	$6c_2$
χ_1	1	1	1	1	1
χ_2	1	1	1	-1	-1

Чтобы найти характер χ_4 одного из трехмерных представлений группы O (а она сама является одним из своих трехмерных представлений), снова будем каждый раз ось вращения принимать за ось Oz. Так мы получим матрицы

$$\Gamma(e)=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\qquad \Gamma(c_3)=\begin{bmatrix}-\frac{1}{2}&-\frac{1}{2}\sqrt{3}&0\\ \frac{1}{2}\sqrt{3}&-\frac{1}{2}&0\\0&0&1\end{bmatrix},$$

$$\Gamma(c_4)=\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix},\quad \Gamma(c_4^2)=\Gamma(c_2)=\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&1\end{bmatrix}.$$

Соответствующий характер

$$\chi_4(e)=3,\quad \chi_4(c_3)=0,\quad \chi_4(c_4^2)=-1,$$
$$\chi_4(c_4)=1,\quad \chi_4(c_2)=-1.$$

Далее, если мы уже имеем одно трехмерное представление группы O, то второе можно получить следующим образом: не изменяя преобразований соответствующих четным подстановкам, все остальные умножим на -1. Покажем, что при этом мы снова получим представление группы O. Пусть первое (трехмерное) представление обозначено через Γ, второе — через Γ', и пусть $a, b \in O$, тогда $\Gamma(ab)=\Gamma(a)\Gamma(b)$. Если оба элемента a и b отвечают четным подстановкам, то и элемент ab тоже отвечает четной подстановке; при этом $\Gamma'(a)=\Gamma(a)$, $\Gamma'(b)=\Gamma(b)$, $\Gamma'(ab)=\Gamma(ab)$, и значит, $\Gamma'(ab)=\Gamma'(a)\Gamma'(b)$.

Если элементы a и b отвечают нечетным подстановкам, то ab по-прежнему четно, и $\Gamma'(a) = -\Gamma(a)$, $\Gamma'(b) = -\Gamma(b)$; $\Gamma'(ab) = \Gamma(ab)$, откуда $\Gamma'(ab) = \Gamma'(a)\,\Gamma'(b)$. Наконец, если элементу a соответствует четная, а элементу b — нечетная подстановка, то элементу ab отвечает нечетная подстановка, и мы имеем $\Gamma'(a) = \Gamma(a)$, $\Gamma'(b) = -\Gamma(b)$, $\Gamma'(ab) = -\Gamma(ab)$, и значит, $\Gamma'(ab) = \Gamma'(a)\,\Gamma'(b)$.

При умножении преобразования на -1 его матрица в любом базисе,— а значит, и ее след — умножаются на -1. Поэтому второе трехмерное представление группы O будет иметь характер

$$\chi_5(e) = 3, \quad \chi_5(c_3) = 0, \quad \chi_5(c_4^2) = -1,$$
$$\chi_5(c_4) = -1, \quad \chi_5(c_2) = 1.$$

Оба трехмерных представления неприводимы, поскольку

$$\frac{1}{24}(9+3+6+6) = 1,$$

и не изоморфны между собой:

$$3\cdot 3 + 3 - 6 - 6 = 0.$$

Наконец, характер χ_3 двумерного представления можно найти алгебраически. Пусть

$$\chi_3(e) = 2, \; \chi_3(c_3) = \alpha, \; \chi_3(c_4^2) = \beta, \; \chi_3(c_4) = \gamma, \; \chi_3(c_2) = \delta.$$

Запишем условия ортогональности этого характера к четырем, уже найденным:

$$\begin{cases} 2 + 8\alpha + 3\beta + 6\gamma + 6\delta = 0, \\ 2 + 8\alpha + 3\beta - 6\gamma - 6\delta = 0, \\ 6 \qquad\quad - 3\beta + 6\gamma - 6\delta = 0, \\ 6 \qquad\quad - 3\beta - 6\gamma + 6\delta = 0. \end{cases}$$

Складывая и вычитая первые два уравнения, получим

$$\begin{cases} 2 + 8\alpha + 3\beta = 0, \\ 6\gamma \qquad\quad + 6\delta = 0. \end{cases}$$

А складывая и вычитая третье и четвертое, будем иметь

$$\begin{cases} 6 - 3\beta = 0, \\ 6\gamma - 6\delta = 0. \end{cases}$$

Следовательно,

$$\beta = 2, \gamma = \delta = 0, \alpha = -1,$$

и значит, характер χ_3 двумерного представления группы O таков:

$$\chi_3(e) = 2, \ \chi_3(c_3) = -1, \ \chi_3(c_4^2) = 2, \ \chi_3(c_4) = 0, \ \chi_3(c_2) = 0.$$

Легко видеть, что и это представление неприводимо:

$$1/24\ (4 + 8 + 3 \cdot 4) = 1.$$

Выпишем полную таблицу характеров группы вращений куба:

O	e	$8c_3$	$3c_4^2$	$6c_4$	$6c_2$
χ_1	1	1	1	1	1
χ_2	1	1	1	-1	-1
χ_3	2	-1	2	0	0
χ_4	3	0	-1	1	-1
χ_5	3	0	-1	-1	1

Группа симметрии тетраэдра T_d изоморфна группе O, и ее таблица характеров идентична таблице характеров группы O:

T_d	e	$8c_3$	$3c_2$	6σ	$6\sigma'$
χ_1	1	1	1	1	1
χ_2	1	1	1	−1	−1
χ_3	2	−1	2	0	0
χ_4	3	0	−1	1	−1
χ_5	3	0	−1	−1	1

§ 11. Тензорное (кронекеровское) произведение матриц

Пусть мы имеем две (квадратные) матрицы

$$A=\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix} \quad \text{и} \quad B=\begin{bmatrix}a & b & c\\ p & q & r\\ x & y & z\end{bmatrix}.$$

Их ***тензорным (кронекеровским) произведением*** называется матрица

$$A\times B=\begin{bmatrix}\alpha B & \beta B\\ \gamma B & \delta B\end{bmatrix}=\begin{bmatrix}\alpha a & \alpha b & \alpha c & \beta a & \beta b & \beta c\\ \alpha p & \alpha q & \alpha r & \beta p & \beta q & \beta r\\ \alpha x & \alpha y & \alpha z & \beta x & \beta y & \beta z\\ \gamma a & \gamma b & \gamma c & \delta a & \delta b & \delta c\\ \gamma p & \gamma q & \gamma r & \delta p & \delta q & \delta r\\ \gamma x & \gamma y & \gamma z & \delta x & \delta y & \delta z\end{bmatrix},$$

В общем случае определение аналогично:

$$A\times B=\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{bmatrix}\times\begin{bmatrix}b_{11} & b_{12} & \dots & b_{1m}\\ b_{21} & b_{22} & \dots & b_{2m}\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ b_{m1} & b_{m2} & \dots & b_{mm}\end{bmatrix}=$$

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

$$=\begin{bmatrix} a_{11}b_{11} & \dots a_{11}b_{1m} & a_{12}b_{11} & \dots a_{12}b_{1m} & \dots a_{1n}b_{11} & \dots a_{1n}b_{1m} \\ a_{11}b_{21} & \dots a_{11}b_{2m} & a_{12}b_{21} & \dots a_{12}b_{2m} & \dots a_{1n}b_{21} & \dots a_{1n}b_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{11}b_{m1} & \dots a_{11}b_{mm} & a_{12}b_{m1} & \dots a_{12}b_{mm} & \dots a_{1n}b_{m1} & \dots a_{1n}b_{mm} \\ a_{21}b_{11} & \dots a_{21}b_{1m} & a_{22}b_{11} & \dots a_{22}b_{1m} & \dots a_{2n}b_{11} & \dots a_{2n}b_{1m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1}b_{m1} & \dots a_{n1}b_{mm} & a_{n2}b_{m1} & \dots a_{n2}b_{mm} & \dots a_{nn}b_{m1} & \dots a_{nn}b_{mm} \end{bmatrix}.$$

Кронекеровское произведение матрицы порядка n и матрицы порядка m будет, очевидно, матрицей порядка mn. Легко проверяются следующие соотношения:

1) $A \times O_m = O_{mn}$,
2) $E_n \times E_m = E_{mn}$,
3) $A \times (B + C) = A \times B + A \times C$,
4) $(A + B) \times C = A \times C + B \times C$,
5) $(\alpha A) \times (\beta B) = (\alpha\beta)(A \times B)$,

где O_k — нулевая матрица, а E_k — единичная матрица порядка k.

Заметим, что *след тензорного произведения матриц* $\operatorname{tr}(A \times B) = a_{11} \operatorname{tr} B + a_{22} \operatorname{tr} B + \dots + a_{nn} \operatorname{tr} B = \operatorname{tr} A \cdot \operatorname{tr} B$ *равен произведению следов сомножителей.*

§ 12. Тензорное произведение векторных пространств

Здесь мы предполагаем, что читатель знаком с понятием тензора, введенным в главе VIII, и с соответствующими обозначениями. Этот и следующий параграфы можно и пропустить без ущерба для понимания дальнейшего; однако от этого несколько пострадает полнота проводимых дальше доказательств.

Если воспользоваться введенными в главе VIII обозначениями, то тензорное произведение матриц $A = [a^i_j]$ порядка n и $B = [b^p_q]$ порядка m, где верхний индекс — номер строки, а нижний — номер столбца, есть матрица $[a^i_j b^p_q]$:

$$\begin{bmatrix} a_1^1b_1^1 & \ldots & a_1^1b_m^1 & a_2^1b_1^1 & \ldots & a_2^1b_m^1 & \ldots & a_n^1b_1^1 & \ldots & a_n^1b_m^1 \\ a_1^1b_1^2 & \ldots & a_1^1b_m^2 & a_2^1b_1^2 & \ldots & a_2^1b_m^2 & \ldots & a_n^1b_1^2 & \ldots & a_n^1b_m^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^1b_1^m & \ldots & a_1^1b_m^m & a_2^1b_1^m & \ldots & a_2^1b_m^m & \ldots & a_n^1b_1^m & \ldots & a_n^1b_m^m \\ a_1^2b_1^1 & \ldots & a_1^2b_m^1 & a_2^2b_1^1 & \ldots & a_2^2b_m^1 & \ldots & a_n^2b_1^1 & \ldots & a_n^2b_m^1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^nb_1^m & \ldots & a_1^nb_m^m & a_2^nb_1^m & \ldots & a_2^nb_m^m & \ldots & a_n^nb_1^m & \ldots & a_n^nb_m^m \end{bmatrix},$$

элементы $a_j^ib_q^p$ которой (их можно обозначить короче: $c_{jq}^{ip} = a_j^ib_q^p$) занумерованы двумя парами индексов i, p и j, q, причем при лексикографическом упорядочении этих пар

$$11,\ 12,\ \ldots,\ 1m,\ 21,\ 22,\ \ldots,\ 2m,\ \ldots,\ n1,\ n2,\ \ldots,\ nm$$

верхняя пара определяет номер строки, а нижняя — номер столбца. Пусть матрицы $A = [a_j^i]$, $B = [b_q^p]$ — невырожденные, и пусть $A^{-1} = [\tilde{a}_j^i]$, $B^{-1} = [\tilde{b}_q^p]$. Покажем, что матрицей, обратной к тензорному произведению $A \times B$, будет матрица, равная тензорному произведению обратных матриц A^{-1} и B^{-1}, т. е. что $(A \times B)(A^{-1} \times B^{-1}) = E_{mn}$. Действительно, элемент (i, p)-й строки и (r, s)-го столбца произведения $(A \times B)(A^{-1} \times B^{-1})$ имеет вид

$$(a_j^ib_q^p)(\tilde{a}_r^j\tilde{b}_s^q) = (a_j^i\tilde{a}_r^j)(b_q^p\tilde{b}_s^q) = \delta_r^i\delta_s^p = \delta_{rs}^{ip}$$

(в тензорных обозначениях — по индексам $j = 1, 2, \ldots, n$ и $q = 1, 2, \ldots, m$ ведется суммирование), где δ_{rs}^{ip} — символ Кронекера (ср. § 1 главы VIII), равный 1, если пары i, p, и r, s тождественны, т. е. если $i = r$ и $p = s$; и равный 0, если $i \neq r$ или если $p \neq s$. Таким образом произведение $(A \times B)(A^{-1} \times B^{-1})$ есть единичная матрица порядка mn, и значит, матрицы $A \times B$ и $A^{-1} \times B^{-1}$ взаимно обратны (откуда, в частности, видно, что если $|A| \neq 0$ и $|B| \neq 0$, то и $|A \times B| \neq 0$).

Дадим теперь определение *тензорного произведения векторных пространств*. Пусть имеются два пространства: R_1 — размерности n и R_2 — размерности m. Выберем в пространствах R_1 и R_2 соответственно базисы $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ и рассмотрим линейное пространство R размерности mn с базисом $e_i f_k$, где $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$. (Заметьте, что мы никак, разумеется, не перемножаем векторы e_i и f_k, взятые из разных пространств — базис пространства R образован просто *парами* векторов, причем пара e_i, f_k обозначена через $e_i f_k$). Элементы этого базиса упорядочим лексикографически:

$$e_1 f_1,\ e_1 f_2,\ \ldots,\ e_1 f_m,\ e_2 f_1,\ e_2 f_2,\ \ldots,\ e_2 f_m,\ \ldots,\ e_n f_1,\ e_n f_2,\ \ldots,\ e_n f_m$$

и обозначим $e_i f_k$ через E_{ik}. (Ясно, что векторы E_{ik}, взятые в этом, лексикографическом, порядке, можно занумеровать и числами $1, 2, \ldots, mn$; однако нам будет удобнее нумеровать их парами чисел.)

Предположим теперь, что в пространстве R_1 мы перешли к новому базису $e'_1, e'_2, \ldots, e'_n$ с матрицей перехода $\mathrm{A} = [\alpha_i^k]$ (и значит, $e'_i = \alpha_i^k e_k$ — не забывайте о суммировании по индексу k!), а в пространстве R_2 — к новому базису $f'_1, f'_2, \ldots, f'_m$ с матрицей перехода $\mathrm{B} = [\beta_p^r]$ (и значит, $f'_p = \beta_p^r f_r$). По определению, будем считать, что при этом в пространстве R совершается переход к новому базису с матрицей перехода, равной тензорному произведению $\mathrm{A} \times \mathrm{B}$ матриц перехода в пространствах R_1 и R_2, т. е. что в пространстве R соответствующий новый базис будет образован векторами

$$E'_{ip} = e'_i f'_p = \alpha_i^k \beta_p^r E_{kr} = \alpha_i^k \beta_p^r e_k f_r.$$

(Заметьте, что правая часть формально является произведением сумм $\alpha_i^k e_k = e'_i$ и $\beta_p^r f_r = f'_p$). *Так определенное пространство R называется тензорным, или кронекеровским, произведением пространств R_1 и R_2 и обозначается $R_1 \otimes R_2$.*

Из доказанного выше вытекает, что если $\mathrm{A} \times \mathrm{B}$ — матрица перехода к новому базису в пространстве R, то обратной к ней будет матрица $\mathrm{A}^{-1} \times \mathrm{B}^{-1}$.

§ 13. Тензорное произведение линейных операторов

Пусть имеются два пространства: R_1 размерности n и R_2 размерности m, и пусть $R = R_1 \otimes R_2$ — их тензорное произведение. Предположим, что в пространстве R_1

задан линейный оператор $\mathcal{A}$ — двухвалентный смешанный тензор a_j^i, при переходе к новому базису преобразующийся по формуле $a_j'^i = \alpha_j^q \tilde{\alpha}_p^i a_q^p$, а в пространстве R_2 — линейный оператор $\mathcal{B}$ — такой же тензор b_s^r, преобразующийся по формуле $b_s'^r = \beta_s^l \tilde{\beta}_m^r b_l^m$. (Здесь $A = [\alpha_j^i]$ — матрица перехода к новому базису в пространстве R_1, $B = [\beta_j^i]$ — матрица перехода в пространстве R_2, $A^{-1} = [\tilde{\alpha}_j^i]$, $B^{-1} = [\tilde{\beta}_j^i]$).

Рассмотрим кронекеровское произведение $A' \times B'$ (где A' — матрица оператора $\mathcal{A}$ в новом базисе пространства R_1, а B' — матрица оператора $\mathcal{B}$ в новом базисе пространства R_2) и покажем, что элементы этой матрицы получаются из элементов матрицы $A \times B$ по правилу преобразования тензора один раз ко- и один раз контравариантного. Действительно, имеем

$$a_j'^i b_s'^r = (\alpha_j^q \tilde{\alpha}_p^i a_q^p)(\beta_s^l \tilde{\beta}_m^r b_l^m) = (\alpha_j^q \beta_s^l)(\tilde{\alpha}_p^i \tilde{\beta}_m^r)(a_q^p b_l^m).$$

(Здесь $\alpha_j^q \beta_s^l$ — элементы матрицы перехода к новому базису в пространстве R, $\tilde{\alpha}_p^i \tilde{\beta}_m^r$ — элементы обратной матрицы, а суммирование ведется по парам индексов q, l и p, m.)

Мы показали, что кронекеровское произведение матриц при переходе к новому базису преобразуется как смешанный двухвалентный тензор, а значит, оно определяет линейный оператор, действующий в пространстве R. *Этот оператор, обозначаемый через* $\mathcal{A} \times \mathcal{B}$, *называется тензорным* (или *кронекеровским*) *произведением операторов* $\mathcal{A}$ *и* $\mathcal{B}$. На базисные векторы пространства R он действует так:

$$(\mathcal{A} \times \mathcal{B}) E_{ij} = a_i^p b_j^q E_{pq} = a_i^p b_j^q e_p f_q.$$

(Заметьте, что правая часть формально является произведением сумм $a_i^p e_p = \mathcal{A} e_i$ и $b_j^q f_q = \mathcal{B} f_j$).

Итак, тензорное произведение линейных операторов не зависит от выбора базисов в пространствах R_1 и R_2. По доказанному в конце § 11 *след кронекеровского произведения линейных операторов равен произведению следов сомножителей.*

Докажем теперь тождество

$$AB \times CD = (A \times C)(B \times D), \qquad (3)$$

где A и B — любые матрицы порядка n, а C и D — матрицы порядка m. Пусть в каких-то фиксированных базисах пространств R_1 и R_2 элементы каждой из матриц обозначены теми же буквами латинского алфавита, что и сами матрицы, но не прописными, а строчными. Тогда (i, k)-й (т. е. стоящий на пересечении i-й строки и k-го столбца) элемент матрицы AB — это свертка $a^i_j b^j_k$, а (p, q)-й элемент матрицы CD равен $c^p_s d^s_q$. Следовательно, элементы тензорного произведения $AB \times CD$ — это

$$\left(a^i_j b^j_k\right)\left(c^p_s d^s_q\right) = \left(a^i_j c^p_s\right)\left(b^j_k d^s_q\right).$$

В правой части стоит свертка произведения матриц $A \times C$ и $B \times D$ по паре индексов j, s, т. е. элементы обычного матричного произведения $(A \times C)(B \times D)$. Формула (3) доказана.

§ 14. Тензорное произведение представлений (представления прямого произведения групп)

Наша ближайшая задача состоит в следующем. Предположим, что группа G равна прямому произведению своих подгрупп A и B: $G = A \times B$. Спрашивается, как, *знаяв все неприводимые представления групп A и B, найти все неприводимые представления их прямого произведения G*?

Пусть Γ^A — какое-нибудь представление группы A и Γ^B — представление группы B. Если $g = ab$, где $a \in A$, $b \in B$ — произвольный элемент группы G, то по определению, положим *)

$$\Gamma(g) = \Gamma^A(a) \times \Gamma^B(b).$$

Проверим, что так мы действительно получим п р е д с т а в л е н и е группы G, т. е. что если $g_1 = a_1 b_1$ и $g_2 = a_2 b_2$ — произвольные элементы группы G (где $a_1, a_2 \in A$, $b_1, b_2 \in B$), то $\Gamma(g_1 g_2) = \Gamma(g_1)\Gamma(g_2)$.

*) Читатель, пропустивший §§ 12—13, может считать, что $\Gamma^A(a)$ и $\Gamma^B(b)$ — это просто м а т р и ц ы соответствующих операторов.

Мы имеем $\Gamma^A(a_1a_2)=\Gamma^A(a_1)\Gamma^A(a_2)$, так как Γ^A — представление группы A, и $\Gamma^B(b_1b_2)=\Gamma^B(b_1)\Gamma^B(b_2)$, поскольку Γ^B — представление группы B. Пользуясь определением оператора Γ и тождеством (3) из предыдущего параграфа *), получаем

$$\Gamma(g_1g_2)=\Gamma(a_1b_1\cdot a_2b_2)=\Gamma(a_1a_2\cdot b_1b_2)=$$
$$=\Gamma^A(a_1a_2)\times\Gamma^B(b_1b_2)=[\Gamma^A(a_1)\Gamma^A(a_2)]\times$$
$$\times[\Gamma^B(b_1)\Gamma^B(b_2)]=[\Gamma^A(a_1)\times\Gamma^B(b_1)][\Gamma^A(a_2)\times\Gamma^B(b_2)]=$$
$$=\Gamma(a_1b_1)\Gamma(a_2b_2)=\Gamma(g_1)\Gamma(g_2),$$

т. е. Γ — действительно представление группы G.

Представление Γ группы G называется *тензорным произведением* *представлений Γ^A и Γ^B* и обозначается символом $\Gamma^A\times\Gamma^B$.

Пусть, далее, χ^A — характер представления Γ^A, χ^B — характер Γ^B и χ — характер их тензорного произведения $\Gamma^A\times\Gamma^B$. Тогда, если $g=ab$, где $a\in A$, $b\in B$, то

$$\chi(g)=\chi(ab)=\operatorname{tr}\Gamma(ab)=\operatorname{tr}[\Gamma^A(a)\times\Gamma^B(b)]=$$
$$=\operatorname{tr}\Gamma^A(a)\operatorname{tr}\Gamma^B(b),$$

т. е.

$$\chi(g)=\chi^A(a)\chi^B(b).$$

Это — важная формула, выражающая характер представления прямого произведения групп через характеры представлений сомножителей.

Пример. Мы знаем, что группа $V=A\times B$, где $A=\{e, a\}$ и $B=\{e, b\}$ — циклические группы второго порядка (см. стр. 297). Зная характеры представлений этих групп:

A	e	a
χ_1^A	1	1
χ_2^A	1	-1

B	e	b
χ_1^B	1	1
χ_2^B	1	-1

*) Читателю, пропустившему §§ 12—13, придется принять эту формулу без доказательства.

мы можем составить таблицу характеров их прямого произведения — группы V:

V	e	a	b	ab
$\chi_1^A\ \chi_1^B$	1	1	1	1
$\chi_1^A\ \chi_2^B$	1	1	-1	-1
$\chi_2^A\ \chi_1^B$	1	-1	1	-1
$\chi_2^A\ \chi_2^B$	1	-1	-1	1

(разумеется, совпадающую с найденной на стр. 327).

Т е о р е м а 6. *Пусть группа G равна прямому произведению $A\times B$. Если представления Γ^A группы A и Γ^B группы B неприводимы, то и их тензорное произведение $\Gamma^A\times\Gamma^B$ неприводимо* *).

Д о к а з а т е л ь с т в о. Пусть χ^A — характер представления Γ^A, χ^B — характер представления Γ^B, k — порядок группы A и l — порядок группы B. Пусть, далее, $a_1, a_2, \ldots, a_k$ — все элементы группы A и $b_1, b_2, \ldots, b_l$ — все элементы группы B. Тогда элементами группы $A\times B = G$ будут всевозможные произведения a_ib_j, где $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, l$.

Пусть χ — характер тензорного произведения $\Gamma^A\times\Gamma^B$. Вычислим его скалярный квадрат:

$$(\chi, \chi) = \frac{1}{kl}\sum_{i,j}\chi(a_ib_j)\overline{\chi(a_ib_j)} = \frac{1}{kl}\sum_{i,j}\chi^A(a_i)\,\chi^B(b_j)\,\overline{\chi^A(a_i)\,\chi^B(b_j)} =$$

$$= \frac{1}{k}\sum_{i=1}^{k}\chi^A(a_i)\,\overline{\chi^A(a_i)}\cdot\frac{1}{l}\sum_{j=1}^{l}\chi^B(b_j)\,\overline{\chi^B(b_j)} = (\chi^A, \chi^A)\,(\chi^B, \chi^B).$$

*) Пусть читатель не удивляется, встретив в литературе и прямо противоположное утверждение. Дело в том, что, помимо определенного здесь тензорного произведения представлений двух р а з н ы х групп существует (несколько похожее на это) определение тензорного произведения двух представлений одной и той же группы (оно тоже будет некоторым представлением этой группы). Для этого, другого, тензорного произведения теорема, аналогичная теореме 6, неверна.

Но $(\chi^A, \chi^A) = 1$ и $(\chi^B, \chi^B) = 1$, так как оба представления Γ^A и Γ^B по условию, неприводимы. Следовательно, $(\chi, \chi) = 1$, и значит, представление $\Gamma^A \times \Gamma^B$ неприводимо.

Теорема 7. *Каждое неприводимое представление группы $G = A \times B$ изоморфно тензорному произведению некоторого неприводимого представления группы A и некоторого неприводимого представления группы B.*

Доказательство. Предположим, что число классов сопряженных элементов группы A равно p, а число классов группы B — равно q, и пусть $\Gamma_1^A, \Gamma_2^A, \ldots, \Gamma_p^A$ — все неприводимые неизоморфные между собой представления группы A, а $\Gamma_1^B, \Gamma_2^B, \ldots, \Gamma_q^B$ — все неприводимые попарно неизоморфные представления группы B. Тогда всевозможные тензорные произведения $\Gamma_r^A \times \Gamma_s^B = \Gamma_{rs}$ являются, по теореме 6, неприводимыми представлениями группы G. Число таких представлений равно pq. Число классов сопряженных элементов группы G тоже равно pq (следствие из теоремы 10 § 10 главы X). Значит, если мы докажем, что произведения Γ_{rs}, не изоморфны между собой, то это и будут все неприводимые представления группы G, чем наше утверждение и будет доказано.

Вычислим скалярное произведение характеров двух таких представлений Γ_{rs} и Γ_{ht}:

$$(\chi_{rs}, \chi_{ht}) = \frac{1}{kl} \sum_{i,j} \chi_{rs}(a_i b_j) \overline{\chi_{ht}(a_i b_j)} =$$

$$= \frac{1}{kl} \sum_{i,j} \chi_r^A(a_i)\, \chi_s^B(b_j)\, \overline{\chi_h^A(a_i)\, \chi_t^B(b_j)} =$$

$$= \frac{1}{k} \sum_{i=1}^{k} \chi_r^A(a_i)\, \overline{\chi_h^A(a_i)}\, \frac{1}{l} \sum_{j=1}^{l} \chi_s^B(b_j)\, \overline{\chi_t^B(b_j)} =$$

$$= (\chi_r^A, \chi_h^A)(\chi_s^B, \chi_t^B).$$

Так как либо Γ_r^A не изоморфно Γ_h^A, либо Γ_s^B не изоморфно Γ_t^B (а, возможно, что имеет место и то, и другое), то по крайней мере одно из скалярных произведений в правой части равно 0. Следовательно, $(\chi_{rs}, \chi_{ht}) = 0$,— и представления Γ_{rs} и Γ_{ht} группы G не изоморфны между собой.

§ 15. Характеры группы симметрии куба

Эта группа является прямым произведением группы O и циклической группы J второго порядка (см. § 8 главы XI). Она имеет 10 классов сопряженных элементов. В соответствии с разложением

$$48 = 24 \cdot 2 = (1^2 + 1^2 + 2^2 + 3^2 + 3^2)2,$$

группа O_h имеет 4 одномерных, 2 двумерных и 4 трехмерных представления. Характеры их определяются следующей таблицей:

	e	$8c_3$	$3c_4^2$	$6c_4$	$6c_2$	j	$8jc_3$	$3jc_4^2$	$6jc_4$	$6jc_2$
χ_1	1	1	1	1	1	1	1	1	1	1
χ_2	1	1	1	−1	−1	1	1	1	−1	−1
χ_3	2	−1	2	0	0	2	−1	2	0	0
χ_4	3	0	−1	1	−1	3	0	−1	1	−1
χ_5	3	0	−1	−1	1	3	0	−1	−1	1
χ_6	1	1	1	1	1	−1	−1	−1	−1	−1
χ_7	1	1	1	−1	−1	−1	−1	−1	1	1
χ_8	2	−1	2	0	0	−2	1	−2	0	0
χ_9	3	0	−1	1	−1	−3	0	1	−1	1
χ_{10}	3	0	−1	−1	1	−3	0	1	1	−1

Задача. Выпишите таблицы характеров прямых произведений $C_2 \times C_3$ (эта группа изоморфна C_6), $C_2 \times C_4$, $C_2 \times C_5$, $C_2 \times C_6$, $V \times C_2$, $D_3 \times C_2$, $D_4 \times C_2$, $D_5 \times C_2$, $D_6 \times C_2$, $T_h = T \times C_2$.

СПИСОК ДОПОЛНИТЕЛЬНОЙ ЛИТЕРАТУРЫ

К главе I

1. *Курош А. Г.* Курс высшей алгебры.— М.: Наука, 1975.

К главам II—VII

2. *Беллман Р.* Введение в теорию матриц.— М.: Наука, 1969.
3. *Беклемишев Д. В.* Курс аналитической геометрии и линейной алгебры.—М.: Наука, 1974.
4. *Гантмахер Ф. Р.* Теория матриц.— М.: Наука, 1966.
5. *Гельфанд И. М.* Лекции по линейной алгебре.— М.: Наука, 1971.
6. *Дьедонне Ж.* Линейная алгебра и элементарная геометрия.— М.: Наука, 1972.
7. *Ефимов Н. В.* Квадратичные формы и матрицы.— М.: Наука, 1975.
8. *Ефимов Н. В., Розендорн Э. Р.* Линейная алгебра и многомерная геометрия.—М.: Наука, 1974.
9. *Карпелевич Ф. И., Садовский Л. Е.* Элементы линейной алгебры и линейного программирования.—М.: Наука, 1967.
10. *Мальцев А. И.* Основы линейной алгебры.—М.: Наука, 1970.
11. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры.—Физматгиз, 1963.
12. *Фрезер Р., Дункан В., Коллар А.* Теория матриц и ее приложения.—М.: ИЛ, 1950.
13. *Халмош П.* Конечномерные векторные пространства.—М.: Физматгиз, 1963.
14. *Шилов Г. Е.* Математический анализ. Конечномерные линейные пространства.—М.: Наука, 1969.

К главе VIII

15. *Акивис М. А., Гольдберг В. В.* Тензорное исчисление.—М.: Наука, 1972.
16. *Кильчевский Н. А.* Элементы тензорного исчисления и его приложения к механике.—М.: Гостехиздат, 1954.

К главе IX

17. *Борн М.* Эйнштейновская теория относительности.—М.: Мир, 1972.
18. *Ландау Л. Д., Румер Ю. Б.* Что такое теория относительности.— М.: Сов. Россия, 1963.
19. *Рашевский П. К.* Риманова геометрия и тензорный анализ.— М.: Наука, 1967.
20. *Шварц Дж.* Как это произошло.—М.: Мир, 1965.

21. *Эйнштейн А.* О специальной и общей теории относительности (общедоступное изложение). Физика и реальность.—М.: Наука, 1963, с. 167—235. Собрание сочинений, т. I.—М.: Наука, 1965, с. 530—600.
22. *Яглом И. М.* Принцип относительности Галилея и неевклидова геометрия.—М.: Наука, 1969.

К главам X—XI

23. *Александров П. С.* Введение в теорию групп.—М.: Учпедгиз, 1951.
24. *Вейль Г.* Симметрия.—М.: Наука, 1968.
25. *Курош А. Г.* Теория групп.—М.: Наука, 1967.
26. *Холл М.* Теория групп.—М.: ИЛ, 1962.

К главам XII—XIII

27. *Вигнер Е.* Теория групп и ее приложения к квантовомеханической теории атомных спектров.—М.: ИЛ, 1961.
28. *Любарский Г. Я.* Теория групп и ее применения в физике.—М.: Физматгиз, 1958.
29. *Мурнаган Ф.* Теория представлений групп.—М.: ИЛ, 1950.
30. *Серр Ж. П.* Линейные представления конечных групп.—М.: Мир, 1970.
31. *Хамермеш М.* Теория групп и ее применение к физическим проблемам.—М.: Мир, 1966.
32. *Хейне В.* Теория групп в квантовой механике.—М.: ИЛ, 1963.
33. *Шмидт О. Ю.* Абстрактная теория групп (сб. «Избранные труды»).—М.: Изд-во АН СССР, 1959.

Задачник

34. *Проскуряков И. В.* Сборник задач по линейной алгебре.—М.: Наука, 1974.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ