В. И. СМИРНОВ

КУРС ВЫСШЕЙ МАТЕМАТИКИ

ТОМ ТРЕТИЙ

ЧАСТЬ ПЕРВАЯ

ИЗДАНИЕ ДЕСЯТОЕ,
СТЕРЕОТИПНОЕ

*Допущено Министерством
высшего и среднего специального образования СССР
в качестве учебного пособия для студентов механико-математических
и физико-математических факультетов университетов*

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1974

517
С 50
УДК 510 (022)

$С \frac{20203-073}{053(02)-74}$ 19-74

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ К ЧЕТВЕРТОМУ ИЗДАНИЮ

В настоящем издании, в связи с добавлением нового материала, третий том разбит на две части. Первая часть содержит весь материал, относящийся к линейной алгебре, теории квадратичных форм и теории групп. В этой части наиболее существенные добавления относятся к теории групп. Большую помощь при составлении этих добавлений мне оказал Д. К. Фаддеев. Ему, в частности, принадлежит изложение материала, относящегося к выяснению простоты группы вращения и группы Лоренца, построение группы по структурным постоянным и интегрированию на группе [**70, 81, 87, 88, 89, 90**]. Приношу ему большую благодарность за помощь в работе над этой книгой.

В. Смирнов

ПРЕДИСЛОВИЕ К ДЕВЯТОМУ ИЗДАНИЮ

Существенные изменения в настоящем издании внесены лишь в материал, относящийся к функциональному пространству L_2. Это связано с тем, что последнее, двадцатое, издание второго тома содержит изложение теории интеграла Лебега и свойств семейства функций, интегрируемых с квадратом.

В. Смирнов

ГЛАВА I

ОПРЕДЕЛИТЕЛИ И РЕШЕНИЕ СИСТЕМ УРАВНЕНИЙ

§ 1. ОПРЕДЕЛИТЕЛЬ И ЕГО СВОЙСТВА

1. Понятие об определителе. Мы начнем настоящий параграф с решения простой алгебраической задачи, а именно задачи о решении систем уравнений первой степени. Рассмотрение этой задачи приведет нас к важному понятию об определителе.

Начнем с рассмотрения наиболее простых частных случаев. Возьмем сначала систему двух уравнений с двумя неизвестными:

$$a_{11}x_1 + a_{12}x_2 = b_1,$$
$$a_{21}x_1 + a_{22}x_2 = b_2.$$

Коэффициенты при неизвестных a_{ik} снабжены двумя значками, первый из которых указывает, в каком уравнении находится этот коэффициент, а второй значок указывает, при каком из неизвестных он стоит.

Решение написанной системы, как известно, имеет вид:

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}; \quad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Возьмем теперь три уравнения с тремя неизвестными:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1,$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2,$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3,$$

причем мы пользуемся прежними обозначениями для коэффициентов. Перепишем первые два уравнения в виде:

$$a_{11}x_1 + a_{12}x_2 = b_1 - a_{13}x_3,$$
$$a_{21}x_1 + a_{22}x_2 = b_2 - a_{23}x_3.$$

Решая их относительно неизвестных x_1 и x_2 по предыдущим формулам, будем иметь:

$$x_1 = \frac{(b_1 - a_{13}x_3)\,a_{22} - a_{12}\,(b_2 - a_{23}x_3)}{a_{11}a_{22} - a_{12}a_{21}};$$
$$x_2 = \frac{a_{11}\,(b_2 - a_{23}x_3) - (b_1 - a_{13}x_3)\,a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Подставляя эти выражения в последнее уравнение системы, получим уравнение для определения неизвестного x_3 и, наконец, решая это уравнение, будем иметь окончательное выражение для этого неизвестного:

$$x_3 = \frac{a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}}. \quad (1)$$

Рассмотрим подробно конструкцию этого выражения. Заметим прежде всего, что его числитель может быть получен из знаменателя простой заменой коэффициентов a_{i3} при определяемом неизвестном свободными членами b_i. Таким образом, остается выяснить закон образования знаменателя, который не содержит уже свободных членов и составлен исключительно из коэффициентов нашей системы. Запишем эти коэффициенты в виде квадратной таблицы, сохраняя тот порядок, в котором они стоят в самой системе

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}. \quad (2)$$

Написанная таблица содержит три строки и три столбца. Числа a_{ik} называются ее элементами. Первый из значков показывает, в какой строке стоит этот элемент, а второй значок указывает номер столбца. Выпишем теперь знаменатель выражения (1):

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \quad (3)$$

Как мы видим, он состоит из шести членов, и каждый его член есть произведение трех элементов таблицы (2), причем в этом произведении участвуют элементы каждой строки и каждого столбца. Действительно, эти произведения имеют вид:

$$a_{1p}a_{2q}a_{3r}, \quad (4)$$

где p, q, r суть целые числа 1, 2, 3, расставленные в некотором определенном порядке. Таким образом, как первые, так и вторые знаки представляют собою совокупность целых чисел 1, 2, 3, и произведения (4) действительно содержат по одному элементу из каждой строки и из каждого столбца. Чтобы получить все члены выражения (3), надо в произведении (4) взять вторые значки p, q, r в различных возможных порядках. Таких возможных перестановок из вторых значков будет, очевидно, шесть:

$$1,\ 2,\ 3;\quad 2,\ 3,\ 1;\quad 3,\ 1,\ 2;\quad 1,\ 3,\ 2;\quad 2,\ 1,\ 3;\quad 3,\ 2,\ 1, \tag{5}$$

и мы получаем, таким образом, все шесть членов выражения (3). Но мы видим, что некоторые из произведений (4) входят в выражение (3) со знаком плюс, а другие со знаком минус, и остается лишь выяснить то правило, согласно которому надо выбирать знак. Со знаком плюс, как мы видим, входят те произведения (4), у которых вторые значки образуют следующие перестановки:

$$1,\ 2,\ 3;\quad 2,\ 3,\ 1;\quad 3,\ 1,\ 2, \tag{5_1}$$

а со знаком минус входят те произведения, вторые значки которых образуют перестановки:

$$1,\ 3,\ 2;\quad 2,\ 1,\ 3;\quad 3,\ 2,\ 1. \tag{5_2}$$

Выясним теперь, чем перестановки (5_1) отличаются от перестановок (5_2). Назовем беспорядком в перестановке тот факт, что большее число стоит впереди меньшего, и подсчитаем число беспорядков в перестановках (5_1). В первой из этих перестановок беспорядков вовсе нет, т. е. число беспорядков равно нулю. Перейдем ко второй перестановке и сравним по величине каждое из чисел, в нее входящих, со всеми следующими. Мы видим, что здесь имеются два беспорядка, а именно число 2 стоит перед числом 1 и число 3 стоит перед числом 1. Точно так же нетрудно убедиться, что и третья из перестановок (5_1) содержит два беспорядка. Одним словом, можно сказать, что все перестановки (5_1) содержат четное число беспорядков. Совершенно так же исследуя перестановки (5_2), мы убеждаемся, что все они содержат нечетное число беспорядков. Мы можем теперь формулировать правило знаков в выражении (3), а именно: те произведения (4), в которых число беспорядков в перестановке, образованной вторыми значками, есть число четное, входят в выражение (3) без всякого изменения. Те же произведения (4), у которых перестановки, образованные вторыми значками, содержат нечетное число беспорядков, входят в выражение (3) с приписанным к ним знаком минус. Выражение (3) называется *определителем третьего порядка*, соответствующим таблице чисел (2). Нетрудно теперь обобщить предыдущее на случай определителей любого порядка.

Пусть имеется n^2 чисел, расставленных в виде квадратной таблицы, имеющей n строк и n столбцов:

$$\left\| \begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{matrix} \right\| . \tag{6}$$

Элементы этой таблицы a_{ik} суть заданные комплексные числа, причем значки i и k указывают номер строки и столбца, на пересечении которых стоит число a_{ik}. Составим всевозможные произведения из чисел таблицы (6) так, чтобы эти произведения содержали по одному числу из каждой строки и из каждого столбца. Эти произведения будут иметь вид:

$$a_{1p_1} a_{2p_2} \ldots a_{np_n}, \tag{7}$$

где $p_1, p_2, \ldots, p_n$ суть числа $1, 2, \ldots, n$, расставленные в некотором порядке. Чтобы получить всевозможные произведения вида (7), нам надо взять всевозможные перестановки вторых значков. Как известно из элементарной алгебры, число таких перестановок будет равно факториалу целого числа n:

$$1 \cdot 2 \cdot 3 \ldots n = n!$$

Каждая из этих перестановок будет иметь некоторое число беспорядков по сравнению с основной перестановкой

$$1, 2, 3, \ldots, n.$$

Те произведения вида (7), вторые значки которых образуют перестановку с четным числом беспорядков, возьмем без всякого изменения, а к тем произведениям вида (7), у которых перестановка вторых значков имеет нечетное число беспорядков, припишем знак минус. Сумма всех полученных таким образом произведений и называется *определителем n-го порядка*, соответствующим таблице (6). Эта сумма будет, очевидно, содержать $n!$ слагаемых. Нетрудно представить это определение в виде формулы. Введем для этого некоторое обозначение. Пусть $p_1, p_2, \ldots, p_n$ — некоторая перестановка из чисел $1, 2, \ldots, n$. Обозначим число беспорядков в этой перестановке символом

$$[p_1, p_2, \ldots, p_n].$$

Тогда данное выше определение определителя, соответствующего таблице (6), может быть записано в виде следующей формулы, причем для обозначения определителя мы пишем таблицу (6) между вертикальными чертами

$$\begin{vmatrix} a_{11} & a_{12} \ldots & a_{1n} \\ a_{21} & a_{22} \ldots & a_{2n} \\ \cdot & \cdot \quad \cdot & \cdot \\ a_{n1} & a_{n2} \ldots & a_{nn} \end{vmatrix} = \sum_{(p_1, p_2, \ldots, p_n)} (-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1} a_{2p_2} \ldots a_{np_n}. \tag{8}$$

Здесь суммирование распространяется на всевозможные перестановки вторых значков, т. е. на всевозможные перестановки $(p_1, p_2, \ldots, p_n)$. Говоря о таблице, как таковой, а не об определителе, из нее составленном, мы ставим эту таблицу между двойными вертикальными чертами.

Заметим, что в выражении (3) мы в каждом произведении расставили сомножители в таком порядке, чтобы первые значки образовывали основную перестановку 1, 2, 3, и таким образом все наши рассуждения относились к перестановкам, образуемым вторыми значками. Можно, наоборот, поставить в каждом произведении сомножители так, чтобы вторые значки всегда шли в возрастающем порядке; при этом выражение (3) перепишется в виде:

$$a_{11}a_{22}a_{33}+a_{31}a_{12}a_{23}+a_{21}a_{32}a_{13}-a_{11}a_{32}a_{23}-a_{21}a_{12}a_{33}-a_{31}a_{22}a_{13}. \tag{9}$$

Здесь первые значки образуют всевозможные перестановки p, q, r, причем легко проверить, что правило знаков у членов выражения (9) может быть формулировано совершенно так же, как и выше, но только по отношению к первым значкам. Это приводит нас к тому, чтобы наряду с суммой (8) рассматривать аналогичную сумму вида:

$$\sum_{(p_1, p_2, \ldots, p_n)} (-1)^{[p_1, p_2, \ldots, p_n]} a_{p_1 1} a_{p_2 2} \ldots a_{p_n n}. \tag{10}$$

Очевидно, что эта последняя сумма состоит из тех же членов, что и сумма (8). В дальнейшем мы увидим, что и знаки ее членов такие же, как и в сумме (8), т. е. так же, как и при $n=3$, сумма (10) совпадает с суммой (8).

Обратимся, наконец, к случаю $n=2$. При этом таблица имеет вид:

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

и формула (8) дает следующее выражение для определителя второго порядка, соответствующего этой таблице:

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

Из предыдущего непосредственно следует, что для выяснения свойств определителя необходимо познакомиться ближе со свойствами перестановок, к чему мы сейчас и переходим.

2. Перестановки. Пусть имеется n каких-нибудь элементов, расставленных в определенном порядке. Назовем это *перестановкой*, образованной из наших элементов. Докажем прежде всего, что таких различных перестановок будет $n!$ При $n=2$ это очевидно, ибо два элемента могут образовывать две различные перестановки. При $n=3$ это непосредственно следует из подсчета перестановок (5), где роль элементов играют числа 1, 2, 3. Без труда можно убедиться, что (5) дают всевозможные перестановки из этих трех элементов. Докажем наше утверждение при любом целом n по закону индукции. Предполагая, что наше утверждение справедливо при некотором n, докажем, что оно справедливо и для числа элементов $(n+1)$. Положим, что n элементов дают $n!$ перестановок и рассмотрим какие-нибудь $(n+1)$ элементов, которые обозначим

$$C_1, C_2, \ldots, C_{n+1}.$$

Обратим сначала внимание на те перестановки, у которых первым элементом будет C_1. Чтобы получить всевозможные такие перестановки, надо поставить на первое место C_1, а затем писать всевозможные перестановки из оставшихся n элементов, и число таких перестановок будет по предположению равно $n!$ Таким образом, число перестановок из элементов C_k, начинающихся с элемента C_1, будет равно $n!$ Совершенно так же число перестановок из элементов C_k, начинающихся с элемента C_2, будет равно также $n!$ В общем, число различных перестановок из элементов C_k будет

$$n! \cdot (n+1) = 1 \cdot 2 \cdot \ldots n \cdot (n+1) = (n+1)!,$$

что и требовалось доказать.

Мы можем, конечно, считать, что за элементы взяты целые числа, начиная с единицы, чего мы и будем придерживаться в дальнейшем. Назовем *транспозицией операцию, которая состоит в том, что в некоторой перестановке мы меняем местами два элемента.* Непосредственно очевидно, что из всякой перестановки мы можем получить всякую другую перестановку, совершая несколько транспозиций. Например, возьмем две перестановки из четырех элементов

$$1,\ 3,\ 4,\ 2, \quad 2,\ 4,\ 1,\ 3.$$

Можно перейти от первой из этих перестановок ко второй при помощи транспозиций путем следующего перехода:

$$1,\ 3,\ 4,\ 2, \rightarrow 2,\ 3,\ 4,\ 1 \rightarrow 2,\ 4,\ 3,\ 1 \rightarrow 2,\ 4,\ 1,\ 3.$$

Здесь нам понадобились три транспозиции, чтобы перейти от первой перестановки ко второй. Если бы мы совершали транспозиции иным образом, то мы могли бы и другим путем перейти от первой перестановки ко второй при помощи транспозиций, т. е., иначе говоря, число транспозиций, необходимых для перехода от одной перестановки к другой, не есть строго определенное число. Можно переходить от одной перестановки к другой при помощи различного числа транспозиций. Но для нас будет существенным доказать, что эти различные числа для двух заданных перестановок будут всегда или все четные или все нечетные. Иначе это выражают, говоря, что эти числа всегда одинаковой четности. Чтобы выяснить это, введем понятие о беспорядке, которым мы уже пользовались в предыдущем номере. Пусть имеются перестановки из n

элементов 1, 2, ..., n. Назовем *основной перестановкой* ту перестановку

$$1,\ 2,\ \ldots,\ n, \tag{12}$$

в которой числа идут возрастающим порядком. Назовем беспорядком в некоторой перестановке тот факт, когда два элемента этой перестановки следуют не в том порядке, в каком они стоят в основной перестановке (12), т. е., иначе говоря, когда большее число стоит левее меньшего. Назовем *перестановками первого класса* те перестановки, в которых число беспорядков есть число четное, и *перестановками второго класса*, — те, где это число есть число нечетное. Основным для дальнейшего будет следующая теорема:

Транспозиция меняет число беспорядков на число нечетное.

Возьмем некоторую перестановку

$$a,\ b,\ \ldots,\ k,\ \ldots,\ p,\ \ldots,\ s \tag{13}$$

и положим, что мы применяем к этой перестановке транспозиции по отношению к элементам k и p, т. е. меняем эти два элемента взаимно местами. После такой транспозиции взаимное расположение элементов k и p относительно элементов, стоящих левее k или правее p, останется прежним. Изменится лишь взаимное расположение элементов k и p относительно тех элементов, которые стоят в перестановке между k и p, а также изменится, конечно, взаимное расположение элементов k и p одного относительно другого. Подсчитаем общее изменение числа беспорядков. Положим, что между элементами k и p в перестановке (13) стоит всего m элементов, и пусть эти средние элементы по сравнению с элементом k дают α порядков и β беспорядков, относительно же элемента p дают α_1 порядков и β_1 беспорядков. Имеем, очевидно:

$$\alpha+\beta=\alpha_1+\beta_1=m. \tag{14}$$

В результате транспозиции порядок перейдет в беспорядок и наоборот, т. е., точнее говоря, если элемент k с некоторым средним элементом до транспозиции был в порядке, то после транспозиции он окажется в беспорядке и наоборот, и то же самое для элемента p. Таким образом общее число беспорядков у элементов k и p относительно средних элементов до транспозиции было $\beta+\beta_1$ и после транспозиции $\alpha+\alpha_1$, т. е. изменение числа беспорядков будет

$$\gamma=(\alpha+\alpha_1)-(\beta+\beta_1).$$

Пользуясь (14), можем переписать это число в виде:

$$\gamma=(\alpha+\alpha_1)-(m-\alpha+m-\alpha_1)=2(\alpha+\alpha_1-m),$$

откуда непосредственно следует, что это число γ будет четным. Остается обратить внимание на взаимное расположение самих элементов k и p. Если до транспозиции они образовывали порядок, то после они образуют беспорядок и наоборот, т. е. здесь изменение числа беспорядков равно единице, и таким образом общее изменение числа беспорядков, происшедших от транспозиции, будет числом нечетным.

Выясним некоторые следствия, вытекающие из доказанной теоремы.

Следствие I. Если выписать все $n!$ перестановок и в каждой из них произвести транспозицию двух определенных элементов, например элементов 1 и 3, то все перестановки первого класса станут перестановками второго класса и наоборот; а в общем получится опять вся совокупность $n!$ перестановок. Отсюда непосредственно следует, что *число перестановок первого и второго класса одинаково.*

Следствие II. Всякая перестановка может быть получена из основной путем транспозиций. Из доказанной теоремы непосредственно следует, что *первый класс образуют те перестановки, которые получаются из основной путем четного числа транспозиций, а второй класс — те перестановки, которые получаются из основной путем нечетного числа транспозиций.*

Следствие III. Выбор основной перестановки совершенно произволен. Мы могли бы вместо перестановки (12) выбрать за основную какую-нибудь другую перестановку, при этом, конечно, при определении беспорядков надо сравнивать перестановку с основной, т. е. исходить из того порядка элементов, в котором они стоят в основной перестановке. Нетрудно видеть, что если мы возьмем вместо перестановки (12) за основную перестановку какую-нибудь из перестановок первого класса, то перестановки первого класса останутся по-прежнему перестановками первого класса, а перестановки второго класса останутся перестановками второго класса. Наоборот, если мы за основную перестановку возьмем какую-нибудь перестановку второго класса, то перестановки второго класса станут перестановками первого класса, и перестановки первого класса станут перестановками второго класса.

Например, если в шести перестановках из элементов 1, 2, 3 мы примем за основную перестановку 2, 1, 3, то перестановками первого класса будут перестановки:

$$2,\ 1,\ 3; \quad 1,\ 3,\ 2; \quad 3,\ 2,\ 1.$$

Во второй из этих перестановок мы имеем два беспорядка: 1 стоит перед 2 и 3 перед 2, а в основной 2 стоит перед 1 и 2 перед 3. Перестановками второго класса будут перестановки:

$$1,\ 2,\ 3; \quad 2,\ 3,\ 1; \quad 3,\ 1,\ 2.$$

В первой из этих перестановок мы имеем один беспорядок по сравнению с основной 2, 1, 3, а именно 1 стоит перед 2.

Принимая во внимание сказанное выше, мы можем формулировать правило знаков в выражении (8) следующим образом: *мы пи-*

шем перед произведением знак плюс, если перестановка из вторых значков принадлежит первому классу, и знак минус, если перестановка из вторых значков принадлежит второму классу, причем за основную перестановку принята перестановка $1, 2, \ldots, n$.

Выясним теперь одно из основных свойств определителя. Переставим в таблице, дающей определитель, первый и второй столбцы. Числа, которые мы раньше обозначали через a_{ik}, мы по-прежнему будем обозначать этими же буквами с теми же самыми значками. Вместо таблицы (6) будем иметь таблицу:

$$\begin{Vmatrix} a_{12} & a_{11} & a_{13} & \ldots & a_{1n} \\ a_{22} & a_{21} & a_{23} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n2} & a_{n1} & a_{n3} & \ldots & a_{nn} \end{Vmatrix}. \tag{15}$$

Мы можем теперь, пользуясь определением, выражаемым формулой (8), составить определитель, соответствующий таблице (15). В этой таблице столбцы пронумерованы в следующем порядке: $2, 1, 3, \ldots, n$, и эту перестановку мы должны считать за основную. Она получилась из прежней основной при помощи одной транспозиции и, следовательно, раньше принадлежала ко второму классу. Таким образом прежние перестановки второго класса станут при новом выборе основной перестановки перестановками первого класса и наоборот. Следовательно, определитель, соответствующий таблице (15), будет суммой тех же слагаемых, которые стоят в формуле (8), но вследствие указанного только что изменения в распределении перестановок вторых значков по классам знаки у членов новой суммы будут противоположны знакам слагаемых суммы (8), т. е. *при перестановке двух столбцов величина определителя меняет знак*. Мы доказали это свойство, переставляя первый и второй столбцы. Точно такое же доказательство годится и при перестановке двух любых столбцов. Так, например, имеет место формула:

$$\begin{vmatrix} 1 & 0 & 3 \\ 2 & 7 & 6 \\ 5 & 3 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 7 \\ 5 & 0 & 3 \end{vmatrix}.$$

Второй определитель получается из первого перестановкой второго и третьего столбцов.

Выясним еще одно свойство определителя. Возьмем некоторое слагаемое суммы (8):

$$(-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1} a_{2p_2} \ldots a_{np_n}. \tag{16$_1$}$$

Переставляя порядок сомножителей, мы можем привести в полный порядок вторые значки, но при этом первые значки будут образовывать некоторую перестановку $q_1, q_2, \ldots, q_n$, и предыдущее выражение запишется в виде:

$$(-1)^{[p_1, p_2, \ldots, p_n]} a_{q_1}{}^1 a_{q_2}{}^2 \ldots a_{q_n}{}^n. \tag{16$_2$}$$

Переход от (16_1) к (16_2) можно совершить при помощи нескольких транспозиций сомножителей. Всякая такая транспозиция будет одновременно транспозицией в перестановке как первых, так и вторых значков. Если число необходимых транспозиций для перехода от (16_1) к (16_2) будет четным, то отсюда будет следовать, что перестановка $p_1, p_2, \ldots, p_n$ принадлежит первому классу, так как она переходит в основную $1, 2, \ldots, n$ при помощи четного числа транспозиций и, следовательно, очевидно, может быть получена из основной также при помощи четного числа транспозиций. Но при этом и перестановка $q_1, q_2, \ldots, q_n$ принадлежит к первому классу, так как она одновременно получается из основной при помощи того же четного числа транспозиций. Точно так же, если $p_1, p_2, \ldots, p_n$ принадлежит второму классу, то и $q_1, q_2, \ldots, q_n$ принадлежит второму классу. Отсюда следует, что $(-1)^{[p_1, p_2, \ldots, p_n]} = (-1)^{[q_1, q_2, \ldots, q_n]}$, и, следовательно, мы можем написать:

$$(-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1} a_{2p_2} \ldots a_{np_n} = (-1)^{[q_1, q_2, \ldots, q_n]} a_{q_1}{}^1 a_{q_2}{}^2 \ldots a_{q_n}{}^n.$$

Итак, если мы сравним соответствующие слагаемые в суммах (8) и (10), то увидим, что эти суммы в точности совпадают. В сумме (10) строки играют ту же роль, что столбцы в сумме (8), и из наших рассуждений непосредственно следует, что *если в таблице заменить все строки столбцами и столбцы строками, не меняя их порядка, то величина определителя от этого не изменится.*

Так, например, мы имеем равенство следующих двух определителей третьего порядка:

$$\begin{vmatrix} 2 & 3 & 5 \\ 7 & 0 & 1 \\ 2 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 7 & 2 \\ 3 & 0 & 1 \\ 5 & 1 & 6 \end{vmatrix}.$$

3. Основные свойства определителя. I. Формулируем прежде всего только что доказанное свойство — *величина определителя не меняется при замене строк столбцами.* В дальнейшем все, что будет доказано для столбцов, будет годиться и для строк и наоборот.

II. В предыдущем номере мы видели, что перестановка двух столбцов меняет лишь знак у определителя, то же относится и к строкам, т. е. *при перестановке двух строк (столбцов) между собой величина определителя меняет лишь знак.*

III. Если определитель имеет две одинаковые строки, то, переставляя их, мы, с одной стороны, ничего не изменим, а с другой стороны, по доказанному переменим знак определителя, т. е., обозначая через Δ величину определителя, $\Delta = -\Delta$ или $\Delta = 0$. Итак, *определитель с двумя одинаковыми строками (столбцами) равен нулю*.

IV. Линейной однородной функцией переменных $x_1, x_2, \ldots, x_n$ называется полином первой степени этих переменных без свободного члена, т. е. выражение вида:

$$\varphi(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

где коэффициенты a_i не зависят от x_s. Такая функция обладает двумя очевидными свойствами:

$$\varphi(kx_1, kx_2, \ldots, kx_n) = k\varphi(x_1, x_2, \ldots, x_n),$$
$$\varphi(x_1+y_1, x_2+y_2, \ldots, x_n+y_n) = \varphi(x_1, x_2, \ldots, x_n) + \varphi(y_1, y_2, \ldots, y_n).$$

Последнее свойство имеет место и для любого числа слагаемых. Обращаясь к формуле (8), мы видим, что каждое слагаемое написанной суммы содержит множителем один и только один элемент из каждой строки. Отсюда следует, что *определитель есть линейная однородная функция элементов какой-нибудь строки (или какого-нибудь из столбцов)*.

Следовательно, *если все элементы некоторой строки (столбца) содержат общий множитель, то его можно вынести за знак определителя*.

Величина определителя, соответствующего таблице (6), часто обозначается, как мы уже указывали выше, в виде:

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

или, более коротко:

$$|a_{ik}| \quad (i, k = 1, 2, \ldots, n).$$

Доказанное свойство можно в частном случае записать, например, в виде

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

Второе из указанных свойств линейных однородных функций дает следующее свойство определителя: *если элементы некоторой строки (столбца) суть суммы равного числа слагаемых, то определитель равен сумме определителей, в которых элементы* *упомянутой строки (столбца) заменены отдельными слагаемыми.* Так, например:

$$\begin{vmatrix} a & b & c+c' \\ d & e & f+f' \\ g & h & i+i' \end{vmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c' \\ d & e & f' \\ g & h & i' \end{vmatrix}.$$

Отметим еще одно очевидное следствие линейности и однородности. *Если все элементы некоторой строки (столбца) равны нулю, то и определитель равен нулю.*

V. Если из таблицы (6) вычеркнуть i-ю строку и k-й столбец, на пересечении которых находится элемент a_{ik}, то останется $(n-1)$ строк и столько же столбцов. Соответствующий определитель $(n-1)$-го порядка называется *минором* основного определителя n-го порядка, соответствующим элементу a_{ik}. Обозначим его через Δ_{ik} и составим произведение

$$A_{ik} = (-1)^{i+k}\Delta_{ik} \tag{17}$$

и назовем его *алгебраическим дополнением* элемента a_{ik}. Покажем теперь, что эти алгебраические дополнения являются коэффициентами той линейной однородной функции, о которой мы говорили в предыдущем свойстве, т. е., что для любой i-й строки имеет место формула:

$$\Delta = A_{i1}a_{i1} + A_{i2}a_{i2} + \ldots + A_{in}a_{in} \quad (i = 1,\ 2,\ \ldots,\ n) \tag{18}$$

и для любого k-го столбца — формула:

$$\Delta = A_{1k}a_{1k} + A_{2k}a_{2k} + \ldots + A_{nk}a_{nk} \quad (k = 1,\ 2,\ \ldots,\ n), \tag{19}$$

где Δ — величина определителя. Иначе говоря, мы должны показать, что если в сумме (8) мы соберем все члены, содержащие некоторый определенный элемент a_{ik}, то коэффициентом при этом элементе будет его алгебраическое дополнение A_{ik}, определяемое формулой (17). Обозначим предварительно этот коэффициент через B_{ik} и заметим прежде всего, что этот коэффициент представляет собою сумму произведений из $(n-1)$ элементов, причем эти произведения не содержат уже элементов i-й строки и k-го столбца.

Возьмем сначала случай $i = k = 1$ и выпишем те слагаемые суммы (8), которые содержат элемент a_{11}:

$$a_{11} \sum_{(p_2,\ \ldots,\ p_n)} (-1)^{[1,\ p_2,\ \ldots,\ p_n]} a_{2p_2} \ldots a_{np_n}.$$

Здесь суммирование должно распространяться на всевозможные перестановки $p_2, p_3, \ldots, p_n$ из чисел $2, 3, \ldots, n$. В полной перестановке $1, p_2, \ldots, p_n$ первый элемент единица по отношению ко всем следующим находится в порядке, а потому для числа беспорядков мы имеем:

$$[1, p_2, \ldots, p_n] = [p_2, \ldots, p_n],$$

причем за основную перестановку в обоих случаях берется та, где числа идут в возрастающем порядке. Мы имеем, таким образом, следующее выражение для коэффициента при a_{11}:

$$B_{11} = \sum_{(p_2, p_3, \ldots, p_n)} (-1)^{[p_2, \ldots, p_n]} a_{2p_2} \ldots a_{np_n}.$$

Эта сумма подходит под определение определителя, но только по сравнению с исходным определителем отсутствуют первая строка и первый столбец. Отсюда видно, что $B_{11} = \Delta_{11} = (-1)^{1+1}\Delta_{11} = A_{11}$, т. е. при $i = k = 1$ наше утверждение доказано. Перейдем к случаю любых i и k. Будем переставлять i-ю строку постепенно с более высокими строками так, чтобы она попала на место первой строки. Для этого придется сделать $(i-1)$ перестановок строк. Совершенно так же постепенной перестановкой приведем k-й столбец на место первого столбца. После этих перестановок элемент a_{ik} попадет в левый верхний угол на место элемента a_{11}. Строка с номером i и столбец с номером k окажутся на первом месте, а порядок остальных строк и столбцов не изменится. Полученный выше результат показывает, что после упомянутых перестановок коэффициент при a_{ik} будет равен Δ_{ik}. Но нам пришлось применить $(i-1)+(k-1)$ перестановок строк и столбцов попарно, и каждая такая перестановка добавляет множитель (-1) к определителю, т. е. в общем мы добавили множитель $(-1)^{(i-1)+(k-1)} = (-1)^{i+k}$, и, следовательно, окончательное выражение для коэффициента B_{ik} будет:

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

что и требовалось доказать. Таким образом мы доказали формулы (18) и (19). Если мы в определителе Δ заменим элементы i-й строки последовательно некоторыми числами $c_1, c_2, \ldots, c_n$, не меняя остальных строк, то в формуле (18) множители A_{is} не изменятся, и величина нового определителя будет

$$\Delta' = A_{i1}c_1 + A_{i2}c_2 + \ldots + A_{in}c_n. \tag{20}$$

В частности, если мы возьмем числа $c_1, c_2, \ldots, c_n$ равными элементам $a_{j1}, a_{j2}, \ldots, a_{jn}$ другой строки с номером j, отличным от i, то определитель Δ' будет иметь две одинаковые строки i-ю и j-ю, и его величина будет равна нулю: $\Delta' = 0$, т. е.

$$A_{i1}a_{j1} + A_{i2}a_{j2} + \ldots + A_{in}a_{jn} = 0 \quad (i \neq j). \tag{21_1}$$

Точно так же для столбцов:

$$A_{1k}a_{1l} + A_{2k}a_{2l} + \ldots + A_{nk}a_{nl} = 0 \quad (k \neq l). \tag{21_2}$$

Формулы (19) и ($21_{1,2}$) приводят нас к следующему важному в дальнейшем свойству определителя.

Если элементы некоторой строки (столбца) умножить на их алгебраические дополнения и эти произведения сложить, то получится величина определителя. Если же элементы некоторой строки (столбца) умножить на алгебраические дополнения соответствующих элементов другой строки (столбца) и эти произведения сложить, то сумма будет равна нулю.

VI. Прибавим к элементам первой строки определителя Δ элементы второй строки, умноженные на некоторый множитель p. Элементы первой строки станут равными $a_{1s}+pa_{2s}$ $(s=1, 2, \ldots, n)$, и в силу свойства IV новый определитель будет суммою двух определителей: прежнего определителя и второго определителя, у которого первая строка состоит из элементов pa_{2s} $(s=1, 2, \ldots, n)$, а остальные строки одинаковы с Δ. Вынося из первой строки p, получим одинаковые первую и вторую строки, и, следовательно, величина этого второго определителя равна нулю, т. е. вообще *величина определителя не изменится, если к элементам некоторой строки (столбца) добавить соответствующие элементы другой строки (столбца), предварительно умножив эти последние на один и тот же множитель.*

Укажем некоторые обозначения, которыми мы будем пользоваться в дальнейшем. Пусть, как и выше, имеется квадратная таблица чисел (6) и пусть l — целое положительное число, не большее, чем n. Введем следующее обозначение определителя порядка l, составленного из строк таблицы (6) с номерами $p_1, p_2, \ldots, p_l$ и столбцов с номерами $q_1, q_2, \ldots, q_l$

$$A\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix} = \begin{vmatrix} a_{p_1q_1} & a_{p_1q_2} & \ldots & a_{p_1q_l} \\ a_{p_2q_1} & a_{p_2q_2} & \ldots & a_{p_2q_l} \\ \cdot & \cdot & \cdot & \cdot \\ a_{p_lq_1} & a_{p_lq_2} & \ldots & a_{p_lq_l} \end{vmatrix}. \tag{22}$$

При этом обычно определителем первого порядка, соответствующим какому-либо числу a, называют само это число, т. е. $A\binom{p}{q}=a_{pq}$. Последовательности целых положительных чисел $p_1, p_2, \ldots, p_l$ и $q_1, q_2, \ldots, q_l$ могут быть расположены и не в порядке возрастания чисел p_s и q_s. Если в обеих этих последовательностях числа идут в возрастающем порядке, то определитель (22) называется *минором* порядка l определителя (8). Этот определитель (22) получается из (8) вычеркиванием $(n-l)$ строк и $(n-l)$ столбцов. Пусть номера этих вычеркнутых строк и столбцов в возрастающем порядке суть: $r_1, r_2, \ldots, r_{n-l}$ и $s_1, s_2, \ldots, s_{n-l}$. Минор

$$A\begin{pmatrix} r_1, & r_2, & \ldots & r_{n-l} \\ s_1, & s_2, & \ldots & s_{n-l} \end{pmatrix}$$

называется минором, дополнительным к минору (22), а выражение:

$$(-1)^{p_1+p_2+\ldots+p_l+q_1+q_2+\ldots+q_l} A\begin{pmatrix} r_1 & r_2 & \ldots & r_{n-l} \\ s_1 & s_2 & \ldots & s_{n-l}\end{pmatrix} \qquad (22_1)$$

называется *алгебраическим дополнением* минора (22). Для отдельного элемента a_{ik} это определение алгебраического дополнения совпадает с прежним определением (17).

Алгебраическое дополнение (22_1) обозначим через

$$A'\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l\end{pmatrix}.$$

Оно вполне определяется заданием определителя (22), т. е. заданием последовательностей номеров его строк $p_1, p_2, \ldots, p_l$ и столбцов $q_1, q_2, \ldots, q_l$.

Фиксируем номера строк. Величина определителя Δ является, очевидно, однородным полиномом степени l элементов этих строк, и она выражается, как можно доказать, формулой (теорема Лапласа):

$$\Delta = \sum_{q_1<q_2<\ldots<q_l} A\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l\end{pmatrix} A'\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l\end{pmatrix}, \qquad (23)$$

где суммирование проводится по всевозможным возрастающим последовательностям чисел $q_1, q_2, \ldots, q_l$, взятых из последовательности $1, 2, \ldots, n$. Число слагаемых в сумме (23) равно числу сочетаний из n элементов по l:

$$C_n^l = \frac{n(n-1)\ldots(n-l+1)}{l!},$$

поскольку порядок чисел q_s не играет роли, ибо они берутся при составлении суммы (23) только в возрастающем порядке. При $l=1$ мы имеем $A\binom{p_1}{q_1} = a_{p_1 q_1}$, и формула (23) переходит в формулу (18) при $i=p_1$. Легко построить формулу, аналогичную (23) и соответствующую разложению Δ по элементам каких-либо выбранных столбцов. Мы не будем в дальнейшем пользоваться формулой (23) и не приводим ее доказательства.

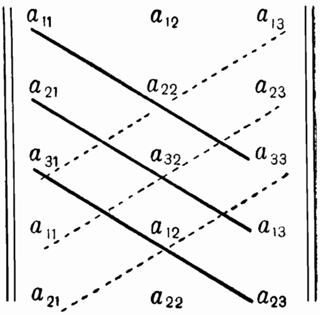
4. Вычисление определителя. Вычисление определителя второго порядка очень просто. Согласно формуле (11) достаточно написать таблицу

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

и взять произведение элементов, стоящих на диагонали, подчеркнутой сплошной чертой, с собственным знаком, и подчеркнутой пунктиром — с обратным знаком.

Перейдем к определителю третьего порядка. В раскрытом виде он записан нами в формуле (3). Нетрудно проверить, что его можно образовать следующим путем: выписываем таблицу, дающую определитель, и подписываем под нею еще раз первую и вторую строки. Будем иметь таблицу, содержащую шесть диагоналей, по три элемента в каждой.

Берем произведение элементов, стоящих на диагоналях, подчеркнутых сплошной чертой, без изменения и на диагоналях, подчеркнутых пунктиром, — со знаком минус. Сумма этих шести произведений и дает определитель (*правило Сарруса*).



Правило это не обобщается на определители более высокого порядка, и там приходится поступать иначе для сокращения вычислений. Полезно, например, пользоваться указанным в предыдущем номере свойством VI определителя. Выясним это на примере. Возьмем определитель четвертого порядка

$$\Delta = \begin{vmatrix} 3 & 5 & 1 & 0 \\ 2 & 1 & 4 & 5 \\ 1 & 7 & 4 & 2 \\ -3 & 5 & 1 & 1 \end{vmatrix}.$$

Умножаем третью строчку на (—2) и складываем со второй; кроме того, умножаем эту же строчку на 3 и складываем с четвертой и вычитаем из первой. Таким образом, в силу упомянутого свойства, придем к определителю, равному вышенаписанному, причем этот определитель будет иметь в первом столбце три нуля

$$\Delta = \begin{vmatrix} 0 & -16 & -11 & -6 \\ 0 & -13 & -4 & 1 \\ 1 & 7 & 4 & 2 \\ 0 & 26 & 13 & 7 \end{vmatrix}.$$

Отсюда, разлагая по элементам первого столбца, согласно формуле (19), получим:

$$\Delta = \begin{vmatrix} -16 & -11 & -6 \\ -13 & -4 & 1 \\ 26 & 13 & 7 \end{vmatrix}.$$

Умножаем третий столбец на 4 и складываем со вторым и затем умножаем его на 13 и складываем с первым. Получаем таким образом:

$$\Delta = \begin{vmatrix} -94 & -35 & -6 \\ 0 & 0 & 1 \\ 117 & 41 & 7 \end{vmatrix} = - \begin{vmatrix} -94 & -35 \\ 117 & 41 \end{vmatrix} = 94 \cdot 41 - 35 \cdot 117 = -241.$$

5. Примеры. I. Пусть требуется вычислить объем параллелепипеда, три ребра которого, выходящих из одной вершины, суть векторы **A**, **B** и **C**. Как известно [II, 117], искомый объем выражается скалярным произведением вектора **A** на векторное произведение ($\mathbf{B} \times \mathbf{C}$):

$$V = \mathbf{A}(\mathbf{B} \times \mathbf{C}). \tag{24}$$

При этом объем получается со знаком плюс, если векторы **A**, **B**, **C** дают ту же ориентировку, что координатные оси, и со знаком минус, если упомянутые ориентировки различны. Составляющие векторного произведения равны

$$B_yC_z - B_zC_y, \quad B_zC_x - B_xC_z, \quad B_xC_y - B_yC_x,$$

и таким образом скалярное произведение, входящее в формулу (24), будет:

$$A_x(B_yC_z - B_zC_y) + A_y(B_zC_x - B_xC_z) + A_z(B_xC_y - B_yC_x).$$

Нетрудно видеть, что эта последняя сумма представляет собою определитель третьего порядка, т. е.

$$V = \begin{vmatrix} A_x & B_x & C_x \\ A_y & B_y & C_y \\ A_z & B_z & C_z \end{vmatrix}. \tag{25}$$

Равенство нулю этого определителя будет нам показывать, что объем равен нулю, иначе говоря, что наши три вектора компланарны, т. е. находятся в одной плоскости. Если мы в определителе переставим две строки (столбца), например первую и вторую, то этим самым порядок векторов **A**, **B**, **C** заменится другим порядком **B**, **A**, **C** и если векторы в прежней последовательности давали ту же ориентировку, что координатные оси, то теперь они будут давать уже иную ориентировку, и наоборот. В соответствии с этим величина определителя изменит знак.

Если аналогичным образом в плоскости XY рассмотрим два вектора с составляющими (A_x, A_y) и (B_x, B_y), то площадь параллелограмма, построенного на этих двух векторах, будет равна определителю второго порядка:

$$P=\begin{vmatrix} A_x & B_x \\ A_y & B_y \end{vmatrix}.$$

Рассмотрим теперь треугольник, координаты вершин которого суть

$$M_1(x_1,\ y_1),\quad M_2(x_2,\ y_2),\quad M_3(x_3,\ y_3).$$

Берем векторы $\mathbf{A}=\overline{M_1M_2}$ и $\mathbf{B}=\overline{M_1M_3}$ с составляющими:

$$\overline{M_1M_2}(x_2-x_1,\quad y_2-y_1),\quad \overline{M_1M_3}(x_3-x_1,\ y_3-y_1),$$

и площадь нашего треугольника может быть выражена в виде:

$$P=\frac{1}{2}\begin{vmatrix} x_2-x_1 & x_3-x_1 \\ y_2-y_1 & y_3-y_1 \end{vmatrix}.$$

Нетрудно показать, что написанный определитель второго порядка можно заменить определителем третьего порядка и написать формулу в следующем виде:

$$P=\frac{1}{2}\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix}.$$

Равенство нулю этого определителя дает условие того, что точки M_1, M_2, M_3 находятся на одной прямой. Иначе говоря, уравнение прямой, проходящей через две заданные точки $(x_1,\ y_1)$ и $(x_2,\ y_2)$, может быть написано в виде:

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

II. Нетрудно составить, пользуясь определителями, уравнения некоторых геометрических мест. Пусть, например, ищется уравнение окружности, проходящей через три заданные точки: $(x_1,\ y_1)$, $(x_2,\ y_2)$ и $(x_3,\ y_3)$. Легко видеть, что это уравнение запишется, при помощи определителя четвертого порядка, следующим образом:

$$\begin{vmatrix} x^2+y^2 & x_1^2+y_1^2 & x_2^2+y_2^2 & x_3^2+y_3^2 \\ x & x_1 & x_2 & x_3 \\ y & y_1 & y_2 & y_3 \\ 1 & 1 & 1 & 1 \end{vmatrix}=0. \tag{26}$$

Действительно, разлагая по элементам первого столбца, убеждаемся, что написанное уравнение есть уравнение второй степени, в котором коэффициенты при x^2 и y^2 одинаковы, и член с произведением xy отсутствует, т. е. уравнению (26) соответствует окружность. Наконец, если мы в этом уравнении подставим $x=x_k$ и $y=y_k$ $(k=1, 2, 3)$, то первый столбец определителя будет совпадать с одним из следующих, и уравнение будет удовлетворено, т. е. окружность действительно проходит через три данные точки. Заметим, что если три заданные точки находятся на одной прямой, то в уравнении (26) коэффициент при (x^2+y^2) окажется равным нулю, и уравнению будет соответствовать не окружность, а прямая.

Совершенно так же уравнение плоскости в пространстве с осями OX, OY, OZ, проходящей через три заданные точки (x_1, y_1, z_1), (x_2, y_2, z_2) и (x_3, y_3, z_3), может быть написано в виде определителя четвертого порядка:

$$\begin{vmatrix} x & x_1 & x_2 & x_3 \\ y & y_1 & y_2 & y_3 \\ z & z_1 & z_2 & z_3 \\ 1 & 1 & 1 & 1 \end{vmatrix}=0. \tag{27}$$

Если три заданные точки лежат на одной прямой, то уравнение (27) превратится в тождество $0=0$.

III. Рассмотрим определитель D_n порядка n, каждая строчка которого состоит из степеней некоторого числа, начиная с $(n-1)$-й степени и до нулевой включительно:

$$D_n=\begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{vmatrix}. \tag{28}$$

При $n=1$ и 2 будем иметь:

$$D_1=1; \quad D_2=x_1-x_2.$$

Для раскрытия определителя D_3 мы заменим в первой строчке этого определителя число x_1 буквой x. Получим определитель вида:

$$D_3(x)=\begin{vmatrix} x^2 & x & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix}.$$

Разлагая его по элементам первой строки, видим, что $D_3(x)$ есть полином второй степени от x. Если подставить в определитель $x=x_2$ или $x=x_3$, то первая строка станет одинаковой со второй или третьей, и величина определителя будет нуль, т. е. квадратный трехчлен $D_3(x)$ имеет корни x_2 и x_3 и может быть представлен в виде:

$$D_3(x) = A_3(x - x_2)(x - x_3),$$

где A_3 есть коэффициент при x^2, т. е. алгебраическое дополнение элемента x^2 определителя $D_3(x)$, стоящего в левом верхнем углу. Отсюда следует

$$A_3 = \begin{vmatrix} x_2 & 1 \\ x_3 & 1 \end{vmatrix},$$

т. е. A_3 есть определитель D_2, составленный из чисел x_2 и x_3. Окончательно:

$$D_3(x) = (x_2 - x_3)(x - x_2)(x - x_3).$$

Подставляя $x = x_1$, получим для D_3 выражение в виде произведения трех множителей:

$$D_3 = \begin{array}{r} (x_1 - x_2)(x_1 - x_3) \\ (x_2 - x_3). \end{array}$$

Совершенно аналогично, имея выражение D_3, можно получить выражение для D_4. Оно будет иметь вид произведения шести множителей:

$$D_4 = \begin{array}{r} (x_1 - x_2)(x_1 - x_3)(x_1 - x_4) \\ (x_2 - x_3)(x_2 - x_4). \\ (x_3 - x_4) \end{array}$$

Точно так же, при любом n получим следующее выражение для определителя D_n, который обычно называется определителем Вандермонда:

$$D_n = \begin{array}{r} (x_1 - x_2)(x_1 - x_3)\ldots(x_1 - x_n) \\ (x_2 - x_3)\ldots(x_2 - x_n). \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ (x_{n-1} - x_n) \end{array} \qquad (29)$$

Написанное выражение имеет интересную связь с основным определением определителя. Любой определитель порядка n может быть записан в виде:

$$\begin{vmatrix} x_{1n} & x_{1\,n-1} & \ldots & x_{11} \\ x_{2n} & x_{2\,n-1} & \ldots & x_{21} \\ \ldots & \ldots & \ldots & \ldots \\ x_{nn} & x_{n\,n-1} & \ldots & x_{n1} \end{vmatrix}. \qquad (30)$$

Заменим в нем чисто формально каждый элемент x_{ik} на x_i^{k-1}. После такой замены определитель (30) перейдет, очевидно, в определитель Вандермонда (28). Отсюда вытекает непосредственно следующее правило образования суммы, дающей величину определителя (30): в выражении (29) открываем скобки и в каждом из полученных после этого членов заменяем x_k^{s-1} на x_{ks}, причем, если такой член не содержит степени некоторого числа x_k, то в соответствующем произведении надо добавить множитель x_k^0, который после упомянутой замены перейдет в x_{k1}. Заметим, что это последнее правило может быть приняго за определение определителя.

IV. Рассмотрим выражение, с которым нам придется иметь дело в последующем:

$$\Delta(x)=\begin{vmatrix} a_{11}+x & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22}+x & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33}+x & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn}+x \end{vmatrix} \tag{31}$$

и разложим его по степеням буквы x. Для этого перепишем его предварительно в следующем виде:

$$\Delta(x)=\begin{vmatrix} a_{11}+x & a_{12}+0 & a_{13}+0 & \dots & a_{1n}+0 \\ a_{21}+0 & a_{22}+x & a_{23}+0 & \dots & a_{2n}+0 \\ a_{31}+0 & a_{32}+0 & a_{33}+x & \dots & a_{3n}+0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}+0 & a_{n2}+0 & a_{n3}+0 & \dots & a_{nn}+x \end{vmatrix}. \tag{32}$$

Каждый из сголбцов этого определителя есть сумма двух слагаемых, и, применяя несколько раз свойства IV определителя, мы представим его в виде суммы 2^n определителей, столбцы которых уже не содержат сумм. Если во всех столбцах выражения (32) вычеркнуть вторые слагаемые, то мы получим член, не содержащий буквы x, т. е. свободный член в разложении $\Delta(x)$:

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{33}$$

Наоборот, если вычеркнуть во всех столбцах первые слагаемые, то получится старший член полинома $\Delta(x)$, равный

$$\begin{vmatrix} x & 0 & 0 & \dots & 0 \\ 0 & x & 0 & \dots & 0 \\ 0 & 0 & x & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & x \end{vmatrix}=x^n.$$

Рассмотрим теперь средние члены полинома. Положим, что в столбцах с номерами $p_1, p_2, \ldots, p_s$ мы удерживаем вторые слагаемые, а в остальных столбцах — первые слагаемые. При этом каждый столбец с номером p_k $(k=1, 2, \ldots, s)$ будет состоять сплошь из нулей, кроме одного элемента, равного x и стоящего на главной диагонали, т. е. на пересечении строки и столбца с одинаковым номером. Разлагая полученный определитель последовательно по элементам столбцов $p_1, p_2, \ldots, p_s$, мы от этих столбцов получим множитель x^s и должны будем вычеркнуть строки с номерами $p_1, p_2, \ldots, p_s$ и столбцы с такими же номерами. После каждого такого вычеркивания алгебраические дополнения соответствующего элемента будут равны в точности минорам, ввиду совпадения номера вычеркиваемой строки и столбца. Итак, при любом выборе номеров столбцов p_k $(k=1, 2, \ldots, s)$ наш определитель будет содержать x^s с коэффициентом, равным определителю порядка $(n-s)$, получаемому из основного определителя (33) вычеркиванием тех строк и столбцов, на пересечении которых стоят элементы главной диагонали $a_{p_1p_1}, a_{p_2p_2}, \ldots, a_{p_sp_s}$. Обозначим этот определитель порядка $(n-s)$ символом $\Delta_{\substack{p_1p_2\ldots p_s\\ p_1p_2\ldots p_s}}$. Он называется обычно *главным минором определителя* Δ, порядка $(n-s)$. Выбирая числа $p_1, p_2, \ldots, p_s$ различным образом, получим в общем окончательный коэффициент при x^s в выражении $\Delta(x)$, в виде суммы всевозможных главных миноров порядка $(n-s)$, т. е.

$$\Delta(x)=x^n+S_1x^{n-1}+S_2x^{n-2}+\ldots+S_{n-1}x+S_n,$$

где S_k есть сумма всех главных миноров определителя Δ порядка k и, в частности, S_n равно Δ. Напишем явное выражение для коэффициента

$$S_k=\sum_{p_1<p_2<\ldots<p_{n-k}}^{(1,\,2,\,\ldots,\,n)}\Delta_{\substack{p_1p_2\ldots p_{n-k}\\ p_1p_2\ldots p_{n-k}}}=$$

$$=\sum_{q_1<q_2<\ldots<q_k}\begin{vmatrix} a_{q_1q_1} & a_{q_1q_2} & \ldots & a_{q_1q_k}\\ a_{q_2q_1} & a_{q_2q_2} & \ldots & a_{q_2q_k}\\ \cdot & \cdot & \cdot & \cdot\\ a_{q_kq_1} & a_{q_kq_2} & \ldots & a_{q_kq_k}\end{vmatrix}. \qquad (34)$$

Здесь суммирование распространяется на всевозможные комбинации из k чисел q_1, q_2, q_k, идущих в возрастающем порядке и взятых среди чисел $1, 2, \ldots, n$. Если бы мы стали суммировать в выражении (34) по каждому значку q_j просто по всем значениям от 1 до n, то в перестановке $q_1, q_2, \ldots, q_k$ целые числа шли бы не только в возрастающем порядке, но и во всех других возможных порядках. Точнее говоря, всякая возрастающая последовательность при суммировании по всем q_j от 1 до n дала бы всего $k!$ перестановок. Заметим теперь, что при перестановке каких-нибудь двух чисел q_i и q_j величина определителя, стоящего в формуле (34), не изменится. Действительно, если мы переставим, например, q_1 и q_2, то, тем самым, в упомянутом определителе переставятся первая и вторая строки и первый и второй столбец, что не повлияет на величину определителя. Таким образом из предыдущего следует, что если мы в выражении (34) будем суммировать по каждому из чисел q_j просто от 1 до n, то каждое слагаемое суммы (34) повторится $k!$ раз, и, следовательно, мы можем написать выражение коэффициента S_k в виде:

$$S_k = \frac{1}{k!} \sum_{q_1=1}^{n} \sum_{q_2=1}^{n} \cdots \sum_{q_k=1}^{n} A\begin{pmatrix} q_1 & q_2 & \dots & q_k \\ q_1 & q_2 & \dots & q_k \end{pmatrix}. \qquad (35)$$

6. Теорема об умножении определителей. В этом номере мы выведем формулу для произведения двух определителей одного и того же порядка.

Пусть имеются два определителя порядка n:

$$\Delta = | a_{ik} |_1^n \qquad (36_1)$$

и

$$\Delta_1 = | b_{ik} |_1^n. \qquad (36_2)$$

Составим новый определитель, элементы которого выражаются формулами:

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{sk} \qquad (i,\ k = 1,\ 2,\ \dots,\ n) \qquad (37)$$

и докажем, что этот определитель равен произведению определителей (36_1) и (36_2). Начнем со случая $n = 2$. Принимая во внимание формулы (37) и разлагая определитель на слагаемые, согласно свойству IV из [**3**], получим:

$$\begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} = \begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{vmatrix} +$$
$$+ \begin{vmatrix} a_{11}b_{11} & a_{12}b_{22} \\ a_{21}b_{11} & a_{22}b_{22} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{11}b_{12} \\ a_{22}b_{21} & a_{21}b_{12} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{vmatrix}.$$

Вынося одинаковые множители из элементов одного и того же столбца, мы получим в первом и четвертом слагаемом определители с одинаковыми столбцами, равные нулю. Переставляя еще в одном из определителей столбцы, будем иметь:

$$\begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = b_{11}b_{22}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + b_{12}b_{21}\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} =$$

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

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix},$$

что и требовалось доказать.

В общем случае порядка n после применения свойства IV из [**3**] будем иметь:

$$|c_{ik}|_1^n = \sum_{(s_1, s_2, \ldots, s_n)} \begin{vmatrix} a_{1s_1}b_{s_11} & a_{1s_2}b_{s_22} & \ldots & a_{1s_n}b_{s_nn} \\ a_{2s_1}b_{s_11} & a_{2s_2}b_{s_22} & \ldots & a_{2s_n}b_{s_nn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{ns_1}b_{s_11} & a_{ns_2}b_{s_22} & \ldots & a_{ns_n}b_{s_nn} \end{vmatrix}, \tag{38}$$

где переменные суммирования $s_1, s_2, \ldots, s_n$ принимают целые значения $1, 2, \ldots, n$. Слагаемые этой суммы могут быть записаны в виде:

$$b_{s_11}b_{s_22} \ldots b_{s_nn}\begin{vmatrix} a_{1s_1} & a_{1s_2} & \ldots & a_{1s_n} \\ a_{2s_1} & a_{2s_2} & \ldots & a_{2s_n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{ns_1} & a_{ns_2} & \ldots & a_{ns_n} \end{vmatrix}. \tag{39}$$

Если среди чисел $s_1, s_2, \ldots, s_n$ есть одинаковые, то написанный определитель имеет одинаковые столбцы, и соответствующее слагаемое равно нулю. Таким образом, мы можем ограничиться рассмотрением таких слагаемых, у которых среди чисел s_k нет одинаковых, и таким образом эта последовательность чисел $s_1, s_2, \ldots, s_n$ представляет собою некоторую перестановку из чисел $1, 2, \ldots, n$. Умножим выражение (39) дважды на $(-1)^{[s_1, s_2, \ldots, s_n]}$, в результате чего оно, очевидно, не изменится, и это выражение можно переписать в виде произведения двух множителей:

$$(-1)^{[s_1, s_2, \ldots, s_n]}\begin{vmatrix} a_{1s_1} & a_{1s_2} & \ldots & a_{1s_n} \\ a_{2s_1} & a_{2s_2} & \ldots & a_{2s_n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{ns_1} & a_{ns_2} & \ldots & a_{ns_n} \end{vmatrix} \cdot (-1)^{[s_1, s_2, \ldots, s_n]} b_{s_11}b_{s_22}\ldots b_{s_nn}.$$

В первом множителе путем нескольких транспозиций перестановку $s_1, s_2, \ldots, s_n$ приведем к виду $1, 2, \ldots, n$. При каждой транспозиции величина $(-1)^{[s_1, s_2, \ldots, s_n]}$ и написанный определитель будут лишь менять знак, а весь первый множитель остается без изменения. Таким образом выражение (39) может быть переписано в виде:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \cdot (-1)^{[s_1, s_2, \ldots, s_n]} b_{s_1 1} b_{s_2 2} \ldots b_{s_n n},$$

и, следовательно, возвращаясь к сумме (38), получим:

$$| c_{ik} |_1^n = \Delta \sum_{(s_1, s_2, \ldots, s_n)} (-1)^{[s_1, s_2, \ldots, s_n]} b_{s_1 1} b_{s_2 2} \ldots b_{s_n n},$$

где суммирование распространяется на все перестановки $s_1, s_2, \ldots, s_n$ из чисел $1, 2, \ldots, n$. Написанная сумма есть определитель Δ_1, т. е. $| c_{ik} |_1^n = \Delta\Delta_1$, что и требовалось доказать. Формула (37) для c_{ik} сводится к следующему: элементы i-й строки определителя Δ умножаются на соответствующие элементы k-го столбца второго определителя, и эти произведения складываются. Мы знаем, что в определителе можно заменить строки столбцами, не меняя его величины. Следовательно, предыдущее правило умножения строки на столбец можно заменить тремя другими правилами: правилом строка на строку, столбец на столбец и столбец на строку.

Формулируем окончательно теорему: *пусть имеются два определителя n-го порядка:*

$$| a_{ik} | \text{ и } | b_{ik} |.$$

Составляем новый определитель

$$| c_{ik} |,$$

элементы которого вычисляются по одной из следующих формул:

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{sk}, \tag{40$_1$}$$

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{ks} \tag{40$_2$}$$

$$c_{ik}=\sum_{s=1}^{n} a_{si}b_{sk}, \tag{40$_3$}$$

$$c_{ik}=\sum_{s=1}^{n} a_{si}b_{ks} \quad (i, k=1, 2, \ldots, n). \tag{40$_4$}$$

При этом величина определителя $|c_{ik}|$ *равна произведению определителей* $|a_{ik}|$ *и* $|b_{ik}|$.

Пример. Наряду с основным определителем

$$\Delta=|a_{ik}|$$

рассмотрим определитель, составленный из алгебраических дополнений его элементов

$$|A_{ik}|.$$

Согласно приведенной выше теореме выразим произведение $|a_{ik}|\cdot|A_{ik}|$ также в виде определителя, умножая строку на строку. Мы будем иметь для этого определителя следующие элементы:

$$c_{ik}=\sum_{s=1}^{n} a_{is}A_{ks}.$$

В силу свойства V определителя мы получим:

$$c_{ik}=0 \text{ при } i\neq k;\ c_{ii}=\Delta,$$

т. е.

$$|a_{ik}|\cdot|A_{ik}|=\begin{vmatrix} \Delta & 0 & 0 & \ldots & 0 \\ 0 & \Delta & 0 & \ldots & 0 \\ 0 & 0 & \Delta & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \Delta \end{vmatrix}$$

или

$$|a_{ik}|_1^n\,|A_{ik}|_1^n=\Delta^n, \text{ т. е. } \Delta\,|A_{ik}|_1^n=\Delta^n.$$

Считая Δ отличным от нуля и сокращая на него, будем иметь:

$$|A_{ik}|_1^n=\Delta^{n-1}. \tag{41}$$

Если элементы $a_{ik}=a_{ik}^{(0)}$ таковы, что определитель Δ равен нулю, то можно указать такие значения a_{ik}, сколь угодно близкие $a_{ik}^{(0)}$, при которых определитель Δ отличен от нуля. Для этих значений a_{ik} будем иметь формулу (41), а потому в пределе при $a_{ik}\to a_{ik}^{(0)}$ эта формула будет иметь место и при $a_{ik}=a_{ik}^{(0)}$, т. е. эта формула имеет место и при $\Delta=0$. Если выразить Δ и A_{ik} через элементы a_{ik}, то она представляет собою тождество относительно a_{ik}.

7. Прямоугольные таблицы. В дальнейшем мы будем встречаться и с такими таблицами чисел, в которых число строк и столбцов может быть и неодинаково. Рассмотрим такую более общую таблицу:

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}. \tag{42}$$

Она содержит m строк и n столбцов, причем числа m и n могут быть или различны, или равны. Вычеркивая из этой таблицы некоторые строки и столбцы так, чтобы число оставшихся строк и столбцов было одинаково, мы сможем из оставшихся строк и столбцов составить определитель. Такие определители назовем определителями, входящими в состав таблицы (42). Наивысший порядок, который они могут иметь, равен, очевидно, наименьшему из двух чисел m и n, а наименьший порядок этих определителей равен единице, причем определители первого порядка суть сами элементы таблицы (42). Положим, что все определители некоторого порядка l, входящие в таблицу, равны нулю. Нетрудно видеть, что тогда и все определители порядка $(l+1)$, входящие в таблицу, также равны нулю. Действительно, всякий такой определитель порядка $(l+1)$ можно представить в виде суммы произведений элементов его некоторой строки на алгебраические дополнения этих элементов. Но последние с точностью до знака совпадают с некоторыми определителями порядка l таблицы, и, следовательно, все равны нулю. Раз все определители порядка $(l+1)$ равны нулю, то так же, как и выше, все определители порядка $(l+2)$ также будут равны нулю и т. д. Итак, если все определители некоторого определенного порядка, входящие в таблицу (42), равны нулю, то и все определители более высокого порядка этой таблицы также равны нулю.

Введем важное для дальнейшего понятие о ранге таблицы, или, как говорят, матрицы (42). *Рангом матрицы* (42) называется наивысший порядок определителя этой таблицы, отличного от нуля, т. е. если ранг таблицы есть k, то среди определителей порядка k, входящих в эту таблицу, есть по крайней мере один, отличный от нуля, но все определители таблицы порядка $(k+1)$ равны нулю.

Пусть наряду с таблицей (42) имеется таблица

$$\begin{Vmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{Vmatrix}, \tag{43}$$

содержащая n строк и m столбцов. Образуем m^2 чисел

$$c_{ik}=\sum_{s=1}^{n} a_{is}b_{sk} \quad (i, k=1, 2, \ldots, m). \tag{44}$$

Квадратная таблица, составленная из чисел c_{ik}, называется обычно произведением прямоугольных таблиц (42) и (43).

Докажем теорему, которая является обобщением теоремы об умножении определителей.

Теорема. *Если $m \leqslant n$, то*

$$|c_{ik}|_1^m=\sum_{r_1<r_2<\ldots<r_m} A\begin{pmatrix}1 & 2 & \ldots & m\\ r_1 & r_2 & \ldots & r_m\end{pmatrix} B\begin{pmatrix}r_1 & r_2 & \ldots & r_m\\ 1 & 2 & \ldots & m\end{pmatrix}, \tag{45}$$

где суммирование распространяется на все значения r_k из ряда чисел $1, 2, \ldots, n$, удовлетворяющие указанному неравенству. Если же $m>n$, то определитель $|c_{ik}|_1^m$ равен нулю.

Смысл символов

$$A\begin{pmatrix}1 & 2 & \ldots & m\\ r_1 & r_2 & \ldots & r_m\end{pmatrix} \text{ и } B\begin{pmatrix}r_1 & r_2 & \ldots & r_m\\ 1 & 2 & \ldots & m\end{pmatrix}$$

указан в [**3**]. Второй из них обозначает определитель, составленный из элементов таблицы (43), принадлежащих строкам $r_1, r_2, \ldots, r_m$ и столбцам $1, 2, \ldots, m$. При $m=n$ сумма, входящая в формулу (45), содержит лишь одно слагаемое, соответствующее $r_1=1$, $r_2=2, \ldots$, $r_m=m$, и формула (45) выражает теорему об умножении определителей.

Рассмотрим случай $m<n$. Доказательство формулы (45) будет аналогично доказательству теоремы об умножении определителей. Как и при этом доказательстве, мы имеем:

$$|c_{ik}|_1^m=\sum_{(s_1, \ldots, s_m)} A\begin{pmatrix}1 & 2 & \ldots & m\\ s_1 & s_2 & \ldots & s_m\end{pmatrix} b_{s_1 1}b_{s_2 2}\ldots b_{s_m m}, \tag{46}$$

причем каждое из чисел s_k может принимать значения $1, 2, \ldots, n$, и можно отбросить те слагаемые, у которых среди чисел s_k есть равные, так как эти слагаемые равны нулю. Возьмем какую-либо определенную последовательность чисел $r_1<r_2<\ldots<r_m$ из ряда чисел $1, 2, \ldots, n$ и выделим из суммы (45) те слагаемые, у которых совокупность чисел $s_1, s_2, \ldots, s_m$ совпадает с совокупностью чисел $r_1, r_2, \ldots, r_m$. Мы получим таким образом часть суммы (46):

$$\sum_{(t_1, t_2, \ldots, t_m)} A\begin{pmatrix}1 & 2 & \ldots & m\\ t_1 & t_2 & \ldots & t_m\end{pmatrix} b_{t_1 1}b_{t_2 2}\ldots b_{t_m m}, \tag{47}$$

где сумма распространяется на всевозможные перестановки $(t_1, t_2, \ldots, t_m)$ из чисел $r_1, r_2, \ldots, r_m$. Умножая каждое слагаемое суммы (47) дважды на $(-1)^{[t_1, t_2, \ldots, t_m]}$, мы докажем совершенно так же, как и в [6], что эта сумма равна

$$A\begin{pmatrix}1 & 2 & \ldots & m\\ r_1 & r_2 & \ldots & r_m\end{pmatrix}B\begin{pmatrix}r_1 & r_2 & \ldots & r_m\\ 1 & 2 & \ldots & m\end{pmatrix}.$$

Чтобы получить всю сумму, входящую в формулу (46), достаточно просуммировать это произведение по всем $r_1 < r_2 < \ldots < r_m$, что и дает формулу (45). Положим, наконец, что $m > n$. Мы можем при этом добавить к таблице (42) $(m-n)$ столбцов, состоящих из нулей, а к таблице (43) — $(m-n)$ строк, состоящих из нулей. Если после такого добавления мы будем вычислять c_{ik} не по формулам (44), а по формулам:

$$c_{ik} = \sum_{s=1}^{m} a_{is}b_{sk} \quad (i,\ k = 1,\ 2,\ \ldots,\ m), \tag{48}$$

то получим прежние значения c_{ik}, так как добавленные слагаемые в правой части (48) равны нулю. С другой стороны, после указанного добавления таблицы (42) и (43) превратились в квадратные таблицы, которым соответствуют определители, равные нулю, а потому, согласно теореме об умножении определителей, и определитель $|c_{ik}|_1^m$ равен нулю, и теорема полностью доказана.

З а м е ч а н и е. Если две прямоугольные таблицы имеют каждая m строк и n столбцов, то, умножая строка на строку:

$$c_{ik} = \sum_{s=1}^{n} a_{is}b_{ks} \quad (i,\ k = 1,\ 2,\ \ldots,\ m),$$

получим определитель $|c_{ik}|_1^m$, величина которого равна нулю при $m > n$, а при $m \leqslant n$ выражается формулой:

$$|c_{ik}|_1^m = \sum_{r_1 < r_2 < \ldots < r_m} A\begin{pmatrix}1 & 2 & \ldots & m\\ r_1 & r_2 & \ldots & r_m\end{pmatrix}B\begin{pmatrix}1 & 2, & \ldots & m\\ r_1 & r_2, & \ldots & r_m\end{pmatrix}.$$

С л е д с т в и е. Пусть имеются две квадратные таблицы порядка n составленные из элементов a_{ik} и b_{ik}, а числа c_{ik} определяются формулами (44). Выразим любой минор $C\begin{pmatrix}p_1 & p_2 & \ldots & p_l\\ q_1 & q_2 & \ldots & q_l\end{pmatrix}$ определителя $|c_{ik}|_1^n$ через миноры определителей $|a_{ik}|_1^n$ и $|b_{ik}|_1^n$. Нетрудно видеть что квадратная таблица, которая образует минор $C\begin{pmatrix} p_1 & p_2 & \dots & p_l \\ q_1 & q_2 & \dots & q_l \end{pmatrix}$, является произведением прямоугольных таблиц:

$$\begin{Vmatrix} a_{p_1 1} & a_{p_1 2} & \dots & a_{p_1 n} \\ a_{p_2 1} & a_{p_2 2} & \dots & a_{p_2 n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{p_l 1} & a_{p_l 2} & \dots & a_{p_l n} \end{Vmatrix} \quad \text{и} \quad \begin{Vmatrix} b_{1q_1} & b_{1q_2} & \dots & b_{1q_l} \\ b_{2q_1} & b_{2q_2} & \dots & b_{2q_l} \\ \cdot & \cdot & \cdot & \cdot \\ b_{nq_1} & b_{nq_2} & \dots & b_{nq_l} \end{Vmatrix}.$$

Применяя доказанную теорему, мы и получим искомое выражение:

$$C\begin{pmatrix} p_1 & p_2 & \dots & p_l \\ q_1 & q_2 & \dots & q_l \end{pmatrix} = \\ = \sum_{r_1 < r_2 < \dots < r_l} A\begin{pmatrix} p_1 & p_2 & \dots & p_l \\ r_1 & r_2 & \dots & r_l \end{pmatrix} B\begin{pmatrix} r_1 & r_2 & \dots & r_l \\ q_1 & q_2 & \dots & q_l \end{pmatrix}, \tag{49}$$

где r_k принимают значения из ряда $1, 2, \dots, n$. Пусть R_A, R_B, R_C — ранги таблиц $\| a_{ik} \|_1^n$, $\| b_{ik} \|_1^n$ и $\| c_{ik} \|_1^n$. Если, например, $R_A < n$, и в формуле (49) мы возьмем любое $l > R_A$, то все $A\begin{pmatrix} p_1 & p_2 & \dots & p_l \\ r_1 & r_2 & \dots & r_l \end{pmatrix}$ будут, в силу определения R_A, равны нулю, а потому и все $C\begin{pmatrix} p_1 & p_2 & \dots & p_l \\ q_1 & q_2 & \dots & q_l \end{pmatrix}$ равны нулю. Отсюда следует, что $R_C < l$, т. е. $R_C \leqslant R_A$. Если ранг таблицы $\| a_{ik} \|_1^n$ равен n, то, очевидно, $R_C \leqslant R_A$, ибо, $R_C \leqslant n$. Совершенно аналогично $R_C \leqslant R_B$. В дальнейшем мы покажем, что если определитель $| b_{ik} |_1^n \neq 0$, то $R_C = R_A$, а если $| a_{ik} |_1^n \neq 0$, то $R_C = R_B$.

§ 2. РЕШЕНИЕ СИСТЕМ УРАВНЕНИЙ

8. Теорема Крамера. Установив понятие об определителе и выяснив его основные свойства, мы переходим теперь к применению этого понятия к решению систем уравнений первой степени. Рассмотрим сначала основной случай, когда число уравнений совпадает с числом неизвестных. Мы можем записать такую систему, содержащую n уравнений и n неизвестных, в виде:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n, \end{array}\right\} \tag{1}$$

причем обозначения коэффициентов такие же, какие мы ввели в [1] для случая трех уравнений с тремя неизвестными. Сделаем одно предположение, а именно будем считать, что определитель системы, т. е. определитель, соответствующий таблице коэффициентов a_{ik} системы, отличен от нуля

$$\Delta = |a_{ik}| \neq 0. \tag{2}$$

Умножим обе части уравнений системы (1) на алгебраические дополнения элементов k-го столбца этого определителя, т. е. обе части первого уравнения системы умножим на A_{1k}, второго на A_{2k} и т. д. Полученные таким образом уравнения сложим. В результате мы придем к уравнению, в правой части которого стоит сумма

$$A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n,$$

а в левой части коэффициент при неизвестном x_l выражается суммой

$$A_{1k}a_{1l} + A_{2k}a_{2l} + \ldots + A_{nk}a_{nl} \quad (l = 1, 2, \ldots, n).$$

Эта последняя сумма равна нулю при $l \neq k$ и определителю Δ при $l = k$, т. е. мы приходим к уравнению вида:

$$\Delta \cdot x_k = A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n.$$

Проделав это для каждого значка k, мы, как следствие уравнений (1), получим систему новых уравнений

$$\Delta \cdot x_k = A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n \quad (k = 1, 2, \ldots, n). \tag{3}$$

Нетрудно показать, что и, наоборот, из системы (3), как следствие, получается система (1). Действительно, умножим обе части уравнения (3) на a_{lk} и просуммируем затем по всем k от 1 до n. Пользуясь опять свойством V определителя, мы придем, как нетрудно видеть, к уравнению

$$\Delta \cdot (a_{l1}x_1 + a_{l2}x_2 + \ldots + a_{ln}x_n) = \Delta \cdot b_l, \tag{4}$$

что, после сокращения на множитель Δ, отличный от нуля, даст нам уравнение с номером l системы (1), причем мы можем это проделать для любого l. Итак, системы (1) и (3) равносильны, и мы можем вместо системы (1) решать систему (3). Последняя система решается непосредственно и дает одно и только одно решение, вычисляемое по формулам

$$x_k = \frac{A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n}{\Delta} \quad (k = 1, 2, \ldots, n). \tag{5}$$

Заметим, что в силу сказанного в [**3**] числитель написанного выражения представляет собою определитель, который получается из определителя Δ заменой элементов k-го столбца, т. е. коэффициентов a_{ik} при x_k, свободными членами b_i. Мы имеем таким образом следующую теорему:

Теорема Крамера. *Если определитель* Δ *системы* (1) *отличен от нуля, то эта система имеет одно определенное решение, выражаемое формулами* (5). *Согласно этим формулам каждое из неизвестных выражается частным двух определителей, причем в знаменателе стоит определитель системы, а в числителе — определитель, который из него получается заменой коэффициентов при определяемом неизвестном соответствующими свободными членами.*

При большом числе уравнений пользование теоремой Крамера неудобно, и существуют другие, приближенные, практические способы решения систем многих уравнений со многими неизвестными, на чем мы останавливаться не будем.

9. Общий случай систем уравнений. Рассмотрим общий случай m уравнений с n неизвестными:

$$\left.\begin{aligned}
X_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1k}x_k + a_{1,k+1}x_{k+1} + \ldots + a_{1n}x_n = b_1\\
X_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2k}x_k + a_{2,k+1}x_{k+1} + \ldots + a_{2n}x_n = b_2\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
X_k &= a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kk}x_k + a_{k,k+1}x_{k+1} + \ldots + a_{kn}x_n = b_k\\
X_{k+1} &= a_{k+1,1}x_1 + a_{k+1,2}x_2 + \ldots + a_{k+1,k}x_k +\\
&\qquad + a_{k+1,k+1}x_{k+1} + \ldots + a_{k+1,n}x_n = b_{k+1}\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
X_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mk}x_k + a_{m,k+1}x_{k+1} + \ldots + a_{mn}x_n = b_m.
\end{aligned}\right\} \quad (6)$$

Для краткости в дальнейших выкладках мы обозначили через X_s всю левую часть уравнения с номером s. Коэффициенты a_{ik} этой системы образуют прямоугольную таблицу, и пусть k — ее ранг. Переставляя строки и столбцы, т. е. переменяя нумерацию уравнений и неизвестных, мы можем достигнуть того, чтобы некоторый определитель порядка k, входящий в таблицу и отличный от нуля, стоял в левом верхнем углу. Назовем его *главным определителем системы*. Он будет иметь вид:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k}\\ a_{21} & a_{22} & \ldots & a_{2k}\\ \cdot & \cdot & \cdot & \cdot\\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}. \quad (7)$$

Составим $(m-k)$ определителей порядка $(k+1)$, которые называются *характеристическими определителями системы* и которые получаются из главного определителя, если к нему добавить одну строку, состоящую из коэффициентов уравнений с номером, бо́льшим чем k, и один столбец, состоящий из свободных членов. Уточняя сказанное, определим эти характеристические определители следующими формулами:

$$\Delta_{k+s} = \begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & b_1 \\ a_{21} & a_{22} & \dots a_{2k} & b_2 \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots a_{kk} & b_k \\ a_{k+s,\,1} & a_{k+s,\,2} & \dots a_{k+s,\,k} & b_{k+s} \end{vmatrix} \tag{8}$$

$$(k+s = k+1,\ k+2,\ \dots,\ m).$$

Если $k = m$, т. е. ранг равен числу уравнений, то характеристических определителей вовсе не будет. Рассмотрим наряду с характеристическими определителями другие определители, которые получаются из них заменой в последнем столбце свободных членов левыми частями уравнений:

$$\begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & X_1 \\ a_{21} & a_{22} & \dots a_{2k} & X_2 \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots a_{kk} & X_k \\ a_{k+s,\,1} & a_{k+s,\,2} & \dots a_{k+s,\,k} & X_{k+s} \end{vmatrix}. \tag{9}$$

Эти последние определители, кроме заданных коэффициентов a_{ik}, содержат x_j. Но нетрудно показать, что определители (9) тождественно равны нулю. Действительно, элементы последнего столбца этих определителей в силу

$$X_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$$

состоят из n слагаемых и, следовательно, всякий определитель (9) может быть представлен, согласно свойству IV из [**3**], в виде суммы членов следующей формы:

$$\begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & a_{1j} \\ a_{21} & a_{22} & \dots a_{2k} & a_{2j} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots a_{kk} & a_{kj} \\ a_{k+s,\,1} & a_{k+s,\,2} & \dots a_{k+s,\,k} & a_{k+s,\,j} \end{vmatrix} \cdot x_j.$$

Легко убедиться, что определитель, стоящий множителем при x_j, равен нулю. Действительно, если $j \leqslant k$, то последний столбец этого определителя совпадает с одним из предыдущих. Если же $j > k$, то упомянутый определитель есть определитель порядка $(k+1)$, входящий в таблицу (5), и, следовательно, он равен нулю, так как по предположению ранг этой таблицы равен k. Вычитая определители (9), равные тождественно нулю, из характеристических определителей и пользуясь свойством IV определителей, мы можем представить характеристические определители в следующем виде:

$$\Delta_{k+s}=\begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & b_1-X_1 \\ a_{21} & a_{22} & \dots a_{2k} & b_2-X_2 \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots a_{kk} & b_k-X_k \\ a_{k+s,\,1} & a_{k+s,\,2} & \dots a_{k+s,\,k} & b_{k+s}-X_{k+s} \end{vmatrix} \tag{10}$$

$$(k+s=k+1,\ k+2,\ \dots,\ m),$$

причем в этой форме они лишь по виду зависят от букв x_j. Предположим теперь, что наша система (6) имеет некоторое решение:

$$x_1=x_1^{(0)},\ x_2=x_2^{(0)},\ \dots,\ x_n=x_n^{(0)}.$$

Подставляя $x_j=x_j^{(0)}$ в последний столбец определителей (10), мы будем иметь в этом последнем столбце нули, т. е. все характеристические определители должны равняться нулю.

Теорема I. *Для того чтобы система* (6) *имела хоть одно решение, необходимо, чтобы все характеристические определители* (8) *были равны нулю.*

Докажем теперь достаточность этого условия и дадим способ нахождения всех решений системы. Итак, предположим, что все характеристические определители равны нулю. Возьмем их в форме (10) и разложим по элементам последнего столбца. Нетрудно видеть, что алгебраическое дополнение элемента $(b_{k+s}-X_{k+s})$ будет равно главному определителю Δ, отличному от нуля, и мы можем записать наше условие, что все характеристические определители равны нулю, в виде:

$$a_1^{(k+s)}(b_1-X_1)+a_2^{(k+s)}(b_2-X_2)+\dots+a_k^{(k+s)}(b_k-X_k)+$$
$$+\Delta(b_{k+s}-X_{k+s})=0 \quad (k+s=k+1,\ k+2,\ \dots,\ m), \tag{11}$$

где $a_p^{(q)}$ — численные коэффициенты, не представляющие для нас никакого интереса.

Положим теперь, что мы имеем какое-нибудь решение первых k уравнений системы, и подставим мысленно это решение вместо x_j в тождества (11). При этом все разности

$$b_1-X_1,\ b_2-X_2,\ \dots,\ b_k-X_k$$

обратятся в нуль, и мы получим после указанной подстановки

$$\Delta\cdot(b_{k+s}-X_{k+s})=0$$

или в силу $\Delta\neq 0$:

$$b_{k+s}-X_{k+s}=0\ \ (k+s=k+1,\ k+2,\ \dots,\ m),$$

т. е. оказывается, что если все характеристические определители равны нулю, то всякое решение первых k уравнений системы будет удовлетворять и всем следующим уравнениям, и нам остается в этом случае решить только первые k уравнений.

Перенесем в этих уравнениях все неизвестные с номером, бо́льшим k, в правую часть, после чего эти уравнения примут вид:

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1k}x_k &= b_1-a_{1,\,k+1}x_{k+1}-\ldots-a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2k}x_k &= b_2-a_{2,\,k+1}x_{k+1}-\ldots-a_{2n}x_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k1}x_1+a_{k2}x_2+\ldots+a_{kk}x_k &= b_k-a_{k,\,k+1}x_{k+1}-\ldots-a_{kn}x_n. \end{aligned}\right\} \tag{12}$$

Будем рассматривать эти уравнения как систему для определения $x_1, x_2, \ldots, x_k$. Ее определитель Δ уже отличен от нуля, и мы можем получить для нее одно определенное решение согласно формуле Крамера. Заметим только, что свободные члены последней системы содержат буквы $x_{k+1}, \ldots, x_n$, которым можно придавать любые значения. Из формул Крамера непосредственно вытекает, что решение системы (12) будет иметь вид:

$$x_j=\alpha_j+\beta_{k+1}^{(j)}x_{k+1}+\ldots+\beta_n^{(j)}x_n \quad (j=1,\ 2,\ \ldots,\ k), \tag{13}$$

где α_s и $\beta_p^{(q)}$ — некоторые численные коэффициенты и $x_{k+1}, \ldots, x_n$ остаются произвольными. Из предыдущего вытекает, что эти формулы и дают самое общее решение системы (6) при сделанном предположении о равенстве нулю всех характеристических определителей.

Теорема II. *Если все характеристические определители системы равны нулю, то достаточно решить лишь те уравнения системы, которые содержат главный определитель относительно тех неизвестных, коэффициенты которых и составляют этот главный определитель. Это решение может быть произведено по формулам Крамера и дает выражение для k неизвестных (где k — ранг таблицы коэффициентов) в виде линейных функций* (13) *остальных* $(n-k)$ *неизвестных, значения которых остаются совершенно произвольными.* Таким образом получаются все решения системы (6).

Сравнивая теоремы I и II, приходим к выводу:

Теорема III. *Необходимым и достаточным условием существования решений системы* (6) *является равенство нулю всех характеристических определителей этой системы.*

Заметим, что если $k=n$, т. е. если ранг равен числу неизвестных, то формулы (13) вовсе не содержат в правых частях x_j, и все неизвестные от x_1 до x_n вполне определяются.

Теорема IV. *Для того чтобы система имела одно определенное решение, необходимо и достаточно, чтобы все характеристические определители были равны нулю и чтобы ранг таблицы ее коэффициентов был равен числу неизвестных.*

Заметим, что все предыдущие рассуждения годятся, очевидно, и для того случая, когда число уравнений равно числу неизвестных, т. е. когда $m = n$.

Пример. Рассмотрим систему четырех уравнений с тремя неизвестными

$$\begin{aligned} x - 3y - 2z &= -1 \\ 2x + y - 4z &= 3 \\ x + 4y - 2z &= 4 \\ 5x + 6y - 10z &= 10. \end{aligned}$$

Напишем таблицу ее коэффициентов:

$$\left\| \begin{matrix} 1 & -3 & -2 \\ 2 & 1 & -4 \\ 1 & 4 & -2 \\ 5 & 6 & -10 \end{matrix} \right\|$$

Нетрудно убедиться, что все определители третьего порядка, входящие в эту таблицу, равны нулю, и что определитель второго порядка, стоящий в левом верхнем углу, отличен от нуля. Таким образом его можно принять за главный определитель, и ранг системы равен двум. Составляем характеристические определители. Их будет в данном случае два:

$$\Delta_3 = \begin{vmatrix} 1 & -3 & -1 \\ 2 & 1 & 3 \\ 1 & 4 & 4 \end{vmatrix} = 0; \qquad \Delta_4 = \begin{vmatrix} 1 & -3 & -1 \\ 2 & 1 & 3 \\ 5 & 6 & 10 \end{vmatrix} = 0.$$

Оба они равны нулю, и, следовательно, данная система совместна. Поэтому достаточно решить первые два уравнения относительно x и y, перенося z направо:

$$\begin{aligned} x - 3y &= 2z - 1 \\ 2x + y &= 4z + 3. \end{aligned}$$

Решение получится в виде:

$$x = \frac{\begin{vmatrix} 2z-1 & -3 \\ 4z+3 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix}} = 2z + \frac{8}{7}, \qquad y = \frac{\begin{vmatrix} 1 & 2z-1 \\ 2 & 4z+3 \end{vmatrix}}{\begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix}} = \frac{5}{7},$$

причем z — произвольно.

10. Однородные системы. Система называется однородной, если в ней все свободные члены b_i равны нулю. Если такая однородная система имеет характеристические определители, то их последний столбец состоит из нулей, и они все равны нулю. Совершенно очевидно, что всякая однородная система имеет решение

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

которое в дальнейшем мы будем называть нулевым. Для однородной системы основным является вопрос о том, имеет ли она решения, отличные от нулевого, и если имеет, то какова будет совокупность всех таких решений. Рассмотрим сначала тот случай, когда число уравнений равно числу неизвестных. Система будет иметь вид:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0 \end{array}\right\}. \tag{14}$$

Если определитель этой системы отличен от нуля, то, согласно теореме Крамера, эта система имеет одно определенное решение, а именно в данном случае нулевое решение. Если же этот определитель равен нулю, то ранг k таблицы коэффициентов будет меньше числа неизвестных n и, следовательно, значения $(n-k)$ неизвестных останутся совершенно произвольными, и мы будем иметь бесчисленное множество решений, отличных от нулевого. Мы приходим таким образом к следующей основной теореме:

Теорема I. *Для того чтобы система* (14) *имела решение, отличное от нулевого, необходимо и достаточно, чтобы ее определитель равнялся нулю.*

Проведем параллель тех результатов, которые мы получили для неоднородной системы (1) и однородной системы (14). Если определитель системы отличен от нуля, то неоднородная система (1) имеет одно определенное решение, и однородная система — только нулевое решение. Если же определитель системы равен нулю, то однородная система (14) имеет решения, отличные от нулевого, но при этом условии неоднородная система (1), вообще говоря, вовсе решения не имеет, ибо для того, чтобы она имела решение, необходимо, чтобы свободные ее члены были выбраны так, чтобы они обращали в нуль все характеристические определители. Приведенный параллелизм результатов будет играть в дальнейшем существенную роль. В вопросах физики однородные системы встретятся при рассмотрении собственных колебаний, а неоднородные — при рассмотрении вынужденных колебаний, и указанный выше случай равенства нулю определителя будет характеризовать для однородной системы наличие собственных колебаний, а для неоднородной системы — явление резонанса.

Переходим теперь к более подробному рассмотрению решений системы (14), когда ее основной определитель равен нулю. Пусть k есть ранг таблицы ее коэффициентов, причем, очевидно, $k < n$. Согласно доказанной в предыдущем номере теореме, мы должны взять те k уравнений, которые содержат главный определитель, и решить их относительно k неизвестных. Положим, не ограничивая общности, что эти неизвестные будут $x_1, \ldots, x_k$. Решения получатся в виде:

$$x_j = \beta_{k+1}^{(j)} x_{k+1} + \ldots + \beta_n^{(j)} x_n \qquad (j = 1, 2, \ldots, k), \tag{15}$$

где $\beta_p^{(q)}$ — определенные численные коэффициенты и $x_{k+1}, \ldots, x_n$ могут принимать произвольные значения.

Отметим одно общее свойство решения системы (14), непосредственно вытекающее из линейности и однородности этой системы, и которое может быть названо принципом наложения решений, а именно — если мы имеем несколько решений системы:

$$x_s = x_s^{(1)};\ x_s = x_s^{(2)};\ x_s = x_s^{(3)};\ \ldots;\ x_s = x_s^{(l)} \quad (s = 1, 2, \ldots, n), \tag{16}$$

то, умножая их на произвольные постоянные и складывая, мы получим также решение системы

$$x_s = C_1 x_s^{(1)} + C_2 x_s^{(2)} + C_3 x_s^{(3)} + \ldots + C_l x_s^{(l)} \quad (s = 1, 2, \ldots, n).$$

Поступая аналогично тому, как это мы делали для линейных дифференциальных уравнений [II, 27], назовем решения (16) линейно-независимыми, если не существует никаких значений постоянных C_i, среди которых есть отличные от нуля, таких, что при всяком s имеют место равенства:

$$\sum_{i=1}^{l} C_i x_s^{(i)} = 0.$$

Нетрудно построить $(n-k)$ линейно-независимых решений системы таких, что, умножая их на произвольные постоянные и складывая, получим все решения системы. Действительно, обратимся к формулам (15), дающим общее решение системы, и построим на основе этих формул решения следующим образом: в первом решении положим $x_{k+1} = 1$, а все остальные x_{k+s} равными нулю; во втором решении положим $x_{k+2} = 1$, а все остальные x_{k+s} равными нулю и т. д. и, наконец, в последнем $(n-k)$-м решении положим $x_n = 1$ и все остальные x_{k+s} равными нулю. Нетрудно видеть, что построенные решения линейно-независимы, так как каждое из них содержит одно из неизвестных равным единице, которое в остальных решениях равно нулю. Обозначим полученные решения следующим образом:

$$x_s = x_s^{(k+1)};\ x_s = x_s^{(k+2)};\ \ldots;\ x_s = x_s^{(m)} \qquad (s = 1, 2, \ldots, n).$$

Возьмем теперь какое-нибудь решение системы (14). Оно получается по формулам (15) при некоторых частных значениях:

$$x_{k+1} = \gamma_{k+1};\quad x_{k+2} = \gamma_{k+2};\ \ldots;\ x_n = \gamma_n.$$

Непосредственно очевидно, что это решение есть линейная комбинация построенных нами решений, а именно:

$$x_s = \gamma_{k+1} x_s^{(k+1)} + \gamma_{k+2} x_s^{(k+2)} + \ldots + \gamma_n x_s^{(n)} \qquad (s = 1, 2, \ldots, n).$$

Мы вернемся к исследованию решений однородной системы (14) и покажем, что при любом выборе линейно-независимых решений их общее число будет равно $(n-k)$.

Обратимся к общему случаю m однородных уравнений с n неизвестными. Если $m < n$, то ранг k, который не может превышать m, также будет меньше n, и $(n-k)$ неизвестных останутся произвольными, т. е. *если число однородных уравнений меньше числа неизвестных, то система имеет решения, отличные от нулевого.*

Вообще $k \leqslant n$ и при $k = n$ система будет иметь только нулевое решение.

11. Линейные формы. С вопросом о решении систем уравнений первой степени тесно связано исследование систем линейных форм. Под линейной формой переменных $x_1, x_2, \ldots, x_n$ мы подразумеваем линейную однородную функцию этих переменных. Пусть имеются m таких линейных форм

$$y_s = a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n \qquad (s = 1, 2, \ldots, m). \tag{17}$$

Написанные формы называются *линейно-зависимыми*, если существуют постоянные $\alpha_1, \alpha_2, \ldots, \alpha_m$, среди которых есть отличные от нуля, такие, что имеет место, тождественно относительно переменных $x_1, x_2, \ldots, x_n$, соотношение:

$$\alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_m y_m = 0.$$

Если таких постоянных нет, то формы (17) называются *линейно-независимыми.* Мы должны в написанном тождестве приравнять нулю коэффициенты при всех переменных x_i. Таким образом написанное тождество равносильно следующей системе n равенств:

$$\begin{aligned} \alpha_1 a_{11} + \alpha_2 a_{21} + \ldots + \alpha_m a_{m1} &= 0, \\ \alpha_1 a_{12} + \alpha_2 a_{22} + \ldots + \alpha_m a_{m2} &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ \alpha_1 a_{1n} + \alpha_2 a_{2n} + \ldots + \alpha_m a_{mn} &= 0. \end{aligned}$$

Формы y_s линейно-независимы тогда и только тогда, когда эта система однородных уравнений относительно $\alpha_1, \alpha_2, \ldots, \alpha_m$ имеет *только нулевое решение.* Полученные выше результаты приводят к ряду следствий, касающихся линейной зависимости форм. Если $m > n$, то написанная однородная система имеет наверно не нулевые решения, и формы линейно-зависимы. Для того чтобы формы были независимыми, необходимо и достаточно, чтобы ранг k таблицы коэффициентов a_{pq} был равен числу форм m. *Если* $m = n$, т. е. *если число форм равно числу переменных, то для линейной независимости необходимо и достаточно, чтобы определитель из всей квадратной* ($m = n$) *таблицы* a_{pq} *был отличен от нуля.* В этом случае говорят, что имеется *полная система линейно-независимых форм.* Если $m \leqslant n$ и формы (17) линейно-независимы (т. е. если $k = m$), то при любых значениях y_s система уравнений (17) разрешима относительно тех переменных x_l, коэффициенты при которых образуют определитель порядка k, отличный от нуля, т. е. *линейно-независимые формы могут принимать любую совокупность значений* y_s. Если $k = m = n$, то при заданных y_s определятся все переменные x_l.

Положим теперь, что $k < m$. Соответственно нумеруя формы y_s и переменные x_l, можем считать, что в таблице a_{pq} в левом верхнем углу стоит определитель порядка k, отличный от нуля. При этом первые k форм: $y_1, y_2, \ldots, y_k$ линейно-независимы, а каждая из остальных форм y_{k+t} выражается линейно через первые k форм. Действительно, для первых k форм ранг таблицы коэффициентов, равный k, совпадает с числом форм, а потому формы линейно независимы. Если мы возьмем $(k+1)$ форму $y_1, y_2, \ldots, y_k, y_{k+t}$, то ранг их таблицы коэффициентов, равный тоже k, меньше числа форм, так что формы линейно зависимы, т. е. существуют постоянные β_s такие, что $\beta_1 y_1 + \ldots + \beta_k y_k + \beta_{k+1} y_{k+1} = 0$. В этом соотношении коэффициент β_{k+1} должен быть отличным от нуля, так как иначе формы $y_1, y_2, \ldots, y_k$ оказались бы линейно-зависимыми, и мы получаем линейное выражение y_{k+1} через первые k форм:

$$y_{k+1} = -\frac{\beta_1}{\beta_{k+1}} y_1 - \frac{\beta_2}{\beta_{k+1}} y_2 - \ldots - \frac{\beta_k}{\beta_{k+1}} y_k.$$

Число k мы назовем *рангом системы форм* (17). Это число, равное рангу таблицы коэффициентов, с другой стороны — равно наибольшему числу линейно-независимых форм из системы форм (17).

Положим, мы имеем k линейно-независимых форм: $y_1, y_2, \ldots, y_k$, где $k < n$. Можно считать, что определитель порядка k, стоящий в верхнем левом углу таблицы a_{pq}, отличен от нуля. Нетрудно видеть, что можно дополнить систему этих k форм до полной системы n линейно-независимых форм. Действительно, достаточно для этого положить, например:

$$y_{k+1} = x_{k+1}; \ \ldots; \quad y_n = x_n.$$

Определитель полученных k форм будет:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} & a_{1\,k+1} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2k} & a_{2\,k+1} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} & a_{k\,k+1} & \ldots & a_{kn} \\ 0 & 0 & \ldots & 0 & 1 \;\; 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 \;\; 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 \;\; 0 & \ldots & 1 \end{vmatrix}.$$

Разлагая этот определитель, начиная с элементов последней строки, потом по элементам предпоследней строки и т. д., видим, что его величина равна определителю порядка k, стоящему в левом верхнем углу, т. е. отлична от нуля. Следовательно, формы $y_1, y_2, \ldots, y_n$, действительно, линейно независимы. Итак, *всякую систему линейно-независимых форм можно дополнить до полной системы линейно-независимых форм.*

12. n-мерное векторное пространство. Мы дадим полученным выше результатам геометрическую формулировку, которой будем пользоваться в дальнейшем. Для этого введем понятия о векторе в n-мерном пространстве, а именно назовем таким вектором совокупность n чисел (комплексных), идущих в определенном порядке. Таким образом *всякий такой вектор $\boldsymbol{x}$ характеризуется последовательностью n комплексных чисел, которые называются составляющими этого вектора: $\boldsymbol{x}(x_1, x_2, \ldots, x_n)$. Совокупность всех таких векторов образует n-мерное векторное пространство R_n.*

Два вектора считаются равными тогда и только тогда, когда все их составляющие одинаковы, т. е. *если имеются два вектора $\boldsymbol{u}(u_1, u_2, \ldots, u_n)$* и $\boldsymbol{v}(v_1, v_2, \ldots, v_n)$, то векторное равенство $\boldsymbol{u}=\boldsymbol{v}$ равносильно следующим n скалярным равенствам: $u_1=v_1$; $u_2=v_2$; $\ldots$; $u_n=v_n$. Определим операцию — умножение вектора на число и сложение векторов. По определению умножение вектора на число сводится к умножению всех составляющих вектора на это число, т. е. если вектор $\boldsymbol{x}$ имеет составляющие $(x_1, x_2, \ldots, x_n)$, то вектор $k\boldsymbol{x}$ имеет составляющие $(kx_1, kx_2, \ldots, kx_n)$. Сложение векторов сводится к сложению составляющих, т. е., если имеются векторы $\boldsymbol{x}(x_1, x_2, \ldots, x_n)$ и $\boldsymbol{y}(y_1, y_2, \ldots, y_n)$, то, по определению, сумма $\boldsymbol{x}+\boldsymbol{y}$ имеет составляющие $(x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$. Нулевым вектором называется вектор $(0, 0, \ldots, 0)$, у которого все составляющие равны нулю. Обозначим этот вектор символом θ. Мы имеем, очевидно, $\theta=0\boldsymbol{x}$, где $\boldsymbol{x}$ — любой вектор, и $\boldsymbol{x}+\theta=\boldsymbol{x}$.

Вычитание векторов определяется так: вектор $\boldsymbol{x}-\boldsymbol{y}$ имеет составляющие $(x_1-y_1, x_2-y_2, \ldots, x_n-y_n)$. Мы имеем, очевидно, $\boldsymbol{x}-\boldsymbol{x}=\theta$ и $\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{x}+(-1)\boldsymbol{y}$, т. е. вычитание вектора $\boldsymbol{y}$ равносильно сложению с вектором $\boldsymbol{y}$, умноженным на число (-1). В дальнейшем нам придется часто писать векторные равенства. Всякое такое равенство равносильно n скалярным равенствам, выражающим тот факт, что соответствующие составляющие обеих частей равенства одинаковы. В дальнейшем мы не будем пользоваться символом θ для нулевого вектора, но надо помнить, что если в векторном равенстве с одной стороны стоит нуль, то этот нуль надо понимать как нулевой вектор. Из данных выше определений непосредственно вытекают обычные свойства сложения и умножения:

$$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}; \quad \boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z};$$

$$(k_1+k_2)\boldsymbol{x}=k_1\boldsymbol{x}+k_2\boldsymbol{x}; \qquad k(\boldsymbol{x}+\boldsymbol{y})=k\boldsymbol{x}+k\boldsymbol{y}; \qquad k_1(k_2\boldsymbol{x})=(k_1k_2)\boldsymbol{x}.$$

В сумме векторов с любым числом слагаемых можно, таким образом, переставлять слагаемые и заключать их в группы. Из равенства $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{z}$ следует $\boldsymbol{x}=\boldsymbol{z}-\boldsymbol{y}$ и $\boldsymbol{y}=\boldsymbol{z}-\boldsymbol{x}$ и, наоборот, из $\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{z}$ следует $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$.

Введем теперь понятие о линейной зависимости и независимости векторов. Векторы

$$\boldsymbol{x}^{(1)}, \ \boldsymbol{x}^{(2)}, \ \ldots, \ \boldsymbol{x}^{(l)} \tag{18}$$

будем называть *линейно-зависимыми*, если существуют такие постоянные $C_1, \ldots, C_l$, не все равные нулю, что

$$C_1\boldsymbol{x}^{(1)}+C_2\boldsymbol{x}^{(2)}+\ldots+C_l\boldsymbol{x}^{(l)}=0. \tag{19}$$

Если таких постоянных не существует, то будем называть векторы (18) *линейно-независимыми*. Обозначим составляющие векторов $\boldsymbol{x}^{(j)}$ через $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)})$. Условие (19), очевидно, равносильно системе n уравнений с неизвестными $C_1, C_2, \ldots, C_l$:

$$\left.\begin{array}{l} x_1^{(1)}C_1+x_1^{(2)}C_2+\ldots+x_1^{(l)}C_l=0, \\ x_2^{(1)}C_1+x_2^{(2)}C_2+\ldots+x_2^{(l)}C_l=0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_n^{(1)}C_1+x_n^{(2)}C_2+\ldots+x_n^{(l)}C_l=0. \end{array}\right\} \tag{20}$$

Пользуясь полученными выше результатами для однородной системы, нетрудно вывести из них ряд следствий и придать этим результатам геометрическую формулировку. Положим сначала, что $l>n$, т. е. что число векторов больше числа измерений пространства. При этом в однородной системе (20) число уравнений будет меньше числа неизвестных, и эта система, как мы знаем, наверно будет иметь для неизвестных C_j решения, отличные от нулевого, т. е. наши векторы будут наверно линейно-зависимыми. Иначе говоря, *число линейно-независимых векторов, самое большее, равно числу измерений пространства.* Рассмотрим теперь случай $l=n$. При этом система (20) будет содержать одинаковое число уравнений и неизвестных и будет иметь решения, отличные от нулевого, тогда и только тогда, когда ее определитель равен нулю, т. е. если в n-мерном пространстве мы имеем n векторов, и составим из n^2 составляющих этих векторов определитель, помещая, например, составляющие каждого вектора в определенном столбце и считая номер строки совпадающим с номером составляющей, то для линейной независимости векторов необходимо и достаточно, чтобы этот определитель был отличен от нуля. Величина этого определителя имеет аналог с объемом параллелепипеда в вещественном трехмерном пространстве.

Мы можем в любом определителе $|b_{ik}|$ порядка n рассматривать элементы каждого столбца $(b_{1k}, b_{2k}, \ldots, b_{nk})$ как составляющие некоторого вектора $\boldsymbol{b}^{(k)}$, и при этом величина определителя будет функцией n векторов $\boldsymbol{b}^{(1)}, \ldots, \boldsymbol{b}^{(n)}$. Равенство нулю этого определителя будет равносильно тому факту, что эти векторы линейно-зависимы.

Обозначим величину определителя как функцию векторов $\boldsymbol{b}^{(k)}$ через

$$|b_{ik}| = \Delta(\boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)}, \ldots, \boldsymbol{b}^{(n)}).$$

Вспоминая, что при перестановке двух столбцов величина определителя меняет знак, мы можем утверждать, что функция Δ изменит лишь знак, если поменять местами ее два аргумента. Такая функция называется обычно *антисимметрической.* Нетрудно видеть, например, что определитель Вандермонда D_n, который мы рассматривали в [5], есть также антисимметрическая функция своих аргументов $x_1, \ldots, x_n$.

Вернемся к рассмотрению системы (20) и к вопросу о линейной независимости векторов $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(l)}$, предполагая $l \leqslant n$. Обозначим через k ранг таблицы, образованной составляющими $x_p^{(q)}$. Если $k=l$, то, как мы видели, система имеет только нулевое решение, т. е. векторы линейно-независимы. Если же $k<l$, то система будет наверно иметь решение, отличное от нулевого, т. е. *для линейной независимости векторов необходимо и достаточно, чтобы их число равнялось рангу таблицы, образованной их составляющими.* Положим теперь, что $k<l$, т. е. что векторы линейно-зависимы. Выделим из них те k векторов (возможно, что это можно будет сделать несколькими способами), составляющие которых содержат определитель порядка k, отличный от нуля. Согласно доказанному выше, эти векторы линейно-независимы. Нетрудно видеть, что каждый из остальных векторов может быть выражен линейно через выделенные векторы. Действительно, пусть $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(k)}$ — линейно-независимые векторы. Присоединяя к ним какой-нибудь вектор $\boldsymbol{x}^{(k+s)}$, получим $(k+1)$ векторов, которые будут линейно зависимы,

так как ранг k таблицы их составляющих меньше их числа $l=k+1$. Итак, будут существовать постоянные C_i ($i=1, 2, \ldots, k, k+s$), среди которых есть отличные от нуля, такие, что

$$C_1x^{(1)}+C_2x^{(2)}+\ldots+C_kx^{(k)}+C_{k+s}x^{(k+s)}=0.$$

При этом наверно $C_{k+s}\neq 0$, ибо, в противном случае, векторы $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$ оказались бы линейно-зависимы. Из написанного равенства следует:

$$x^{(k+s)}=-\frac{C_1}{C_{k+s}}x^{(1)}-\frac{C_2}{C_{k+s}}x^{(2)}-\ldots-\frac{C_k}{C_{k+s}}x^{(k)},$$

т. е. $x^{(k+s)}$ выражается линейно через $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$. Пусть $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ — n каких-нибудь линейно-независимых векторов. В качестве примера таких векторов мы можем указать векторы:

$$(1, 0, 0, \ldots, 0);\ (0, 1, 0, \ldots, 0);\ \ldots;\ (0, 0, 0, \ldots, 1). \tag{21}$$

Если мы возьмем какой угодно вектор x, то $(n+1)$ векторов $x^{(1)}, x^{(2)}, \ldots, x^{(n)}, x$, как мы видели выше, наверно линейно-зависимы:

$$C_1x^{(1)}+C_2x^{(2)}+\ldots+C_nx^{(n)}+Cx=0,$$

причем постоянная C наверно отлична от нуля, ибо в противном случае векторы $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ оказались бы линейно-зависимы. Из предыдущего следует, что *любой вектор x выражается через n линейно-независимых векторов:*

$$x=\alpha_1x^{(1)}+\alpha_2x^{(2)}+\ldots+\alpha_nx^{(n)} \qquad \left(\alpha_s=-\frac{C_s}{C}\right). \tag{22}$$

Нетрудно видеть, что *существует лишь одно определенное выражение x через $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$*. Действительно, если бы, кроме написанного выражения, существовало еще одно выражение:

$$x=\beta_1x^{(1)}+\beta_2x^{(2)}+\ldots+\beta_nx^{(n)},$$

в котором есть коэффициенты β_s, отличные от соответствующих α_s, то, вычитая два написанных соотношения почленно, мы получили бы:

$$(\alpha_1-\beta_1)x^{(1)}+(\alpha_2-\beta_2)x^{(2)}+\ldots+(\alpha_n-\beta_n)x^{(n)}=0,$$

т. е. векторы $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ оказались бы линейно-зависимыми. Если за векторы $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ принять векторы (21), то числа α_s формулы (22) будут, очевидно, совпадать с составляющими x_s вектора $x(x_1, x_2, \ldots, x_n)$. В общем случае их можно назвать также *составляющими вектора x, если $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ приняты за орты.* Придавая числам α_s всевозможные комплексные значения, получим все векторы нашего n-мерного пространства. Положим теперь, что мы имеем k линейно-независимых векторов

$$x^{(1)}, x^{(2)}, \ldots, x^{(k)}, \tag{23}$$

причем $k<n$. Говорят, что совокупность векторов, получаемых по формуле

$$y=C_1x^{(1)}+C_2x^{(2)}+\ldots+C_kx^{(k)}, \tag{23$_1$}$$

где C_s — произвольные постоянные, образует некоторое *подпространство* L_k измерения k. Как и выше, можно показать, что всякий вектор, принадлежащий L_k, единственным образом выражается через $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(k)}$. Говорят иначе, что *векторы* (23) *образуют подпространство* L_k.

Заметим, что если какой-нибудь вектор z принадлежит L_k, т. е. выражается формулой вида (23_1), то вектор cz, где c — любая постоянная, также выражается, очевидно, формулой вида (23_1), т. е. также принадлежит L_k. Точно так же, если $\boldsymbol{z}^{(1)}$ и $\boldsymbol{z}^{(2)}$ принадлежат L_k, то их сумма $\boldsymbol{z}^{(1)}+\boldsymbol{z}^{(2)}$ также принадлежит L_k. Отсюда непосредственно следует и более общее свойство: *если некоторые векторы $\boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}, \ldots, \boldsymbol{z}^{(p)}$ принадлежат L_k, то их любая линейная комбинация $\gamma_1\boldsymbol{z}^{(1)}+\gamma_2\boldsymbol{z}^{(2)}+\ldots+\gamma_p\boldsymbol{z}^{(p)}$ также принадлежит L_k.*

Возьмем m каких-нибудь векторов, принадлежащих L_k:

$$y^{(s)}=C_1^{(s)}\boldsymbol{x}^{(1)}+C_2^{(s)}\boldsymbol{x}^{(2)}+\ldots+C_k^{(s)}\boldsymbol{x}^{(k)} \qquad (s=1,\ 2,\ \ldots,\ m). \quad (24)$$

В силу линейной независимости векторов (23) соотношение вида

$$\alpha_1\boldsymbol{y}^{(1)}+\alpha_2\boldsymbol{y}^{(2)}+\ldots+\alpha_m\boldsymbol{y}^{(m)}=0$$

равносильно системе k однородных уравнений

$$\alpha_1 C_q^{(1)}+\alpha_2 C_q^{(2)}+\ldots+\alpha_m C_q^{(m)}=0 \qquad (q=1,\ 2,\ \ldots,\ k).$$

Если эта система имеет решения, отличные от нулевого, то векторы (24) линейно-зависимы. В частности, если $m>k$, то наверно имеются решения, отличные от нулевого, т. е. всякая совокупность векторов числом больше k из подпространства, образованного векторами (23), будет совокупностью линейно-зависимых векторов. Отсюда непосредственно следует, что подпространство, образованное линейно-независимыми векторами (23), не может быть образовано никакой совокупностью линейно-независимых векторов $\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(l)}$, число которых $l<k$. Действительно, согласно вышеуказанному, в таком подпространстве не может существовать больше, чем l, линейно-независимых векторов, а с другой стороны, этому подпространству должны принадлежать линейно-независимые векторы (23), число которых k больше l. *Если мы возьмем k каких-нибудь линейно-независимых векторов $\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \ldots, \boldsymbol{u}^{(k)}$, принадлежащих L_k, то они образуют, в указанном выше смысле, это же подпространство L_k.*

Действительно, в силу определения подпространства всякая линейная комбинация

$$C_1\boldsymbol{u}^{(1)}+C_2\boldsymbol{u}^{(2)}+\ldots+C_k\boldsymbol{u}^{(k)}$$

принадлежит L_k. С другой стороны, возьмем какой-нибудь вектор $\boldsymbol{y}$ из L_k. Векторы $\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(k)}, \boldsymbol{y}$, числом $(k+1)$, принадлежат L_k и, по предыдущему, должны быть линейно-зависимыми

$$\beta_1\boldsymbol{u}^{(1)}+\beta_2\boldsymbol{u}^{(2)}+\ldots+\beta_k\boldsymbol{u}^{(k)}+\gamma\boldsymbol{y}=0,$$

и, поскольку $\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(k)}$ — линейно-независимы, коэффициент γ

должен быть отличным от нуля, т. е. всякий вектор $\boldsymbol{y}$ из L_k выражается через $\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(k)}$, т. е. эти последние векторы действительно образуют L_k. Если в формулах (24) $m=k$ и определитель из коэффициентов $C_p^{(q)}$ отличен от нуля, то векторы $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k)}$ будут линейно-независимыми векторами из L_k. В общем случае нетрудно показать, что число линейно-независимых векторов, даваемых формулами (24), равно рангу таблицы $C_p^{(q)}$.

Выше мы видели, что если вектор $\boldsymbol{z}$ принадлежит некоторому подпространству L, то вектор $c\boldsymbol{z}$ при любом постоянном c также принадлежит L, и если $\boldsymbol{z}^{(1)}$ и $\boldsymbol{z}^{(2)}$ принадлежат L, то $(\boldsymbol{z}^{(1)}+\boldsymbol{z}^{(2)})$ также принадлежат L. Мы могли бы дать новое определение подпространства, а именно назвать *подпространством* такую совокупность векторов, что если $\boldsymbol{z}$ принадлежит L, то и $c\boldsymbol{z}$ принадлежит L, и если $\boldsymbol{z}^{(1)}$ и $\boldsymbol{z}^{(2)}$ принадлежат L, то $(\boldsymbol{z}^{(1)}+\boldsymbol{z}^{(2)})$ также принадлежит L. Отсюда непосредственно следует, что всякая линейная комбинация векторов, принадлежащих L, также принадлежит L. Мы только что видели, что из прежнего определения подпространства вытекают, как следствие, те свойства, которые сформулированы в новом определении. Покажем и наоборот, что из нового определения вытекает прежнее, т. е. что эти *два определения равносильны.*

Возьмем некоторый вектор $\boldsymbol{x}^{(1)}$, принадлежащий L. По определению подпространства L векторы $C_1\boldsymbol{x}^{(1)}$, при произвольном C_1, также принадлежат L. Если этими векторами исчерпывается всё L, то мы имеем L_1 в прежнем смысле. В противном случае в L входит некоторый вектор $\boldsymbol{x}^{(2)}$, линейно-независимый с $\boldsymbol{x}^{(1)}$, и векторы $C_1\boldsymbol{x}^{(1)}+C_2\boldsymbol{x}^{(2)}$, при произвольных C_1 и C_2, принадлежат L. Если этими векторами исчерпывается всё L, то L совпадает с некоторым L_2 в прежнем смысле. В противном случае в L входит некоторый вектор $\boldsymbol{x}^{(3)}$ такой, что $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$, $\boldsymbol{x}^{(3)}$ линейно-независимы. Продолжая так и дальше, мы, путем присоединения конечного числа линейно-независимых векторов $\boldsymbol{x}^{(s)}$, исчерпаем всё L, так как не существует более n линейно-независимых векторов. Общее число k этих векторов $\boldsymbol{x}^{(s)}$ дает нам измерение подпространства L. Если окажется, что $k=n$, то L совпадает со всем n-мерным пространством.

Отметим одно обстоятельство, связанное с образованием подпространства. Положим, что векторы $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(k)}$ — линейно-зависимы. При этом по-прежнему говорят, что формула (23_1) определяет некоторое подпространство L. Положим, что среди указанных векторов первые l: $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(l)}$ — линейно-независимы, а каждый из следующих векторов: $\boldsymbol{x}^{(l+1)}, \ldots, \boldsymbol{x}^{(k)}$ выражается линейно через первые l. При этом совокупность векторов, определяемых формулой (23_1), будет, очевидно, совпадать с совокупностью векторов, определяемых формулой:

$$\boldsymbol{y}=C_1\boldsymbol{x}^{(1)}+C_2\boldsymbol{x}^{(2)}+\ldots+C_l\boldsymbol{x}^{(l)},$$

т. е. подпространство L, образуемое линейно-зависимыми векторами $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$, имеет размерность $l (l < k)$.

Рассмотрим вещественное трехмерное пространство и условимся откладывать векторы от некоторой определенной точки O (начало). В данном случае $n = 3$. При $k = 1$ подпространство L_1 есть некоторая прямая, проходящая через O, а L_2 есть некоторая плоскость, проходящая через O.

13. Скалярное произведение. Условимся относительно одного обозначения. Если α — некоторое комплексное число, то символом $\bar{\alpha}$ будем обозначать число, сопряженное с α, и символом $|\alpha|$ — модуль числа α, так что $\alpha\bar{\alpha} = |\alpha|^2$. Если α вещественно, то $\bar{\alpha} = \alpha$ и $|\alpha|^2 = \alpha^2$. Введем теперь новое понятие, которое в дальнейшем будет играть большую роль.

Определение. *Скалярным произведением двух векторов*

$$x(x_1, x_2, \ldots, x_n) \quad \text{и} \quad y(y_1, y_2, \ldots, y_n)$$

называется число, равное следующей сумме:

$$\sum_{s=1}^{n} x_s \bar{y}_s.$$

Мы будем обозначать скалярное произведение символом (x, y). Имеем:

$$(x, y) = \sum_{s=1}^{n} x_s \bar{y}_s; \qquad (y, x) = \sum_{s=1}^{n} y_s \bar{x}_s,$$

откуда следует

$$(y, x) = \overline{(x, y)}.$$

Назовем *два вектора взаимно перпендикулярными или взаимно ортогональными, если их скалярное произведение равно нулю.* Поскольку число, сопряженное с нулем, есть также нуль, в условии ортогональности порядок векторов в скалярном произведении роли не играет. Нетрудно видеть, что нулевой вектор $(0, 0, \ldots, 0)$ ортогонален к любому вектору x.

Из определения скалярного произведения непосредственно вытекают следующие его свойства:

$$(\alpha x, y) = \alpha (x, y); \quad (x, \alpha y) = \bar{\alpha} (x, y),$$

где α — численный множитель. Кроме того:

$$(x + y, z) = (x, z) + (y, z); \quad (x, y + z) = (x, y) + (x, z),$$

и этот распределительный закон имеет место для любого числа слагаемых. Из него следует, между прочим:

$$(x + y, u + v) = (x, u) + (x, v) + (y, u) + (y, v).$$

Составим скалярное произведение вектора $\boldsymbol{x}(x_1, x_2, \ldots, x_n)$ на этот же вектор:

$$(\boldsymbol{x}, \boldsymbol{x}) = \sum_{s=1}^{n} x_s \bar{x}_s = \sum_{s=1}^{n} |x_s|^2.$$

Мы получаем таким образом вещественное число, которое будет положительным для вектора $\boldsymbol{x}$, отличного от нулевого вектора $(0, 0, \ldots, 0)$, и равным нулю для нулевого вектора. *Корень квадратный* (*арифметическое значение*) *из вещественного числа* $(\boldsymbol{x}, \boldsymbol{x})$ *называется нормой или длиной вектора* $\boldsymbol{x}$. Обозначая эту норму символом $\|\boldsymbol{x}\|$, можем написать:

$$\|\boldsymbol{x}\|^2 = (\boldsymbol{x}, \boldsymbol{x}) = \sum_{s=1}^{n} |x_s|^2; \quad \|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})} = \sqrt{\sum_{s=1}^{n} |x_s|^2}.$$

Равенство $\|\boldsymbol{x}\| = 0$ равносильно тому, что $\boldsymbol{x}$ есть нулевой вектор. Положим, что имеются три взаимно перпендикулярных вектора $\boldsymbol{x}$, $\boldsymbol{y}$ и $\boldsymbol{z}$, т. е.

$$(\boldsymbol{x}, \boldsymbol{y}) = 0; \ (\boldsymbol{x}, \boldsymbol{z}) = 0; \ (\boldsymbol{y}, \boldsymbol{z}) = 0.$$

Применяя распределительный закон для скалярного произведения и принимая во внимание написанные равенства, получим:

$$(\boldsymbol{x} + \boldsymbol{y} + \boldsymbol{z}, \boldsymbol{x} + \boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) + (\boldsymbol{z}, \boldsymbol{z})$$

или

$$\|\boldsymbol{x} + \boldsymbol{y} + \boldsymbol{z}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + \|\boldsymbol{z}\|^2.$$

Эта формула выражает *теорему Пифагора.* Она справедлива для любого числа слагаемых. Существенно лишь, что эти слагаемые векторы попарно ортогональны. Покажем, что *если векторы* $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(l)}$ *попарно ортогональны и среди них нет нулевого вектора, то они линейно независимы.* Действительно, положим

$$\sum_{s=1}^{l} C_s \boldsymbol{x}^{(s)} = 0$$

и покажем, что все числа C_s должны равняться нулю. Умножим обе части этого равенства скалярно на $\boldsymbol{x}^{(k)}$, где k — одно из чисел $1, 2, \ldots, l$:

$$\sum_{s=1}^{l} C_s (\boldsymbol{x}^{(s)}, \boldsymbol{x}^{(k)}) = 0.$$

В силу того, что векторы $\boldsymbol{x}^{(s)}$ попарно ортогональны, $(\boldsymbol{x}^{(s)}, \boldsymbol{x}^{(k)}) = 0$ при $s \neq k$, так что последняя формула дает: $C_k (\boldsymbol{x}^{(k)}, \boldsymbol{x}^{(k)}) = 0$, т. е. $C_k \|\boldsymbol{x}^{(k)}\|^2 = 0$, откуда, в силу $\|\boldsymbol{x}^{(k)}\|^2 > 0$, следует $C_k = 0$, и это при любом выборе k.

14. Геометрическая интерпретация однородных систем. Рассмотрим однородную систему

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0 \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0. \end{array}\right\} \tag{25}$$

Введем векторы:

$$a^{(1)}(\bar{a}_{11}, \bar{a}_{12}, \ldots, \bar{a}_{1n}); \ldots; \quad a^{(n)}(\bar{a}_{n1}, \bar{a}_{n2}, \ldots, \bar{a}_{nn}). \tag{26}$$

При этом система (25) может быть записана в следующем сжатом виде:

$$(x, a^{(1)})=0; \quad \ldots; \quad (x, a^{(n)})=0, \tag{27}$$

т. е. дело сводится к нахождению вектора x, перпендикулярного ко всем векторам $a^{(j)}$. Если определитель $|a_{ik}|$ отличен от нуля, то, очевидно, и определитель $|\bar{a}_{ik}|$, с ним сопряженный по величине, также отличен от нуля. В этом случае векторы $a^{(j)}$ линейно-независимы, и система (27) имеет только нулевое решение, т. е. не существует вектора (кроме нуля), который был бы одновременно перпендикулярен к n линейно-независимым векторам (в n-мерном пространстве).

Рассмотрим теперь другой случай, когда определитель системы (25) равен нулю. Положим, что ранг системы будет k. Если составить таблицу из сопряженных элементов, то входящие в нее определители будут по величине сопряженными с определителями, входящими в таблицу a_{ik}, и ранг сопряженной таблицы будет также очевидно k. Таким образом, по доказанному выше, среди векторов $a^{(j)}$ будет k линейно-независимых, а остальные векторы будут их линейными комбинациями. Не ограничивая общности, можем считать, что эти линейно-независимые векторы суть

$$a^{(1)}, \ldots, a^{(k)}, \tag{28}$$

а для остальных будем иметь выражения вида:

$$a^{(k+s)}=\beta_1^{(k+s)}a^{(1)}+\ldots+\beta_k^{(k+s)}a^{(k)} \quad (k+s=k+1, k+2, \ldots, n),$$

где $\beta_p^{(q)}$ — численные коэффициенты. Из последнего непосредственно следует, что если x перпендикулярен векторам (28), то тем самым он будет перпендикулярен и ко всем векторам $a^{(j)}$. Действительно:

$$(x, a^{(k+s)})=\sum_{i=1}^{k}(x, \beta_i^{(k+s)}a^{(i)})$$

и вся сумма равна нулю, так как каждое из отдельных слагаемых по условию обращается в нуль. Итак, достаточно решить первые k уравнений системы. Считая, как всегда, что определитель порядка k, отличный от нуля, стоит в левом верхнем углу, для искомого вектора $\boldsymbol{x}$ мы получим $(n-k)$ линейно-независимых решений $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n-k)}$ тем способом, который был указан в [12], и всякое решение будет представлять собою линейную комбинацию этих $(n-k)$ векторов. Можно сказать, что в данном случае векторы, определяемые формулой

$$\boldsymbol{y}=C_1\boldsymbol{a}^{(1)}+\ldots+C_k\boldsymbol{a}^{(k)},$$

где C_i — произвольные постоянные, образуют пространство L_k измерения k, которое является подпространством для всего пространства n измерений. Совершенно так же найденные векторы $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(n-k)}$ образуют некоторое подпространство M_{n-k} измерения $(n-k)$. *Подпространство M_{n-k} ортогонально подпространству L_k* в том смысле, что всякий вектор из M_{n-k} ортогонален всякому вектору из L_k (и, очевидно, наоборот). Подпространство M_{n-k} состоит из векторов, которые удовлетворяют системе (27), т. е. ортогональны $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(k)}$. Нетрудно видеть, что n векторов $\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(k)}, \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n-k)}$ — линейно-независимы. Действительно, положим, что между ними существует соотношение:

$$(c_1\boldsymbol{a}^{(1)}+\ldots+c_k\boldsymbol{a}^{(k)})+(d_1\boldsymbol{x}^{(1)}+\ldots+d_{n-k}\boldsymbol{x}^{(n-k)})=0. \qquad (29)$$

Первая скобка дает некоторый вектор $\boldsymbol{a}$ из L_k, а вторая — некоторый вектор $\boldsymbol{x}$ из M_{n-k}, и мы имеем $\boldsymbol{a}+\boldsymbol{x}=0$ или $\boldsymbol{a}=-\boldsymbol{x}$. Но векторы $\boldsymbol{a}$ и $\boldsymbol{x}$ взаимно-ортогональны, т. е. оказывается, что вектор $\boldsymbol{a}$ ортогонален сам себе, иначе говоря $(\boldsymbol{a}, \boldsymbol{a})=0$ или $\|\boldsymbol{a}\|=0$, откуда следует, что вектор $\boldsymbol{a}$ есть нулевой вектор; то же можно сказать и о векторе $\boldsymbol{x}$. Итак:

$$c_1\boldsymbol{a}^{(1)}+\ldots+c_k\boldsymbol{a}^{(k)}=0 \quad \text{и} \quad d_1\boldsymbol{x}^{(1)}+\ldots+d_{n-k}\boldsymbol{x}^{(n-k)}=0.$$

Но векторы $\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(k)}$, по условию, линейно-независимы, и, следовательно, все постоянные c_s должны быть равны нулю; то же можно утверждать и относительно d_s. Итак, в соотношении (29) все коэффициенты должны быть равны нулю, т. е. векторы $\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(k)}, \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n-k)}$ действительно линейно-независимы.

Всякий вектор $\boldsymbol{x}$ единственным образом может быть представлен в виде:

$$\boldsymbol{x}=(\gamma_1\boldsymbol{a}^{(1)}+\ldots+\gamma_k\boldsymbol{a}^{(k)})+(\delta_1\boldsymbol{x}^{(1)}+\ldots+\delta_{n-k}\boldsymbol{x}^{(n-k)}),$$

причем первая скобка дает вектор, принадлежащий L_k, а вторая — вектор, принадлежащий M_{n-k}. Векторы, входящие в состав M_{n-k}, суть, как мы уже упоминали, всевозможные решения системы (27), и таким образом при любом выборе полной системы линейно-независимых решений число таких решений будет равно $(n-k)$, т. е. числу измерений M_{n-k}. Приведенное выше исследование однородной системы приводит нас к следующему важному результату:

Если имеется подпространство L_k измерения $k\,(k<n)$, то векторы, ортогональные к этому подпространству, образуют

подпространство M_{n-k} *измерения* $(n-k)$, *и всякий вектор* $\boldsymbol{x}$ *из* R_n *может быть представлен в виде суммы* $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$, *где* $\boldsymbol{y}$ *принадлежит* L_k *и* $\boldsymbol{z}$ *принадлежит* M_{n-k}.

Покажем, что *такое представление единственно.* Пусть, кроме указанного представления, имеется еще одно: $\boldsymbol{x}=\boldsymbol{u}+\boldsymbol{v}$, где $\boldsymbol{u}$ из L_k и $\boldsymbol{v}$ — из M_{n-k}. Надо доказать, что $\boldsymbol{u}=\boldsymbol{y}$ и $\boldsymbol{v}=\boldsymbol{z}$. Мы имеем: $\boldsymbol{y}+\boldsymbol{z}=\boldsymbol{u}+\boldsymbol{v}$, откуда $\boldsymbol{y}-\boldsymbol{u}=\boldsymbol{v}-\boldsymbol{z}$. Разность $\boldsymbol{y}-\boldsymbol{u}$ принадлежит L_k, а разность $\boldsymbol{v}-\boldsymbol{z}$ принадлежит M_{n-k}, откуда следует, что вектор $\boldsymbol{y}-\boldsymbol{u}$ ортогонален сам себе, т. е. $(\boldsymbol{y}-\boldsymbol{u}, \boldsymbol{y}-\boldsymbol{u})=0$ или $\|\boldsymbol{y}-\boldsymbol{u}\|=0$, откуда $\boldsymbol{y}-\boldsymbol{u}=0$ и $\boldsymbol{y}=\boldsymbol{u}$. При этом из $\boldsymbol{y}-\boldsymbol{u}=\boldsymbol{v}-\boldsymbol{z}$ следует, что $\boldsymbol{v}=\boldsymbol{z}$. Вектор $\boldsymbol{y}$, входящий в представление $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$, называется *проекцией* $\boldsymbol{x}$ *в подпространство* L_k. В указанном представлении векторы $\boldsymbol{y}$ и $\boldsymbol{z}$ ортогональны, и теорема Пифагора дает: $\|\boldsymbol{x}\|^2=\|\boldsymbol{y}\|^2+\|\boldsymbol{z}\|^2$, откуда следует, что $\|\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|$, причем знак равенства имеет место в том и только в том случае, когда $\boldsymbol{z}$ есть нулевой вектор, т. е. когда $\boldsymbol{x}$ принадлежит L_k, так что $\boldsymbol{y}=\boldsymbol{x}$. Аналогично $\|\boldsymbol{z}\|\leqslant\|\boldsymbol{x}\|$, и знак равенства имеет место только в том случае, когда $\boldsymbol{x}$ ортогонально L_k, т. е. $\boldsymbol{z}=\boldsymbol{x}$. Подпространства L_k и M_{n-k} называются, обычно, *дополнительными ортогональными подпространствами.* Если $k=n$, то L_n есть все R_n, а M_0 сводится к одному нулевому вектору.

Положим, что мы имеем вещественное трехмерное пространство, о котором мы говорили выше, и пусть $k=2$, так что $n-k=3-2=1$. Подпространство L_2 есть некоторая плоскость P, проходящая через точку O, а M_1 есть прямая, проходящая через O и перпендикулярная P. Всякий вектор может быть единственным образом представлен как сумма двух векторов, из которых один лежит в плоскости P, а другой — на прямой K. Мы провели геометрическую интерпретацию решения однородной системы в том случае, когда число уравнений равно числу неизвестных. Совершенно так же можно провести рассуждения и в общем случае, когда число векторов $\boldsymbol{a}^{(s)}$ не обязательно равно числу n. Аналогичное замечание относится и к следующему номеру.

15. Случай неоднородной системы. Рассмотрим неоднородную систему:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2 \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=b_n. \end{array}\right\} \tag{30}$$

Ее можно толковать как задачу нахождения вектора $\boldsymbol{x}(x_1, \ldots, x_n)$ при заданных векторах (26) из системы

$$(\boldsymbol{x}, \boldsymbol{a}^{(1)})=b_1; \ldots; (\boldsymbol{x}, \boldsymbol{a}^{(n)})=b_n. \tag{31}$$

Если определитель системы отличен от нуля, то теорема Крамера дает один определенный ответ. Положим, что определитель системы равен нулю и ранг таблицы ее коэффициентов равен k, причем определитель порядка k, отличный от нуля, стоит, как всегда, в левом верхнем углу. Наряду с системой (30) напишем систему однородных уравнений, коэффициенты которых получаются из коэффициентов данной системы заменой строк столбцами и всех чисел — сопряженными. Эта новая система будет, таким образом, иметь вид:

$$\left.\begin{array}{l}\bar{a}_{11}y_1+\bar{a}_{21}y_2+\ldots+\bar{a}_{n1}y_n=0\\ \bar{a}_{12}y_1+\bar{a}_{22}y_2+\ldots+\bar{a}_{n2}y_n=0\\ \cdots\cdots\cdots\cdots\cdots\\ \bar{a}_{1n}y_1+\bar{a}_{2n}y_2+\ldots+\bar{a}_{nn}y_n=0.\end{array}\right\} \qquad (32)$$

Таблица ее коэффициентов по-прежнему имеет ранг k, и определитель k-го порядка, стоящий в левом верхнем углу, по-прежнему отличен от нуля. Назовем эту однородную систему *союзной* с системой (30). Как мы видели выше, ее общее решение есть линейная комбинация $(n-k)$ решений (векторов), которые можно получить, например, если решить по теореме Крамера первые k уравнений относительно $y_1, \ldots, y_k$, полагая остальные y_{k+s} равными нулю, кроме одного, которое надо считать равным единице. Мы придем таким путем при $y_{k+1}=1$ к системе:

$$\begin{array}{l}\bar{a}_{11}y_1+\bar{a}_{21}y_2+\ldots+\bar{a}_{k1}y_k=-\bar{a}_{k+1,1},\\ \bar{a}_{12}y_1+\bar{a}_{22}y_2+\ldots+\bar{a}_{k2}y_k=-\bar{a}_{k+1,2},\\ \cdots\cdots\cdots\cdots\cdots\\ \bar{a}_{1k}y_1+\bar{a}_{2k}y_2+\ldots+\bar{a}_{kk}y_k=-\bar{a}_{k+1,k}.\end{array}$$

Решая эту систему и переходя к сопряженным величинам, получим:

$$\bar{y}_m=-\frac{\Delta'_m}{\Delta'}\quad (m=1,\,2,\,\ldots,\,k) \qquad (33)$$

$$\bar{y}_{k+1}=1;\quad \bar{y}_{k+2}=\bar{y}_{k+3}=\ldots=\bar{y}_n=0,$$

где

$$\Delta'=\begin{vmatrix}a_{11}, & a_{21}, & \ldots, & a_{k1}\\ a_{12}, & a_{22}, & \ldots, & a_{k2}\\ \cdot & \cdot & \cdot & \cdot\\ a_{1k}, & a_{2k}, & \ldots, & a_{kk}\end{vmatrix}\neq 0,$$

и Δ'_m получается из Δ' заменой элементов m-го столбца на $a_{k+1,1}\ldots, a_{k+1,k}$. Составим условие перпендикулярности вектора $\boldsymbol{b}$ с составляющими $(b_1, \ldots, b_n)$ с вектором $\boldsymbol{y}$, полученным только что при решении системы (32)

$$(\boldsymbol{b},\boldsymbol{y})=-\sum_{m=1}^{k}\frac{\Delta'_m}{\Delta'}b_m+b_{k+1}=0$$

или

$$-\sum_{m=1}^{k} \Delta'_m b_m + \Delta' b_{k+1} = 0. \tag{34}$$

Заменяя в определителе Δ'_m строки столбцами и переставляя затем m-ю строку на последнее место при помощи $(k-m)$ перестановок последних строк, получим:

$$-\Delta'_m = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m-1,1} & a_{m-1,2} & \dots & a_{m-1,k} \\ a_{m+1,1} & a_{m+1,2} & \dots & a_{m+1,k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} \\ a_{k+1,1} & a_{k+1,2} & \dots & a_{k+1,k} \end{vmatrix} \cdot (-1)^{k+1+m}.$$

Это есть как раз алгебраическое дополнение элемента b_m в характеристическом определителе:

$$\Delta_{k+1} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_k \\ a_{k+1,1} & a_{k+1,2} & \dots & a_{k+1,k} & b_{k+1} \end{vmatrix},$$

и условие (34) выражает как раз равенство нулю этого характеристического определителя. Аналогичным образом при $y_{k+s} = 1$ мы получим условие $\Delta_{k+s} = 0$. Приходим, таким образом, к следующему результату: *если определитель системы* (30) *равен нулю, то для существования решения этой системы необходимо и достаточно, чтобы вектор* $(b_1, \dots, b_n)$ *был ортогонален ко всем векторам, дающим решение однородной союзной системы* (32).

Общее решение системы (30) будет суммой какого-нибудь частного решения этой системы и общего решения соответствующей однородной системы, которая получается из (30) заменой всех b_j нулями. Это общее решение однородной системы будет содержать $(n-k)$ произвольных постоянных.

Укажем еще одну, важную в дальнейшем, геометрическую интерпретацию основной теоремы о решении систем. Пусть имеется n линейных форм с n независимыми переменными:

$$\begin{gathered} y_1 = a_{11}x_1 + \dots + a_{1n}x_n, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ y_n = a_{n1}x_1 + \dots + a_{nn}x_n. \end{gathered}$$

Будем считать, что x_s могут принимать любые комплексные значения, а $(y_1, \ldots, y_n)$ — пусть составляющие некоторого вектора. Если определитель $|a_{ik}|$ отличен от нуля, то при любых значениях y_k мы получаем определенные значения x_k, и предыдущие формулы дают всё n-мерное пространство y. Положим теперь, что таблица $\|a_{ik}\|$ имеет ранг $r<n$. Не ограничивая общности, можем считать, что определитель r-го порядка, стоящий в левом верхнем углу, отличен от нуля. При этом основная теорема о решении систем дает нам следующее: совокупность значений $(y_1, \ldots, y_n)$, получаемых по предыдущим формулам, обладает тем свойством, что значения $y_1, \ldots, y_r$ могут быть произвольными, но если фиксировать эти значения, то значения $y_{r+1}, \ldots, y_n$ получаются вполне определенными, а именно эти последние значения получаются из условия равенства нулю характеристических определителей. На языке геометрии это значит, что предыдущие формулы дают подпространство измерения r, которое образовано теми векторами, которые получатся, если положить одно из y_s $(s=1, 2, \ldots, r)$ равным единице, а остальные — нулю. Итак, если ранг таблицы $\|a_{ik}\|$ равен r, то предыдущие формулы дают совокупность значений $(y_1, \ldots, y_n)$, определяющих некоторое подпространство измерения r.

Мы рассмотрели тот случай, когда число линейных форм равно числу переменных x_s. Можно взять общий случай

$$\begin{gathered}
y_1 = a_{11}x_1 + \ldots + a_{1n}x_n, \\
\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\
y_m = a_{m1}x_1 + \ldots + a_{mn}x_n.
\end{gathered}$$

В этом случае эти формулы при произвольных x_s определяют в пространстве m измерений подпространство, число измерений которого равно рангу таблицы $\|a_{ik}\|$. Доказательство не отличается вовсе от предыдущего.

16. Определитель Грамма. Неравенство Адамара. Пусть имеется m векторов:

$$x^{(s)}\,(x_1^{(s)},\ x_2^{(s)},\ \ldots,\ x_n^{(s)}) \qquad (s=1, 2, \ldots, m).$$

Составим определитель порядка m из скалярных произведений $(x^{(i)}, x^{(k)})$ и введем для него специальное обозначение:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) = |(x^{(i)}, x^{(k)})|_1^m =$$

$$= \begin{vmatrix}
(x^{(1)}, x^{(1)}) & (x^{(1)}, x^{(2)}) & \ldots & (x^{(1)}, x^{(m)}) \\
(x^{(2)}, x^{(1)}) & (x^{(2)}, x^{(2)}) & \ldots & (x^{(2)}, x^{(m)}) \\
\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot \\
(x^{(m)}, x^{(1)}) & (x^{(m)}, x^{(2)}) & \ldots & (x^{(m)}, x^{(m)})
\end{vmatrix}. \qquad (35)$$

Этот определитель называется *определителем Грамма векторов*

$$x^{(1)}, x^{(2)}, \ldots, x^{(m)}.$$

Разберем отдельно случаи $m=n$, $m<n$ и $m>n$. Общий член написанного определителя имеет вид:

$$(x^{(i)},\ x^{(k)}) = \sum_{s=1}^{n} x_s^{(i)} \overline{x_s^{(k)}}.$$

При $m=n$ определитель (35) равен произведению определителей

$$\begin{vmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} \\ \dots & \dots & \dots & \dots \\ x_1^{(n)} & x_2^{(n)} & \dots & x_n^{(n)} \end{vmatrix} \cdot \begin{vmatrix} \overline{x_1^{(1)}} & \overline{x_1^{(2)}} & \dots & \overline{x_1^{(n)}} \\ \overline{x_2^{(1)}} & \overline{x_2^{(2)}} & \dots & \overline{x_2^{(n)}} \\ \dots & \dots & \dots & \dots \\ \overline{x_n^{(1)}} & \overline{x_n^{(2)}} & \dots & \overline{x_n^{(n)}} \end{vmatrix},$$

причем применяется правило строка на столбец. Принимая во внимание неизменность величин определителя при замене строк столбцами, можем утверждать, что второй множитель будет комплексно сопряженным с первым, и, слецовательно, величина определителя Грамма (35) при $m=n$ равна квадрату модуля определителя $|x_k^{(i)}|_1^n$, образованного составляющими $x_k^{(i)}$ векторов $x_1, x_2, \dots, x_n$. Таким образом, определитель (35) положителен, если указанные векторы линейно независимы, и равен нулю, если они линейно зависимы [12]. При $m \neq n$ мы имеем две прямоугольные таблицы:

$$\left\| \begin{matrix} x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} \\ \dots & \dots & \dots & \dots \\ x_1^{(m)} & x_2^{(m)} & \dots & x_n^{(m)} \end{matrix} \right\| \quad (36_1) \quad \text{и} \quad \left\| \begin{matrix} \overline{x_1^{(1)}} & \overline{x_1^{(2)}} & \dots & \overline{x_1^{(m)}} \\ \overline{x_2^{(1)}} & \overline{x_2^{(2)}} & \dots & \overline{x_2^{(m)}} \\ \dots & \dots & \dots & \dots \\ \overline{x_m^{(1)}} & \overline{x_m^{(2)}} & \dots & \overline{x_m^{(m)}} \end{matrix} \right\|, \qquad (36_2)$$

и таблица, входящая в определитель (35), является произведением последних двух таблиц [7]. В силу теоремы, доказанной в [7], определитель (35) равен нулю при $m > n$. Но в этом случае векторы $x^{(1)}, x^{(2)}, \dots, x^{(m)}$ — линейно-зависимы [12]. При $m < n$, в силу указанной выше теоремы,

$$G(x^{(1)}, x^{(2)}, \dots, x^{(m)}) = \sum_{r_1 < r_2 < \dots < r_m} X\begin{pmatrix} 1, & 2, & \dots, & m \\ r_1, & r_2, & \dots, & r_m \end{pmatrix} Y\begin{pmatrix} r_1, & r_2, & \dots, & r_m \\ 1, & 2, & \dots, & m \end{pmatrix},$$

где $X\begin{pmatrix} 1, & 2, & \dots, & m \\ r_1, & r_2, & \dots, & r_m \end{pmatrix}$ означает миноры таблицы (36_1) и $Y\begin{pmatrix} r_1, & r_2, & \dots, & r_m \\ 1, & 2, & \dots, & m \end{pmatrix}$ — миноры таблицы (36_2). Как и выше, $Y\begin{pmatrix} r_1, & r_2, & \dots, & r_m \\ 1, & 2, & \dots, & m \end{pmatrix}$ есть число, сопряженное с $X\begin{pmatrix} 1, & 2, & \dots, & m \\ r_1, & r_2, & \dots, & r_m \end{pmatrix}$, и последняя формула переписывается в виде:

$$G(x^{(1)}, x^{(2)}, \dots, x^{(m)}) = \sum_{r_1 < r_2 < \dots < r_m} \left| X\begin{pmatrix} 1, & 2, & \dots, & m \\ r_1, & r_2, & \dots, & r_m \end{pmatrix} \right|^2. \qquad (37)$$

Если векторы $x^{(1)}, x^{(2)}, \dots, x^{(m)}$ — линейно-независимы, то ранг таблицы (36_1) равен m [12], и среди неотрицательных слагаемых, стоящих в правой части равенства (37), по крайней мере одно слагаемое положительно. Если же упомянутые векторы линейно-зависимы, то ранг таблицы (36_1) меньше m, все определители порядка m, содержащиеся в этой таблице, равны нулю, и из (37) следует, что в этом случае $G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) = 0$. Таким образом, рассмотрение всех трех случаев: $m = n$, $m > n$ и $m < n$ приводит нас к следующей общей теореме:

Теорема. *Определитель Грамма $G(x^{(1)}, x^{(2)}, \ldots, x^{(m)})$ положителен, если векторы $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$ линейно-независимы, и равен нулю, если они линейно-зависимы.*

Сейчас мы докажем еще одну формулу для определителя Грамма. Предварительно условимся в некоторых обозначениях. Пусть x — любой вектор из R_n, и для него имеет место разложение $x = y + z$, где y принадлежит подпространству, определяемому векторами $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$, а z — ортогонален к этому подпространству. Формула, которую мы хотим доказать, имеет вид:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) = \| z \|^2 G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}). \tag{38}$$

Принимая во внимание равенства: $(x^{(s)}, x) = (x^{(s)}, y)$; $(x, x^{(s)}) = (y, x^{(s)})$, которые следуют из ортогональности z ко всем $x^{(s)}$, и формулу $(x, x) = (y, y) + (z, z)$ [13], можем написать:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) =$$
$$= \begin{vmatrix} (x^{(1)}, x^{(1)}) & (x^{(1)}, x^{(2)}) & \ldots & (x^{(1)}, x^{(m)}) & (x^{(1)}, y) \\ (x^{(2)}, x^{(1)}) & (x^{(2)}, x^{(2)}) & \ldots & (x^{(2)}, x^{(m)}) & (x^{(2)}, y) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (x^{(m)}, x^{(1)}) & (x^{(m)}, x^{(2)}) & \ldots & (x^{(m)}, x^{(m)}) & (x^{(m)}, y) \\ (y, x^{(1)}) & (y, x^{(2)}) & \ldots & (y, x^{(m)}) & (y, y) + (z, z) \end{vmatrix}.$$

Представляя элементы последней строки в виде:

$$(y, x^{(1)}) + 0,\ (y, x^{(2)}) + 0,\ \ldots,\ (y, x^{(m)}) + 0,\ (y, y) + (z, z)$$

и представляя определитель в виде суммы двух определителей согласно свойству IV из [3], можем написать:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) = G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, y) +$$
$$+ \begin{vmatrix} (x^{(1)}, x^{(1)}) & (x^{(1)}, x^{(2)}) & \ldots & (x^{(1)}, x^{(m)}) & (x^{(1)}, y) \\ (x^{(2)}, x^{(1)}) & (x^{(2)}, x^{(2)}) & \ldots & (x^{(2)}, x^{(m)}) & (x^{(2)}, y) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (x^{(m)}, x^{(1)}) & (x^{(m)}, x^{(2)}) & \ldots & (x^{(m)}, x^{(m)}) & (x^{(m)}, y) \\ 0 & 0 & \ldots & 0 & \| z \|^2 \end{vmatrix}. \tag{39}$$

Вектор y принадлежит подпространству, определяемому векторами $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$, т. е. линейно выражается через $x^{(s)}$, и потому, в силу доказанной теоремы:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, y) = 0.$$

Разлагая определитель формулы (39) по элементам последней строки, мы и получаем формулу (38).

Из этой формулы непосредственно следует неравенство

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) \leqslant \| x \|^2 G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}). \tag{40}$$

Отметим, что если векторы $x^{(s)}$ линейно-зависимы, то

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) = G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) = 0.$$

Если $x^{(s)}$ — линейно-независимы, то в неравенстве (40) имеет место знак равенства в том и только в том случае, когда $y=0$, т. е. когда x ортогонален ко всем $x^{(s)}$.

Если мы к исходному определителю Грамма $G(x^{(1)}, x^{(2)}, \ldots, x^{(m)})$ применим несколько раз неравенство (40), то получим для него следующую оценку:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) \leqslant \| x^{(1)} \|^2 \| x^{(2)} \|^2 \ldots \| x^{(m)} \|^2. \tag{41}$$

При этом надо иметь в виду, что $G(x^{(1)}) = \| x^{(1)} \|^2$.

Знак равенства в (41) имеет место в том и только в том случае, когда векторы попарно ортогональны. Считается при этом, что ни один из векторов не есть нулевой вектор. Доказанное неравенство легко приводит к оценке любого определителя. Пусть Δ — определитель порядка n с элементами a_{ik}. Будем рассматривать элементы i-й строки этого определителя $(a_{i1}, a_{i2}, \ldots, a_{in})$ как составляющие некоторого вектора $x^{(i)}$ из R_n. Составим новый определитель с элементами $\bar{a}_{ik}$ — сопряженными с a_{ik}. Он равен, очевидно, $\bar{\Delta}$. Умножая Δ на $\bar{\Delta}$ по правилу строка на строку, получим определитель Грамма $G(x^{(1)}, x^{(2)}, \ldots, x^{(n)})$, величина которого, в силу теоремы об умножении определителей, равна $\Delta\bar{\Delta}$, т. е. равна $|\Delta|^2$. Применяя неравенство (41), получим следующую оценку для модуля определителя, принадлежащую Адамару:

$$|\Delta|^2 \leqslant \sum_{k=1}^{n} |a_{1k}|^2 \cdot \sum_{k=1}^{n} |a_{2k}|^2 \ldots \sum_{k=1}^{n} |a_{nk}|^2. \tag{42}$$

Если определитель Δ имеет вещественные элементы, то можем написать:

$$\Delta^2 \leqslant \sum_{k=1}^{n} a_{1k}^2 \cdot \sum_{k=1}^{n} a_{2k}^2 \ldots \sum_{k=1}^{n} a_{nk}^2. \tag{43}$$

Если для элементов определителя Δ имеет место оценка $|a_{ik}| \leqslant M$ $(i, k = 1, 2, \ldots, n)$, то, очевидно:

$$\sum_{k=1}^{n} |a_{ik}|^2 \leqslant nM^2,$$

и (42) дает оценку:

$$|\Delta| \leqslant n^{\frac{n}{2}} M^n. \tag{44}$$

Знак равенства в (42) имеет место, в силу сказанного выше, в том и только в том случае, когда векторы $x^{(i)}$ попарно ортогональны.

Можем получить и другие оценки для определителя Грамма. Они основываются на неравенстве, которое является обобщением неравенства (40).

Пусть каждая из букв X, Y, Z обозначает последовательность некоторых векторов из R_n. Неравенство, о котором мы упоминали, имеет вид:

$$G(X, Y, Z)\, G(X) \leqslant G(X, Z)\, G(X, Y). \tag{45}$$

В этом неравенстве не исключен случай пустой последовательности векторов, т. е. такой последовательности, которая не содержит ни одного вектора. Если W — такая последовательность, то надо считать $G(W) = 1$.

На основе этого неравенства может быть получена следующая оценка определителя Грамма:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) \leqslant \left[\prod_{k=1}^{m} G(x^{(1)}, \ldots, x^{(k-1)}, x^{(k+1)}, \ldots, x^{(m)})\right]^{\frac{1}{m-1}},$$

где Π — знак произведения. Многократное применение этой оценки приводит к новым оценкам, в которых определитель Грамма содержит меньшее число векторов. Во всех этих оценках знак равенства имеет место в том и только в том случае, когда векторы попарно ортогональны. Указанные только что оценки определителя Грамма даны М. К. Фаге (Доклады Академии наук СССР, 1946 г., т. LIV, № 9).

17. Системы линейных дифференциальных уравнений с постоянными коэффициентами. Применим полученные результаты к задаче интегрирования систем линейных дифференциальных уравнений с постоянными коэффициентами. Напишем такую систему в виде:

$$\left.\begin{array}{l} x_1' = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ x_2' = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n' = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n, \end{array}\right\}, \tag{46}$$

где x_j — искомые функции от t, x_j' — их производные и a_{ik} — заданные постоянные. Будем искать решение в виде:

$$x_1 = b_1 e^{\lambda t}; \quad x_2 = b_2 e^{\lambda t}; \quad \ldots; \quad x_n = b_n e^{\lambda t}. \tag{47}$$

Подставляя в систему (46) и сокращая на множитель $e^{\lambda t}$, будем иметь для определения постоянных $b_1, \ldots, b_n$ систему уравнений

$$\left.\begin{array}{l} (a_{11} - \lambda) b_1 + a_{12}b_2 + \ldots + a_{1n}b_n = 0 \\ a_{21}b_1 + (a_{22} - \lambda) b_2 + \ldots + a_{2n}b_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}b_1 + a_{n2}b_2 + \ldots + (a_{nn} - \lambda) b_n = 0. \end{array}\right\} \tag{48}$$

Так как для неизвестных b_j мы должны получить решение, отличное от нулевого, определитель написанной системы должен равняться нулю, т. е. для постоянной λ мы получаем уравнение вида:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{49}$$

Уравнение такого вида называется обычно *вековым уравнением*. Оно хорошо известно при рассмотрении колебаний механических систем без сопротивления в частном случае, когда таблица коэффициентов a_{ik} симметрична, т. е. когда $a_{ik} = a_{ki}$, и все коэффициенты вещественны, о чем мы будем говорить дальше при рассмотрении малых колебаний. Сейчас мы займемся исследованием системы в общем случае. Уравнение (49) есть алгебраическое уравнение степени n со старшим членом $(-\lambda)^n$. Если это уравнение имеет n различных корней

$$\lambda=\lambda_1, \ldots, \lambda=\lambda_n,$$

то, подставляя каждый из них $\lambda=\lambda_j$ в коэффициенты системы (48), мы будем иметь n однородных уравнений для соответствующих неизвестных $b_1, \ldots, b_n$ с определителем, равным нулю, и сможем получить для этих неизвестных решение, отличное от нулевого. Таким образом, по формулам (47) мы получим n линейно-независимых решений нашей системы (46), и их линейная комбинация даст общий интеграл системы. Если же вековое уравнение (49) имеет кратные корни, то решение задачи будет более сложным, а именно — каждому корню уравнения (49) кратности k должны соответствовать k линейно-независимых решений системы (46), причем одно из этих решений будет наверно иметь вид (47), а остальные будут, вообще говоря, содержать еще множителем полином от t. Но в данном случае в отличие от одного уравнения с постоянными коэффициентами [II, **39**] может случиться, что не одно, а несколько решений (или даже все), соответствующих упомянутому кратному корню, будут иметь вид (47). Мы не будем более подробно останавливаться на исследовании этого обстоятельства, так как в дальнейшем, пользуясь теорией функций комплексного переменного, применим для решения системы (46) другой метод.

Вернемся к рассмотрению векового уравнения (49), которое играет основную роль в нашей задаче. Решение этого уравнения, хотя бы и приближенное, представляет при больших n практические трудности, вызванные тем обстоятельством, что искомое неизвестное λ стоит не в каком-нибудь одном столбце или строке, а входит в диагональные члены. Разложение левой части этого уравнения по степеням λ требует больших вычислений, которые нами указаны выше в [**4**]. Укажем прием преобразования уравнения (49) к практически более удобному виду, при котором неизвестное λ попадает в один столбец. Прием этот принадлежит акад. А. Н. Крылову и изложен им в работе „О численном решении уравнения, которым в технических вопросах определяются частоты малых колебаний материальных систем" (Известия Академии наук СССР, 1931 г.).

Составим линейную комбинацию искомых величин

$$\xi=\alpha_{01}x_1+\alpha_{02}x_2+\ldots+\alpha_{0n}x_n, \tag{50}$$

где α_{0j} — взятые каким-нибудь образом численные коэффициенты. Будем дифференцировать уравнение (50) n раз по t, заменяя каждый раз в правой части производные x_j' их выражениями из системы (46). Мы получим таким образом $(n+1)$ уравнений вида:

$$\left.\begin{array}{l}
\xi=\alpha_{01}x_1+\alpha_{02}x_2+\ldots+\alpha_{0n}x_n \\
\xi'=\alpha_{11}x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\xi^{(n-1)}=\alpha_{n-1,1}x_1+\alpha_{n-1,2}x_2+\ldots+\alpha_{n-1,n}x_n \\
\xi^{(n)}=\alpha_{n1}x_1+\alpha_{n2}x_2+\ldots+\alpha_{nn}x_n.
\end{array}\right\} \tag{51}$$

Положим, что определитель, составленный из коэффициентов a_{ik}, входящих в первые n уравнений, отличен от нуля. При этом первые n уравнений дадут нам выражение x_j через $\xi, \xi', \ldots, \xi^{(n-1)}$, и, подставляя их в последнее уравнение, будем иметь уравнение n-го порядка для ξ. Можно непосредственно произвести это исключение x_j из $(n+1)$ уравнений (51) при помощи определителей. Действительно, перепишем эти уравнения следующим образом:

$$\begin{gathered}
\xi x_0 + a_{01}x_1 + a_{02}x_2 + \ldots + a_{0n}x_n = 0, \\
\xi' x_0 + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\xi^{(n)} x_0 + a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0,
\end{gathered}$$

где $x_0 = -1$, и будем рассматривать эти $(n+1)$ уравнений как однородные уравнения относительно величин $x_0, x_1, \ldots, x_n$.

Определитель написанной системы должен равняться нулю, что и дает нам искомый результат исключения

$$\begin{vmatrix} \xi & a_{01} & a_{02} & \ldots & a_{0n} \\ \xi' & a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \xi^{(n)} & a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = 0. \tag{52}$$

Будем искать решение этого уравнения в виде:

$$\xi = e^{\lambda t}.$$

Подставляя это в первый столбец определителя (52) и вынося из первого столбца за знак определителя множитель $e^{\lambda t}$, мы придем, сокращая на этот множитель, к следующему уравнению для определения λ:

$$\begin{vmatrix} 1 & a_{01} & a_{02} & \ldots & a_{0n} \\ \lambda & a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda^n & a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = 0. \tag{53}$$

Нетрудно показать, что при нашем предположении уравнение (53) имеет те же корни, что и уравнение (49). Действительно, пусть $\lambda = \lambda_0$ есть некоторый корень уравнения (53). Мы имеем при этом для (52) решение вида:

$$\xi = Ce^{\lambda_0 t}, \tag{54}$$

где C — произвольная постоянная. При этом первые n уравнений системы (51) дадут нам и для x_j решения вида (47) при $\lambda = \lambda_0$, т. е. $\lambda = \lambda_0$ будет и корнем уравнения (49). Наоборот, если $\lambda = \lambda_0$ есть корень уравнения (49), то мы имеем для x_j решение вида (47) при $\lambda = \lambda_0$, где b_j — численные постоянные, среди которых есть отличные от нуля. Подставляя эти выражения x_j в первое из уравнений системы (51), мы получим и для ξ решение вида (54), причем это решение будет наверно отлично от нуля. Действительно, в противном случае мы имели бы $\xi = \xi' = \ldots = \xi^{(n-1)} = 0$, откуда, в силу первых n уравнений системы (51) непосредственно следовало бы $x_1 = x_2 = \ldots = x_n = 0$.

Таким образом, действительно всякий корень $\lambda = \lambda_0$ уравнения (49) будет и корнем уравнения (53). Итак, при нашем предположении уравнение (53) имеет те же корни, что и уравнение (49). Применение этого метода к численным примерам, а также рассмотрение случая, когда наше предположение не имеет места, можно найти в вышеупомянутой статье акад. А. Н. Крылова.

Наиболее простые вычисления получатся, если формулу (50) взять в виде $\xi = x_1$. В этом случае уравнение (53) будет иметь вид:

$$\begin{vmatrix} 1 & 1 & 0 & \ldots & 0 \\ \lambda & a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda^n & a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = 0.$$ [1]

Вместо системы (46) рассмотрим систему уравнений второго порядка вида:

$$\left.\begin{aligned} x_1'' &= a_{11}x_1 + a_{11}x_2 + \ldots + a_{1n}x_n \\ x_2'' &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_n'' &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n. \end{aligned}\right\} \tag{55}$$

Такие системы часто встречаются в механике. Если искать их решение в виде

$$x_j = b_j \cos(\lambda t + \varphi),$$

то получим для λ уравнение вида:

$$\begin{vmatrix} a_{11}+\lambda^2 & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}+\lambda^2 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn}+\lambda^2 \end{vmatrix} = 0. \tag{56}$$

Постоянные b_j определятся из системы, аналогичной системе (48), и φ останется произвольным.

Наконец, для систем вида:

$$\left.\begin{aligned} x_1'' &= a_{11}x_1 + \ldots + a_{1n}x_n + c_{11}x_1' + \ldots + c_{1n}x_n' \\ x_2'' &= a_{21}x_1 + \ldots + a_{2n}x_n + c_{21}x_1' + \ldots + c_{2n}x_n' \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_n'' &= a_{n1}x_1 + \ldots + a_{nn}x_n + c_{n1}x_1' + \ldots + c_{nn}x_n', \end{aligned}\right\} \tag{57}$$

содержащих и первые производные, отыскивая опять решения в форме (47), мы приходим к вековому уравнению вида:

$$\begin{vmatrix} a_{11}+c_{11}\lambda-\lambda^2, & a_{12}+c_{12}\lambda & \ldots & a_{1n}+c_{1n}\lambda \\ a_{21}+c_{21}\lambda, & a_{22}+c_{22}\lambda-\lambda^2 & \ldots & a_{2n}+c_{2n}\lambda \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1}+c_{n1}\lambda, & a_{n2}+c_{n2}\lambda & \ldots & a_{nn}+c_{nn}\lambda-\lambda^2 \end{vmatrix} = 0. \tag{58}$$

---

[1]) Удобный метод преобразования векового уравнения был дан А. Данилевским (Математический сборник, т. 2, вып. 1).

Вводя добавочные неизвестные

$$x_{n+1} = x_1'; \quad x_{n+2} = x_2'; \ \ldots; \quad x_{2n} = x_n', \tag{59}$$

мы приведем систему (57) к $2n$ уравнениям первого порядка, из которых n уравнений получатся из (57) заменой

$$x_j'' = x_{n+j}' \text{ и } x_j' = x_{n+j} \quad (j = 1, 2, \ldots, n),$$

а остальные n суть уравнения (59).

18. Функциональные определители. Пусть имеются n функций от n переменных $x_1, x_2, \ldots, x_n$:

$$\varphi_1(x_1, x_2, \ldots, x_n); \ \varphi_2(x_1, x_2, \ldots, x_n); \ \ldots; \ \varphi_n(x_1, x_2, \ldots, x_n). \tag{60}$$

Функциональным определителем от этих функций по переменным x_s называется определитель n-го порядка, элементы которого вычисляются по формулам: $a_{ik} = \frac{\partial \varphi_i}{\partial x_k}$. Введем специальное обозначение для функционального определителя:

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)} = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & \ldots & \frac{\partial \varphi_1}{\partial x_n} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} & \ldots & \frac{\partial \varphi_2}{\partial x_n} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial \varphi_n}{\partial x_1} & \frac{\partial \varphi_n}{\partial x_2} & \ldots & \frac{\partial \varphi_n}{\partial x_n} \end{vmatrix}. \tag{61}$$

Мы встречались уже раньше с такими определителями при замене переменных в крагных интегралах [II, **63** и **80**]. Если мы имеем замену переменных на плоскости

$$x = \varphi(u, v); \quad y = \psi(u, v), \tag{62}$$

при которой точка (u, v) переходит в точку (x, y), то абсолютное значение функционального определителя

$$\frac{D(\varphi, \psi)}{D(u, v)} \tag{63}$$

дает коэффициент изменения площади в данной точке (u, v) при точечном преобразовании (61), причем мы считаем, что в той области, в которой применяется точечное преобразование (62), частные производные от функций (62) по u и v непрерывны, и определитель (63) в нуль не обращается. Аналогичным образом, если имеется точечное преобразование в трехмерном пространстве:

$$x = \varphi(q_1, q_2, q_3); \quad y = \psi(q_1, q_2, q_3); \quad z = \omega(q_1, q_2, q_3),$$

при котором точка с координатами (q_1, q_2, q_3) переходит в точку

(x, y, z), а объем (V_1) переходит в объем (V), то формула замены переменных в тройном интеграле имеет вид [II, 63]:

$$\iiint\limits_{V} f(x, y, z)\,dx\,dy\,dz = \iiint\limits_{(V_1)} f[\varphi, \psi, \omega]\,|D|\,dq_1\,dq_2\,dq_3,$$

где

$$D = \frac{D(\varphi, \psi, \omega)}{D(q_1, q_2, q_3)},$$

и $|D|$ дает коэффициент изменения объема в данном месте при переходе от (q_1, q_2, q_3) к (x, y, z).

Совершенно так же мы могли бы рассматривать одну функцию от одной независимой переменной $u = f(x)$ как точечное преобразование на оси OX, при котором точка с абсциссой x переходит в новое положение с абсциссой u. При этом, очевидно, абсолютное значение производной $|f'(x)|$ характеризует изменения линейного размера в данном месте. Все сказанное выше распространяется и на случай точечного преобразования в n-мерном пространстве и при замене переменных в n-кратном интеграле [II, 100].

Выясненная на случаях двух и трех измерений аналогия между функциональным определителем и производной имеет своим следствием и некоторую аналогию между их формальными свойствами, которую мы сейчас и покажем.

Пусть имеется система функций

$$\varphi_1(y_1, \ldots, y_n), \ldots, \varphi_n(y_1, \ldots, y_n),$$

и положим, что $y_1, \ldots, y_n$ не суть независимые переменные, а сами функции от $x_1, \ldots, x_n$, так что в конечном счете функции φ_i будут функциями переменных x_i. Мы можем составить три функциональных определителя:

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(y_1, \ldots, y_n)}; \quad \frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)}; \quad \frac{D(y_1, \ldots, y_n)}{D(x_1, \ldots, x_n)}.$$

Элементы этих функциональных определителей будут соответственно

$$\frac{\partial \varphi_i}{\partial y_k}; \quad \frac{\partial \varphi_i}{\partial x_k}; \quad \frac{\partial y_i}{\partial x_k}.$$

Но по правилу дифференцирования сложных функций мы имеем:

$$\frac{\partial \varphi_i}{\partial x_k} = \frac{\partial \varphi_i}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_k} + \ldots + \frac{\partial \varphi_i}{\partial y_n} \cdot \frac{\partial y_n}{\partial x_k},$$

и применяя теорему об умножении определителей по схеме строка на столбец, получим равенство, выражающее первое свойство функционального определителя:

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)} = \frac{D(\varphi_1, \ldots, \varphi_n)}{D(y_1, \ldots, y_n)} \cdot \frac{D(y_1, \ldots, y_n)}{D(x_1, \ldots, x_n)}. \tag{64}$$

Это равенство аналогично правилу дифференцирования сложных функций одной независимой переменной.

Выясним еще одно свойство функциональных определителей. Систему функций φ_i можно рассматривать как преобразование от переменных x_i к новым переменным φ_i:

$$\varphi_i = \varphi_i(x_1, \ldots, x_n) \quad (i = 1, 2, \ldots, n). \tag{65}$$

Отметим сначала один частный случай, а именно случай так называемого тождественного преобразования:

$$\varphi_1 = x_1; \quad \varphi_2 = x_2; \ \ldots; \ \varphi_n = x_n.$$

Его функциональный определитель будет

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

Представим себе, что уравнения (65) решены относительно x_i так, что x_i выражены через φ_i:

$$x_i = x_i(\varphi_1, \ldots, \varphi_n) \quad (i = 1, 2, \ldots, n). \tag{66}$$

Преобразование (66) естественно назвать обратным преобразованию (65). Если подставить выражения (66) в правые части равенств (65), то получим тождества $\varphi_1 = \varphi_1; \ldots; \varphi_n = \varphi_n$, или, иначе говоря, получим тождественное преобразование. Применим теперь к этому частному случаю формулу (64), причем нам надо будет считать $y_i = x_i$ и $x_i = \varphi_i$, а в левой части формулы (64) мы получим функциональный определитель тождественного преобразования:

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(\varphi_1, \ldots, \varphi_n)} = \frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)} \frac{D(x_1, \ldots, x_n)}{D(\varphi_1, \ldots, \varphi_n)}$$

или

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)} \frac{D(x_1, \ldots, x_n)}{D(\varphi_1, \ldots, \varphi_n)} = 1, \tag{67}$$

т. е. *произведение функциональных определителей прямого и обратного преобразований равно единице*. Это свойство аналогично свойству производной обратной функции для случая одного независимого переменного.

Выясним теперь смысл того условия, что функциональный определитель

$$\frac{D(\varphi_1, \varphi_2, \ldots, \varphi_n)}{D(x_1, x_2, \ldots, x_n)} \tag{68}$$

от функций $\varphi_1(x_1, \ldots, x_n)$; $\varphi_2(x_1, \ldots, x_n)$; ... ; $\varphi_n(x_1, \ldots, x_n)$ по переменным x_s тождественно равен нулю. Положим, что эти функции связаны функциональной зависимостью:

$$F(\varphi_1, \ldots, \varphi_n) = 0, \tag{69}$$

причем это равенство есть тождество относительно независимых переменных x_j. Дифференцируя его по всем независимым переменным, получим n тождеств:

$$\left.\begin{array}{c} \dfrac{\partial F}{\partial \varphi_1}\dfrac{\partial \varphi_1}{\partial x_1} + \ldots + \dfrac{\partial F}{\partial \varphi_n}\dfrac{\partial \varphi_n}{\partial x_1} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial F}{\partial \varphi_1}\dfrac{\partial \varphi_1}{\partial x_n} + \ldots + \dfrac{\partial F}{\partial \varphi_n}\dfrac{\partial \varphi_n}{\partial x_n} = 0. \end{array}\right\} \tag{70}$$

Мы можем рассматривать написанные тождества как линейные уравнения относительно n величин:

$$\frac{\partial F}{\partial \varphi_1}, \ \ldots, \ \frac{\partial F}{\partial \varphi_n},$$

причем, очевидно, величины эти не могут одновременно равняться нулю тождественно, так как в противном случае F не содержало бы ни одной из функций φ_i. Таким образом, определитель однородной системы (70) должен равняться нулю, что и сводится к тому, что функциональный определитель (68) равен нулю. Итак, из наличия функциональной зависимости (69) вытекает тождественное равенство нулю функционального определителя (68). Можно доказать и обратное предложение, на чем мы останавливаться не будем, т. е. *тождественное равенство нулю функционального определителя* (68) *является необходимым и достаточным условием зависимости между функциями* $\varphi_i(x_1, \ldots, x_n)$. [1])

Рассмотрим в качестве примера три функции от трех независимых переменных

$$\varphi_1 = x_1^2 + x_2^2 + x_3^2; \quad \varphi_2 = x_1 + x_2 + x_3; \quad \varphi_3 = x_1x_2 + x_1x_3 + x_2x_3. \tag{71}$$

Нетрудно проверить, что между ними существует следующая зависимость:

$$\varphi_2^2 - \varphi_1 - 2\varphi_3 = 0.$$

Составим для системы функций (71) функциональный определитель:

$$\frac{D(\varphi_1, \varphi_2, \varphi_3)}{D(x_1, x_2, x_3)} = \begin{vmatrix} 2x_1 & 2x_2 & 2x_3 \\ 1 & 1 & 1 \\ x_2 + x_3 & x_1 + x_3 & x_1 + x_2 \end{vmatrix}.$$

Предоставляем читателю показать, что этот определитель обращается тождественно в нуль.

---

[1]) Заметим, что наши рассуждения относительно системы (70) носят формальный характер и не являются, строго говоря, доказательством.

19. Неявные функции. Мы доказали в томе I теорему о существовании неявной функции [1, 15, 7], определяемой одним уравнением. Обобщим теперь эту теорему на случай системы уравнений. Формулируем сначала доказанную теорему: пусть $x=x_0$, $y=y_0$ — решение уравнения

$$F(x, y)=0 \tag{72}$$

и пусть $F(x, y)$ и ее частные производные первого порядка — непрерывные функции при $x=x_0$, $y=y_0$ и при всех значениях x, y, достаточно близких к ним, и пусть, наконец, частная производная $F'_y(x, y)$ отлична от нуля при $x=x_0$, $y=y_0$. При этом уравнение (72) определяет при x, достаточно близких к x_0, единственную функцию $y(x)$, непрерывную, имеющую производную и удовлетворяющую условию $y(x_0)=y_0$. Как мы уже упоминали, совершенно так же можно доказать, что уравнение

$$F(x, y, z)=0,$$

имеющее решение $x=x_0$, $y=y_0$, $z=z_0$, при условии непрерывности функции $F(x, y, z)$ и ее частных производных первого порядка в окрестности указанных значений и при условии $F'_z(x_0, y_0, z_0)\neq 0$, определяет единственную функцию $z(x, y)$, непрерывную в окрестности $x=x_0$, $y=y_0$, имеющую производные по x и y и удовлетворяющую условию $z(x_0, y_0)=z_0$. Рассмотрим теперь систему двух уравнений:

$$\varphi(x, y, z)=0; \quad \psi(x, y, z)=0. \tag{73}$$

Пусть эта система имеет решение $x=x_0, y=y_0, z=z_0$, *функции* $\varphi(x, y, z)$, $\psi(x, y, z)$ *и их частные производные непрерывны в окрестности указанных значений, и функциональный определитель:*

$$\frac{D(\varphi, \psi)}{D(y, z)}=\begin{vmatrix} \frac{\partial\varphi}{\partial y}, & \frac{\partial\varphi}{\partial z} \\ \frac{\partial\psi}{\partial y}, & \frac{\partial\psi}{\partial z}\end{vmatrix}=\frac{\partial\varphi}{\partial y}\frac{\partial\psi}{\partial z}-\frac{\partial\varphi}{\partial z}\frac{\partial\psi}{\partial y} \tag{74}$$

отличен от нуля при указанных значениях переменных. При этом система (73) *определяет при* x, *достаточно близких к* x_0, *единственную систему функций* $y(x)$, $z(x)$, *непрерывных, имеющих производные первого порядка и удовлетворяющих условию* $y(x_0)=y_0$, $z(x_0)=z_0$.

Так как выражение (74) отлично от нуля при $x=x_0$, $y=y_0$, $z=z_0$, то по крайней мере одна из частных производных $\frac{\partial\psi}{\partial y}$ или $\frac{\partial\psi}{\partial z}$ отлична от нуля. Положим, например, что $\frac{\partial\psi}{\partial z}$ отлична от нуля при указанных значениях переменных. Согласно формулированной выше теореме, второе из уравнений (73) определяет единственным образом функцию $z(x, y)$. Подставляя эту функцию в первое из уравнений системы, получим уравнение с переменными x и y:

$$\varphi[x, y, z(x, y)]=0. \tag{75}$$

Чтобы доказать теорему, нам остается только показать, что частная производная от левой части уравнения (75) по y отлична от нуля при $x=x_0$, $y=y_0$. Эта частная производная выразится формулой:

$$\left(\frac{\partial\varphi}{\partial y}\right)=\frac{\partial\varphi}{\partial y}+\frac{\partial\varphi}{\partial z}\cdot\frac{\partial z}{\partial y}, \tag{76}$$

где $\left(\frac{\partial\varphi}{\partial y}\right)$ есть полная производная от $\varphi(x, y, z)$ по аргументу y. Функция

$z(x, y)$ есть решение второго из уравнений (73), что приводит к тождеству:

$$\psi[x, y, z(x, y)] \equiv 0.$$

Дифференцируя это тождество по y, получим:

$$\frac{\partial\psi}{\partial y} + \frac{\partial\psi}{\partial z} \cdot \frac{\partial z}{\partial y} \equiv 0. \tag{77}$$

Умножим обе части (76) на $\frac{\partial\psi}{\partial z}$ и сложим с тождеством (77), предварительно умножив обе его части на $-\frac{\partial\varphi}{\partial z}$. Мы получим таким образом после элементарного преобразования:

$$\frac{\partial\psi}{\partial z} \cdot \left(\frac{\partial\varphi}{\partial y}\right) = \frac{D(\varphi, \psi)}{D(y, z)}.$$

При $x = x_0$, $y = y_0$ функция $z(x, y)$ обращается в z_0, и при указанных значениях переменных $\frac{\partial\psi}{\partial z}$ и (74) отличны от нуля, а потому $\left(\frac{\partial\varphi}{\partial y}\right)$ также отлична от нуля. Следовательно, уравнение (75) определяет единственную функцию $y(x)$. Подставляя ее в $z(x, y)$, получаем и z как функцию от x. Это доказательство справедливо для случая нескольких независимых переменных.

В общем случае теорема о неявных функциях читается так: *пусть имеется система n уравнений:*

$$F_1(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0; \ \ldots; \quad F_n(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0, \tag{78}$$

имеющая решение

$$x_k = x_k^{(0)}, \quad y_l = y_l^{(0)} \qquad \begin{pmatrix} k = 1, \ldots, & m \\ l = 1, \ldots, & n \end{pmatrix}; \tag{79}$$

пусть F_l — непрерывные функции с непрерывными частными производными первого порядка в окрестности значений (79) *и пусть, наконец, функциональный определитель*

$$\frac{D(F_1, \ldots, F_n)}{D(y_1, \ldots, y_n)} = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} & \cdots & \frac{\partial F_2}{\partial y_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_n}{\partial y_1} & \frac{\partial F_n}{\partial y_2} & \cdots & \frac{\partial F_n}{\partial y_n} \end{vmatrix} \tag{80}$$

отличен от нуля при значениях (79). *При этом уравнения* (78) *при x_k, достаточно близких к $x_k^{(0)}$, определяют единственную систему функций $y_l(x_1, \ldots, x_n)$ непрерывных, имеющих производные первого порядка и удовлетворяющих условию $y_l(x_1^{(0)}, \ldots, x_k^{(0)}) = y_l^{(0)}$.*

Наметим доказательство этой теоремы. Считая теорему справедливой для системы $(n-1)$ уравнений (при $n = 1$ и $n = 2$ она действительно справедлива), докажем ее для системы n уравнений. Разлагая определитель (80) по элементам первого столбца, можем утверждать, что хоть одно из алгебраических дополнений этих элементов отлично от нуля при значениях (79), ибо сам определитель при этих значениях по условию отличен от нуля. Соответственно нумеруя функции F_l, можем считать, что отлично от нуля алгебраическое дополнение элемента $\frac{\partial F_1}{\partial y_1}$. Это алгебраическое дополнение представляет собою функциональный определитель от функций $F_2, \ldots, F_n$, по переменным $y_2, \ldots, y_n$. Согласно теореме для систем $(n-1)$ уравнений, уравнения:

$$F_2(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0; \ldots; \quad F_n(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0 \tag{81}$$

определяют единственным образом функции

$$y_2 = \varphi_2(x_1, \ldots, x_m, y_1); \ldots; \quad y_n = \varphi_n(x_1, \ldots, x_m, y_1). \tag{82}$$

Подставляя эти функции в первое из уравнений системы (78), получим уравнение для определения y_1:

$$F_1(x_1, \ldots, x_m, y_1, \varphi_2, \ldots, \varphi_n) = 0. \tag{83}$$

Нам остается проверить, что полная производная от левой части этого уравнения по y_1 отлична от нуля при значениях (79). Эта производная выражается формулой:

$$\left(\frac{\partial F_1}{\partial y_1}\right) = \frac{\partial F_1}{\partial y_1} + \sum_{s=2}^{n} \frac{\partial F_1}{\partial \varphi_s} \cdot \frac{\partial \varphi_s}{\partial y_1}. \tag{84}$$

Подставляя функции (82) в левые части уравнений (81), получим тождества, которые дифференцируем по y_1:

$$\frac{\partial F_l}{\partial y_1} + \sum_{s=2}^{n} \frac{\partial F_l}{\partial \varphi_s} \cdot \frac{\partial \varphi_s}{\partial y_1} = 0 \quad (l = 2, \ldots, n). \tag{85}$$

Обозначим через $A_1, A_2, \ldots, A_n$ алгебраические дополнения элементов первого столбца определителя (80). Умножим (84) на A_1, (85) на A_l и вычтем последние тождества из первого. Таким путем мы получим равенство:

$$A_1\left(\frac{\partial F_1}{\partial y_1}\right) = \sum_{l=1}^{n} \frac{\partial F_l}{\partial y_1} A_l + \sum_{s=2}^{n} \left[\sum_{l=1}^{n} \frac{\partial F_l}{\partial \varphi_s} A_l\right] \frac{\partial \varphi_s}{\partial y_1}.$$

В правой части первая сумма дает определитель (80), который мы для краткости обозначим просто буквой D, а суммирование по l во второй части представляет собою сумму произведений элементов некоторого столбца D, с номером, отличным от единицы, на алгебраические дополнения соответствующих элементов первого столбца, т. е. эта сумма равна нулю. Заметим при этом, что дифференцирование по φ_s, все равно что дифференцирование по y_s. Таким образом предыдущая формула переписывается в виде:

$$A_1\left(\frac{\partial F_1}{\partial y_1}\right) = D.$$

При значениях (79) A_1 и D отличны от нуля, и, следовательно, то же можно утверждать и относительно производной левой части уравнения (83) по y_1, и это уравнение дает единственную функцию $y_1(x_1, x_2, \ldots, x_m)$, подставив которую в функции (82), получаем окончательный результат.

Частным случаем теоремы о неявных функциях является *теорема об обращении системы функций. Пусть имеются уравнения*

$$y_k = f_k(x_1, \ldots, x_n) \quad (k = 1, 2, \ldots, n). \tag{86}$$

Положим, что функции f_k непрерывны вместе со своими производными первого порядка в окрестности значений $x_k = x_k^{(0)}$ $(k = 1, 2, \ldots, n)$, и при этих значениях функциональный определитель

$$\frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)} \tag{87}$$

отличен от нуля. При этом уравнения (86) *определяют единственным образом $x_k(y_1, \ldots, y_n)$, как функции $y_1, \ldots, y_n$ в окрестности значений $y_k^{(0)} = f_k(x_k^{(0)}, \ldots, x_n^{(0)})$, причем эти функции непрерывны, имеют производные первого порядка и $x_k(y_1^{(0)}, \ldots, y_n^{(0)}) = x_k^{(0)}$.*

Чтобы доказать эту теорему, достаточно рассмотреть уравнения:

$$f_k(x_1, \ldots, x_n) - y_k = 0 \quad (k = 1, 2, \ldots, n)$$

и воспользоваться теоремой о неявных функциях, причем в данном случае роль y_l играют x_k.

Если f_k суть линейные однородные функции переменных x_k, то система (86) имеет вид:

$$y_k = a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n.$$

Определитель (87) в данном случае сводится к определителю $|a_{kl}|$ из коэффициентов a_{ki}, и возможность однозначного решения системы дается теоремой Крамера.

ГЛАВА II

ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ И КВАДРАТИЧНЫЕ ФОРМЫ

§ 3. ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ

20. Преобразование координат в трехмерном пространстве. Линейным преобразованием с n переменными называется преобразование вида:

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n. \end{aligned}\right\} \quad (1)$$

Можно толковать это преобразование как переход от некоторого вектора $(x_1, \ldots, x_n)$ n-мерного пространства к другому вектору $(x_1', \ldots, x_n')$. Можно иначе толковать $(x_1, \ldots, x_n)$ как координаты точки в n-мерном пространстве, а преобразование (1) — как переход от одной точки к другой.

Можно толковать (1) еще и иначе, а именно: считать, что $(x_1, \ldots, x_n)$ и $(x_1', \ldots, x_n')$ суть составляющие одного и того же вектора (координаты одной и той же точки), но при различном выборе осей Формулы (1) дают при этом формулы преобразования составляющих (координат) при переходе от одной координатной системы к другой. Раньше мы неоднократно встречались с формулами вида (1) при $n=2$ и $n=3$.

Настоящая глава в своей первой части и будет посвящена подробному рассмотрению линейного преобразования вида (1). Для большей ясности мы начнем с рассмотрения вещественного трехмерного пространства, а затем перейдем к общему случаю n-мерного пространства и комплексных составляющих. В случае трехмерного пространства мы начнем рассмотрение с того наиболее простого случая, когда преобразованию (1) соответствует переход от одних прямоугольных осей к другим таким же. Откладывая векторы от начала координат, мы можем считать, очевидно, (x_1, x_2, x_3) или составляющими вектора, или координатами его конца.

Как известно из аналитической геометрии, формулы преобразования прямолинейных прямоугольных координат имеют вид:

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned}\right\}, \tag{2}$$

где a_{ik} суть косинусы углов, образованных новыми осями с прежними; a_{ik} определяются по следующей таблице:

	X_1	X_2	X_3
X_1'	a_{11}	a_{12}	a_{13}
X_2'	a_{21}	a_{22}	a_{23}
X_3'	a_{31}	a_{32}	a_{33}

(3)

Как известно, таблица коэффициентов (3) обладает в данном случае тем свойством, что сумма квадратов элементов каждой строки и столбца равна единице и сумма произведений соответствующих элементов двух разных строк или разных столбцов равна нулю. Величина определителя $|a_{ik}|$ равна, очевидно [**5**], объему прямоугольного параллелепипеда с ребрами, равными единице и направленными по новым координатным осям, т. е. эта величина равна единице, если ориентировки осей одинаковы, и (-1), если они различны. Преобразование, дающее обратный переход от (x_1', x_2', x_3') к (x_1, x_2, x_3), будет, очевидно, иметь вид:

$$\left.\begin{aligned} x_1 &= a_{11}x_1' + a_{21}x_2' + a_{31}x_3' \\ x_2 &= a_{12}x_1' + a_{22}x_2' + a_{32}x_3' \\ x_3 &= a_{13}x_1' + a_{23}x_2' + a_{33}x_3'. \end{aligned}\right\}. \tag{2_1}$$

Иначе говоря, преобразование, обратное (2), получается простой перестановкой строк и столбцов таблицы его коэффициентов. Определитель этого обратного преобразования, очевидно, равен определителю преобразования (2).

Покажем теперь, что упомянутое выше свойство коэффициентов преобразования (2) может быть получено при выполнении одного требования, которое непосредственно вытекает из геометрической сущности вопроса, а именно — поставим себе следующую задачу: найти все вещественные преобразования вида (2) такие, что

$$x_1'^2 + x_2'^2 + x_3'^2 = x_1^2 + x_2^2 + x_3^2. \tag{4}$$

Такая постановка задачи дает возможность обобщить рассмотренные нами выше преобразования на случай пространства с любым числом измерений. Покажем, что наши преобразования, требуемые новой задачей, совпадают с теми, которые нами рассмотрены были выше, т. е. покажем, что требование (4) приведет к упомянутым выше соотношениям для коэффициентов a_{ik}. Подставляя выражения (2) в левую часть соотношения (4), раскрывая скобки и приравнивая коэффициенты при квадратах переменных единице, а при произведении различных коэффициентов — нулю, мы получим как раз шесть соотношений вида:

$$a_{1k}a_{1l}+a_{2k}a_{2l}+a_{3k}a_{3l}=\delta_{kl} \quad (k,\ l=1,\ 2,\ 3), \tag{5}$$

где

$$\delta_{kl}=0 \quad \text{при } k\neq l \quad \text{и} \quad \delta_{kk}=1, \tag{6}$$

т. е. сумма квадратов элементов каждого столбца равна единице и сумма произведений соответствующих элементов разных столбцов равна нулю. Эти условия называются обычно *условиями ортогональности по столбцам.* Таким образом, элементы каждого столбца равны направляющим косинусам некоторого направления и направления, соответствующие различным столбцам, взаимно ортогональны. Отсюда непосредственно следует, что в рассматриваемом случае преобразование (2) совпадает с рассмотренным выше, и свойство ортогональности имеет место не только по отношению к столбцам, но и по отношению к строкам.

Формулы (2) мы можем толковать не как преобразование координат в неизменном пространстве, а как преобразование пространства при неизменности координатных осей. Положим сначала, что определитель преобразования равен $(+1)$, т. е. обе системы координатных осей имеют одну и ту же ориентировку. Мы можем при этом повернуть пространство, как твердое целое, вокруг начала вместе с осями $(X_1',\ X_2',\ X_3')$ так, чтобы эти оси совпали с осями $(X_1,\ X_2,\ X_3)$, которые мы считаем неподвижными при этом вращении и к которым мы относим координаты всякой точки как до вращения, так и после него. Если некоторая точка M имела до вращения координаты $(x_1,\ x_2,\ x_3)$, то в результате вращения эта точка займет новое положение M' и будет иметь новые координаты $(x_1',\ x_2',\ x_3')$. Поскольку точка M движется вместе с координатными осями $(X_1',\ X_2',\ X_3')$, координаты $(x_1',\ x_2',\ x_3')$ точки M' относительно осей $(X_1,\ X_2,\ X_3)$, с которыми совпали оси $(X_1',\ X_2',\ X_3')$ в результате вращения, будут совпадать с координатами точки M относительно осей $(X_1',\ X_2',\ X_3')$ до вращения. Мы видим, таким образом, что формулы (2) в случае определителя $(+1)$ представляют собой преобразование координат любой точки в результате произведенного вращения пространства.

Положим теперь, что определитель $|a_{ik}|$ равен (-1). Вместо преобразования (2) рассмотрим преобразование

$$x_i''=-a_{i1}x_1-a_{i2}x_2-a_{i3}x_3 \qquad (i=1,\ 2,\ 3).$$

Его коэффициенты по-прежнему обладают свойствами (5), но определитель из коэффициентов уже равен $(+1)$, т. е. этому преобразованию соответствует вращение пространства вокруг начала. Чтобы получить координаты (x_1', x_2', x_3'), нам надо совершить еще преобразование:

$$x_1' = -x_1''; \quad x_2' = -x_2''; \quad x_3' = -x_3'',$$

а такое изменение знаков у всех координат есть преобразование симметрии относительно начала. Итак, в случае определителя (-1) преобразованию (2) соответствует некоторое вращение пространства вокруг начала с последующей симметрией относительно начала.

Мы видели выше, что девять коэффициентов a_{ik} должны удовлетворять шести соотношениям (5). Отсюда следует, что они должны выражаться через три независимых параметра. Укажем на один из возможных выборов этих параметров в случае вращения пространства вокруг начала.

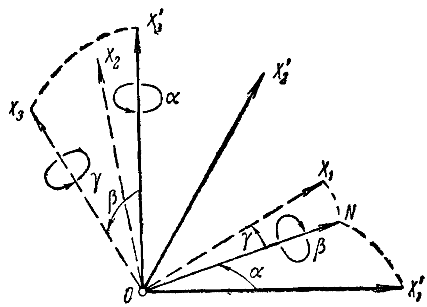


Рис. 1.

Введем в рассмотрение две системы координатных осей: одну X_1', X_2', X_3' — неподвижную, к которой относятся все координаты, и другую X_1, X_2, X_3, неизменно связанную с вращающимся пространством. Чтобы определить вращение, надо установить три параметра, которые определяли бы положение второй системы осей относительно первой. Пусть ON (рис. 1) — пересечение плоскости $X_1'OX_2'$ с плоскостью X_1OX_2. Возьмем на этой прямой определенным образом выбранное направление, и пусть α — угол $X_1'ON$, отсчитываемый от OX_1'. Введем еще угол $\beta = X_3'OX_3$ и угол $\gamma = NOX_1$. Эти три угла вполне характеризуют положение второй системы осей относительно первой, т. е. вполне характеризуют произведенное вращение. Будем обозначать его следующим символом: $\{\alpha, \beta, \gamma\}$. Из предыдущего непосредственно следует, что наше движение есть результат последовательно проведенных трех следующих движений: 1) вращение вокруг оси X_3' на угол α; 2) вращение вокруг нового положения оси X_1' на угол β; 3) вращение вокруг новой оси X_3 на угол γ. Эти три угла называются обычно углами Эйлера. Заметим, что для этих углов мы можем написать следующие пределы их изменения: $0 \leqslant \alpha < 2\pi$; $0 \leqslant \beta < \pi$; $0 \leqslant \gamma < 2\pi$.

Если $\beta = 0$, то движение $\{\alpha, \beta, \gamma\}$ сводится просто к повороту вокруг оси X_3 на угол $\alpha + \gamma$, и в этом смысле при любом δ мы имеем:

$$\{\alpha,\ 0,\ \gamma\} = \{\alpha + \delta,\ 0,\ \gamma - \delta\}.$$

Это обстоятельство указывает на то, что в некоторых случаях параметры $\{\alpha, \beta, \gamma\}$ соответствуют вращениям пространства вокруг начала не вполне однозначно. Различным значениям параметров соответствуют одни и те же вращения. Нетрудно вывести формулы, выражающие коэффициенты a_{ik} ерез тригонометрические функции углов α, β, γ [см. **62**]. В дальнейшем мы будем иметь и другой выбор параметров, характеризующих вращение пространства вокруг начала, и вернемся еще к углам Эйлера.

21. Общие линейные преобразования вещественного трехмерного пространства. Будем рассматривать теперь линейные вещественные преобразования вида (2) с любыми коэффициентами, но будем всегда считать, что определитель преобразования отличен от нуля:

$$|a_{ik}| \neq 0. \tag{7}$$

В этом случае преобразование обычно называется *собственным преобразованием*. Если оно не удовлетворяет условиям (5), то оно связано с деформацией пространства [II, **125**]. Заметим, что в преобразовании (2) характерным является таблица его коэффициентов, из которой вполне определяется закон перехода от любого вектора с составляющими (x_1, x_2, x_3) к новому вектору с составляющими (x_1', x_2', x_3'). Обозначим всю таблицу коэффициентов одной буквой

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}, \tag{8}$$

причем мы ставим эту таблицу, как и раньше, между двойными чертами, чтобы отличить ее от определителя. Она называется иначе *матрицей*. Определитель таблицы (8) будем обозначать символом $D(A)$. Это есть некоторое определенное число. Преобразование (2) мы запишем символически в следующем виде:

$$\boldsymbol{x}' = A\boldsymbol{x}, \tag{9}$$

где $\boldsymbol{x}'$ — вектор с составляющими (x_1', x_2', x_3') и $\boldsymbol{x}$ — вектор с составляющими (x_1, x_2, x_3).

Преобразование, при котором каждый вектор остается неизменным, назовем *тождественным преобразованием*. Ему соответствует таблица

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \tag{10}$$

которая называется обычно *единичной таблицей*, или *единичной матрицей*, и обозначается символом I.

Считая $D(A)\neq 0$, можем решить уравнения (2) относительно (x_1, x_2, x_3) и получим формулы:

$$\left.\begin{aligned} x_1&=\frac{A_{11}}{D(A)}x_1'+\frac{A_{21}}{D(A)}x_2'+\frac{A_{31}}{D(A)}x_3'\\ x_2&=\frac{A_{12}}{D(A)}x_1'+\frac{A_{22}}{D(A)}x_2'+\frac{A_{32}}{D(A)}x_3\\ x_3&=\frac{A_{13}}{D(A)}x_1'+\frac{A_{23}}{D(A)}x_2'+\frac{A_{33}}{D(A)}x_3'\end{aligned}\right\}, \tag{11}$$

где A_{ik} — алгебраические дополнения элементов a_{ik} в определителе $D(A)$. Это линейное преобразование называется обычно преобразованием, обратным (2), и, если матрица этого последнего обозначена через A, то матрицу преобразования (11) обозначают через A^{-1}. Введем теперь важное для дальнейшего понятие о произведении двух преобразований или произведении двух матриц. Положим, что мы имеем два линейных преобразования, от (x_1, x_2, x_3) к (x_1', x_2', x_3'):

$$\left.\begin{aligned} x_1'&=a_{11}x_1+a_{12}x_2+a_{13}x_3\\ x_2'&=a_{21}x_1+a_{22}x_2+a_{23}x_3\\ x_3'&=a_{31}x_1+a_{32}x_2+a_{33}x_3\end{aligned}\right\} \text{ или } x'=Ax \tag{12}$$

и затем от (x_1', x_2', x_3') к (x_1'', x_2'', x_3''):

$$\left.\begin{aligned} x_1''&=b_{11}x_1'+b_{12}x_2'+b_{13}x_3'\\ x_2''&=b_{21}x_1'+b_{22}x_2'+b_{23}x_3'\\ x_3''&=b_{31}x_1'+b_{32}x_2'+b_{33}x_3'\end{aligned}\right\} \text{ или } x''=Bx'. \tag{13}$$

Этот последовательный переход от (x_1, x_2, x_3) к (x_1', x_2', x_3') и затем от (x_1', x_2', x_3') к (x_1'', x_2'', x_3'') можно заменить непосредственным переходом от (x_1, x_2, x_3) к (x_1'', x_2'', x_3''), который также будет линейным преобразованием

$$x_k''=c_{k1}x_1+c_{k2}x_2+c_{k3}x_3 \qquad (k=1,\ 2,\ 3). \tag{14}$$

Это последнее линейное преобразование называется произведением преобразований (12) и (13), причем здесь существенно отметить порядок, в котором производились преобразования. Подставляя выражения (12) в правые части формулы (13), мы и получим формулы (14). Отсюда непосредственно получается выражение элементов c_{ik} таблицы произведения преобразований через элементы таблиц перемножаемых преобразований:

$$c_{ik}=\sum_{s=1}^{3} b_{is}a_{sk} \qquad (i,\ k=1,\ 2,\ 3). \tag{15}$$

Преобразование (14) обычно записывается следующим образом:

$$x'' = BAx. \tag{16}$$

Матрицу C с элементами c_{ik}, получаемыми по формулам (15), называют произведением матриц A и B и записывают это в следующем виде:

$$C = BA, \tag{17}$$

причем в смысле порядка следования преобразований надо читать это произведение справа налево. Принимая во внимание теорему об умножении определителей и формулы (15), мы можем написать для определителей преобразований очевидное равенство:

$$D(C) = D(B)\,D(A), \tag{18}$$

т. е. *определитель произведения преобразований равен произведению определителей этих преобразований.* Нетрудно доказать следующие соотношения, имеющие простой геометрический смысл:

$$AA^{-1} = A^{-1}A = I. \tag{19}$$

Заметим, кроме того, что из самого процесса образования обратного преобразования следует, что преобразование, обратное A^{-1}, будет преобразование A. Действительно, решая систему (11) относительно x'_k, получим, очевидно, опять формулы (2). Это можно записать следующим образом:

$$(A^{-1})^{-1} = A. \tag{20}$$

Понятие произведения преобразований может быть распространено на случай любого числа сомножителей, а именно: результат последовательного преобразования с некоторыми матрицами A, B и C будет новое преобразование с матрицей D:

$$D = CBA. \tag{21}$$

Если матрицы A, B и C имеют элементы a_{ik}, b_{ik} и c_{ik}, то матрица D, являющаяся их произведением, будет иметь элементы, определяемые по формулам:

$$d_{ik} = \sum_{p,\, q=1}^{3} c_{iq} b_{qp} a_{pk}. \tag{22}$$

Действительно, для элементов матрицы $E = BA$ имеем формулы

$$e_{ik} = \sum_{p=1}^{3} b_{ip} a_{pk}$$

и, наконец, для матрицы CE будем иметь, согласно (15), формулы

$$d_{ik}=\sum_{q=1}^{3} c_{iq}e_{qk},$$

т. е. как раз формулы (22). Заметим, что в дальнейшем мы будем часто обозначать элементы матрицы A символом

$$\{A\}_{ik}.$$

Произведение матриц не подчиняется, вообще говоря, *переместительному закону*, т. е. вообще $BA \neq AB$, но, как нетрудно видеть, подчиняется сочетательному закону — сомножители можно соединять в группы:

$$C(BA)=(CB)A. \tag{23}$$

В левой его части мы должны матрицу A умножить на B слева и затем полученный результат умножить на матрицу C. В правой части мы сначала должны умножить матрицу B на C, а затем матрицу A помножить на матрицу, полученную в результате только что произведенного умножения. Нетрудно видеть, что в обоих случаях элементы матрицы, получающиеся в окончательном произведении, будут выражаться по формулам (22). Действительно, для левой части это было показано выше. Что же касается правой части, то мы имеем, производя последовательно указанные умножения:

$$\{CB\}_{ik}=\sum_{q=1}^{3}\{C\}_{iq}\{B\}_{qk}$$

и

$$\{(CB)A\}_{ik}=\sum_{p=1}^{3}\{CB\}_{ip}\{A\}_{pk}=\sum_{p,\,q=1}^{3}\{C\}_{iq}\{B\}_{qp}\{A\}_{pk},$$

что совпадает, очевидно, с (22) при наших новых обозначениях.

Отметим еще один важный тип линейных преобразований, а именно рассмотрим преобразования

$$x_1'=k_1x_1;\quad x_2'=k_2x_2;\quad x_3'=k_3x_3, \tag{24}$$

которые сводятся к растяжению вдоль координатных осей, причем численные коэффициенты k_1, k_2, k_3 характеризуют величину этого растяжения (или сжатия). Матрица указанного преобразования имеет, очевидно, вид:

$$\begin{Vmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{Vmatrix},$$

т. е. все ее элементы, стоящие вне главной диагонали, равны нулю. Такая матрица называется *диагональной матрицей,* и мы ее будем обозначать символом

$$[k_1,\ k_2,\ k_3].$$

В частности, если числа одинаковы, т. е. если $k_1 = k_2 = k_3 = k$, то преобразование сведется к умножению всех составляющих вектора на одно и то же число k, и это будет, очевидно, преобразование подобия с центром в начале координат. Всякий вектор, не меняя своего направления (считаем $k > 0$), изменится лишь по длине, причем его длина умножится на k. Такое преобразование будем обозначать

$$x' = kx,$$

т. е. будем число k считать частным случаем матрицы, а именно будем диагональную матрицу с одинаковыми элементами на главной диагонали

$$\begin{Vmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{Vmatrix} \tag{25}$$

считать числом k. Нетрудно видеть, пользуясь (15), что произведение таких матриц сводится к обычному перемножению чисел, т. е.

$$[k,\ k,\ k]\cdot[l,\ l,\ l] = [kl,\ kl,\ kl].$$

Вообще легко проверить, что для диагональных матриц имеет место простой закон умножения

$$[k_1,\ k_2,\ k_3]\ [l_1,\ l_2,\ l_3] = [k_1 l_1,\ k_2 l_2,\ k_3 l_3], \tag{26}$$

т. е. два растяжения вдоль координатных осей равносильны одному растяжению с коэффициентами, равными произведению соответствующих коэффициентов составляющих растяжений. Из формулы (26), между прочим, непосредственно следует, что произведение двух диагональных матриц не меняется при перестановке сомножителей. Пользуясь представлением числа в виде диагональной матрицы (25) и формулой (15), нетрудно видеть, что произведение kA сводится к умножению всех элементов матрицы A на число k. Это произведение не зависит от порядка сомножителей, т. е.

$$\{kA\}_{ik} = \{Ak\}_{ik} = k\,\{A\}_{ik}. \tag{27}$$

Мы рассматривали основное линейное преобразование (2) как деформацию пространства, при которой вектор с составляющими $(x_1,\ x_2,\ x_3)$ переходит в новый вектор с составляющими $(x'_1,\ x'_2,\ x'_3)$. Можно, конечно, как мы уже указывали и раньше, толковать это преобразование и как точечное преобразование, при котором точка с координатами (x_1, x_2, x_3) переходит в точку с координатами (x'_1, x'_2, x'_3).

При определении составляющих вектора мы могли пользоваться любой системой осей, иначе говоря, любыми ортами, т. е. мы могли взять любые три некомпланарных вектора $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ за основные векторы (за орты), и при этом для всякого вектора $\boldsymbol{x}$ мы, как известно, имеем единственное представление вида [II, 114]:

$$\boldsymbol{x} = x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}. \tag{28}$$

Числа x_1, x_2, x_3 и называются составляющими вектора $\boldsymbol{x}$ во взятой системе координат, определяемой ортами $\boldsymbol{i}$, $\boldsymbol{j}$ и $\boldsymbol{k}$. Нашей задачей сейчас является исследовать влияние различного выбора ортов на вид линейного преобразования.

Точнее говоря, если в системе координат, определяемой ортами $\boldsymbol{i}$, $\boldsymbol{j}$ и $\boldsymbol{k}$, некоторое линейное преобразование имеет вид (12), то каким образом будет выглядеть то же самое линейное преобразование пространства в другой системе координат, определяемой ортами $\boldsymbol{i}_1$, $\boldsymbol{j}_1$, $\boldsymbol{k}_1$? Положим, что новые орты выражаются через старые по формулам:

$$\left.\begin{aligned} \boldsymbol{i}_1 &= t_{11}\boldsymbol{i} + t_{12}\boldsymbol{j} + t_{13}\boldsymbol{k} \\ \boldsymbol{j}_1 &= t_{21}\boldsymbol{i} + t_{22}\boldsymbol{j} + t_{23}\boldsymbol{k} \\ \boldsymbol{k}_1 &= t_{31}\boldsymbol{i} + t_{32}\boldsymbol{j} + t_{33}\boldsymbol{k}. \end{aligned}\right\} \tag{29}$$

Заметим, что определитель, составленный из коэффициентов t_{ik}, должен быть отличным от нуля. В противном случае векторы $\boldsymbol{i}_1$, $\boldsymbol{j}_1$ и $\boldsymbol{k}_1$ оказались бы линейно зависимыми, т. е. компланарными. В новой системе координат наш вектор, определяемый формулой (28), будет иметь уже новые составляющие

$$y_1\boldsymbol{i}_1 + y_2\boldsymbol{j}_1 + y_3\boldsymbol{k}_1.$$

Установим прежде всего формулу, выражающую новые составляющие вектора через его прежние составляющие. Мы имеем, очевидно, используя выражения (29) новых ортов:

$$\sum_{s=1}^{3} y_s(t_{s1}\boldsymbol{i} + t_{s2}\boldsymbol{j} + t_{s3}\boldsymbol{k}) = x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}.$$

Сравнивая коэффициенты при $\boldsymbol{i}$, $\boldsymbol{j}$ и $\boldsymbol{k}$, получим формулы, выражающие прежние составляющие через новые:

$$\left.\begin{aligned} x_1 &= t_{11}y_1 + t_{21}y_2 + t_{31}y_3 \\ x_2 &= t_{12}y_1 + t_{22}y_2 + t_{32}y_3 \\ x_3 &= t_{13}y_1 + t_{23}y_2 + t_{33}y_3. \end{aligned}\right\} \tag{30}$$

Таблица этого преобразования отличается от таблицы преобразования (29) тем, что строки заменены столбцами. Действительно, в каждой строке таблицы (29) не меняется первый значок, а в таблице (30) это имеет место для второго значка. Обозначая таблицу преобразования (29) через T, мы будем таблицу преобразования (30) обозначать через $T^{(*)}$ и называть ее транспонированной по отношению к T. При этом можно записать формулы (30) коротко в следующей форме:

$$(x_1,\ x_2,\ x_3) = T^{(*)}(y_1, y_2, y_3), \tag{31}$$

причем (x_1, x_2, x_3) обозначает три составляющие вектора в прежних ортах, а (y_1, y_2, y_3) — в новых ортах. Выражение новых составляющих через старые будет иметь вид:

$$(y_1, y_2, y_3) = T^{(*)-1}(x_1, x_2, x_3),$$

где $T^{(*)-1}$ есть линейное преобразование, обратное $T^{(*)}$. Оно называется обычно *контраградиентным* по отношению к T. Для краткости письма обозначим соответствующую ему таблицу особой буквой

$$U = T^{(*)-1}. \tag{32}$$

Мы можем, таким образом, утверждать, что при перемене основных ортов, согласно формуле (29), составляющие всякого вектора испытывают линейное преобразование с таблицей U, определяемой по формуле (32). Таким образом, наши два вектора $\boldsymbol{x}(x_1, x_2, x_3)$ и $\boldsymbol{x}'(x_1', x_2', x_3')$, входящие в преобразование (9), после преобразования ортов будут иметь уже другие составляющие, определяемые через прежние по формулам:

$$(y_1, y_2, y_3) = U(x_1, x_2, x_3);\quad (y_1', y_2', y_3') = U(x_1', x_2', x_3'). \tag{33}$$

Нашей задачей, таким образом, является установить линейную зависимость между составляющими (y_1, y_2, y_3) и (y_1', y_2', y_3'). Мы можем совершить переход от вектора (y_1, y_2, y_3) к новому вектору (y_1', y_2', y_3') следующим образом: сначала перейти от вектора (y_1, y_2, y_3) к вектору (x_1, x_2, x_3), что в силу (33) совершается при помощи таблицы U^{-1}. Затем перейти от вектора (x_1, x_2, x_3) к вектору (x_1', x_2', x_3') при помощи таблицы A преобразования (9) и, наконец, от вектора (x_1', x_2', x_3') к вектору (y_1', y_2', y_3') при помощи таблицы преобразования U. Окончательно мы будем иметь линейное преобразование следующего вида:

$$\boldsymbol{y}' = UAU^{-1}\boldsymbol{y}. \tag{34}$$

Это преобразование называется подобным преобразованию (9), и его матрица UAU^{-1} называется подобной матрице A.

Формулируем окончательно полученный результат. *Если формулы* (33) *выражают линейное преобразование составляющих* *вектора, вызываемое переменой основных ортов, то всякое линейное преобразование пространства, которое в прежних основных ортах имело вид:*

$$x' = Ax,$$

будет иметь в новой координатной системе вид:

$$y' = UAU^{-1}y.$$

22. Ковариантные и контравариантные афинные векторы. Положим, что линейное преобразование (9) выражает просто переход от одной системы декартовых координатных осей к другой, т. е. что его коэффициенты суть направляющие косинусы, определяемые по таблице (3). В этом случае, как мы видели в [20], транспонированная таблица $A^{(*)}$ совпадает с обратной таблицей A^{-1} и, следовательно, контраградиентная таблица $A^{(*)-1}$ будет совпадать с основной таблицей A, т. е.

$$A^{(*)} = A^{-1}; \quad A^{(*)-1} = A. \tag{35}$$

Рассматривая неизменный по длине и направлению вектор, мы можем, конечно, утверждать, что его составляющие преобразуются по тем же самым формулам (9), что и координаты, т. е.

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{aligned}\right\} \tag{36}$$

Таким образом, вектор вполне характеризуется тремя числами во всякой фиксированной декартовой координатной системе, и при переходе от одной декартовой системы к другой эти три числа (составляющие вектора) преобразуются по тем же формулам (36), что и координаты. Предположим теперь, что мы принимаем в расчет не только переход от одной декартовой системы к другой, но вообще всевозможные линейные преобразования координат с определителем, отличным от нуля, что соответствует, как мы выше видели, произвольному выбору трех некомпланарных векторов за основные орты. Будем, как и выше, наряду с таблицей A преобразования (36) рассматривать и контраградиентную таблицу $V = A^{(*)-1}$. В общем случае они будут различными, и мы имеем таким образом возможность двоякого определения вектора при любом линейном преобразовании координат. Во-первых, можно определить вектор как тройку чисел, которая преобразуется при переходе от одной системы координат к другой по тем же формулам, что и сами координаты, т. е. по формулам

$$(x_1', x_2', x_3') = A(x_1, x_2, x_3). \tag{37}$$

Такой вектор назовем *контравариантным афинным вектором,* причем общее линейное преобразование (36) называется иногда афинным преобразованием. Иначе, можно определить вектор так, чтобы при всяком линейном преобразовании (36) его составляющие испытывали соответствующее контраградиентное преобразование, т. е.

$$(x_1', x_2', x_3') = V(x_1, x_2, x_3). \tag{38}$$

Такой вектор называется *ковариантным афинным вектором.*

В обоих случаях, имея составляющие вектора в какой-нибудь координатной системе, мы тем самым имеем его составляющие и во всех других системах координат, получаемых из упомянутой при помощи любого афинного преобразования. Приведем примеры векторов обоего рода. Прежде всего радиус-вектор, соединяющий две определенные точки пространства, является, очевидно, контравариантным вектором, так как его составляющие в вышеуказанном смысле слова (разности координат его концов) преобразуются по тем же линейным формулам, что и сами координаты. Приведем еще пример контравариантного вектора. Положим, что координаты точки (x_1, x_2, x_3) суть функции некоторого параметра t, и определим вектор скорости, имеющий составляющие

$$\left(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \frac{dx_3}{dt}\right).$$

Дифференцируя основные формулы (36) по t, убеждаемся непосредственно в том, что вектор скорости также является контравариантным вектором.

Дадим теперь пример ковариантного вектора. Рассмотрим для этого некоторую функцию точки $f(x_1, x_2, x_3)$ в пространстве и определим в любой координатной системе вектор, называемый *градиентом* этой функции и определяемый следующими составляющими:

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}.$$

Согласно правилу дифференцирования сложных функций и формулам (36), будем иметь:

$$\frac{\partial f}{\partial x_s} = a_{1s}\frac{\partial f}{\partial x_1'} + a_{2s}\frac{\partial f}{\partial x_2'} + a_{3s}\frac{\partial f}{\partial x_3'} \qquad (s = 1, 2, 3),$$

т. е. составляющие градиента в осях (x_1, x_2, x_3) выражаются через составляющие градиента в осях (x_1', x_2', x_3') линейным преобразованием с таблицей $A^{(*)}$ и, следовательно, составляющие в осях (x_1', x_2', x_3') через составляющие в осях (x_1, x_2, x_3) выражаются линейным преобразованием с таблицей $A^{(*)-1} = V$, т. е. градиент функции есть действительно ковариантный вектор.

Нетрудно выразить формулы (37) и (38) через частные производные новых координат по старым и наоборот. Введем обозначения, несколько отличные от предыдущих, которые являются обычными в теории векторов. Будем у составляющих контравариантных векторов приписывать значок сверху, а у ковариантных снизу. В соответствии с этим у самих координат будем писать значок сверху.

Коэффициенты преобразования (36) можно представить в виде частных производных следующим образом:

$$a_{ik} = \frac{\partial x'^{(i)}}{\partial x^{(k)}}. \tag{39}$$

Элементы контраградиентной матрицы V будут:

$$V_{ik} = \frac{A_{ik}}{D(A)},$$

и те же элементы имеет матрица $(A^{-1})^{(*)}$, т. е.

$$(A^{(*)})^{-1} = (A^{-1})^{(*)},$$

т. е. можно сначала перейти к обратной матрице, а потом переставить строки и столбцы. При переходе к обратной матрице коэффициент c_{ik}

будет $\frac{\partial x^{(i)}}{\partial x'^{(k)}}$ и после транспонирования получим для элементов матрицы V выражения:

$$V_{ik}=\frac{\partial x^{(k)}}{\partial x'^{(i)}}. \tag{40}$$

Пусть $u^{(s)}$ — составляющие контравариантного вектора в координатах $x^{(k)}$ и $u'^{(s)}$ — в координатах $x'^{(s)}$. По определению имеем:

$$u'^{(i)}=\sum_{s=1}^{3}\frac{\partial x'^{(i)}}{\partial x^{(s)}}u^{(s)} \quad (i=1, 2, 3). \tag{41}$$

Точно так же для ковариантного вектора по определению получим:

$$u_i'=\sum_{s=1}^{3}\frac{\partial x^{(s)}}{\partial x'^{(i)}}u_s. \tag{42}$$

Заметим, что мы можем пользоваться этими формулами для определения составляющих вектора не только при линейном преобразовании координат, но и в случае самого общего преобразования, когда одни координаты выражаются через другие при помощи любых, вообще нелинейных, функций.

Укажем еще на одну возможность определения ковариантного вектора, причем контравариантный вектор определяется просто как такой вектор, составляющие которого преобразуются по тем же формулам, что и координаты. Итак, пусть мы имеем некоторый контравариантный вектор $u^{(s)}$ и ковариантный вектор v_s.

Составим сумму произведений:

$$u^{(1)}v_1+u^{(2)}v_2+u^{(3)}v_3. \tag{43}$$

Нетрудно видеть, что она остается неизменной, или, как говорят, есть скаляр, если $u^{(s)}$ и v_s изменяются по соответствующим им формулам (41) и (42).

Действительно, будем иметь непосредственно в силу правила дифференцирования сложных функций

$$\sum_{s=1}^{3}u'^{(s)}v_s'=\sum_{s=1}^{3}\left[\sum_{k=1}^{3}\frac{\partial x'^{(s)}}{\partial x^{(k)}}u^{(k)}\right]\left[\sum_{l=1}^{3}\frac{\partial x^{(l)}}{\partial x'^{(s)}}v_l\right]=u^{(1)}v_1+u^{(2)}v_2+u^{(3)}v_3.$$

Таким образом, определив контравариантный вектор указанным выше способом, мы можем определить закон изменения составляющих ковариантного вектора из требования, чтобы выражение (43) оставалось неизменным. Применяя буквально те же вычисления, которыми мы пользовались в предыдущем номере, мы придем к тому, что при инвариантности выражения (43) составляющие v_s должны испытывать линейное преобразование, контрградиентное тому, которое испытывают составляющие $u^{(s)}$. Предоставляем читателю показать, что при любом преобразовании координат (не только линейном) вектор скорости является контравариантным вектором и градиент функции — ковариантным вектором.

В заключение сделаем одно замечание по поводу разницы между контравариантными и ковариантными векторами, которые были определены выше чисто формальным образом — формулами перехода от одной системы к другой. Пусть x — некоторый вектор, заданный своей длиной и направлением. Имея орты, мы образуем составляющие вектора по формуле (28). Назовем эти составляющие сейчас *контравариантными составляющими* и напишем формулу (28) в виде:

$$\boldsymbol{x} = x^{(1)}\boldsymbol{i} + x^{(2)}\boldsymbol{j} + x^{(3)}\boldsymbol{k}. \tag{44}$$

Назовем *ковариантной составляющей* вектора $\boldsymbol{x}$ на орт $\boldsymbol{i}$ величину прямоугольной проекции $\boldsymbol{x}$ на $\boldsymbol{i}$, умноженную на длину $\boldsymbol{i}$, и аналогично для двух других ортов. Таким образом имеем для каждой системы ортов три ковариантные составляющие (x_1, x_2, x_3). Можно показать, что они преобразуются при переходе от одних ортов к другим как составляющие ковариантного вектора. Действительно, можно показать, на чем мы не останавливаемся, что в данном случае выражение $x^{(1)}x_1 + x^{(2)}x_2 + x^{(3)}x_3$ будет давать квадрат длины вектора $\boldsymbol{x}$ и таким образом будет оставаться неизменным при преобразовании ортов.

23. Понятие тензора. Мы переходим теперь к некоторому обобщению понятия вектора, причем сначала будем рассматривать только линейные преобразования координат. Пусть в некоторой координатной системе задана таблица девяти чисел:

$$b_{ik} \quad (i, k = 1, 2, 3).$$

Составим выражение вида:

$$\sum_{i,k=1}^{3} b_{ik} u^{(i)} v^{(k)}, \tag{45}$$

где $u^{(i)}$ и $v^{(k)}$ суть составляющие двух контравариантных векторов. Совершая переход к новым координатам, мы можем в выражении (45) выразить $u^{(i)}$ и $v^{(k)}$ через новые составляющие $u'^{(i)}$ и $v'^{(k)}$ и таким образом преобразуем выражение (45) к следующему виду:

$$\sum_{i,k=1}^{3} b_{ik} u^{(i)} v^{(k)} = \sum_{i,k=1}^{3} b'_{ik} u'^{(i)} v'^{(k)}. \tag{46}$$

Мы будем иметь, таким образом, и в новой координатной системе таблицу девяти чисел с элементами b'_{ik}. Такая таблица, определенная в любой координатной системе из требования инвариантности выражения (45), называется *ковариантным тензором второго ранга*. Точно так же, взяв два ковариантных вектора с составляющими u_i и v_k и образовав выражение

$$\sum_{i,k=1}^{3} b^{(i,k)} u_i v_k, \tag{47}$$

мы, задав в какой-либо координатной системе таблицу девяти чисел $b^{(i,k)}$, будем иметь из требования инвариантности выражения (47) таблицу девяти чисел и в любой другой координатной системе. Это даст нам так называемый *контравариантный тензор второго ранга*. Наконец, взяв один контравариантный вектор с составляющими $u^{(i)}$ и один ковариантный вектор с составляющими v_k и образовав выражение

$$\sum_{i,k=1}^{3} b_i^{(k)} u^{(i)} v_k, \tag{48}$$

мы таким же точно путем придем к понятию *смешанного тензора второго ранга.*

Покажем теперь, каким образом, имея коэффициенты линейного преобразования координат (36), можно составить формулы, выражающие составляющие некоторого тензора в новых координатах через его составляющие в прежних координатах. Остановимся сначала на случае ковариантного тензора второго ранга. Составляющие $u^{(i)}$ и $v^{(k)}$ контравариантных векторов в прежней координатной системе выражаются через составляющие $u'^{(i)}$ и $v'^{(k)}$ в новой координатной системе при помощи линейного преобразования с таблицей A^{-1}. Обозначая элементы этой таблицы через $\{A^{-1}\}_{ik}$, мы будем иметь таким образом:

$$u^{(i)} = \sum_{k=1}^{3} \{A^{-1}\}_{ik} u'^{(k)}; \quad v^{(i)} = \sum_{k=1}^{3} \{A^{-1}\}_{ik} v'^{(k)}.$$

Подставляя в выражение (45) и определяя коэффициент при произведении $u'^{(i)}v'^{(k)}$, мы получим выражение для составляющей b'_{ik} тензора в новой координатной системе:

$$b'_{ik} = \sum_{p,\,q=1}^{3} b_{pq} \{A^{-1}\}_{pi} \{A^{-1}\}_{qk}. \tag{49}$$

Для случая контравариантного тензора второго ранга нам совершенно так же надо будет выразить составляющие ковариантных векторов u_i и v_k через новые составляющие. Согласно определению ковариантного вектора, u'_i выражается через u_i при помощи таблицы $A^{(*)-1}$ и, следовательно, u_i выражается через u'_i при помощи таблицы $A^{(*)}$, транспонированной по отношению к A, и аналогично v_i, т. е.:

$$u_i = \sum_{k=1}^{3} \{A\}_{ki} u'_k; \quad v_i = \sum_{k=1}^{3} \{A\}_{ki} v'_k.$$

Подставляя это в выражение (47), получим формулы преобразования для составляющих контравариантного тензора второго ранга:

$$b'^{(i,k)} = \sum_{p,\,q=1}^{3} b^{(p,\,q)} \{A\}_{ip} \{A\}_{kq}. \tag{50}$$

Точно так же, для составляющих смешанного тензора второго ранга будем иметь следующую формулу преобразования:

$$b'^{(k)}_{i} = \sum_{p,\,q=1}^{3} b^{(q)}_{p} \{A^{-1}\}_{pi} \{A\}_{kq}. \tag{51}$$

Если мы выразим коэффициенты линейного преобразования через частные производные

$$\frac{\partial x'^{(i)}}{\partial x^{(k)}} \text{ и } \frac{\partial x^{(i)}}{\partial x'^{(k)}}$$

и подставим эти выражения в предыдущие формулы, то будем иметь формулы преобразования тензоров второго ранга для случая любого преобразования координат. Совершенно аналогично предыдущему можно определить понятие и о тензоре ранга выше второго, на чем мы останавливаться не будем.

В предыдущем мы все время имели дело с таблицей, выражающей линейное преобразование трехмерного пространства в некоторых координатных осях. Пусть это будет таблица B и положим, что мы совершили афинное преобразование координат по формулам

$$(y_1, y_2, y_3) = A(x_1, x_2, x_3),$$

где A есть некоторая таблица с определителем, отличным от нуля. Как было показано выше, в новых координатах наше преобразование пространства будет иметь таблицу

$$ABA^{-1}.$$

Нетрудно видеть, что такое преобразование таблицы B совпадает с указанным выше преобразованием для смешанного тензора второго ранга. Действительно, применяя правила для перемножения таблиц, будем иметь следующие формулы:

$$\{BA^{-1}\}_{qi} = \sum_{p=1}^{3} \{B\}_{qp} \{A^{-1}\}_{pi},$$

и дальше отсюда

$$\{A(BA^{-1})\}_{ki} = \sum_{q=1}^{3} \{A\}_{kq} \{BA^{-1}\}_{qi} = \sum_{p,\, q=1}^{3} \{B\}_{qp} \{A^{-1}\}_{pi} \{A\}_{kq}.$$

Обозначая $b_k^{(i)}$ вместо $\{B\}_{ik}$, получим как раз формулы вида (51).

Упомянем еще о некоторых тензорах частного вида. Положим, что в некоторой координатной системе ковариантный тензор обладает тем свойством, что

$$b_{ik} = b_{ki} \quad (i,\ k = 1,\ 2,\ 3). \tag{52}$$

Нетрудно видеть, что он будет обладать этим же свойством и в любой другой координатной системе. Действительно, согласно (49):

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{pq} \{A^{-1}\}_{pk} \{A^{-1}\}_{qi}$$

или в силу (52):

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{qp} \{A^{-1}\}_{pk} \{A^{-1}\}_{qi},$$

или, изменяя обозначения переменных суммирования:

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{pq} \{A^{-1}\}_{qk} \{A^{-1}\}_{pi},$$

откуда непосредственно видно, что b'_{ki} действительно совпадает с b'_{ik}. Такой тензор называется *симметричным ковариантным тензором*. Совершенно аналогично определение *симметричного контравариантного тензора*. Точно так же, если в некоторой координатной системе $b_{ik} = -b_{ki}$ или $b^{(i,\,k)} = -b^{(k,\,i)}$, то то же будет иметь место и в любой другой координатной системе, и соответствующие тензоры называются *кососимметрическими*. Для смешанного тензора указанное обстоятельство уже не будет иметь места, так что, например, соотношение $b_i^{(k)} = b_k^{(i)}$ не будет инвариантным при преобразовании координат. Мы переходим теперь к рассмотрению некоторых частных случаев тензоров.

24. Примеры афинных ортогональных тензоров. В дальнейших примерах мы ограничим линейные преобразования координат лишь теми преобразованиями, которые были рассмотрены в [20] и которые соответствуют переходу от одной декартовой системы к другой. Такие преобразования называются обычно ортогональными преобразованиями трехмерного пространства. Для них, как мы видели выше, контраградиентное преобразование $A^{(*)-1}$ совпадает с A и исчезает разница ковариантного и контравариантного вектора. Точно так же для этих преобразований координат будем очевидно иметь одно только понятие тензора второго ранга. Обозначая таблицу коэффициентов ортогонального преобразования координат, как и выше, через $\{A\}_{ik}$, получим для формулы преобразования тензора второго ранга следующую формулу:

$$b'_{ik} = \sum_{p,\,q=1}^{3} b_{pq} \{A\}_{ip} \{A\}_{kq}, \tag{53}$$

которая непосредственно вытекает из формул предыдущего параграфа. Будем толковать элементы каждого столбца таблицы $\| b_{ik} \|$ как составляющие некоторого вектора. Мы имели, таким образом, три вектора:

$$\boldsymbol{b}^{(1)}\,(b_{11}, b_{21}, b_{31});\quad \boldsymbol{b}^{(2)}\,(b_{12}, b_{22}, b_{32});\quad \boldsymbol{b}^{(3)}\,(b_{13}, b_{23}, b_{23}).$$

Будем говорить, что первый из них соответствует оси x_1, второй — оси x_2 и третий — оси x_3. Сопоставим теперь любому направлению (n) вектор $\boldsymbol{b}^{(n)}$ согласно формуле

$$\boldsymbol{b}^{(n)} = \cos(n, x_1)\,\boldsymbol{b}^{(1)} + \cos(n, x_2)\,\boldsymbol{b}^{(2)} + \cos(n, x_3)\,\boldsymbol{b}^{(3)}. \tag{54}$$

Возьмем теперь какие-нибудь декартовы оси (x'_1, x'_2, x'_3) вместо прежних (x_1, x_2, x_3), и составим векторы согласно формуле (54), соответствующие направлениям новых координатных осей:

$$\boldsymbol{b}'^{(k)} = \cos(x'_k, x_1)\,\boldsymbol{b}^{(1)} + \cos(x'_k, x_2)\,\boldsymbol{b}^{(2)} + \cos(x'_k, x_3)\,\boldsymbol{b}^{(3)}. \tag{55}$$

Рассматривая проекции этих векторов на новые координатные оси x'_1, x'_2, x'_3, будем иметь таблицу девяти чисел $\| b'_{ik} \|$, аналогичную таблице $\| b_{ik} \|$. Покажем, что элементы новой таблицы выражаются через элементы таблицы b_{ik} как раз по формулам преобразования составляющих тензоров второго ранга. Действительно, рассмотрим, например, элемент b'_{12}. Согласно определению это есть составляющая вектора $\boldsymbol{b}'^{(2)}$ на новую ось x'_1. Формула (55) дает:

$$\boldsymbol{b}'^{(2)} = \cos(x'_2, x_1)\,\boldsymbol{b}^{(1)} + \cos(x'_2, x_2)\,\boldsymbol{b}^{(2)} + \cos(x'_2, x_3)\,\boldsymbol{b}^{(3)}, \tag{56}$$

откуда видно, что $\boldsymbol{b}'^{(2)}$ есть линейная функция векторов $\boldsymbol{b}^{(i)}$, и, чтобы получить b'_{12}, достаточно в правой части формулы (56) заменить векторы $\boldsymbol{b}^{(i)}$ их проекциями на ось x'_1, т. е. заменить эти векторы следующими выражениями: $\boldsymbol{b}^{(i)}$ — на $b_{1i}\cos(x'_1, x_1) + b_{2i}\cos(x'_1, x_2) + b_{3i}\cos(x'_1, x_3)$ $(i = 1, 2, 3)$. Заметим, кроме того, что, согласно таблице (2): $\cos(x'_i, x_k) = a_{ik} = \{A\}_{ik}$.

Подставляя эти выражения вместо упомянутых векторов в правую часть формулы (56), будем иметь:

$$b'_{12} = \sum_{p,q=1}^{3} b_{pq} \{A\}_{1p} \{A\}_{2q},$$

что как раз и совпадает с формулой (53). Таким образом мы можем утверждать, что если для трех взаимно перпендикулярных направлений определены три вектора $b^{(1)}$, $b^{(2)}$, $b^{(3)}$ и по формуле (54) определен вектор для любого направления (n), то таблица девяти чисел, дающих проекции векторов $b'^{(k)}$ $(k=1, 2, 3)$ на оси $x'^{(k)}$ в любой декартовой координатной системе, определяет афинный ортогональный тензор второго ранга, т. е. тензор второго ранга, определенный для всевозможных ортогональных преобразований.

Заметим, что когда мы говорим, что $b^{(1)}$ соответствует направлению некоторой оси x_1, то это не значит, что $b^{(1)}$ должен иметь направление оси x_1. Существенным является лишь формула (54), которая сопоставляет всякому направлению (n) вектор $b^{(n)}$, направление которого, вообще говоря, не совпадает с (n).

Приведем теперь два примера афинного ортогонального тензора второго ранга. Первый из этих примеров есть известный из теории упругости тензор напряжения. Рассмотрим деформированное упругое тело и проведем в некоторой фиксированной его точке M бесконечно малую площадку $d\sigma$ с нормалью (n). В теории упругости принимают, что воздействие на упомянутую площадку той части упругой среды, которая находится со стороны определяемой направлением нормали, будет равносильно произведению некоторого вектора $b^{(n)}$, зависящего от направления нормали (n), на величину площадки $d\sigma$. Из рассмотрения условий равновесия бесконечно малого тетраэдра, выделенного из упругого тела, получается равенство (54), из которого непосредственно вытекает, что напряжение есть тензор второго ранга. В любой декартовой координатной системе этот тензор будет характеризоваться таблицей девяти чисел $\| b_{ik} \|$, причем, как доказывается в теории упругости, этот тензор будет симметричным, т. е. $b_{ik}=b_{ki}$. Иначе говоря, проекция на ось x_i напряжения, действующего на площадку, перпендикулярную к оси x_k, равна проекции на ось x_k напряжения, действующего на площадку, перпендикулярную к оси x_i.

Перейдем теперь к другому примеру тензора. Рассмотрим некоторое векторное поле $\boldsymbol{C}(M)$. Если выбрать некоторую декартову координатную систему (x_1, x_2, x_3) и взять производные от составляющих поля (c_1, c_2, c_3) по координатам, то мы получим следующую таблицу девяти величин:

$$\left\| \begin{matrix} \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial c_1}{\partial x_2} & \dfrac{\partial c_1}{\partial x_3} \\[2ex] \dfrac{\partial c_2}{\partial x_1} & \dfrac{\partial c_2}{\partial x_2} & \dfrac{\partial c_2}{\partial x_3} \\[2ex] \dfrac{\partial c_3}{\partial x_1} & \dfrac{\partial c_3}{\partial x_2} & \dfrac{\partial c_3}{\partial x_3} \end{matrix} \right\|. \tag{57}$$

Для любого направления (n) определим вектор, соответствующий этому направлению, как производную $\dfrac{\partial c}{\partial n}$, так что элементы таблицы (57), находящиеся в k-м столбце, будут давать составляющие вектора, соответствующего направлению оси x_k. Для любого направления (n) будем иметь формулу [II; 120]:

$$\frac{\partial c_i}{\partial n}=\cos(n, x_1)\frac{\partial c_i}{\partial x_1}+\cos(n, x_2)\frac{\partial c_i}{\partial x_2}+\cos(n, x_3)\frac{\partial c_i}{\partial x_3} \quad (i=1, 2, 3),$$

т. е. определенная нами таблица представляет собою тензор второго ранга. Этот тензор не будет, вообще говоря, ни симметричным, ни антисимметричным. Но нетрудно его представить в виде суммы симметричного и антисимметричного тензора, понимая под суммой двух таблиц сумму соответствующих элементов этих таблиц.

Предварительно сделаем некоторые общие замечания. Из линейного характера формул (53) вытекает, что если $\| b_{ik} \|$ и $\| c_{ik} \|$ будут два тензора, то сумма $\| b_{ik} + c_{ik} \|$ также будет тензором. Кроме того, эта же формула сохраняется при перестановке значков, т. е.

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{qp} \{A\}_{ip} \{A\}_{kq},$$

так что если некоторые таблицы, определенные для всех осей, дают тензор, то и транспонированные таблицы дают тензор. Положим теперь, что мы имеем некоторый тензор $\| b_{ik} \|$. Мы можем представить его в виде суммы

$$\| b_{ik} \| = \left\| \frac{b_{ik} + b_{ki}}{2} \right\| + \left\| \frac{b_{ik} - b_{ki}}{2} \right\|.$$

Первое слагаемое представляет собою, очевидно, симметричный тензор, а второе — антисимметричный тензор.

Применяя это разбиение к тензору, определяемому таблицей (57), получим его симметричную часть в виде:

$$\left\| \begin{matrix} \dfrac{\partial c_1}{\partial x_1}, & \dfrac{1}{2}\left(\dfrac{\partial c_1}{\partial x_2} + \dfrac{\partial c_2}{\partial x_1}\right), & \dfrac{1}{2}\left(\dfrac{\partial c_1}{\partial x_3} + \dfrac{\partial c_3}{\partial x_1}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial c_1}{\partial x_2} + \dfrac{\partial c_2}{\partial x_1}\right), & \dfrac{\partial c_2}{\partial x_2}, & \dfrac{1}{2}\left(\dfrac{\partial c_2}{\partial x_3} + \dfrac{\partial c_3}{\partial x_2}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial c_1}{\partial x_3} + \dfrac{\partial c_3}{\partial x_1}\right), & \dfrac{1}{2}\left(\dfrac{\partial c_2}{\partial x_3} + \dfrac{\partial c_3}{\partial x_2}\right), & \dfrac{\partial c_3}{\partial x_3} \end{matrix} \right\|. \tag{58}$$

Если имеется деформация сплошной среды и $\overline{MC}$ есть вектор смещения, т. е. тот вектор, на который сместилась точка M среды, то таблица (58) определяет так называемый *тензор деформации*. Антисимметрическая часть тензора будет:

$$\left\| \begin{matrix} 0, & \dfrac{1}{2}\left(\dfrac{\partial c_1}{\partial x_2} - \dfrac{\partial c_2}{\partial x_1}\right), & \dfrac{1}{2}\left(\dfrac{\partial c_1}{\partial x_3} - \dfrac{\partial c_3}{\partial x_1}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial c_2}{\partial x_1} - \dfrac{\partial c_1}{\partial x_2}\right), & 0, & \dfrac{1}{2}\left(\dfrac{\partial c_2}{\partial x_3} - \dfrac{\partial c_3}{\partial x_2}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial c_3}{\partial x_1} - \dfrac{\partial c_1}{\partial x_3}\right), & \dfrac{1}{2}\left(\dfrac{\partial c_3}{\partial x_2} - \dfrac{\partial c_2}{\partial x_3}\right), & 0 \end{matrix} \right\|. \tag{59}$$

Мы уже раньше производили разбиение тензора на две части для частного случая линейной однородной деформации [II, 125] и видели, что в этом случае антисимметрической части соответствовало вращение пространства, как целого (без деформации), вокруг некоторой оси.

25. Случай *n*-мерного комплексного пространства. Обратимся сейчас к общему случаю n-мерного пространства. Вектором в таком пространстве мы раньше уже назвали последовательность n чисел, вещественных или комплексных [12]:

$$\boldsymbol{x}(x_1, x_2, \ldots, x_n),$$

причем эти числа называются составляющими вектора $\boldsymbol{x}$. Мы считаем при этом, что пространство отнесено к определенным ортам:

$$\boldsymbol{a}^{(1)}(1, 0, \ldots, 0);\ \boldsymbol{a}^{(2)}(0, 1, \ldots, 0);\ \ldots;\ \boldsymbol{a}^{(n)}(0, 0, \ldots, 1),$$

так что

$$x = x_1 a^{(1)} + x_2 a^{(2)} + \ldots + x_n a^{(n)}. \tag{60}$$

Условие равенства векторов и простейшие действия над ними были определены в [**12**].

Линейным преобразованием n-мерного пространства назовем переход от вектора $x(x_1, x_2, \ldots, x_n)$ к вектору $y(y_1, y_2, \ldots, y_n)$ по формулам:

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \qquad (i = 1, 2, \ldots, n), \tag{61}$$

или иначе

$$y = Ax, \tag{62}$$

где A есть таблица или матрица $\| a_{ik} \|_1^n$ преобразования. Если ее определитель $D(A)$ отличен от нуля, то преобразование (62) называется неособым преобразованием, а матрица A — неособой матрицей (таблицей). В этом случае, решая уравнение (61) относительно x_i, получим преобразование, обратное (61) или (62):

$$x = A^{-1}y, \tag{63}$$

где таблица A^{-1} имеет элементы

$$\{A^{-1}\}_{ik} = \frac{A_{ki}}{D(A)}, \tag{64}$$

причем через $D(A)$ мы обозначаем определитель таблицы A и через A_{ik} — алгебраические дополнения его относительно элементов a_{ik}.

Дальше, аналогично предыдущему [**21**], определяется произведение двух преобразований, а именно последовательное применение двух преобразований

$$y = Ax; \quad z = By$$

равносильно одному линейному преобразованию

$$z = BAx,$$

которое называется *произведением* преобразований A и B и таблица которого определяется по формуле:

$$\{BA\}_{ik} = \sum_{s=1}^{n} \{B\}_{is} \{A\}_{sk}. \tag{65}$$

Это произведение зависит, вообще говоря, от порядка сомножителей, т. е., кроме исключительных случаев, мы имеем

$$BA \neq AB.$$

Нетрудно распространить определение произведения на случай любого числа сомножителей, причем имеет место сочетательный закон, т. е. сомножители можно соединять в группы:

$$(CB)A = C(BA). \tag{66}$$

Обратное преобразование удовлетворяет соотношениям

$$AA^{-1}=A^{-1}A=I; \quad (A^{-1})^{-1}=A, \tag{67}$$

где символом I мы обозначили так называемую единичную матрицу, у которой элементы, стоящие на главной диагонали, равны единице, а остальные элементы — нулю. Этой матрице соответствует тождественное преобразование $y_i=x_i$ $(i=1, 2, \ldots, n)$.

Так же, как и выше, определим диагональную матрицу порядка n:

$$[k_1, k_2, \ldots, k_n]=\begin{Vmatrix} k_1 & 0 & 0 & \ldots & 0 \\ 0 & k_2 & 0 & \ldots & 0 \\ 0 & 0 & k_3 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & k_n \end{Vmatrix}. \tag{68}$$

Ей соответствует преобразование: $y_i=k_ix$ $(i=1, 2, \ldots, n)$.

Произведение диагональных матриц не зависит от порядка сомножителей и определяется по формуле

$$[k_1, k_2, \ldots, k_n][l_1, l_2, \ldots, l_n]=[l_1, l_2, \ldots, l_n][k_1, k_2, \ldots, k_n]= \\ =[k_1l_1, k_2l_2, \ldots, k_nl_n].$$

В частном случае $k_1=k_2=\ldots=k_n=k$ мы получим матрицу

$$[k, k, \ldots, k]=\begin{Vmatrix} k & 0 & 0 & \ldots & 0 \\ 0 & k & 0 & \ldots & 0 \\ 0 & 0 & k & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & k \end{Vmatrix}, \tag{69}$$

которой будет соответствовать умножение всех составляющих вектора на число k. В соответствии со сказанным в начале настоящего параграфа мы будем считать, что матрица (69) является просто числом k, т. е. будем считать число k частным случаем матрицы. Нетрудно видеть, пользуясь формулой (65), что произведение числа k, трактуемого как матрица (69), на любую матрицу A не зависит от порядка сомножителей и сводится к умножению всех элементов матрицы A на число k:

$$\{[k, k, \ldots, k]A\}_{ik}=\{kA\}_{ik}=k\{A\}_{ik}. \tag{70}$$

Положим теперь, что мы взяли за основные орты не векторы $\boldsymbol{a}^{(k)}$, а новые векторы $\boldsymbol{b}^{(k)}$, которые выражаются через $\boldsymbol{a}^{(k)}$ по формулам:

$$\left.\begin{array}{l} \boldsymbol{b}^{(1)} = t_{11}\boldsymbol{a}^{(1)} + t_{12}\boldsymbol{a}^{(2)} + \ldots + t_{1n}\boldsymbol{a}^{(n)} \\ \boldsymbol{b}^{(2)} = t_{21}\boldsymbol{a}^{(1)} + t_{22}\boldsymbol{a}^{(2)} + \ldots + t_{2n}\boldsymbol{a}^{(n)} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \boldsymbol{b}^{(n)} = t_{n1}\boldsymbol{a}^{(1)} + t_{n2}\boldsymbol{a}^{(2)} + \ldots + t_{nn}\boldsymbol{a}^{(n)} \end{array}\right\}, \tag{71}$$

причем определитель, составленный из элементов t_{ik}, не равен нулю. При этом векторы $\boldsymbol{a}^{(k)}$, наоборот, выражаются линейно через векторы $\boldsymbol{b}^{(k)}$, и всякая линейная комбинация векторов $\boldsymbol{a}^{(k)}$ есть в то же самое время линейная комбинация векторов $\boldsymbol{b}^{(k)}$ и наоборот. Иначе говоря, векторы $\boldsymbol{b}^{(k)}$, как орты, образуют то же пространство, что и векторы $\boldsymbol{a}^{(k)}$. Если некоторый вектор $\boldsymbol{x}$ в системе координат, определяемой ортами $\boldsymbol{a}^{(k)}$, имеет составляющие $(x_1, \ldots, x_n)$, то в системе координат, определяемой ортами $\boldsymbol{b}_k$, он будет иметь другие составляющие $(x'_1, \ldots, x'_n)$, которые выражаются через прежние при помощи линейного преобразования, контраградиентного преобразованию (71), что можно записать так:

$$(x'_1, \ldots, x'_n) = T^{(*)-1}(x_1, \ldots, x_n), \tag{72}$$

где таблица $T^{(*)}$ есть транспонированная таблица по отношению к таблице T, соответствующей преобразованию (71).

Если мы имели некоторое преобразование пространства, которое в первоначальной координатной системе выражалось формулой (62), то в новой координатной системе это же преобразование будет выражаться формулой

$$\boldsymbol{y}' = UAU^{-1}\boldsymbol{x}', \tag{73}$$

где

$$U = T^{(*)-1}.$$

Матрица

$$UAU^{-1}$$

называется *подобной* матрице A.

Основными понятиями в предыдущем изложении являлись понятия вектора и матрицы. Заметим, что иногда рассматривают вектор $\boldsymbol{x}(x_1, \ldots, x_n)$ тоже как матрицу, один из столбцов которой, безразлично какой, заполнен числами $(x_1, \ldots, x_n)$, а остальные элементы матрицы равны нулю. Положим, например, что мы ставим составляющие вектора в первый столбец. Таким образом мы будем иметь представление нашего вектора в виде матрицы:

$$\left\| \begin{array}{cccc} x_1 & 0 & \ldots & 0 \\ x_2 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ x_n & 0 & \ldots & 0 \end{array} \right\|.$$

Иногда такую матрицу, у которой только один столбец содержит элементы, отличные от нуля, обозначают символом

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{Vmatrix} x_1 & 0 & \dots & 0 \\ x_2 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ x_n & 0 & \dots & 0 \end{Vmatrix}. \tag{74}$$

Покажем теперь, что линейное преобразование (62) может быть записано в виде произведения матрицы (74) на матрицу преобразования A. Действительно, перемножая матрицу (74) на матрицу A по правилу (65) и принимая во внимание, что у матрицы (74) только элементы первого столбца отличны от нуля, мы получим произведение в виде матрицы, у которой тоже только элементы первого столбца будут отличными от нуля, и эти элементы, как нетрудно видеть, будут

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n,$$

т. е. они дают как раз линейное преобразование (62). Мы можем таким образом записать это преобразование в виде:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \tag{75}$$

где справа стоит произведение двух матриц.

В заключение настоящего номера отметим еще раз общие законы, которым подчиняются действия с векторами в n-мерном пространстве

$$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}; \qquad (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}).$$

Если $\boldsymbol{x}$ и $\boldsymbol{y}$ — какие-нибудь два вектора, то вектор $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{x}$ с составляющими $(y_k - x_k)$ является единственным вектором, удовлетворяющим условию $\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{y}$.

Пусть a и b — какие-нибудь числа. Мы имеем:

$$(a + b)\,\boldsymbol{x} = a\boldsymbol{x} + b\boldsymbol{x}; \qquad a\,(b\boldsymbol{x}) = (ab)\,\boldsymbol{x}; \qquad a\,(\boldsymbol{x} + \boldsymbol{y}) = a\boldsymbol{x} + a\boldsymbol{y}.$$

Для числа единица мы имеем $1\boldsymbol{x} = \boldsymbol{x}$ и $0\boldsymbol{x} = 0$, где нуль, стоящий справа, обозначает вектор, у которого все составляющие равны нулю.

26. Основы матричного исчисления. В формулах, приведенных в предыдущем номере, матрица входила в качестве нового символа, над которым мы могли производить и некоторые действия, аналогичные действиям над обычными числами. Это приводит нас к естественной мысли построить новую алгебру, которая годилась бы для символов, под которыми мы подразумеваем матрицы. Иначе говоря, мы будем толковать матрицу как новый вид числа, как некоторое *гиперкомплексное* число. Совершенно так же, как с помощью двух вещественных чисел мы пришли выше к построению чисел новой природы, а именно комплексных чисел вида $a+ib$, так и теперь мы с помощью n^2 комплексных чисел a_{ik}, расставленных в виде квадратной таблицы, приходим к понятию нового числа — *матрицы*. Но только надо отметить их существенную разницу. А именно, мы видели, что над буквами, изображающими комплексные числа, можно производить все формальные операции алгебры, известные для вещественных чисел. Для матриц мы получим алгебру, существенно отличную от известной нам алгебры комплексных чисел. Существенным моментом, который вызывает это отличие, является *некоммутативность умножения*, т. е. зависимость результата умножения от порядка сомножителей. Мы переходим сейчас к установлению основных правил алгебры матриц, причем во многих отношениях руководящим путем для нас будут служить те результаты, которые мы получили выше, толкуя матрицу как таблицу линейного преобразования.

Везде в дальнейшем, если не будет оговорено особо, мы будем рассматривать квадратные матрицы одного и того же порядка n. Если A — такая матрица, то, как и выше, ее элементы будем обозначать через $\{A\}_{ik}$.

Две матрицы A и B считаются равными тогда и только тогда, когда

$$\{A\}_{ik}=\{B\}_{ik} \qquad (i,\ k=1,\ 2,\ldots,\ n), \tag{76}$$

т. е. когда все их соответствующие элементы одинаковы.

Сложение матриц определяется по формуле:

$$\{A+B\}_{ik}=\{A\}_{ik}+\{B\}_{ik}, \tag{77}$$

т. е. сводится к сложению соответствующих элементов.

Произведение определяется по формуле:

$$\{BA\}_{ik}=\sum_{s=1}^{n}\{B\}_{is}\{A\}_{sk}. \tag{78}$$

Как мы выше видели, вообще говоря, $BA \neq AB$, но имеет место сочетательный закон [21]:

$$(CB)\,A=C\,(BA). \tag{79}$$

Определитель произведения равен произведению определителей перемножаемых матриц:

$$D\,(BA)=D\,(B)\cdot D\,(A). \tag{80}$$

Имеет место, очевидно, и распределительный закон:

$$(A+B)C=AC+BC \text{ и } C(A+B)=CA+CB. \tag{81}$$

Отметим еще одну особенность умножения, а именно произведение матриц может обращаться в нуль, т. е. в матрицу, у которой все элементы равны нулю, хотя все сомножители и отличны от нуля. В качестве примера приведем произведение двух одинаковых матриц второго порядка

$$\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}.$$

Совершенно так, как было указано в предыдущем номере, вводится понятие обратной матрицы A^{-1}, если A — неособая матрица, т. е. если $D(A)\neq 0$. Если $C=BA$ и R_A, R_B и R_C — ранги матриц A, B, C, то, как мы видели, $R_C\leqslant R_A$ [7]. Если B — неособая матрица, то $A=B^{-1}C$, и, как и выше, можно утверждать, что $R_A\leqslant R_C$ и, следовательно, $R_C=R_A$, т. е. *при умножении матрицы A на неособую матрицу B (справа или слева) ее ранг не меняется.* Для единичной матрицы I имеет место соотношение

$$BI=IB=B, \tag{82}$$

где B — любая матрица.

Нетрудно видеть, что матрица A^{-1} является единственным решением уравнений

$$AX=I \text{ и } XA=I, \tag{83}$$

где I — единичная матрица. Действительно, умножая, например, первое из этих равенств слева на A^{-1} и принимая во внимание (79) и (67), получим $X=A^{-1}$ и аналогично для второго уравнения. Отметим, что если $D(A)=0$, то уравнения (83) вовсе не имеют решений, т. е. матрица A не имеет обратной. Действительно, как следствие уравнений (83), имеем: $D(A)D(X)=1$, что противоречит условию $D(A)=0$.

Напомним также понятие диагональной матрицы, введенное в предыдущем номере, а также того, что всякое число k можно рассматривать как частный случай матрицы. Нетрудно ввести понятие целой положительной степени матрицы

$$A^p=A\cdot A\cdots A.$$

Целые отрицательные степени матрицы вводятся как целые положительные степени обратной матрицы, т. е.

$$A^{-p}=(A^{-1})^p. \tag{84}$$

Имеем, очевидно:

$$A^{-p}=(A^p)^{-1}, \text{ т. е. } A^{-p}A^p=A^pA^{-p}=I. \tag{85}$$

Символ частного двух матриц

$$\frac{A}{B}$$

не имеет определенного смысла. Мы можем толковать его двояко — или как произведение AB^{-1}, или как произведение $B^{-1}A$, причем эти два произведения, вообще говоря, различны, и только в тех частных случаях, когда они совпадают, символ частного имеет определенный смысл.

Далее, основным понятием является понятие подобных матриц, которое мы также ввели в предыдущем номере. Отметим некоторые очевидные формулы

$$(CBA)^{-1} = A^{-1}B^{-1}C^{-1}, \tag{86}$$

$$CBAC^{-1} = (CBC^{-1})(CAC^{-1}). \tag{87}$$

Если через $A^{(*)}$ обозначить матрицу, транспонированную по отношению к A, то имеет место формула:

$$(CBA)^{(*)} = A^{(*)}B^{(*)}C^{(*)}, \tag{88}$$

которую нетрудно проверить, пользуясь определением умножения. Введем два обозначения. Обозначим через $\overline{A}$ матрицу, элементы которой суть числа, сопряженные с элементами матрицы A, т. е.

$$\{\overline{A}\}_{ik} = \overline{\{A\}_{ik}}, \tag{89}$$

причем символом $\overline{\alpha}$ обозначаем комплексное число, сопряженное с α, и через $\tilde{A}$ матрицу, которая получается из A, если строки заменить столбцами и все элементы — сопряженными числами, т. е.

$$\{\tilde{A}\}_{ik} = \overline{\{A\}_{ki}}. \tag{90}$$

Матрица $\tilde{A}$ называется иногда сопряженной или эрмитовски сопряженной с матрицей A[1]). Нетрудно проверить формулу

$$\widetilde{CBA} = \tilde{A}\tilde{B}\tilde{C}. \tag{91}$$

Предлагаем проверить также следующую элементарную формулу:

$$(A^{(*)})^{-1} = A^{-1(*)},$$

т. е. что знак обратной матрицы и транспонирования можно менять местами, о чем мы уже упоминали и выше [**20**].

Отметим еще одну полезную в дальнейшем формулу. Из соотношения (67) непосредственно вытекает

$$D(A)D(A^{-1}) = 1,$$

[1]) По имени французского математика второй половины XIX века — Ш. Эрмита.

т. е.

$$D(A^{-1}) = D(A)^{-1}. \tag{92}$$

Иначе говоря, определитель обратной матрицы обратен по величине определителю основной матрицы.

Введем еще понятие *квазидиагональной матрицы*, являющейся обобщением диагональной матрицы. Выясним это понятие сначала на частном случае. Пусть имеется матрица седьмого порядка вида:

$$\begin{Vmatrix} b_{11} & b_{12} & b_{13} & 0 & 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 0 & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{11} & c_{12} & 0 & 0 \\ 0 & 0 & 0 & c_{21} & c_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & d_{11} & d_{12} \\ 0 & 0 & 0 & 0 & 0 & d_{21} & d_{22} \end{Vmatrix}.$$

Обозначим через B матрицу третьего порядка с элементами b_{ik}, а через C и D — матрицы второго порядка с элементами c_{ik} и d_{ik}. Предыдущая матрица седьмого порядка называется *квазидиагональной* структуры $\{3, 2, 2\}$ и обозначается символом

$$[B, C, D].$$

Положим вообще, что главная диагональ матрицы порядка n, состоящая из элементов a_{ii}, разбита на m частей, причем первая часть состоит из первых k_1 элементов, вторая из следующих k_2 элементов и так далее, так что $k_1 + \ldots + k_m = n$. Мы можем рассматривать первые k_1 элементов как главную диагональ некоторой матрицы X_1 порядка k_1; следующие k_2 элементов — как главную диагональ некоторой матрицы X_2 порядка k_2 и так далее. Положим, что все элементы нашей матрицы A, не принадлежащие к матрицам X_s, равны нулю. При этом матрица A называется *квазидиагональной матрицей структуры* $\{k_1, \ldots, k_m\}$ и обозначается следующим образом:

$$A = [X_1, X_2, \ldots, X_m].$$

Правила действия с квазидиагональными матрицами одинаковой структуры отличаются большой простотой. Мы приведем соответствующие формулы, не останавливаясь на их доказательстве. Оно может быть проведено совершенно элементарно на основе определения действий. Для сложения квазидиагональных матриц одинаковой структуры мы имеем формулу

$$[X_1, X_2, \ldots, X_m] + [Y_1, Y_2, \ldots, Y_m] = \\ = [X_1 + Y_1, X_2 + Y_2, \ldots, X_m + Y_m], \tag{93}$$

причем одинаковость структуры равносильна тому, что порядок всякой матрицы X_k совпадает с порядком соответствующей матрицы Y_k.

Точно так же для умножения и возвышения в степень имеем:

$$[Y_1, Y_2, \ldots, Y_m][X_1, X_2, \ldots, X_m] = [Y_1X_1, Y_2X_2, \ldots, Y_mX_m], \tag{94}$$

$$[X_1, X_2, \ldots, X_m]^p = [X_1^p, X_2^p, \ldots, X_m^p], \tag{95}$$

где p — любое целое положительное или отрицательное число, причем, если p есть целое отрицательное число, то необходимо, конечно, требовать, чтобы определители $D(X_k)$ были отличны от нуля.

Правило подобного преобразования матрицы $[X_1, X_2, \ldots, X_m]$ при помощи матрицы той же структуры выражается формулой:

$$[Y_1, \ldots, Y_m][X_1, \ldots, X_m][Y_1, \ldots, Y_m]^{-1} = \\ = [Y_1X_1Y_1^{-1}, \ldots, Y_mX_mY_m^{-1}]. \tag{96}$$

Отметим геометрический смысл тех линейных преобразований которые доставляются квазидиагональными матрицами. Рассмотрим для простоты случай указанной выше квазидиагональной матрицы седьмого порядка, структуры $\{3, 2, 2\}$. Рассмотрим линейное преобразование, соответствующее этой матрице. Если в первоначальном векторе $(x_1, \ldots, x_7)$ мы имеем: $x_4 = x_5 = x_6 = x_7 = 0$, то и в преобразованном, очевидно, будет: $y_4 = y_5 = y_6 = y_7 = 0$, т. е., иными словами, все векторы, которые принадлежат подпространству, образованному тремя первыми основными ортами, и после преобразования будут принадлежать этому подпространству, и само преобразование будет определяться матрицей третьего порядка B. То же относится к подпространству, образованному следующими двумя ортами, и, наконец, к подпространству, образованному последними двумя ортами.

Напомним при этом, что *подпространством, образованным векторами* $x^{(1)}, \ldots, x^{(l)}$, мы называем совокупность векторов, определяемых формулой

$$c_1x^{(1)} + \ldots + c_lx^{(l)},$$

где $c_1, \ldots, c_l$ — произвольные постоянные.

27. Характеристические числа матриц и приведение матриц к каноническому виду. Подобные матрицы не равны, конечно, между собою в смысле (76), но в геометрическом смысле равносильны в том отношении, что они осуществляют одно и то же линейное преобразование пространства, но выраженное в различных координатных системах. Мы займемся сейчас разысканием инвариантов этих матриц, т. е. таких выражений, составленных из элементов, которые имели бы одинаковое значение для всех подобных матриц. Нетрудно составить один из таких инвариантов. Это будет определитель матрицы. Действительно, пусть A — некоторая матрица и UAU^{-1} — ей подобная, причем U — любая матрица с определителем, отличным от нуля. Мы имеем в силу (80) и (92):

$$D(UAU^{-1}) = D(U)\,D(A)\,D(U^{-1}) = D(U)\,D(A)\,D(U)^{-1} = D(A).$$

Чтобы построить еще другие инварианты, образуем полином $\varphi(\lambda)$ степени n от некоторого параметра λ, равный определителю матрицы, которая получается из матрицы A, если мы вычтем из всех ее диагональных членов параметр λ, т. е. положим

$$\varphi(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix}, \tag{97}$$

где a_{ik} — элементы матрицы A. Иначе мы можем записать это так:

$$\varphi(\lambda) = D(A-\lambda) = D(A-\lambda I), \tag{98}$$

ибо по условию λ или λI есть диагональная матрица, у которой все элементы, стоящие на главной диагонали, равны λ. Подставляя UAU^{-1} вместо A и принимая во внимание, что любая матрица переместительна с числом λ и, следовательно, $U\lambda U^{-1} = \lambda$, будем иметь:

$$D(UAU^{-1}-\lambda) = D[U(A-\lambda)\,U^{-1}] = D(A-\lambda),$$

т. е.

$$D(UAU^{-1}-\lambda) = D(A-\lambda). \tag{99}$$

Мы видим, таким образом, что полином (97), составленный для матрицы UAU^{-1}, совпадает с таким же полиномом, составленным для матрицы A. Иными словами, все коэффициенты полинома (97) суть инварианты по отношению к подобным матрицам. Старший коэффициент написанного полинома, как легко видеть, равен $(-1)^n$. Отметим особо два его коэффициента, а именно свободный член и коэффициент при $(-1)^{n-1}\lambda^{n-1}$. Первый из них равен, очевидно, определителю, и этот инвариант мы уже отметили и раньше. Что же касается коэффициента при $(-1)^{n-1}\lambda^{n-1}$, то, пользуясь результатами [**5**], мы видим, что он совпадает с суммой диагональных элементов. Эта сумма называется обычно *следом матрицы* и обозначается следующим образом:

$$\mathrm{Sp}(A) = \{A\}_{11} + \{A\}_{22} + \dots + \{A\}_{nn} = a_{11} + a_{22} + \dots + a_{nn},$$

где Sp есть начальные буквы немецкого слова «Spur», что значит по-русски «след» (французское «trace»). Итак, подобные матрицы имеют одинаковый определитель и одинаковый след.

Напишем теперь уравнение

$$D(A-\lambda) = 0, \tag{100}$$

которое называется *характеристическим уравнением матрицы* A, а его корни — *характеристическими числами, или собственными значениями матрицы* A. Согласно предыдущему, мы можем сказать, что подобные матрицы имеют одинаковые характеристические числа. Мы уже имели уравнение вида (100) раньше.

Поставим теперь следующий вопрос: нельзя ли найти матрицу V, совершающую преобразование подобия от заданной матрицы A к матрице $V^{-1}AV$ так, чтобы последняя матрица была диагональной матрицей. Иными словами, с точки зрения линейных преобразований пространства, нельзя ли выбрать такие координатные оси, в которых линейное преобразование, характеризуемое в первоначальной системе координат матрицей A, сводилось бы просто в новых осях к преобразованию вида $y_k = \lambda_k x_k$. Заметим, что мы пишем подобную матрицу в виде $V^{-1}AV$ вместо прежнего вида UAU^{-1}, что не имеет, конечно, существенного значения.

Мы можем записать наше условие в виде:

$$V^{-1}AV = [\lambda_1, \lambda_2, \ldots, \lambda_n], \tag{101}$$

где искомыми являются элементы матрицы V и числа λ_k. Можно, очевидно, переписать это условие, умножая обе части слева на V, в виде:

$$AV = V[\lambda_1, \lambda_2, \ldots, \lambda_n]. \tag{102}$$

Определим по формуле (65) элементы обеих частей написанного равенства со значками i и k. Таким образом мы получим n^2 уравнений:

$$\sum_{s=1}^{n} a_{is} v_{sk} = v_{ik} \lambda_k,$$

где a_{ik} и v_{ik} — элементы матриц A и V.

Фиксируя второй значок k и полагая $i = 1, 2, \ldots, n$, получим n уравнений, содержащих только элементы $v_{1k}, \ldots, v_{nk}$ столбца матрицы V с номером k и число λ_k:

$$\sum_{s=1}^{n} a_{is} v_{sk} = \lambda_k v_{ik} \qquad (i = 1, 2, \ldots, n). \tag{103}$$

Если мы будем считать элементы $(v_{1k}, \ldots, v_{nk})$ как составляющие некоторого вектора $\boldsymbol{v}^{(k)}$, то можем записать предыдущее равенство в виде одного векторного равенства:

$$A\boldsymbol{v}^{(k)} = \lambda_k \boldsymbol{v}^{(k)}. \tag{104}$$

Мы видим таким образом, что разыскание матрицы V, которая приводит матрицу A к диагональной форме, сводится к разысканию таких векторов $\boldsymbol{v}^{(k)}$, которые воспроизводятся с точностью до численного множителя в результате линейного преобразования, определяемого матрицей A. Этот факт является алгебраическим аналогом того факта современной квантовой механики, согласно которому матричная механика Гейзенберга по существу равносильна волновой механике Шрёдингера. Согласно первой точке зрения, существенным вопросом является задача приведения некоторой матрицы (бесконечной) к диагональной форме. Что же касается волновой механики, то здесь основным вопросом является задача отыскания таких векторов (в пространстве с бесчисленным множеством измерений), которые бы воспроизводились с точностью до численного множителя в результате некоторого линейного преобразования. Предыдущие соображения мы назвали алгебраическим аналогом потому, что ограничиваясь пространством с конечным числом измерений, мы приводим наши задачи к чисто алгебраическим задачам. В более же сложных случаях пространства с бесчисленным множеством измерений мы существенным образом выходим из рамок обычной алгебры и нуждаемся в аппарате анализа. Более подробно все эти вопросы будут выяснены впоследствии, причем заметим, что для приложений к физике в рассматриваемом случае конечного числа измерений достаточно ограничиться матрицами A частного типа (эрмитовские матрицы, у которых $a_{ki} = \bar{a}_{ik}$) и, кроме того, матрица U также должна быть определенного типа (унитарная матрица, определение которой будет дано ниже). Здесь мы рассматриваем общий вопрос для любой конечной матрицы, причем ограничимся лишь приведением окончательных результатов, не приводя полностью доказательство. Для тех задач, которые будут интересны в приложениях, вопрос будет решен полностью.

Переходим к решению системы (103) или (104). В раскрытом виде мы можем записать ее так:

$$\left.\begin{aligned}
&(a_{11} - \lambda_k) v_{1k} + a_{12} v_{2k} + \ldots + a_{1n} v_{nk} = 0\\
&a_{21} v_{1k} + (a_{22} - \lambda_k) v_{2k} + \ldots + a_{2n} v_{nk} = 0\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{n1} v_{1k} + a_{n2} v_{2k} + \ldots + (a_{nn} - \lambda_k) v_{nk} = 0.
\end{aligned}\right\} \qquad (105)$$

Для того чтобы получить для $(v_{1k}, \ldots, v_{nk})$ решение, отличное от нулевого, необходимо и достаточно, чтобы определитель написанной системы был равен нулю, т. е. необходимо и достаточно, чтобы число λ_k было корнем характеристического уравнения. Разберем подробно лишь тот случай, когда это уравнение имеет различные корни. Обозначим эти корни через $\lambda_1, \ldots, \lambda_n$. Подставляя в коэффициенты системы (105) вместо λ_k первый корень λ_1, мы сможем определить из нее элементы первого столбца таблицы V, причем не рассматриваем вопроса о том, насколько широк будет выбор этих величин v_{i1}. Выберем решение системы каким-нибудь одним определенным образом, лишь так, чтобы оно было отлично от нулевого.

Точно так же, заменяя в коэффициентах системы (105) $\lambda_k = \lambda_2$, мы сможем определить элементы второго столбца матрицы V и т. д. до n-го столбца. Равенства (105) равносильны (102), и чтобы перейти к основному равенству (101), нам надо только, чтобы V имела обратную матрицу V^{-1}, т. е. чтобы определитель V был отличен от нуля. Будем доказывать это от противного. Положим, что он равен нулю. Как мы знаем [12], это равносильно тому, что между векторами $\boldsymbol{v}^{(k)}$, определяемыми столбцами матрицы V, существует линейное соотношение:

$$C_1\boldsymbol{v}^{(1)} + \ldots + C_n\boldsymbol{v}^{(n)} = 0,$$

где не все коэффициенты C_k равны нулю. Применим к обеим частям этого равенства $(n-1)$ раз преобразование, определяемое матрицей A. Пользуясь (104), будем иметь n равенств:

$$\begin{gathered} C_1\boldsymbol{v}^{(1)} + \ldots + C_n\boldsymbol{v}^{(n)} = 0 \\ \lambda_1 C_1\boldsymbol{v}^{(1)} + \ldots + \lambda_n C_n\boldsymbol{v}^{(n)} = 0 \\ \cdots\cdots\cdots\cdots\cdots \\ \lambda_1^{n-1} C_1\boldsymbol{v}^{(1)} + \ldots + \lambda_n^{n-1} C_n\boldsymbol{v}^{(n)} = 0. \end{gathered}$$

Принимая во внимание, что не все векторы $C_k\boldsymbol{v}^{(k)}$ равны нулю, можем утверждать, что определитель написанной системы должен равняться нулю:

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \cdot & \cdot & \cdot & \cdot \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{vmatrix} = 0,$$

где числа λ_k по условию различны. Но последнее равенство противоречит тому, что определитель Вандермонда от неравных чисел отличен от нуля. Таким образом мы доказали возможность приведения матрицы преобразованием подобия к диагональной форме для того случая, когда все характеристические числа матрицы различны. В том случае, когда среди характеристических чисел имеются равные, может случиться, что матрица не может быть приведена преобразованием подобия к диагональной форме. Все же и в этом случае существует наиболее простое или, как говорят, каноническое представление матрицы. Это каноническое представление в том случае, когда матрица приводится к диагональной форме, имеет вид:

$$[\lambda_1, \lambda_2, \ldots, \lambda_n],$$

где λ_k — характеристические числа матрицы. Для общего случая сформулируем лишь результат [1]). Пусть $\lambda = a$ является корнем уравнения

[1]) Его доказательство приводится в специальном дополнении в конце второй части этого тома.

(100) кратности k. Положим далее, что для всех определителей порядка $(n-1)$ таблицы, стоящей в левой части уравнения (100), число $\lambda=a$ является корнем кратности k_1, но не выше, т. е. все эти определители делятся на $(\lambda-a)^{k_1}$, но хоть один из них не делится на $(\lambda-a)^{k_1+1}$. Положим далее, что все определители упомянутой таблицы порядка $(n-2)$ имеют $\lambda=a$ корнем кратности k_2, но не выше, и так дальше, и, наконец, все определители порядка $(n-m)$ имеют упомянутый корень кратности k_m, а хоть один из определителей порядка $(n-m-1)$ уже вовсе не обращается в нуль при $\lambda=a$. То же самое будет, очевидно, иметь место и для определителей более низкого порядка. Можно доказать, что числа k_s убывают: $k>k_1>k_2>\ldots>k_m$. Введем следующие положительные целые числа:

$$l_1=k-k_1;\quad l_2=k_1-k_2;\quad \ldots;\quad l_{m+1}=k_m,$$

причем, очевидно, $l_1+l_2+\ldots+l_{m+1}=k$.

Биномы:

$$(\lambda-a)^{l_1};\quad (\lambda-a)^{l_2};\quad \ldots;\quad (\lambda-a)^{l_{m+1}}$$

называются *элементарными делителями матрицы* A, соответствующими корню $\lambda=a$. Мы можем таким образом определить элементарные делители для всех характеристических чисел матрицы A, в результате чего получим *совокупность элементарных делителей:*

$$(\lambda-\lambda_1)^{\rho_1};\quad (\lambda-\lambda_2)^{\rho_2};\quad \ldots;\quad (\lambda-\lambda_p)^{\rho_p}, \tag{106}$$

где

$$\rho_1+\rho_2+\ldots+\rho_p=n \tag{107}$$

и среди чисел λ_k могут быть и одинаковые.

Выше мы видели, что характеристические числа не меняются при преобразовании подобия. Оказывается, что таким же свойством обладает и совокупность элементарных делителей матрицы. Введем теперь некоторые новые, простейшие матрицы $I_\rho(a)$. Мы подразумеваем под этим матрицу порядка ρ, у которой на главной диагонали везде стоит число a, на следующей нижестоящей диагонали стоит везде единица, а остальные элементы равны нулю:

$$I_\rho(a)=\begin{vmatrix} a & 0 & 0 & \ldots & 0 & 0 \\ 1 & a & 0 & \ldots & 0 & 0 \\ 0 & 1 & a & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & a & 0 \\ 0 & 0 & 0 & \ldots & 1 & a \end{vmatrix}. \tag{108}$$

Основным фактом в вопросе о представлении матриц в канонической форме является следующий результат: если матрица A имеет элементарные делители (106), то существует такая матрица U с определителем, отличным от нуля, что

$$UAU^{-1} = [I_{\rho_1}(\lambda_1),\ I_{\rho_2}(\lambda_2),\ \ldots,\ I_{\rho_p}(\lambda_p)]. \tag{109}$$

Заметим, что если известны все характеристические числа матрицы A, то нахождение матрицы U сводится к элементарным алгебраическим операциям. Если $\rho = 1$, то под $I_\rho(a)$ мы подразумеваем просто число a. Может случиться, что и при наличии одинаковых характеристических чисел все элементарные делители (106) будут простыми, т. е. будут иметь вид:

$$(\lambda - \lambda_1);\quad (\lambda - \lambda_2);\quad \ldots;\quad (\lambda - \lambda_n).$$

В этом случае квазидиагональная матрица

$$[I_{\rho_1}(\lambda_1),\ I_{\rho_2}(\lambda_2),\ \ldots,\ I_{\rho_p}(\lambda_p)]$$

превращается просто в диагональную матрицу $[\lambda_1, \ldots, \lambda_n]$, и матрица приводится к диагональной форме.

Нетрудно доказать, что для возможности приведения матрицы к диагональной форме необходимо и достаточно, чтобы ранг таблицы коэффициентов в системе (105) был равен $(n - \mu_k)$, где μ_k — кратность корня λ_k векового уравнения. При выполнении этого условия система (105) определит μ_k линейно-независимых векторов $(v_{1k}, v_{2k}, \ldots, v_{nk})$ [14].

Отметим, что матрица U, входящая в формулу (109), определяется не единственным образом. В частности, если d — величина определителя матрицы U, то в формуле (109) можно заменить

$$U \text{ на } \frac{1}{\sqrt[n]{d}}\,U \text{ и } U^{-1} \text{ на } \sqrt[n]{d}\,U,$$

т. е. можно считать, что в формуле (109) определитель матрицы U равен единице. Этими указаниями мы и ограничимся в общей задаче приведения матриц к каноническому виду. Мы вернемся к этой задаче в специальном добавлении ко второй части третьего тома. Как уже упоминалось выше, мы будем дальше рассматривать подробно эту задачу для матриц особого типа.

28. Унитарные и ортогональные преобразования. В этом и следующих номерах мы будем пользоваться понятиями скалярного произведения и нормы (длины) вектора, которое было введено в [13]. Напомним, что квадрат нормы (длины) определяется формулой:

$$\|x\|^2 = (x,\ x) = \sum_{s=1}^{n} |x_s|^2, \tag{110}$$

или, в случае вещественных составляющих,

$$\|x\|^2 = \sum_{s=1}^{n} x_s^2.$$

Это определение нормы связано с определенным выбором основных ортов, т. е. координатных осей. Мы будем называть координатную систему с указанным выше определением нормы нормальной, или декартовой, системой. Помимо длины вектора было определено еще *скалярное произведение двух векторов* следующей формулой:

$$(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n. \tag{111}$$

В случае вещественных векторов эта формула принимает более симметричный вид

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

Из (111) вытекает, что при перестановке порядка векторов величина скалярного произведения переходит в сопряженную, т. е.

$$(y, x) = \overline{(x, y)}. \tag{112}$$

Два вектора мы назвали *перпендикулярными* или *ортогональными* если их скалярное произведение равно нулю.

В дальнейшем всегда будем считать, если не оговорено особо противоположное, что мы имеем дело с декартовой системой координат. В связи с этим приобретают особое значение те линейные преобразования, которые соответствуют переходу одной декартовой системы в другую. Мы знаем, что всякому переходу от одних ортов к другим соответствует линейное преобразование составляющих. Пусть имеется такое преобразование

$$(y_1, \ldots, y_n) = U(x_1, \ldots, x_n), \tag{113}$$

причем первоначальная система координат была декартовой. Для того чтобы и новая система была декартовой, необходимо и достаточно, чтобы длина вектора и в новой системе выражалась суммой квадратов модулей составляющих, т. е.

$$|y_1|^2 + \ldots + |y_n|^2 = |x_1|^2 + \ldots + |x_n|^2. \tag{114}$$

Покажем, что при этом величина скалярного произведения и в новой системе координат выразится формулой, аналогичной (111). Действительно, положим, мы имели в первоначальной системе координат два вектора

$$x(x_1, \ldots, x_n) \text{ и } x'(x'_1, \ldots, x'_n),$$

причем в новой системе им соответствуют векторы

$$y(y_1, \ldots, y_n) \text{ и } y'(y'_1, \ldots, y'_n).$$

Составим два новых вектора $z = x + x'$ и $u = x + ix'$, которые имеют составляющие $(x_k + x'_k)$ и $(x_k + ix'_k)$. Считая условие (114) выполненным, будем иметь:

$$\sum_{k=1}^{n} (y_k + y'_k)(\bar{y}_k + \bar{y}'_k) = \sum_{k=1}^{n} (x_k + x'_k)(\bar{x}_k + \bar{x}'_k),$$

откуда, опять-таки в силу (114), получаем окончательно

$$\sum_{k=1}^{n} (y_k \bar{y}'_k + y'_k \bar{y}_k) = \sum_{k=1}^{n} (x_k \bar{x}'_k + x'_k \bar{x}_k), \tag{115_1}$$

ибо

$$\sum_{k=1}^{n} |y_k|^2 = \sum_{k=1}^{n} |x_k|^2 \quad \text{и} \quad \sum_{k=1}^{n} |y'_k|^2 = \sum_{k=1}^{n} |x'_k|^2.$$

Точно так же:

$$\sum_{k=1}^{n} (y_k + iy'_k)(\bar{y}_k - i\bar{y}'_k) = \sum_{k=1}^{n} (x_k + ix'_k)(\bar{x}_k - i\bar{x}'_k)$$

и отсюда

$$\sum_{k=1}^{n} (y'_k \bar{y}_k - y_k \bar{y}'_k) = \sum_{k=1}^{n} (x'_k \bar{x}_k - x_k \bar{x}'_k). \tag{115_2}$$

Равенства (115_1) и (115_2) дают

$$\sum_{k=1}^{n} y_k \bar{y}'_k = \sum_{k=1}^{n} x_k \bar{x}'_k, \tag{116}$$

т. е. скалярное произведение действительно выражается прежней формулой. Таким образом, если преобразование (113) удовлетворяет условию (114), то оно удовлетворяет и условию (116), т. е. оставляет неизменным величину скалярного произведения. Наоборот, из условия (116) вытекает (114), если положить в (116) $x'_k = x_k$, так как скалярное произведение двух одинаковых векторов сводится, очевидно, к квадрату длины вектора. Линейные преобразования, удовлетворяющие условию (114) или условию (116), называются обычно *унитарными преобразованиями.*

Если рассматривать вещественное пространство и вещественные матрицы линейных преобразований, то условие (114) сведется просто к условию

$$y_1^2 + y_2^2 + \ldots + y_n^2 = x_1^2 + x_2^2 + \ldots + x_n^2, \tag{117}$$

и соответственные вещественные преобразования называются *ортогональными.* Они являются, очевидно, частным случаем унитарных.

Переходим теперь к выяснению основных свойств унитарных преобразований. Напишем для преобразования (113) условие (114) в явной форме, обозначая через u_{ik} элементы матрицы U:

$$\sum_{k=1}^{n} |u_{k1}x_1 + \ldots + u_{kn}x_n|^2 = \sum_{k=1}^{n} |x_k|^2$$

или

$$\sum_{k=1}^{n} (u_{k1}x_1 + \ldots + u_{kn}x_n)(\bar{u}_{k1}\bar{x}_1 + \ldots + \bar{u}_{kn}\bar{x}_n) = \sum_{k=1}^{n} x_k\bar{x}_k. \quad (118)$$

Раскрывая скобки в левой части формулы и приравнивая коэффициенты при $x_p\bar{x}_p$ единице, а при $x_p\bar{x}_q\,(p\neq q)$ нулю, будем иметь необходимое и достаточное условие для элементов унитарного преобразования в следующей форме:

$$\left.\begin{aligned} &\sum_{k=1}^{n} |u_{kp}|^2 = 1 \qquad (p = 1, 2, \ldots, n), \\ &\sum_{k=1}^{n} u_{kp}\bar{u}_{kq} = 0 \qquad (p \neq q), \end{aligned}\right\} \quad (119)$$

т. е. сумма квадратов модулей элементов каждого столбца должна равняться единице и сумма произведений элементов некоторого столбца на величины, сопряженные с соответствующими элементами другого столбца, должна равняться нулю. Иногда эти условия еще записывают так:

$$\sum_{k=1}^{n} u_{kp}\bar{u}_{kq} = \delta_{pq}, \quad (120)$$

где δ_{pq} суть элементы единичной матрицы, т. е.

$$\delta_{pq} = \begin{cases} 0 & (p \neq q) \\ 1 & (p = q) \end{cases}. \quad (121)$$

Выше мы применили к тождеству (118) метод неопределенных коэффициентов. Это является, конечно, достаточным для выполнения тождества. Нетрудно показать, придавая x_k частные значения, что тождественность коэффициентов при подобных членах является и необходимым условием.

Возьмем определитель $D(A)$ и другой определитель $D(\overline{A})$, составленный из сопряженных элементов. Умножая их по схеме столбец на столбец [6], мы получим, в силу (119), определитель единичной матрицы, т. е. единицу. С другой стороны, очевидно, что оба упомянутых определителя будут выражаться комплексными сопряженными числами, и из сказанного непосредственно следует

$$|D(A)|^2 = 1,$$

т. е. *квадрат модуля определителя унитарной матрицы равен единице*. Иначе говоря, определитель унитарной матрицы по модулю равен единице, т. е. выражается комплексным числом вида $e^{i\varphi}$, где φ — вещественно.

Введем в рассмотрение матрицу $U^{(*)}$, транспонированную с U. Условия (119), которые называются обычно условиями ортогональности по столбцам, могут быть записаны в виде следующего матричного равенства:

$$\overline{U}^{(*)} U = I, \tag{122}$$

что равносильно

$$U^{-1} = \overline{U}^{(*)} = \tilde{U}, \tag{123}$$

т. е. если матрица унитарна, то обратная ей матрица совпадает с эрмитовски-сопряженной матрицей.

Преобразование U^{-1}, обратное U, выражает переход от вектора $\boldsymbol{y}$ к вектору $\boldsymbol{x}$. Оно также, очевидно, удовлетворяет условию унитарности (114), т. е. если U — унитарная матрица, то и обратная U^{-1} будет унитарной. Иными словами, в силу (123), матрица $\tilde{U}$ будет унитарной, и ее столбцы удовлетворяют условию ортогональности. Но столбцы $\tilde{U}$ суть строки $\overline{U}$. Мы можем таким образом утверждать, что в унитарной матрице не только столбцы, но и строки удовлетворяют условиям ортогональности, т. е. наряду с формулами (120) будем иметь также формулы

$$\sum_{k=1}^{n} u_{pk}\overline{u}_{qk} = \delta_{pq}. \tag{124}$$

Аналогично предыдущему, если матрицы U_1 и U_2 удовлетворяют условию (114), то и их произведение U_2U_1 также, очевидно, будет удовлетворять этому условию, т. е. произведение двух унитарных матриц есть также унитарная матрица.

Укажем две различные формы, в которых можно представить определение унитарной матрицы:

$$|U\boldsymbol{x}|^2 = |\boldsymbol{x}|^2 \text{ или } (U\boldsymbol{x},\ U\boldsymbol{x}') = (\boldsymbol{x},\ \boldsymbol{x}'), \tag{125_1}$$

причем в последнем равенстве $\boldsymbol{x}$ и $\boldsymbol{x}'$ — любые векторы.

Отметим теперь те обстоятельства, которые будут иметь место, если унитарная матрица имеет вещественные элементы. В этом случае, как мы уже говорили, она называется ортогональной и соответствующее ей преобразование — ортогональным преобразованием.

В данном случае вместо формул (120) и (124) мы будем иметь формулы

$$\sum_{k=1}^{n} u_{kp} u_{kq} = \delta_{pq}; \qquad \sum_{k=1}^{n} u_{pk} u_{qk} = \delta_{pq}. \tag{125$_1$}$$

Кроме того, определитель преобразования должен быть, очевидно, вещественным числом, а потому его величина может равняться лишь ± 1. Эти вещественные ортогональные преобразования в n-мерном пространстве являются полным аналогом тех преобразований трехмерного пространства, которые мы рассматривали в [**20**]. В этом вещественном случае, кроме того, $\tilde{U}$ совпадает с $U^{(*)}$, т. е. обратное преобразование U^{-1} получается из U заменой строк столбцами.

Отметим еще, что всякое комплексное число $e^{i\varphi}$, где φ — вещественно, рассматриваемое как матрица $[e^{i\varphi}, e^{i\varphi}, \ldots, e^{i\varphi}]$, представляет собою унитарную матрицу, и если U есть унитарная матрица, то и произведение $e^{i\varphi}U$ есть унитарная матрица. Смысл произведения числа на матрицу указан нами в [**25**].

29. Неравенство Коши — Буняковского. Установим в настоящем параграфе одно неравенство, которым нам придется пользоваться в дальнейшем. Оно состоит в следующем: каковы бы ни были вещественные числа $\alpha_1, \alpha_2, \ldots, \alpha_m$ и $\beta_1, \beta_2, \ldots, \beta_m$, имеем:

$$\left(\sum_{k=1}^{m} \alpha_k \beta_k\right)^2 \leqslant \sum_{k=1}^{m} \alpha_k^2 \cdot \sum_{k=1}^{m} \beta_k^2, \tag{126}$$

причем знак равенства будет иметь место тогда и только тогда, когда α_k и β_k пропорциональны:

$$\frac{\beta_1}{\alpha_1} = \frac{\beta_2}{\alpha_2} = \ldots = \frac{\beta_m}{\alpha_m}. \tag{127}$$

Пусть ξ — любое вещественное число. Составим сумму:

$$S = \sum_{k=1}^{m} (\xi\alpha_k - \beta_k)^2,$$

которая, очевидно, неотрицательна. Знак равенства будет иметь место тогда и только тогда, когда

$$\frac{\beta_1}{\alpha_1} = \frac{\beta_2}{\alpha_2} = \ldots = \frac{\beta_m}{\alpha_m} = \xi,$$

и в этом случае, очевидно:

$$\left(\sum_{k=1}^{m} \alpha_k \beta_k\right)^2 = \sum_{k=1}^{m} \alpha_k^2 \cdot \sum_{k=1}^{m} \beta_k^2.$$

Вообще же говоря, раскрывая в выражении S скобки, получим трехчлен второй степени

$$S = A\xi^2 - 2B\xi + C,$$

где

$$A = \sum_{k=1}^{m} \alpha_k^2; \quad B = \sum_{k=1}^{m} \alpha_k \beta_k; \quad C = \sum_{k=1}^{m} \beta_k^2.$$

Написанный трехчлен при всех вещественных ξ остается неотрицательным, откуда следует: $AC - B^2 \geqslant 0$, т. е. $B^2 \leqslant AC$, что и приводит к неравенству (126).

Если $AC - B^2 = 0$, то трехчлен при некотором вещественном ξ должен обращаться в нуль, а при этом, как мы видели, должно выполняться условие (127). Наоборот, при выполнении этого условия в формуле (126) имеет место знак равенства. Положим теперь, что α_k и β_k — комплексные числа. Принимая во внимание, что

$$\left| \sum_{k=1}^{m} \alpha_k \beta_k \right| \leqslant \sum_{k=1}^{m} |\alpha_k| \, |\beta_k|,$$

получим, применяя к последней сумме, состоящей из положительных слагаемых, неравенство (126):

$$\left| \sum_{k=1}^{m} \alpha_k \beta_k \right|^2 \leqslant \sum_{k=1}^{m} |\alpha_k|^2 \cdot \sum_{k=1}^{m} |\beta_k|^2. \qquad (126_1)$$

Нетрудно показать, что в данном случае, при комплексных α_k и β_k, знак равенства будет иметь место тогда и только тогда, когда $|\alpha_k|$ и $|\beta_k|$ пропорциональны и все произведения $\alpha_k \beta_k$ имеют одинаковый аргумент. Неравенство (126) применимо не только к суммам, но и к интегралам, как мы об этом уже упоминали раньше [**II, 161**]. Если $f_1(x)$ и $f_2(x)$ — две вещественные функции в промежутке $a \leqslant x \leqslant b$, то неравенство для интегралов имеет вид:

$$\left[\int_a^b f_1(x) f_2(x)\, dx \right]^2 \leqslant \int_a^b f_1^2(x)\, dx \cdot \int_a^b f_2^2(x)\, dx. \qquad (126_2)$$

Действительно, составим выражение

$$\int_a^b [\xi f_1(x) - f_2(x)]^2\, dx = \xi^2 \int_a^b f_1^2(x)\, dx - 2\xi \int_a^b f_1(x) f_2(x)\, dx + \int_a^b f_2^2(x)\, dx,$$

где ξ — любое вещественное число. Из вида левой части следует, что это выражение ни при каких вещественных ξ не может быть отрицательным. Но если трехчлен $A\xi^2 - 2\xi B + C$ при всех вещественных ξ не отрицателен, то, как известно из элементарной алгебры, $AC -$

$-B^2 \geqslant 0$. В применении к предыдущему трехчлену это и дает неравенство (126_2). Это неравенство для интегралов впервые было доказано В. Л. Буняковским. Для сумм оно встречалось у Коши.

30. Свойства скалярного произведения и нормы. Отметим теперь некоторые свойства скалярного произведения и нормы. Применяя неравенство (126_1) и принимая во внимание, что $|\bar{y}_k| = |y_k|$, можем написать:

$$|(\boldsymbol{x}, \boldsymbol{y})|^2 = \left|\sum_{k=1}^{n} x_k \bar{y}_k\right|^2 \leqslant \sum_{k=1}^{n} |x_k|^2 \cdot \sum_{k=1}^{n} |y_k|^2,$$

т. е.

$$|(\boldsymbol{x}, \boldsymbol{y})| \leqslant \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|. \tag{128}$$

Докажем теперь так называемое правило треугольника

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|. \tag{129}$$

Мы имеем:

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = (\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) + (\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{y}, \boldsymbol{x}),$$

или, принимая во внимание (128), получим:

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 \leqslant \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| = (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2,$$

откуда и следует (129).

В заключение настоящего номера рассмотрим, какое влияние оказывает выбор системы координат на метрику пространства, т. е. на выражение квадрата длины вектора. Положим, что вместо основной декартовой мы берем новую систему координат, причем за основные орты принимаем некоторые независимые векторы

$$\boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}, \ldots, \boldsymbol{z}^{(n)}.$$

Для любого вектора будем иметь:

$$\boldsymbol{x} = z_1 \boldsymbol{z}^{(1)} + \ldots + z_n \boldsymbol{z}^{(n)},$$

где z_k — его составляющие в новой координатной системе.

Квадрат длины этого вектора будет выражаться скалярным произведением вектора на самого себя, т. е.

$$|\boldsymbol{x}|^2 = (z_1 \boldsymbol{z}^{(1)} + \ldots + z_n \boldsymbol{z}^{(n)}, z_1 \boldsymbol{z}^{(1)} + \ldots + z_n \boldsymbol{z}^{(n)}).$$

Раскрывая это, согласно вышеуказанным формулам, будем иметь следующее выражение для квадрата длины вектора:

$$|\boldsymbol{x}|^2 = \sum_{i, k=1}^{n} a_{ik} z_i \bar{z}_k, \tag{130}$$

где коэффициенты α_{ik} определяются по формулам

$$\alpha_{ik}=(z^{(i)},\ z^{(k)}).$$

При перестановке значков они, очевидно, переходят в сопряженные, т. е.

$$\alpha_{ki}=\overline{\alpha}_{ik}. \tag{131}$$

Сумма вида (130) с коэффициентами, удовлетворяющими условию (131), называется обычно *формой Эрмита*. Непосредственно очевидно, что всякое выражение вида (130) при условии (131) будет иметь при всевозможных комплексных z_k лишь вещественные значения, так как при $i \neq k$ два члена суммы (130) будут сопряженными, а в членах вида $\alpha_{kk}|z_k|^2$, в силу условия (131), коэффициенты α_{kk} будут вещественными. Кроме того, по самому построению формы Эрмита (130), в данном случае мы можем утверждать, что сумма (130) будет неотрицательной и будет обращаться в нуль только тогда, когда все z_k равны нулю. Формула (130) и определяет метрику пространства в новой координатной системе.

Метрика (130) будет совпадать с метрикой (110) в соответствующей декартовой системе, если $\alpha_{ik}=0$ при $i \neq k$ и $\alpha_{kk}=1$, или $(z^{(i)},\ z^{(k)})=0$ при $i \neq k$ и $(z^{(k)},\ z^{(k)})=1$, т. е., иначе говоря, если векторы $z^{(k)}$, принятые нами за орты, будут взаимно ортогональными единичными векторами (длины единица).

В дальнейшем всякую систему взаимно ортогональных и единичных векторов $z^{(k)}$ $(k=1,\ 2,\dots,\ l)$ мы будем называть *ортонормированной* системой.

Заметим еще, что если формула (113) определяет унитарное преобразование для составляющих вектора, то соответствующее преобразование для перехода от прежних ортов к новым будет даваться таблицей

$$U^{(*)-1},$$

контраградиентной U. В данном случае, в силу (123), эта таблица будет совпадать с таблицей $\overline{U}$, а для вещественных ортогональных преобразований она будет просто совпадать с U.

31. Процесс ортогонализации векторов. Положим, что нам даны m каких-нибудь линейно-независимых векторов $x^{(1)},\dots,\ x^{(m)}$. Совокупность векторов вида:

$$C_1x^{(1)}+\dots+C_mx^{(m)},$$

где C_k — произвольные коэффициенты, определяет все наше пространство, если $m=n$, и некоторое подпространство R_m измерения m, если $m<n$. Покажем, что мы всегда можем построить ортонормированную систему векторов $z^{(k)}$ $(k=1,\ 2,\dots,\ m)$, которая образует

то же подпространство R_m, что и векторы $x^{(k)}$. Иначе говоря, $z^{(k)}$ должны выражаться линейно через $x^{(k)}$ и, наоборот, $x^{(k)}$ должны выражаться через $z^{(k)}$. Эти векторы мы можем построить по следующей схеме:

$$\left.\begin{array}{l} y^{(1)} = x^{(1)} \\ y^{(2)} = x^{(2)} - (x^{(2)},\ z^{(1)})\,z^{(1)} \\ y^{(3)} = x^{(3)} - (x^{(3)},\ z^{(1)})\,z^{(1)} - (x^{(3)},\ z^{(2)})\,z^{(2)} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\}, \tag{132}$$

где

$$z^{(1)} = \frac{y^{(1)}}{|y^{(1)}|};\quad z^{(2)} = \frac{y^{(2)}}{|y^{(2)}|};\ \ldots;\ z^{(m)} = \frac{y^{(m)}}{|y^{(m)}|}. \tag{133}$$

Вектор $z^{(1)}$ получается из $y^{(1)}$ простым делением на длину $y^{(1)}$ и, следовательно, длина $z^{(1)}$ равна единице. Затем строится вектор $y^{(2)}$ по указанной выше формуле. Из самого его определения непосредственно вытекает, что он ортогонален с $z^{(1)}$:

$$(y^{(2)},\ z^{(1)}) = (x^{(2)},\ z^{(1)}) - (x^{(2)},\ z^{(1)})(z^{(1)},\ z^{(1)}) = 0.$$

Деля вектор $y^{(2)}$ на его длину, получаем вектор $z^{(2)}$. Затем дальше строим вектор $y^{(3)}$ по указанной выше формуле. Из нее непосредственно вытекает, что он ортогонален с $z^{(1)}$ и $z^{(2)}$.

Действительно, в силу ортогональности $z^{(1)}$ и $z^{(2)}$, получим, например:

$$(y^{(3)},\ z^{(2)}) = (x^{(3)},\ z^{(2)}) - (x^{(3)},\ z^{(2)})(z^{(2)},\ z^{(2)}) = 0.$$

Деля вектор $y^{(3)}$ на его длину, получаем вектор $z^{(3)}$ и т. д.

Все вновь построенные векторы выражаются линейно через $x^{(k)}$. Нетрудно видеть и наоборот, что $x^{(k)}$ выражаются через $z^{(k)}$. Для этого достаточно только постепенно решать предыдущие равенства относительно $x^{(1)}$, $x^{(2)}$ и т. д.

Заметим также, что ни один из вновь построенных векторов $y^{(k)}$ не может обратиться в нуль. Действительно, если бы мы на некотором шаге вычислений получили вектор $y^{(k)}$, равный нулю, то, так как он выражается линейно через $x^{(s)}$, причем коэффициент при $x^{(k)}$ в этом линейном выражении равен единице, то мы получили бы линейную зависимость между векторами $x^{(s)}$, что противоречит тому условию, что эти векторы линейно-независимы.

Напомним, что если имеется некоторая совокупность попарно ортогональных векторов, отличных от нулевого вектора, то эти векторы линейно-независимы.

Если $m = n$, то $z^{(k)}$ дают полную ортонормированную систему. Если же $m < n$, то для получения полной системы декартовых координат мы должны в дополнение к построенным векторам $z^{(k)}$ достроить еще $(n - m)$ векторов, которые были бы ортогональны и между собой и к векторам $z^{(k)}$. Эти новые единичные векторы должны, таким образом, образовывать подпространство R'_{n-m} измерения $(n-m)$, ортогональное подпространству R_m [12]. Новые искомые векторы u должны удовлетворять системе уравнений

$$(u, x^{(1)})=0, \ldots, (u, x^{(m)})=0.$$

Здесь мы имеем систему m однородных уравнений с n неизвестными, причем ранг этой системы равен m, поскольку векторы $x^{(k)}$ — линейно-независимы [12]. Эта система имеет $(n-m)$ линейно-независимых решений, т. е. мы получаем $(n-m)$ линейно-независимых векторов. Применяя к ним указанный выше процесс ортогонализации и приводя их длину к единице, мы и получим совместно с $z^{(1)}, \ldots, z^{(m)}$ полную ортонормированную систему. Сделаем еще одно замечание. Подпространство R_m, образованное ортонормированной системой векторов $z^{(k)}$, может быть образовано и другой такой же системой. Действительно, достаточно для этого применить к системе векторов $z^{(k)}$ унитарное преобразование. Мы видим, таким образом, что процесс ортогонализации системы векторов может совершаться различным образом, и указанный выше прием дает лишь одну из таких возможностей.

§ 4. КВАДРАТИЧНЫЕ ФОРМЫ

32. Преобразование квадратичной формы к сумме квадратов. Рассмотрим в пространстве некоторую поверхность второго порядка, имеющую центр в начале координат

$$Ax^2+By^2+Cz^2+2Dxy+2Exz+2Fyz+G=0.$$

Всегда можно выбрать так новые координатные оси (x', y', z'), чтобы в преобразованном уравнении остались лишь члены, содержащие квадраты координат, т. е. так, чтобы преобразованное уравнение имело вид:

$$\lambda_1 x'^2+\lambda_2 y'^2+\lambda_3 z'^2+G=0.$$

Задача сводится к отысканию такого ортогонального преобразования, связывающего переменные (x', y', z') и (x, y, z), чтобы совокупность членов второго измерения относительно координат в левой части уравнения привелась к сумме квадратов. Поставим аналогичную задачу для случая вещественного пространства n измерений. Пусть у нас имеется вещественная квадратичная форма от n переменных:

$$\varphi(x_1, \ldots, x_n)=\sum_{i, k=1}^{n} a_{ik} x_i x_k, \tag{134}$$

причем a_{ik} — вещественные коэффициенты, удовлетворяющие условию

$$a_{ki}=a_{ik}. \tag{135}$$

В предыдущем примере мы можем считать $x=x_1$, $y=x_2$, $z=x_3$ и $a_{11}=A$; $a_{22}=B$; $a_{33}=C$; $a_{12}=a_{21}=D$; $a_{13}=a_{31}=E$; $a_{23}=a_{32}=F$.

Назовем *матрицей квадратичной формы* (134) матрицу, составленную из элементов a_{ik}. Эта матрица будет симметрична, т. е. будет совпадать со своей транспонированной.

Положим, что мы преобразуем форму (134) к новым переменным, вводя вместо x_k переменные x'_k, причем преобразование это записано в виде:

$$(x_1, \ldots, x_n)=B(x'_1, \ldots, x'_n), \tag{136}$$

где B есть матрица с элементами b_{ik}. Внося выражения (136) в формулу (134), будем иметь:

$$\varphi=\sum_{i,k=1}^{n} a_{ik}(b_{i1}x'_1+\ldots+b_{in}x'_n)(b_{k1}x'_1+\ldots+b_{kn}x'_n).$$

Раскрывая скобки, получим коэффициент при $x'_p x'_q$ $(p\neq q)$:

$$\sum_{i,k=1}^{n} a_{ik}(b_{ip}b_{kq}+b_{iq}b_{kp}).$$

Пользуясь (135), легко видеть, что половина написанного выражения будет равна просто сумме:

$$\sum_{i=k}^{n} b_{ip}\sum_{k=1}^{n} a_{ik}b_{kq}.$$

Таким образом, разбирая аналогично и случай $p=q$, мы увидим, что в новых переменных квадратичная форма будет иметь вид:

$$\varphi=\sum_{i,k=1}^{n} c_{ik}x'_i x'_k, \tag{137}$$

где

$$c_{ik}=c_{ki}=\sum_{t=1}^{n} b_{ti}\sum_{s=1}^{n} a_{ts}b_{sk}.$$

Суммирование по s дает $\{AB\}_{tk}$. Если в множителе b_{ti} будем считать t — номером столбца и i — номером строки, то b_{ti} будет элементом транспонированной матрицы $\{B^{(*)}\}_{it}$, откуда

$$c_{ik}=c_{ki}=\sum_{t=1}^{n}\{B^{(*)}\}_{it}\{AB\}_{tk},$$

т. е. преобразованная форма (137) будет иметь матрицу, определяемую через матрицу A формы в прежних переменных и матрицу B преобразования (136) следующим образом:

$$C=B^{(*)}AB. \tag{138}$$

Если преобразование (136) ортогонально, то для ортогональной матрицы B транспонированная $B^{(*)}$ совпадает с обратной B^{-1}, и мы имеем в этом случае, вместо формулы (138), формулу:

$$C = B^{-1}AB. \tag{139}$$

Таким образом, наша задача о построении ортогонального преобразования (136), приводящего квадратичную форму (134) к сумме квадратов, равносильна задаче построения такой ортогональной матрицы B, чтобы матрица (139) была просто диагональной матрицей $[\lambda_1, \ldots, \lambda_n]$, ибо матрица такой формы, которая приводится к сумме квадратов, и есть диагональная матрица, причем ее элементы λ_k и суть коэффициенты при квадратах x'^2_k. Итак, мы должны иметь, как и выше:

$$B^{-1}AB = [\lambda_1, \ldots, \lambda_n]$$

или

$$AB = B[\lambda_1, \ldots, \lambda_n]. \tag{140}$$

Заметим, что в данном случае матрица A не какая угодно матрица, а вещественная симметричная матрица и B должна быть ортогональной матрицей. Будем поступать совершенно так же, как это мы делали выше в [**27**] при рассмотрении общего случая. Перепишем уравнение (140) в виде:

$$\sum_{s=1}^{n} a_{is}b_{sk} = \lambda_k b_{ik}. \tag{141}$$

Отсюда для элементов k-го столбца матрицы B имеем n уравнений. Вводя вектор $x^{(k)}$ с составляющими $(b_{1k}, \ldots, b_{nk})$, можем переписать последнее уравнение так:

$$Ax^{(k)} = \lambda_k x^{(k)}. \tag{142}$$

Перенеся все члены (141) в одну сторону, будем иметь для определения $b_{1k}, \ldots, b_{nk}$ систему n однородных уравнений

$$\left.\begin{array}{r}(a_{11} - \lambda_k) b_{1k} + a_{12}b_{2k} + \ldots + a_{1n}b_{nk} = 0 \\ a_{21}b_{1k} + (a_{22} - \lambda_k) b_{2k} + \ldots + a_{2n}b_{nk} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}b_{1k} + a_{n2}b_{2k} + \ldots + (a_{nn} - \lambda_k) b_{nk} = 0\end{array}\right\}. \tag{143}$$

Определитель этой системы должен равняться нулю, и для чисел λ_k мы получаем алгебраическое уравнение степени n:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{144}$$

Это, как мы знаем, и есть характеристическое уравнение матрицы A.

Докажем прежде всего, что для вещественной симметричной матрицы A уравнение (144) имеет все корни вещественные. Предварительно дадим новую форму записи для квадратичной формы. Пусть x — вектор с составляющими $(x_1, \ldots, x_n)$ — вещественными или комплексными и A — матрица с любыми элементами a_{ik}. Составим скалярное произведение:

$$(Ax, x) = \sum_{i=1}^{n} \bar{x}_i (a_{i1}x_1 + \ldots + a_{in}x_n).$$

Мы видим, что оно может быть записано в виде:

$$(Ax, x) = \sum_{i,k=1}^{n} a_{ik}\bar{x}_i x_k. \tag{145}$$

Если выполнено условие

$$a_{ki} = \bar{a}_{ik} \quad (a_{kk} \text{ — вещественны}), \tag{146}$$

то это есть форма Эрмита, значения которой обязательно вещественны. Тот случай, когда A есть вещественная симметричная матрица, есть частный случай условий (146). Если при этом еще составляющие вектора x вещественны, то формула (145) и дает нам квадратичную форму (134).

Перейдем теперь к доказательству вещественности корней уравнения (144). Пусть λ_k — некоторый корень этого уравнения. Система (143) дает нам при этом составляющие вектора $x^{(k)}$ (вещественного или комплексного), удовлетворяющего уравнению (142). Умножим обе части этого уравнения справа скалярно на $x^{(k)}$. Мы получим:

$$|x^{(k)}|^2 \lambda_k = (Ax^{(k)}, x^{(k)}).$$

Выражение, стоящее справа, как мы видели, есть число вещественное, и, следовательно, λ_k есть также число вещественное. Мы доказали таким образом вещественность корней уравнения (144) не только для вещественных симметричных матриц, но и для матриц, элементы которых удовлетворяют условию (146). Такие матрицы называются обычно *эрмитовскими матрицами.*

В рассматриваемом случае коэффициенты системы (143) суть вещественные числа, и мы можем считать, что и составляющие $x^{(k)}$ — вещественны. Покажем теперь, что если λ_p и λ_q — два различных корня уравнения (144), то соответствующие векторы $x^{(p)}$ и $x^{(q)}$, удовлетворяющие уравнению (142), — взаимно ортогональны. Мы имеем по условию:

$$Ax^{(p)} = \lambda_p x^{(p)}; \quad Ax^{(q)} = \lambda_q x^{(q)}.$$

Умножая первое из этих уравнений скалярно на $x^{(q)}$, а второе на $x^{(p)}$, и вычитая, получим:

$$(Ax^{(p)},\ x^{(q)}) - (x^p,\ Ax^{(q)}) = (\lambda_p - \lambda_q)(x^{p)},\ x^{(q)}). \tag{147}$$

Покажем теперь, что для любых двух векторов x и y (вещественных и комплексных) имеет место формула

$$(Ax,\ y) = (x,\ Ay), \tag{148}$$

если только элементы матрицы A удовлетворяют условиям (146). Действительно, левая часть формулы (148) дает:

$$(Ax,\ y) = \sum_{k=1}^{n}(a_{k1}x_1 + \ldots + a_{kn}x_n)\,\bar{y}_k = \sum_{i,k=1}^{n} a_{ki}x_i\bar{y}_k,$$

или, в силу (146):

$$(Ax,\ y) = \sum_{i,k=1}^{n} \overline{a}_{ik}x_i\bar{y}_k.$$

Такой же результат даст нам и правая часть формулы (148). Эта формула справедлива, таким образом, и для вещественных симметричных матриц, которые являются частным случаем эрмитовских. В силу (148) левая часть (147) равна нулю, и в силу $\lambda_p \neq \lambda_q$ мы имеем $(x^{(p)},\ x^{(q)}) = 0$, т. е. векторы $x^{(p)}$ и $x^{(q)}$ действительно ортогональны. В данном случае это вещественные векторы, и условие их ортогональности сводится к тому, что сумма произведений их составляющих равна нулю.

Если уравнение (144) имеет различные корни, то мы будем иметь, таким образом, n взаимно ортогональных вещественных векторов $x^{(k)}$. Уравнение (142) линейное, однородное относительно $x^{(k)}$, и мы можем умножить решение этого уравнения на любую постоянную, так что можно считать, что упомянутые выше векторы $x^{(k)}$ имеют длину, равную единице.

Составляющие этих векторов образуют столбцы матрицы B. Иначе говоря, эта матрица удовлетворяет условию ортонормированности по столбцам и является ортогональной матрицей. Таким образом наша задача приведения квадратичной формы ортогональным преобразованием к сумме квадратов или — что то же — приведения матрицы A к диагональной форме окончена в предположении, что уравнение (144) имеет различные корни. Числа λ_k называют иногда *собственными значениями матрицы* A, а векторы $x^{(k)}$ — *собственными векторами этой матрицы*.

33. Случай кратных корней характеристического уравнения. Перейдем теперь к рассмотрению общего случая, когда уравнение (144) может иметь и одинаковые корни. Возьмем некоторый корень $\lambda = \lambda_1$ уравнения (144) и построим соответствующее ему решение уравнения (142). Это будет некоторый вещественный вектор $x^{(1)}$ длины единица. Присоединим к нему еще $(n-1)$ вещественных единичных векторов, которые бы вместе с ним образовывали полную ортонормированную систему [**31**]. Переход от основных ортов к этим новым ортам будет, как мы знаем, выражаться некоторым ортогональным преобразованием над составляющими векторами, и матрица A перейдет в подобную матрицу $A_1 = B_1^{-1}AB_1$. Соответствующее этой новой матрице уравнение

$$A_1x = \lambda x \tag{149}$$

будет иметь в качестве решения, соответствующего собственному значению $\lambda = \lambda_1$ (собственные значения не меняются при преобразовании подобия), вектор $x^{(1)}$, который мы приняли за первый орт и который имеет, следовательно, составляющие $(1, 0, \ldots, 0)$. Подставляя это решение в уравнение (149), будем иметь:

$$A_1(1, 0, \ldots, 0) = (\lambda_1, 0, \ldots, 0),$$

откуда непосредственно следует для элементов первого столбца

$$\{A_1\}_{11} = \lambda_1; \quad \{A_1\}_{21} = \{A_1\}_{31} = \ldots = \{A_1\}_{n1} = 0. \tag{150}$$

Покажем теперь, что вещественная матрица A_1 будет также симметрична, т. е. будет совпадать со своей транспонированной. Действительно:

$$A_1^{(*)} = (B_1^{-1}AB_1)^{(*)} = B_1^{(*)}A^{(*)}B_1^{(*)-1}.$$

Но в силу ортогональности матрицы B_1:

$$B_1^{(*)} = B_1^{-1} \quad \text{и} \quad B_1^{(*)-1} = B_1,$$

откуда следует

$$A_1^{(*)} = A_1.$$

Принимая во внимание формулы (150) и симметричность матрицы A_1, мы можем написать:

$$\{A_1\}_{11} = \lambda_1; \quad \{A_1\}_{21} = \{A_1\}_{12} = \ldots = \{A_1\}_{n1} = \{A_1\}_{1n} = 0,$$

т. е. у матрицы A_1 все элементы первой строки и первого столбца обращаются в нуль, кроме, может быть, элемента $\{A_1\}_{11} = \lambda_1$, т. е. эта матрица A_1 имеет форму:

$$A_1 = \begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{Vmatrix},$$

где через $a_{ik}^{(1)}$ мы обозначили элементы A_1.

Квадратичная форма φ в новых переменных будет иметь вид:

$$\varphi = \lambda_1 y'^2_1 + \sum_{i,k=2}^{n} a^{(1)}_{ik} y'_i y'_k.$$

Таким образом мы выделили один квадрат и пришли к рассмотрению квадратичной формы с $(n-1)$ переменными

$$\sum_{i,\,k=2}^{n} a^{(1)}_{ik} y'_i y'_k$$

или, что аналогично, к рассмотрению соответствующей ей матрицы C_1 порядка $(n-1)$, являющейся частью матрицы A_1. Здесь мы можем рассуждать совершенно так же, как и выше, и в подпространстве $(n-1)$-го измерения, образованном последними $(n-1)$ ортами, можем выделить некоторый единичный вектор $x^{(2)}$, являющийся решением уравнения

$$C_1 x^{(2)} = \lambda_2 x^{(2)}.$$

Этот вектор будет, очевидно, ортогонален вектору $x^{(1)}$. В результате этого второго преобразования орт $x^{(1)}$ сохранится, а остальные орты перейдут в другие, взаимно ортогональные единичные орты, первым из которых будет $x^{(2)}$. В этих новых переменных квадратичная форма φ будет иметь вид:

$$\varphi = \lambda_1 y''^2_1 + \lambda_2 y''^2_2 + \sum_{i,\,k=3}^{n} a^{(2)}_{i,k} y''_i y''_k.$$

Продолжая так и дальше, мы приведем, наконец, квадратичную форму к сумме квадратов, т. е. соответствующую ей матрицу к диагональной форме. Это получится в результате применения нескольких ортогональных преобразований, что, очевидно, равносильно применению одного ортогонального преобразования B, являющегося их произведением.

Окончательная диагональная матрица

$$B^{-1}AB = [\lambda_1, \ldots, \lambda_n] \tag{151}$$

будет подобна основной матрице A, и, следовательно, ее характеристическое уравнение

$$\begin{vmatrix} \lambda_1 - \lambda & 0 & \ldots & 0 \\ 0 & \lambda_2 - \lambda & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n - \lambda \end{vmatrix} = 0$$

будет совпадать с уравнением (144), иными словами, коэффициенты λ_k при квадратах в приведенной квадратичной форме

$$\varphi = \lambda_1 x_1'^2 + \ldots + \lambda_n x_n'^2 \tag{152}$$

будут корнями уравнения (144), причем каждый кратный корень будет повторяться столько раз, сколько единиц содержит его кратность.

Каждый столбец окончательного ортогонального преобразования B дает, как мы знаем, вектор, являющийся решением уравнения (142), причем из самого закона построения следует, что соответствующее значение λ_k совпадает с тем коэффициентом в квадратичной форме (152), который стоит при соответствующей переменной. Укажем более точно это соотношение. В силу (136) ортогональное преобразование B, удовлетворяющее условию (140), переводит переменные $(x_1', \ldots, x_n')$ в переменные $(x_1, \ldots, x_n)$.

Обратное преобразование B^{-1} будет транспонированным по отношению к B, т. е. мы будем иметь:

$$x'_k = b_{1k} x_1 + \ldots + b_{nk} x_n \quad (k = 1, 2, \ldots n) \tag{153}$$

и вектор $\boldsymbol{x}^{(k)}$, имеющий составляющие $(b_{1k}, \ldots, b_{nk})$, будет решением уравнения (142), при $\lambda = \lambda_k$.

Покажем, наконец, что мы нашли *все решения уравнения* (142). Прежде всего, из предыдущих рассуждений вытекает, что значение λ_k должно быть корнем уравнения (144). Возьмем какой-нибудь корень этого уравнения λ и положим для определенности, что его кратность равна трем, причем можно, конечно, считать, что $\lambda = \lambda_1 = \lambda_2 = \lambda_3$. Предыдущие рассуждения дают нам для уравнения

$$A\boldsymbol{x} = \lambda_1 \boldsymbol{x} \tag{154}$$

три решения:

$$\boldsymbol{x}^{(1)}(b_{11}, \ldots, b_{n1}); \quad \boldsymbol{x}^{(2)}(b_{12}, \ldots, b_{n2}); \quad \boldsymbol{x}^{(3)}(b_{13}, \ldots, b_{n3}).$$

Покажем, что всякое решение уравнения (154) должно быть их линейной комбинацией. Действительно, если бы это было не так, то мы имели бы еще некоторое решение этого уравнения $\boldsymbol{y}$, линейно-независимое с $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$ и $\boldsymbol{x}^{(3)}$. Вектор $\boldsymbol{y}$ может оказаться и комплексным, но в этом последнем случае его вещественная и мнимая части в отдельности должны удовлетворять уравнению (154), так как это уравнение имеет вещественные коэффициенты. Очевидно, что по крайней мере одна из этих частей должна представлять собою вектор, линейно-независимый с $\boldsymbol{x}^{(k)}$ $(k = 1, 2, 3)$. Мы можем таким образом считать, что тот вектор $\boldsymbol{y}$, о котором мы упоминали выше, есть вещественный вектор. Как мы доказали раньше, он должен быть ортогонален ко всем векторам $\boldsymbol{x}^{(k)}$ при $k > 3$, так как этим последним соответствуют значения λ_k, отличные от λ_1. Таким образом выходит, что вектор $\boldsymbol{y}$ будет линейно-независимым со всей совокупностью векторов $x^{(k)}$, т. е. мы имеем $(n+1)$ линейно-независимых векторов, что невозможно. Итак, для всякого корня уравнения (144) кратности m уравнение (154) имеет ровно m линейно-независимых вещественных решений.

Подставив в коэффициенты системы (143) вместо λ_k некоторый корень $\lambda=\lambda_0$ кратности m, мы получим однородную систему, имеющую m линейно-независимых решений, т. е. ранг этой системы должен равняться $(n-m)$. Иначе говоря, эта система сведется к $(n-m)$ уравнениям. Возьмем какое-нибудь решение этой системы и умножим его на такой множитель, чтобы сумма квадратов чисел, входящих в это решение, была равна единице. Таким образом получим один вектор, соответствующий взятому корню уравнения $\lambda=\lambda_0$. Чтобы найти следующий вектор, добавим к $(n-m)$ уравнениям нашей системы еще одно уравнение, выражающее ортогональность нового искомого вектора к уже построенному. Таким образом для нахождения составляющих нового вектора будем иметь однородную систему из $(n-m+1)$ уравнений. Взяв какое-нибудь решение этой системы и нормировав его опять к единице (длина вектора равна единице), перейдем к нахождению третьего вектора, соответствующего тому же корню уравнения $\lambda=\lambda_0$. Для этого добавим к основным $(n-m)$ уравнениям системы еще два уравнения, выражающих ортогональность нового искомого вектора к двум, уже полученным, и т. д. — пока не построим ортонормированную систему, состоящую из m векторов, соответствующих корню уравнения $\lambda=\lambda_0$ кратности m. Из указанного построения непосредственно вытекает некоторая произвольность построения основных решений уравнения (142). Если все корни уравнения простые, то эта произвольность сводится лишь к тому, что все составляющие вектора $x^{(k)}$ можно помножить на (-1). Положим теперь, что уравнение (144) имеет корень кратности m. В этом случае соответствующие m единичных ортогональных векторов, являющихся решением уравнения (142), образуют некоторое подпространство R_m измерения m. Мы можем, очевидно, в этом подпространстве выбирать любым образом взаимно ортогональные единичные орты, и все они будут также решениями уравнения (142) при $\lambda=\lambda_0$, т. е. мы можем переходить от одной ортонормированной системы решений к другой, совершая ортогональное преобразование подпространства R_m. Все сказанное относится и к любому другому кратному корню уравнения (144).

Для пояснения сказанного обратимся к той задаче, с которой мы начали предыдущий номер, а именно к задаче приведения уравнения поверхности второго порядка к осям симметрии. Положим для определенности, что эта поверхность есть эллипсоид. Случай разных корней уравнения (144) соответствует тому факту, что все полуоси этого эллипсоида различны. В этом случае единственный произвол в выборе окончательных осей координат сводится к изменению направления этих осей. Если уравнение (144), которое в рассматриваемом случае будет уравнением третьей степени, имеет два одинаковых корня, то эллипсоид будет эллипсоидом вращения, и две оси симметрии могут лежать как угодно в плоскости, проходящей через центр и перпендикулярной к оси вращения, лишь бы они были взаимно ортогональны, т. е. в данном случае произвол в выборе окончательных осей состоит еще в произвольном ортогональном преобразовании в указанной выше плоскости. Наконец, если уравнение (144) имеет все три одинаковых корня, то наш эллипсоид есть сфера, и наше уравнение не содержит членов с произведениями координат. В этом случае мы вообще можем совершенно произвольно выбирать прямолинейные, прямоугольные координатные оси в пространстве.

34. Примеры. Рассмотрим два численных примера.

1. Привести к осям симметрии уравнение поверхности

$$x_1^2+5x_2^2+x_3^2+2x_1x_2+6x_1x_3+2x_2x_3=5.$$

Соответствующая квадратичная форма будет иметь вид:

$$\varphi=x_1^2+x_1x_2+3x_1x_3+x_2x_1+5x_2^2+x_2x_3+3x_3x_1+x_3x_2+x_3^2.$$

Характеристическое уравнение ее матрицы будет

$$\begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix}=0,$$

откуда, разлагая по элементам первой строки:

$$(1-\lambda)[(5-\lambda)(1-\lambda)-1]-(1-\lambda-3)+3[1-3(5-\lambda)]=0$$

или

$$\lambda^3-7\lambda^2+36=0.$$

Это уравнение, как нетрудно проверить, имеет корни

$$\lambda_1=-2;\quad \lambda_2=3;\quad \lambda_3=6,$$

и уравнение нашей поверхности, отнесенное к осям симметрии, будет

$$-2x_1'^2+3x_2'^2+6x_3'^2=5.$$

Будем теперь определять элементы ортогональной матрицы:

$$B=\begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{Vmatrix}.$$

Мы имеем для них систему

$$\left.\begin{aligned} (1-\lambda)b_{1k}+b_{2k}+3b_{3k}=0 \\ b_{1k}+(5-\lambda)b_{2k}+b_{3k}=0 \\ 3b_{1k}+b_{2k}+(1-\lambda)b_{3k}=0. \end{aligned}\right\} \qquad (155)$$

Подставляем сначала $\lambda=\lambda_1=-2$, что приводит нас к двум уравнениям

$$3b_{11}+b_{21}+3b_{31}=0$$
$$b_{11}+7b_{21}+b_{31}=0.$$

Решение этой системы имеет вид:

$$b_{11}=-k_1;\quad b_{21}=0;\quad b_{31}=k_1,$$

где k_1 — произвольное число. Выбираем его так, чтобы сумма квадратов чисел, составляющих решение, была равна единице. Окончательно получаем:

$$b_{11}=\frac{1}{\sqrt{2}};\quad b_{21}=0;\quad b_{31}=-\frac{1}{\sqrt{2}},$$

причем можно взять и решение с обратным знаком.

Подставляем теперь в коэффициенты системы (155) $\lambda=\lambda_2=3$. Получаем систему, в которой третье уравнение есть разность первых двух и, таким образом, приходим к двум уравнениям:

$$-2b_{12}+b_{22}+3b_{32}=0$$
$$b_{12}+2b_{22}+b_{32}=0.$$

Нетрудно найти решение этой системы, нормированное к единице:

$$b_{12}=\frac{1}{\sqrt{3}};\quad b_{22}=-\frac{1}{\sqrt{3}};\quad b_{32}=\frac{1}{\sqrt{3}}.$$

Подставляем, наконец, в коэффициенты системы (155) третий корень. Получаем опять систему, в которой одно из уравнений есть следствие двух других. Решая оставшиеся два уравнения и нормируя полученное решение к единице, будем иметь:

$$b_{13}=\frac{1}{\sqrt{6}};\quad b_{23}=\frac{2}{\sqrt{6}};\quad b_{33}=\frac{1}{\sqrt{6}}.$$

В данном случае формулы преобразования переменных имеют вид:

$$x'_1=\frac{1}{\sqrt{2}}x_1-\frac{1}{\sqrt{2}}x_3$$
$$x'_2=\frac{1}{\sqrt{3}}x_1-\frac{1}{\sqrt{3}}x_2+\frac{1}{\sqrt{3}}x_3$$
$$x'_3=\frac{1}{\sqrt{6}}x_1+\frac{2}{\sqrt{6}}x_2+\frac{1}{\sqrt{6}}x_3.$$

2. Привести к осям симметрии уравнение поверхности

$$2x_1^2+6x_2^2+2x_3^2+8x_1x_3=1.$$

В данном случае квадратичная форма запишется в виде:

$$\varphi=2x_1^2+0x_1x_2+4x_1x_3+0x_2x_1+6x_2^2+0x_2x_3+4x_3x_1+0x_3x_2+2x_3^2,$$

характеристическое уравнение ее матрицы будет

$$\begin{vmatrix} 2-\lambda & 0 & 4 \\ 0 & 6-\lambda & 0 \\ 4 & 0 & 2-\lambda \end{vmatrix}=0.$$

Раскрывая этот определитель, придем к уравнению вида:

$$\lambda^3 - 10\lambda^2 + 12\lambda + 72 = 0.$$

Его корни будут $\lambda_1 = -2$; $\lambda_2 = \lambda_3 = 6$, т. е. это уравнение имеет двойной корень.

Переходим к определению коэффициентов ортогонального преобразования, для которых имеем систему

$$\left.\begin{aligned} (2-\lambda)\,b_{1k} + \quad 4b_{3k} &= 0,\\ (6-\lambda)\,b_{2k} &= 0,\\ 4b_{1k} + (2-\lambda)\,b_{3k} &= 0. \end{aligned}\right\} \tag{155$_1$}$$

Подставляя $\lambda = -2$, получим, как нетрудно вычислить, нормированное к единице решение

$$b_{11} = \frac{1}{\sqrt{2}}; \quad b_{21} = 0; \quad b_{31} = -\frac{1}{\sqrt{2}}.$$

Подставим теперь в коэффициенты системы (155) двойной корень $\lambda = 6$, для которого мы должны получить два линейно-независимых и взаимно-ортогональных решения. При указанной подстановке система сведется к одному уравнению

$$-b_{12} + b_{32} = 0. \tag{155$_2$}$$

Возьмем нормированное к единице решение этого уравнения

$$b_{12} = \frac{1}{\sqrt{2}}; \quad b_{22} = 0; \quad b_{32} = \frac{1}{\sqrt{2}}.$$

Для нахождения второго решения заметим, что оно должно удовлетворять как решению (155$_2$), так и условию ортогональности с уже найденным решением. Таким образом, мы получаем для его нахождения два уравнения

$$\begin{aligned} -b_{13} + \qquad b_{33} &= 0\\ \frac{1}{\sqrt{2}}\,b_{13} + \frac{1}{\sqrt{2}}\,b_{33} &= 0 \end{aligned}$$

или $b_{13} = b_{33} = 0$, откуда нормированное к единице решение будет $b_{13} = 0$; $b_{23} = 1$; $b_{33} = 0$.

Окончательно ортогональное преобразование будет иметь вид:

$$\begin{aligned} x'_1 &= \frac{1}{\sqrt{2}}\,x_1 - \frac{1}{\sqrt{2}}\,x_3\\ x'_2 &= \frac{1}{\sqrt{2}}\,x_1 + \frac{1}{\sqrt{2}}\,x_3\\ x'_3 &= \qquad x_2 \end{aligned}$$

и уравнение поверхности в осях симметрии приведется к виду:

$$-2x'^2_1 + 6\,(x'^2_2 + x'^2_3) = 1.$$

35. Классификация квадратичных форм. Задачу о приведении квадратичной формы к сумме квадратов можно поставить и в более общем виде, чем это мы делали выше, не требуя обязательно, чтобы линейное преобразование от новых переменных к старым было ортогональным, а именно мы можем поставить задачу следующим образом: требуется привести вещественную квадратичную форму (134) к виду:

$$\varphi = \mu_1 X_1^2 + \mu_2 X_2^2 + \ldots + \mu_n X_n^2, \tag{156}$$

где X_k суть какие-нибудь n линейно-независимых вещественных линейных форм переменных x_k. При такой постановке задачи коэффициенты μ_k не являются какими-либо определенными числами, как это мы имели выше, но мы можем все-таки высказать некоторое утверждение относительно этих коэффициентов, а именно: число этих коэффициентов, отличных от нуля, должно всегда равняться рангу таблицы, составленной из коэффициентов a_{ik} квадратичной формы. Иначе говоря, при любом приведении квадратичной формы к сумме квадратов линейных линейно-независимых форм число квадратов равно рангу упомянутой таблицы. Кроме того, имеет место и еще одно свойство, которое обычно называется *законом инерции квадратичных форм*, а именно: при любом преобразовании вещественной квадратичной формы к виду (156), где линейные формы X_k также вещественны, *число положительных коэффициентов μ_k (и число отрицательных коэффициентов μ_k') будет всегда одним и тем же*. Высказанные соображения будут нами доказаны в конце настоящего номера.

Поставленная общая задача о приведении квадратичной формы к виду (156) решается весьма просто выделением полных квадратов. Проведем это на частном примере:

$$\varphi = x_1^2 + 4x_2^2 + x_3^2 + 2x_1x_2 - 6x_1x_3 + 8x_2x_3.$$

Добавляя к членам $(x_1^2 + 2x_1x_2 - 6x_1x_3)$ слагаемые $(x_2^2 + 9x_3^2 - 6x_2x_3)$, мы получаем полный квадрат и можем записать φ в виде:

$$\varphi = (x_1 + x_2 - 3x_3)^2 + 3x_2^2 - 8x_3^2 + 14x_2x_3.$$

Точно так же, выделяя еще один квадрат, мы приходим окончательно к представлению квадратичной формы в виде (156):

$$\varphi = (x_1 + x_2 - 3x_3)^2 - 2\left(2x_3 - \frac{7}{4}x_2\right)^2 + \frac{73}{8}(x_2)^2.$$

Линейные формы, стоящие в круглых скобках, будут, очевидно, линейно-независимыми.

В случае отсутствия в выражении φ квадратов переменных вычисление надо проводить несколько иначе. Пусть мы имеем:

$$\varphi = ax_1x_2 + Px_1 + Qx_2 + R,$$

где a — численный коэффициент, отличный от нуля, P и Q — линейные формы переменных, не содержащие x_1 и x_2, R — квадратичная форма, также не содержащая x_1 и x_2. Мы можем написать:

$$\varphi = a\left(x_1 + \frac{Q}{a}\right)\left(x_2 + \frac{P}{a}\right) + R - \frac{PQ}{a^2}.$$

Если положить:

$$X_1 = \frac{1}{2}\left(x_1 + x_2 + \frac{P+Q}{a}\right); \quad X_2 = \frac{1}{2}\left(x_1 - x_2 - \frac{P-Q}{a}\right)$$

и

$$\varphi_1 = R - \frac{PQ}{a^2},$$

то получим:

$$\varphi = aX_1^2 - aX_2^2 + \varphi_1,$$

где φ_1 — квадратичная форма, не содержащая x_1 и x_2. Выделив два квадрата, мы освободились от двух переменных.

Приведение квадратичной формы к виду (156) дает возможность естественной классификации таких форм. Рассмотрим ряд случаев.

I. Положим, что все коэффициенты μ_k в формуле (156) положительны. В этом случае форма называется *определенно положительной*. Нетрудно показать, что она имеет положительные значения при всех вещественных значениях x_k и может обращаться в нуль только тогда, когда все x_k равны нулю. Действительно, для того чтобы правая часть формулы (156) обратилась в нуль, необходимо и достаточно, ввиду положительности всех μ_k, чтобы все линейные формы x_k были равны нулю. Мы получаем, таким образом, для x_k систему n однородных уравнений с определителем, отличным от нуля (формы линейно-независимы), и эта система имеет, следовательно, только нулевое решение.

II. Если все коэффициенты μ_k отрицательны, то квадратичная форма называется *определенно отрицательной*. Как и выше, можно показать, что она имеет при всяких вещественных x_k только отрицательные значения, причем обращается в нуль только тогда, когда все x_k равны нулю.

III. Рассмотрим теперь тот случай, когда среди коэффициентов μ_k есть равные нулю, а все не равные нулю определенного знака, например, положительны. В этом случае мы будем иметь для формы φ представление вида:

$$\varphi = \mu_1 X_1^2 + \ldots + \mu_m X_m^2 \qquad (m < n), \tag{156_1}$$

где все μ_k положительны. Здесь опять значения нашей формы не могут быть отрицательными ни при каких значениях x_k, но могут равняться нулю и тогда, когда значения x_k отличны от нуля. Действительно, чтобы получить нулевое значение формы, мы должны написать систему m однородных уравнений для x_k:

$$X_1 = X_2 = \ldots = X_m = 0,$$

и так как $m < n$, то эта система наверно имеет решения, отличные от нулевого. Точно так же, если в формуле (156_1) все коэффициенты μ_k отрицательны, то квадратичная форма не может иметь положительных значений, но может обращаться в нуль и при значениях x_k, отличных от нуля. В рассматриваемом случае форма называется *знакопостоянной* — положительной или отрицательной.

IV. Наконец, если среди коэффициентов μ_k формулы (156) имеются как положительные, так и отрицательные, то, как нетрудно видеть, квадратичная форма может принимать как положительные, так и отрицательные значения при вещественных значениях x_k. В этом случае она называется *знакопеременной*.

Предыдущая классификация вещественных квадратичных форм имеет непосредственное приложение к задаче на maxima и minima функции от нескольких переменных. Пусть имеется функция n независимых переменных $x_1, \ldots, x_n$:

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

причем при значениях $x_1 = \ldots = x_n = 0$ выполнены необходимые условия maxima и minima, т. е. все частные производные от функции ψ по независимым переменным обращаются в нуль. Разлагая нашу функцию в ряд Маклорена, будем иметь:

$$\psi(x_1, \ldots, x_n) - \psi(0, \ldots, 0) = \varphi(x_1, \ldots, x_n) + \omega,$$

где через $\varphi(x_1, \ldots, x_n)$ мы обозначили квадратичную форму переменных x_k, а через ω — совокупность членов измерения выше второго относительно x_k. Если квадратичная форма φ определенно положительна, то мы имеем минимум функции в точке $x_1 = \ldots = x_n = 0$. Если она определенно отрицательна, то мы имеем максимум. Если она знакопеременна, то мы не имеем ни минимума, ни максимума, и, наконец, если φ — знакопостоянная форма, то мы имеем дело с сомнительным случаем. Этот результат является естественным дополнением к тому, который мы имели в [I, **163**] для случая функции от двух независимых переменных.

Переходим к доказательству предложений, высказанных в начале настоящего параграфа. Пусть имеется квадратичная форма:

$$\varphi = \sum_{i, k=1}^{n} a_{ik} x_i x_k \qquad (a_{ik} = a_{ki}),$$

причем r есть ранг таблицы ее коэффициентов. Составим систему n линейных форм:

$$\frac{1}{2} \frac{\partial \varphi}{\partial x_s} = \sum_{l=1}^{n} a_{sl} x_l \qquad (s = 1, 2, \ldots, n). \tag{157}$$

При составлении выражений этих частных производных мы пользовались условиями $a_{ik} = a_{ki}$. Число r есть, очевидно, ранг системы форм (157) в смысле [11].

Положим, что φ приводится к сумме m квадратов линейно-независимых форм:

$$y_s = \beta_{s1}x_1 + \beta_{s2}x_2 + \ldots + \beta_{sn}x_n \quad (s = 1, 2, \ldots, m), \tag{158}$$

т. е.

$$\varphi = \mu_1 y_1^2 + \mu_2 y_2^2 + \ldots + \mu_m y_m^2, \tag{159}$$

где μ_s отличны от нуля. Нам надо доказать, что $m = r$. Пользуясь выражением (159), составим линейные формы (157):

$$\frac{1}{2}\frac{\partial\varphi}{\partial x_s} = \mu_1\beta_{1s}y_1 + \mu_2\beta_{2s}y_2 + \ldots + \mu_m\beta_{ms}y_m \quad (s = 1, 2, \ldots, n). \tag{157$_1$}$$

Переменные y_s могут принимать любые значения, поскольку формы (158) линейно-независимы. Поэтому при определении линейной зависимости форм (157_1) y_s можно считать за независимые переменные, и наибольшее число линейно-независимых форм в системе (157_1) должно быть равно рангу таблицы коэффициентов $\mu_k\beta_{ki}$, где номер столбца k принимает возможные значения: $k = 1, 2, \ldots, m$, и номер строки $i = 1, 2, \ldots, n$. Элементы каждого столбца этой таблицы содержат общий множитель μ_k, отличный от нуля, и потому ранг таблицы $\mu_k\beta_{ki}$ совпадает с рангом таблицы β_{ki}. Поскольку система m форм (158) есть система линейно-независимых форм, этот ранг равен m, т. е. наибольшее число линейно-независимых форм в системе (157_1) или (157) равно m. С другой стороны, по условию, это число равно r, откуда и следует, что $m = r$.

Покажем теперь, что при любом способе представления φ формулой вида (159), где y_s — вещественные линейно-независимые формы, число положительных и отрицательных коэффициентов μ_s всегда одно и то же. Будем доказывать это от противного. Положим, что мы имеем два представления φ формулами вида (159), причем число положительных коэффициентов в этих представлениях различно:

$$\left.\begin{aligned} \varphi &= \lambda_1 y_1^2 + \ldots + \lambda_p y_p^2 - \lambda_{p+1}y_{p+1}^2 - \ldots - \lambda_m y_m^2, \\ \varphi &= \lambda_1' y_1'^2 + \ldots + \lambda_q' y_q'^2 - \lambda_{q+1}' y_{q+1}'^2 - \ldots - \lambda_m' y_m'^2. \end{aligned}\right\} \tag{160}$$

В этих формулах λ_s и λ_s' мы считаем положительными. Формы $y_1, \ldots, y_m$ — линейно-независимы, то же можно сказать о формах $y_1', \ldots, y_m'$. Раз $p \neq q$, то всегда можно считать, что, например, $p < q$. Покажем, что это приведет нас к нелепости. Присоединим к формам $y_1, \ldots, y_m$ формы $y_{m+1}, \ldots, y_n$ так, чтобы получилась полная система линейно-независимых форм [11]. Напишем систему линейных однородных уравнений для $x_1, x_2, \ldots, x_n$:

$$y_1 = 0; \ldots; y_p = 0; \ y_{q+1}' = 0; \ldots; y_m' = 0; \ y_{m+1} = 0; \ldots; y_n = 0. \tag{161}$$

Число этих однородных уравнений $p + (m - q) + (n - m) = n - (q - p)$, и так как $q - p > 0$, то это число меньше n. Следовательно, написанная однородная система имеет вещественные решения, отличные от нулевого. Возьмем какое-нибудь из этих решений: $x_s = x_s^{(0)}$ $(s = 1, 2, \ldots, n)$. При этих значениях x_s мы будем иметь, в силу (161):

$$\varphi = -\lambda_{p+1}y_{p+1}^2 - \ldots - \lambda_m y_m^2 = \lambda_1' y_1'^2 + \ldots + \lambda_q' y_q'^2.$$

Отсюда видно, что при $x_s = x_s^{(0)}$ квадратичная форма φ должна обращаться в нуль, и, следовательно, $x_s = x_s^{(0)}$, кроме уравнений (161), должны удовлетворять уравнениям:

$$y_{p+1} = 0; \ldots; y_m = 0.$$

Окончательно получаем, что $x_s = x_s^{(0)}$ должны обращать в нуль все формы полной системы линейно-независимых форм: $y_1, y_2, \ldots, y_n$. Но этого не может быть, поскольку в однородной системе относительно $x_1, x_2, \ldots, x_n$:

$$y_1 = 0;\ y_2 = 0; \ldots;\ y_n = 0$$

определитель отличен от нуля, ибо формы y_s линейно-независимы. Мы пришли к противоречию, что и доказывает закон инерции.

36. Формула Якоби. Приведем без доказательства формулу Якоби, дающую в удобном виде приведение квадратичной формы к сумме квадратов.

Для этого введем сначала некоторые обозначения. Положим:

$$A_i(x) = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, \ldots, n),$$

$$\Delta_0 = 1;\ \Delta_1 = a_{11};\ \Delta_k = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix},$$

$$(k = 2, 3, \ldots, n)$$

$$X_1 = A_1(x);\ X_k = \begin{vmatrix} a_{11} & \ldots & a_{1k-1} & A_1(x) \\ a_{21} & \ldots & a_{2k-1} & A_2(x) \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & \ldots & a_{kk-1} & A_k(x) \end{vmatrix}.$$

Если ранг таблицы коэффициентов a_{ik} равен r и определители $\Delta_1, \Delta_2, \ldots, \Delta_r$ отличны от нуля, то формула Якоби имеет вид:

$$\varphi = \sum_{i,k=1}^{n} a_{ik} x_i x_k = \sum_{k=1}^{r} \frac{X_k^2}{\Delta_k \Delta_{k-1}}, \tag{162}$$

причем линейные формы X_k $(k = 1, 2, \ldots, r)$ — линейно-независимы. Последняя формула дает возможность по знакам Δ_k определить, к какому типу в отношении закона инерции принадлежит форма φ.

В частности, если все определители $\Delta_1, \Delta_2, \ldots, \Delta_n$ положительны (при этом $r = n$), из (162) следует, что φ — определенно положительна. Можно доказать и обратное предложение — если φ определенно положительная форма, то все указанные определители должны быть положительны. При применении формулы (162) можно, конечно, нумеровать переменные x_s в любом порядке. При перемене нумерации будут меняться, конечно, и указанные выше определители Δ_k, и каждый из главных миноров матрицы $| a_{ik} |_1^n$ может быть определителем из последовательности Δ_k при определенной нумерации переменных x_s. Из сказанного выше следует, что у определенно положительной формы φ все главные миноры положительны, но при этом достаточно убедиться в положительности определителей

$$\Delta_s \ (s = 1, 2, \ldots, n).$$

Можно показать, что для того, чтобы форма φ была положительной, необходимо и достаточно, чтобы все главные миноры были неотрицательны, т. е. были больше нуля или равны нулю. Здесь недостаточно определения знаков только определителей Δ_s, а нужно определять знаки всех главных миноров.

Для того чтобы форма φ была определенно отрицательной, необходимо и достаточно соблюдение неравенств $(-1)^k \Delta_k > 0$ $(k=1, 2, \ldots, n)$. Для того чтобы φ была отрицательной, необходимо и достаточно, чтобы главные миноры имели знак $(-1)^k$, где k — порядок минора, или были равны нулю.

Доказательство указанных в настоящем номере предложений можно найти в книге Ф. Р. Гантмахера „Теория матриц" (1953 г.).

37. Одновременное приведение двух квадратичных форм к сумме квадратов. Пусть имеются две квадратичные формы:

$$\varphi_1 = \sum_{i,\,k=1}^{n} a_{ik} x_i x_k; \quad \varphi_2 = \sum_{i,\,k=1}^{n} b_{ik} x_i x_k,$$

причем φ_1 определенно положительна, т. е. приводится к сумме n положительных квадратов. Требуется найти такое линейное преобразование (не обязательно ортогональное), чтобы в результате его обе формы перешли в сумму квадратов.

Прежде всего, введем такие новые переменные y_k, чтобы форма φ_1 перешла в сумму квадратов. Это можно сделать, например, элементарным приемом, указанным в предыдущем номере. В новых переменных будем иметь следующие представления квадратичных форм:

$$\varphi_1 = \sum_{k=1}^{n} \mu_k y_k^2; \quad \varphi_2 = \sum_{i,\,k=1}^{n} b'_{ik} y_i y_k.$$

По условию все числа μ_k положительны, и мы можем ввести новые вещественные переменные $z_k = \sqrt{\mu_k} y_k$. При этом получим формулы вида:

$$\varphi_1 = \sum_{k=1}^{n} z_k^2; \quad \varphi_2 = \sum_{i,\,k=1}^{n} b''_{ik} z_i z_k.$$

Совершим ортогональное преобразование от переменных z_k к новым переменным z'_k, приводящее форму φ_2 к сумме квадратов.

При этом φ_1 останется суммой квадратов, так как преобразование ортогонально, и мы будем иметь окончательно обе формы в виде суммы квадратов:

$$\varphi_1 = \sum_{k=1}^{n} z_k'^2; \qquad \varphi_2 = \sum_{k=1}^{n} \lambda_k z_k'^2.$$

Числа λ_k называются иногда *характеристическими числами формы* φ_2 *по отношению к форме* φ_1.

Установим теперь то уравнение, которому должны удовлетворять эти числа λ_k и которое будет вполне аналогично уравнению (144) из [**32**]. Для этого введем понятие о *дискриминанте* квадратичной формы, а именно: *дискриминантом квадратичной формы называется определитель, составленный из ее коэффициентов*.

Положим, что мы преобразуем форму φ с матрицей коэффициентов A к новым переменным при помощи преобразования

$$(x_1, \ldots, x_n) = B(x'_1, \ldots, x'_n).$$

Матрица новой формы будет, как известно [**32**]:

$$C = B^{(*)} AB,$$

и ее определитель вычисляется по формуле

$$D(C) = D(B^{(*)}) D(A) D(B).$$

Определители $D(B^{(*)})$ и $D(B)$, очевидно, равны, так как соответствующие таблицы получаются одна из другой лишь заменой строк столбцами. Мы имеем таким образом

$$D(C) = D(A) D(B)^2,$$

т. е. при линейном преобразовании переменных в квадратичной форме дискриминант формы умножается на квадрат определителя преобразования от новых переменных к первоначальным.

Вернемся теперь к нашим квадратичным формам φ_1, φ_2 и составим квадратичную форму

$$\omega = \varphi_2 - \lambda\varphi_1 = \sum_{i, k=1}^{n} (b_{ik} - \lambda a_{ik}) x_i x_k,$$

коэффициенты которой содержат параметр λ.

В результате преобразования к новым переменным эта форма будет иметь вид:

$$\omega = \sum_{k=1}^{n} (\lambda_k - \lambda) z_k'^2,$$

и ее дискриминант в новых переменных выражается, очевидно, произведением

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda), \tag{163}$$

а в старых переменных этот дискриминант будет равен определителю с элементами $(b_{ik} - \lambda a_{ik})$. Как мы показали, эти два дискриминанта будут отличаться лишь множителем — квадратом определителя преобразования, не содержащим λ и отличным от нуля. Отсюда непосредственно следует, что оба дискриминанта имеют одинаковые корни относительно параметра λ. Принимая во внимание (163), видим, что числа λ_k суть корни уравнения

$$\begin{vmatrix} b_{11} - \lambda a_{11} & b_{12} - \lambda a_{12} \ldots b_{1n} - \lambda a_{1n} \\ b_{21} - \lambda a_{21} & b_{22} - \lambda a_{22} \ldots b_{2n} - \lambda a_{2n} \\ \ldots & \ldots \\ b_{n1} - \lambda a_{n1} & b_{n2} - \lambda a_{n2} \ldots b_{nn} - \lambda a_{nn} \end{vmatrix} = 0.$$

38. Малые колебания. Мы видели выше [**II, 20**], что движение механической системы, имеющей n степеней свободы, связи которой не содержат времени и которая находится под воздействием сил, имеющих потенциал, определяется системой дифференциальных уравнений вида:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial q_k'}\right) - \frac{\partial T}{\partial q_k} = \frac{\partial U}{\partial q_k}\ (k = 1,\ 2, \ldots,\ n), \tag{164}$$

где T — кинетическая энергия системы и U — заданная функция (силовая функция) от q_k, которую мы считаем не зависящей от t. Как мы упоминали выше, T есть квадратичная форма от производных q'_k от q_k по времени

$$T = \sum_{i,\ k=1}^{n} a_{ik} q_i' q_k' \qquad (a_{ki} = a_{ik}), \tag{165}$$

причем коэффициенты суть заданные функции от q_k. Положим, что значения $q_k = 0$ обращают в нуль частные производные

$$\frac{\partial U}{\partial q_k} = 0 \text{ при } q_1 = \ldots = q_n = 0 \qquad (k = 1, 2, \ldots, n). \tag{166}$$

При этом система дифференциальных уравнений (164) имеет очевидное решение $q_1 = \ldots = q_n = 0$, которому соответствует некоторое положение равновесия системы. Функция U определена лишь с точностью до постоянного слагаемого, и мы можем всегда считать, что она обращается в нуль при $q_1 = \ldots = q_n = 0$. В силу условия (166) можно, таким образом, утверждать, что разложение функции U по степеням q_k начинается лишь с членов второго измерения. Положим, что квадратичная форма, получаемая от этих членов второго измерения, будет определенно отрицательной, откуда следует, что U имеет максимум при $q_1 = \ldots = q_n = 0$, или — что то же — потенциальная энергия $(-U)$ имеет минимум. Как мы доказали [**II, 20**] при этом положение равновесия $q_1 = \ldots = q_n = 0$ будет устойчивым, и при малых начальных возмущениях система будет совершать малые колебания около упомянутого положения равновесия, так что q_k во все время движения будут оставаться малыми. Мы можем поэтому при исследовании этих малых колебаний считать, что U сводится лишь к членам второго измерения, т. е. имеет вид:

$$-U = \sum_{i,\ k=1}^{n} b_{ik} q_i q_k \qquad (b_{ki} = b_{ik}). \tag{167}$$

Точно так же мы можем в коэффициентах a_{ik} выражения (165) положить приближенно $q_k = 0$, после чего эти коэффициенты окажутся заданными числами. Подставляя все это в систему (164), будем иметь систему n линейных уравнений с постоянными коэффициентами:

$$\left.\begin{array}{l}
a_{11}q_1'' + a_{12}q_2'' + \ldots + a_{1n}q_n'' + b_{11}q_1 + b_{12}q_2 + \ldots + b_{1n}q_n = 0 \\
a_{21}q_1'' + a_{22}q_2'' + \ldots + a_{2n}q_n'' + b_{21}q_1 + b_{22}q_2 + \ldots + b_{2n}q_n = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}q_1'' + a_{n2}q_2'' + \ldots + a_{nn}q_n'' + b_{n1}q_1 + b_{n2}q_2 + \ldots + b_{nn}q_n = 0.
\end{array}\right\} \tag{168}$$

Если мы будем искать решение этой системы в форме гармонических колебаний одной и той же частоты и начальной фазы, но с разными амплитудами

$$q_k = A_k \cos(\lambda t + \varphi) \qquad (k = 1, 2, \ldots, n), \tag{169}$$

то, подставляя в систему (168), будем иметь систему уравнений для A_k и λ:

$$\left.\begin{array}{l} (b_{11} - \lambda^2 a_{11}) A_1 + (b_{12} - \lambda^2 a_{12}) A_2 + \ldots + (b_{1n} - \lambda^2 a_{1n}) A_n = 0 \\ (b_{21} - \lambda^2 a_{21}) A_1 + (b_{22} - \lambda^2 a_{22}) A_2 + \ldots + (b_{2n} - \lambda^2 a_{2n}) A_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ (b_{n1} - \lambda^2 a_{n1}) A_1 + (b_{n2} - \lambda^2 a_{n2}) A_2 + \ldots + (b_{nn} - \lambda^2 a_{nn}) A_n = 0. \end{array}\right\} \tag{170}$$

Для того чтобы эта система имела для A_k решение, отличное от нулевого, мы должны приравнять ее определитель нулю:

$$\begin{vmatrix} b_{11} - \lambda^2 a_{11} & b_{12} - \lambda^2 a_{12} & \ldots & b_{1n} - \lambda^2 a_{1n} \\ b_{21} - \lambda^2 a_{21} & b_{22} - \lambda^2 a_{22} & \ldots & b_{2n} - \lambda^2 a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} - \lambda^2 a_{n1} & b_{n2} - \lambda^2 a_{n2} & \ldots & b_{nn} - \lambda^2 a_{nn} \end{vmatrix} = 0. \tag{171}$$

Взяв некоторый корень этого уравнения и подставив в коэффициенты системы (170), мы получим для A_k решения — одно или несколько, которые мы затем можем умножить на произвольную постоянную. Кроме того, формула (169) содержит еще произвольную постоянную φ.

Мы получаем более отчетливое решение задачи, применяя теорию квадратичных форм. Заметим прежде всего, что по самому существу дела квадратичная форма (165) переменных q'_k, выражающая кинетическую энергию при движении, будет определенно положительной формой. Кроме того, в данном случае по условию задачи и форма (167) будет определенно положительной. Как мы видели, можно ввести вместо переменных q_k такие новые переменные p_k, связанные с прежним линейным преобразованием с постоянными коэффициентами, чтобы в новых переменных квадратичные формы T и $(-U)$ одновременно привелись к сумме квадратов, причем для формы T это должно быть чистой суммой квадратов с коэффициентами, равными единице. Заметим при этом, что линейная зависимость для p_k и q_k приводит к такой же зависимости для p'_k и q'_k. Мы будем, таким образом, иметь:

$$T = \sum_{s=1}^{n} p'^2_s; \quad -U = \sum_{s=1}^{n} \lambda^2_s p^2_s, \tag{172}$$

где все коэффициенты при p^2_s положительны, так что мы имели возможность обозначить их через квадраты. Вместо системы (168) мы можем написать уравнения Лагранжа (164) для новых переменных

$$\frac{d}{dt}\left[\frac{\partial T}{\partial p'_k}\right] = \frac{\partial U}{\partial p_k}.$$

Подставляя (172), будем иметь чрезвычайно простую систему

$$p''_k + \lambda^2_k p_k = 0 \qquad (k = 1, 2, \ldots, n).$$

Решения этой системы будут:

$$p_k = C_k \cos(\lambda_k t + \psi_k) \qquad (k = 1, 2, \ldots, n), \tag{173}$$

где C_k и ψ_k — произвольные постоянные. Обобщенные координаты p_k называются главными координатами нашей механической системы.

Основные координаты q_k выражаются через них линейным образом с постоянными коэффициентами. Из результатов предыдущего номера непосредственно следует, что числа λ_k должны быть корнями уравнения (170). Заметим, что среди них могут оказаться и одинаковые, но и в этом случае формулы (169) дают общее решение задачи малых колебаний в рассматриваемом случае.

39. Экстремальные свойства собственных значений квадратичной формы. Рассмотрим задачу приведения вещественной квадратичной формы к сумме квадратов с новой точки зрения. Для простоты ограничимся случаем трех переменных

$$\varphi = \sum_{i,\, k=1}^{3} a_{ik} x_i x_k = \sum_{k=1}^{3} \lambda_k x'^2_k, \tag{174}$$

где x'_k связаны с x_k некоторым ортогональным преобразованием

$$\left.\begin{aligned} x_1 &= b_{11}x'_1 + b_{12}x'_2 + b_{13}x'_3 \\ x_2 &= b_{21}x'_1 + b_{22}x'_2 + b_{23}x'_3 \\ x_3 &= b_{31}x'_1 + b_{32}x'_2 + b_{33}x'_3. \end{aligned}\right\} \tag{175}$$

Для определенности будем считать, что числа λ_k идут в убывающем порядке, т. е.

$$\lambda_1 > \lambda_2 > \lambda_3. \tag{176}$$

Нашей задачей будет определение чисел λ_k и коэффициентов b_{ik} по значениям формы φ на единичной сфере K, т. е. на сфере с центром в начале координат и радиусом единица:

$$x_1^2 + x_2^2 + x_3^2 = 1 \quad \text{или} \quad x'^2_1 + x'^2_2 + x'^2_3 = 1. \tag{177}$$

Каждая точка такой сферы характеризует некоторое направление в пространстве, определяемое единичным вектором, идущим из начала в упомянутую точку. Мы можем написать формулу (174) в виде:

$$\varphi = \lambda_1 (x'^2_1 + x'^2_2 + x'^2_3) + (\lambda_2 - \lambda_1) x'^2_2 + (\lambda_3 - \lambda_1) x'^2_3,$$

откуда видно, что на упомянутой единичной сфере K будем иметь:

$$\varphi = \lambda_1 + (\lambda_2 - \lambda_1) x'^2_2 + (\lambda_3 - \lambda_1) x'^2_3.$$

Отсюда непосредственно следует, что λ_1 есть максимум φ на K. Этот максимум достигается, очевидно, в точке

$$x'_1 = 1; \qquad x'_2 = x'_3 = 0,$$

или, в силу (175), в прежних координатах это будет точка сферы K с координатами:

$$x_1 = b_{11}; \qquad x_2 = b_{21}; \qquad x_3 = b_{31}.$$

Эта точка определяет вектор, соответствующий первому столбцу ортогонального преобразования (175), т. е. вектор, являющийся решением уравнения

$$Ax = \lambda x \tag{178}$$

при $\lambda = \lambda_1$. Итак, первое по величине собственное значение квадратичной формы (174) равно максимуму значения этой формы на единичной сфере, а соответствующий собственный вектор $x^{(1)}$, являющийся решением уравнения (178), есть вектор, идущий из начала в ту точку единичной сферы, где упомянутый максимум достигается.

Перейдем теперь к определению второго собственного значения и соответствующего собственного вектора. Положим в формуле $x_1'=0$. Этому уравнению соответствует плоскость, проходящая через начало и перпендикулярная к вектору $\boldsymbol{x}^{(1)}$. Сечение этой плоскости с единичной сферой даст окружность

$$x_2'^2+x_3'^2=1.$$

На этой окружности мы имеем:

$$\varphi=\lambda_2 x_2'^2+\lambda_3 x_3'^2,$$

откуда непосредственно видно, что λ_2 есть максимум значений φ на единичной сфере при условии перпендикулярности соответствующих векторов уже найденному вектору $\boldsymbol{x}^{(1)}$. Точно так же, как и выше, докажем, что собственный вектор $\boldsymbol{x}^{(2)}$, т. е. решение уравнения (178) при $\lambda=\lambda_2$, есть вектор, идущий в ту точку, где этот максимум достигается.

Имея два вектора, мы получим и третий $\boldsymbol{x}^{(3)}$, как перпендикулярный к ним обоим, а собственное значение λ_3 будет значением формы φ в той точке единичной сферы, в которой она пересекается с вектором $\boldsymbol{x}^{(3)}$.

Если бы, например, мы имели $\lambda_1=\lambda_2$, то при отыскании первого максимума формы φ на единичной сфере мы получили бы не точку, а целую окружность, где этот максимум достигается.

Предыдущее рассуждение легко переносится и на случай любого числа измерений. Мы приведем для этого общего случая лишь результат, вполне аналогичный предыдущему. Пусть имеется вещественная квадратичная форма n переменных:

$$\varphi=\sum_{i,\ k=1}^{n} a_{ik}x_i x_k. \tag{179}$$

Единичный вектор в вещественном n-мерном пространстве будет изображаться совокупностью вещественных чисел, сумма квадратов которых равна единице. Мы будем говорить, что концы таких векторов лежат на единичной сфере, и уравнение этой единичной сферы будет, очевидно:

$$x_1^2+x_2^2+\ldots+x_n^2=1. \tag{180}$$

Наибольшее характеристическое число формы φ будет равно максимуму этой формы на единичной сфере (180), и соответственный собственный вектор будет определяться вектором $\boldsymbol{x}^{(1)}$, идущим из начала в ту точку сферы, где φ достигает максимума Для получения второго по величине характеристического числа будем рассматривать единичные векторы, перпендикулярные к уже найденному вектору $\boldsymbol{x}^{(1)}$. Среди них найдется такой $\boldsymbol{x}^{(2)}$,который даст наибольшее значение формы φ. Этот второй максимум λ_2 и будет равен второму собственному значению формы, а упомянутый вектор $\boldsymbol{x}^{(2)}$ будет соответствующим собственным вектором. Рассмотрим теперь единичные векторы, перпендикулярные к $\boldsymbol{x}^{(1)}$ и $\boldsymbol{x}^{(2)}$, что равносильно присоединению к условию (180) еще двух условий:

$$(\boldsymbol{x}^{(1)},\ \boldsymbol{x})=0 \text{ и } (\boldsymbol{x}^{(2)},\ \boldsymbol{x})=0.$$

Среди них найдется такой, который опять даст форме φ наибольшее значение. Это значение и будет третьим по величине собственным значением формы λ_3, а упомянутый вектор будет соответствующим собственным вектором и т. д.

Мы могли бы расположить собственные значения квадратичной формы не в убывающем, а в возрастающем порядке, так что первое собственное значение было бы наименьшим, следующее — вторым в порядке возрастания и т. д. При этом мы получили бы задачи, совершенно аналогичные предыдущим, но только везде, где говорилось о наибольшем значении, нам пришлось бы говорить о наименьшем значении.

Все предыдущие рассуждения обобщаются также и на случай одновременного приведения двух квадратичных форм к сумме квадратов. Пусть две квадратичные формы:

$$\varphi = \sum_{i,\ k=1}^{n} a_{ik} x_i x_k; \qquad \psi = \sum_{i,\ k=1}^{n} b_{ik} x_i x_k$$

приводятся к сумме квадратов

$$\varphi = \sum_{k=1}^{n} x'^2_k; \qquad \psi = \sum_{k=1}^{n} \lambda_k x'^2_k$$

при помощи линейного преобразования

$$(x_1, \ldots, x_n) = B(x'_1, \ldots, x'_n),$$

причем мы считаем, что числа λ_k идут в убывающем порядке.

При этом λ_1 будет наибольшим значением ψ при условии, что $\varphi = 1$, причем это наибольшее значение будет достигаться как раз при

$$x_1 = b_{11};\ x_2 = b_{21};\ \ldots;\ x_n = b_{n1}.$$

Аналогично определяются и следующие собственные значения

40. Эрмитовские матрицы и формы Эрмита. В предыдущих параграфах мы рассматривали вещественные симметричные матрицы и отметили, что они являются частным случаем эрмитовских матриц, элементы которых суть комплексные числа, удовлетворяющие соотношению

$$a_{ki} = \bar{a}_{ik}. \tag{181}$$

При $i = k$ это соотношение показывает, что диагональные элементы a_{kk} должны быть вещественными.

Иначе можно формулировать определение эрмитовской матрицы так: эрмитовская матрица не меняется, если в ней строки заменить столбцами и все элементы сопряженными, т. е. в обозначениях из [**26**]:

$$\bar{A}^{(*)} = A \text{ или } \tilde{A} = A. \tag{182}$$

Матрица $\tilde{A}$, как мы знаем, называется эрмитовски сопряженной с A. Поэтому эрмитовские матрицы иначе называются самосопряженными.

Выше мы показали [**32**], что эрмитовская матрица A удовлетворяет при любых векторах $\boldsymbol{x}$ и $\boldsymbol{y}$ соотношению

$$(A\boldsymbol{x},\ \boldsymbol{y}) = (\boldsymbol{x},\ A\boldsymbol{y}). \tag{183}$$

Это соотношение, как и два предыдущих, может служить определением эрмитовской матрицы.

Отметим еще одно свойство эрмитовских матриц.

Пусть A — эрмитовская матрица и U — любая унитарная матрица. Нетрудно показать, что $U^{-1}AU$ будет, как и A, эрмитовской матрицей. По условию $\bar{A}^{(*)}=A$. Надо доказать, что таким же свойством обладает и $U^{-1}AU$. Мы имеем [26]:

$$\overline{(U^{-1}AU)}^{(*)}=\bar{U}^{(*)}A^{(*)}\bar{U}^{(*)-1}$$

или, принимая во внимание условие для A и унитарный характер U, откуда следует $\bar{U}^{(*)}=U^{-1}$, получим:

$$\overline{(U^{-1}AU)}^{(*)}=U^{-1}AU,$$

что и требовалось доказать.

Можно сказать, что при унитарном преобразовании координат, которое осуществляется для составляющих вектора формулой

$$(x_1, \ldots, x_n)=U(x'_1, \ldots, x'_n),$$

эрмитовская матрица A, как оператор линейного преобразования пространства, будет выглядеть в новых координатах в виде $U^{-1}AU$, и, следовательно, доказанное предложение можно еще формулировать так: унитарные преобразования пространства не меняют эрмитовского характера матрицы как оператора.

Поставим теперь задачу о приведении эрмитовской матрицы к диагональной форме при помощи унитарного преобразования

$$U^{-1}AU=[\lambda_1, \ldots, \lambda_n]. \tag{184}$$

Как и выше для вещественных симметричных матриц, наша задача равносильна задаче решения уравнения вида:

$$A\boldsymbol{x}=\lambda\boldsymbol{x}, \tag{185}$$

где λ есть одно из чисел λ_k и составляющие вектора $\boldsymbol{x}$ дают элементы соответствующего столбца матрицы U.

Эти числа λ_k и соответствующие им векторы $\boldsymbol{x}^{(k)}$ называются собственными значениями и собственными векторами матрицы A.

Собственные значения, как мы знаем, должны быть обязательно корнями уравнения

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix}=0. \tag{186}$$

Пусть $\lambda=\lambda_1$ — некоторый корень этого уравнения и $\boldsymbol{x}^{(1)}$ — некоторое решение уравнения (185) при $\lambda=\lambda_1$.

Это уравнение будет линейным и однородным, так что его решение можно умножать на любую постоянную, и мы можем поэтому считать длину вектора $\boldsymbol{x}^{(1)}$ равной единице. Возьмем этот вектор за первый орт новой координатной системы и достроим каким-нибудь образом еще $(n-1)$ единичных векторов так, чтобы в общем получить ортонормированную систему векторов. Примем эти векторы за новые орты, и пусть U_1 — то унитарное преобразование, которое соответствует переходу к этим новым ортам. В новой координатной системе наша эрмитовская матрица A перейдет в новую эрмитовскую матрицу $A_1 = U_1^{-1}AU_1$, причем соответствующее уравнение

$$A_1 x = \lambda x$$

должно при $\lambda = \lambda_1$ иметь в качестве решения вектор с составляющими $(1, 0, \ldots, 0)$. Как и в [33], это обстоятельство покажет нам, что у матрицы A_1 все элементы первой строки и первого столбца обратятся в нуль, кроме элемента, стоящего на пересечении первой строки и первого столбца и равного λ_1.

Из эрмитовского характера матрицы A_1 непосредственно следует, что этот элемент λ_1 должен быть числом вещественным, и отсюда, между прочим, непосредственно следует, что все корни уравнения (186) должны быть вещественными, что мы видели и раньше. Итак, матрица A_1 будет иметь вид:

$$\begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{Vmatrix},$$

т. е. это будет квазидиагональная матрица вида

$$[\lambda_1, C_1],$$

где через C_1 мы обозначили эрмитовскую матрицу порядка $(n-1)$ с элементами $a_{ik}^{(1)}$. Повторяя предыдущее рассуждение, мы сможем при помощи некоторого унитарного преобразования U_2, произведенного над ортами, кроме первого орта, привести матрицу C_1 к такому виду, при котором все элементы ее первой строки и первого столбца будут нули, кроме элемента, стоящего на их пересечении.

Упомянутое унитарное преобразование мы можем рассматривать как унитарное преобразование во всем нашем n-мерном пространстве. Это будет квазидиагональная унитарная матрица вида:

$$[1, U_2].$$

В результате этого преобразования наша эрмитовская матрица перейдет в эрмитовскую матрицу вида:

$$U_2]^{-1}[\lambda_1, C_1][1, U_2] = [\lambda_1, U_2^{-1}C_1U_2],$$

и в раскрытом виде эта матрица будет

$$\left\| \begin{matrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{matrix} \right\|.$$

Продолжая рассуждать так и дальше, мы окончательно приведем нашу эрмитовскую матрицу к диагональной форме, причем то общее унитарное преобразование U, которое входит в форму (184), будет произведением тех унитарных преобразований, которые нам придется совершать при указанном выше процессе последовательного приведения матрицы к диагональной форме.

Обратимся теперь к уравнению (185). В [**33**] мы показали, что его решения, соответствующие различным значениям λ, обязательно взаимно ортогональны.

Совершенно так же, как и в [**33**], мы можем показать, что векторы, представляемые столбцами матрицы U, вместе с соответствующими значениями λ дают все решения уравнения (185). Необходимо только при этом иметь в виду одно важное обстоятельство, касающееся кратных корней уравнения (186). Если, например, уравнение (186) имеет корень $\lambda = \lambda_1$ кратности m, то при $\lambda = \lambda_1$ уравнение (185) будет иметь m линейно-независимых решений $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(m)}$. Всякая их линейная комбинация с произвольными коэффициентами будет, очевидно, также решением уравнения (185), т. е. уравнение

$$A\boldsymbol{x} = \lambda_1 \boldsymbol{x}$$

будет иметь совокупность решений, представляемую подпространством, образованным векторами $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(m)}$, т. е. определяемую суммой

$$x = C_1\boldsymbol{x}^{(1)} + \dots + C_m\boldsymbol{x}^{(m)}$$

с произвольными коэффициентами $C_1, \dots, C_m$. Мы можем выбирать в этом подпространстве любым образом ортонормированную систему m векторов, составляющие которых и дадут нам те столбцы матрицы U, которые соответствуют собственному значению $\lambda = \lambda_1$. Здесь мы имеем, следовательно, такой же произвол в выборе матрицы U, какой мы имели и в [**33**] для B. Кроме того, очевидно, мы можем составляющие всякого вектора $\boldsymbol{x}^{(s)}$, который у нас получился в результате решения уравнения (185), умножить на численный множитель, по модулю равный единице, т. е. на численный множитель вида $e^{i\varphi}$ (фазовый множитель). При этом длина вектора останется равной единице и сохранится условие ортогональности этого вектора к остальным векторам, входящим в полную систему решений уравнения (185). Наконец, мы можем в матрице U произвольным образом менять порядок столбцов. Это несущественное преобразование сводится, очевидно, к перемене нумерации ортов в новой координатной системе и оно влечет за собой лишь перестановку чисел λ_k в диагональной форме матрицы. В дальнейшем всегда будем считать, что эти числа идут в возрастающем порядке.

Перейдем теперь к рассмотрению форм Эрмита. Мы будем говорить, что эрмитовской матрице A соответствует *форма Эрмита* вида:

$$(A\boldsymbol{x}, \boldsymbol{x}) = \sum_{i, k=1}^{n} a_{ik}\bar{x}_i x_k, \tag{187}$$

где $x_1, \ldots, x_n$ суть составляющие вектора $\boldsymbol{x}$. Число $(A\boldsymbol{x}, \boldsymbol{x})$ вещественно [**32**].

Положим теперь, что мы совершили над нашим пространством некоторое унитарное преобразование, причем старые составляющие вектора выражаются через новые по формуле $\boldsymbol{x} = U\boldsymbol{x}'$. В новых координатах форма Эрмита (187) будет иметь вид:

$$(AU\boldsymbol{x}', U\boldsymbol{x}').$$

Пользуясь свойством (125_1) унитарных преобразований, мы можем оба вектора, образующих скалярное произведение, помножить слева на унитарную матрицу U^{-1} и, таким образом, получим в новых координатах следующее выражение для формы Эрмита (187):

$$(U^{-1}AU\boldsymbol{x}', \boldsymbol{x}'). \tag{188}$$

В частности, если унитарное преобразование U преобразует матрицу A к диагональной форме, т. е. имеет место (184), то в новых переменных в нашей форме Эрмита останутся лишь члены, содержащие произведения $\bar{x}'_i \cdot x'_i$, и мы будем иметь приведение формы Эрмита к сумме квадратов модулей:

$$(U^{-1}AU\boldsymbol{x}', \boldsymbol{x}') = \lambda_1\bar{x}'_1x'_1 + \lambda_2\bar{x}'_2x'_2 + \ldots + \lambda_n\bar{x}'_nx'_n.$$

Таким образом, здесь, как и в [**32**], задача преобразования матрицы A к диагональной форме равносильна задаче приведения соответствующей формы Эрмита к сумме квадратов модулей.

Вместо формы Эрмита рассматривают иногда так называемую *билинейную форму*, определяемую следующим образом:

$$(A\boldsymbol{y}, \boldsymbol{x}) = \sum_{i, k=1}^{n} a_{ik}\bar{x}_i y_k.$$

Если применить опять к пространству унитарное преобразование так, что новые составляющие будут выражаться через старые по прежним формулам, то в новых координатах мы получим:

$$(A\boldsymbol{y}, \boldsymbol{x}) = (AU\boldsymbol{y}', U\boldsymbol{x}')$$

или, в силу свойства унитарного преобразования:

$$(U^{-1}AUy', \ x').$$

Наконец, если U приводит A к диагональной форме, то в соответствующих координатах билинейная форма приведется к следующему простейшему виду:

$$\sum_{k=1}^{n} \lambda_k \bar{x}'_k y'_k.$$

Заметим, что всякая диагональная матрица с вещественными элементами есть эрмитовская матрица, а потому и матрица $U^{-1}[\lambda_1, \ldots, \lambda_n]\,U$, где U — любая унитарная матрица, будет также эрмитовской. Выше мы видели, что и, наоборот, всякую эрмитовскую матрицу мы можем написать в таком виде.

Формы Эрмита делятся, как и вещественные квадратичные формы [**35**], по знаку характеристических чисел λ_k. Если, например, все λ_k положительны, то форма Эрмита называется определенно положительной. Характерным ее свойством является то свойство, что ее значения положительны при всяких x_k, и она может обратиться в нуль лишь при $x_1 = \ldots = x_n = 0$. Аналогично определяются знакопостоянные и знакопеременные формы Эрмита. Исследование совершенно аналогично случаю вещественных квадратичных форм и основано на формуле

$$(Ax, \ x) = \lambda_1 \bar{x}'_1 x'_1 + \ldots + \lambda_n \bar{x}'_n x'_n.$$

Формула (183) справедлива для эрмитовских матриц. Если A — любая матрица и $\tilde{A} = \bar{A}^{(*)}$ — сопряженная с ней матрица, то вместо (183) будем иметь:

$$(Ax, \ y) = (x, \ \tilde{A}y). \qquad (183_1)$$

Если a_{ik} — элементы матрицы A, то у матрицы $\tilde{A}$ элементы будут $\{\tilde{A}\}_{ik} = \overline{a_{ki}}$, и формула (183_1) проверяется непосредственной подстановкой, как и формула (183).

41. Коммутирующие эрмитовские матрицы. Пусть A и B — две эрмитовские матрицы. Посмотрим, при каких условиях и их произведение BA будет эрмитовской матрицей. Составим матрицу, эрмитовски сопряженную с произведением BA:

$$(\overline{BA})^{(*)} = \bar{A}^{(*)}\bar{B}^{(*)}$$

или, в силу эрмитовского характера A и B:

$$(\overline{BA})^{(*)} = AB.$$

Для того чтобы BA было эрмитовской матрицей, необходимо и достаточно, чтобы AB совпадало с BA, т. е. чтобы *матрицы коммутировали.* Положим, что эрмитовские матрицы A и B приводятся к диагональной форме при помощи одного унитарного преобразования U:

$$A = U^{-1}[\lambda_1, \ldots, \lambda_n]U; \quad B = U^{-1}[\mu_1, \ldots, \mu_n]U.$$

Нетрудно видеть, что в этом случае они будут коммутировать

$$AB = BA = U^{-1}[\lambda_1\mu_1, \ldots, \lambda_n\mu_n]U.$$

Покажем теперь, что и наоборот: *если две эрмитовские матрицы коммутируют, то их можно при помощи одного и того же унитарного преобразования одновременно привести к диагональной форме,* т. е. переместительность эрмитовских матриц является не только необходимым, но и достаточным условием возможности их одновременного приведения при помощи унитарного преобразования к диагональной форме. Итак, положим, что $AB = BA$. Заметим при этом, что и подобные им матрицы будут также коммутировать. Действительно:

$$(C^{-1}AC)(C^{-1}BC) = C^{-1}ABC = C^{-1}BAC,$$

и точно такое же выражение получится для произведения

$$(C^{-1}BC)(C^{-1}AC).$$

Положим, что за C мы выбрали унитарное преобразование, приводящее A к диагональной форме, и подвергли такому же преобразованию и матрицу B. Новые матрицы будут также коммутировать, и мы можем, таким образом, при доказательстве нашего предложения считать просто, что матрица A уже имеет диагональную форму, т. е. что элементы a_{ik} удовлетворяют условию

$$a_{ik} = 0 \text{ при } i \neq k. \tag{189}$$

Обозначим через b_{ik} элементы матрицы B и запишем условие, что наши матрицы коммутируют:

$$\sum_{s=1}^{n} a_{is}b_{sk} = \sum_{s=1}^{n} b_{is}a_{sk} \quad (i,\ k = 1,\ 2,\ \ldots,\ n).$$

В силу (189) условия эти будут иметь вид:

$$(a_{ii} - a_{kk})\,b_{ik} = 0 \quad (i,\ k = 1,\ 2,\ \ldots,\ n). \tag{190}$$

Если все числа a_{ii} различны, то из последних равенств непосредственно следует, что $b_{ik} = 0$ при $i \neq k$, т. е. что матрица B также имеет диагональную форму, и предложение доказано.

Перейдем теперь к рассмотрению общего случая, когда среди чисел a_{ii} имеются одинаковые. Для определенности предположим, что числа эти распадаются на две группы одинаковых между собой чисел:

$$a_{11} = \ldots = a_{mm}; \; a_{m+1,\, m+1} = \ldots = a_{nn}.$$

Из формулы (190) непосредственно следует, что в этом случае элементы b_{ik} могут быть отличны от нуля только тогда, когда или оба значка i и k больше m или оба они не больше m. Таким образом, в данном случае матрица B имеет квазидиагональную форму:

$$B=[B_1, B_2],$$

где B_1 — некоторая эрмитовская матрица порядка m и B_2 — эрмитовская матрица порядка $(n-m)$. В раскрытой форме мы можем написать матрицу B в виде:

$$\begin{Vmatrix} b_{11} & \dots & b_{mm} & 0 & & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{m1} & \dots & b_{mm} & 0 & & \dots & 0 \\ 0 & \dots & 0 & b_{m+1,\ m+1} & & \dots & b_{m+1,\ n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & b_{n,\ m+1} & & \dots & b_{nn} \end{Vmatrix}.$$

Мы можем не меняя диагональной формы A, подвергать подпространство, образованное первыми m ортами, любому унитарному преобразованию, и то же самое относится к подпространству, образованному последними $(n-m)$ ортами. Выберем эти унитарные преобразования V_1 и V_2 так, чтобы матрицы B_1 и B_2 записались в диагональной форме. В общем мы будем иметь унитарное преобразование всего n-мерного пространства, имеющее квазидиагональную форму

$$[V_1, V_2].$$

В силу сказанного выше, в новых координатах матрица A сохранит диагональную форму, а матрица B примет вид:

$$[V_1, V_2]^{-1}[B_1, B_2][V_1, V_2]=[V_1^{-1}B_1V_1, V_2^{-1}B_2V_2],$$

т. е. будет также иметь диагональную форму, и наше предложение доказано.

Если мы теперь построим для наших коммутирующих матриц уравнения

$$A\boldsymbol{x}=\lambda\boldsymbol{x}; \quad B\boldsymbol{x}=\mu\boldsymbol{x}, \tag{191}$$

то из предыдущего непосредственно следует, что для обоих этих уравнений мы можем построить одну и ту же систему n линейно-независимых решений. Эти решения и будут давать столбцы той унитарной матрицы U, которая приводит обе наши матрицы к диагональной форме. Иначе говоря, мы можем для двух коммутирующих эрмитовских матриц построить одну и ту же полную систему n линейно-независимых собственных векторов. Что же касается собственных значений, т. е. значений параметров λ и μ, то они будут, конечно, вообще говоря, различными. Заметим, что из предыдущего еще не следует, что всякий собственный вектор матрицы A будет и собственным вектором матрицы B. Если все собственные значения A и B различны, так что каждому значению λ_k и μ_k соответствует с точностью до численного множителя только один вектор, то это будет, конечно, так. Но это уже не будет, вообще говоря, иметь места, если среди собственных значений есть одинаковые. Пусть $x^{(k)}$ есть полная система собственных векторов матриц A и B, а λ_k и μ_k — соответствующие собственные значения. Положим, например, что $\lambda_1 = \lambda_2$, но $\mu_1 \neq \mu_2$. При этом векторы $C_1 x^{(1)} + C_2 x^{(2)}$ при любом выборе постоянных C_1 и C_2 будут собственными векторами для A, но уже не будут собственными векторами для B.

Все предыдущие рассуждения легко переносятся и на случай нескольких матриц, а именно: если имеется несколько эрмитовских матриц $A_1, \ldots, A_l$, то для того, чтобы их можно было одновременно привести к диагональной форме при помощи унитарного преобразования, необходимо и достаточно, чтобы они попарно коммутировали, т. е. $A_i A_k = A_k A_i$ при любых i и k от 1 до l.

42. Приведение унитарных матриц к диагональной форме. Унитарные матрицы в отношении приведения к диагональной форме обладают свойством, совершенно аналогичным эрмитовским матрицам, а именно: если V есть некоторая унитарная матрица, то всегда можно найти такую унитарную матрицу U, что матрица

$$U^{-1} V U$$

будет диагональной матрицей. Мы можем записать нашу задачу в следующем виде:

$$VU = U[\lambda_1, \ldots, \lambda_k], \tag{192}$$

где U — искомая унитарная матрица и λ_k — искомые числа.

Как и раньше для эрмитовской матрицы, столбцам матрицы U будут соответствовать некоторые векторы $x^{(k)}$ и эти векторы должны быть решениями уравнения

$$Vx = \lambda x, \tag{193}$$

где λ совпадают с числами λ_k. Отсюда, как и выше, непосредственно следует, что эти числа λ должны быть корнями характеристического уравнения

$$\begin{vmatrix} v_{11} - \lambda & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} - \lambda & \ldots & v_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ v_{n1} & v_{n2} & \ldots & v_{nn} - \lambda \end{vmatrix} = 0, \tag{194}$$

где v_{ik} — элементы матрицы V.

Заметим прежде всего, что если матрицы V_1 и U_1 унитарны, то и матрица $U_1^{-1} V_1 U_1$ будет унитарной. Действительно, из унитарности U_1 следует унитарность U_1^{-1}, и произведение унитарных матриц есть также унитарная матрица.

Возьмем некоторый корень $\lambda = \lambda_1$ уравнения (194) и, подставив его вместо λ в уравнение (193), определим единичный вектор $x^{(1)}$, удовлетворяющий этому уравнению, примем этот вектор за новый орт и присоединим к нему еще $(n-1)$ единичных векторов так, чтобы получить n единичных взаимно ортогональных векторов. Переход от прежних ортов к новым будет равносилен некоторому унитарному преобразованию U_1, и наша унитарная матрица V перейдет в подобную

$$V_1 = U_1^{-1} V U_1.$$

Соответствующее уравнение

$$V_1 x = \lambda x$$

будет иметь, при $\lambda = \lambda_1$, в качестве решения вектор с составляющими $(1, 0, \ldots, 0)$, откуда будет, как и раньше, непосредственно следовать, что элементы первого столбца матрицы V_1 равны все нулю, кроме первого элемента, который равен λ_1. Но поскольку в унитарной матрице сумма квадратов модулей элементов каждого столбца равна единице, мы можем утверждать, что число λ_1 по модулю равно единице. Напомним теперь, что в унитарной матрице V_1 и сумма квадратов модулей элементов каждой строки также равна единице. Но, как только что показано, первый элемент первой строки λ_1 равен по модулю единице, и, следовательно, остальные элементы этой строки должны равняться нулю. Итак, в результате первого унитарного преобразования мы привели нашу унитарную матрицу к такому виду, что в ее первой строке и первом столбце все элементы равны нулю, кроме первого:

$$\begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & v_{22}^{(1)} & \ldots & v_{2n}^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & v_{n2}^{(1)} & \ldots & v_{nn}^{(1)} \end{Vmatrix}.$$

Совершенно аналогичное обстоятельство мы имели выше для эрмитовских матриц. Далее, элементы $v_{ik}^{(1)}$ образуют унитарную матрицу порядка $(n-1)$. Применяя еще унитарное преобразование, мы сможем и в этой матрице получить нули в первой строке и столбце, кроме первого элемента, который по модулю будет равен единице. В окончательном счете, в результате произведенных двух унитарных преобразований, наша унитарная матрица приведется к виду:

$$\begin{Vmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & v_{33}^{(2)} & \ldots & v_{3n}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & v_{n3}^{(2)} & \ldots & v_{nn}^{(2)} \end{Vmatrix}.$$

Продолжая наше рассуждение так же и дальше, мы приведем нашу унитарную матрицу при помощи некоторого унитарного преобразования к диагональной форме. Отметим, что из предыдущих соображений непосредственно следует, что *все характеристические числа унитарной матрицы будут по модулю равны единице.*

Так же, как и в [41], можно показать, что если некоторые унитарные матрицы попарно коммутируют, то их можно при помощи одного и того же унитарного преобразования привести к диагональной форме.

Отметим еще следующий факт. Пусть унитарная матрица U приводит некоторую матрицу A к диагональной форме, т. е. пусть $U^{-1}AU$ есть диагональная матрица. Как известно, модуль определителя U равен единице, и мы можем подобрать вещественное число ω так, чтобы определитель унитарной матрицы $e^{i\omega}U$ равнялся единице. При этом унитарная матрица $e^{i\omega}U$ также будет приводить A к диагональной форме, ибо

$$(e^{i\omega}U)^{-1}A(e^{i\omega}U) = e^{i\omega}e^{-i\omega}U^{-1}AU = U^{-1}AU.$$

Таким образом, всегда можно считать, что определитель унитарной матрицы U, приводящей некоторую матрицу к диагональной форме, равен единице.

Пример. Рассмотрим в качестве примера приведение к диагональной форме некоторой вещественной ортогональной матрицы третьего порядка

$$V = \begin{Vmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{Vmatrix}. \tag{195}$$

Будем считать, что определитель этой матрицы равен $(+1)$, так что этой матрице соответствует некоторое движение трехмерного пространства, как целого, вокруг начала. Характеристическое уравнение для матрицы (195) будет по условию иметь свободный член, равный единице, ибо этот свободный член совпадает, очевидно, с определителем матрицы. С другой стороны, мы видели, что все корни нашего характеристического уравнения должны иметь модуль, равный единице. Старший член характеристического уравнения будет $(-\lambda)^3 = -\lambda^3$, и, следовательно, свободный член уравнения — единица — будет равняться просто произведению корней этого уравнения. Поскольку это уравнение имеет вещественные коэффициенты, возможны лишь два случая, а именно: или это уравнение имеет один корень, равный единице, а два других мнимые, сопряженные с модулем, равным единице, т. е. два других корня будут вида $e^{\pm i\varphi}$, или уравнение будет иметь корень единица, и оба других корня будут равняться (-1). Второй случай является частным случаем первого при $\varphi = \pi$.

Собственному значению $\lambda = 1$ соответствует вещественный вектор $x^{(1)}$, который должен являться решением уравнения

$$Vx^{(1)} = x^{(1)}. \tag{196}$$

Иначе говоря, этот вектор не должен меняться при том повороте пространства, который определяется матрицей V. Этот вектор, отвечающий вещественному значению $\lambda = 1$, будет вещественным вектором, и он будет определять, очевидно, ту ось, вокруг которой повернулось пространство (всякий поворот пространства вокруг начала равносилен вращению вокруг некоторой оси, проходящей через начало). Для определения составляющих этого вектора $x^{(1)}$ через элементы V перепишем уравнение (196) в виде:

$$V^{-1}x^{(1)}=x^{(1)}$$

или, поскольку матрица V вещественна и унитарна, мы можем написать:

$$V^{(*)}x^{(1)}=x^{(1)}.$$

Вычитая это из (196), будем иметь:

$$(V-V^{(*)})\,x^{(1)}=0.$$

Напишем это равенство в раскрытом виде, причем составляющие вектора $x^{(1)}$ обозначим через $(u_{11},\ u_{21},\ u_{31})$. Получим систему уравнений:

$$\begin{aligned}
&& (v_{12}-v_{21})\,u_{21} &+(v_{13}-v_{31})\,u_{31}=0\\
(v_{21}-v_{12})\,u_{11} && &+(v_{23}-v_{32})\,u_{31}=0\\
(v_{31}-v_{13})\,u_{11} &+& (v_{32}-v_{23})\,u_{21} & \qquad\qquad\qquad=0
\end{aligned}$$

и из нее непосредственно следуют формулы, определяющие направление оси вращения:

$$u_{11}:u_{21}:u_{31}=(v_{23}-v_{32}):(v_{31}-v_{13}):(v_{12}-v_{21}).$$

Два других собственных вектора $x^{(2)}$ и $x^{(3)}$ должны, очевидно, удовлетворять уравнениям

$$Vx^{(2)}=e^{i\varphi}x^{(2)} \quad\text{и}\quad Vx^{(3)}=e^{-i\varphi}x^{(3)}, \tag{197}$$

и это будут уже векторы с комплексными составляющими. Мы сможем определить φ из того условия, что сумма корней характеристического уравнения равна, очевидно, сумме диагональных членов, т. е. следу матрицы V:

$$1+e^{-i\varphi}+e^{i\varphi}=1+2\cos\varphi=v_{11}+v_{22}+v_{33},$$

причем можно считать, что φ лежит между 0 и π.

Из уравнений (197) следует, что составляющие вектора $x^{(2)}$ и $x^{(3)}$ мы можем считать мнимыми сопряженными, поскольку значения λ в уравнениях (197) суть мнимые сопряженные числа. Составим новую унитарную матрицу:

$$U_0=\begin{Vmatrix}1 & 0 & 0\\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{i}{\sqrt{2}}\\ 0 & \dfrac{1}{\sqrt{2}} & -\dfrac{i}{\sqrt{2}}\end{Vmatrix}. \tag{198}$$

Нетрудно убедиться непосредственно, что элементы столбцов матрицы $W=UU_0$ будут равны составляющим векторов

$$x^{(1)};\qquad \frac{x^{(2)}+x^{(3)}}{\sqrt{2}};\qquad i\frac{x^{(2)}-x^{(3)}}{\sqrt{2}},$$

т. е. будут вещественны. Кроме того, матрица W, как произведение двух унитарных матриц, также будет унитарной матрицей, т. е. W будет ортогональной матрицей. Применим теперь к матрице V преобразование подобия, пользуясь унитарной вещественной матрицей W. Мы получим:

$$W^{-1}VW=U_0^{-1}U^{-1}VUU_0=U_0^{-1}\,[1,\ e^{i\varphi},\ e^{-i\varphi}]\,U_0.$$

Производя фактически перемножение матриц, получим:

$$W^{-1}VW=\begin{Vmatrix}1 & 0 & 0\\ 0 & \cos\varphi & -\sin\varphi\\ 0 & \sin\varphi & \cos\varphi\end{Vmatrix}. \qquad (199)$$

Мы можем всегда считать, что определитель ортогональной матрицы W равен $(+1)$, ибо в противном случае мы могли бы умножить эту матрицу на (-1), от чего соотношение (199) не изменилось бы. Таким образом, матрице W будет соответствовать также некоторое движение трехмерного пространства. Матрица (199), полученная в результате преобразования координат $\boldsymbol{x}=W\boldsymbol{x}'$, подобна матрице V и дает в новых координатах то же преобразование, которое первоначальная матрица V давала в прежних координатах. Из вида матрицы (199) непосредственно следует, что этой матрице (199) соответствует вращение вокруг новой оси $\boldsymbol{x}'^{(1)}$ на угол φ, и сущность нашего преобразования сводится к тому, что мы приняли за ось $\boldsymbol{x}'^{(1)}$, упомянутую выше, ось вращения, изображаемую вектором $\boldsymbol{x}^{(1)}$.

Из предыдущего непосредственно вытекает еще одно важное обстоятельство, а именно: все вещественные матрицы, которым соответствует поворот пространства на некоторый определенный угол φ, могут быть приведены при помощи преобразования подобия (различного для различных матриц) к одному и тому же виду (199), и, следовательно, все такие матрицы будут подобны между собой.

Матрицы, соответствующие различным углам вращения, не могут быть между собой подобны, так как характеристические числа таких матриц 1, $e^{i\varphi}$ и $e^{-i\varphi}$ будут наверно различны между собой при различных значениях угла φ. Все эти свойства имеют очень простое геометрическое значение.

43. Матрицы проектирования. Мы переходим теперь к рассмотрению некоторого частного случая эрмитовских матриц. Пусть R_m — некоторое подпространство измерения m, образованное m линейно-независимыми векторами $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}$. Это подпространство R_m представляет собою совокупность векторов вида:

$$C_1\boldsymbol{y}^{(1)}+\ldots+C_m\boldsymbol{y}^{(m)},$$

где C_k — произвольные численные коэффициенты. Ортогонализируя векторы $\boldsymbol{y}^{(k)}$, мы можем построить ортонормированную систему векторов, $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}$, которая образует то же подпространство R_m. Мы можем далее дополнить ее до полной ортонормированной системы n векторов, построив еще ортонормированную систему $\boldsymbol{x}^{(m+1)}, \ldots, \boldsymbol{x}^{(n)}$.

Эти последние векторы образуют некоторое подпространство R'_{n-m} измерения $(n-m)$, причем два подпространства R_m и R'_{n-m} взаимно ортогональны в том смысле, что любой вектор подпространства R_m ортогонален любому вектору подпространства R'_{n-m} [14]. Разлагая произвольный вектор $\boldsymbol{x}$ по ортам $\boldsymbol{x}^{(k)}$:

$$\boldsymbol{x}=x_1\boldsymbol{x}^{(1)}+\ldots+x_n\boldsymbol{x}^{(n)}, \qquad (200)$$

мы можем представить его в виде суммы двух векторов:

$$\boldsymbol{x}=[x_1\boldsymbol{x}^{(1)}+\ldots+x_m\boldsymbol{x}^{(m)}]+[x_{m+1}\boldsymbol{x}^{(m+1)}+\ldots+x_n\boldsymbol{x}^{(n)}]=\boldsymbol{u}+\boldsymbol{v}, \qquad (201)$$

из которых один принадлежит R_m, а второй R'_{n-m}. Нетрудно видеть, что такое разложение любого вектора $\boldsymbol{x}$ на два составляющих единственно. Действительно, положим, что, кроме разложения (201), мы имеем второе разложение $\boldsymbol{x}=\boldsymbol{u}'+\boldsymbol{v}'$ указанного выше свойства. Отсюда $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{u}'+\boldsymbol{v}'$ или $\boldsymbol{u}-\boldsymbol{u}'=\boldsymbol{v}'-\boldsymbol{v}$.

Вектор, стоящий слева, принадлежит R_m, а справа R'_{n-m}, и, следовательно, $\boldsymbol{u}-\boldsymbol{u}'$ и $\boldsymbol{v}-\boldsymbol{v}'$ должны быть ортогональны.

Но каждый вектор, ортогональный сам себе, равен, очевидно, нулю [**14**] и, следовательно, $\boldsymbol{u}-\boldsymbol{u}'=0$, т. е. $\boldsymbol{u}$ совпадает с $\boldsymbol{u}'$, а $\boldsymbol{v}$ — с $\boldsymbol{v}'$, т. е. векторы $\boldsymbol{u}$ и $\boldsymbol{v}$ определяются по вектору $\boldsymbol{x}$ единственным образом. Вектор $\boldsymbol{u}$ *называется проекцией* вектора $\boldsymbol{x}$ в *подпространство* R_m. Та матрица, которая осуществляет переход от вектора $\boldsymbol{x}$ к вектору $\boldsymbol{u}$, называется *матрицей проектирования* в подпространство R_m, и мы ее обозначим через P_{R_m}. Вид этой матрицы зависит, конечно, от выбора координатных осей.

Если за основные орты мы выберем $\boldsymbol{x}^{(k)}$, то вектор $\boldsymbol{x}$ представляется формулой (201), а вектор $\boldsymbol{u}$ — формулой

$$\boldsymbol{u}=x_1\boldsymbol{x}^{(1)}+\ldots+x_m\boldsymbol{x}^{(m)},$$

и в данном случае операция проектирования сводится просто к тому, что первые m составляющих остаются прежними, а остальные делаются равными нулю. Соответствующая матрица проектирования будет, очевидно, диагональной матрицей вида: $P_{R_m}=[1, 1, \ldots, 1, 0, 0, \ldots, 0]$, где первые m мест заняты единицами, а остальные нулями. Если бы мы иначе пронумеровали орты, то получили бы только другой порядок элементов, но по-прежнему имели бы диагональную матрицу, состоящую из единиц и нулей. В общем случае при любом выборе декартовых осей матрица проектирования имеет вид:

$$P_{R_m}=U^{-1}[1, \ldots, 1, 0, \ldots, 0]\,U, \tag{202}$$

где U есть некоторая унитарная матрица, и собственные значения P_{R_m} равны или нулю, или единице. Наоборот, всякая эрмитовская матрица указанного вида есть матрица проектирования в некоторое подпространство, число измерений которого равно числу собственных значений P_{R_m}, равных единице.

Можно определить матрицу проектирования и иным образом, а именно: *матрица проектирования есть такая эрмитовская матрица, которая удовлетворяет соотношению*

$$P^2=P. \tag{203}$$

Действительно, принимая во внимание, что $1^2=1$ и $0^2=0$, нетрудно проверить, что матрицы вида (202) удовлетворяют соотношению (203). Наоборот, если некоторая эрмитовская матрица удовлетворяет соотношению (203), и мы представим ее в виде: $P=U^{-1}[\lambda_1, \ldots, \lambda_n]\,U$, то в силу (203): $U^{-1}[\lambda_1^2, \ldots, \lambda_n^2]\,U=U^{-1}[\lambda_1, \ldots, \lambda_n]\,U$, т. е. $\lambda_k^2=\lambda_k$ $(k=1, 2, \ldots, n)$, откуда непосредственно следует, что $\lambda_k=1$ или 0. Если все характеристические числа матрицы равны единице, то эта матрица есть единичная матрица, и ей соответствует тождественное преобразование, т. е., иначе говоря, проектирование вектора во все пространство (вектор остается неизменным). Исключая этот тривиальный случай, мы будем иметь у матрицы проектирования хоть одно характеристическое число равным нулю и, следовательно, определитель этой матрицы, равный произведению характеристических чисел, будет также равен нулю, и мы не можем, конечно, говорить об обратной матрице P^{-1}. Заметим еще, что непосредственно из определения следует, что матрица проектирования P_{R_m} не меняет вектора, если он принадлежит подпространству R_m, и уменьшает длину вектора, если он не принадлежит R_m.

После этих предварительных сведений перейдем к рассмотрению некоторых действий с матрицами проектирования. Пусть имеются две матрицы проектирования P_R и P_S, такие, что их произведение равно нулю, т. е. матрице, все элементы которой равны нулю:

$$P_S P_R = 0. \tag{204}$$

Возьмем некоторый вектор $\boldsymbol{x}$ из подпространства R, так что $P_R \boldsymbol{x} = \boldsymbol{x}$. Формула (204) даст нам:

$$P_S \boldsymbol{x} = 0.$$

Но отсюда непосредственно следует, что $\boldsymbol{x}$ ортогонален к любому вектору из подпространства S. Действительно, в противном случае мы могли бы найти в подпространстве S единичный вектор $\boldsymbol{y}$, не ортогональный к $\boldsymbol{x}$, и, принимая его за первый орт, мы имели бы для первой составляющей вектора $\boldsymbol{x}$ величину, отличную от нуля и при проектировании $\boldsymbol{x}$ в S эта составляющая осталась бы неизменной. Таким образом мы видим, что при выполнении условия (204) всякий вектор из R ортогонален всякому вектору из S, а следовательно, и наоборот. Но тогда наряду с (204) мы имеем и

$$P_R P_S = 0. \tag{205}$$

Действительно, для любого вектора $\boldsymbol{y}$, вектор $P_S \boldsymbol{y}$ принадлежит S, а потому ортогонален ко всем векторам из R, т. е. для любого вектора $\boldsymbol{y}$ мы имеем: $P_R P_S \boldsymbol{y} = 0$, что и равносильно (205). Наоборот, если два подпространства R и S взаимно ортогональны в указанном выше смысле, то имеют место (204) и (205).

Рассмотрим теперь сумму двух матриц проектирования:

$$P = P_R + P_S \tag{206}$$

и положим, что выполнены условия (204) и (205). Покажем, что матрица (206), которая является, очевидно, эрмитовской матрицей, будет также матрицей проектирования. Для этого убедимся, что квадрат ее совпадает с ней самой:

$$P^2 = (P_R + P_S)(P_R + P_S) = P_R^2 + P_R P_S + P_S P_R + P_S^2,$$

откуда, в силу сделанного условия и того, что P_R и P_S суть матрицы проектирования, имеем: $P^2 = P_R + P_S = P$.

Нетрудно показать, что в рассматриваемом случае матрице P соответствует операция проектирования в подпространство $(R+S)$, которое образуется соединением подпространств R и S в том смысле, что подпространство $(R+S)$ образовано совокупностью всех тех векторов, которые служили для образования подпространств R и S, т. е. если векторы $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(p)}$

образовывали R и $y^{(1)}, \ldots, y^{(q)}$ образовывали S, то подпространство $(R+S)$ будет представляться совокупностью векторов:

$$C_1 x^{(1)} + \ldots + C_p x^{(p)} + D_1 y^{(1)} + \ldots + D_q y^{(q)},$$

где C_k и D_k — произвольные постоянные. Предыдущее свойство обобщается и на любое число слагаемых:

$$P = P_{S_1} + \ldots + P_{S_m}. \tag{207}$$

Если подпространства S_k попарно взаимно ортогональны, т. е. если любой вектор из S_i ортогонален к любому вектору из S_j при различных i и j, то сумма (207) представляет собою матрицу проектирования в подпространство $(S_1 + \ldots + S_m)$, образованное всеми теми векторами, которые служили для образования подпространств S_k. В частном случае эта сумма может быть равной единичной матрице $I = P_{S_1} + \ldots + P_{S_m}$, и в этом случае обычно говорят о разложении единицы на матрицы проектирования или просто о разложении единицы.

Рассмотрим еще произведение двух матриц проектирования

$$P = P_S P_R. \tag{208}$$

Для того чтобы это произведение также было матрицей проектирования, необходимо прежде всего, чтобы это произведение было эрмитовской матрицей, а для этого, как известно [**41**], необходимо, чтобы наши матрицы коммутировали

$$P_R P_S = P_S P_R. \tag{209}$$

Покажем, что это условие достаточно, т. е. что в данном случае квадрат матрицы P^2 совпадает с самой матрицей P:

$$P^2 = P_S P_R P_S P_R$$

или, переставляя матрицы в силу условия (209):

$$P^2 = P_S^2 P_R^2 = P_S P_R,$$

что и требовалось доказать. Нетрудно проверить, что при условии коммутирования (209) матрице (208) соответствует проектирование в подпространство, образованное векторами, общими тем двум совокупностям векторов, которые образуют R и S.

Отметим еще один результат, не останавливаясь на его доказательстве, которое не представляет никакого труда, а именно: если подпространство S составляет часть подпространства R, то разность

$$P = P_R - P_S \tag{210}$$

есть также матрица проектирования. Если $x^{(k)}$ суть основные векторы, образующие S, то для получения основных векторов, образующих R, мы должны добавить к вышеуказанным векторам еще один или несколько линейно независимых векторов. Эти последние векторы сами по себе образуют некоторое подпространство T, и матрица (210) в рассматриваемом случае и будет матрицей проектирования в это подпространство.

Пользуясь матрицами проектирования, можно формулировать задачу приведения эрмитовской матрицы к диагональной форме вполне однозначным образом и при наличии кратных собственных значений.

Положим, например, что мы имеем эрмитовскую матрицу

$$A = U[\lambda_1, \ldots, \lambda_n]U^{-1},$$

где U — некоторая унитарная матрица. Положим для определенности, что числа λ_k распадаются на две части, одинаковые между собой, причем первые m чисел равны μ, а остальные $(n-m)$ равны ν:

$$A = U[\mu, \ldots, \mu, \nu, \ldots, \nu]U^{-1}.$$

Мы можем, очевидно, переписать нашу матрицу в виде:

$$A = \mu U[1, \ldots, 1, 0, \ldots, 0]U^{-1} + \nu U[0, \ldots, 0, 1, \ldots, 1]U^{-1}.$$

Введем в рассмотрение матрицы проектирования

$$P_R = U[1, \ldots, 1, 0, \ldots, 0]U^{-1}; \quad P_S = U[0, \ldots, 0, 1, \ldots, 1]U^{-1}.$$

Соответствующие подпространства R и S, очевидно, взаимно ортогональны, и наши матрицы проектирования в сумме дают единичную матрицу. Мы имеем таким образом в данном случае $A = \mu P_R + \nu P_S$, причем $\lambda_1 = \ldots = \lambda_m = \mu$ и $\lambda_{m+1} = \ldots = \lambda_n = \nu$.

В общем случае задача приведения эрмитовской матрицы к диагональной форме сводится к такому разложению единичной матрицы

$$I = P_{S_1} + \ldots + P_{S_m}, \tag{211}$$

чтобы наша матрица A могла быть представлена в виде:

$$A = \mu_1 P_{S_1} + \ldots + \mu_m P_{S_m}, \tag{212}$$

где μ_k — различные собственные значения нашей матрицы. Таким образом всякой эрмитовской матрице соответствует определенное разложение единицы (211) такое, что эта матрица представляется в виде (212).

Нетрудно перевести все предыдущие результаты на язык не матриц, а форм Эрмита. Всякой матрице проектирования P_R с элементами p_{ik} соответствует некоторая форма Эрмита

$$P_R(\boldsymbol{x}) = (P_R\boldsymbol{x}, \boldsymbol{x}) = \sum_{i,\,k=1}^{n} p_{ik}\bar{x}_i x_k, \tag{213}$$

которая называется иногда *особой* формой (Einzelform). Символ $P_R(\boldsymbol{x})$ — краткая запись $(P_R\boldsymbol{x}, \boldsymbol{x})$.

Если соответствующее подпространство R имеет l измерений и мы выберем за первые l ортов l единичных взаимно ортогональных векторов подпространства R, то в такой координатной системе наша форма (213) будет иметь вид:

$$(P_R\boldsymbol{x}', \boldsymbol{x}') = \bar{x}'_1 x'_1 + \bar{x}'_2 x'_2 + \ldots + \bar{x}'_l x'_l.$$

Заметим далее, что если матрицы P_{S_k} являются разложением единицы, согласно (211), то, выбирая за орты единичные взаимно ортогональные векторы в каждом из подпространств S_k, мы будем, очевидно, иметь:

$$\sum_{k=1}^{m} P_{S_k}(\boldsymbol{x}') = \sum_{i=1}^{n} x'_i x'_i,$$

и, следовательно, при всяком выборе координатных осей сумма

$$\sum_{k=1}^{m} P_{S_k}(x)$$

выражает квадрат длины вектора. Мы можем, таким образом, сказать, что *задача приведения формы Эрмита A к сумме квадратов равносильна следующим двум равенствам:*

$$(Ax, x) = \sum_{k=1}^{m} \mu_k P_{S_k}(x), \tag{214}$$

$$|x|^2 = \sum_{k=1}^{m} P_{S_k}(x). \tag{215}$$

Введение матриц проектирования позволяет, таким образом, формулировать задачу приведения эрмитовской матрицы к диагональной форме без всякого специального выбора координатных осей. Это дает в свою очередь возможность перенести предыдущие результаты, с соответственными изменениями, и на случай пространства с бесчисленным множеством измерений, что и является основной задачей математического аппарата современной квантовой механики. Мы будем говорить об этом лишь значительно позже. Это распространение на случай бесчисленного множества измерений выводит нас из рамок алгебры и существенным образом связано с введением аппарата анализа.

44. Функции от матриц. Матрица может играть роль аргумента некоторой функции. Мы ограничимся здесь рассмотрением наиболее элементарных функций, а именно *полинома от матрицы и рациональной дроби.* Более подробное рассмотрение теории функций матриц мы сделаем впоследствии, после изложения теории функций комплексного переменного. Полином $f(A)$ степени m от переменной матрицы A имеет вид:

$$f(A) = c_0 + c_1 A + \ldots + c_m A^m, \tag{216}$$

где c_k — некоторые численные коэффициенты. Значение функции в данном случае есть тоже некоторая матрица, элементы которой выражаются, очевидно, по формулам:

$$\{f(A)\}_{ik} = c_0 \delta_{ik} + c_1 \{A\}_{ik} + \ldots + c_m \{A^m\}_{ik},$$

где

$$\delta_{ik} = 0 \text{ при } i \neq k \text{ и } \delta_{ii} = 1.$$

Можно рассматривать полином и от нескольких матриц, но при этом надо иметь в виду некоммутативность этих матриц при умножении. Общий вид полинома второй степени от двух переменных матриц A и B будет:

$$f(A, B) = c_0 + c_1 A + c_2 B + c_3 A^2 + c_4 B^2 + c_5 AB + c_6 BA.$$

Заменим в формуле (216) матрицу A некоторой подобной ей матрицей $U^{-1}AU$. Принимая во внимание, что $(U^{-1}AU)^k = U^{-1}A^kU$, будем иметь:

$$f(U^{-1}AU) = c_0 + c_1U^{-1}AU + \ldots + c_mU^{-1}A^mU = \\ = U^{-1}(c_0 + c_1A + \ldots + c_mA^m)U,$$

т. е.

$$f(U^{-1}AU) = U^{-1}f(A)U. \qquad (217)$$

Аналогичная формула будет иметь место и для полинома от нескольких матриц

$$f(U^{-1}AU,\ U^{-1}BU) = U^{-1}f(A,\ B)U. \qquad (218)$$

Остановимся теперь несколько подробнее на случае эрмитовских матриц. Если A есть эрмитовская матрица, то непосредственно из определения следует, что любая целая положительная степень A^k, а также произведение cA, где c — вещественная постоянная, суть также эрмитовские матрицы. Кроме того, сумма эрмитовских матриц есть также эрмитовская матрица. Отсюда непосредственно следует, что если в формуле (216) A есть эрмитовская матрица и коэффициенты c_k — вещественные числа, то и значение функции $f(A)$ будет эрмитовской матрицей. Эта эрмитовская матрица $f(A)$, очевидно, коммутирует с A, и их можно одновременно привести к диагональной форме при помощи некоторого унитарного преобразования. Заметим прежде всего, что если мы подставим в функцию (216) вместо A некоторую диагональную матрицу $[\lambda_1, \ldots, \lambda_n]$, то в результате получим, очевидно, также диагональную матрицу

$$\sum_{k=0}^{m} c_k[\lambda_1^k, \ldots, \lambda_n^k] = [f(\lambda_1), \ldots, f(\lambda_n)], \qquad (219)$$

где $f(\lambda_k)$ есть численное значение нашего полинома при подстановке вместо A числа λ_k.

Положим теперь, что V есть унитарное преобразование, преобразующее матрицу A к диагональной форме

$$A = V[\lambda_1, \ldots, \lambda_n]V^{-1}.$$

В силу (217) и (219) будем иметь:

$$f(A) = V[f(\lambda_1), \ldots, f(\lambda_n)]V^{-1},$$

т. е. V преобразует и $f(A)$ к диагональной форме, причем $f(\lambda_k)$ суть характеристические числа этой последней матрицы.

Перейдем теперь к рассмотрению рациональных дробей. Пусть $f_1(A)$ и $f_2(A)$ — два полинома от матрицы A. Рассмотрим их частное

$$\frac{f_1(A)}{f_2(A)}. \qquad (220)$$

Как мы раньше видели, частное двух матриц не имеет, вообще говоря, определенного значения [**26**], но в данном случае, как нетрудно показать, мы получим для частного (220) одно определенное значение, если только определитель матрицы $f_2(A)$ отличен от нуля. Частное (220) можно записать двояко:

$$f_1(A) f_2(A)^{-1} \text{ или } f_2(A)^{-1} f_1(A).$$

Покажем, что эти два произведения равны между собой:

$$f_1(A) f_2(A)^{-1} = f_2(A)^{-1} f_1(A),$$

или, что равносильно:

$$f_2(A) f_1(A) = f_1(A) f_2(A). \tag{221}$$

Так как наши полиномы содержат только одну матрицу A, то они коммутируют, т. е. (221) действительно имеет место, и наше частное (220) имеет определенное значение. Нетрудно проверить дальше, что рациональные дроби в случае одной матрицы перемножаются, как и обычные дроби. Действительно:

$$\frac{f_1(A)}{f_2(A)} \frac{f_3(A)}{f_4(A)} = f_1(A) f_2(A)^{-1} f_3(A) f_4(A)^{-1}$$

или, принимая во внимание коммутативность:

$$\frac{f_1(A)}{f_2(A)} \frac{f_3(A)}{f_4(A)} = f_1(A) f_3(A) \cdot [f_2(A) f_4(A)]^{-1} = \frac{f_1(A) f_3(A)}{f_2(A) f_4(A)}.$$

В качестве примера рассмотрим рациональную дробь вида:

$$U = \frac{1 + iA}{1 - iA}, \tag{222}$$

где A — некоторая эрмитовская матрица, т. е. $\bar{A}^{(*)} = A$. Легко показать, что матрица U будет унитарной, т. е. что

$$\overline{U}^{(*)} = U^{-1}. \tag{223}$$

Действительно, мы имеем:

$$\overline{U} = \frac{1 - i\bar{A}}{1 + i\bar{A}} = (1 - i\bar{A})(1 + i\bar{A})^{-1},$$

откуда, переходя к транспонированной матрице, получим [**26**]:

$$\overline{U}^{(*)} = (1 + i\bar{A})^{(*)-1} (1 - i\bar{A})^{(*)} = (1 + i\bar{A}^{(*)})^{-1} (1 - i\bar{A}^{(*)}),$$

или, в силу того, что $\bar{A}^{(*)} = A$:

$$\overline{U}^{(*)} = (1 + iA)^{-1} (1 - iA) = \frac{1 - iA}{1 + iA} = U^{-1},$$

т. е. (223) выполнено, и U есть действительная унитарная матрица.

Формулу (222) мы можем записать в виде:

$$U(1 - iA) = (1 + iA),$$

причем, в силу (222), U коммутирует с A, и значит

$$A = i\,\frac{U-1}{U+1}. \tag{224}$$

Совершенно так же, как и выше, можно показать, что если U есть унитарная матрица и определитель матрицы $U+1$ отличен от нуля, то A, определяемая формулой (224), будет эрмитовской матрицей. Таким образом, всякую унитарную матрицу, для которой $D(U+1) \neq 0$, можно представить через эрмитовскую матрицу A по формуле (222).

45. Пространство с бесчисленным множеством измерений. Мы переходим теперь к введению понятия о пространстве с бесчисленным множеством измерений. Предварительно нам надо ввести понятие о пределе комплексного переменного. Положим, что комплексное переменное $z = x + yi$ принимает последовательно значения:

$$z_1 = x_1 + y_1 i; \quad z_2 = x_2 + y_2 i; \; \ldots; \quad z_n = x_n + y_n i; \; \ldots \tag{225}$$

Говорят, что комплексное число $\alpha = a + bi$ есть предел последовательности (225), если модуль разности $(\alpha - z_n)$ стремится к нулю при беспредельном возрастании n, т. е. $|\alpha - z_n| \to 0$ при $n \to \infty$, и пишут $\alpha = \lim z_n$ или $z_n \to \alpha$. Но

$$|\alpha - z_n| = |(a - x_n) + (b - y_n)\,i| = \sqrt{(a - x_n)^2 + (b - y_n)^2}.$$

Поскольку под радикалом оба слагаемых не отрицательны, условие, $|\alpha - z_n| \to 0$ равносильно двум условиям: $x_n \to a$ и $y_n \to b$. Итак:

$$x_n + y_n i \to a + bi \tag{226}$$

равносильно $x_n \to a$ и $y_n \to b$. Рассмотрим ряд с комплексными членами:

$$\sum_{k=1}^{\infty} (a_k + b_k i). \tag{227}$$

Он называется сходящимся, если сумма его первых n членов

$$S_n = \sum_{k=1}^{n} (a_k + b_k i) = (a_1 + a_2 + \ldots + a_n) + (b_1 + b_2 + \ldots + b_n)\,i$$

стремится к пределу: $S_n \to a + bi$ при беспредельном возрастании n и этот предел $(a + bi)$ называется суммою ряда. Из определения предела следует, что сходимость ряда (227) равносильна сходимости рядов

$$a = \sum_{k=1}^{\infty} a_k \quad \text{и} \quad b = \sum_{k=1}^{\infty} b_k, \tag{228}$$

составленных из вещественных и мнимых частей членов ряда (227).

Положим, что сходится ряд

$$\sum_{k=1}^{\infty} |a_k + ib_k| = \sum_{k=1}^{\infty} \sqrt{a_k^2 + b_k^2}, \tag{229}$$

составленный из модулей членов ряда (227). В силу очевидных неравенств

$$|a_k| \leqslant \sqrt{a_k^2 + b_k^2} \quad \text{и} \quad |b_k| \leqslant \sqrt{a_k^2 + b_k^2}, \tag{230}$$

при этом и ряды (228) будут сходящимися и притом абсолютно сходящимися, а потому и ряд (227) будет также сходящимся, т. е. если сходится ряд (229), то ряд (227) и подавно сходится. В этом случае ряд (227) называется *абсолютно сходящимся.* Применяя обычный признак Коши, мы можем формулировать необходимое и достаточное условие абсолютной сходимости следующим образом: при любом малом положительном ε существует такое N, что

$$\sum_{k=n}^{n+p} |a_k + ib_k| < \varepsilon, \tag{231}$$

если только $n > N$ и p — любое целое положительное число.

Применим теперь сказанное выше к некоторым частным случаям, которые играют существенную роль в дальнейшем. Рассмотрим ряд вида:

$$\sum_{k=1}^{\infty} \alpha_k \beta_k, \tag{232}$$

где α_k и β_k — некоторые комплексные числа, относительно которых известно, что ряды

$$\sum_{k=1}^{\infty} |\alpha_k|^2 \quad \text{и} \quad \sum_{k=1}^{\infty} |\beta_k|^2 \tag{233}$$

сходятся. Применим доказанное в [**29**] неравенство

$$\left\{\sum_{k=n}^{n+p} |\alpha_k \beta_k|\right\}^2 \leqslant \sum_{k=n}^{n+p} |\alpha_k|^2 \sum_{k=n}^{n+p} |\beta_k|^2.$$

Принимая во внимание сходимость рядов (233), мы получаем отсюда, что сумма

$$\sum_{k=n}^{n+p} |\alpha_k \beta_k|$$

будет сколько угодно малой при больших n и любых p, т. е. сходимость рядов (233) обеспечивает абсолютную сходимость ряда (232).

Рассмотрим теперь ряд

$$\sum_{k=1}^{\infty} |\alpha_k + \beta_k|^2 = \sum_{k=1}^{\infty} (\alpha_k + \beta_k)(\bar{\alpha}_k + \bar{\beta}_k), \tag{234}$$

причем по-прежнему будем считать, что ряды (233) сходятся. Ряд (234) можно представить в виде суммы четырех рядов:

$$\sum_{k=1}^{\infty} |\alpha_k|^2; \quad \sum_{k=1}^{\infty} |\beta_k|^2; \quad \sum_{k=1}^{\infty} \alpha_k \bar{\beta}_k; \quad \sum_{k=1}^{\infty} \bar{\alpha}_k \beta_k.$$

Первые два из них сходятся по условию, сходимость же последних двух рядов вытекает из доказанного выше предложения, т. е. сходимость рядов (233) обеспечивает и сходимость ряда (234).

Обратимся теперь к рассмотрению *пространства с бесчисленным множеством измерений*. Мы назовем *вектором в таком пространстве* последовательность бесчисленного множества комплексных чисел

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

причем всегда будем считать, что эти числа подчиняются некоторому условию, а именно ряд

$$\sum_{k=1}^{\infty} |x_k|^2 \tag{235}$$

должен быть сходящимся рядом. Совокупность таких векторов называется обычно *пространством Гильберта*, который впервые изучал такое пространство. В дальнейшем мы будем для краткости называть его пространством l_2.

Для векторов пространства l_2 мы введем, как и выше, основные операции умножения вектора на число и сложение векторов. Если x_k — составляющие x, то составляющие вектора cx, где c — некоторое комплексное число, считаются равными cx_k. Если x_k и y_k — составляющие векторов x и y, то составляющие вектора $(x+y)$ считаются равными (x_k+y_k). Разность $x-y$ есть сумма x и $(-1)y$ (ср. [**12**]). Раз ряд (235) сходится, то, очевидно, ряд $\sum\limits_{k=1}^{\infty} |cx_k|^2$ также сходится. Точно так же, если ряды

$$\sum_{k=1}^{\infty} |x_k|^2 \text{ и } \sum_{k=1}^{\infty} |y_k|^2$$

сходятся, то из вышесказанного вытекает, что и ряд

$$\sum_{k=1}^{\infty} |x_k + y_k|^2$$

также сходится, т. е. последовательности чисел $(cx_1, cx_2, \ldots)$ и $(x_1+y_1, x_2+y_2, \ldots)$ определяют векторы $\boldsymbol{cx}$ и $\boldsymbol{x}+\boldsymbol{y}$, если $\boldsymbol{x}$ и $\boldsymbol{y}$ принадлежат l_2. Нулевой вектор есть вектор, все составляющие которого равны нулю. В векторных равенствах он обычно обозначается числом нуль. Операции над векторами подчиняются обычным правилам (ср. [**12**]):

$$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}; \quad (\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z}),$$

$$(a+b)\,\boldsymbol{x}=a\boldsymbol{x}+b\boldsymbol{x}; \quad a\,(\boldsymbol{x}+\boldsymbol{y})=a\boldsymbol{x}+a\boldsymbol{y}; \quad a\,(b\boldsymbol{x})=(ab)\,\boldsymbol{x}.$$

Точно так же, в силу сказанного выше, мы можем для векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ определить скалярное произведение

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{\infty} x_k \bar{y}_k,$$

откуда следуют формулы [**13**]

$$(\boldsymbol{y}, \boldsymbol{x}) = \overline{(\boldsymbol{x}, \boldsymbol{y})}; \quad (a\boldsymbol{x}, \boldsymbol{y}) = a\,(\boldsymbol{x}, \boldsymbol{y}); \quad (\boldsymbol{x}, a\boldsymbol{y}) = \bar{a}\,(\boldsymbol{x}, \boldsymbol{y}),$$

$$(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z}); \quad (\boldsymbol{z}, \boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{z}, \boldsymbol{x}) + (\boldsymbol{z}, \boldsymbol{y}).$$

Сумма

$$(\boldsymbol{x}, \boldsymbol{x}) = \sum_{k=1}^{\infty} |x_k|^2 \tag{236}$$

определяет квадрат *нормы* (длины) вектора $\boldsymbol{x}$. Введем следующее обозначение нормы:

$$\sum_{k=1}^{\infty} |x_k|^2 = \|\boldsymbol{x}\|^2, \tag{237}$$

т. е. $\|\boldsymbol{x}\|$ обозначает норму вектора $\boldsymbol{x}$.

Для скалярного произведения имеет место неравенство [**30**]

$$|(\boldsymbol{x}, \boldsymbol{y})| \leqslant \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| \tag{238}$$

и совершенно так же, как в [**30**], выводится правило треугольника

$$\|\boldsymbol{x}+\boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|. \tag{239}$$

Норма положительна для всякого вектора, кроме нулевого вектора, у которого она равна нулю. Два вектора $\boldsymbol{u}$ и $\boldsymbol{v}$ называются *взаимно ортогональными* или просто *ортогональными*, если их скалярное произведение равно нулю, т. е. $(\boldsymbol{u}, \boldsymbol{v}) = 0$ и $(\boldsymbol{v}, \boldsymbol{u}) = 0$, причем одно из этих равенств есть следствие другого.

Если векторы $x^{(k)}$ $(k=1, 2, \ldots, m)$ попарно ортогональны, т. е. если $(x^{(i)}, x^{(k)})=0$ при $i \neq k$, то имеем, очевидно:

$$(x^{(1)}+\ldots+x^{(m)},\ x^{(1)}+\ldots+x^{(m)})=(x^{(1)},\ x^{(1)})+\ldots+(x^{(m)},\ x^{(m)})$$

или, что то же самое:

$$\| x^{(1)}+\ldots+x^{(m)} \|^2=\| x^{(1)} \|^2+\ldots+\| x^{(m)}\|^2, \tag{240}$$

т. е. *квадрат нормы суммы попарно ортогональных векторов равен сумме квадратов норм слагаемых*. Естественно назвать это предложение *теоремой Пифагора*. Из определения нормы непосредственно следует, что если c — комплексное число, то для нормы вектора cx имеем

$$\| cx \| = | c | \cdot \| x \|.$$

Говоряг, что векторы

$$x^{(1)},\ x^{(2)},\ \ldots,\ x^{(m)} \tag{241}$$

образуют ортонормированную систему, если эти векторы попарно ортогональны и норма каждого равна единице, т. е.

$$(x^{(i)},\ x^{(k)})=\delta_{ik},$$

где $\delta_{ik}=0$ при $i \neq k$ и $\delta_{ii}=1$. Отметим, что при этом формула (240) дает

$$\| \alpha_1 x^{(1)}+\ldots+\alpha_m x^{(m)} \|^2=| \alpha_1 |^2+\ldots+| \alpha_m |^2,$$

где α_k — любые комплексные числа. Ортонормированные системы могут состоять и из бесчисленного множества векторов. В качестве примера приведем основные орты пространства l_2:

$$a^{(1)}(1,\ 0,\ 0,\ 0,\ \ldots); \quad a^{(2)}(0,\ 1,\ 0,\ 0,\ \ldots); \quad a^{(3)}(0,\ 0,\ 1,\ 0,\ \ldots); \quad \ldots$$

Для составляющих любого вектора $y(y_1,\ y_2,\ \ldots)$ имеем:

$$y_k=(y,\ a^{(k)}).$$

Вернемся к конечной ортонормированной системе (241). Скалярное произведение $(y,\ x^{(k)})$ называют часто *коэффициентом Фурье* вектора y относительно ортонормированной системы (241) или *величиной проекции y на ось $x^{(k)}$*. Сумма

$$\sum_{k=1}^{m} (y,\ x^{(k)})\, x^{(k)}$$

вообще говоря, отлична от y. Представим y в виде

$$v=\sum_{k=1}^{m} (y,\ x^{(k)})\, x^{(k)}+u. \tag{242}$$

Умножая обе части этого равенства скалярно на $x^{(l)}$ и принимая во внимание ортонормированность системы (241), получим

$$(y,\ x^{(l)}) = (y,\ x^{(l)}) + (u,\ x^{(l)}),$$

т. е. $(u,\ x^{(l)}) = 0$ $(l = 1,\ 2,\ \ldots,\ m)$ или, иначе говоря, вектор u ортогонален ко всем векторам $x^{(k)}$ $(k = 1,\ 2,\ \ldots,\ m)$. Мы можем таким образом применить к правой части равенства (242) теорему Пифагора

$$\|y\|^2 = \sum_{k=1}^{m} |(y,\ x^{(k)})|^2 + \|u\|^2, \tag{243}$$

откуда непосредственно следует неравенство

$$\sum_{k=1}^{m} |(y,\ x^{(k)})|^2 \leqslant \|y\|^2, \tag{244}$$

которое называется *неравенством Бесселя*. В этом неравенстве будет иметь место знак равенства в том и только в том случае, когда $u = 0$, т. е. когда u — нулевой вектор. Это непосредственно следует из (243).

Для перенесения последних результатов на случай бесконечных ортонормированных систем нам необходимо ввести понятие предела последовательности векторов и рассматривать бесконечные ряды, члены которых — векторы l_2.

46. Сходимость векторов. Пусть имеется бесконечная последовательность векторов $v^{(k)}$ $(k = 1,\ 2,\ \ldots)$. Будем говорить, что эта последовательность *стремится к вектору* v, или что вектор v есть *предел* этой последовательности, если при $k \to \infty$

$$\|v - v^{(k)}\| \to 0, \text{ т. е. } \|v - v^{(k)}\|^2 \to 0. \tag{245}$$

Обозначая через $v_1^{(k)},\ v_2^{(k)},\ \ldots$ составляющие $v^{(k)}$, а через $v_1,\ v_2,\ \ldots$ — составляющие v, можем написать условие (245) в раскрытом виде:

$$\lim_{k \to \infty} \sum_{k=1}^{\infty} [\,|v_1 - v_1^{(k)}|^2 + |v_2 - v_2^{(k)}|^2 + \ldots] = 0. \tag{246}$$

Раз сумма неотрицательных слагаемых должна стремиться к нулю, то то же можно утверждать и о каждом слагаемом, т. е. из (246) следует

$$\lim_{k \to \infty} |v_m - v_m^{(k)}| = 0 \qquad (m = 1,\ 2,\ \ldots), \tag{247}$$

т. е. каждая составляющая $v_m^{(k)}$ должна стремиться к соответствующей составляющей v_m. Подробнее говоря, вещественная и мнимая части $v_m^{(k)}$ должны стремиться к вещественной и мнимой частям v_m [45].

Заметим, что из (246) следует (247), но обратное не верно, т. е. из (247) не следует (246). В качестве примера положим, что вектор $\boldsymbol{v}^{(k)}$ имеет составляющие $(0, \ldots, 0, 1, 0 \ldots)$, причем единица стоит на месте, номер которого равен k. При беспредельном возрастании k каждая из составляющих станет равной нулю, т. е. при любом целом m мы имеем: $v_m^{(k)} \to 0$, т. е. $v_m = 0$ $(m = 1, 2, \ldots)$, но в то же время сумма (246) все время остается равной единице.

Если последовательность $\boldsymbol{v}^{(k)}$ стремится к $\boldsymbol{v}$, то пишут $\boldsymbol{v}^{(k)} \Rightarrow \boldsymbol{v}$. Рассмотрим еще пример, когда сходимость имеет место. Пусть $\boldsymbol{v}(v_1, v_2, \ldots)$ — некоторый вектор. Определим векторы $\boldsymbol{v}^{(k)}$ следующим образом: вектор $\boldsymbol{v}^{(k)}$ имеет первые k составляющие те же, что и $\boldsymbol{v}$, а остальные его составляющие равны нулю, т. е.:

$$\boldsymbol{v}^{(k)}(v_1, v_2, \ldots, v_k, 0, 0, \ldots).$$

Нетрудно показать, что $\boldsymbol{v}^{(k)} \Rightarrow \boldsymbol{v}$. Действительно, в рассматриваемом случае

$$\| \boldsymbol{v} - \boldsymbol{v}^{(k)} \|^2 = \sum_{n=k+1}^{\infty} | v_n |^2,$$

и, в силу сходимости ряда с общим членом $| v_n |^2$, написанная сумма стремится к нулю при беспредельном возрастании k. Отметим некоторые простые правила, связанные с понятием предела. Если $\boldsymbol{u}^{(k)} \Rightarrow \boldsymbol{u}$ и $\boldsymbol{v}^{(k)} \Rightarrow \boldsymbol{v}$, то

$$\boldsymbol{u}^{(k)} + \boldsymbol{v}^{(k)} \Rightarrow \boldsymbol{u} + \boldsymbol{v} \text{ и } (\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)}) \to (\boldsymbol{u}, \boldsymbol{v}).$$

Отметим, что скалярное произведение есть комплексное число и в последней формуле в связи с этим мы написали $\to$, а не $\Rightarrow$. Можно сказать, что эта формула выражает непрерывность скалярного произведения. Мы имеем, в силу (234):

$$\| (\boldsymbol{u} + \boldsymbol{v}) - (\boldsymbol{u}^{(k)} + \boldsymbol{v}^{(k)}) \| = \| (\boldsymbol{u} - \boldsymbol{u}^{(k)}) + (\boldsymbol{v} - \boldsymbol{v}^{(k)}) \| \leqslant \\ \leqslant \| \boldsymbol{u} - \boldsymbol{u}^{(k)} \| + \| \boldsymbol{v} - \boldsymbol{v}^{(k)} \|,$$

причем, в силу определения предела, $\| \boldsymbol{u} - \boldsymbol{u}^{(k)} \| \to 0$ и $\| \boldsymbol{v} - \boldsymbol{v}^{(k)} \| \to 0$. Из написанного неравенства следует, что и

$$\| (\boldsymbol{u} + \boldsymbol{v}) - (\boldsymbol{u}^{(k)} + \boldsymbol{v}^{(k)}) \| \to 0,$$

т. е. действительно $\boldsymbol{u}^{(k)} + \boldsymbol{v}^{(k)} \Rightarrow \boldsymbol{u} + \boldsymbol{v}$. Далее, в силу определения предела, следует:

$$\boldsymbol{u}^{(k)} = \boldsymbol{u} + \boldsymbol{s}^{(k)}; \quad \boldsymbol{v}^{(k)} = \boldsymbol{v} + \boldsymbol{t}^{(k)},$$

где $\| \boldsymbol{s}^{(k)} \| \to 0$ и $\| \boldsymbol{t}^{(k)} \| \to 0$. Для скалярного произведения имеем:

$$(\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)}) = (\boldsymbol{u} + \boldsymbol{s}^{(k)}, \boldsymbol{v} + \boldsymbol{t}^{(k)}) = \\ = (\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{u}, \boldsymbol{t}^{(k)}) + (\boldsymbol{s}^{(k)}, \boldsymbol{v}) + (\boldsymbol{s}^{(k)}, \boldsymbol{t}^{(k)}),$$

откуда

$$|(u, v)-(u^{(k)}, v^{(k)})| \leqslant |(u, t^{(k)})| + |(s^{(k)}, v)| + |(s^{(k)}, t^{(k)})|,$$

или, в силу (238):

$$|(u, v)-(u^{(k)}, v^{(k)})| \leqslant \|u\| \cdot \|t^{(k)}\| + \|s^{(k)}\| \cdot \|v\| + \|s^{(k)}\| \cdot \|t^{(k)}\|.$$

Правая часть стремится к нулю, а потому и

$$|(u, v)-(u^{(k)}, v^{(k)})| \to 0, \text{ т. е. } (u^{(k)}, v^{(k)}) \to (u, v).$$

В частности $(u^{(k)}, u^{(k)}) \to (u, u)$, т. е. $\|u^{(k)}\|^2 \to \|u\|^2$ или $\|u^{(k)}\| \to \|u\|$.

Легко доказать также, что если последовательность числа c_k имеет предел c, то $c_k u^{(k)} \Rightarrow cu$.

Имеет место также необходимое и достаточное условие существования предела, выражаемое обычным признаком Коши. Формулируем этот признак в данном случае. Пусть имеется последовательность векторов

$$v^{(k)} \quad (k=1, 2, \ldots). \tag{248}$$

Для того чтобы эта последовательность имела предел, необходимо и достаточно выполнение следующего условия: для любого малого и положительного ε существует такое N, что

$$\|v^{(n)}-v^{(m)}\| \leqslant \varepsilon, \tag{249}$$

если только n и $m > N$.

Покажем прежде всего необходимость условия. Пусть последовательность (248) имеет предел v. Мы можем при этом написать

$$v^{(n)}-v^{(m)}=(v^{(n)}-v)+(v-v^{(m)}),$$

и отсюда по правилу треугольника

$$\|v^{(n)}-v^{(m)}\| \leqslant \|v^{(n)}-v\| + \|v-v^{(m)}\|.$$

Из определения предела непосредственно вытекает, что оба слагаемых в правой части стремятся к нулю при возрастании n и m, а следовательно, то же самое должно иметь место и для левой части, т. е. условие (249) при этом обязательно должно быть выполнено. Переходим теперь к доказательству достаточности условия (249). Положим, что это условие выполнено, и докажем, что последовательность (248) стремится к пределу. Условие (249) в раскрытом виде может быть записано так:

$$\sum_{s=1}^{\infty} |v_s^{(n)}-v_s^{(m)}|^2 \leqslant \varepsilon^2 \text{ при } n \text{ и } m > N, \tag{250}$$

где $v_s^{(j)}$ — составляющие $\boldsymbol{v}^{(j)}$. Отсюда непосредственно следует, что при любом s мы имеем: $|v_s^{(n)} - v_s^{(m)}| \leqslant \varepsilon$ при n и $m > N$, или, разлагая на вещественную и мнимую части

$$v_s^{(n)} = a_s^{(n)} + ib_s^{(n)},$$

можем написать

$$|a_s^{(n)} - a_s^{(m)}| < \varepsilon \quad \text{и} \quad |b_s^{(n)} - b_s^{(m)}| < \varepsilon \quad \text{при} \quad n \quad \text{и} \quad m > N.$$

Применяя признак Коши [I; 31], можем утверждать, что $a_s^{(n)}$ и $b_s^{(n)}$ имеют пределы при $n \to \infty$. Обозначая их через a_s и b_s, можем утверждать, что $v_s^{(n)}$ имеет пределом комплексное число $a_s + ib_s$. Покажем прежде всего, что ряд

$$\sum_{s=1}^{\infty} |v_s|^2$$

сходится, т. е. что v_s являются составляющими некоторого вектора. Удерживая в сумме (250) конечное число первых слагаемых и переходя в этой конечной сумме к пределу по n, будем иметь:

$$\sum_{s=1}^{M} |v_s - v_s^{(m)}|^2 \leqslant \varepsilon^2,$$

где M — любое целое число. Переходя в этом последнем равенстве к пределу при $M \to \infty$, получим:

$$\sum_{s=1}^{\infty} |v_s - v_s^{(m)}|^2 \leqslant \varepsilon^2, \tag{251}$$

откуда непосредственно следует, что числа $v_s - v_s^{(m)}$ образуют составляющие некоторого вектора. То же самое мы знаем и относительно чисел $v_s^{(m)}$, а следовательно, то же мы можем утверждать и об их сумме, т. е. о числах v_s. Таким образом, эти числа суть составляющие некоторого вектора $\boldsymbol{v}$, и неравенство (251) можно записать в виде:

$$\| \boldsymbol{v} - \boldsymbol{v}^{(m)} \| \leqslant \varepsilon,$$

при $m > N$, т. е. $\boldsymbol{v}^{(m)} \to \boldsymbol{v}$, и, следовательно, последовательность (248) действительно имеет предел. Каждая составляющая v_s вектора $\boldsymbol{v}$ определяется, очевидно, как предел $v_s^{(m)}$, откуда непосредственно следует, что этот предел может быть только один. Рассмотрим теперь бесконечную сумму векторов

$$\boldsymbol{u}^{(1)} + \boldsymbol{u}^{(2)} + \ldots \tag{252}$$

Она называется сходящейся, если сумма первых n слагаемых

$$s^{(n)} = u^{(1)} + \ldots + u^{(n)}$$

имеет предел в указанном выше смысле при $n \to \infty$. В силу признака Коши, необходимым и достаточным условием сходимости является выполнение неравенства

$$\| s^{(n+p)} - s^{(n)} \| = \| u^{(n+1)} + \ldots + u^{(n+p)} \| \leqslant \varepsilon \tag{253}$$

при $n > N$ и любом p.

Отметим, что из непрерывности скалярного произведения [46] непосредственно вытекает следующее: если ряд (252) сходится и s — его сумма, а z — любой элемент l_2, то

$$(s, z) = \sum_{k=1}^{\infty} (u^{(k)}, z) \text{ и } (z, s) = \sum_{k=1}^{\infty} (z, u^{(k)}), \tag{254}$$

т. е., кратко говоря, скалярные произведения (s, z) и (z, s) можно определять почленно.

Установим теперь необходимое и достаточное условие сходимости ряда (252), составленного из попарно ортогональных векторов $u^{(k)}$. Согласно признаку Коши мы должны составить выражение (253), квадрат которого, в силу теоремы Пифагора, равен

$$\| u^{(n+1)} \|^2 + \ldots + \| u^{(n+m)} \|^2.$$

Отсюда непосредственно следует, что *для сходимости ряда, составленного из попарно ортогональных векторов, необходимо и достаточно, чтобы сходился ряд из квадратов норм членов ряда.* Этот результат можно сформулировать и иначе. Пусть $x^{(k)}$ $(k = 1, 2, \ldots)$ — бесконечная ортонормированная система. Составим ряд

$$\sum_{k=1}^{\infty} \alpha_k x^{(k)}, \tag{255}$$

где α_k — некоторые комплексные числа. Из вышесказанного непосредственно следует, что *необходимым и достаточным условием сходимости ряда* (255) *является сходимость ряда*

$$\sum_{k=1}^{\infty} | \alpha_k |^2.$$

Между прочим, из этого непосредственно следует, что перестановка слагаемых ряда (255) не нарушает его сходимости. Можно показать, что при этом и сумма ряда не меняется. Положим, что ряд (255) сходится, и обозначим через s его сумму

$$s = \sum_{k=1}^{\infty} \alpha_k x^{(k)}. \tag{255_1}$$

Умножая обе части на $x^{(l)}$ и принимая во внимание ортонормированность системы $x^{(k)}$, получим: $\alpha_l = (s, x^{(l)})$ $(l = 1, 2, \ldots)$. Заменяя α_k на $(s, x^{(k)})$ и умножая обе части (255_1) скалярно на s, получим:

$$\| s \|^2 = \sum_{k=1}^{\infty} |(s, \alpha^{(k)})|^2, \tag{256}$$

т. е. *коэффициенты сходящегося ряда* (255_1) *суть коэффициенты Фурье его суммы, и для этой суммы имеет место формула замкнутости* (256) *по отношению к ортонормированной системе* $x^{(k)}$.

Вернемся к ряду (252) и положим, что члены ряда попарно ортогональны и что ряд сходится. Мы можем при этом написать $u^{(k)} = \alpha_k x^{(k)}$, где $\alpha_k = \| u^{(k)} \|$ — числа и $x^{(k)}$ образуют ортонормированную систему. Полученный выше результат может быть записан в виде:

$$\| s \|^2 = \sum_{k=1}^{\infty} \| u^{(k)} \|^2,$$

т. е. теорема Пифагора верна и для сходящихся рядов, члены которых попарно ортогональны.

47. Ортонормированные системы. Пусть имеется бесконечная ортонормированная система

$$x^{(k)} \quad (k = 1, 2, \ldots), \tag{257}$$

т. е.

$$(x^{(i)}, x^{(k)}) = \delta_{ik}, \tag{257_1}$$

и y — какой-либо вектор l_2. Составим «ряд Фурье» y относительно системы (257):

$$\sum_{k=1}^{\infty} (y, x^{(k)}) x^{(k)}. \tag{258}$$

Как мы видели [45], для любого конечного m

$$\sum_{k=1}^{m} |(y, x^{(k)})|^2 \leqslant \| y \|^2,$$

и в пределе при $m \to \infty$ получим:

$$\sum_{k=1}^{\infty} |(\mathbf{y}, \mathbf{x}^{(k)})|^2 \leqslant \|\mathbf{y}\|^2, \tag{259}$$

откуда следует сходимость ряда, стоящего в левой части этого неравенства (неравенство Бесселя), и, в силу сказанного в [**46**], ряд (258) сходится. Положим

$$\mathbf{y} = \sum_{k=1}^{\infty} (\mathbf{y}, \mathbf{x}^{(k)})\, \mathbf{x}^{(k)} + \mathbf{u}. \tag{260}$$

Как и в [**45**] можно показать, что $\mathbf{u}$ ортогонален ко всем $\mathbf{x}^{(k)}$ и теорема Пифагора дает

$$\|\mathbf{y}\|^2 = \sum_{k=1}^{\infty} |(\mathbf{y}, \mathbf{x}^{(k)})|^2 + \|\mathbf{u}\|^2.$$

Мы можем утверждать [**45**], что *знак равенства в неравенстве* (259) *равносилен тому, что вектор* $\mathbf{u}$ *в формуле* (260) *есть нулевой вектор.*

Ортонормированная система (257) называется *полной*, если не существует вектора, отличного от нулевого и ортогонального ко всем $\mathbf{x}^{(k)}$; она называется *замкнутой*, если для любого вектора $\mathbf{y}$ из l_2 имеет место уравнение замкнутости:

$$\|\mathbf{y}\|^2 = \sum_{y=1}^{\infty} |(\mathbf{y}, \mathbf{x}^{(k)})|^2. \tag{261}$$

Если система (257) полна, то вектор $\mathbf{u}$, входящий в формулу (260), есть нулевой вектор, и мы имеем разложение:

$$\mathbf{v} = \sum_{k=1}^{\infty} (\mathbf{y}, \mathbf{x}^{(k)})\, \mathbf{x}^{(k)}, \tag{262}$$

откуда, как мы видели [**46**], следует уравнение замкнутости (261). Наоборот, пусть система замкнута и вектор $\mathbf{u}$ ортогонален ко всем $\mathbf{x}^{(k)}$, т. е. $(\mathbf{u}, \mathbf{x}^{(k)}) = 0$ $(k = 1, 2, \ldots)$. При этом из формулы (261) при замене $\mathbf{y}$ на $\mathbf{u}$ следует $\|\mathbf{u}\|^2 = 0$, т. е. $\mathbf{u}$ — нулевой вектор. Итак, *полнота и замкнутость эквивалентны.* Из вышесказанного легко следует, что *полнота (замкнутость) системы* $\mathbf{x}^{(k)}$ *равносильна возможности представления любого вектора* $\mathbf{y}$ *рядом* (262).

Положим, что система (257) замкнута и выведем для любых двух векторов $\mathbf{y}$ и $\mathbf{z}$ из l_2 обобщенное уравнение замкнутости. Нетрудно видеть, что для векторов $\mathbf{y} + \mathbf{z}$ и $\mathbf{y} + i\mathbf{z}$ коэффициенты Фурье равны соответственно $(y, x^{(k)})+(z, x^{(k)})$ и $(y, x^{(k)})+i(z, x^{(k)})$. Применяем к $y+z$ и $y+iz$ уравнение замкнутости:

$$\sum_{k=1}^{\infty}[(y, x^{(k)})+(z, x^{(k)})][\overline{(y, x^{(k)})}+\overline{(z, x^{(k)})}]=(y+z, y+z),$$

$$\sum_{k=1}^{\infty}[(y, x^{(k)})+i(z, x^{(k)})][\overline{(y, x^{(k)})}-i\overline{(z, x^{(k)})}]=(y+iz, y+iz).$$

Пользуясь уравнением замкнутости для y и z, получим:

$$\sum_{k=1}^{\infty}(y, x^{(k)})\overline{(z, x^{(k)})}+\sum_{k=1}^{\infty}(z, x^{(k)})\overline{(y, x^{(k)})}=(y, z)+(z, y).$$

$$\sum_{k=1}^{\infty}(y, x^{(k)})\overline{(z, x^{(k)})}-\sum_{k=1}^{\infty}(z, x^{(k)})\overline{(y, x^{(k)})}=(y, z)-(z, y),$$

откуда и следует обобщенное уравнение замкнутости для полной ортонормированной системы (257):

$$(y, z)=\sum_{k=1}^{\infty}(y, x^{(k)})\overline{(z, x^{(k)})}. \qquad (261_1)$$

Если y совпадает с z, то эта формула переходит в (261). Обозначим через $x_s^{(k)}$ составляющие векторов $x^{(k)}$ ($k, s=1, 2, \ldots$) ортонормированной системы (257) и выпишем их в виде бесконечной матрицы:

$$\left\| \begin{matrix} x_1^{(1)}, & x_1^{(2)}, & x_1^{(3)}, & \ldots \\ x_2^{(1)}, & x_2^{(2)}, & x_2^{(3)}, & \ldots \\ x_3^{(1)}, & x_3^{(2)}, & x_3^{(3)}, & \ldots \\ \cdot & \cdot & \cdot & \cdot \end{matrix} \right\|. \qquad (263)$$

Из условий ортонормированности (257_1) непосредственно следуют формулы

$$\sum_{s=1}^{\infty} x_s^{(i)}\overline{x_s^{(k)}}=\delta_{ik} \qquad (i, k=1, 2, \ldots),$$

т. е. следует *ортонормированность столбцов матрицы* (263). Выясним теперь условия, при которых ортонормированная система (257) будет полной.

Сначала установим необходимые условия полноты. Положим, что система (257) — полная, и применим к векторам $a^{(k)}$ $(0, 0, \ldots, 1, 0, \ldots)$, которые мы ввели в [45], уравнения замкнутости (261) и (261_1)

Принимая во внимание, что $(a^{(k)}, x^{(l)}) = \overline{x_k^{(l)}}$ и что $a^{(k)}$ образуют ортонормированную систему, получим:

$$\sum_{s=1}^{\infty} x_i^{(s)} \overline{x_k^{(s)}} = \delta_{ik},$$

т. е. для полноты ортонормированной системы $x^{(k)}$ необходима ортонормированность строк матрицы (263). Покажем, что для полноты достаточно потребовать лишь нормированности строк:

$$\sum_{s=1}^{\infty} | x_k^{(s)} |^2 = 1. \tag{264}$$

Положим, что эти условия выполнены, и докажем полноту системы $x^{(k)}$. Формулы (264) являются, в силу $(a^{(k)}, x^{(s)}) = \overline{x_k^{(s)}}$ и $\| a^{(k)} \| = 1$, уравнением замкнутости для векторов $a^{(k)}$, и следовательно, имеем:

$$a^{(k)} = \sum_{s=1}^{\infty} (a^{(k)}, x^{(s)}) x^{(s)}. \tag{265}$$

Вместе с тем коэффициенты Фурье любого вектора $y(y_1, y_2, \ldots)$ по отношению к ортонормированной системе $a^{(k)}$ равны составляющим $y^{(k)}$ $(k = 1, 2, \ldots)$ и уравнение замкнутости для y совпадает просто с определением нормы y (236), т. е. система $a^{(k)}$ — полная. Пусть v — вектор, ортогональный ко всем $x^{(k)}$. Докажем, что это нулевой вектор. Из (265) следует, что $(a^{(k)}, v) = 0$ $(k = 1, 2, \ldots)$. Поскольку $a^{(k)}$ образуют полную систему, отсюда и следует, что v — нулевой вектор. Таким образом, при соблюдении условий (264) доказана полнота ортонормированной системы $x^{(k)}$. Сформулируем полученный результат: *для того чтобы ортонормированная система $x^{(k)}$ была полной (замкнутой), необходимо и достаточно, чтобы имели место формулы* (264) (*нормированность по строкам матрицы* (263)). При соблюдении этих условий будет иметь место и ортогональность строк матрицы (263).

48. Линейные преобразования с бесчисленным множеством переменных. Рассмотрим в кратких чертах линейные преобразования с бесчисленным множеством переменных:

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \ldots, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + \ldots, \\ &\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{266}$$

или

$$x' = Ax, \tag{267}$$

где A есть бесконечная матрица с элементами a_{ik}. Поставим прежде всего условие, чтобы бесконечные ряды, входящие в правые части равенств (266), были сходящимися для любого вектора $\boldsymbol{x}$ из пространства l_2. Как мы знаем, это условие будет выполнено, если ряды

$$\sum_{k=1}^{\infty} |a_{ik}|^2 \qquad (i=1,\ 2,\ \ldots)$$

будут сходящимися при всяком i. Можно показать, что это условие не только достаточно, но и необходимо. Если это условие не выполнено, то ряды, стоящие в правых частях равенств (266), будут сходящимися не для всего пространства l_2, но лишь для некоторой его части.

Естественно также поставить условие того, чтобы числа x'_k, получаемые в результате преобразования (266), также являлись составляющими некоторого вектора пространства l_2, если x_k суть составляющие некоторого вектора, т. е. чтобы ряд

$$\sum_{k=1}^{\infty} |x'_k|^2$$

был сходящимся, если только сходится ряд

$$\sum_{k=1}^{\infty} |x_k|^2.$$

Если матрица A удовлетворяет обоим вышеуказанным условиям, то соответствующее преобразование A называется *ограниченным преобразованием*. Смысл этого термина заключается в том, что для такого преобразования можно показать существование положительного числа M такого, что

$$\|\boldsymbol{x}'\|^2 \leqslant M \|\boldsymbol{x}\|^2, \tag{268}$$

или в раскрытом виде:

$$\sum_{k=1}^{\infty} |x'_k|^2 \leqslant M \sum_{k=1}^{\infty} |x_k|^2. \tag{269}$$

Остановимся на одном частном случае линейных преобразований. Рассмотрим линейное преобразование

$$\left.\begin{aligned} x'_1 &= u_{11}x_1 + u_{12}x_2 + \ldots, \\ x'_2 &= u_{21}x_1 + u_{22}x_2 + \ldots, \\ &\ldots\ldots\ldots\ldots\ldots, \end{aligned}\right\} \tag{270}$$

причем, как всегда, считаем, что ряды

$$\sum_{k=1}^{\infty} |u_{ik}|^2$$

сходятся при всяком i. Введем в рассмотрение векторы $\boldsymbol{u}^{(k)}$ с составляющими: $\bar{u}_{k1}, \bar{u}_{k2}, \ldots$, и положим, что коэффициенты u_{ik} таковы, что векторы $\boldsymbol{u}^{(k)}$ образуют полную ортонормированную систему. Как мы показали выше, это равносильно ортогональности и нормированности таблицы u_{ik} по строкам и столбцам, т. е.

$$\sum_{s=1}^{\infty} u_{sp}\bar{u}_{sq} = \delta_{pq}, \quad \sum_{s=1}^{\infty} u_{ps}\bar{u}_{qs} = \delta_{pq}. \tag{271}$$

Соответствующее преобразование (270) называется в этом случае *унитарным.*

Равенства (270) мы можем записать в виде:

$$\left.\begin{aligned} (\boldsymbol{x}, \boldsymbol{u}^{(1)}) &= x_1', \\ (\boldsymbol{x}, \boldsymbol{u}^{(2)}) &= x_2', \\ &\ldots\ldots \end{aligned}\right\} \tag{272}$$

Формула замкнутости дает:

$$\sum_{k=1}^{\infty} |x_k'|^2 = \|\boldsymbol{x}\|^2 = \sum_{k=1}^{\infty} |x_k|^2,$$

т. е., как и в случае конечного числа измерений, унитарное преобразование не меняет длины вектора, и в формуле (268) мы можем считать $M=1$.

Систему (270) или, что то же, (272) нетрудно решить относительно x_k, если задачные числа x_k' таковы, что ряд из квадратов их модулей сходится. Принимая во внимание, что векторы $\boldsymbol{u}^{(k)}$ $(\bar{u}_{k1}, \bar{u}_{k2}, \ldots)$ образуют по условию полную ортонормированную систему, получаем в силу (272)

$$\boldsymbol{x} = x_1'\boldsymbol{u}^{(1)} + x_2'\boldsymbol{u}^{(2)} + \ldots \tag{273}$$

или

$$\left.\begin{aligned} x_1 &= \bar{u}_{11}x_1' + \bar{u}_{21}x_2 + \ldots, \\ x_2 &= \bar{u}_{12}x_1' + \bar{u}_{22}x_2' + \ldots, \\ &\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{274}$$

Формулы эти показывают, что преобразование, обратное унитарному, получается заменой строк столбцами и всех элементов сопряженными, т. е. здесь имеет место полная аналогия со случаем конечного числа измерений.

В общем случае даже ограниченных матриц вопрос об обратной матрице и о приведении матрицы к диагональной форме представляет большие трудности и приводит к результатам, которые не имеют своего точного аналога в пространстве с конечным числом измерений. Подробное рассмотрение теории бесконечных матриц проведено в пятом томе. Здесь мы ограничимся указанием лишь некоторых результатов. Приведем необходимое и достаточное условие для a_{ik}, при котором формула (266) дает ограниченное преобразование. Оно формулируется так: *существует такое положительное число M, что при любом целом положительном l и для любых комплексных чисел x_s $(s=1, 2, \ldots)$ выполняется неравенство*

$$\left|\sum_{i,k=1}^{l} a_{ik} x_i \bar{x}_k\right| \leqslant M \sum_{k=1}^{l} |x_k|^2.$$

Доказывается и следующее достаточное условие ограниченности преобразования (266): *существует такое положительное число p (не зависящее от i и k), что выполняются неравенства:*

$$\sum_{i=1}^{\infty} |a_{ik}| \leqslant p \qquad \sum_{k=1}^{\infty} |a_{ik}| \leqslant p$$

$$(k=1, 2, \ldots). \qquad (i=1, 2, \ldots).$$

Если матрица A определяет ограниченное преобразование (266), то существует единственная матрица $\tilde{A}$, также определяющая ограниченное преобразование и такая, что для любых $\boldsymbol{x}$ и $\boldsymbol{y}$ выполняется равенство

$$(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \tilde{A}\boldsymbol{y}),$$

и элементы этой матрицы $\tilde{A}$ выражаются через элементы A по формулам $\tilde{a}_{ik} = \bar{a}_{ki}$. Если $\tilde{A}$ совпадает с A, т. е. $a_{ik} = \bar{a}_{ki}$, то ограниченное преобразование (266) (или матрица A) называется *самосопряженным* (*самосопряженной*).

Для ограниченных преобразований имеет место формула:

$$(A\boldsymbol{x}, \boldsymbol{y}) = \sum_{n=1}^{\infty} \Big(\sum_{m=1}^{\infty} a_{nm} x_m\Big) \bar{y}_n = \sum_{m=1}^{\infty} x_m \Big(\sum_{n=1}^{\infty} a_{nm} \bar{y}_n\Big) =$$

$$= \lim_{\substack{k \to \infty \\ l \to \infty}} \sum_{m=1}^{k} \sum_{n=1}^{l} a_{nm} x_m \bar{y}_n.$$

Отметим один важный частный случай ограниченных операторов, а именно тот случай, когда сходится двойной ряд

$$\sum_{n, m=1}^{\infty} |a_{nm}|^2. \tag{275}$$

При этом двойной ряд

$$\sum_{n, m=1}^{\infty} a_{nm} x_m \bar{y}_n$$

абсолютно сходится при любом выборе векторов $x(x_1, x_2, \ldots)$ и $y(y_1, y_2, \ldots)$. Если, кроме сходимости ряда (275), мы имеем $a_{ik} = \bar{a}_{ki}$, то имеет место возможность приведения формы Эрмита к сумме квадратов при помощи унитарного преобразования

$$\sum_{n, m=1}^{\infty} a_{nm} x_m \bar{x}_n = \sum_{k=1}^{\infty} \lambda_k z_k \bar{z}_k,$$

где λ_k — вещественные числа и вектор $z(z_1, z_2, \ldots)$ получается путем применения к вектору $x(x_1, x_2, \ldots)$ некоторого унитарного преобразования: $z = Ux$. При этом $\lambda_k \to 0$ при $k \to \infty$ (может случиться, что $\lambda_k = 0$ при всех достаточно больших k). Если A и B — две бесконечные матрицы, дающие ограниченные преобразования, то их последовательное применение дает также ограниченное преобразование, коэффициенты которого вычисляются по обычным формулам

$$\{BA\}_{ik} = \sum_{s=1}^{\infty} \{B\}_{is} \{A\}_{sk}.$$

Отметим еще, что если последовательность векторов $x^{(k)}$ имеет предел x, т. е. если $x^{(k)} \Rightarrow x$, то $Ax^{(k)} \Rightarrow Ax$, если A — матрица ограниченного преобразования.

Существенно важную роль в приложениях к математической физике играют и неограниченные линейные преобразования. Их исследование изложено в пятом томе.

49. Функциональное пространство L_2. Мы рассмотрели пространство l_2, в котором вектор определился бесчисленным множеством составляющих, которые мы нумеровали целыми числами: первая составляющая — x_1, вторая — x_2 и т. д.

Переходим теперь к рассмотрению семейства $L(\mathcal{E})$ комплексных функций $f(x) = f_1(x) + if_2(x)$, определенных на измеримом множестве $\mathcal{E}$ и таких, что вещественные функции $f_1(x)$ и $f_2(x)$ измеримы на $\mathcal{E}$ и принадлежат $L_2(\mathcal{E})$ [II, 161—163], откуда следует, что $|f(x)|$ есть измеримая на $\mathcal{E}$ функция и принадлежит $L_2(\mathcal{E})$. Если вещественные и мнимые части двух функций соответственно эквивалентны, то функции называются *эквивалентными*, и они, как элементы функционального пространства $L_2(\mathscr{E})$, отождествляются, т. е. элементом этого пространства является все множество эквивалентных функций и любая функция этого множества может быть представителем этого множества. Нулевой элемент $L_2(\mathscr{E})$ есть функция, эквивалентная нулю на $\mathscr{E}$, — например, функция, тождественно равная нулю на $\mathscr{E}$. В дальнейшем вместо $L_2(\mathscr{E})$ мы будем писать просто L_2.

Свойства функционального пространства L_2 аналогичны свойствам l_2. Элементы L_2 можно умножать на комплексные числа и складывать [II, 161]. Скалярное произведение двух элементов $f(x)$ и $g(x)$ определяется формулой

$$(f, g) = \int_{\mathscr{E}} f(x)\overline{g(x)}\,dx \tag{276}$$

и квадрат нормы элемента $f(x)$ — формулой

$$\|f\|^2 = (f, f) = \int_{\mathscr{E}} |f(x)|^2\,dx \tag{277}$$

Определение ортогональности — то же, что и для l_2, и без изменения повторяются результаты [45]. Сходимость последовательности элементов $f_n(x)$ к элементу $f(x)$ определяется формулой

$$\lim_{n\to\infty} \int_{\mathscr{E}} |f(x) - f_n(x)|^2\,dx = 0, \tag{278}$$

и повторяются без изменения определения и результаты [46] и [47], причем надо иметь в виду [II, 161—163]. Отметим, что если a_k и b_k — коэффициенты Фурье элементов $f(x)$ и $g(x)$ относительно полной ортонормированной системы $\varphi_k(x)$ $(k = 1, 2, \ldots)$, т. е.

$$a_k = \int_{\mathscr{E}} f(x)\overline{\varphi_k(x)}\,dx; \quad b_k = \int_{\mathscr{E}} g(x)\overline{\varphi_k(x)}\,dx, \tag{279}$$

то имеет место, как и в [47], обобщенное уравнение замкнутости

$$\int_{\mathscr{E}} f(x)\overline{g(x)}\,dx = \sum_{k=1}^{\infty} a_k\overline{b}_k. \tag{280}$$

Из сказанного выше следует, что сходимость ряда в L_2

$$\psi(x) = \psi_1(x) + \psi_2(x) + \ldots$$

есть сходимость в среднем

$$\lim_{n\to\infty} \int_{\mathscr{E}} |f(x) - s_n(x)|^2 dx = 0,$$

где $s_n(x)$ — сумма первых n членов ряда.

50. Связь между пространствами l_2 и L_2. Пусть, как и выше,

$$\varphi_k(x) \quad (k = 1, 2, \ldots) \tag{281}$$

— некоторая полная ортонормированная в L_2 система. При этом любому элементу из L_2 соответствует бесконечная последовательность a_k его коэффициентов Фурье, сумма квадратов модулей которых сходится, и наоборот, любой бесконечной последовательности комплексных чисел c_k $(k = 1, 2, \ldots)$, сумма квадратов модулей которых сходится, соответствует определенный элемент L_2, для которого c_k суть коэффициенты Фурье относительно системы (281) [**II, 163**]. Таким образом, эта система приводит в биоднозначное соответствие элементы L_2 и элементы l_2: всякому элементу L_2 соответствует определенный элемент l_2, и наоборот [**II, 163**]. При этом сохраняются сложение, умножение на число, скалярное произведение (в силу (280)) и норма. Если последовательность $f_n(x)$ элементов L_2 сходится к элементу $f(x)$, т. е. $\| f - f_n \| \to 0$ при $n \to \infty$, то то же будет иметь место и для соответствующих элементов l_2, что следует из равенства норм:

$$\sum_{k=1}^{\infty} | a_k - a_k^{(n)} |^2 = \int_{\mathscr{E}} | f(x) - f_n(x) |^2 dx,$$

где a_k и $a_k^{(n)}$ — составляющие элементов l_2, соответствующих $f(x)$ и $f_n(x)$.

При установлении соответствия между пространствами L_2 и l_2 мы исходили из определенной полной ортонормированной системы (281). Если взять другую такую же систему

$$\psi_k(x) \quad (k = 1, 2, \ldots), \tag{282}$$

то, конечно, закон соответствия будет уже иным. Всякая функция системы (282) разлагается в ряд Фурье по $\varphi_k(x)$:

$$\psi_k(x) = \sum_{i=1}^{\infty} u_{ik}\varphi_i(x) \quad (k = 1, 2, \ldots), \tag{283}$$

причем ряд, стоящий справа, сходится в L_2. Принимая во внимание обобщенное уравнение замкнутости, получим

$$\sum_{s=1}^{\infty} u_{si}\bar{u}_{sk} = \delta_{ik} \quad (i, k = 1, 2, \ldots),$$

т. е. матрица U с элементами u_{ik} ортонормирована по строкам. Принимая во внимание, что

$$u_{ik} = \int\limits_{\mathscr{E}} \psi_k(x) \overline{\varphi_i(x)}\, dx$$

и полноту системы (281), получаем

$$\sum_{i=1}^{\infty} |u_{ik}|^2 = \int\limits_{\mathscr{E}} |\psi_k(x)|^2\, dx = 1,$$

откуда следует, что таблица U есть таблица унитарного преобразования. Нетрудно показать, что и наоборот, всякая таблица унитарного преобразования U приводит согласно (283) к полной ортонормированной системе (282), если такою же была система (281). Приведем в качестве примера для пространства L_2 функций одной независимой переменной на отрезке $(-\pi, \pi)$ систему функций [**II, 174**]:

$$\varphi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx} \quad (k = 1, 2, \ldots). \tag{284}$$

Нетрудно доказать, что это — ортонормированная система. Отметим, что нумерация по k производится здесь не от 1 до ∞, а от $(-\infty)$ до $(+\infty)$. Такое изменение нумерации несущественно. Пользуясь результатами [**II, 169**] можно показать, что система (284) — полная.

51. Линейные операторы в L_2. Положим, что имеет место определенный закон, согласно которому всякой функции $f(x)$ из $L_2(\mathscr{E})$ соответствует некоторая другая функция $F(x)$ из того же $L_2(\mathscr{E})$:

$$F(x) = A[f(x)], \tag{285}$$

где A — символическое обозначение этого закона соответствия. Мы имеем здесь как бы обобщенное понятие функции: роль аргумента играет не любое число из некоторого множества, например, промежутка, а любая функция из $L_2(\mathscr{E})$, и значением функции является также некоторая функция из $L_2(\mathscr{E})$. Такое соответствие, устанавливаемое формулой (285), называется обычно *функциональным оператором*. Оператор A называется *линейным*, если

$$\left.\begin{aligned} A[f(x)+g(x)] &= A[f(x)] + A[g(x)]; \\ A[\alpha f(x)] &= \alpha A[f(x)], \end{aligned}\right\} \tag{286}$$

где α — любое комплексное число, и *ограниченным*, если существует такое положительное число M, что

$$\| Af \| \leqslant M \| f \| \tag{287}$$

при любом $f(x)$ из $L_2(\mathscr{E})$.

Если последовательность $f_n(x)$ сходится к $f(x)$ в L_2 и A — линейный ограниченный оператор, то $Af_n(x)$ сходится к $Af(x)$ в L_2. Это непосредственно следует из неравенства

$$\| A[f(x)-f_n(x)] \| \leqslant M \| f-f_n \|,$$

т. е.

$$\| Af-Af_n \| \leqslant M \| f-f_n \|.$$

Нетрудно сформулировать для рассматриваемого случая определения эрмитовского и унитарного преобразований. Линейный ограниченный оператор A называется *эрмитовским* (*самосопряженным*), если

$$(Af,\ g)=(f,\ Ag) \tag{288}$$

для любых $f(x)$ и $g(x)$ из L_2. Линейный оператор U называется *унитарным*, если он биоднозначно преобразует L_2 в себя,

$$U[f(x)]=F(x), \tag{289}$$

и не меняет нормы:

$$\| Uf(x) \| = \| f \|.$$

Указанная биоднозначность преобразования (289) сводится к следующему: не только всякому элементу $f(x)$ из L_2 соответствует определенный элемент $F(x)$, но и всякому $F(x)$ из L_2 соответствует один определенный прообраз $f(x)$. Из этого следует, что для U имеется обратный оператор U^{-1}, который по $F(x)$ восстанавливает $f(x)$. Нетрудно видеть, что U^{-1} — также унитарный оператор. Для унитарного оператора неравенство (287) можно заменить равенством, положив $M=1$. Легко показать, что унитарное преобразование не меняет не только нормы, но и скалярного произведения, т. е.

$$(Uf,\ Ug)=(f,\ g)$$

для любых $f(x)$ и $g(x)$ из L_2, и что U преобразует всякую полную ортонормированную систему из L_2 в такую же систему. Определим еще понятие сопряженного оператора. Оператор A^* называется *сопряженным* с линейным ограниченным оператором A, если для любых $f(x)$ и $g(x)$ из L_2 имеет место равенство:

$$(Af,\ g)=(f,\ A^*g).$$

Можно показать, что для всякого линейного ограниченного оператора A существует единственный сопряженный оператор A^*. Рассмотрим пространство L_2 функций $f(x)$ одного независимого переменного, определенных на конечном или бесконечном промежутке $(a,\ b)$. В таком L_2 линейные операторы часто определяются формулой

$$F(x)=\int_a^b K(x,\ t)f(t)\,dt, \tag{290}$$

где $K(x, t)$ — функция, определенная (измеримая) в квадрате $a \leqslant x \leqslant b$, $a \leqslant t \leqslant b$. Операторы вида (290) называются обычно *интегральными* операторами, а функция $K(x, t)$ — *ядром* оператора. Если

$$\int_a^b \int_a^b |K(x, t)|^2 \, dx \, dt < +\infty,$$

или если существует такое число p, что $\int_a^b |K(x, t)| \, dt \leqslant p$ и $\int_a^b |K(x, t)| \, dx \leqslant p$ при всех t, x из (a, b), то формула (290) определяет линейный ограниченный оператор. Можно показать, что оператор, сопряженный с ограниченным оператором (290), есть также интегральный оператор с ядром $K^*(x, t) = \overline{K(t, x)}$. Примером унитарного преобразования в L_2 на промежутке $(-\infty, +\infty)$ является преобразование Фурье:

$$F(x) = Uf(x) = \lim_{\substack{n \to \infty \\ m \to \infty}} \frac{1}{\sqrt{2\pi}} \int_{-n}^{m} e^{-ixt} f(t) \, dt,$$

где lim есть предел функции от x, полученной в результате интегрирования в L_2 на промежутке $(-\infty, +\infty)$, т. е. если обозначим

$$\varphi_{n, m}(x) = \frac{1}{\sqrt{2\pi}} \int_{-n}^{m} e^{-ixt} f(t) \, dt,$$

то

$$\lim_{\substack{n \to \infty \\ m \to \infty}} \int_{-\infty}^{+\infty} |F(x) - \varphi_{n, m}(x)|^2 \, dx = 0.$$

Если $f(x)$ не только принадлежит L_2 на промежутке $(-\infty, +\infty)$, но и суммируема на этом промежутке, то преобразование Фурье можно записать в обычной форме

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ixt} f(t) \, dt.$$

Неизменность нормы при преобразовании Фурье имеет вид:

$$\int_{-\infty}^{+\infty} |F(x)|^2 \, dx = \int_{-\infty}^{+\infty} |f(x)|^2 \, dx.$$

Рассмотрим теперь один частный пример. Пусть в пространстве L_2 на промежутке $(-\pi, \pi)$ задан линейный оператор умножения на независимую переменную:

$$A[f(x)] = xf(x). \tag{291}$$

Имеем, очевидно:

$$\|Af\|^2 = \int_{-\pi}^{\pi} |x|^2 |f(x)|^2 dx \leqslant \pi^2 \|f\|^2,$$

т. е. для оператора (291) в формуле (287) можно считать $M = \pi$. Построим линейное преобразование, выражающее оператор (291) в l_2, если принять за основу полную ортонормированную систему (284). Пусть c_k — коэффициенты Фурье $f(x)$ относительно системы (284):

$$c_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-ikx} f(x)\, dx \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

и c_k' — коэффициенты Фурье $xf(x)$:

$$c_k' = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-ikx} xf(x)\, dx \qquad (k = 0, \pm 1, \pm 2, \ldots).$$

Нам надо построить линейное преобразование (бесконечную матрицу), выражающее c_k' через c_k. Определим коэффициенты Фурье функции $\frac{1}{\sqrt{2\pi}} e^{ikx} x$:

$$b_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(k-m)} x\, dx \qquad (m = 0, \pm 1, \pm 2, \ldots).$$

Интегрируя по частям, получим

$$b_m = \frac{(-1)^{k-m}}{i(k-m)} \quad \text{при} \quad m \neq k; \quad b_k = 0.$$

Переписывая выражение c_k' в виде:

$$c_k' = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \overline{e^{ikx} x} f(x)\, dx$$

и применяя обобщенное уравнение замкнутости, получим

$$c_k' = i \sum_{m=-\infty}^{+\infty}{}' \frac{(-1)^{k-m}}{k-m} c_m \qquad (k = 0, \pm 1, \pm 2, \ldots), \tag{292}$$

где штрих над знаком суммы показывает, что надо исключить слагаемое, соответствующее $m=k$. Формула (292) и дает линейное преобразование в l_2, соответствующее оператору (291) в L_2, если за координатные функции в пространстве L_2 взяты функции (284). Нетрудно показать, что оператор (291) — самосопряженный.

Проведем общее рассуждение для любого линейного ограниченного самосопряженного оператора A, причем будем записывать коэффициенты Фурье в виде скалярного произведения и учтем определение (288) самосопряженного оператора. Введем коэффициенты Фурье c_k и c_k' для $f(x)$ и $Af(x)$ относительно полной ортонормированной системы $\varphi_k(x)$:

$$c_k=(f,\ \varphi_k);\quad c_k'=(Af,\ \varphi_k)=(f,\ A\varphi_k). \tag{293}$$

Введем теперь коэффициенты Фурье функций $A\varphi_m(x)$

$$a_{km}=(A\varphi_m,\ \varphi_k)=(\varphi_m,\ A\varphi_k), \tag{294}$$

откуда следует, что $a_{mk}=\overline{a_{km}}$, ибо по определению

$$a_{mk}=(A\varphi_k,\ \varphi_m).$$

Принимая во внимание обобщенное уравнение замкнутости, формулу для c_k' и (294), получим:

$$c_k'=\sum_{m=1}^{\infty}(f,\ \varphi_m)\overline{(A\varphi_k,\ \varphi_m)}=\sum_{m=1}^{\infty}(f,\ \varphi_m)(\varphi_m,\ A\varphi_k),$$

т. е.

$$c_k'=\sum_{m=1}^{\infty}a_{km}c_m\qquad (k=1,\ 2,\ \ldots).$$

Это преобразование и выражает оператор A в l_2, если $\varphi_k(x)$ приняты за координатные функции в L_2.

Мы рассматривали в предыдущих параграфах три случая, когда линейный оператор задан на всем L_2 и ограничен. Если мы возьмем, например, оператор дифференцирования $A[f(x)]=f'(x)$, то он задан не на всем L_2, ибо не всякая функция из L_2 имеет производную. Кроме того, указанный оператор не ограничен на том множестве функций, на котором он задан.

Более подробное и строгое изложение материала последних параграфов будет дано в пятом томе.

ГЛАВА III

ОСНОВЫ ТЕОРИИ ГРУПП И ЛИНЕЙНЫЕ ПРЕДСТАВЛЕНИЯ ГРУПП

§ 5. ОСНОВЫ ОБЩЕЙ ТЕОРИИ ГРУПП

52. Группы линейных преобразований. Рассмотрим совокупность всех унитарных преобразований в n-мерном пространстве. Все эти преобразования имеют определитель, отличный от нуля, так что для каждого унитарного преобразования Ux, которое вполне характеризуется соответствующей матрицей U, существует вполне определенное обратное преобразование $U^{-1}x$, которое также будет унитарным [28]. Кроме того, если U_1x и U_2x — два унитарных преобразования, то и их произведение U_2U_1x также будет унитарным преобразованием. Все эти свойства совокупности всех унитарных преобразований выражают коротко тем, что говорят, что *совокупность унитарных преобразований образует группу.*

Вообще совокупность некоторых линейных преобразований с определителем, отличным от нуля, образует группу, если выполнены следующие два условия: во-первых, *если некоторое преобразование принадлежит нашей совокупности, то и обратное преобразование также принадлежит совокупности*, *и*, во-вторых, *произведение двух преобразований, принадлежащих нашей совокупности (при любом порядке сомножителей) также принадлежит нашей совокупности, причем множители могут быть и одинаковыми.*

Принимая во внимание, что произведение всякого преобразования на обратное ему есть гождественное преобразование, мы можем утверждать, что *группа обязательно должна содержать тождественное преобразование,* т. е. *единичную матрицу.*

Вообще линейное преобразование вполне определяется своей матрицей, и во всем предыдущем так же, как и впоследствии, мы можем говорить или о *группе линейных преобразований,* или о *группе матриц.*

Приведем еще примеры групп линейных преобразований. Нетрудно видеть, что совокупность всех вещественных ортогональных преобразований образует группу. Как известно, эти вещественные ортогональные преобразования имеют определитель, равный (± 1). Если взять совокупность вещественных ортогональных преобразований с определителем ($+1$), то они также образуют группу. Но если мы возьмем совокупность вещественных ортогональных преобразований с определителем (-1), то они уже не образуют группу, так как произведение двух матриц с определителем (-1) дает матрицу с определителем ($+1$).

В частности, если мы рассмотрим группу вещественных ортогональных преобразований с тремя переменными, то это будет группа, состоящая из чистых вращений пространства вокруг начала и из преобразований, которые получаются в результате такого вращения и преобразований симметрии относительно начала. Если же мы возьмем группу линейных ортогональных преобразований с тремя переменными, с определителем ($+1$), то это будет группа вращения пространства вокруг начала.

Во всех рассмотренных случаях группа содержала бесчисленное множество преобразований, в частности группа вращения трехмерного пространства вокруг начала зависела от трех произвольных вещественных параметров — углов Эйлера, о которых мы говорили выше.

В качестве следующего примера рассмотрим вращение пространства вокруг оси Z на угол φ. Соответствующие формулы имеют вид:

$$\left.\begin{aligned} x' &= x\cos\varphi - y\sin\varphi \\ y' &= x\sin\varphi + y\cos\varphi. \end{aligned}\right\} \tag{1}$$

При возможных значениях вещественного параметра φ в промежутке $(0, 2\pi)$ мы получаем, очевидно, группу, содержащую бесчисленное множество преобразований и зависящую от одного вещественного параметра. Введем следующее обозначение для матрицы преобразования:

$$Z_\varphi = \begin{Vmatrix} \cos\varphi, & -\sin\varphi \\ \sin\varphi, & \cos\varphi \end{Vmatrix}. \tag{2}$$

Непосредственно очевидно, что произведение двух вращений на угол φ_1 и φ_2 дает вращение на угол $(\varphi_1 + \varphi_2)$:

$$Z_{\varphi_2} Z_{\varphi_1} = Z_{\varphi_2 + \varphi_1} \tag{3}$$

и точно так же

$$Z_{\varphi_1} Z_{\varphi_2} = Z_{\varphi_1 + \varphi_2}.$$

Мы видим, таким образом, что в данном случае все преобразования группы или, как говорят, элементы группы попарно коммутируют. Такая группа называется *абелевой группой*. Кроме того, в последнем примере перемножение двух элементов группы сводится просто к сложению значений параметра φ, соответствующих перемножаемым матрицам.

Мы можем несколько расширить последнюю группу, взяв не только вращение плоскости XY вокруг начала, но и зеркальное отображение, т. е. симметрию, относительно оси Y, причем, очевидно, безразлично, в каком порядке производить эти операции — сначала вращение вокруг начала, а затем симметрию относительно оси Y или наоборот. Перемена порядка повлияет на результат, но общая совокупность преобразований будет одна и та же при обоих способах. Это будут вещественные ортогональные преобразования с двумя переменными. Общий вид соответствующих матриц будет

$$\{\varphi, d\} = \begin{Vmatrix} d\cos\varphi, & -d\sin\varphi \\ \sin\varphi, & \cos\varphi \end{Vmatrix}, \tag{4}$$

где φ — прежний параметр, а d — число, равное ± 1. При $d = 1$ мы получаем простое вращение плоскости XY вокруг начала, а при $d = -1$ получается вращение, после которого производится упомянутая выше симметрия. Нетрудно проверить, что для произведения матриц (4) мы будем иметь следующее правило:

$$\{\varphi_2, d_2\}\ \{\varphi_1, d_1\} = \{\varphi_1 + d_1\varphi_2, d_1 d_2\}. \tag{5}$$

В данном случае произведение уже может зависеть от порядка сомножителей, т. е. группа уже не будет абелевой. Точно так же, очевидно, группа вещественных ортогональных преобразований в трехмерном пространстве или даже группа вращения трехмерного пространства вокруг начала не будет абелевой.

До сих пор мы приводили примеры групп, содержащих бесчисленное множество преобразований (элементов), и соответствующие матрицы содержали произвольные вещественные параметры. Сейчас приведем некоторые примеры групп, содержащих конечное число элементов. Пусть m — некоторое целое положительное число. Рассмотрим совокупность вращений плоскости XY вокруг начала на углы

$$0, \frac{2\pi}{m}, \frac{4\pi}{m}, \ldots, \frac{2(m-1)\pi}{m}.$$

Мы будем иметь здесь всего m преобразований, матрицы которых будут

$$Z_{\frac{2k\pi}{m}} = \begin{Vmatrix} \cos\frac{2k\pi}{m}, & -\sin\frac{2k\pi}{m} \\ \sin\frac{2k\pi}{m}, & \cos\frac{2k\pi}{m} \end{Vmatrix} \qquad (k = 0, 1, \ldots, m-1).$$

Эти преобразования образуют, очевидно, группу, и элементы этой группы суть целые положительные степени одного и того же преобразования, а именно

$$Z_{\frac{2k\pi}{m}} = \left(Z_{\frac{2\pi}{m}}\right)^k \qquad (k = 0, 1, \ldots, m-1). \tag{6}$$

Такая конечная группа, состоящая из степеней некоторого преобразования, называется обычно *циклической*.

Если мы возьмем некоторый угол φ_0, не соизмеримый с π, то преобразования (матрицы)

$$Z^k_{\varphi_0} = Z_{k\varphi_0} \qquad (k = 0, \pm 1, \pm 2, \ldots) \tag{7}$$

образуют также, очевидно, группу. Но эта группа будет состоять уже из бесчисленного множества элементов, так как ни при каких целых показателях матрица $Z^k_{\varphi_0}$ не совпадает с $Z^0_{\varphi_0} = I$. Группа (7) будет бесконечной группой, но ее матрицы не содержат никакого непрерывно меняющегося параметра. В данном случае, как говорят, *число элементов в группе будет счетным*, т. е. мы можем пронумеровать все элементы группы целыми числами, снабдив всякий элемент группы значком, равным целому числу, так что разным элементам будут соответствовать разные значки, и всякое целое число будет знаком у некоторого элемента. Этого нельзя сделать в случае групп, содержащих непрерывно меняющиеся параметры.

53. Группы правильных многогранников. Приведем еще примеры конечных групп, причем образуем их из вращений трехмерного пространства вокруг начала. Такие вращения в определенной координатной системе выражаются, как мы знаем, некоторыми линейными преобразованиями координат. Заметим, что когда мы говорим о вращении пространства вокруг начала, то подразумеваем под этим лишь окончательный эффект перехода из начального положения в преобразованное. Каким путем этот переход совершается, это совершенно не входит в наше рассмотрение. Действительно, всякое линейное преобразование определяет координаты преобразованной точки, но, конечно, ничего не говорит о самом пути преобразования. Рассмотрение самого пути преобразования совершенно при этом не входит в рассуждения.

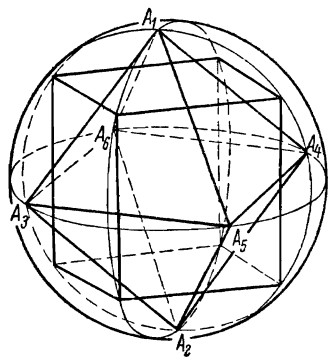


Рис. 2.

Рассмотрим сферу с центром в начале и радиусом единица. Впишем в эту сферу какой-нибудь правильный многогранник, например октаэдр (рис. 2). Поверхность этого многогранника состоит, как известно, из восьми равносторонних треугольников. Рассмотрим теперь совокупность тех вращений трехмерного пространства вокруг начала, при которых взятый нами октаэдр совмещается сам с собой. Нетрудно видеть, что совокупность этих вращений образует группу и что эта группа содержит конечное число элементов. Подсчитаем число элементов группы. Возьмем какую-нибудь ось октаэдра, соединяющую две противоположные вершины. Октаэдр совместится сам с собой, если мы повернем пространство вокруг упомянутой оси на угол 0, $\frac{\pi}{2}$, π, $\frac{3\pi}{2}$. Вращению на угол 0 соответствует, очевидно, тождественное преобразование, т. е. единичная матрица. Обозначим упомянутые четыре вращения вокруг взятой оси через

$$S_0 = I, \ S_1, \ S_2, \ S_3. \tag{8}$$

Пусть A — одна из вершин октаэдра, лежащая на взятой оси. Введем в рассмотрение пять линейных преобразований

$$T_1, \ T_2, \ T_3, \ T_4, \ T_5,$$

при которых октаэдр совмещается сам с собой, а его вершина A совпадает с одной из остальных пяти вершин октаэдра. Наряду с четырьмя вращениями (8) составим еще 20 вращений пространства вокруг начала следующего вида:

$$T_kS_0, \ T_kS_1, \ T_kS_2, \ T_kS_3 \qquad (k = 1, 2, 3, 4, 5). \tag{9}$$

Нетрудно проверить, что 24 вращения (8) и (9) различны. Это совершенно очевидно и геометрически, а также может быть доказано следующим путем: пусть

$$T_pS_q = T_{p_1}S_{q_1}. \tag{10}$$

Преобразования S_i соответствуют повороту вокруг оси, проходящей через вершину A, и при этих преобразованиях A остается на месте. Преобразования T_p и T_{p_1} при различных значках p и p_1 переводят вершину A в различные вершины, и, следовательно, из равенства (10) вытекает, что значки p и p_1 должны быть одинаковы, но тогда, очевидно, из этого же равенства будет путем умножения слева на $T_p^{-1} = T_{p_1}^{-1}$ следовать, что и значки q и q_1 также должны быть одинаковы, т. е. равенство (10) может иметь место только тогда, когда правая и левая части состоят из одинаковых сомножителей. Следовательно, выражения (8) и (9) дают нам 24 различных вращения, при которых октаэдр совмещается сам с собой. Покажем теперь, что этим и исчерпываются все вращения, обладающие этим свойством. Действительно, пусть V — некоторое вращение, при котором наш октаэдр переходит сам в себя. Положим, что при этом вершина A совмещается с некоторой другой вершиной A_j и пусть T_j — то из преобразований T_k, которое преобразует A также в A_j. Составим преобразование $T_j^{-1}V$. При этом октаэдр переходит в себя, и вершина A остается на месте. Следовательно, остается на месте и противоположная вершина, и составленное нами преобразование есть одно из вращений S_i вокруг оси, проходящей через вершину A, т. е. $T_j^{-1}V = S_i$, и отсюда $V = T_jS_i$. Иначе говоря, всякое вращение, при котором октаэдр переходит в себя, должно заключаться среди тех 24 вращений, которые мы составили выше. Итак, окончательно, ***группа вращения, при которой октаэдр переходит в себя, содержит* 24 *элемента.***

Мы можем, очевидно, вписать в сферу единичного радиуса куб таким образом, чтобы радиусы, идущие в центры грани октаэдра, имели своими концами вершины куба. Отсюда непосредственно следует, что группа вращения для куба будет той же самой, что и для октаэдра. Положим, что мы иначе выбрали положение октаэдра, а именно, что новое положение октаэдра получается из первоначального при помощи вращения, осуществляемого некоторой матрицей U. Если V есть некоторое вращение, при котором прежний октаэдр переходил сам в себя, то, очевидно, UVU^{-1} будет давать такое вращение, при котором новый октаэдр будет переходить в себя, и наоборот. Таким образом, если группа вращения прежнего октаэдра состояла из матриц V_k $(k=1, 2, \ldots, 24)$, то группа вращения нового октаэдра будет просто состоять из подобных матриц UV_kU^{-1}. Иначе говоря, получается *подобная группа*. Вообще, ***если совокупность некоторых матриц V_k образует группу, то совокупность подобных матриц UV_kU^{-1} при любой фиксированной матрице U также образует группу.*** Это нетрудно непосредственно доказать из определения группы, что мы и предлагаем проделать читателю. Вторая группа называется обычно *подобной* первой.

Рассмотрим тетраэдр, поверхность которого состоит из четырех равносторонних треугольников и имеет четыре вершины. Возьмем какую-нибудь ось тетраэдра, соединяющую его вершину A с центром противолежащей грани. Тетраэдр совместится сам с собою, если повернем пространство вокруг упомянутой оси в некотором направлении на угол $0, \frac{2\pi}{3}, \frac{4\pi}{3}$. Пусть S_0, S_1, S_2 — эти вращения. Далее вводим три линейных преобразования T_1, T_2, T_3, при которых тетраэдр совмещается сам с собою, а его вершина A совпадает с одной из остальных трех его вершин. Наряду с вращениями S_0, S_1, S_2 составим девять вращений T_kS_0, T_kS_1, T_kS_2 $(k=1, 2, 3)$. Полученные 12 вращений различны, и это суть все вращения, при которых тетраэдр переходит в себя.

Рассмотрим теперь икосаэдр, поверхность которого состоит из 20 равносторонних треугольников и имеет 12 вершин. Возьмем, как и выше, какую-нибудь ось икосаэдра, соединяющую его вершину A с противоположной вершиной. Икосаэдр совместится сам с собою, если повернем пространство на угол $\frac{2k\pi}{5}$ $(k=0, 1, 2, 3, 4)$. Пусть S_k — эти вращения. Далее имеем 11 вращений T_l $(l=1, 2, \ldots, 11)$, при которых икосаэдр совмещается сам с собою, а вершина A переходит в одну из остальных вершин. Полная группа вращений, при которых икосаэдр переходит в себя, состоит из пяти вращений S_k и 55 вращений $T_l S_k$. Таким образом, эта группа содержит 60 вращений. Такой же будет и группа додекаэдра, поверхность которого состоит из 12 правильных пятиугольников и содержит 20 вершин. Чтобы убедиться в этом, надо расположить додекаэдр относительно икосаэдра так же, как это выше мы сделали для куба относительно октаэдра.

Рассмотрим еще одну группу, состоящую из вращений трехмерного пространства. Пусть в плоскости XY находится правильный n-угольник, центр которого совпадает с началом координат. Возьмем какую-нибудь ось n-угольника, соединяющую его вершину A с противоположной вершиной (если n — четно), или с серединой противоположной стороны (если n — нечетно). При вращении плоскости XY вокруг этой оси на угол 0 и π, n-угольник совмещается сам с собою. Первое вращение есть тождественное преобразование I, а второе мы обозначим через S.

Кроме того, мы имеем вращения T_k вокруг оси Z на угол $\frac{2k\pi}{n}$ $(k = 1, 2, \ldots, n-1)$, при которых n-угольник тоже совмещается сам с собою, а его вершина A переходит в одну из других вершин. При $k = 0$ получаем тождественное преобразование $T_0 = I$. Полная группа преобразований, при которых n-угольник переходит в себя, будет состоять из следующих $2n$ преобразований: T_k и $T_k S$ $(k = 0, 1, 2, \ldots, n-1)$.

Указанный n-угольник, поверхность которого считается дважды (верх и низ), называется обычно диэдром, а построенная группа — группой диэдра.

54. Преобразования Лоренца. Все примеры групп линейных преобразований, которые мы приводили выше, состояли из унитарных преобразований или из вращений трехмерного пространства (частный случай унитарных преобразований). Сейчас мы изучим некоторую новую группу линейных преобразований, элементы которой уже не являются унитарными преобразованиями. Эта группа играет важную роль в принципе относительности, электродинамике и той части квантовой механики, которая связана с принципом относительности.

Рассмотрим четыре переменные x_1, x_2, x_3, x_4, из которых первые три суть пространственные координаты точки, а последняя переменная есть время. В связи с основным требованием специального принципа относительности о неизменности некоторой определенной скорости c (скорость света) в случае относительного движения, возникает вопрос о таких линейных преобразованиях упомянутых выше четырех переменных, при которых выражение

$$x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2$$

остается неизменным, т. е., подробнее говоря, мы должны найти такие линейные преобразования, выражающие новые переменные x'_k через прежние x_k, чтобы имело место тождество:

$$x_1'^2 + x_2'^2 + x_3'^2 - c^2 x_4'^2 = x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2.$$

Рассмотрим сначала тот случай, когда координаты x_2 и x_3 остаются неизменными, и в линейном преобразовании участвуют лишь переменные x_1 и x_4. Мы должны, таким образом, найти такие линейные преобразования

$$x'_1 = a_{11}x_1 + a_{14}x_4, \qquad x'_4 = a_{41}x_1 + a_{44}x_4, \tag{11}$$

чтобы

$$x_1'^2 - c^2 x_4'^2 = x_1^2 - c^2 x_4^2. \tag{12}$$

Введем вместо x_4 новую чисто мнимую переменную y_1, согласно формуле

$$y_1 = icx_4.$$

Искомые линейные преобразования должны иметь вид:

$$x'_1 = \alpha_{11}x_1 + \alpha_{12}y_1, \qquad y'_1 = \alpha_{21}x_1 + \alpha_{22}y_1, \tag{13}$$

где

$$\alpha_{11} = a_{11}; \quad \alpha_{12} = \frac{a_{14}}{ic}; \quad \alpha_{21} = ica_{41}; \quad \alpha_{22} = a_{44},$$

а условие (12) переписывается при этом следующим образом:

$$x_1'^2 + y_1'^2 = x_1^2 + y_1^2. \tag{14}$$

Коэффициенты α_{11} и α_{22} должны быть вещественными, а α_{12} и α_{21} — чисто мнимыми. Обозначим поэтому $\alpha_{12} = i\beta_{12}$ и $\alpha_{21} = i\beta_{21}$. Условие (14) равносильно, очевидно, условию ортогональности преобразования (13), и, следовательно, сумма квадратов элементов каждой строки и столбца должна равняться единице. Это сразу дает $\beta_{12}^2 = \beta_{21}^2 = \alpha_{11}^2 - 1 = \alpha_{22}^2 - 1$ и $\alpha_{11}^2 = \alpha_{22}^2$. Положим $\alpha_{22} = \alpha$ и $\beta_{12} = \alpha\beta$. Будем считать положительными коэффициенты α_{11} и α_{22}, что соответствует неизменности в направлении отсчета x_1 и x_4. Мы получим таким образом вместо (13) в силу предыдущих соотношений:

$$x'_1 = \alpha x_1 + i\alpha\beta y_1, \qquad y'_1 = \alpha_{21}x_1 + \alpha y_1.$$

Условие ортогональности строк

$$\alpha\alpha_{21} + i\alpha^2\beta = 0$$

даст нам $\alpha_{21} = -i\alpha\beta$, т. е. β_{12} и β_{21} должны быть противоположных знаков. Наконец, условие

$$\alpha_{11}^2 + \alpha_{12}^2 = 1$$

даст

$$\alpha^2 - \alpha^2\beta^2 = 1 \text{ или } \alpha = \frac{1}{\sqrt{1-\beta^2}} \quad (\beta^2 < 1),$$

и окончательно мы приходим к следующим формулам:

$$x_1' = \frac{x_1 + i\beta y_1}{\sqrt{1-\beta^2}}; \quad y_1' = \frac{-i\beta x_1 + y_1}{\sqrt{1-\beta^2}},$$

или, возвращаясь вновь от переменной $y_1 = icx_4$ к прежней переменной x_4:

$$x_1' = \frac{x_1 - \beta c x_4}{\sqrt{1-\beta^2}}; \quad x_4' = \frac{-\frac{\beta}{c} x_1 + x_4}{\sqrt{1-\beta^2}}. \tag{15}$$

Из этих равенств непосредственно следует, что координатная система, которой соответствуют переменные со штрихами, двигается по отношению к первоначальной координатной системе со скоростью

$$v = \beta c \tag{16}$$

в направлении оси x_1. Действительно, если принять x_1' постоянным, то получим:

$$dx_1 - \beta c\, dx_4 = 0, \quad \text{т. е.} \quad \frac{dx_1}{dx_4} = \beta c.$$

Вводя вместо β скорость v согласно формуле (16) и заменяя x_1 на x и x_4 на t, получим обычную форму *преобразования Лоренца* с двумя переменными

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad t' = \frac{-\frac{v}{c^2} x + t}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{17}$$

В предельном случае при $c \to \infty$ мы получаем обычные формулы относительного движения классической механики

$$x' = x - vt; \quad t' = t.$$

Нетрудно проверить, что преобразования Лоренца (17), зависящие от одного вещественного параметра v, образуют группу. Решая уравнения (17) относительно x и t, получим преобразование, обратное (17). Покажем, что это будет то же преобразование Лоренца, которое получается из преобразования (17) заменой v на $(-v)$. Действительно, решая уравнения (17), будем иметь:

$$\left(1 - \frac{v^2}{c^2}\right) x = \sqrt{1 - \frac{v^2}{c^2}}\,(x' + vt'); \quad \left(1 - \frac{v^2}{c^2}\right) t = \sqrt{1 - \frac{v^2}{c^2}} \left(\frac{v}{c^2} x' + t'\right),$$

откуда непосредственно следует

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad t = \frac{\frac{v}{c^2} x' + t'}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Рассмотрим теперь два преобразования Лоренца L_1 и L_2, соответствующих значениям параметра $v = v_1$ и $v = v_2$. Составим их произведение L_2L_1 и покажем, что это тоже будет преобразование Лоренца. Нам надо составить произведение из двух матриц

$$\left\|\begin{matrix} \frac{1}{\sqrt{1-\beta_2^2}} & -\frac{\beta_2 c}{\sqrt{1-\beta_2^2}} \\ -\frac{\frac{\beta_2}{c}}{\sqrt{1-\beta_2^2}} & \frac{1}{\sqrt{1-\beta_2^2}} \end{matrix}\right\| \left\|\begin{matrix} \frac{1}{\sqrt{1-\beta_1^2}} & -\frac{\beta_1 c}{\sqrt{1-\beta_1^2}} \\ -\frac{\frac{\beta_1}{c}}{\sqrt{1-\beta_1^2}} & \frac{1}{\sqrt{1-\beta_1^2}} \end{matrix}\right\|,$$

где

$$\beta_1=\frac{v_1}{c}; \qquad \beta_2=\frac{v_2}{c}.$$

Применяя обычные правила умножения матриц, получим для произведения следующую матрицу:

$$\frac{1+\beta_1\beta_2}{\sqrt{1-\beta_2^2}\,\sqrt{1-\beta_1^2}}\left\|\begin{matrix} 1 & -\frac{\beta_1 c+\beta_2 c}{1+\beta_1\beta_2} \\ -\frac{\frac{\beta_1}{c}+\frac{\beta_2}{c}}{1+\beta_1\beta_2} & 1 \end{matrix}\right\|. \tag{18}$$

Введем новую величину

$$v_3=\frac{v_1+v_2}{1+\frac{v_1v_2}{c^2}}. \tag{19}$$

Нетрудно проверить справедливость тождества:

$$\frac{1+\frac{v_1v_2}{c^2}}{\sqrt{1-\frac{v_2^2}{c^2}}\,\sqrt{1-\frac{v_1^2}{c^2}}}=\frac{1}{\sqrt{1-\frac{v_3^2}{c^2}}},$$

и в результате матрица (18) может быть написана в следующем виде:

$$\left\|\begin{matrix} \frac{1}{\sqrt{1-\beta_3^2}} & -\frac{\beta_3 c}{\sqrt{1-\beta_3^2}} \\ -\frac{\frac{\beta_3}{c}}{\sqrt{1-\beta_3^2}} & \frac{1}{\sqrt{1-\beta_3^2}} \end{matrix}\right\| \quad \left(\beta_3=\frac{v_3}{c}\right),$$

т. е. ей соответствует также преобразование Лоренца со значением параметра $v=v_3$. Таким образом, формула (19) дает *правило сложения скоростей в специальном принципе относительности.* Если в формуле (19) положить $v_1=c$, то, как легко проверить, и для результирующей скорости v_3 мы будем иметь $v_3=c$, т. е. *действительно скорость c не меняется при наложении двух движений.*

При выводе формул (15) мы фиксировали определенным образом знаки коэффициентов линейного преобразования (11), считая коэффициенты a_{11} и a_{44} положительными. Можно заменить это требование другим, а именно: положительностью коэффициента a_{44} и положительностью определителя

$$a_{11}a_{44} - a_{14}a_{41}. \tag{20}$$

Нетрудно видеть, что отсюда, как следствие, получится и положительность коэффициента a_{11}, и наоборот. Действительно, определитель преобразования (17) равен $(+1)$, т. е. при $a_{11} > 0$ и определитель (20) положителен. Если бы мы взяли $a_{11} = -\alpha$ и $a_{44} = \alpha$, где $\alpha > 0$, то получилось бы преобразование с определителем (-1). Условие положительности коэффициента a_{11} равносильно тому, что при фиксированном x_1 и $x_4 \to \infty$ мы имеем $x'_4 \to \infty$. Можно сказать, что это соответствует неизменности в направлении отсчета времени. Таким образом, формулы дают не все линейные преобразования, удовлетворяющие условию (12), но лишь те, для которых определитель (20) положителен и которые не меняют направления отсчета времени.

Обратимся теперь к рассмотрению *общего преобразования Лоренца* для случая четырех переменных x_k $(k = 1, 2, 3, 4)$, причем должно быть выполнено условие:

$$x_1'^2 + x_2'^2 + x_3'^2 - c^2 x_4'^2 = x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2. \tag{21}$$

Будем рассматривать x_k $(k = 1, 2, 3)$ и x'_k $(k = 1, 2, 3)$ как декартовы координаты в двух различных трехмерных пространствах R и R'. Покажем, что, выбирая соответственным образом координатные оси в этих двух пространствах, мы можем привести общее преобразование Лоренца к тому частному случаю, который был нами рассмотрен выше. Обозначим через T общее преобразование Лоренца и через S частное преобразование Лоренца рассмотренного выше типа. Наше утверждение равносильно тому обстоятельству, что мы можем представить T в виде

$$T = VSU, \tag{22}$$

где U и V — вещественные ортогональные преобразования, соответствующие упомянутым выше преобразованиям координат в пространствах R и R'.

Введем, как и выше, четыре новые переменные:

$$y_1 = x_1;\ y_2 = x_2;\ y_3 = x_3;\ y_4 = icx_4,$$

и аналогичным образом:

$$y'_1 = x'_1;\ y'_2 = x'_2;\ y'_3 = x'_3;\ y'_4 = icx'_4.$$

Вместо условия (21) получим для новых переменных обычные условия ортогонального преобразования:

$$y_1'^2 + y_2'^2 + y_3'^2 + y_4'^2 = y_1^2 + y_2^2 + y_3^2 + y_4^2. \tag{23}$$

Искомое линейное преобразование будет иметь вид:

$$y'_k = \alpha_{k1}y_1 + \alpha_{k2}y_2 + \alpha_{k3}y_3 + \alpha_{k4}y_4 \quad (k = 1, 2, 3, 4). \tag{24}$$

Принимая во внимание, что y_4 и y'_4 должны быть чисто мнимыми, мы можем утверждать, что коэффициенты α_{k1}, α_{k2}, α_{k3}, при $k=1, 2, 3$, а также α_{44} должны быть вещественными, а коэффициенты α_{41}, α_{42}, α_{43} и α_{k4} при $k=1, 2, 3$ должны быть чисто мнимыми. Перемена координатных осей в пространстве R' равносильна вещественному ортогональному преобразованию над переменными y'_1, y'_2, y'_3. Рассмотрим коэффициенты:

$$\alpha_{14}=i\beta_{14}; \quad \alpha_{24}=i\beta_{24}; \quad \alpha_{34}=i\beta_{34}.$$

Три вещественных числа β_{14}, β_{24}, β_{34} определяют некоторый вектор, и если мы направление этого вектора примем за новую первую ось пространства R', то в результате соответствующего ортогонального преобразования коэффициенты α_{24} и α_{34} обратятся в нуль. Чтобы убедиться в этом, достаточно только заметить, что в силу формул (24) ортогональное преобразование над переменными y'_1, y'_2, y'_3 сводится к такому же преобразованию над β_{14}, β_{24}, β_{34}. Итак, будем считать, что это преобразование координат в пространстве R' уже совершено, так что мы имеем $\alpha_{24}=\alpha_{34}=0$. Условие (23) показывает, что коэффициенты преобразования (24) должны удовлетворять обычным условиям ортогонального преобразования. Принимая во внимание равенство нулю упомянутых выше коэффициентов, получаем, рассматривая вторую и третью строчки, следующие условия:

$$\alpha_{k1}^2+\alpha_{k2}^2+\alpha_{k3}^2=1 \qquad (k=2,\ 3),$$
$$\alpha_{21}\alpha_{31}+\alpha_{22}\alpha_{32}+\alpha_{23}\alpha_{33}=0,$$

где все входящие коэффициенты вещественны. В силу написанных условий два вектора с составляющими $(\alpha_{21}, \alpha_{22}, \alpha_{23})$ и $(\alpha_{31}, \alpha_{32}, \alpha_{33})$ будут единичными по длине и взаимно ортогональными по направлению. Если мы в пространстве R выберем эти два вектора за основные орты, направленные по осям X_2 и X_3, то две суммы

$$\alpha_{k1}y_1+\alpha_{k2}y_2+\alpha_{k3}y_3 \quad (k=2,\ 3),$$

выражающие скалярные произведения упомянутых выше двух векторов на переменный вектор (y_1, y_2, y_3), выразятся просто в виде y_2 и y_3, т. е. при таком выборе координатных осей мы будем иметь:

$$\alpha_{22}=\alpha_{33}=1; \quad \alpha_{21}=\alpha_{23}=\alpha_{31}=\alpha_{32}=0.$$

Таким образом, окончательно, при сделанном выборе осей в обоих пространствах матрица преобразования (24) будет иметь вид:

$$\left\| \begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{matrix} \right\|. \qquad (25)$$

Эта матрица получилась в результате умножения первоначальной матрицы на два ортогональных преобразования, которые касались только первых трех переменных, но которые можно, конечно, рассматривать и как ортогональные преобразования с четырьмя переменными, причем четвертая переменная остается без изменения. Принимая во внимание, что произведение двух ортогональных преобразований также должно быть ортогональным, мы можем утверждать, что элементы матрицы (25) также должны удовлетворять условию ортогональности. Написав условие ортогональности первой строки со второй и третьей, получим

$$\alpha_{12}=\alpha_{13}=0,$$

и точно так же условие ортогональности четвертой строки со второй и третьей даст нам

$$\alpha_{43}=\alpha_{42}=0.$$

В результате приходим к следующей матрице:

$$\begin{Vmatrix} \alpha_{11} & 0 & 0 & \alpha_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \alpha_{41} & 0 & 0 & \alpha_{44} \end{Vmatrix},$$

т. е. в данном случае мы имеем линейное преобразование

$$y_1'=\alpha_{11}y_1+\alpha_{14}y_4,$$
$$y_4'=\alpha_{41}y_1+\alpha_{44}y_4,$$

которое должно удовлетворять условию

$$y_1'^2+y_4'^2=y_1^2+y_4^2.$$

Именно таким преобразованием мы и занимались выше, и оно нас и привело к специальным преобразованиям Лоренца вида (15), и, таким образом, формулу (22) можно считать установленной. Заметим только, что правило выбора знаков при определении преобразования S будет тем же самым, что и выше, если мы потребуем, чтобы общее преобразование Лоренца T не меняло направления отсчета времени и имело определитель больше нуля. Ортогональные преобразования U и V мы всегда можем считать вращениями трехмерного пространства, так что и их определитель будет больше нуля, причем они вовсе не затрагивают четвертой переменной. Мы придем, таким образом, к необходимости того, чтобы и у преобразования S определитель был больше нуля и чтобы это преобразование не меняло отсчета времени, т. е. при сделанном предположении относительно общего преобразования T мы для частного преобразования придем как раз к тому условию, при котором формулы были выведены. Общие преобразования Лоренца, удовлетворяющие поставленным выше двум условиям, называются обычно *положительными преобразованиями Лоренца*. Из предыдущих рассуждений следует, что соответствующие им матрицы получаются по формуле (22), где S — специальное преобразование Лоренца вида (15), а U и V — матрицы вращения трехмерного пространства. Можно показать, что положительные преобразования Лоренца образуют группу, как и преобразования (15).

Предыдущие рассуждения показывают, что матрица наиболее общего преобразования Лоренца, определяемого лишь условием (21), может быть представлена по формуле (22), где U и V — вращения и S — общее преобразование Лоренца с двумя переменными. Если это положительное преобразование, то из формул (15) непосредственно следует, что $D(S)=1$, и определитель всякого положительного преобразования Лоренца будет также равен единице, поскольку определители U и V равны единице, причем матрицы U, S и V мы рассматриваем как матрицы четвертого порядка. В общем случае преобразования Лоренца второго порядка, как нетрудно показать, определитель может равняться (± 1) и, следовательно, общее преобразование Лоренца также будет иметь определитель (± 1).

55. Перестановки. До сих пор мы рассматривали примеры групп, элементами которых являлись линейные преобразования. Понятие группы не связано обязательно операциями линейного преобразования и может быть построено и для операций другого рода. Мы переходим сейчас к рассмотрению операций, с которыми уже встречались раньше [**2**], а именно переходим к рассмотрению *перестановок*. Выясним сначала некоторые основные факты и понятия, относящиеся к перестановкам.

Пусть имеется n некоторых объектов, которые мы, как и в [**2**], пронумеруем, т. е. будем просто считать, что эти объекты суть целые числа $1, 2, \ldots, n$. Из этих чисел мы, как известно, можем составить $n!$ перестановок. Возьмем одну из этих перестановок

$$p_1, p_2, \ldots, p_n. \tag{26}$$

Числа p_k в своей совокупности дают все целые числа от 1 до n, причем в перестановке (26) они расставлены в определенном порядке. Сравним перестановку (26) с основной перестановкой $1, 2, \ldots, n$:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix} (P). \tag{27}$$

Переход от основной перестановки к перестановке (26) совершается путем замены 1 на p_1, 2 на p_2 и так далее. Обозначим эту операцию одной буквой P и будем в дальнейшем называть ее перестановкой. Определим теперь понятие об *обратной перестановке* P^{-1}.

Это будет такая операция, которая переводит (26) в основную последовательность, т. е. заменяет p_1 на 1, p_2 на 2 и так далее. Разъясним это на частном примере: возьмем $n=5$ и рассмотрим перестановку

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} (P).$$

Обратная перестановка будет иметь вид:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 5 & 3 \end{pmatrix} (P^{-1}).$$

Нетрудно видеть, что

$$(P^{-1})^{-1} = P. \tag{28}$$

Введем теперь понятие о *произведении перестановок*. Пусть P_1 и P_2 — какие-нибудь две перестановки. Назовем произведением перестановок P_2P_1 такую перестановку, которая получается в результате применения сначала перестановки P_1, а затем перестановки P_2. Например, если мы имеем две перестановки

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \end{pmatrix} (P_2) \text{ и } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix} (P_1),$$

то их произведение P_2P_1 будет давать перестановку вида:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 1 & 3 \end{pmatrix} (P_2P_1).$$

Непосредственно очевидно, что обратная перестановка P^{-1} вполне определяется из условия

$$P^{-1}P = PP^{-1} = I, \tag{29}$$

где через I мы обозначили единичную перестановку, при которой каждый элемент заменяется сам собой.

Последовательно применяя несколько перестановок, мы можем составить произведение нескольких перестановок $P_3P_2P_1$. Нетрудно видеть, что такое произведение удовлетворяет сочетательному закону, т. е.

$$P_3(P_2P_1) = (P_3P_2)P_1. \tag{30}$$

Действительно, произведя перестановку P_1, мы можем затем последовательно производить P_2 и P_3, или вместо этого мы можем заменить это последовательное применение P_2 и P_3 применением одной перестановки (P_3P_2), которая равносильна последовательному применению P_2 и P_3. Заметим, наконец, что единичная перестановка удовлетворяет, очевидно, следующим условиям:

$$IP = PI = P, \tag{31}$$

где P — любая перестановка. Произведение перестановок не будет, вообще говоря, удовлетворять переместительному закону, т. е. произведения P_2P_1 и P_1P_2 будут, вообще говоря, различными перестановками. Предлагаем проверить это на предыдущем примере.

Мы установили, таким образом, основные понятия произведения, обратной перестановки и единичной перестановки, совершенно так же, как это сделали раньше для линейных преобразований (матриц). Можно теперь продолжить эту аналогию и дальше и установить понятие о *группе*, а именно: совокупность перестановок образует группу, если выполнены следующие два условия: во-первых, если некоторая перестановка принадлежит нашей совокупности, то и обратная перестановка также принадлежит нашей совокупности, и, во-вторых, произведение двух перестановок, принадлежащих нашей совокупности (при любом порядке сомножителей), также принадлежит нашей совокупности. Как и в случае линейных преобразований, единичная перестановка должна обязательно принадлежать группе.

Совокупность всех $n!$ перестановок образует, очевидно, группу. Перейдем теперь к установлению еще другой группы, которая составляет лишь часть предыдущей. Заметим для этого, что всякая перестановка может быть выполнена при помощи нескольких транспозиций [2], причем для заданной перестановки число транспозиций может быть различным, но, как мы доказали выше, оно будет всегда для заданной перестановки или четным, или нечетным. Перестановки, состоящие из четного числа транспозиций, образуют сами по себе группу. Группа, образованная всеми перестановками, называется обычно *симметрической группой*, а группа, состоящая из четных перестановок, т. е. из перестановок, сводящихся к четному числу транспозиций, называется *знакопеременной*.

Рассмотрим теперь перестановки особого типа. Пусть $l_1, l_2, \ldots, l_m$ — какие-нибудь m различных чисел из первых n чисел. Положим, что наша перестановка состоит в замене l_1 на l_2, l_2 на l_3 и так далее, l_{m-1} на l_m и, наконец, l_m на l_1. Перестановку такого типа назовем *циклом* и обозначим ее символом $(l_1, l_2, \ldots, l_m)$. Совершая круговую перестановку чисел внутри скобки, мы будем получать циклы:

$$(l_2, l_3, \ldots, l_m, l_1), \quad (l_3, l_4, \ldots, l_m, l_1, l_2), \ldots,$$

которые, очевидно, дают ту же перестановку, что и $(l_1, l_2, \ldots, l_m)$. Если $m=1$, т. е. если мы имеем цикл (l_1), то такой цикл равносилен, очевидно, единичной перестановке, и его не имеет смысла рассматривать. Цикл из двух чисел (l_1, l_2) равносилен, очевидно, транспозиции элементов l_1 и l_2.

Если имеются два цикла без общих элементов, то их произведение не зависит от порядка сомножителей.

Пусть, например, $n=5$, и мы имеем произведение двух циклов без общих элементов (1, 3) (2, 4, 5) и (2, 4, 5) (1, 3).

Оба эти произведения дают, очевидно, одну и ту же перестановку

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

Мы можем всякую перестановку P представить в виде *произведения циклов*, не имеющих общих элементов. Чтобы сделать это, возьмем элемент 1 и примем его за первый элеменг в цикле. За второй элемент цикла возьмем тот элемент, который получается из 1 при помощи P. Пусть это будет l_2. За третий элемент возьмем тот, который получается из l_2 при помощи P и так далее, пока, наконец, не дойдем до такого элемента, который переходит в 1 при помощи P. Это и будет последний элемент составленного цикла. Нетрудно видеть, что этот цикл не может содержать одинаковых элементов. Составленный таким образом цикл не исчерпает, вообще говоря, все n элементов. Из оставшихся элементов возьмем какой-нибудь за первый элемент нового цикла и, как и выше, составим второй цикл и т. д.

В качестве примера возьмем перестановку при $n=6$:

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

Применяя предыдущий прием, можем представить ее в виде произведения циклов

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 4 & 1 & 2 & 5 \end{pmatrix} = (1,\ 3,\ 4,)\ (2,\ 6,\ 5),$$

причем порядок сомножителей справа не играет роли.

Нетрудно видеть, что произведение двух транспозиций можно представить в виде произведения трехчленных циклов. Если двухчленные циклы не имеют общих элементов, то мы имеем, как легко проверить:

$$(l_3,\ l_4)(l_1,\ l_2) = (l_1,\ l_3,\ l_4)(l_1,\ l_2,\ l_4),$$

а при наличии общих элементов:

$$(l_1,\ l_3)(l_1,\ l_2) = (l_1,\ l_2,\ l_3).$$

Таким образом, всякую перестановку из знакопеременной группы можно представить в виде произведения трехчленных циклов.

Отметим еще, что при перестановке можно в первой строке вместо натурального ряда чисел писать эти числа в любом порядке. Важно лишь, чтобы под каждым числом стояло то число, в которое оно переходит в результате взятой перестановки. Приведем для примера две записи одной и той же перестановки:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 5 & 4 & 2 \\ 5 & 3 & 4 & 1 & 2 \end{pmatrix}.$$

Пусть имеется некоторая перестановка

$$P=\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{pmatrix}.$$

Обратную перестановку мы можем, очевидно, записать в виде:

$$P^{-1}=\begin{pmatrix} b_1 & b_2 & \dots & b_n \\ a_1 & a_2 & \dots & a_n \end{pmatrix}.$$

Пусть имеются две перестановки, причем вторую мы запишем двояко:

$$P=\begin{pmatrix} 1 & 2 & \dots & n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}; \quad Q=\begin{pmatrix} 1 & 2 & \dots & n \\ d_1 & d_2 & \dots & d_n \end{pmatrix}=\begin{pmatrix} c_1 & c_2 & \dots & c_n \\ f_1 & f_2 & \dots & f_n \end{pmatrix}.$$

Мы имеем:

$$PQ^{-1}=\begin{pmatrix} 1 & 2 & \dots & n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}\begin{pmatrix} d_1 & d_2 & \dots & d_n \\ 1 & 2 & \dots & n \end{pmatrix}=\begin{pmatrix} d_1 & d_2 & \dots & d_n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}$$

и, следовательно,

$$QPQ^{-1}=\begin{pmatrix} c_1 & c_2 & \dots & c_n \\ f_1 & f_2 & \dots & f_n \end{pmatrix}\begin{pmatrix} d_1 & d_2 & \dots & d_n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}=\begin{pmatrix} d_1 & d_2 & \dots & d_n \\ f_1 & f_2 & \dots & f_n \end{pmatrix}.$$

Из написанного вытекает следующее правило: чтобы получить перестановку QPQ^{-1}, надо в обеих строчках перестановки

$$P=\begin{pmatrix} 1 & 2 & \dots & n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}$$

совершить перестановку Q.

56. Абстрактные группы. При определении группы мы можем совершенно отвлечься от конкретного значения тех операций, которые в своей совокупности образуют группу и которые в предыдущем были линейными преобразованиями или перестановками. Мы придем, таким образом, к понятию *абстрактной группы*.

Абстрактная группа есть совокупность некоторых символов, — таких, что для этих символов определено умножение в том смысле, что дается определенное правило, согласно которому из двух элементов P и Q совокупности (разных или одинаковых) получается третий элемент, также принадлежащий совокупности, который называется их произведением и обозначается через QP. При этом должны быть выполнены следующие три условия:

1. *Перемножение должно подчиняться сочетательному закону*, т. е. $(RQ)P=R(QP)$, откуда вообще будет следовать, что мы можем в любом произведении, не меняя, конечно, порядка сомножителей, соединять любые сомножители в одну группу.

2. *В нашей совокупности должен существовать один и только один такой элемент E, который, будучи помножен на любой*

другой элемент с той или с иной стороны, воспроизводит этот же самый элемент, т. е.

$$EP = PE = P. \tag{32}$$

Будем называть этот элемент *единичным элементом.*

3. *Для любого элемента нашей совокупности P существует в нашей же совокупности единственный другой элемент Q, который удовлетворяет условию*

$$QP = PQ = E \qquad (Q = P^{-1}). \tag{33}$$

Из (32) при $P = E$ вытекает $EE = E$, т. е., в силу определения обратного элемента, элементом, обратным E, будет сам элемент E ($E^{-1} = E$).

Можно поставить эти условия, определяющие абстрактную группу, в более узкой форме, причем из этих более узких требований остальные уже будут вытекать в качестве необходимых формальных следствий, но мы на этом останавливаться не будем. Вообще ограничимся лишь простейшими и основными фактами, связанными с понятием абстрактной группы. Более подробное рассмотрение теории групп дает материал, который сам по себе может заполнить целую книгу. Нашей целью является лишь сообщить читателю основные понятия и этим облегчить чтение физической литературы, в которой зачастую применяется понятие группы и где часто пользуются основными свойствами групп. В дальнейшем вместо E мы будем писать иногда I. Элемент Q, определяемый соотношениями (33), называется *обратным* P и обозначается P^{-1}. Имеет место, очевидно, соотношение (28), ибо из (33) следует и то, что P обратно Q.

Установив понятие абстрактной группы, перейдем теперь к выяснению некоторых новых понятий, а также к доказательству некоторых свойств абстрактных групп. Заметим прежде всего, что число элементов в группе, как это мы видели выше, может быть как конечным, так и бесконечным. Рассмотрим некоторое произведение элементов группы

$$RQP.$$

Это будет также некоторый элемент группы. Обратный элемент будет получаться совершенно так же, как и в группе линейных преобразований, а именно он будет:

$$(RQP)^{-1} = P^{-1}Q^{-1}R^{-1}.$$

В этом нетрудно убедиться, совершая перемножение и пользуясь сочетательным законом. Пусть P — некоторый элемент группы. Его целые положительные степени

$$P^0 = I,\ P^1,\ P^2\ \ldots$$

также будут элементами группы. Если существует такое целое положительное число m, что $P^m = I$, то говорят, что элемент будет *конечного порядка,* причем *порядком элемента* будет наименьшее значение целого положительного числа m, при котором $P^m = I$. При этом среди элементов

$$I,\ P,\ P^2,\ \ldots,\ P^{m-1}$$

уже не может быть одинаковых. Действительно, из условия $P^k = P^l\,(k < l)$ непосредственно вытекает $P^{l-k} = I$. Для конечной группы все элементы будут, очевидно, конечного порядка.

Обозначим элементы нашей группы через P_α. Если группа конечна, то можно считать, что значок α пробегает конечное число целых положительных значений. Если группа бесконечна, то он может пробегать все целые значения [52], может меняться непрерывно и даже он может быть равносилен нескольким значкам, которые непрерывно изменяются. Пусть U — некоторый фиксированный элемент нашей группы. Составим всевозможные произведения UP_α. Нетрудно видеть, что при соответствующем изменении значка α написанное произведение даст нам опять все элементы группы и притом по одному разу.

Действительно, из равенства

$$UP_{\alpha_1} = UP_{\alpha_2}$$

умножением на U^{-1} слева получаем $P_{\alpha_1} = P_{\alpha_2}$, т. е. при разных α и произведение UP_α различно. Покажем теперь, что это произведение будет обращаться в любой элемент нашей группы. Действительно, равенство $UP_\alpha = P_{\alpha_0}$ равносильно $P_\alpha = U^{-1}P_{\alpha_0}$, т. е. произведение UP_α даст нам элемент P_{α_0}, когда сомножитель P_α будет равняться элементу $U^{-1}P_{\alpha_0}$ нашей группы. Тот же самый результат мы получили бы, если бы приписали фиксированный элемент U не слева, а справа. Итак, мы получаем следующий результат: *если P_α пробегает все элементы группы и U — некоторый фиксированный элемент группы, то произведение UP_α (или $P_\alpha U$) также пробегает все элементы группы и притом по одному разу.*

Рассмотрим частный пример группы. Положим, что группа состоит из шести элементов (группа шестого порядка), и обозначим эти элементы следующим образом:

$$E,\ A,\ B,\ C,\ D,\ F.$$

Закон умножения определим при помощи следующей таблицы:

	E	A	B	C	D	F
E	E	A	B	C	D	F
A	A	E	D	F	B	C
B	B	F	E	D	C	A
C	C	D	F	E	A	B
D	D	C	A	B	F	E
F	F	B	C	A	E	D

(34)

Этой таблицей надо пользоваться для определения умножения следующим образом. Если мы хотим, например, составить произведение DB, то должны в первой строке найти B, в первом столбце D и на пересечении соответствующих строк и столбцов найдем элемент A, который и будет являться произведением DB. Нетрудно проверить, что при этом будут удовлетворены все те условия, которые мы упоминали при определении абстрактной группы, причем элемент E будет играть роль единичного элемента.

В предыдущих номерах мы имели примеры конкретного осуществления абстрактного понятия группы. В одном случае роль элемента играло линейное преобразование (его матрица) и перемножение двух элементов сводилось к последовательному применению двух линейных преобразований, т. е. к перемножению соответствующих этим преобразованиям матриц. В другом случае роль элемента играла перестановка, и перемножение двух элементов сводилось к последовательному выполнению двух перестановок. Приведем еще пример конкретного осуществления группы.

Пусть элементами являются всевозможные комплексные числа и перемножение двух элементов сводится к сложению соответствующих комплексных чисел. В данном случае роль единичного элемента играет число нуль, и элементом, обратным комплексному числу α, является число $(-\alpha)$. Вместо комплексных чисел мы могли бы взять за элементы всевозможные векторы $\boldsymbol{x}(x_1, x_2, \ldots, x_n)$ комплексного n-мерного пространства R_n и определить умножение элементов, как сложение соответствующих векторов. При этом роль единичного элемента играет нулевой вектор. Иначе можно сказать, что элементами группы являются векторы из R_n, а групповым действием — сложение векторов. Отметим, что в последних двух примерах результаты перемножения двух элементов группы не зависят от порядка сомножителей, т. е., как говорят, любые два элемента группы коммутируют. Такие группы называются *абелевыми группами* [ср. **45**]. Простейшим примером абелевой группы является так называемая *циклическая группа*, которая состоит из единичного элемента E и степеней некоторого элемента P. Если m — наименьшее целое положительное число, при котором $P^m = E$, то циклическая группа содержит m элементов: $E, P, P^2, \ldots, P^{m-1}$. Если такого целого положительного m нет, то циклическая группа бесконечна: $E, P, P^2, \ldots$

57. Подгруппа. Пусть имеется некоторая группа G и положим, что совокупность H элементов, содержащая лишь часть элементов группы G, также образует группу, при сохранении прежнего определения умножения. В этом случае группа H называется *подгруппой* группы G. Нетрудно видеть, что совокупность, состоящая из одного единичного элемента группы G, есть всегда подгруппа. Это — тривиальная подгруппа. В дальнейшем, говоря о подгруппе мы будем разуметь не тривиальную подгруппу.

Обозначим через H_α элементы подгруппы H, и пусть G_1 — некоторый элемент общей группы G, причем G_1 не принадлежит H. Произведения $G_1 H_\alpha$, как мы видели выше, дадут нам различные элементы группы G, причем эти элементы не принадлежат H. Действительно, в противном случае мы имели бы при некоторых значениях α_1 и α_2 значка α: $G_1 H_{\alpha_1} = H_{\alpha_2}$, откуда $G_1 = H_{\alpha_2} H_{\alpha_1}^{-1}$, т. е. G_1 должно принадлежать H, что противоречит поставленному условию. Положим теперь, что G_1 и G_2 суть два различных элемента общей группы G, не принадлежащих подгруппе H. Покажем, что совокупности элементов $G_1 H_\alpha$ и $G_2 H_\alpha$ или вовсе не имеют общих элементов, или совпадают, т. е. состоят из одних и тех же элементов. Действительно, если при некоторых значениях значка α мы имеем $G_2 H_{\alpha_2} = G_1 H_{\alpha_1}$, то отсюда следует $G_2 = G_1 H_{\alpha_1} H_{\alpha_2}^{-1} = G_1 H_{\alpha_3}$, т. е. G_2 принадлежит совокупности элементов $G_1 H_\alpha$, и точно так же G_1 принадлежит совокупности элементов $G_2 H_\alpha$. Отсюда следует, что произведения $G_1 H_\alpha$ и $G_2 H_\alpha$ определяют одну и ту же совокупность элементов.

Возьмем все элементы H_α подгруппы H. Они не исчерпывают всех элементов группы G. Рассмотрим некоторый элемент G_1, не принадлежащий H, и составим всевозможные произведения $G_1 H_\alpha$, которые, как мы видели выше, все различны между собою и отличны от H_α.

Может случиться, что элементы H_α и $G_1 H_\alpha$ также не исчерпывают всей группы. Возьмем некоторый элемент группы G_2, не принадлежащий H_α и $G_1 H_\alpha$, и составим всевозможные произведения $G_2 H_\alpha$. Как мы видели выше, элементы $G_2 H_\alpha$ будут все различны между собой и будут отличны от элементов H_α и $G_1 H_\alpha$. Если элементы H_α, $G_1 H_\alpha$ и $G_2 H_\alpha$ не исчерпывают всех элементов группы G, то возьмем некоторый элемент G_3, не принадлежащий указанным выше трем совокупностям элементов, и составим произведения $G_3 H_\alpha$. Мы получим таким образом новые элементы группы и т. д. Положим, что путем конечного числа таких операций мы исчерпаем все элементы группы G. Пусть для этого потребуется $(m-1)$ элементов G_k. При этом все элементы группы G будут представлены в следующем виде:

$$H_\alpha,\ G_1 H_\alpha,\ G_2 H_\alpha,\ \ldots,\ G_{m-1} H_\alpha, \tag{35}$$

где значок α пробегает значения, соответствующие подгруппе H. Если положим $G'_k = G_k H_{\alpha_0}$, где α_0 как-нибудь фиксировано, то совокупность элементов $G'_k H_\alpha$, как показано выше, будет совпадать с совокупностью $G_k H_\alpha$. Иначе говоря, в каждой совокупности $G_k H_\alpha$ $(G_0 = I)$ любой элемент совокупности может играть роль представителя G_k. Отсюда непосредственно следует, что при заданной подгруппе H_α разбиение элементов группы G на совокупности вида (35) вполне определено. Совокупности $G_k H_\alpha$ называются ***совокупностями, сопряженными относительно подгруппы*** H_α.

В рассматриваемом случае (35) подгруппа H называется *подгруппой конечного индекса*, а именно *подгруппой индекса m*. Если группа G конечна, то индекс подгруппы H будет равен, очевидно, частному от деления порядка всей группы G на порядок подгруппы H, причем *порядком конечной группы называется число содержащихся в ней элементов*. Заметим, что из совокупностей (35) только первая совокупность образует подгруппу. Каждая из остальных совокупностей G_kH_α не содержит единичного элемента, а потому не может образовать подгруппы.

При построении схемы (35) мы умножали элементы H_α подгруппы H на элементы G_k группы G слева. Можно было бы производить это умножение и справа. Вводя вместо G_k другое обозначение G'_k, мы пришли бы таким же образом к представлению элементов группы G вместо схемы (35) в следующем виде:

$$H_\alpha,\ H_\alpha G'_1,\ H_\alpha G'_2,\ \ldots,\ H_\alpha G'_{m-1}, \tag{36}$$

при этом индекс подгруппы m не меняется. Совокупности элементов G_kH_α называются иногда сопряженными совокупностями слева, а $H_\alpha G'_k$ — сопряженными совокупностями справа.

Заметим прежде всего, что если α пробегает все значения, соответствующие подгруппе H, то элементы H_α^{-1} дают все элементы H. Это непосредственно следует из того, что элемент, обратный некоторому элементу, принадлежащему H, также принадлежит H. Переходим теперь к доказательству совпадения индексов у совокупностей, сопряженных справа и сопряженных слева. Возьмем какие-нибудь две различные совокупности из (35), G_pH_α и G_qH_α $(p \neq q)$, причем для первой из совокупностей (35) мы можем считать, например, $G_p = E$. Возьмем обратные элементы:

$$(G_pH_\alpha)^{-1} = H_\alpha^{-1}G_p^{-1} \text{ и } (G_qH_\alpha)^{-1} = H_\alpha^{-1}G_q^{-1}.$$

Принимая во внимание сделанные выше замечания, мы можем переписать эти совокупности элементов в виде $H_\alpha G_p^{-1}$ и $H_\alpha G_q^{-1}$. Нетрудно видеть, что они не имеют общих элементов. Действительно, если бы мы имели:

$$H_{\alpha_1}G_p^{-1} = H_{\alpha_2}G_q^{-1},$$

то отсюда вытекало бы

$$G_p^{-1}G_q = H_{\alpha_1}^{-1}H_{\alpha_2} = H_{\alpha_3}, \text{ или } G_q = G_pH_{\alpha_3},$$

и оказалось бы, что G_q принадлежит совокупности $G_pH_{\alpha_3}$, чего не может быть. Таким образом, оказывается, что совокупности

$$H_\alpha,\ H_\alpha G_1^{-1},\ H_\alpha G_2^{-1},\ \ldots,\ H_\alpha G_{m-1}^{-1}$$

будут сопряженными совокупностями справа, так что в (36) можно просто брать $G'_s = G_s^{-1}$.

Рассмотрим некоторые примеры подгрупп. Пусть G — совокупность вещественных ортогональных преобразований с тремя переменными и H — совокупность вещественных ортогональных преобразований с тремя переменными и с определителем $(+1)$. Всякое вещественное ортогональное преобразование будет или вращением, т. е. принадлежащим H, или произведением вращения на симметрию относительно начала, которая выражается формулами:

$$x' = -x; \quad y' = -y; \quad z' = -z. \quad (S) \tag{37}$$

В данном случае группа G может быть представлена следующей схемой:

$$H_\alpha, \ SH_\alpha \tag{38}$$

или

$$H_\alpha, \ H_\alpha S, \tag{39}$$

где H_α обозначает совокупность всех элементов группы H. В данном случае H_α будет подгруппой индекса 2.

Пусть G — симметрическая группа перестановок из n элементов, H — знакопеременная группа, состоящая из четных перестановок. Пусть далее S — какая-нибудь определенная нечетная перестановка, например перестановка, состоящая из одного цикла (1, 2), т. е. сводящаяся к транспозиции элементов 1 и 2. Мы можем, очевидно, и здесь представить G по схеме (38) или (39). В обоих случаях умножение слева приводит к тому же результату, что и умножение справа.

В данном случае знакопеременная группа будет подгруппой симметрической группы с индексом два.

Рассмотрим еще конечную группу правильного октаэдра, о которой мы говорили выше. Пусть A — некоторая вершина октаэдра и l — ось, проходящая через эту вершину. Пусть S_0, S_1, S_2, S_3 — вращения на угол 0, $\frac{\pi}{2}$, π и $\frac{3\pi}{2}$ вокруг этой оси. Эти вращения образуют подгруппу всей полной группы вращения октаэдра. Обозначим через T_k $(k = 1, 2, 3, 4, 5)$ вращения, переводящие вершину A в остальные пять вершин октаэдра. Мы можем представить полную группу октаэдра по следующей схеме:

$$S_\alpha, \ T_1 S_\alpha, \ T_2 S_\alpha, \ T_3 S_\alpha, \ T_4 S_\alpha, \ T_5 S_\alpha,$$

т. е. подгруппа S_α будет подгруппой индекса шесть.

Пусть G_s, G_s^{-1} $(s = 1, 2, \ldots, k)$ — какие-либо элементы группы G. Рассмотрим множество всех тех элементов группы G, которые можно представить в виде произведения элементов G_s, G_s^{-1} $(s = 1, 2, \ldots, k)$.

Это множество элементов образует, очевидно, группу, которая является подгруппой для G или совпадает с G.

Говорят, что эта подгруппа порождается данной совокупностью элементов G_s, G_s^{-1} $(s = 1, 2, \ldots, k)$.

58. Классы и нормальный делитель. Пусть U и V — некоторые элементы группы. Элемент $W = VUV^{-1}$ называется *сопряженным элементу* U. Нетрудно видеть, что и наоборот U будет сопряженным с W. Действительно, $U = V^{-1}WV$. Два элемента U_1 и U_2, сопряженные с одним и тем же третьим W:

$$U_1 = V_1 W V_1^{-1}; \quad U_2 = V_2 W V_2^{-1},$$

будут сопряженными и между собой

$$U_2 = V_2 V_1^{-1} U_1 (V_2 V_1^{-1})^{-1}.$$

Совокупность всех взаимно сопряженных элементов группы образует то, что называется *классом группы.* Класс вполне определяется одним своим элементом U. Действительно, задав U, мы получим весь класс по формуле $G_a U G_a^{-1}$, где G_a пробегает все элементы группы. Таким образом мы можем разбить всю группу на классы. Принимая во внимание основное свойство единичного элемента, формулированное нами в [**56**], мы имеем:

$$G_a I G_a^{-1} = I,$$

т. е. единичный элемент сам по себе составляет класс.

Если элемент U будет порядка m, т. е. если m есть наименьшее целое положительное число, при котором $U^m = I$, то и всякий сопряженный элемент $G_a U G_a^{-1}$ имеет тот же порядок m, что непосредственно следует из равенства:

$$(G_a U G_a^{-1})^m = G_a U^m G_a^{-1} = I.$$

Иначе говоря, все элементы одного и того же класса имеют одинаковый порядок.

Заметим, что когда G_a пробегает все элементы группы G, то произведение $G_a U G_a^{-1}$ может давать элементы класса и по нескольку раз. Так, например, если $U = I$, то, как мы видели, указанное произведение всегда дает I.

В качестве примера возьмем опять группу вращений октаэдра. Пусть U есть вращение на угол $\frac{\pi}{2}$ вокруг некоторой оси $A_p A_q$ октаэдра. Если вращение T_k, принадлежащее нашей группе октаэдра, преобразует ось l в ось l_1, причем вершина A_p переходит в A_r и вершина A_q в A_s, то элемент группы $T_k U T_k^{-1}$ будет давать вращение на угол $\frac{\pi}{2}$ вокруг оси $A_r A_s$. Если, например, T_k преобразует A_p в A_q, то упомянутое произведение будет давать вращение на угол $\frac{\pi}{2}$ вокруг оси $A_q A_p$ или, иначе говоря, вращение на угол $\frac{3\pi}{2}$ вокруг оси $A_p A_q$. Если T_k преобразует ось $A_p A_q$ в саму себя, т. е. есть поворот вокруг этой оси, то произведение $T_k U T_k^{-1}$ совпадает с U. Таким образом, в данном случае класс элементов, сопряженных с U, будет представлять собою совокупность вращений на угол $\frac{\pi}{2}$ вокруг осей октаэдра.

Совершенно так же, если мы возьмем группу вращений трехмерного пространства вокруг начала, то, как мы знаем, всякий элемент U этой группы будет представлять собою вращение на некоторый угол φ вокруг некоторой оси. В данном случае класс элементов, сопряженных с U, будет совокупностью вращений на угол φ вокруг всевозможных осей, проходящих через начало.

В тесной связи с понятием класса стоит и другое важное понятие, а именно: понятие о *нормальном делителе*, к которому мы сейчас и переходим. Пусть G — некоторая группа и H — ее подгруппа. Пусть G_1 — некоторый фиксированный элемент группы G. Рассмотрим совокупность элементов этой группы, представляемых произведением

$$G_1 H_a G_1^{-1}, \tag{40}$$

где через H_a мы обозначили переменный элемент подгруппы H, т. е., иначе говоря, H_a пробегает все элементы подгруппы H. Нетрудно видеть, что произведения (40) образуют также подгруппу. Действительно, если взять, например, произведение двух элементов, принадлежащих совокупности (40), то оно также будет принадлежать этой совокупности:

$$(G_1 H_{a_2} G_1^{-1})(G_1 H_{a_1} G_1^{-1}) = G_1 H_{a_2} H_{a_1} G_1^{-1} = G_1 H_{a_3} G_1^{-1},$$

и аналогично выполняются остальные условия для образования группы.

Подгруппа (40) называется *подобной подгруппе* H, и если G_1 принадлежит подгруппе H, то подгруппа (40) также состоит из элементов, принадлежащих H, и, как нетрудно видеть, просто совпадает с H.

Всякий элемент H_{a_0} подгруппы H может быть получен в этом случае по формуле (40), если мы возьмем

$$H_a = G_1^{-1} H_{a_0} G_1.$$

Если элемент G_1 не принадлежит подгруппе H, то подгруппа (40) может быть и отличной от подгруппы H.

Подгруппа H называется *нормальным делителем* полной группы G, если при любом выборе элемента G_1 из полной группы G подгруппа (40) совпадает с подгруппой H. Мы дальше приведем примеры нормальных делителей группы, а сейчас перейдем к выяснению некоторых новых понятий, связанных с понятием нормального делителя.

Положим, что подгруппа H есть нормальный делитель полной группы G. Для простоты письма будем считать, что эта подгруппа имеет конечный индекс m. В этом случае все элементы группы G могут быть представлены по схеме

$$H_a,\ G_1 H_a,\ G_2 H_a,\ \ldots,\ G_{m-1} H_a, \tag{41}$$

где H_a, как всегда, — переменный элемент подгруппы H. Раз H есть нормальный делитель, то совокупность элементов $G_k H_a G_k^{-1}$ совпадает с совокупностью элементов H_a, т. е. совокупность элементов $G_k H_a$ совпадает с совокупностью элементов $H_a G_k$.

Таким образом, если *H есть нормальный делитель*, то разбиение элементов полной группы на сопряженные совокупности по схеме (41) совпадает с разбиением элементов на сопряженные совокупности по схеме

$$H_a,\ H_a G_1,\ H_a G_2,\ \ldots,\ H_a G_{m-1}. \tag{42}$$

Иначе говоря, в этом случае *сопряженные совокупности справа совпадают с сопряженными совокупностями слева.*

Если H_{a_0} есть некоторый элемент нормального делителя, то при любом G_0 из G элемент $G_0 H_{a_0} G_0^{-1}$ также принадлежит нормальному делителю, т. е. если некоторый элемент принадлежит нормальному делителю, то и весь класс, в который этот элемент входит в основной группе, также принадлежит нормальному делителю. Нетрудно показать и наоборот, что если некоторая подгруппа обладает тем свойством, что, содержа некоторый элемент, она содержит и весь класс, к которому этот элемент принадлежит в основной группе, то такая подгруппа есть нормальный делитель.

Обратимся теперь к рассмотрению сопряженных совокупностей в схеме (41) или (42), где элементы H_a образуют нормальный делитель H. Рассмотрим произведения $G_l H_a G_k H_{a'}$ элементов некоторой сопряженной совокупности $G_l H_a$ на элементы сопряженной совокупности $G_k H_{a'}$.

Мы можем совокупность этих произведений написать в виде:

$$G_l (H_a G_k) H_{a'}, \quad G_l G_k H_a H_{a'}.$$

Элементы H_a и $H_{a'}$ содержатся в нормальном делителе H, и то же самое можно сказать и об их произведении. Таким образом, предыдущие произведения мы можем написать в виде:

$$G_l G_k H_a.$$

Все элементы этого вида заключаются в одной и той же сопряженной совокупности (41), а именно в той сопряженной совокупности, к которой принадлежит элемент $G_l G_k$. Нетрудно также показать, что мы получим таким образом все элементы этой сопряженной совокупности. Короче говоря, если подгруппа есть нормальный делитель, то перемножение одной сопряженной совокупности на другую дает также некоторую сопряженную совокупность. Будем рассматривать каждую из сопряженных совокупностей как некий новый элемент, причем первую из сопряженных совокупностей в схеме (41), (38) будем считать единичным элементом. Предыдущий результат об умножении сопряженных совокупностей даст нам правила умножения этих новых, введенных нами элементов, причем, как нетрудно проверить, что мы предоставляем сделать читателю, это правило умножения удовлетворяет всем условиям, которые требуются для образования группы, т. е. введенные нами новые элементы при указанном правиле перемножения сами образуют группу, в которой первая из сопряженных совокупностей схемы играет роль единичного элемента. Эта новая группа, порядок которой равен индексу нормального делителя H, называется дополнительной к H группой или фактор-группой относительно H.

Всякая группа G имеет два тривиальных делителя: один состоит из одного единичного элемента и другой совпадает со всей группой.

В дальнейшем, говоря о нормальном делителе, мы будем считать, что он отличен от упомянутых двух тривиальных нормальных делителей. Может случиться, что группа не имеет ни одного нормального делителя.

Такая группа называется *простой*.

59. Примеры. 1. Рассмотрим группу G вещественных ортогональных преобразований в трехмерном пространстве. Пусть H — подгруппа движения, т. е. совокупность ортогональных преобразований с определителем $(+1)$. Пусть, далее, S есть симметрия относительно начала, определяемая формулой (37). Если H_a есть переменный элемент из H, то полная группа G может быть представлена по схеме:

$$H_a,\ SH_a \quad \text{или} \quad H_a,\ H_aS. \tag{43}$$

Если G_1 есть любое преобразование из G, то $G_1H_aG_1^{-1}$ имеет определитель $(+1)$, т. е. принадлежит H, и H есть нормальный делитель индекса два. Рассмотрим группу, дополнительную к H. Первой из совокупностей (43) соответствует единичный элемент E этой группы. Произведение двух элементов из второй совокупности, т. е. двух ортогональных преобразований с определителем (-1), дает ортогональное преобразование с определителем $(+1)$, которое принадлежит первой совокупности. Если K — элемент, соответствующий второй совокупности, то из сказанного следует, что $K^2=E$. Итак, дополнительная к H группа состоит из двух элементов E и K, и $K^2=E$, т. е. это есть циклическая группа порядка два. Это будет вообще иметь место при нормальных делителях индекса два.

2. Для симметрической группы перестановок знакопеременная группа будет нормальным делителем индекса два.

Выпишем элементы симметрической группы с тремя элементами и обозначим каждый из них одной буквой, пользуясь обозначениями из [55]:

$$E;\ A=(2,\ 3);\ B=(1,\ 2);\ C=(1,\ 3);\ D=(1,\ 3,\ 2);\ F=(1,\ 2,\ 3).$$

Знакопеременная группа, состоящая из перестановок E, D и F, будет циклической группой третьего порядка ($F=D^2$ и $D=F^2$), причем $D^3=F^3=E$. Вся симметрическая группа состоит из трех классов: I E; II A, B и C; III D и F.

Знакопеременная группа также состоит из трех классов: I E; II D; III F. Нетрудно проверить, что закон умножения для элементов рассматриваемой симметрической группы совпадает с тем законом, который определяется таблицей (34) из [56].

Знакопеременная группа при $n=4$ содержит 12 элементов, которые распределяются на четыре класса:

I E; II $A_1=(1, 2)$ $(3, 4)$; $A_2=(1, 3)$ $(2, 4)$; $A_3=(1, 4)$ $(2, 3)$;
III $B_1=(1, 2, 3)$; $B_2=(2, 1, 4)$; $B_3=(3, 4, 1)$; $B_4=(4, 3, 2)$;
IV $C_1=(1, 2, 4)$; $C_2=(2, 1, 3)$; $C_3=(3, 4, 2)$; $C_4=(4, 3, 1)$.

Второй класс содержит три элемента второго порядка, а третий и четвертый классы — по четыре элемента третьего порядка. Произведение двух элементов второго класса дает, как нетрудно проверить, опять элемент второго класса, и, поскольку все элементы второго порядка попали во второй класс, можно утверждать, что эти три элемента совместно с единичным элементом образуют нормальный делитель рассматриваемой знакопеременной группы. Порядок его равен четырем, а индекс трем. Легко проверить, что элементы B_i третьего класса попадут в одну из сопряженных совокупностей элементов по указанному нормальному делителю, а элементы C_i — в другую сопряженную совокупность. Далее нетрудно проверить, что произведение двух элементов третьего класса дает некоторый элемент четвертого класса, а произведение двух элементов четвертого класса дает элемент третьего класса. В дополнительной группе указанному нормальному делителю соответствует единичный элемент E. Пусть A и B — два других элемента дополнительной группы. Из сказанного выше непосредственно следует, что $A^2=B$ и $B^2=A$, и непосредственно очевидно, что дополнительная группа, состоящая из элементов E, A и A^2, причем $A^3=E$, есть циклическая группа третьего порядка.

Отметим, что элементы E (1, 2, 3) и (2, 1, 3) нашей основной знакопеременной группы образуют циклическую подгруппу третьего порядка, но эта подгруппа не является элементарным делителем.

Если мы пронумеруем вершины тетраэдра в каком-либо порядке, то легко непосредственно проверить, что указанная выше знакопеременная группа при $n=4$ соответствует тем вращениям, при которых тетраэдр переходит в себя. Всякая перестановка определяет переход одних вершин в другие. Перестановкам третьего класса соответствует вращение вокруг одной из осей тетраэдра на угол $\frac{2\pi}{3}$, а перестановкам четвертого класса — вращения в противоположном направлении вокруг тех же осей и на тот же угол. Так, например, перестановкам (1, 2, 3) и (2, 1, 3) соответствуют вращения вокруг оси, проходящей через вершину с номером 4. Перестановкам второго класса соответствуют такие вращения тетраэдра, при которых ни одна из вершин не остается неизменной.

Можно доказать, что знакопеременная группа при $n>4$ является простой.

3. Если имеется абелева группа G и ее какая-либо подгруппа H, то при любом выборе элементов H_a из H и G_1 из G мы имеем $G_1H_a=H_aG_1$, т. е. $G_1H_aG_1^{-1}=H_a$, откуда непосредственно видно, что H есть нормальный делитель, т. е. всякая подгруппа абелевой группы есть ее нормальный делитель. В качестве примера рассмотрим группу G сложения векторов в R_n, о которой мы говорили в [49].

В качестве подгруппы H возьмем векторы, принадлежащие некоторому подпространству L_k из R_n $(0<k<n)$. Сопряженные совокупности получаются путем добавления к какому-либо вектору x из R_n всех векторов подпространства L_k.

Если x принадлежит L_k, то сопряженная совокупность совпадает с подгруппой H. Введем некоторые орты $x^{(1)}$, $x^{(2)}$, ..., $x^{(k)}$ в L_k и орты $x^{(k+1)}$, ..., $x^{(n)}$ в дополнительном подпространстве M_{n-k}. Всякая сопряженная совокупность элементов будет состоять, в силу сказанного выше, из векторов:

$$c_1x^{(1)}+c_2x^{(2)}+\ldots+c_kx^{(k)}+c_{k+1}x^{(k+1)}+\ldots+c_nx^{(n)},$$

причем c_{k+1}, ..., c_n имеют фиксированные значения, а c_1, c_2, ..., c_k — могут принимать любые значения.

Таким образом, всякой сопряженной совокупности мы можем сопоставить определенный вектор из M_{n-k}, и, наоборот, всякому вектору из M_{n-k} соответствует определенная сопряженная совокупность. Сложению двух векторов из каких-либо двух сопряженных совокупностей соответствует сложение соответствующих этим совокупностям векторов из M_{n-k}. Иначе говоря, элементами дополнительной группы можно считать векторы из M_{n-k} при прежней групповой операции (сложение векторов).

В этом примере порядок нормального делителя H и его индекс равны бесконечности.

60. Изоморфные и гомоморфные группы. Две группы A и B называются *изоморфными*, если между их элементами можно установить такое соответствие, что каждому элементу из A соответствует определенный элемент из B и, наоборот, всякому элементу из B соответствует определенный элемент A (биоднозначное соответствие), причем это соответствие таково, что произведению двух каких-либо элементов из A соответствует произведение соответствующих элементов из B. Если A и B — изоморфные абстрактные группы, то они имеют совершенно одинаковую структуру, т. е. по существу не отличаются одна от другой.

Перейдем теперь к установлению нового понятия, которое является обобщением понятия изоморфных групп. Группа B называется *гомоморфной группе* A, если каждому элементу A соответствует определенный элемент B, причем каждый элемент из B соответствует хоть одному элементу из A, и это соответствие таково, что произведению двух элементов из A соответствует произведение соответствующих элементов из B. В данном случае в отличие от изоморфных групп соответствие не должно быть обратно-однозначным, т. е. один и тот же элемент группы B может соответствовать нескольким различным элементам группы A. Если группа B гомоморфна группе A, и каждому элементу из B будет соответствовать один определенный элемент из A, то эти группы будут и изоморфными. Заметим, кроме того, что если элементам A_1 и A_2 из A соответствуют элементы B_1 и B_2 из B, то по определению элементу A_2A_1 из A соответствует элемент B_2B_1 из B.

Пусть A_0 — единичный элемент из A и B_0 — соответствующий элемент из B. Нетрудно показать, что и B_0 будет единичным элементом. Действительно, для любого A_1 из A имеем равенство

$$A_0A_1 = A_1A_0 = A_1,$$

которое приводит к равенству соответствующих элементов из B:

$$B_0B_1 = B_1B_0 = B_1,$$

причем по определению гомоморфизма B_1 можно считать любым элементом из B. Последнее равенство показывает, что B_0 есть единичный элемент группы B. Итак, в изоморфных и гомоморфных группах единичному элементу из A соответствует единичный элемент из B. Возьмем теперь два обратных элемента A_1 и A_1^{-1} из A, и пусть B_1 и B_2 — соответствующие элементы из B. Равенство $A_1A_1^{-1}=A_1^{-1}A_1=A_0$, где A_0 — единичный элемент, дает, по определению гомоморфных групп, $B_1B_2=B_2B_1=B_0$, где по предыдущему B_0 — единичный элемент, а потому $B_2=B_1^{-1}$, т. е. обратным элементам из A соответствуют и обратные элементы из B.

Положим, что наши группы только гомоморфны, но не изоморфны. Рассмотрим совокупность элементов C_α в группе A, которым соответствует единичный элемент B_0 из B. Если C_α соответствует B_0, то по предыдущему C_α^{-1} соответствует $B_0^{-1}=B_0$, и всякому произведению $C_{\alpha_2}C_{\alpha_1}$ соответствует также $B_0B_0=B_0$, т. е. совокупность элементов C_α группы A, которым соответствует единичный элемент из B, образует некоторую подгруппу C группы A.

Покажем, что эта подгруппа будет нормальным делителем. Действительно, пусть A_1 — какой-нибудь элемент группы A и B_1 — соответствующий элемент из B. Всякому элементу вида $A_1C_\alpha A_1^{-1}$ соответствует в B элемент $B_1B_0B_1^{-1}$, или в силу основного свойства единичного элемента можно утверждать, что всякому элементу вида $A_1C_\alpha A_1^{-1}$ соответствует единичный элемент из B, т. е. всякий элемент вида $A_1C_\alpha A_1^{-1}$ есть один из элементов C_α, т. е. он принадлежит подгруппе C, а следовательно, эта подгруппа C есть нормальный делитель. Рассмотрим теперь разбиение группы A на сопряженные совокупности по схеме

$$C_\alpha,\quad A_1C_\alpha,\quad A_2C_\alpha,\ \dots \tag{44}$$

Пусть B_k — элемент, соответствующий A_k. Возьмем два элемента $A_kC_{\alpha 1}$ и $A_kC_{\alpha 2}$, принадлежащих одной и той же сопряженной совокупности. Им будут соответствовать элементы B_kB_0 и B_kB_0, т. е. один и тот же элемент B_k из B.

Элементам из разных сопряженных совокупностей A_kC_α и A_lC_α соответствуют элементы B_k и B_l. Покажем, что эти последние элементы различны. Действительно, если бы они были одинаковы, то элементу $A_k^{-1}A_l$ соответствовал бы единичный элемент B_0 из B, т. е. элемент $A_k^{-1}A_l$ должен был бы быть одним из элементов C_α, т. е. мы имели бы $A_k^{-1}A_l=C_{\alpha_0}$, т. е. $A_l=A_kC_{\alpha_0}$, что противоречит схеме (44). Итак, *если группа B гомоморфна группе A, то совокупность элементов C из A, соответствующих единичному элементу из B, образует нормальный делитель, и всякая сопряженная совокупность с этим нормальным делителем представляет собою совокупность всех элементов A, которым соответствует один и тот же элемент из B.* Из определения гомоморфных групп, следует непосредственно, что произведению двух каких-либо элементов из различных (или одинаковых) сопряженных совокупностей соответствует произведение тех элементов группы B, которые соответствуют упомянутым сопряженным совокупностям, т. е., короче говоря, всякой сопряженной совокупности из A соответствует определенный элемент из B, различным сопряженным совокупностям соответствуют различные элементы из B, и это соответствие устанавливает изоморфность группы B с дополнительной для C_a группой в A.

В качестве примера возьмем опять группу вещественных ортогональных преобразований в трехмерном пространстве и сопоставим каждому такому преобразованию число, равное определителю преобразования, определив умножение в области этих чисел обычным образом, как умножения чисел. В данном случае наша группа будет гомоморфна группе, состоящей из двух элементов $(+1)$ и (-1), причем умножение для этих двух элементов определяется обычным образом, как умножение чисел. Роль единичного элемента будет играть $(+1)$. Для этого примера нормальный делитель будет представлять собою группу вращения.

Если группа B гомоморфна, но не изоморфна группе A, то совокупность элементов группы A, которым соответствует единичный элемент из B, называется обычно *ядром гомоморфизма*. Мы видели, что ядро гомоморфизма является нормальным делителем группы A.

61. Примеры. 1. Возьмем группу G вещественных ортогональных преобразований в трехмерном пространстве, сопоставим каждому такому преобразованию число, равное определителю этого преобразования, и определим групповую операцию для этих чисел как их обычное умножение. При этом группа G', состоящая из чисел $(+1)$ и (-1) при обычном умножении этих чисел, будет гомоморфной группе G. Единичный элемент $(+1)$ группы G' соответствует вращениям трехмерного пространства из G. Эти вращения образуют нормальный делитель, а дополнительной группой является циклическая группа порядка два [59].

2. Возьмем на плоскости XY правильный треугольник с вершинами:

$$(1,\ 0);\quad (\cos 120°,\ \sin 120°);\quad (\cos 240°,\ \sin 240°)$$

и образуем группу G, состоящую из вращений плоскости вокруг начала на угол 0°, 120°, 240°, при которых треугольник переходит в себя, и из симметрии плоскости относительно оси X с последующим вращением на угол 0°, 120°, 240°. Это будет группа диэдра при $n=3$.

Выпишем все матрицы, соответствующие элементам этой группы:

$$E=\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix};\quad A=\begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix};\quad B=\begin{Vmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{Vmatrix};$$

$$C=\begin{Vmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{Vmatrix};\quad D=\begin{Vmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{Vmatrix};\quad F=\begin{Vmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{Vmatrix}.$$

Если мы обратимся к схеме умножения, определенной таблицей (34) из [56], то увидим, что эта схема умножения как раз и соответствует умножению матриц, образующих нашу группу. Выше мы видели [59], что упомянутая схема умножения соответствует также симметрической группе перестановок из трех элементов:

$$E;\ A=(2,3);\quad B=(1,2);\quad C=(1,3);\quad D=(1,3,2);\quad F=(1,2,3). \qquad (45)$$

Таким образом, если мы будем считать элементы этих двух групп, обозначенные одной и той же буквой, соответствующими, то эти две группы будут изоморфными. Перестановки группы (45) соответствуют перестановкам вершин упомянутого выше треугольника, если их занумеровать соответствующим образом.

Совершенно так же, как об этом мы уже упоминали в [59], группа тетраэдра изоморфна знакопеременной группе при $n=4$.

3. Можно указать общий прием построения групп перестановок, гомоморфных данной группе G. Пусть H — какая-либо подгруппа группы G конечного индекса n. Напишем соответствующие ей сопряженные совокупности элементов:

$$H,\ HS_1,\ HS_2,\ \ldots,\ HS_{n-1}.$$

Если мы умножим каждую из этих совокупностей справа на некоторый элемент S из G, то произойдет лишь некоторая перестановка порядка этих совокупностей, и будем считать, что эта перестановка и соответствует взятому элементу S из G. Нетрудно показать, что таким образом и получится группа G' перестановок, гомоморфная группе G.

Для того чтобы элементу S из G соответствовал единичный элемент из G', необходимо и достаточно, чтобы при умножении справа на S всякая сопряженная совокупность переходила в себя, т. е. чтобы

$$H_\alpha S = H_\beta \quad \text{и} \quad H_\alpha S_k S = H_\beta S_k \qquad (k=1,2,\ldots,n-1),$$

где H_α — любой элемент H и H_β также принадлежит H. Написанные равенства можно переписать в виде:

$$S = H_\alpha^{-1} H_\beta;\quad S = (S_k^{-1} H_\alpha S_k)^{-1}(S_k^{-1} H_\beta S_k),$$

и из них следует, что для того чтобы элементу S соответствовал единичный элемент из G', необходимо и достаточно, чтобы S одновременно принадлежало H и всем подобным подгруппам $S_k^{-1} H S_k$.

Если H — нормальный делитель G, то указанное требование сводится к тому, что S принадлежит H, и группа G' в этом случае изоморфна дополнительной группе. Если H сводится к одному единичному элементу, то группа G изоморфна группе перестановок G', которая получается, если элементы группы G:

$$E,\ S_1,\ S_2,\ \ldots,\ S_n$$

умножим справа на любой элемент S из G, что приведет к некоторой перестановке элементов G. В дальнейшем мы подробно рассмотрим построение групп линейных преобразований, изоморфных заданной группе.

62. Стереографическая проекция. Закончив основы общей теории групп, мы переходим теперь к рассмотрению некоторого частного примера соответствия между группами, играющего важную роль в физике. Предварительно выясним понятие о *стереографической проекции*, дающей определенный закон соответствия между точками сферы и плоскости.

Рассмотрим трехмерное пространство с координатными осями XYZ и сферу C с центром в начале и радиусом единица. Пусть S — точка сферы, имеющая координаты $(0, 0, -1)$, и M — переменная точка на сфере (рис. 3). Прямая SM пересечет плоскость XY в некоторой точке P, и мы имеем таким образом вполне определенный закон соответствия между точками сферы C и точками плоскости XY, причем точке сферы S с координатами $(0,\ 0,\ -1)$ соответствует бесконечно далекая точка плоскости. Установленное соответствие точек и дает нам стереографическую проекцию сферы на плоскость.

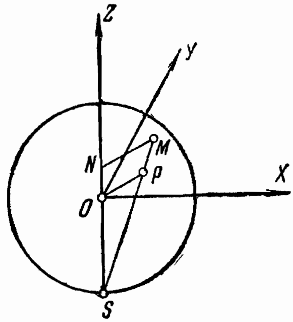


Рис. 3.

Обратимся теперь к выводу формул, дающих стереографическую проекцию. Пусть MN — перпендикуляр из точки M на ось Z. Мы имеем из подобия треугольников, принимая во внимание, что $SO=1$:

$$NM=(1+ON)\,OP. \tag{46}$$

Обозначая через $(x,\ y,\ z)$ координаты точки M и через $(\alpha,\ \beta)$ координаты P, сможем написать:

$$NM=(1+z)\,OP,$$

или, проектируя параллельные отрезки OP и NM на оси X и Y:

$$x=(1+z)\,\alpha;\quad y=(1+z)\,\beta. \tag{47}$$

Уравнение $x^2+y^2+z^2=1$ дает нам квадратное уравнение для z:

$$(\alpha^2+\beta^2)(1+z)^2+z^2=1,$$

и, решая его, получим:

$$z=\frac{\pm 1-(\alpha^2+\beta^2)}{1+(\alpha^2+\beta^2)}.$$

Но для всех точек $(\alpha,\ \beta)$ на конечном расстоянии мы должны иметь $z>-1$, и, следовательно, в предыдущей формуле мы должны брать $(+1)$. Пользуясь еще формулами (47), получим окончательные выражения $(x,\ y,\ z)$ через $(\alpha,\ \beta)$:

$$x=\frac{2\alpha}{1+\alpha^2+\beta^2};\quad y=\frac{2\beta}{1+\alpha^2+\beta^2};\quad z=\frac{1-(\alpha^2+\beta^2)}{1+\alpha^2+\beta^2}. \tag{48}$$

Вместо двух вещественных координат α и β на плоскости введем одну комплексную координату $\zeta=\alpha+i\beta$. Обозначая, как всегда, через $\bar{\zeta}$ комплексное число, сопряженное с ζ, мы можем переписать предыдущие формулы в следующем виде:

$$x+iy=\frac{2\zeta}{1+\zeta\bar{\zeta}};\quad x-iy=\frac{2\bar{\zeta}}{1+\zeta\bar{\zeta}};\quad z=\frac{1-\zeta\bar{\zeta}}{1+\zeta\bar{\zeta}}. \tag{49}$$

Представим комплексное переменное ζ в виде отношения двух других комплексных ξ и η:

$$\zeta=\frac{\eta}{\xi}. \tag{50}$$

Пары значений ξ и η, отличающиеся общим множителем, т. е. пары вида $k\xi$, $k\eta$ и ξ, η, будут давать одно и то же ζ, т. е. одну и ту же точку плоскости, и пара значений $\eta \neq 0$, $\xi = 0$ будет давать бесконечно далекую точку плоскости. Комплексные числа ξ и η называются однородными комплексными координатами на плоскости. Формулы (49) мы можем, пользуясь (50) и отделяя вещественную и мнимую части, переписать в виде:

$$x = \frac{\bar{\xi}\eta + \xi\bar{\eta}}{\xi\bar{\xi} + \eta\bar{\eta}}; \quad y = \frac{1}{i}\frac{\bar{\xi}\eta - \xi\bar{\eta}}{\xi\bar{\xi} + \eta\bar{\eta}}; \quad z = \frac{\xi\bar{\xi} - \eta\bar{\eta}}{\xi\bar{\xi} + \eta\bar{\eta}}. \tag{51}$$

Для любых комплексных значений ξ и η последние формулы дают вещественные (x, y, z), удовлетворяющие соотношению

$$x^2 + y^2 + z^2 - 1 = 0, \tag{51_1}$$

как это и следовало ожидать, ибо точка (x, y, z) находится на единичной сфере.

63. Унитарная группа и группа движения. Рассмотрим теперь некоторое унитарное преобразование над переменными (ξ, η):

$$\xi' = a\xi + b\eta, \quad \eta' = c\xi + d\eta, \tag{52}$$

причем в силу унитарности должно быть:

$$\xi'\bar{\xi}' + \eta'\bar{\eta}' = \xi\bar{\xi} + \eta\bar{\eta}. \tag{53}$$

Новые значения переменных (ξ', η') дадут нам и новую точку на сфере

$$x' = \frac{\bar{\xi}'\eta' + \xi'\bar{\eta}'}{\xi'\bar{\xi}' + \eta'\bar{\eta}'}; \quad y' = \frac{1}{i}\frac{\bar{\xi}'\eta' - \xi'\bar{\eta}'}{\xi'\bar{\xi}' + \eta'\bar{\eta}'}; \quad z' = \frac{\xi'\bar{\xi}' - \eta'\bar{\eta}'}{\xi'\bar{\xi}' + \eta'\bar{\eta}'}. \tag{54}$$

Определитель унитарного преобразования (52), равный, как известно, по модулю единице, будет выражаться некоторым числом вида $e^{i\varphi}$. Умножая все коэффициенты преобразования (52) на $e^{-i\frac{\varphi}{2}}$, получим унитарное преобразование с определителем единица. Но при этом ξ' и η' умножатся также на $e^{-i\frac{\varphi}{2}}$. Этот дополнительный множитель совершенно не повлияет на величину ζ. Мы можем, таким образом, ограничиться рассмотрением унитарных преобразований (52) при условии, что определитель преобразования равен единице, т. е.

$$ad - bc = 1. \tag{55}$$

Даже при этом ограничительном условии два преобразования, коэффициенты которых отличаются знаком, дадут нам значения ξ' и η', отличающиеся знаком, и мы придем при обоих этих преобразованиях к одной и той же точке ζ'.

Если мы подставим в формулы (54) вместо ξ' и η' их выражения (52) и примем во внимание условие (53), то увидим, пользуясь (51), что переменные (x', y', z') выразятся в виде линейных однородных полиномов через (x, y, z). В силу (53) знаменатель в выражениях (51) и (54) оказывается одинаковым, и переменные (x, y, z) испытывают то же самое линейное преобразование, какое испытывают выражения:

$$u = \xi\bar{\eta} + \bar{\xi}\eta; \qquad v = \frac{1}{i}(\xi\bar{\eta} - \bar{\xi}\eta); \qquad w = \xi\bar{\xi} - \eta\bar{\eta} \tag{56}$$

при унитарном преобразовании (52). Дальше мы установим точно вид этого линейного преобразования.

Установим прежде всего общий вид унитарных преобразований (52) с определителем единица. Обычные условия унитарности дают нам [28]:

$$a\bar{c} + b\bar{d} = 0; \quad c\bar{c} + d\bar{d} = 1.$$

Умножая условие (55) на $\bar{c}$ и пользуясь первым из написанных условий, получим:

$$-bd\bar{d} - bc\bar{c} = \bar{c},$$

откуда в силу второго условия будем иметь $\bar{c} = -b$ или $c = -\bar{b}$, и совершенно аналогично можно показать, что $d = \bar{a}$. Мы можем, таким образом, написать все унитарные преобразования с определителем, равным единице, в следующем виде:

$$\left.\begin{aligned} \xi' &= a\xi + b\eta, \\ \eta' &= -\bar{b}\xi + \bar{a}\eta, \end{aligned}\right\} \tag{57}$$

где a и b — любые комплексные числа, удовлетворяющие условию

$$a\bar{a} + b\bar{b} = 1. \tag{58}$$

Составим теперь выражения (56) с новыми переменными

$$u' + iv' = 2\bar{\xi}'\eta'; \quad u' - iv' = 2\xi'\bar{\eta}'; \quad w' = \xi'\bar{\xi}' - \eta'\bar{\eta}',$$

или, пользуясь (57):

$$\begin{aligned} u' + iv' &= \bar{a}^2 2\bar{\xi}\eta - \bar{b}^2 2\xi\bar{\eta} - 2\bar{a}\bar{b}(\xi\bar{\xi} - \eta\bar{\eta}), \\ u' - iv' &= -b^2 2\bar{\xi}\eta + a^2 2\xi\bar{\eta} - 2ab(\xi\bar{\xi} - \eta\bar{\eta}), \\ w' &= \bar{a}b 2\bar{\xi}\eta + a\bar{b} 2\xi\bar{\eta} + (a\bar{a} - b\bar{b})(\xi\bar{\xi} - \eta\bar{\eta}). \end{aligned}$$

Производя замену

$$2\bar{\xi}\eta = u + iv; \qquad 2\xi\bar{\eta} = u - iv; \qquad \xi\bar{\xi} - \eta\bar{\eta} = w$$

и складывая и вычитая первые два уравнения, получим выражения

(u', v', w') через (u, v, w), или, что то же, выражения (x', y', z') через (x, y, z):

$$\left.\begin{aligned} x' &= \frac{1}{2}(a^2+\bar{a}^2-b^2-\bar{b}^2)\,x+ \\ &\quad+\frac{i}{2}(\bar{a}^2+\bar{b}^2-a^2-b^2)\,y-(ab+\bar{a}\bar{b})\,z, \\ y' &= \frac{i}{2}(a^2+\bar{b}^2-\bar{a}^2-b^2)\,x+ \\ &\quad+\frac{1}{2}(a^2+\bar{a}^2+b^2+\bar{b}^2)\,y+i(\bar{a}\bar{b}-ab)\,z, \\ z' &= (a\bar{b}+\bar{a}b)\,x+i(a\bar{b}-\bar{a}b)\,y+(a\bar{a}-b\bar{b})\,z. \end{aligned}\right\} \tag{59}$$

Всякому унитарному преобразованию (57) соответствует некоторое преобразование плоскости XY, а это, в свою очередь, в силу соответствия, устанавливаемого стереографической проекцией, дает некоторое преобразование сферы.

В соответствии с этим (59) есть вещественное преобразование, в силу которого уравнение

$$x^2+y^2+z^2=1$$

переходит в уравнение

$$x'^2+y'^2+z'^2=1.$$

Но линейное однородное преобразование (59) не меняет свободного члена 1, и, следовательно, оно должно оставлять неизменной и левую часть уравнения, т. е.

$$x'^2+y'^2+z'^2=x^2+y^2+z^2.$$

Все эти обстоятельства можно непосредственно получить и из самого вида преобразования (59). Итак, формулы (59) дают вещественные ортогональные преобразования с тремя переменными. Покажем теперь, что определитель преобразования (59) всегда равен $(+1)$. Этот определитель есть непрерывная функция вещественных и мнимых частей комплексных переменных a и b, которые должны удовлетворять соотношению (58). Но величина определителя может быть только $(+1)$ или (-1), и в силу упомянутой выше непрерывности эта величина должна быть все время $(+1)$ или все время (-1). Но при $a=1$ и $b=0$ формулы (59) дают нам тождественное преобразование с определителем $(+1)$, т. е. действительно определитель преобразования (59) всегда равен $(+1)$. Итак, линейные преобразования (59) представляют собою вращение пространства вокруг начала.

Докажем теперь, что всякое вращение пространства может быть представлено в виде (59). Если мы положим:

$$a=e^{-\frac{i}{2}\varphi};\quad \bar{a}=e^{\frac{i}{2}\varphi};\quad b=\bar{b}=0,$$

т. е. возьмем матрицу унитарного преобразования

$$A_\varphi = \begin{Vmatrix} e^{-\frac{i}{2}\varphi} & 0 \\ 0 & e^{\frac{i}{2}\varphi} \end{Vmatrix}, \tag{60}$$

то формулы (59) дадут нам:

$$\left.\begin{aligned} x' &= x\cos\varphi - y\sin\varphi, \\ y' &= x\sin\varphi + y\cos\varphi, \\ z' &= z, \end{aligned}\right\} \tag{61}$$

т. е. мы получим вращение вокруг оси Z на угол φ.

Если теперь возьмем:

$$a = \bar{a} = \cos\frac{\psi}{2}; \quad b = -i\sin\frac{\psi}{2}; \quad \bar{b} = i\sin\frac{\psi}{2},$$

т. е. матрицу унитарного преобразования определим следующим образом:

$$B_\psi = \begin{Vmatrix} \cos\frac{\psi}{2} & -i\sin\frac{\psi}{2} \\ -i\sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{Vmatrix}, \tag{62}$$

то формулы (59) дадут нам:

$$\left.\begin{aligned} x' &= x, \\ y' &= y\cos\psi - z\sin\psi, \\ z' &= y\sin\psi + z\cos\psi. \end{aligned}\right\} \tag{63}$$

Это будет вращение вокруг оси X на угол ψ.

Но, как мы знаем [**20**], всякое вращение с углами Эйлера $\{\alpha, \beta, \gamma\}$ может быть получено в результате поворота вокруг Z на угол α, последующего поворота вокруг новой оси X на угол β и последующего затем поворота вокруг новой оси Z на угол γ.

Обозначим через Z_φ матрицу третьего порядка, соответствующую преобразованию (61), и через X_ψ — матрицу преобразования (63). Поворот вокруг оси Z на угол α будет осуществляться матрицей Z_α. При этом новая ось X получится из прежней оси X при помощи этой же матрицы, и поворот вокруг новой оси X на угол β будет осуществляться, как нетрудно видеть, матрицей $Z_\alpha X_\beta Z_\alpha^{-1}$, и первые два поворота осуществляются матрицей

$$Z_\alpha X_\beta Z_\alpha^{-1} Z_\alpha = Z_\alpha X_\beta.$$

Как и выше, поворот вокруг новой оси Z на угол γ будет осуществляться матрицей

$$(Z_\alpha X_\beta)\, Z_\gamma\, (Z_\alpha X_\beta)^{-1},$$

и окончательно вращение $\{\alpha, \beta, \gamma\}$ будет осуществляться матрицей

$$(Z_\alpha X_\beta) Z_\gamma (Z_\alpha X_\beta)^{-1} (Z_\alpha X_\beta)$$

или

$$Z_\alpha X_\beta Z_\gamma. \tag{64}$$

В предыдущих рассуждениях мы пользовались тем очевидным фактом, что если Z_φ есть матрица, дающая вращение вокруг некоторой оси l, проходящей через начало, на угол φ, и матрица M переводит l в ось l_1, то вращение вокруг l_1 на угол φ будет определяться подобной матрицей:

$$MZ_\varphi M^{-1}.$$

Заметим теперь, что если A_1 и A_2 — два унитарных преобразования (57), которым соответствуют ортогональные преобразования (59) U_1 и U_2, то произведению A_2A_1 будет, очевидно, соответствовать также произведение U_2U_1. Таким образом, вращение пространства $\{\alpha, \beta, \gamma\}$ будет получаться в силу (64) от унитарной матрицы, являющейся произведением трех унитарных матриц:

$$\begin{Vmatrix} e^{-i\frac{\alpha}{2}} & 0 \\ 0 & e^{i\frac{\alpha}{2}} \end{Vmatrix} \cdot \begin{Vmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{Vmatrix} \cdot \begin{Vmatrix} e^{-i\frac{\gamma}{2}} & 0 \\ 0 & e^{i\frac{\gamma}{2}} \end{Vmatrix}. \tag{65}$$

Итак, всякому унитарному преобразованию соответствует определенное вращение трехмерного пространства, и таким образом получаются все вращения. Произведению двух унитарных преобразований будет соответствовать произведение соответствующих вращений. Мы можем сказать, что формулы (59) определяют гомоморфизм группы унитарных преобразований с определителем единица с группой вращения трехмерного пространства.

Посмотрим теперь, какие унитарные преобразования дают тождественное преобразование, т. е. единичный элемент в группе вращения. Третья из формул (59) дает нам при этом

$$a\bar{b} = 0; \quad a\bar{a} - b\bar{b} = 1,$$

откуда $|a| = 1$ и $b = 0$. Пусть $a = e^{i\vartheta}$. Первая из формул (59) дает:

$$\frac{1}{2}(e^{i2\vartheta} + e^{-i2\vartheta}) = 1.$$

Отсюда непосредственно следует $\vartheta = 0$ или π, т. е. $a = \pm 1$.

Мы получаем, таким образом, два унитарных преобразования с матрицами

$$E = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}; \quad S = \begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix} = -E,$$

которым соответствует единичный элемент в группе вращения.

Положим теперь, что две унитарные матрицы U и V дают одно и то же вращение. При этом $V^{-1}U$ будет давать тождественное преобразование в группе вращения пространства, т. е. $V^{-1}U=E$ или $(-E)$, так что $U=V$ или $U=-V$. Заметим при этом, что знак $(-)$ перед матрицей означает, что у всех элементов матрицы надо изменить знак. Предыдущие рассуждения показывают, что унитарные преобразования (57) приводят к одинаковому вращению пространства только тогда, когда они отличаются лишь знаком. Наоборот, если они отличаются только знаком, то, как мы уже упоминали выше и как это следует из формулы (59), они дают одно и то же вращение пространства. Окончательно можем сказать, что *группа вращения пространства будет гомоморфна группе унитарных преобразований* (57) *с определителем единица, причем одинаковые вращения получаются тогда и только тогда, когда унитарные матрицы отличаются лишь знаком.*

Матрицы E и $(-E)$ образуют нормальный делитель H группы G — унитарных преобразований (57) с определителем единица. Всякая сопряженная совокупность по этому нормальному делителю H состоит из двух элементов G_1 и $(-G_1)$, где G_1 — любой элемент группы. Из сказанного выше непосредственно следует, что *группа вращения изоморфна дополнительной к H группе.*

Формулы (59) содержат два комплексных параметра a и b, которые должны удовлетворять соотношению (58). Каждый из комплексных параметров содержит два вещественных параметра

$$a=a_1+ia_2;\quad b=b_1+ib_2,$$

и соотношение (58) равносильно следующему:

$$a_1^2+a_2^2+b_1^2+b_2^2=1.$$

Таким образом, формулы (59) содержат четыре вещественных параметра, которые должны удовлетворять одному соотношению, т. е. формулы (59) содержат три независимых вещественных параметра, как это и должно быть для группы вращения. Параметры a и b называются обычно *параметрами Кейли — Клейна.* Нетрудно получить их выражение через углы Эйлера. Действительно, перемножая три унитарные матрицы (65), получим, как мы видели выше, ту унитарную матрицу, которая будет соответствовать вращению с углами Эйлера $\{\alpha, \beta, \gamma\}$. Умножая, получим для соответствующих параметров a и b следующие выражения:

$$a=e^{-i\frac{\alpha+\gamma}{2}}\cos\frac{\beta}{2};\quad b=-ie^{i\frac{\gamma-\alpha}{2}}\sin\frac{\beta}{2}. \tag{66}$$

Если прибавить 2π к α или γ, то a и b переменят знак, а вращение по существу останется тем же. Это обстоятельство было отмечено уже выше.

64. Общая линейная группа и группа Лоренца. Мы установили только что тесную связь унитарной группы с двумя переменными и группы вращения трехмерного пространства. Совершенно аналогичным образом можно установить связь между общей линейной группой с двумя переменными и с определителем, равным единице, и группой Лоренца.

Введем четыре переменные x_1, x_2, x_3, x_0, и, обращаясь к формулам (51), выражающим стереографическую проекцию, положим в них

$$x=\frac{x_1}{x_0};\quad y=\frac{x_2}{x_0};\quad z=\frac{x_3}{x_0}. \tag{67}$$

Это дает нам следующие формулы:

$$\frac{x_1}{x_0}=\frac{\bar{\xi}\eta+\xi\bar{\eta}}{\xi\bar{\xi}+\eta\bar{\eta}};\quad \frac{x_2}{x_0}=\frac{1}{i}\,\frac{\bar{\xi}\eta-\xi\bar{\eta}}{\xi\bar{\xi}+\eta\bar{\eta}};\quad \frac{x_3}{x_0}=\frac{\xi\bar{\xi}-\eta\bar{\eta}}{\xi\bar{\xi}+\eta\bar{\eta}}.$$

Они определяют x_k с точностью до произвольного общего множителя, и мы можем положить:

$$\left.\begin{aligned} &x_0=\xi\bar{\xi}+\eta\bar{\eta}; && x_1=\bar{\xi}\eta+\xi\bar{\eta};\\ &x_2=\frac{1}{i}(\bar{\xi}\eta-\xi\bar{\eta}); && x_3=\xi\bar{\xi}-\eta\bar{\eta}. \end{aligned}\right\} \tag{68}$$

Прежние переменные удовлетворяли соотношению (51_1), и, следовательно, в силу (67), новые переменные, определяемые по формулам (68), будут при любых комплексных значениях ξ и η удовлетворять соотношению

$$x_1^2+x_2^2+x_3^2-x_0^2=0. \tag{69}$$

При унитарности преобразования с ξ и η выражение $(\xi\bar{\xi}+\eta\bar{\eta})$ оставалось неизменным, т. е., согласно (68), оставалась неизменной переменная x_0, выражающая сейчас время, и мы получали таким образом вращение трехмерного пространства. Откажемся теперь от унитарности преобразования и рассмотрим общую группу линейных преобразований

$$\xi'=a\xi+b\eta;\quad \eta'=c\xi+d\eta. \tag{70}$$

Будем поступать дальше аналогично тому, как мы поступили в случае унитарных преобразований. Составим выражения:

$$\left.\begin{aligned} &x_1+ix_2=2\bar{\xi}\eta; && x_1-ix_2=2\xi\bar{\eta};\\ &x_0+x_3=2\xi\bar{\xi}; && x_0-x_3=2\eta\bar{\eta}. \end{aligned}\right\} \tag{71}$$

Для новых переменных ξ', η' получим новые x_k':

$$x_1'+ix_2'=2\bar{\xi}'\eta';\quad x_1'-ix_2'=2\xi'\bar{\eta}';$$
$$x_0'+x_3=2\xi'\bar{\xi}';\quad x_0'-x_3'=2\eta'\bar{\eta}'.$$

Подставляя выражения (70) и пользуясь (71), получим:

$$\left.\begin{aligned}
x_1' + ix_2' &= \bar{a}d(x_1 + ix_2) + \bar{b}c(x_1 - ix_2) + \\
&\quad + \bar{a}c(x_0 + x_3) + \bar{b}d(x_0 - x_3), \\
x_1' - ix_2' &= b\bar{c}(x_1 + ix_2) + a\bar{d}(x_1 - ix_2) + \\
&\quad + a\bar{c}(x_0 + x_3) + b\bar{d}(x_0 - x_3), \\
x_0' + x_3' &= \bar{a}b(x_1 + ix_2) + a\bar{b}(x_1 - ix_2) + \\
&\quad + a\bar{a}(x_0 + x_3) + b\bar{b}(x_0 - x_3), \\
x_0' - x_3' &= \bar{c}d(x_1 + ix_2) + c\bar{d}(x_1 - ix_2) + \\
&\quad + c\bar{c}(x_0 + x_3) + d\bar{d}(x_0 - x_3),
\end{aligned}\right\} \tag{72}$$

откуда непосредственно получаются линейные выражения с вещественными коэффициентами x_k' через x_k, которые мы выписывать не будем. Заметим только, что если сложить последние два из уравнений (72), то в выражении для x_0' коэффициент при x_0 окажется равным $\frac{1}{2}(a\bar{a} + b\bar{b} + c\bar{c} + d\bar{d})$, т. е. этот коэффициент будет положительным.

Новые переменные x_k', как и первоначальные, удовлетворяют уравнению:

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = 0. \tag{73}$$

Если в левой части этого уравнения заменить x_k' их выражениями через x_k, то должно получиться уравнение (69). Но левая часть уравнения (73) может при этом оказаться отличающейся от левой части уравнения (69) на постоянный множитель, т. е. мы будем иметь в данном случае:

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = k(x_1^2 + x_2^2 + x_3^2 - x_0^2),$$

где k — некоторая постоянная. Пользуясь предыдущими формулами и принимая во внимание, что

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = (x_1' + ix_2')(x_1' - ix_2') - (x_0' + x_3')(x_0' - x_3'),$$

нетрудно показать, что $k = (ad - bc)(\bar{a}\bar{d} - \bar{b}\bar{c}) = |ad - bc|^2$. Если мы хотим получить $k = 1$, т. е. преобразование Лоренца:

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = x_1^2 + x_2^2 + x_3^2 - x_0^2, \tag{74}$$

то должны брать линейные преобразования (70) с определителем по модулю, равным единице, т. е. выражающимся числом вида $e^{i\varphi}$. Умножая, как и раньше, все коэффициенты преобразования (70) на $e^{-i\frac{\varphi}{2}}$, мы, с одной стороны, не изменим величин x_1', x_2', x_3', определяемых по формулам (68), если в них заменить ξ и η на ξ' и η', ибо эти формулы содержат произведение одной из величин (ξ', η') на одну из величин $(\bar{\xi}', \bar{\eta}')$, и, с другой стороны, приведем определитель преобразования к единице.

Будем, таким образом, рассматривать преобразования (70) с определителем единица:

$$ad - bc = 1. \tag{75}$$

Как и в предыдущем номере, мы можем показать, что линейное преобразование, выражающее x'_k через x_k, имеет определитель $(+1)$. Напомним, кроме того, что в нем коэффициент при x_0 в выражении x'_0 — положителен, т. е. это преобразование имеет определитель $(+1)$ и не меняет направления отсчета времени, т. е. преобразования (72) суть положительные преобразования Лоренца.

Итак, окончательно, линейные преобразования при условии (75) дают положительные преобразования Лоренца, которые мы определили в [**54**].

Как и в предыдущем номере, поставим теперь вопрос — можно ли получить по формулам (72) любое положительное преобразование Лоренца. Отметим прежде всего, что, как и в предыдущем номере, произведению двух линейных преобразований (70) отвечает произведение соответствующих преобразований Лоренца, т. е., точнее говоря: если A и B — два линейных преобразования (70), которые приводят, согласно (72), к преобразованиям Лоренца T_1 и T_2, то линейному преобразованию BA будет соответствовать преобразование Лоренца T_2T_1. Как мы видели в [**54**], всякое положительное преобразование Лоренца может быть представлено в виде:

$$T = VSU,$$

где U и V суть простые вращения трехмерного пространства и S — положительное преобразование Лоренца с двумя переменными. Согласно результатам предыдущего номера, мы можем получить любое вращение при помощи некоторого унитарного преобразования вида (70) с определителем единица. Таким образом, нам остается показать, что мы можем получить и любое положительное преобразование S Лоренца с двумя переменными по формулам (72) при соответствующем выборе линейного преобразования (70). Сравнивая (74) с (21) из [**54**], видим, что сейчас мы считаем $c = 1$, так что формулы (17) из [**54**], дающие положительные преобразования Лоренца с двумя переменными, перепишутся в виде:

$$\left.\begin{aligned} x'_3 &= \frac{-vx_0 + x_3}{\sqrt{1-v^2}}; & x'_0 &= \frac{x_0 - vx_3}{\sqrt{1-v^2}} \\ x_1 &= x_1; & x_2 &= x_2. \end{aligned}\right\} \tag{76}$$

Введем величину

$$u = \frac{1}{\sqrt{1-v^2}} > 1$$

и рассмотрим линейное преобразование (70) частного вида:

$$\xi' = l\xi; \quad \eta' = \frac{1}{l}\eta,$$

где l — вещественная постоянная. Определитель его, очевидно, равен единице. В данном случае $a = l$, $d = \frac{1}{l}$ и $b = c = 0$. Подставляя это в формулы (72), мы получим как раз преобразование (76), если l удовлетворяет условиям:

$$\frac{l^2}{2} + \frac{1}{2l^2} = u; \quad \frac{l^2}{2} - \frac{1}{2l^2} = -vu.$$

Это дает непосредственно $l^2 = u \pm \sqrt{u^2 - 1}$. Второе из условий показывает, что при $v > 0$ надо брать корень для l^2, меньший единицы, а при $v < 0$ — больший единицы, и при этом второе условие будет выполняться. Извлекая корень, получаем для l два значения, отличающихся лишь знаком. Мы можем, таким образом, окончательно утверждать, что *группа линейных преобразований* (70) *с определителем единица гомоморфна группе положительных преобразований Лоренца, причем этот гомоморфизм осуществляется формулами* (72). Как и в предыдущем номере, этот гомоморфизм не будет изоморфизмом, т. е. различные преобразования (70) могут приводить к одному и тому же преобразованию Лоренца. Из формул (72) непосредственно вытекает, что тождественное преобразование в группе Лоренца получается от двух линейных преобразований с матрицами

$$E = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}; \quad S = \begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix} = -E,$$

и совершенно так же, как и в предыдущем номере, можно показать, что *всякое преобразование из группы Лоренца может быть получено лишь из двух линейных преобразований* (70), *коэффициенты которых отличаются только знаком.*

Совершенно так же, как и в [**63**], элементы E и $(-E)$ образуют нормальный делитель H группы линейных преобразований с определителем единица, и *группа положительных преобразований Лоренца изоморфна дополнительной к H группе.*

Линейные преобразования (70) содержат четыре комплексных коэффициента, связанных условием (75). Таким образом, формулы (72) содержат три произвольных комплексных параметра, или, иначе говоря, шесть произвольных вещественных параметров.

§ 6. ЛИНЕЙНЫЕ ПРЕДСТАВЛЕНИЯ ГРУПП

65. Представление группы линейными преобразованиями. Пусть G — некоторая группа с элементами G_α, и положим, что каждому элементу G_α соответствует определенная матрица A_α, причем все матрицы A_α имеют один и тот же порядок и их определители отличны от нуля. Положим далее, что это соответствие таково, что всякому произведению $G_{\alpha_2}G_{\alpha_1}$ соответствует матрица $A_{\alpha_2}A_{\alpha_1}$, которая является произведением A_{α_2} и A_{α_1}. В этом случае говорят, что *матрицы A_α или соответствующие им линейные преобразования дают линейное представление группы G.* Пусть G_0 — единичный элемент группы и A_0 — соответствующая матрица. Поскольку $G_0G_\alpha = G_\alpha$, мы должны иметь $A_0A_\alpha = A_\alpha$, откуда, умножая справа на A_α^{-1}, имеем $A_0 = I$, т. е. единичному элементу должна соответствовать единичная матрица. Пусть G_{α_1} и G_{α_2} — обратные элементы и A_{α_1} и A_{α_2} — соответствующие матрицы. Из равенства $G_{\alpha_2} \cdot G_{\alpha_1} = G_0$ следует $A_{\alpha_2}A_{\alpha_1} = I$, т. е. обратным элементам соответствуют и обратные матрицы. Из предыдущего непосредственно следует, что матрицы A_α (или соответствующие линейные преобразования) *образуют группу A, гомоморфную группе G.* Если различным элементам G соответствуют и различные матрицы, то A будет не только гомоморфна, но и изоморфна G. В этом случае говорят, что она дает *биоднозначное линейное представление группы G.* Если это не так, то совокупность элементов G, которым соответствует единичная матрица в A, образует нормальный делитель группы G, и группа A будет изоморфна дополнительной к этому нормальному делителю группе [**57**].

Если основная группа G сама есть группа линейных преобразований, то, очевидно, она сама и дает одно из возможных своих линейных представлений.

Сделаем одно замечание по поводу данного определения линейного представления.

Пусть нам известно, что всякому элементу G_α соответствует определенная матрица A_α, причем произведению элементов соответствует произведение матриц, но неизвестно, будут ли определители матриц A_α отличны от нуля. Покажем, что если один определитель $D(A_{\alpha_0})$ равен нулю, то и все $D(A_\alpha)$ равны нулю. Действительно, совокупность матриц $A_{\alpha_0}A_\alpha$ при переменном α содержит все матрицы, соответствующие элементам группы [**56**]. Но $D(A_{\alpha_0}A_\alpha) = D(A_{\alpha_0})\,D(A_\alpha)$, и произведение равно нулю, так как первый множитель по условию равен нулю. Таким образом, имея закон соответствия, при котором произведению соответствует произведение, нам достаточно проверить, что один из определителей $D(A_\alpha)$ отличен от нуля; например, достаточно проверить, что единичному элементу из G соответствует единичная матрица из A.

Пусть X — некоторая матрица того же порядка, что и матрицы A_α, с определителем, отличным от нуля. Мы имеем:

$$(XA_{\alpha_2}X^{-1})(XA_{\alpha_1}X^{-1}) = XA_{\alpha_2}A_{\alpha_1}X^{-1},$$

и, следовательно, матрицы $XA_\alpha X^{-1}$ также дают линейное представление нашей группы G. Такие *два подобных представления называются обычно эквивалентными представлениями.* Положим, что порядок матриц A_α равен n и что $(x_1, \ldots, x_n)$ суть составляющие вектора в n-мерном пространстве, над которым совершаются преобразования A_α, так что группа A будет

$$\boldsymbol{x}' = A_\alpha \boldsymbol{x}. \tag{77}$$

Эквивалентное линейное представление

$$\boldsymbol{y}' = XA_\alpha X^{-1}\boldsymbol{y}, \tag{78}$$

как мы знаем [**25**], имеет тот смысл, что в упомянутом пространстве вводятся новые оси, причем новые составляющие выражаются через прежние по формулам

$$(y_1, \ldots, y_n) = X(x_1, \ldots, x_n). \tag{79}$$

При этих новых осях линейные преобразования пространства будут уже выражаться по формулам (78), т. е. эквивалентные линейные представления могут быть получены в результате простой замены координатных осей согласно формулам (79). Назовем переменные $(x_1, \ldots, x_n)$, входящие в формулы (77), *объектами линейного представления.* Переход к эквивалентному линейному представлению равносилен, таким образом, замене объектов линейных представлений другими при помощи некоторого линейного преобразования (79) с определителем, отличным от нуля.

Пусть имеется линейное представление группы G при помощи матриц A_α порядка n и другое линейное представление той же группы при помощи матриц B_α порядка m. Составим квазидиагональные матрицы порядка $(n+m)$:

$$[A_\alpha, B_\alpha] = \begin{Vmatrix} A_\alpha & 0 \\ 0 & B_\alpha \end{Vmatrix}. \tag{80}$$

Согласно правилу перемножения квазидиагональных матриц, мы имеем:

$$[A_{\alpha_2}, B_{\alpha_2}][A_{\alpha_1}, B_{\alpha_1}] = [A_{\alpha_2}A_{\alpha_1}, B_{\alpha_2}B_{\alpha_1}].$$

Таким образом, матрицы (80) также дают некоторое линейное представление группы G. Вообще, имея несколько представлений группы G при помощи матриц A_α, B_α, C_α, мы можем составить и новое представление, пользуясь квазидиагональной матрицей

$$D_\alpha = [A_\alpha, B_\alpha, C_\alpha] = \begin{Vmatrix} A_\alpha & 0 & 0 \\ 0 & B_\alpha & 0 \\ 0 & 0 & C_\alpha \end{Vmatrix}. \tag{81}$$

Заметим теперь, что если мы перейдем к эквивалентному представлению при помощи матриц $XD_\alpha X^{-1}$, то квазидиагональный характер матриц, вообще говоря, нарушится, и по виду этого нового представления нельзя уже будет сразу сказать, что оно с точностью до эквивалентного представления составлено из других представлений с меньшим числом измерений по закону (81). Если наше линейное представление D_α имеет чисто квазидиагональный вид (80), то оно распадается на несколько линейных представлений A_α и B_α с меньшим числом измерений, т. е. с матрицами меньшего порядка. В этом случае линейное представление называется *приведенным*. Если некоторое линейное представление E_α не имеет квазидиагонального вида, но некоторое эквивалентное ему представление $XE_\alpha X^{-1}$ имеет такой вид, то представление E_α называется *приводимым*. Наконец, если не только само представление, но и все эквивалентные ему представления не имеют квазидиагонального вида, т. е. не являются приведенными, то такое представление называется *неприводимым представлением*.

Отметим некоторые условия, при наличии которых можно утверждать, что представление будет приводимым. Пусть линейное представление состоит из матриц A_α порядка n, которые дают линейные преобразования с переменными $(x_1, x_2, \ldots, x_n)$. Предположим, что все матрицы A_α — унитарные и что подпространство R', образованное первыми k ортами, переходит само в себя в результате преобразования A_α, т. е. если $x_{k+1} = x_{k+2} = \ldots = x_n = 0$, то и $x'_{k+1} = x'_{k+2} = \ldots = x'_n = 0$. Иначе говоря, все матрицы A_α имеют вид:

$$\begin{Vmatrix} A'_\alpha & N_\alpha \\ 0 & A''_\alpha \end{Vmatrix}, \tag{82}$$

где A'_α — некоторая матрица порядка k, A''_α — некоторая матрица порядка $(n-k)$, и в левом нижнем углу, имеющем $(n-k)$ строк и k столбцов, стоят везде нули. Рассмотрим пространство R'', образованное последними $(n-k)$ ортами. Оно будет состоять из векторов, ортогональных ко всем векторам подпространства R', указанного выше. Так как каждое преобразование A_α переводит подпространство R' в себя и в силу унитарности сохраняет свойство ортогональности векторов, то всякий вектор подпространства R'' должен в результате преобразования A_α перейти в вектор, также принадлежащий этому подпространству. Иначе говоря, если $x_1 = \ldots = x_k = 0$, то и $x'_1 = \ldots = x'_k = 0$. Отсюда непосредственно следует, что в матрицах (82) и все элементы, стоящие в правом верхнем углу, имеющем k строк и $(n-k)$ столбцов, также должны быть все равны нулю, т. е. матрицы рассматриваемого линейного представления будут:

$$\begin{Vmatrix} A'_\alpha & 0 \\ 0 & A''_\alpha \end{Vmatrix} = [A'_\alpha, A''_\alpha],$$

и, следовательно, представление будет приведенное. Положим теперь, что все унитарные преобразования A_α оставляют вообще инвариантным некоторое подпространство R_1, измерения $k (k < n)$, где n — порядок матриц A_α. Преобразуем координатные оси так, чтобы подпространство R_1 было образовано первыми k ортами, что равносильно переходу к эквивалентному линейному представлению и может быть осуществлено при помощи унитарного преобразования. После такого преобразования в силу предыдущего представление окажется приведенным. Мы имеем, таким образом, следующую теорему:

Теорема *Если линейное представление группы состоит из унитарных матриц, и эти матрицы оставляют неизменным некоторое подпространство, то такое представление есть приводимое представление.*

Вопрос о приводимости или неприводимости представления тесно связан с вопросом о переходе от матриц A_α к подобным матрицам $XA_\alpha X^{-1}$. Отметим некоторые частные случаи перехода к эквивалентным представлениям, которые получаются при специальном выборе матрицы X. Построим матрицу X, у которой в первой строке на втором месте стоит единица, а на других местах нули; во второй строке стоит на первом месте единица, а на остальных местах нули, и, начиная с третьей строки, стоят на главной диагонали единицы, а на остальных местах нули, т. е.

$$X = \begin{Vmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{Vmatrix}.$$

Непосредственно разлагая, начиная с последней строки, увидим, что $D(X) = -1$. Применяя обычные правила умножения матриц, нетрудно проверить, что если Y есть некоторая матрица, то подобная матрица XYX^{-1} будет получаться из Y взаимной перестановкой первой и второй строки а также первого и второго столбца. Точно так же всякая перестановка строк, сопровождаемая такой же перестановкой столбцов, равносильна переходу к некоторой подобной матрице с помощью преобразования X, которое не зависит, очевидно, от матрицы Y. Таким образом, *если мы во всех матрицах A_α, дающих некоторое линейное представление группы, совершим одну и ту же перестановку строк и столбцов, то это будет равносильно переходу к эквивалентному представлению.*

Если можно распределить целые числа $1, 2, \ldots, n$ на два класса так, что на пересечении любой строки каждой матрицы A_α с номером из одного класса с любым столбцом с номером из другого класса стоит нуль, то такое представление будет приводимым. Действительно, чтобы совершить его приведение, т. е. сделать его приведенным, достаточно переставить строки и столбцы так, чтобы строки и столбцы одного класса стояли сплошь сверху и слева, а строки и столбцы другого класса — снизу и справа.

В заключение настоящего номера отметим еще тот случай, когда линейное представление группы G будет первого порядка, т. е. когда все матрицы A_α будут матрицами первого порядка, иначе говоря, обыкновенными числами. В этом случае каждому элементу G_α нашей группы соответствует преобразование $x' = m_\alpha x$, или, проще говоря, число m_α, и произведению G_2G_1 соответствует обычное произведение чисел m_2m_1.

66. Основные теоремы. Пусть имеется конечная группа G, содержащая m элементов $G_1, \ldots, G_m$, и пусть $A_1, \ldots, A_m$ — матрицы некоторого порядка n, которые дают линейное представление этой группы. Объекты этого представления обозначим через $\boldsymbol{x}(x_1, \ldots, x_n)$. Рассмотрим выражение:

$$\varphi(x_1, \ldots, x_n) = \sum_{s=1}^{m} |A_s \boldsymbol{x}|^2. \tag{83}$$

В раскрытом виде это будет выражение:

$$\varphi = \sum_{s=1}^{m} \sum_{i=1}^{n} (a_{i1}^{(s)} x_1 + \ldots + a_{in}^{(s)} x_n)(\bar{a}_{i1}^{(s)} \bar{x}_1 + \ldots + \bar{a}_{in}^{(s)} \bar{x}_n), \tag{84}$$

где через $a_{ik}^{(s)}$ мы обозначили элементы матрицы A_s. Нетрудно проверить, что выражение (84) будет формой Эрмита, т. е. в этом выражении коэффициенты при $\bar{x}_p x_q$ и $x_p \bar{x}_q$ суть комплексные сопряженные числа. Кроме того, из формулы (83) следует, что эта форма Эрмита представляет собою сумму квадратов длин некоторых векторов, т. е. это будет определенно положительная форма Эрмита [**40**]. Иначе говоря, совершая некоторое унитарное преобразование

$$\boldsymbol{y} = U\boldsymbol{x},$$

приводящее нашу форму к сумме квадратов

$$\varphi = \sum_{j=1}^{n} \lambda_j \bar{y}_j y_j,$$

мы будем иметь все коэффициенты λ_j положительными. Совершая еще преобразование $z_j = \sqrt{\lambda_j} y_j$, в новых переменных будем иметь для нашей формы Эрмита φ выражение в виде суммы чистых квадратов:

$$\varphi = \bar{z}_1 z_1 + \ldots + \bar{z}_n z_n. \tag{85}$$

Совершим над переменными $(x_1, \ldots, x_n)$ некоторое преобразование

$$x' = A_k x, \tag{86}$$

принадлежащее к линейному представлению нашей группы. Нетрудно видеть, что при этом форма Эрмита φ не изменится. Действительно:

$$\varphi(x_1', \ldots, x_n') = \sum_{s=1}^{m}{}' \, | A_s A_k x |^2.$$

Но, как мы знаем [**56**], совокупность преобразований (матриц)

$$A_1 A_k, \; A_2 A_k, \; \ldots, \; A_m A_k$$

совпадает с совокупностью матриц

$$A_1, \; A_2, \; \ldots, \; A_m;$$

поэтому, если мы выразим преобразование (86) в новых переменных $(z_1, \ldots, z_n)$, которые связаны с прежними формулами вида:

$$(z_1, \ldots, z_n) = B_0 (x_1, \ldots, x_n),$$

где B_0 — некоторая матрица, то вместо группы A_k получим подобную группу $B_0 A_k B_0^{-1}$, и все преобразования этой подобной группы не будут менять выражения (85), т. е. не будут менять суммы квадратов модулей, и, следовательно, будут все унитарными преобразованиями. Мы показали, таким образом, для случая конечных групп, что всякое линейное представление эквивалентно некоторому унитарному представлению, т. е. представлению, состоящему из унитарных преобразований. При некоторых дополнительных условиях это свойство сохраняется и при линейных представлениях бесконечных групп, зависящих от параметров, и в дальнейшем, когда будем говорить о линейном представлении группы, мы будем подразумевать всегда унитарное представление. Мы имеем, таким образом, следующую теорему:

Теорема I. *Всякое линейное представление группы (конечной) имеет эквивалентное унитарное представление.*

Выведем теперь необходимое и достаточное условие приводимости линейного представления. Введем предварительно один новый термин, а именно назовем диагональную матрицу $[k, \ldots, k]$, содержащую на диагонали одинаковые элементы, *кратной* единичной матрице. *Такую матрицу можно обозначить в виде kI.* Как мы видели выше, в отношении алгебраических операций она эквивалентна числу k.

Положим теперь, что у нас имеется приводимое линейное представление некоторой группы. Такое представление осуществляется, например, матрицами вида:

$$D_a = X[A_a, B_a, C_a]X^{-1},$$

где X — некоторая матрица и внутренняя матрица квазидиагональна. Составим матрицу

$$Y = X[kI, lI, mI]X^{-1},$$

где средний квазидиагональный член имеет ту же конструкцию, что и в матрицах D_a. Нетрудно видеть, что матрица Y коммутирует со всеми матрицами D_a. Действительно:

$$D_aY = X[A_ak, B_al, C_am]X^{-1}$$

и точно так же

$$YD_a = X[kA_a, lB_a, mC_a]X^{-1}.$$

Но при умножении любой матрицы на число порядок множителей не играет роли. Кроме того, если числа k, l и m различны, что мы и предполагаем, то матрица Y не кратна единичной матрице. Действительно, она имеет, очевидно, различные характеристические числа k, l и m. Таким образом мы приходим к следующей теореме:

Теорема II. *Если линейное представление приводимо, то существует матрица, отличная от кратной единичной матрицы и коммутирующая со всеми матрицами, входящими в упомянутое приводимое линейное представление.*

Покажем теперь, что имеет место и обратная теорема, т. е.

Теорема III. *Если существует матрица Y, отличная от кратной единичной матрицы и коммутирующая со всеми матрицами D_a линейного представления, то такое линейное представление будет приводимым.*

Итак, по условию теоремы мы имеем для любого значка a:

$$D_aY = YD_a. \tag{87}$$

Пусть Z — такая матрица, с определителем, отличным от нуля, что все матрицы ZD_aZ^{-1} унитарны: $ZD_aZ^{-1} = U_a$. Перепишем предыдущее равенство в виде:

$$Z^{-1}U_aZY = YZ^{-1}U_aZ.$$

Умножая слева на Z и справа на Z^{-1}, получим:

$$U_a(ZYZ^{-1}) = (ZYZ^{-1})U_a,$$

т. е. матрица ZYZ^{-1} коммутирует со всеми матрицами унитарного представления. Эта матрица, очевидно, не кратна единичной, ибо если $ZYZ^{-1} = kI$, то и $Y = kI$. Нам достаточно доказать приводимость эквивалентного линейного представления U_a. Таким образом доказательство нашей теоремы свелось к тому случаю, когда линейное представление унитарно. Мы будем для простоты письма считать, что само линейное представление, состоящее из матриц D_a, унитарно.

Пусть λ_1 — некоторое характеристическое число матрицы Y. Матрица $\lambda_1 = \lambda_1 I$, как известно [**25**], коммутирует с любой матрицей и, следовательно, матрица $Y - \lambda_1 I$ так же, как и Y, удовлетворяет условиям (87), т. е. коммутирует со всеми матрицами D_a. Нетрудно видеть, что по крайней мере одно из характеристических чисел матрицы $Y_1 = Y - \lambda_1 I$ будет равно нулю. Действительно, характеристическое уравнение для матрицы Y_1 будет:

$$D(Y_1 - \lambda I) = D[Y - (\lambda + \lambda_1) I] = 0,$$

т. е. оно получается из характеристического уравнения для Y заменой λ на $(\lambda + \lambda_1)$, и так как среди характеристических чисел матрицы Y было характеристическое число, равное λ_1, то среди характеристических чисел Y_1 будет хоть одно равное нулю. Следовательно, определитель матрицы Y_1, равный произведению ее характеристических чисел, также будет равен нулю. Мы можем, таким образом, при доказательстве нашей теоремы предполагать все матрицы D_a унитарными, и определитель матрицы Y, входящей в формулы (87), равным нулю.

Рассмотрим совокупность векторов, имеющих составляющие:

$$\left.\begin{aligned} x_1 &= y_{11} u_1 + y_{12} u_2 + \ldots + y_{1n} u_n, \\ x_2 &= y_{21} u_1 + y_{22} u_2 + \ldots + y_{2n} u_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= y_{n1} u_1 + y_{n2} u_2 + \ldots + y_{nn} u_n, \end{aligned}\right\} \qquad (88)$$

где u_s принимают любые значения и где y_{ik} — элементы матрицы Y. Поскольку определитель матрицы Y равен нулю, ранг таблицы $\| y_{jk} \|$ будет меньше n. Пусть он равен некоторому числу $r < n$. В этом случае, как мы знаем [**16**], формулы (88) определяют некоторое подпространство R' измерения r.

Рассмотрим левую часть равенства:

$$D_a Y\boldsymbol{u} = Y D_a \boldsymbol{u}. \qquad (89)$$

Вектор $Y\boldsymbol{u}$ имеет как раз составляющие (88) и, следовательно, $D_a Y\boldsymbol{u}$ есть результат применения преобразования D_a к некоторому произвольному вектору из подпространства R'. В правой части формулы (89) мы имеем результат применения преобразования Y к вектору $D_a \boldsymbol{u}$, т. е. составляющие правой части выражаются тоже по формулам (88), где только вместо $u_1, \ldots, u_n$ поставлены составляющие векторы $D_a \boldsymbol{u}$, т. е. правая часть формулы (89) представляет собою вектор, принадлежащий подпространству R'. Таким образом мы видим, сравнивая левую часть с правой, что применение преобразования D_a к любому вектору из подпространства R' дает вектор, также принадлежащий этому подпространству. Но, как мы знаем [**65**], если унитарные преобразования оставляют инвариантным некоторое подпространство, то они образуют приводимое представление, и таким образом теорема доказана.

Теоремы II и III показывают, что *необходимым и достаточным условием неприводимости линейного представления является тот факт, что не существует матрицы, отличной от матрицы вида kI, которая коммутировала бы со всеми матрицами, входящими в линейное представление.*

Из теоремы I непосредственно следует, что в теореме из [**65**] нет необходимости упоминать об унитарности представления, и можно вообще утверждать, что если все матрицы некоторого представления оставляют неизменным некоторое подпространство, то такое представление приводимо. Очевидно и обратное утверждение.

67. Абелевы группы и представления первого порядка. Группа G называется *абелевой*, если все ее элементы попарно коммутируют, т. е. если при любых значках $G_{a_2}G_{a_1}=G_{a_1}G_{a_2}$ [**56**]. Пусть A_{a_1} и A_{a_2} — матрицы, соответствующие G_{a_1} и G_{a_2} в некотором линейном представлении. Произведению $G_{a_2}G_{a_1}$ соответствует $A_{a_2}A_{a_1}$, а произведению $G_{a_1}G_{a_2}$ соответствует $A_{a_1}A_{a_2}$. Но упомянутые произведения совпадают, и, следовательно, мы должны иметь:

$$A_{a_2}A_{a_1}=A_{a_1}A_{a_2},$$

т. е. *матрицы, образующие линейное представление абелевой группы, попарно коммутируют.*

Положим, что представление унитарно, т. е. что все матрицы унитарны. При этом, как известно, существует такое унитарное преобразование U, что все матрицы UA_aU^{-1} имеют чисто диагональную форму [**42**], т. е. в данном случае некоторое эквивалентное линейное представление состоит из чисто диагональных матриц

$$UA_aU^{-1}=[k_1^{(a)}, \ldots, k_n^{(a)}].$$

Мы видим, таким образом, что в данном случае линейное представление распадается на n представлений первого порядка

$$B_a^{(s)}=k_s^{(a)} \quad (s=1, 2, \ldots, n)$$

Итак, *всякое унитарное представление абелевой группы эквивалентно некоторой совокупности представлений первого порядка, причем переход к эквивалентному представлению совершается также с помощью унитарного преобразования.*

Рассмотрим теперь ряд примеров представлений абелевых групп, а также некоторые примеры линейных представлений первого порядка и не абелевых групп.

Пример 1. В качестве первого примера рассмотрим циклическую (абелеву) группу порядка m, состоящую из элементов

$$S^0 = I,\ S,\ S^2,\ \ldots,\ S^{m-1} \quad (S^m = I). \tag{90}$$

Если элементу S соответствует линейное преобразование $x' = \omega x$, или, что то же, число ω, то элементам (90) будут соответствовать следующие числа:

$$1,\ \omega,\ \omega^2,\ \ldots,\ \omega^{m-1}.$$

Поскольку $S^m = I$, мы должны иметь $\omega^m = 1$, т. е.

$$\omega = e^{\frac{2k\pi i}{m}},$$

где k — некоторое целое число, которое мы можем, очевидно, принимать равным одному из чисел ряда $0, 1, 2, \ldots, m-1$.

Рассмотрим подробнее случай $m = 2$. При этом будет:

$$I,\ S \text{ и } S^2 = I,$$

т. е. $S = S^{-1}$. При $k = 0$ обоим элементам I и S соответствует одно и то же тождественное преобразование $x' = x$ или число 1; при $k = 1$ элементу I соответствует преобразование $x' = x$, а элементу S — преобразование $x' = -x$, или, проще говоря, элементу I — число 1, а элементу S — число (-1). В физических применениях представляется важным тот случай, когда группа состоит из тождественного преобразования трехмерного пространства и преобразования симметрии относительно начала:

$$x' = -x;\ y' = -y;\ z' = -z\ (S).$$

Мы имеем, очевидно, $m = 2$. Указанные выше два представления можно назвать тождественным и знакопеременным представлением симметрии относительно начала.

Пример 2. Рассмотрим группу вращения вокруг оси Z. Матрицы этой группы имеют вид:

$$Z_\varphi = \begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix} \tag{91}$$

и, кроме того, как мы видели раньше, удовлетворяют очевидному соотношению

$$Z_{\varphi_2} Z_{\varphi_1} = Z_{\varphi_1} Z_{\varphi_2} = Z_{\varphi_1 + \varphi_2}.$$

Такому же соотношению удовлетворяет также функция $e^{l\varphi}$.

Но надо заметить, что если $\varphi = 2\pi$, то поворот равносилен тождественному преобразованию, и, следовательно, мы должны иметь $e^{2\pi l} = 1$, т. е число l должно быть вида $l = ki$, где k — любое целое число Мы имеем, таким образом, бесчисленное множество линейных представлений нашей группы вращения, причем матрице (91) соответствует число $e^{\varphi k i}$.

Придавая целому числу k всевозможные значения

$$k = 0,\ \pm 1,\ \pm 2,\ \ldots,$$

мы и будем иметь бесчисленное множество линейных представлений группы вращения.

Пример 3. Рассмотрим теперь группу, состоящую из $n!$ перестановок над n элементами. Мы можем каждой перестановке сопоставить число $(+1)$. Так получится то, что называется симметрическим представлением группы перестановок. Иначе, мы можем всякой перестановке первого класса, состоящей из четного числа транспозиций, сопоставить число $(+1)$, а всякой перестановке второго класса — число (-1). Таким образом получится то, что называется антисимметрическим представлением группы перестановок. В этом представлении всякой перестановке из знакопеременной подгруппы соответствует число $(+1)$, а остальным перестановкам соответствует число (-1). Можно показать, на чем мы не останавливаемся, что указанными двумя случаями и исчерпываются все возможности линейных представлений первого порядка для группы перестановок. Эта группа имеет и другие представления выше первого порядка.

Пример 4. Рассмотрим теперь группу всех вещественных ортогональных преобразований на плоскости, т. е. группу, образованную вращениями плоскости вокруг начала, соединенными с симметрией относительно оси Y. Как мы видели выше [**52**], матрицы этой группы будут иметь вид:

$$\{\varphi,\ d\} = \begin{Vmatrix} d\cos\varphi & -d\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix}, \tag{92}$$

где $d = 1$ для чистого вращения и $d = -1$ для вращения, соединенного с симметрией. Кроме очевидного линейного представления первого порядка, при котором каждой матрице (92) соответствует число $(+1)$, мы можем построить еще линейное представление первого порядка, при котором матрице (92) соответствует число $(+1)$, если $d = 1$, и число (-1), если $d = -1$. Это даст нам действительно линейное представление, так как произведение двух матриц вида (92) соответствует чистому вращению, если d имеет одинаковые знаки, и вращению с симметрией, если d имеет разные знаки в перемножаемых матрицах.

68. Линейные представления унитарной группы с двумя переменными. Рассмотрим линейные представления унитарной группы с двумя переменными. Эта группа, как мы знаем, имеет вид:

$$\left.\begin{aligned} x_1' &= ax_1 + bx_2 \\ x_2' &= -\bar{b}x_1 + \bar{a}x_2 \end{aligned}\right\} \tag{93}$$

где комплексные числа a и b должны подчиняться условию

$$a\bar{a}+b\bar{b}=1. \tag{94}$$

Построим $(m+1)$ величин:

$$\xi_0=x_1^m;\ \xi_1=x_1^{m-1}x_2;\ \ldots;\ \xi_m=x_2^m. \tag{95}$$

Если мы возьмем $\xi'_k=x_1'^{m-k}x_2'^k$ и подставим вместо x'_1 и x'_2 их выражения (93), то, очевидно, каждое ξ'_k выразится линейно через ξ_k и, следовательно, всякому преобразованию из группы (93) будет соответствовать линейное преобразование от переменных ξ к переменным ξ'_k. Очевидно, что произведению преобразований будет соответствовать произведение преобразований, и мы будем иметь, таким образом, линейное представление группы (93) порядка $(m+1)$. Но, как оказывается, это представление не будет унитарным. Чтобы построить унитарное представление, достаточно в каждое из переменных (95) ввести некоторый дополнительный постоянный множитель, и именно вместо формул (95) мы определим переменные по следующим формулам:

$$\eta_k=\frac{x_1^{m-k}x_2^k}{\sqrt{(m-k)!k!}}\qquad (k=0,\ 1,\ \ldots,\ m) \tag{96$_1$}$$

и точно так же:

$$\eta'_k=\frac{x_1'^{m-k}x_2'^k}{\sqrt{(m-k)!\,k!}}\qquad (k=0,\ 1,\ \ldots,\ m), \tag{96$_2$}$$

где, как всегда, считается $0!=1$.

Проверим, что при таком определении переменных наше представление будет унитарным, т. е. что

$$\sum_{k=0}^m \eta'_k\bar{\eta}'_k=\sum_{k=0}^m \eta_k\bar{\eta}_k. \tag{97}$$

Действительно, применяя формулу бинома Ньютона, имеем:

$$m!\cdot\sum_{k=1}^m \eta'_k\bar{\eta}'_k=m!\sum_{k=0}^m \frac{x_1'^{m-k}\overline{x_1'^{m-k}}\,x_2'^k\overline{x_2'^k}}{(m-k)!\,k!}=(x'_1\bar{x}'_1+x'_2\bar{x}'_2)^m,$$

и точно так же

$$m!\cdot\sum_{k=0}^m \eta_k\bar{\eta}_k=(x_1\bar{x}_1+x_2\bar{x}_2)^m.$$

Но, в силу унитарности преобразования (93),

$$x'_1\bar{x}'_1+x'_2\bar{x}'_2=x_1\bar{x}_1+x_2\bar{x}_2,$$

и, следовательно, имеет место и соотношение (97).

Выведем теперь формулы, которые дают в явном виде коэффициенты построенного унитарного представления группы (93). С этой целью несколько изменим предыдущее обозначение, а именно положим:

$$\eta_l = \frac{x_1^{j+l} x_2^{j-l}}{\sqrt{(j+l)!\,(j-l)!}} \quad (l = -j, -j+1, \ldots, j-1, j). \quad (98)$$

В наших прежних обозначениях $m = 2j$, и число j будет целым, если m — четно, и будет равно половине целого нечетного числа, если m — нечетное число. Если, например, $m = 5$, то формулы (98) дают нам следующие шесть переменных:

$$\eta_{-\frac{5}{2}} = \frac{x_2^5}{\sqrt{5!}}; \quad \eta_{-\frac{3}{2}} = \frac{x_1 x_2^4}{\sqrt{1!\,4!}}; \quad \eta_{-\frac{1}{2}} = \frac{x_1^2 x_2^3}{\sqrt{2!\,3!}};$$

$$\eta_{\frac{1}{2}} = \frac{x_1^3 x_2^2}{\sqrt{3!\,2!}}; \quad \eta_{\frac{3}{2}} = \frac{x_1^4 x_2}{\sqrt{4!\,1!}}; \quad \eta_{\frac{5}{2}} = \frac{x_1^5}{\sqrt{5!}}.$$

В данном случае наши переменные пронумерованы не первыми шестью целыми числами, а дробными числами, которые отличаются друг от друга на единицу и идут от $\left(-\frac{5}{2}\right)$ до $\left(+\frac{5}{2}\right)$. Если, например, $m = 4$, то мы имеем по формулам (98) пять переменных:

$$\eta_{-2} = \frac{x_2^4}{\sqrt{4!}}; \quad \eta_{-1} = \frac{x_1 x_2^3}{\sqrt{1!\,3!}}; \quad \eta_0 = \frac{x_1^2 x_2^2}{\sqrt{2!\,2!}};$$

$$\eta_1 = \frac{x_1^3 x_2}{\sqrt{3!\,1!}}; \quad \eta_2 = \frac{x_1^4}{\sqrt{4!}}.$$

Здесь нумерация переменных совершается целыми числами от (-2) до $(+2)$. При всяком фиксированном $m = 2j$ мы будем иметь совершенно такую же нумерацию строк и столбцов в матрицах, которые будут давать линейное представление порядка $(2j+1)$ группы (93).

Перейдем теперь к определению элементов этих матриц. Мы имеем:

$$\eta_l' = \frac{x_1'^{j+l} x_2'^{j-l}}{\sqrt{(j+l)!\,(j-l)!}} = \frac{(ax_1 + bx_2)^{j+l}(-\bar{b}x_1 + \bar{a}x_2)^{j-l}}{\sqrt{(j+l)!\,(j-l)!}},$$

и нам надо представить правую часть в виде линейной комбинации величин η_l. Применение формулы бинома Ньютона дает:

$$\eta' = \sum_{k=0}^{j+l} \sum_{k'=0}^{j-l} (-1)^{j-l-k'} \frac{\sqrt{(j+l)!\,(j-l)!}}{k!\,k'!\,(j+l-k)!\,(j-l-k')!} \times$$

$$\times \bar{a}^{k'} a^{j+l-k} \bar{b}^{j-l-k'} b^k x_1^{2j-k-k'} x_2^{k+k'}.$$

Если мы будем считать $p! = \infty$, когда p есть целое отрицательное число, то можем в предыдущей формуле производить суммирование по k и k' от $(-\infty)$ до $(+\infty)$, ибо лишние слагаемые будут содержать в знаменателе множитель, равный бесконечности, и обратятся в нуль. Введем вместо k' новую переменную суммирования $s = j - k - k'$, по которой тоже можно производить суммирование от $(-\infty)$ до $(+\infty)$ по целым значениям или по половинам целых значений, смотря по тому — будет ли j целым числом или половиной целого числа. Мы получаем таким образом

$$\eta_l' = \sum_k \sum_s (-1)^{k+s-l} \frac{\sqrt{(j+l)!\,(j-l)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times \bar{a}^{j-k-s} a^{j+l-k} \bar{b}^{k+s-l} b^k x_1^{j+s} x_2^{j-s}.$$

Но, согласно (98), мы имеем:

$$x_1^{j+s} x_2^{j-s} = \sqrt{(j+s)!\,(j-s)!}\,\eta_s,$$

и окончательно получаем искомую линейную зависимость в следующем виде:

$$\eta_l' = \sum_k \sum_s (-1)^{k+s-l} \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times \bar{a}^{j-k-s} a^{j+l-k} \bar{b}^{k+s-l} b^k \eta_s.$$

Таким образом, при заданном фиксированном j элементы матрицы линейного преобразования порядка $(2j+1)$, соответствующего унитарному преобразованию (93) с матрицей

$$\left\| \begin{matrix} a & b \\ -\bar{b} & \bar{a} \end{matrix} \right\|,$$

будут:

$$D_j \begin{Bmatrix} a, & b \\ -\bar{b}, & \bar{a} \end{Bmatrix}_{ls} =$$
$$= (-1)^{s-l} \sum_k (-1)^k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times \bar{a}^{j-k-s} a^{j+l-k} \bar{b}^{k+s-l} b^k. \quad (99)$$

Здесь значки l и s пробегают следующий ряд значений:

$$l,\ s = -j,\ -j+1,\ \ldots,\ j-1,\ j,$$

причем напомним еще раз, что если j есть половина целого числа, то это дает нумерацию строк и столбцов матриц также по половинам целых чисел. Принимая во внимание, что $p! = \infty$, если p есть целое отрицательное число, мы получаем следующие пределы суммирования по k в формулах (99):

$$k \geqslant 0;\ k \geqslant l - s;\ k \leqslant j - s;\ k \leqslant j + l. \quad (100)$$

Отметим некоторое упрощение в формулах (99), которого можно достигнуть, переходя к подобному представлению. Пусть A — некоторая матрица с элементами a_{pq} и пусть $S=[\delta_1, \ldots, \delta_n]$ — диагональная матрица.

Применяя обычное правило умножения, нетрудно проверить, что матрица SAS^{-1} будет иметь следующие элементы:

$$\{SAS^{-1}\}_{pq}=\delta_p a_{pq}\delta_q^{-1}.$$

Если мы применим теперь это правило для матрицы

$$D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a}\end{Bmatrix}$$

и примем $\delta_k=(-1)^k$, то в формулах (99) исчезнет множитель $(-1)^{s-l}$, и в дальнейшем мы будем считать, что этот множитель отсутствует.

Перейдем теперь к доказательству того, что *линейное представление унитарной группы* (93), *определяемое матрицей с элементами* (99), *будет неприводимым.* Предварительно докажем две леммы.

Лемма I. *Если некоторая диагональная матрица, все диагональные элементы которой попарно различны, коммутирует с матрицей A, то A есть также диагональная матрица.*

По условию мы имеем:

$$A[\delta_1, \ldots, \delta_n]=[\delta_1, \ldots, \delta_n]A,$$

где числа δ_k попарно различны. Пусть a_{pq} — элементы матрицы A. Применяя правило умножения, мы получим из предыдущего условия:

$$a_{pq}\delta_q=\delta_p a_{pq} \quad \text{или} \quad a_{pq}(\delta_q-\delta_p)=0,$$

и, следовательно, $a_{pq}=0$, если $p\neq q$, т. е. матрица A — действительно диагональная матрица.

Лемма II. *Если некоторая диагональная матрица* $[\delta_1, \ldots, \delta_n]$ *коммутирует с матрицей A, в которой по крайней мере один столбец не содержит ни одного нуля, то* $\delta_1=\ldots=\delta_n$.

Переставляя строки и столбцы, т. е. переходя к подобным матрицам, мы можем достигнуть того, чтобы столбец, не содержащий нулей, стоял на первом месте. При этом диагональная матрица остается по-прежнему диагональной, и наши матрицы по-прежнему будут коммутировать. Таким образом, мы можем считать, обозначая через a_{pq} элементы матрицы A, что

$$a_{i1}\neq 0 \qquad (i=1, 2, \ldots, n),$$

и, кроме того, по условию, как и выше:

$$a_{i1}(\delta_1-\delta_i)=0 \qquad (i=1, 2, \ldots, n),$$

откуда имеем $\delta_1=\ldots=\delta_n$, и, следовательно, лемма доказана.

Переходим теперь к доказательству неприводимости линейного представления, определяемого матрицами (99). Пусть Y есть некоторая матрица порядка $(2j+1)$, которая коммутирует со всеми матрицами

$$D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a}\end{Bmatrix},$$

получающимися при различных a и b, удовлетворяющих условию (94). Для доказательства неприводимости нам надо показать, что Y должна быть кратной единичной матрице. Рассмотрим сначала тот случай, когда $b=0$ и $a=e^{i\alpha}$. Эти комплексные числа удовлетворяют, очевидно, условию (94).

Пользуясь формулами (99), мы получаем прежде всего, что при этом

$$D_j\begin{Bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha}\end{Bmatrix}_{ls}=0 \text{ при } l\neq s,$$

а диагональные элементы будут в данном случае

$$D_j\begin{Bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha}\end{Bmatrix}_{ll}=e^{i2l\alpha} \quad (l=-j, -j+1, \ldots, j-1, j),$$

и наша матрица будет иметь вид:

$$D_j\begin{Bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha}\end{Bmatrix}=\begin{Vmatrix} e^{-i2j\alpha} & 0 & 0 & \ldots & 0 \\ 0 & e^{-i2(j-1)\alpha} & 0 & \ldots & 0 \\ 0 & 0 & e^{-i2(j-2)\alpha} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & e^{i2j\alpha}\end{Vmatrix}, \tag{101}$$

т. е. это диагональная матрица с различными элементами на главной диагонали, при подходящем выборе α. Пользуясь первой леммой, мы можем утверждать, что матрица Y, которая должна коммутировать и с матрицей (101), также должна быть диагональной матрицей, т. е.

$$Y=[\delta_1, \ldots, \delta_n]. \tag{102}$$

Рассмотрим теперь тот случай, когда числа a и b оба отличны от нуля, и возьмем первый столбец матрицы $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a}\end{Bmatrix}$. Его элементы определятся по формулам (99), если мы там положим $s=-j$. Неравенства (100) дают при этом:

$$k\geqslant 0;\ k\geqslant l+j;\ k\leqslant 2j;\ k\leqslant j+l \quad (l=-j, -j+1, \ldots, j-1, j),$$

откуда видно, что в данном случае вся сумма, входящая в формулу (99), приведется к одному слагаемому, которое получится при $k=j+l$ и будет отлично от нуля. Таким образом, в данном случае действительно первый столбец матрицы $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$ не содержит нулей. Но раз диагональная матрица (102) должна коммутировать и с такой матрицей, то, согласно лемме II, все числа δ_k одинаковы, т. е. Y кратна единичной матрице. Итак, матрицы $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$ дают, действительно, неприводимое линейное представление унитарной группы (93). Придавая j ряд значений

$$j = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ 2,\ \ldots$$

мы получаем бесчисленное множество этих линейных представлений. При $j=0$ получается тривиальное тождественное представление, при котором всякому элементу группы (93) соответствует число единица. Рассмотрим теперь при $j>0$, каким преобразованиям группы (93) соответствует тождественное преобразование в группе представлений $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$, которое определяется равенствами $\eta_l' = \eta_l$ или, что то же, равенствами

$$(ax_1 + bx_2)^{j+l}(-\bar{b}x_1 + \bar{a}x_2)^{j-l} = x_1^{j+l}x_2^{j-l}$$
$$(l = -j,\ -j+1,\ \ldots,\ j-1,\ j).$$

Полагая $j=l$, получаем $(ax_1 + bx_2)^{2j} = x_1^{2j}$, откуда следует, что $b=0$, и предыдущее тождество записывается в виде:

$$a^{j+l}\bar{a}^{j-l}x_1^{j+l}x_2^{j-l} = x_1^{j+l}x_2^{j-l} \qquad (l = -j,\ -j+1,\ \ldots,\ j-1,\ j)$$

откуда $a^{j+l}\bar{a}^{j-l} = 1$. Но $|a| = 1$ при $b=0$, и последнее равенство переписывается в виде:

$$a^{2l} = 1 \qquad (l = -j,\ -j+1,\ \ldots,\ j-1,\ j).$$

Если j — половина нечетного числа, то мы можем положить $l = \frac{1}{2}$, что дает $a=1$. Если j — целое число, то равенства $a^{2l}=1$ сводятся к одному $a^2 = 1$, откуда $a = \pm 1$.

Таким образом, если j — половина нечетного числа, то тождественное преобразование в группе $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$ соответствует только тождественному преобразованию в группе (93), т. е. в этом случае $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$ будет биоднозначным представлением группы (93). Если же j — число целое, то тождественному преобразованию в группе $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ соответствуют в группе (93) два преобразования с матрицами

$$E=\begin{Vmatrix}1 & 0\\ 0 & 1\end{Vmatrix};\ S=\begin{Vmatrix}-1 & 0\\ 0 & -1\end{Vmatrix}=-E.$$

Эти преобразования образуют циклическую группу второго порядка, и $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ является биоднозначным представлением дополнительной группы [**58**]. Иначе можно сказать, что всякому преобразованию в представлении $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ при целом j соответствуют два преобразования из группы (93), у которых числа a и b отличаются лишь знаком.

69. Линейные представления группы вращения. Предыдущие результаты представляются особенно важными потому, что унитарная группа (93) тесно связана с группой вращения трехмерного пространства, и полученный выше результат приводит нас к неприводимым линейным представлениям группы вращения.

Всякому унитарному преобразованию (93) соответствует определенное вращение, причем одновременная перемена знака у a и b дает унитарное преобразование, которому соответствует то же самое вращение. Параметры a и b связаны с углами Эйлера соответствующего вращения формулами [**63**]:

$$a=e^{-\frac{1}{2}i(\gamma+\alpha)}\cos\frac{1}{2}\beta;\quad b=-ie^{\frac{1}{2}i(\gamma-\alpha)}\sin\frac{1}{2}\beta. \tag{103}$$

Рассмотрим сначала тот случай, когда j есть целое число. В этом случае формулы (99) показывают, что одновременная перемена знаков у a и b не меняет слагаемых, стоящих в правой части, так как сумма показателей у a, $\bar{a}$ b и $\bar{b}$ будет в данном случае равна четному числу $2j$. Таким образом, при этом тем двум унитарным преобразованиям, которые дают одно и то же вращение, соответствует одинаковая матрица в линейном представлении. Иначе говоря, при целом j каждому вращению с углами Эйлера $\{\alpha, \beta, \gamma\}$ соответствует определенная матрица в линейном представлении D_j. Вместо $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ обозначим теперь эту матрицу через

$$D_j\{\alpha,\ \beta,\ \gamma\}. \tag{104}$$

Если j есть половина целого числа, то одновременная перемена знаков у a и b приводит к перемене знака и у всех выражений (99), т. е. в данном случае тем унитарным преобразованиям, которые приводят к одному и тому же движению, соответствуют различные матрицы, а именно матрицы, все элементы которых отличаются знаком. В данном случае каждому вращению будут соответствовать также две матрицы, отличающиеся знаком, т. е. в данном случае в (104) перед D_j мы должны поставить два знака. Таким образом, при целом j матрицы (104) дают линейное представление группы вращения. При j, равном половине целого числа, мы не получаем, точно говоря, линейного представления. В данном случае говорят о двузначном линейном представлении.

Чтобы найти выражение элементов матриц (104), достаточно в выражениях (99) заменить a и b по формулам (103). Мы получаем, отбрасывая предварительно множитель $(-1)^{s-l}$:

$$D_j\{\alpha, \beta, \gamma\}_{ls} = i^{s-l}\sum_k (-1)^k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$

$$\times e^{-il\alpha - is\gamma}\cos^{2j+l-2k-s}\frac{1}{2}\beta\,\sin^{2k+s-l}\frac{1}{2}\beta. \qquad (105)$$

Если воспользоваться переходом к эквивалентному представлению при помощи матрицы

$$X = \begin{Vmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{Vmatrix},$$

то дело сведется к перестановке строк и столбцов в обратном порядке, т. е. в данном случае к замене l и s на $(-l)$ и $(-s)$. Таким образом, вместо формул (105) мы можем написать другие формулы:

$$D'_j\{\alpha, \beta, \gamma\}_{ls} = i^{l-s}\sum_k (-1)^k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k+s)!\,(j-l-k)!\,(k-s+l)!} \times$$

$$\times e^{il\alpha + is\gamma}\cos^{2j-l-2k+s}\frac{1}{2}\beta\,\sin^{2k-s+l}\frac{1}{2}\beta, \qquad (106)$$

причем, пользуясь теми же соображениями, что и в [68], мы можем отбросить множитель i^{l-s}.

Отметим простейшие частные случаи. При $j=0$ мы имеем линейное представление первого порядка

$$\eta' = \eta.$$

Это есть тривиальное тождественное представление. При $j = \frac{1}{2}$ мы имеем $2j+1=2$, и величины $\eta_{-\frac{1}{2}}$ и $\eta_{\frac{1}{2}}$ будут просто равны

x_2 и x_1, т. е. в данном случае унитарная группа (93) и является своим собственным линейным представлением (с точностью до перестановки строк и столбцов).

Для группы движения мы получим двузначное представление второго порядка, определяемое матрицами

$$D'_{\frac{1}{2}}\{\alpha, \beta, \gamma\} = \begin{Vmatrix} e^{-\frac{1}{2}i(\gamma+\alpha)}\cos\frac{1}{2}\beta & ie^{\frac{1}{2}i(\gamma-\alpha)}\sin\frac{1}{2}\beta \\ ie^{-\frac{1}{2}i(\gamma-\alpha)}\sin\frac{1}{2}\beta & e^{\frac{1}{2}i(\gamma+\alpha)}\cos\frac{1}{2}\beta \end{Vmatrix}.$$

При $j=1$ мы будем иметь линейное представление третьего порядка:

$$D'_1\{\alpha, \beta, \gamma\} = \begin{Vmatrix} e^{-i(\gamma+\alpha)}\frac{1+\cos\beta}{2} & -e^{-i\alpha}\frac{\sin\beta}{\sqrt{2}} & e^{i(\gamma-\alpha)}\frac{1-\cos\beta}{2} \\ e^{-i\gamma}\frac{\sin\beta}{\sqrt{2}} & \cos\beta & -e^{i\gamma}\frac{\sin\beta}{\sqrt{2}} \\ e^{-i(\gamma-\alpha)}\frac{1-\cos\beta}{2} & e^{i\alpha}\frac{\sin\beta}{\sqrt{2}} & e^{i(\gamma+\alpha)}\frac{1+\cos\beta}{2} \end{Vmatrix}.$$

Линейные представления $D'_j\{\alpha, \beta, \gamma\}$ при целом j дают биоднозначное представление группы вращения. Это следует непосредственно из того, что каждому $D_j\{\alpha, \beta, \gamma\}$ соответствуют две матрицы из группы (93), отличающиеся лишь знаками у a и b, а таким матрицам, как мы упоминали выше, соответствует одно и то же вращение. Если j — половина нечетного числа, то каждому вращению соответствуют две матрицы из представления $D_j\{\alpha, \beta, \gamma\}$, отличающиеся лишь знаком. В частности тождественному преобразованию из группы вращения соответствуют из $D_j\{\alpha, \beta, \gamma\}$ матрицы $\pm E$, где E — единичная матрица порядка $(2j+1)$. *Если ограничиться преобразованиями из $D_j\{\alpha, \beta, \gamma\}$, достаточно близкими к тождественному преобразованию, то $D_j\{\alpha, \beta, \gamma\}$ будут однозначным представлением группы вращения.* При этом в общих формулах (106) можно ограничиться значениями α, β и γ, достаточно близкими к нулю. Но если мы прибавим 2π к α или γ, то ввиду того, что s и l суть половины нечетных чисел, все элементы матрицы $D_j\{\alpha, \beta, \gamma\}$ изменят знак, и мы получим второе представление того же самого по существу вращения. Покажем дальше, что *указанные представления суть все изоморфные неприводимые представления группы вращения.*

Так как представления $D_j\{\alpha, \beta, \gamma\}$ суть все неприводимые представления группы вращения, то матрица $D'_1\{\alpha, \beta, \gamma\}$ должна быть подобна матрице $D\{\alpha, \beta, \gamma\}$, соответствующей вращению пространства с углами

Эйлера $\{\alpha, \beta, \gamma\}$. В [**63**] мы видели, что $D = Z_\alpha X_\beta Z_\gamma$, и, производя перемножение матриц, стоящих слева, получим:

$$D = \begin{vmatrix} \cos\alpha\cos\gamma - \sin\alpha\cos\beta\sin\gamma & -\cos\alpha\sin\gamma - \sin\alpha\cos\beta\cos\gamma & \sin\alpha\sin\beta \\ \sin\alpha\cos\gamma + \cos\alpha\cos\beta\sin\gamma & -\sin\alpha\sin\gamma + \sin\alpha\cos\beta\cos\gamma & -\cos\alpha\cos\beta \\ \sin\beta\sin\gamma & \sin\beta\cos\gamma & \cos\beta \end{vmatrix},$$

и нетрудно проверить формулу:

$$AD_1'\{\alpha, \beta, \gamma\}A^{-1} = D\{\alpha, \beta, \gamma\},$$

где

$$A = \begin{Vmatrix} 1 & 0 & 1 \\ i & 0 & -i \\ 0 & \sqrt{2}i & 0 \end{Vmatrix}.$$

70. Теорема о простоте группы вращения. Докажем сейчас, что группа вращения есть простая группа, т. е. что она не имеет нормальных делителей [**58**]. Если бы такой делитель имелся, то в силу сказанного в [**63**] ему соответствовал бы нормальный делитель группы G преобразований (57) с определителем единица, отличный от нормального делителя H, состоящего из E и $(-E)$. Таким образом, остается показать, что группа G не имеет нормальных делителей, отличных от H, т. е. надо показать, что если элементарный делитель H_1 группы G содержит матрицу A, отличную от E и $(-E)$, то H_1 совпадает с G. Отметим прежде всего, что если H_1 содержит некоторую матрицу B, то в силу определения нормального делителя, H_1 содержит все матрицы $U^{-1}BU$, где U — любая матрица группы G. Выбирая соответствующим образом матрицу U, мы можем получить таким путем любую матрицу группы G, имеющую те же характеристические числа, что и матрица B. Следовательно, чтобы установить, что H_1 совпадает с G, достаточно показать, что H_1 содержит матрицы с любыми допустимыми характеристическими числами. Эти числа должны иметь вид $e^{i\omega}$ и $e^{-i\omega}$, где ω — вещественное число, ибо матрица унитарна, и ее определитель равен единице.

В силу сказанного выше мы можем вместо матрицы A взять матрицу $U^{-1}AU$ и, таким образом, можем считать, что A есть диагональная матрица.

Итак, дано, что H_1 содержит матрицу $A = [e^{i\varphi}, e^{-i\varphi}]$, где φ — вещественно, и $e^{i\varphi} \neq \pm 1$. При этом $A^{-1} = [e^{-i\varphi}, e^{i\varphi}]$. Возьмем произвольную матрицу группы G:

$$U = \begin{Vmatrix} x & y \\ -\bar{y} & \bar{x} \end{Vmatrix} \qquad (\bar{x}x + y\bar{y} = 1).$$

При этом

$$U^{-1} = \begin{Vmatrix} \bar{x} & -y \\ \bar{y} & x \end{Vmatrix}.$$

Поскольку подгруппа H_1 содержит A и является нормальным делителем, она должна содержать и матрицу:

$$Y = A(UA^{-1}U^{-1}).$$

Производя перемножение матриц и учитывая равенство $x\bar{x} + y\bar{y} = 1$, получим следующее выражение следа s матрицы Y:

$$s = 2 - 4y\bar{y}\sin^2\varphi = 2 - 4\rho^2\sin^2\varphi,$$

где $\rho = |y|$ может принимать любое значение из промежутка: $0 \leqslant \rho \leqslant 1$, и $\sin\varphi \neq 0$. Характеристические числа $(e^{i\alpha}, e^{-i\alpha})$ матрицы Y суть корни уравнения:

$$\lambda^2 - s\lambda + 1 = 0, \text{ т. е. } \lambda^2 + (4\rho^2\sin^2\varphi - 2)\lambda + 1 = 0.$$

При изменении ρ от $\rho = 0$ до $\rho = 1$ значения α будут изменяться от $\alpha = 0$ до $\alpha = 2\varphi$. Введем следующее обозначение:

$$U_\beta = [e^{i\beta}, e^{-i\beta}].$$

Из сказанного выше следует, что H_1 содержит все матрицы H_α при $0 \leqslant \alpha \leqslant 2\varphi$. Теперь уже нетрудно показать, что H содержит любую матрицу $U_\beta\,(\beta > 0)$. Действительно, выберем целое положительное число n так, чтобы выполнялось неравенство $0 < \frac{\beta}{n} < 2\varphi$. При этом H_1 содержит $U_{\frac{\beta}{n}}$, а потому содержит и

$$U^n_{\frac{\beta}{n}} = U_\beta.$$

Таким образом, H_1 содержит матрицы с любыми характеристическими числами и тем самым, в силу сказанного выше, совпадает с G. Таким образом доказано, что *группа вращения есть простая группа.*

Из этого непосредственно следует, что группа вращения не может иметь гомоморфных (не изоморфных) представлений. Действительно, если бы имелось такое представление, то тождественному преобразованию в группе представлений должны были бы соответствовать в группе вращений преобразования, образующие нормальный делитель, которого, как показано выше, не существует.

71. Уравнение Лапласа и линейные представления группы вращения. Выясним сейчас связь между линейными представлениями групп и дифференциальными уравнениями. Эта связь лежит в основе применения линейных представлений к вопросам современной физики. Мы начнем с наиболее простого случая уравнения Лапласа [II, 202], который не даст нам ничего нового и послужит лишь к выяснению общего вопроса. Предварительно установим некоторые общие факты, которые играют большую роль в вопросах линейных представлений групп и которые в частных случаях уже известны нам из предыдущих примеров.

Пусть группа G, линейное представление которой строится, есть группа линейных преобразований порядка n:

$$x'_k = g^{(\alpha)}_{k1} x_1 + \ldots + g^{(\alpha)}_{kn} x_n \qquad (k = 1, 2, \ldots, n), \tag{107}$$

где значок α, характеризующий элемент группы G, пробегает конечную или бесконечную совокупность значений. Положим далее, что существует m функций

$$\varphi_s(x_1, \ldots, x_n) \qquad (s = 1, 2, \ldots, m), \tag{108}$$

таких, что при замене независимых переменных по формулам (107) эти функции испытывают также некоторое линейное преобразование

$$\varphi_s(x'_1, \ldots, x'_n) = a^{(\alpha)}_{s1} \varphi_1(x_1, \ldots, x_n) + \ldots + a^{(\alpha)}_{sm} \varphi_m(x_1, \ldots, x_n) \tag{109}$$
$$(s = 1, 2, \ldots, m).$$

Мы имеем здесь матрицу A_α с элементами $a^{(\alpha)}_{ik}$, соответствующую преобразованию (107) группы G. Рассмотрим два преобразования группы

$$(x'_1, \ldots, x'_n) = G_{\alpha_1}(x_1, \ldots, x_n); \; (x''_1, \ldots, x''_n) = G_{\alpha_2}(x'_1, \ldots, x'_n);$$
$$G_{\alpha_3} = G_{\alpha_2} G_{\alpha_1}.$$

Соответствующее преобразование функций (108) будет:

$$\varphi_s(x'_1, \ldots, x'_n) = a^{(\alpha_1)}_{s1} \varphi_1(x_1, \ldots, x_n) + \ldots + a^{(\alpha_1)}_{sm} \varphi_m(x_1, \ldots, x_n) \tag{110_1}$$

и

$$\varphi_s(x''_1, \ldots, x''_n) = a^{(\alpha_2)}_{s1} \varphi_1(x'_1, \ldots, x'_n) + \ldots + a^{(\alpha_2)}_{sm} \varphi_m(x'_1, \ldots, x'_n). \tag{110_2}$$

Подставляя в (110_2) вместо $\varphi_s(x'_1, \ldots, x'_n)$ их выражения (110_1), будем иметь непосредственную зависимость $\varphi_s(x''_1, \ldots, x'_n)$ через $\varphi_s(x_1, \ldots, x_n)$, дающую матрицу A_{α_3}. Мы получим таким образом

$$\{A_{\alpha_3}\}_{ik} = \sum_{s=1}^{m} a^{(\alpha_2)}_{is} a^{(\alpha_1)}_{sk}, \text{ т. е. } A_{\alpha_3} = A_{\alpha_2} A_{\alpha_1},$$

и формулы (109) определяют, очевидно, некоторое линейное представление порядка m для группы G. В предыдущих рассуждениях мы считали, что функции φ_s линейно-независимы. При этом линейные преобразования (109) определяются вполне однозначно и $D(A_\alpha) \neq 0$, ибо в противном случае $\varphi_s(x'_1, \ldots, x'_n)$ были бы связаны линейной зависимостью.

В частном случае, при построении линейных представлений унитарной группы, роль функций φ_s играли функции (96_1).

Положим, что G есть группа вращения трехмерного пространства, так что $n=3$, и положим, что функции φ_s ортогональны и нормированы в некотором шаре K с центром в начале, т. е.

$$\iiint\limits_K \varphi_p(x_1, x_2, x_3)\overline{\varphi_q(x_1, x_2, x_3)}\,dx_1\,dx_2\,dx_3 = \delta_{pq}. \tag{111}$$

Покажем, что при этом линейное представление (109) группы вращения будет унитарным. Действительно, в результате вращения G_α сфера K перейдет в себя, и определитель G_α, как известно, равен единице. Условие (111) дает нам:

$$\iiint\limits_K \varphi_p(x'_1, x'_2, x'_3)\overline{\varphi_q(x'_1, x'_2, x'_3)}\,dx'_1\,dx'_2\,dx'_3 = \delta_{pq},$$

или в силу (109):

$$\iiint\limits_K \left[\sum_{i=1}^m a_{pi}^{(\alpha)}\varphi_i(x_1, x_2, x_3)\cdot\sum_{j=1}^m \overline{a_{qj}^{(\alpha)}\varphi_j(x_1, x_2, x_3)}\right]dx'_1\,dx'_2\,dx'_3 = \delta_{pq}.$$

Переходя к переменным (x_1, x_2, x_3), мы должны будем, согласно правилу замены переменных в тройном интеграле, заменить просто $dx'_1\,dx'_2\,dx'_3$ на $dx_1\,dx_2\,dx_3$ и затем интегрировать по тому же самому шару K. В силу условий (111) мы получим:

$$\sum_{i=1}^m a_{pi}^{(\alpha)}\overline{a_{qi}^{(\alpha)}} = \delta_{pq} \qquad (p, q = 1, 2, \ldots, m),$$

где, как всегда, $\delta_{pq}=0$ при $p\neq q$ и $\delta_{pp}=1$, т. е. каждая из матриц A обладает в данном случае свойством ортогональности по строкам, так что транспонированная матрица будет обладать ортогональностью по столбцам, а следовательно [**28**], и по строкам, т. е. основная матрица будет обладать ортогональностью и по строкам и по столбцам, или, иначе говоря, матрица A_α будет действительно унитарной матрицей для любого значка α.

Рассмотрим теперь уравнение Лапласа с двумя переменными

$$\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}=0 \tag{112}$$

или, пользуясь векторными обозначениями,

$$\operatorname{div}\operatorname{grad} U = 0. \tag{113}$$

Возьмем однородный полином от x и y степени l:

$$\varphi_l(x, y) = a_0x^l + a_1x^{l-1}y + \ldots + a_kx^{l-k}y^k + \ldots + a_ly^l. \tag{114}$$

Покажем, что существуют два линейно-независимых полинома вида (114), которые являются решением уравнения (112), и всякое решение уравнения (112), представляющееся однородным полиномом степени l, должно быть линейной комбинацией этих двух полиномов с постоянными коэффициентами. Действительно, коэффициенты полинома (114) выражаются по формулам

$$a_k = \frac{1}{(l-k)!\,k!}\,\frac{\partial^l \varphi_l(x, y)}{\partial x^{l-k}\partial y^k}.$$

Но раз этот полином должен удовлетворять уравнению (112), то мы можем двукратное дифференцирование по y заменить двукратным дифференцированием по x при одновременном изменении знака, так как уравнение (112) можно переписать в виде:

$$\frac{\partial^2 U}{\partial y^2} = -\frac{\partial^2 U}{\partial x^2}.$$

Таким образом, мы получим для коэффициентов a_k выражения следующего вида:

$$a_k = \pm\frac{1}{(l-k)!\,k!}\,\frac{\partial^l \varphi_l}{\partial x^l} \quad \text{или} \quad a_k = \pm\frac{1}{(l-k)!\,k!}\,\frac{\partial^l \varphi_l}{\partial x^{l-1}\partial y},$$

т. е. все коэффициенты полинома (114) выразятся через коэффициенты a_0 и a_1. Это рассуждение показывает нам, что существует не больше двух линейно-независимых однородных полиномов, удовлетворяющих уравнению (112). Покажем теперь, что такие два различных полинома действительно существуют. Рассмотрим для этого однородный полином

$$\omega_l(x, y) = (x + iy)^l.$$

Раскрывая скобки и отделяя вещественную и мнимую части, получим:

$$\omega_l(x, y) = \varphi_l(x, y) + i\psi_l(x, y),$$

где полиномы $\varphi_l(x, y)$ и $\psi_l(x, y)$ — вещественные однородные полиномы степени l, линейно-независимые друг от друга. Дифференцируя $\omega_l(x, y)$, получим:

$$\frac{\partial^2\omega_l(x, y)}{\partial x^2} = l(l-1)(x+iy)^{l-2};$$

$$\frac{\partial^2\omega_l(x, y)}{\partial y^2} = -l(l-1)(x+iy)^{l-2},$$

т. е. $\omega_l(x, y)$ удовлетворяет уравнению (112). То же самое, следовательно, можно сказать и о вещественной и мнимой части этого выражения, т. е. полиномы $\varphi_l(x, y)$ и $\psi_l(x, y)$ и дают нам два искомых решения уравнения (112). Введем полярные координаты

$$x = r\cos\varphi; \qquad y = r\sin\varphi,$$

откуда

$$\omega_l(x, y) = r^l e^{il\varphi}.$$

При этом полиномы φ_l и ψ_l примут очень простой вид:

$$\varphi_l(x, y) = r^l\cos l\varphi; \qquad \psi_l(x, y) = r^l\sin l\varphi.$$

Повернем плоскость XY вокруг начала на некоторый угол ϑ:

$$x' = x\cos\vartheta - y\sin\vartheta, \quad y' = x\sin\vartheta + y\cos\vartheta \tag{115}$$

Нетрудно видеть, что при этом уравнение (112) останется инвариантным, т. е., точнее говоря, в новых переменных уравнение будет выглядеть совершенно так же:

$$\frac{\partial^2 U}{\partial x'^2} + \frac{\partial^2 U}{\partial y'^2} = 0. \tag{116}$$

Это можно непосредственно проверить, пользуясь формулами (115) и правилом дифференцирования сложных функций. Кроме того, указанное обстоятельство непосредственно вытекает из того, что левая часть уравнения (113) имеет определенное значение, не зависящее от выбора осей и, следовательно, имеет одну и ту же форму при любом выборе прямолинейных прямоугольных осей. Полиномы $\varphi_l(x', y')$ и $\psi_l(x', y')$ должны удовлетворять уравнению (116), а следовательно, и уравнению (112), а потому они должны линейно выражаться через $\varphi_l(x, y)$ и $\psi_l(x, y)$. Это и дает нам линейное представление группы вращения плоскости.

Вместо указанных двух полиномов возьмем два других полинома, которые являются их линейными комбинациями:

$$\varphi_l'(x, y) = \varphi_l(x, y) - i\psi_l(x, y); \quad \psi_l'(x, y) = \varphi_l(x, y) + i\psi_l(x, y)$$

или

$$\varphi_l'(x, y) = (x - iy)^l = r^l e^{-il\varphi}; \quad \psi_l'(x, y) = (x + iy)^l = r^l e^{il\varphi}.$$

Эти полиномы испытывают следующее преобразование:

$$\varphi_l'(x', y') = r^l e^{-il(\varphi+\vartheta)} = e^{-il\vartheta}\varphi_l'(x, y),$$
$$\psi_l'(x', y') = r^l e^{il(\varphi+\vartheta)} = e^{il\vartheta}\psi_l'(x, y),$$

т. е. преобразованию (115) соответствует в линейном представлении матрица

$$\begin{Vmatrix} e^{-il\vartheta} & 0 \\ 0 & e^{il\vartheta} \end{Vmatrix},$$

где угол ϑ может иметь любое значение. Из вида матриц непосредственно вытекает, что это линейное представление имеет приведенную форму. Оно дает два линейных представления первого порядка, определяемых числами $e^{-il\vartheta}$ и $e^{il\vartheta}$. Во всех предыдущих рассуждениях целое число l могло иметь любое значение. Мы получили таким образом те же самые линейные представления группы вращения плоскости, которые мы имели и раньше в [69].

Перейдем теперь к уравнению Лапласа с тремя переменными

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0 \tag{117}$$

или

$$\operatorname{div}\operatorname{grad} U = 0. \tag{118}$$

Рассмотрим и здесь однородные полиномы степени l с тремя переменными

$$\varphi_l(x, y, z) = a_0 z^l + X_1(x, y) z^{l-1} + X_2(x, y) z^{l-2} + \ldots + X_{l-1}(x, y) z + X_l(x, y), \tag{119}$$

где $X_k(x, y)$ — однородный полином степени k своих аргументов. Каждый такой полином $X_k(x, y)$ содержит $(k+1)$ произвольных коэффициентов, так что в общем однородный полином $\varphi_l(x, y, z)$ степени l с тремя переменными будет содержать следующее число произвольных коэффициентов:

$$1 + 2 + 3 + \ldots + (l+1) = \frac{(l+1)(l+2)}{2}.$$

Подставляя выражение (119) в уравнение (117), получим слева однородный полином степени $(l-2)$, и, приравнивая его коэффициенты нулю, получим $\frac{(l-1)l}{2}$ однородных уравнений с $\frac{(l+1)(l+2)}{2}$ неизвестными коэффициентами полинома $\varphi_l(x, y, z)$. Мы имеем:

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

и, следовательно, по крайней мере $(2l+1)$ коэффициентов в полиноме $\varphi_l(x, y, z)$ останутся произвольными, т. е. будет существовать по крайней мере $(2l+1)$ линейно-независимых однородных полиномов степени l, удовлетворяющих уравнению (117). Пользуясь тем же методом, что и для двух переменных можно показать, что их будет и не больше, чем $(2l+1)$, т. е. их будет в точности $(2l+1)$. Обозначим эти полиномы через

$$\psi_s^{(l)}(x, y, z) \qquad (s = 1, 2, \ldots, 2l+1).$$

Если

$$(x', y', z') = U \cdot (x, y, z)$$

есть некоторое вращение трехмерного пространства вокруг начала, то при этом уравнение (117) остается инвариантным, и полиномы $\psi_s^{(l)}(x, y, z)$ дают некоторое линейное представление группы вращения трехмерного пространства, порядка $(2l+1)$.

В дальнейшем мы подробно изложим теорию этих гармонических полиномов и выведем для них явные выражения. Мы увидим, что их всегда можно выбрать так, что они будут ортогональными и нормированными в любой сфере с центром в начале. При этом доставляемое ими линейное представление группы вращения будет унитарным. Можно показать, что это и будет как раз линейное представление, эквивалентное представлению $D_l\{\alpha, \beta, \gamma\}$, которое мы построили в [69]. Позже мы вернемся к этому вопросу.

72. Прямое произведение матриц. Пусть имеются две матрицы

$$A=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix} \text{ и } B=\begin{Vmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{Vmatrix}; \tag{120}$$

первая — порядка n и вторая — порядка m. Составим новую матрицу C с элементами $C_{ij;\,kl}$, которые получаются от перемножения каждого элемента матрицы A на каждый элемент матрицы B:

$$\{C\}_{ij;\,kl}=c_{ij;\,kl}=a_{ik}b_{jl}. \tag{121}$$

В данном случае роль первого значка играет совокупность двух целых чисел (i, j), а роль второго значка — совокупность целых чисел (k, l), причем

$$i \text{ и } k=1,\ 2,\ \dots,\ n;$$
$$j \text{ и } l=1,\ 2,\ \dots,\ m.$$

Иными словами, мы имеем здесь особый способ наименования строк и столбцов, а именно: строки и столбцы именуются совокупностью двух целых чисел, причем первое число принимает значение от 1 до n, а второе — от 1 до m. Мы можем, конечно, перенумеровать строки и столбцы обычным образом, т. е. простыми целыми числами, которые идут от 1 до nm, причем каждой паре чисел (i, j) или (k, l) соответствует определенное целое число при новой нумерации, и если эти пары одинаковы, то и соответствующие целые числа одинаковы. Эту нумерацию простыми целыми числами можно делать различным способом. При переходе от одного способа к другому дело сведется к некоторой перестановке одновременно строк и столбцов, т. е. к переходу к подобной матрице, что в дальнейшем не будет иметь никакого значения.

Матрица C называется *прямым произведением матриц* A и B, и обозначается это обычно следующим образом:

$$C=A\times B. \tag{122}$$

Порядок множителей в этом последнем произведении нового типа не играет никакой роли.

Положим, например, что обе матрицы (120) будут матрицами второго порядка. В этом случае их прямое произведение будет матрицей четвертого порядка, которую мы можем написать, например, в следующем виде:

$$C=\begin{Vmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{Vmatrix}=\begin{Vmatrix} c_{11;\,11} & c_{11;\,12} & c_{11;\,21} & c_{11;\,22} \\ c_{12,\,11} & c_{12,\,12} & c_{12;\,21} & c_{12;\,22} \\ c_{21;\,11} & c_{21;\,12} & c_{21;\,21} & c_{21;\,22} \\ c_{22;\,11} & c_{22;\,12} & c_{22;\,21} & c_{22;\,22} \end{Vmatrix}$$

или, иначе, одновременно переставляя строки и столбцы.

Положим, что A и B суть диагональные матрицы

$$A=[\gamma_1, \ldots, \gamma_n]; \quad B=[\delta_1, \ldots, \delta_m].$$

В этом случае $a_{ik}=0$ и $b_{jl}=0$ при $i \neq k$ и $j \neq l$, и, следовательно, согласно (121), $c_{ij;\,kl}$ отлично от нуля только, если пара чисел (i, j) совпадает с парой чисел (k, l), т. е. если матрица C также будет диагональной. На ее главной диагонали будут стоять всевозможные произведения чисел γ_k на числа δ_l. Если все γ_k и все δ_l равны единице, то C будет также единичной матрицей. Мы имеем, таким образом, следующую теорему:

Теорема I. *Прямое произведение двух диагональных матриц есть диагональная матрица, и прямое произведение двух единичных матриц есть единичная матрица.*

Докажем также следующую теорему:

Теорема II. *Если $A^{(1)}$ и $A^{(2)}$ — две матрицы одного и того же порядка n, а $B^{(1)}$ и $B^{(2)}$ — две матрицы также одного и того же порядка m, то имеет место формула:*

$$(A^{(2)} \times B^{(2)})(A^{(1)} \times B^{(1)})=A^{(2)}A^{(1)} \times B^{(2)}B^{(1)}. \qquad (123)$$

Заметим, что когда мы пишем две матрицы одного и того же порядка рядом, без всякого знака, то это, как всегда, обозначает обычное произведение этих двух матриц. Обозначая элементы матриц соответствующими малыми буквами с двумя значками внизу, мы имеем, согласно определению прямого произведения:

$$\{A^{(t)} \times B^{(t)}\}_{ij,\,kl}=a_{ik}^{(t)}b_{jl}^{(t)} \qquad (t=1,\ 2),$$

и, пользуясь правилом обычного умножения матриц, получим для элементов левой части равенства (123) следующие формулы:

$$d_{ij;\,kl}=\sum_{p=1}^{n}\sum_{q=1}^{m} a_{ip}^{(2)}b_{jq}^{(2)}a_{pk}^{(1)}b_{ql}^{(1)}. \qquad (124)$$

Покажем, что те же формулы получаются и для элементов правой части. Мы имеем по определению обычного умножения:

$$\{A^{(2)}A^{(1)}\}_{ik}=\sum_{p=1}^{n} a_{ip}^{(2)}a_{pk}^{(1)}; \quad \{B^{(2)}B^{(1)}\}_{jl}=\sum_{q=1}^{m} b_{iq}^{(2)}b_{ql}^{(1)},$$

и, по определению прямого произведения:

$$d_{ij;\,kl}=\sum_{p=1}^{n} a_{ip}^{(2)}a_{pk}^{(1)} \cdot \sum_{q=1}^{m} b_{jq}^{(2)}b_{ql}^{(1)},$$

что и совпадает с (124). Перейдем теперь к доказательству последней теоремы о прямом произведении.

Теорема III. *Если матрицы A и B унитарны, то и их прямое произведение $C = A \times B$ также есть унитарная матрица.*

По условию теоремы мы имеем:

$$\sum_{s=1}^{n} a_{sp}\bar{a}_{sq} = \delta_{pq}; \quad \sum_{s=1}^{m} b_{sp}\bar{b}_{sq} = \delta_{pq}. \tag{125}$$

Проверим для матрицы C условия ортогональности и нормальности по столбцам и обозначим:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij;\,p_1q_1}\bar{c}_{ij;\,p_2q_2} = \delta_{p_1q_1;\,p_2q_2},$$

т. е. в силу (121):

$$\delta_{p_1q_1;\,p_2q_2} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ip_1}\bar{a}_{ip_2}b_{jq_1}\bar{b}_{jq_2} = \sum_{i=1}^{n} a_{ip_1}\bar{a}_{ip_2}\sum_{j=1}^{m} b_{jq_1}\bar{b}_{jq_2}. \tag{126}$$

Если пары чисел (p_1, q_1) и (p_2, q_2) различны, то хоть один из множителей, стоящих в правой части (126), будет равен нулю, а если эти пары совпадают, оба множителя равны единице в силу (125). Таким образом, $\delta_{p_1q_1;\,p_2q_2}$ равно нулю, если упомянутые пары не совпадают, и равно единице, если эти пары совпадают, что и доказывает нашу теорему.

Мы можем, очевидно, прямое произведение двух матриц умножить еще, в смысле прямого произведения, на третью матрицу и получить прямое произведение трех матриц

$$A^{(1)} \times A^{(2)} \times A^{(3)}.$$

Удерживая прежнее обозначение, мы будем иметь для элементов этой новой матрицы следующие выражения:

$$c_{ikl;\,i'k'l'} = a_{ii'}^{(1)}a_{kk'}^{(2)}a_{ll'}^{(3)}.$$

Аналогичным образом составляется прямое произведение любого конечного числа матриц, причем это прямое произведение представляет собою матрицу, порядок которой равен произведению порядков перемножаемых матриц. Последовательность множителей не играет роли.

73. Композиция двух линейных представлений группы. Пусть имеется некоторая группа G с элементами G_α и положим, что построены два линейных представления этой группы:

$$x'_i = a_{i1}^{(\alpha)}x_1 + \ldots + a_{in}^{(\alpha)}x_n \qquad (i = 1, 2, \ldots, n) \tag{127}$$

и

$$y'_k = b_{k1}^{(\alpha)}y_1 + \ldots + b_{km}^{(\alpha)}y_m \qquad (k = 1, 2, \ldots, m) \tag{128}$$

где значок α пробегает конечную или бесконечную совокупность значений. Обозначим через $A^{(\alpha)}$ и $B^{(\alpha)}$ матрицы преобразований (127) и (128) и составим их прямое произведение

$$C^{(\alpha)} = A^{(\alpha)} \times B^{(\alpha)}. \tag{129}$$

Покажем, что матрицы $C^{(\alpha)}$ также дают некоторое линейное представление нашей группы G. Действительно, всякому элементу G_α группы G соответствует матрица $C^{(\alpha)}$; произведению $G_{\alpha_2}G_{\alpha_1} = G_{\alpha_3}$ будет соответствовать матрица $C^{(\alpha_2)}C^{(\alpha_1)}$, которая определяется в силу (123) формулой:

$$C^{(\alpha_2)}C^{(\alpha_1)} = (A^{(\alpha_2)} \times B^{(\alpha_2)})(A^{(\alpha_1)} \times B^{(\alpha_1)}) = (A^{(\alpha_2)}A^{(\alpha_1)}) \times (B^{(\alpha_2)}B^{(\alpha_1)}).$$

Но раз матрицы $A^{(\alpha)}$ и $B^{(\alpha)}$ дают линейное представление группы, то

$$A^{(\alpha_2)}A^{(\alpha_1)} = A^{(\alpha_3)} \quad \text{и} \quad B^{(\alpha_2)}B^{(\alpha_1)} = B^{(\alpha_3)},$$

и, следовательно:

$$C^{(\alpha_2)}C^{(\alpha_1)} = A^{(\alpha_3)} \times B^{(\alpha_3)},$$

т. е. согласно (129):

$$C^{(\alpha_2)}C^{(\alpha_1)} = C^{(\alpha_3)}.$$

Таким образом, произведению элементов G_α соответствует как раз произведение соответствующих матриц $C^{(\alpha)}$, и эти матрицы дают новое линейное представление группы G. Заметим при этом, что единичному элементу из G соответствует при этом прямое произведение единичных матриц $A^{(\alpha)}$ и $B^{(\alpha)}$, т. е. единичная матрица $C^{(\alpha)}$.

Составим nm произведений $x_i y_k$ и подвергнем каждый из сомножителей преобразованиям (127) и (128). Мы будем иметь:

$$x'_i y'_k = (a_{i1}^{(\alpha)} x_1 + \ldots + a_{in}^{(\alpha)} x_n) \cdot (b_{k1}^{(\alpha)} y_1 + \ldots + b_{km}^{(\alpha)} y_m),$$

или, раскрывая скобки:

$$x'_i y'_k = \sum_{p=1}^{n} \sum_{q=1}^{m} c_{ik;\,pq}^{(\alpha)} x_p y_q, \quad \text{где} \quad c_{ik;\,pq}^{(\alpha)} = a_{ip}^{(\alpha)} b_{kq}^{(\alpha)},$$

т. е. *если x_i и y_k суть объекты в линейных представлениях, определяемых матрицами $A^{(\alpha)}$ и $B^{(\alpha)}$, то $x_i y_k$ будут объектами в линейном представлении той же группы, определяемом матрицами $C^{(\alpha)}$*. Если матрицы $A^{(\alpha)}$ и $B^{(\alpha)}$ давали неприводимые линейные представления, то матрица $C^{(\alpha)}$ не обязательно будет давать неприводимое линейное представление. В дальнейшем мы подробно рассмотрим тот случай, когда группа G есть группа вращения трехмерного пространства, а матрицы $A^{(\alpha)}$ и $B^{(\alpha)}$ суть различные неприводимые линейные представления этой группы, построенные нами в [69]. Покажем что в этом случае произведение

$$D_{j1}\{\alpha, \beta, \gamma\} \times D_{j2}\{\alpha, \beta, \gamma\}$$

будет приводимым, и определим, из каких неприводимых представлений оно будет состоять.

В качестве примера рассмотрим уравнение Шредингера для случая двух электронов, находящихся в поле положительного ядра. Это уравнение имеет вид:

$$\left[-\frac{h^2}{8\pi^2 m}\sum_{s=1}^{2}\left(\frac{\partial^2}{\partial x_s^2}+\frac{\partial^2}{\partial y_s^2}+\frac{\partial^2}{\partial z_s^2}\right)+V\right]\psi=E\psi, \tag{130}$$

где

$$V=\sum_{s=1}^{2}-\frac{e^2 e_0}{\sqrt{x_s^2+y_s^2+z_s^2}}+\frac{1}{2}\frac{e^2}{\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}} \tag{131}$$

причем постоянные имеют обычные значения. Второе слагаемое в выражении V происходит от взаимодействия электронов. Если мы пренебрежем в первом приближении этим взаимодействием, то уравнение будет:

$$(H_1+H_2)\,\psi=E\psi, \tag{132}$$

где

$$H_s=-\frac{h^2}{8\pi^2 m}\left(\frac{\partial^2}{\partial x_s^2}+\frac{\partial^2}{\partial y_s^2}+\frac{\partial^2}{\partial z_s^2}\right)-\frac{e^2 e_0}{\sqrt{x_s^2+y_s^2+z_s^2}}\quad (s=1,\ 2).$$

Положим, что отдельные уравнения:

$$H_1\psi=E\psi;\quad H_2\psi=E\psi \tag{133}$$

имеют собственные значения E_1 и E_2 и соответствующие собственные функции

$$\psi_1\,(x_1,\ y_1,\ z_1)\quad \text{и}\quad \psi_2\,(x_2,\ y_2,\ z_2),$$

т. е.

$$H_1\psi_1=E_1\psi_1,\quad H_2\psi_2=E_2\psi_2. \tag{134}$$

Если мы подставим в уравнение (132):

$$\psi=\psi_1\,(x_1,\ y_1,\ z_1)\cdot\psi_2\,(x_2,\ y_2,\ z_2),$$

то получим, очевидно, в силу (134):

$$(H_1+H_2)\,\psi=\psi_2 H_1\psi_1+\psi_1 H_2\psi_2=(E_1+E_2)\,\psi_1\psi_2=(E_1+E_2)\,\psi,$$

т. е. уравнение (132) будет иметь собственную функцию $\psi_1\psi_2$, которой будет соответствовать значение (E_1+E_2). Левая часть уравнений (133) содержит оператор Лапласа и расстояние точки до начала координат, и, следовательно, эти левые части не меняются, если мы совершим вращение пространства вокруг начала. Может случиться, что характеристическому числу $E=E_1$ в первом из уравнений (133) соответствует несколько собственных функций ψ_1. Все эти функции, являясь решением уравнения, дадут некоторое линейное представление группы вращения, совершенно так же, как в [69] однородные гармонические полиномы давали нам также представление группы вращения. Пусть это будет некоторое представление $D_{j1}\{\alpha,\ \beta,\ \gamma\}$. Совершенно так же решения второго из уравнений (133) при заданном собственном значении $E=E_2$ дадут нам некоторое представление $D_{j2}\{\alpha,\ \beta,\ \gamma\}$ группы вращения. Произведение $\psi_1\psi_2$, согласно предыдущему, даст нам линейное представление группы вращения, совпадающее с прямым произведением $D_{j1}\times D_{j2}$, и для физической характеристики соответствующего собственного значения (E_1+E_2) уравнения (132) представляется существенным выделить из этого представления те неприводимые представления, на которые оно распадается. Это обстоятельство играет существенную роль в теории возмущений.

74. Прямое произведение групп и его линейные представления. Понятие о прямом произведении матриц играет роль и в другом вопросе, к которому мы сейчас и переходим. Пусть имеются две группы G и H, элементы которых обозначим через G_α и H_β, причем значки α и β пробегают, независимо один от другого, вообще говоря, различные совокупности значений. Определим новую группу F, элементы которой определяются как пары элементов из G и H:

$$F_{\alpha\beta} = (G_\alpha, H_\beta),$$

причем первый элемент пары есть элемент G, а второй — элемент H. Назовем единичным элементом этой группы элемент $F_{\alpha\beta}$ в том случае, когда G_α и H_β суть единичные элементы из G и H, и аналогичным образом определим обратные элементы из группы F. Закон умножения для группы F определим естественно формулой:

$$F_{\alpha_2\beta_2}F_{\alpha_1\beta_1} = (G_{\alpha_2}G_{\alpha_1}, H_{\beta_2}H_{\beta_1}).$$

Нетрудно проверить, что совокупность элементов $F_{\alpha\beta}$ действительно образует группу. Эту группу F назовем прямым произведением групп G и H. Пусть имеется некоторое линейное представление группы G, осуществляемое матрицами $A^{(\alpha)}$, и некоторое линейное представление группы H, осуществляемое матрицами $B^{(\beta)}$. Пользуясь формулой (123), как и в предыдущем номере, можно показать, что *прямые произведения*

$$C^{(\alpha,\beta)} = A^{(\alpha)} \times B^{(\beta)}$$

дают линейное представление группы F. Кроме того, если представления $A^{(\alpha)}$ и $B^{(\beta)}$ унитарны, то и представление $C^{(\alpha,\beta)}$ группы F будет унитарным [**72**].

Покажем теперь, что *если представления $A^{(\alpha)}$ и $B^{(\beta)}$ неприводимы, то и представление $C^{(\alpha,\beta)}$ группы F будет неприводимым.* Пусть n — порядок матриц $A^{(\alpha)}$ и m — порядок матриц $B^{(\beta)}$. Матрицы $C^{(\alpha,\beta)}$ будут иметь порядок nm. Пусть существует некоторая матрица X порядка nm, которая коммутирует со всеми матрицами $C^{(\alpha,\beta)}$. Обозначим элементы матриц соответствующими малыми буквами. Мы имеем, следовательно, для любых значков i, j, p и q, а также для любых α и β:

$$\sum_{l=1}^{m}\sum_{k=1}^{n} x_{ij;\,kl}\,a_{kp}^{(\alpha)}b_{lq}^{(\beta)} = \sum_{l=1}^{m}\sum_{k=1}^{n} a_{ik}^{(\alpha)}b_{jl}^{(\beta)}x_{kl;\,pq}, \tag{135}$$

где

$$a_{kp}^{(\alpha)}b_{lq}^{(\beta)} = c_{kl;\,pq}^{(\alpha,\beta)} \quad \text{и} \quad a_{ik}^{(\alpha)}b_{jl}^{(\beta)} = c_{ij;\,kl}^{(\alpha,\beta)}.$$

Если мы положим, что $G^{(\alpha)}$ есть единичный элемент группы G, то $A^{(\alpha)}$ будет единичной матрицей, т. е. $a_{kp}^{(\alpha)} = 0$ при $k \neq p$ и $a_{pp}^{(\alpha)} = 1$, и формула (135) даст:

$$\sum_{l=1}^{m} x_{ij;\,pl}\,b_{lq}^{(\beta)} = \sum_{l=1}^{m} b_{jl}^{(\beta)}x_{il;\,pq}, \tag{136}$$

и точно так же, полагая, что $B^{(\beta)}$ есть единичный элемент группы B, получим:

$$\sum_{k=1}^{m} x_{ij;\,kq} a_{kp}^{(\alpha)} = \sum_{k=1}^{n} a_{ik}^{(\alpha)} x_{kj;\,pq}. \tag{137}$$

Если мы возьмем nm элементов $x_{ij;\,kl}$ и закрепим значки i и k, то получится m^2 элементов

$$x_{ij;\,kl} \qquad (j,\ l = 1,\ 2,\ \ldots,\ m),$$

которые дадут некоторую матрицу порядка m. Обозначим эту матрицу через $X_1^{(i,k)}$. Точно так же, закрепляя в $x_{ij;\,kl}$ значки j и l, получим матрицу $X_2^{(j,l)}$ порядка n. В силу (136) все матрицы $X_1^{(i,k)}$ коммутируют со всеми матрицами $B^{(\beta)}$, образующими неприводимое представление группы B, и, следовательно, все матрицы $X_1^{(i,k)}$ кратны единичной матрице, т. е. элементы $x_{ij;\,kl}$ при фиксированных i и k имеют одинаковое значение, если $j=l$, и, кроме того, $x_{ij;\,kl}=0$, если $j \neq l$. Мы можем записать это следующим образом

$$x_{ij;\,kl} = x_{i1;\,k1}\delta_{jl}. \tag{138_1}$$

Точно так же из рассмотрения матриц $X_2^{(j,l)}$ будет следовать:

$$x_{ij;\,kl} = x_{1j;\,1l}\delta_{ik}, \tag{138_2}$$

где, как всегда,

$$\delta_{pq} = 0 \text{ при } p \neq q \text{ и } \delta_{pp} = 1.$$

Из сравнения (138_1) и (138_2) вытекает, что $x_{ij;\,kl}$ отлично от нуля только в том случае, если $i=k$ и $j=l$, причем в этом последнем случае все числа $x_{ij;\,ij}$ одинаковы между собой, т. е. матрица X, коммутирующая со всеми матрицами $C^{(\alpha,\beta)}$, будет обязательно кратной единичной матрице. Отсюда и вытекает непосредственно, что линейное представление группы F, определяемое прямым произведением $A^{(\alpha)} \times B^{(\beta)}$, будет неприводимым. Можно показать, что *таким образом получаются все неприводимые представления группы F.*

Положим, что G и H суть группы линейных преобразований с одним и тем же числом переменных, и предположим, что любые две матрицы G_α и H_β попарно коммутируют, т. е.

$$G_\alpha H_\beta = H_\beta G_\alpha. \tag{139}$$

В предыдущих рассуждениях мы считали, что элемент группы F определяется парой элементов (G_α, H_β), и построили определенный закон перемножения внутри группы F, который описан нами выше. В данном случае мы можем считать элементом группы F просто произведение матриц (139), которое не зависит от порядка. Эта новая группа F изоморфна прежней F. Если G_{α_0} и H_{β_0} — единичные матрицы, то и произведение $G_{\alpha_0}H_{\beta_0} = H_{\beta_0}G_{\alpha_0}$ будет единичной матрицей. Матрица $G_\alpha^{-1}H_\beta^{-1} = H_\beta^{-1}G_\alpha^{-1}$ будет, очевидно, обратной произведению $G_\alpha H_\beta$, и мы имеем в силу (139) следующий закон умножения:

$$G_{\alpha_2}H_{\beta_2} \cdot G_{\alpha_1}H_{\beta_1} = (G_{\alpha_2}G_{\alpha_1})(H_{\beta_2}H_{\beta_1}),$$

т. е. все упомянутые выше при образовании группы F свойства в данном случае выполнены, так что произведения (139) можно считать переменным элементом группы F. В качестве частного случая возьмем тот случай, когда G есть группа вращения трехмерного пространства и H — группа второго порядка, состоящая из тождественного преобразования I и симметрии S относительно начала [57]. В данном случае условия (139) выполняются. Если G_α есть любое вращение пространства, то очевидно $G_\alpha S = SG_\alpha$. Группа F в данном случае будет группой всех вещественных ортогональных преобразований трехмерного пространства. Для группы H мы имели [67] два линейных представления первого порядка: одно тождественное, состоящее из чисел $(+1)$, и другое антисимметрическое, при котором матрице I соответствует $(+1)$ и матрице S соответствует (-1). Если мы возьмем теперь некоторое линейное представление $D_j\{\alpha, \beta, \gamma\}$ группы вращения, то можем брать прямое произведение матриц этого представления с обоими представлениями группы симметрии относительно начала. В одном случае мы получим линейное представление полной группы ортогональных преобразований, при котором всякому вращению с углами Эйлера $\{\alpha, \beta, \gamma\}$, взятому в чистом виде или соединенному с симметрией относительно начала, соответствует одна и та же матрица $D_j\{\alpha, \beta, \gamma\}$. Обозначим это представление группы ортогональных преобразований через $D_j^+\{\alpha, \beta, \gamma\}$. В другом случае чистому вращению будет соответствовать матрица $D_j\{\alpha, \beta, \gamma\}$, а вращению, соединенному с симметрией, матрица $-D_j\{\alpha, \beta, \gamma\}$. Такое представление группы ортогональных преобразований обозначим через $D_j^-\{\alpha, \beta, \gamma\}$.

Разберем еще один пример прямого произведения двух групп. Пусть у нас имеются две точки: (x_1, y_1, z_1) и (x_2, y_2, z_2). Положим, что группа G есть группа вращения трехмерного пространства. Наши переменные испытывают при этом линейные преобразования:

$$\begin{aligned} x'_k &= g_{11}x_k + g_{12}y_k + g_{13}z_k, \\ y'_k &= g_{21}x_k + g_{22}y_k + g_{23}z_k, \qquad (k = 1,\ 2) \\ z'_k &= g_{31}x_k + g_{32}y_k + g_{33}z_k, \end{aligned} \tag{140}$$

где таблица g_{ik} есть матрица некоторого вращения. Положим дальше, что группа H есть группа, состоящая из тождественного преобразования и из преобразования, которому соответствует перестановка номеров 1 и 2 у наших точек. Это последнее преобразование будет иметь вид:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad (S). \tag{141}$$

Мы имеем, очевидно, $S^2 = I$, и следовательно, эта последняя группа H будет состоять из двух преобразований: I и S. Если G_α есть некоторое вращение, то, очевидно, $G_\alpha S = SG_\alpha$, так как безразлично — производить ли нумерацию точек до или после вращения. В данном случае мы получим те же линейные представления для общей группы F, что и выше. Если бы мы вместо двух взяли n точек, то группа H, состоящая из перемен нумерации этих точек, имела бы своими элементами линейные преобразования с n переменными, а эта группа H была бы изоморфна группе перестановок из n элементов. В данном случае точно так же операция вращения и операция перестановки номеров точек будут коммутировать друг с другом, и, взяв прямое произведение матрицы линейного представления группы вращения и матриц некоторого линейного представления группы перестановок, мы получим линейное представление общей группы F.

75. Разбиение композиции $D_j \times D_{j'}$ линейных представлений группы вращения. Возвратимся сейчас к тому, что мы говорили в [**73**]. Мы видели там, что если рассмотрим уравнение Шредингера для двух электронов и будем пренебрегать взаимодействием электронов, то собственные функции уравнения Шредингера будут давать линейное представление группы вращения, которое получается путем композиции двух линейных представлений группы вращения. Результаты предыдущего номера показывают, что представляется важным уметь разбить такое линейное представление на неприводимые части. В настоящем номере мы и займемся этим вопросом. Задача математически формулируется следующим образом. Пусть имеются два неприводимых представления $D_j\{\alpha, \beta, \gamma\}$ и $D_{j'}\{\alpha, \beta, \gamma\}$ группы вращения. Составим их композицию $D_j \times D_{j'}$, которая также будет давать [**73**] некоторое линейное представление группы вращения. Требуется выделить те неприводимые части, из которых это представление состоит.

Объектами линейного представления D_j порядка $(2j+1)$ будут величины

$$U_m = \frac{u_1^{j+m} u_2^{j-m}}{\sqrt{(j+m)!\,(j-m)!}} \quad (m = -j,\ -j+1,\ \ldots,\ j-1,\ j) \tag{142}$$

и объектами линейного представления $D_{j'}$ будут величины

$$V_{m'} = \frac{v_1^{j'+m'} v_2^{j'-m'}}{\sqrt{(j'+m')!\,(j'-m')!}} \quad (m' = -j',\ -j'+1,\ \ldots,\ j'-1,\ j'), \tag{143}$$

причем (u_1, u_2) и (v_1, v_2) подвергаются одинаковым унитарным преобразованиям с определителем $(+1)$ [68] Если мы составим $(2j+1)(2j'+1)$ величин

$$W_{mm'} = U_m V_{m'} = \frac{u_1^{j+m} u_2^{j-m} v_1^{j'+m'} v_2^{j'-m'}}{\sqrt{(j+m)!\,(j-m)!\,(j'+m')!\,(j'-m')!}} \tag{144}$$

$$\begin{pmatrix} m = -j, \ -j+1, \ \ldots, \ j-1, \ j \\ m' = -j', \ -j'+1, \ \ldots, \ j'-1, \ j' \end{pmatrix},$$

то эти величины будут объектами в том линейном представлении группы вращения, которое определяется композицией $D_j \times D_{j'}$.

В дальнейшем мы будем считать j и j' или целым числом, или половиной целого числа, т. е., строго говоря, будем брать линейные представления унитарной группы с двумя переменными и определителем, равным единице.

Пусть k — целое число (или половина целого числа), удовлетворяющее неравенству

$$|j - j'| \leqslant k \leqslant j + j'. \tag{145}$$

Покажем, что мы можем составить из величин (144) такие линейные комбинации, числом $2k+1$, которые дают линейное представление D_k группы вращения.

Составим для доказательства этого утверждения выражение вида:

$$L = (u_1 v_2 - u_2 v_1)^l \, (u_1 x_1 + u_2 x_2)^{2j-l} \, (v_1 x_1 + v_2 x_2)^{2j'-l}, \tag{146}$$

где l — некоторое фиксированное целое число, удовлетворяющее неравенствам:

$$l \geqslant 0; \ l \leqslant 2j; \ l \leqslant 2j'. \tag{147}$$

Если переменные (u_1, u_2) и (v_1, v_2) претерпевают одно и то же линейное преобразование

$$u_1' = a_{11} u_1 + a_{12} u_2; \qquad v_1' = a_{11} v_1 + a_{12} v_2$$

$$u_2' = a_{21} u_1 + a_{22} u_2; \qquad v_2' = a_{21} v_1 + a_{22} v_2$$

с определителем $(+1)$, т. е. $a_{11} a_{22} - a_{12} a_{21} = 1$, то нетрудно видеть, что первый из множителей выражения (146) остается неизменным. Действительно,

$$u_1' v_2' - u_2' v_1' = (a_{11} a_{22} - a_{12} a_{21})(u_1 v_2 - u_2 v_1).$$

Выражение (146) есть, очевидно, однородный полином от x_1 и x_2 степени $2(j+j'-l)$. Он состоит, следовательно, из членов вида:

$$a_s x_1^s x_2^{2(j+j'-l)-s} \qquad (s = 0, 1, \ldots, 2(j+j'-l)).$$

Вводя обозначение

$$k = j + j' - l, \tag{148}$$

$$y_{m''} = \frac{x_1^{k+m''} x_2^{k-m''}}{\sqrt{(k+m'')!\,(k-m'')!}} \tag{149}$$

$$(m'' = -k, \ -k+1, \ \ldots, \ k-1, \ k),$$

мы можем написать выражение (146) следующим образом:

$$L = \sum_{m''=-k}^{+k} c_{m''} y_{m''}. \tag{150}$$

Коэффициенты $c_{m''}$ будут зависеть от переменных $(u_1, u_2)(v_1, v_2)$.

Из выражения (146) непосредственно следует, что $c_{m''}$ есть однородный полином от (u_1, u_2) степени $2j$ и однородный полином от (v_1, v_2) степени $2j'$, т. е. $c_{m''}$ будет состоять из слагаемых вида:

$$a'_{pq} \ u_1^p \ u_2^{2j-p} \ v_1^q \ v_2^{2j'-q},$$

или, принимая во внимание (142) и (143), мы можем утверждать, что $c_{m''}$ будет линейной комбинацией произведений

$$c_{m''} = \sum_m \sum_{m'} d_{mm'}^{(m'')} U_m V_{m'} \ (m'' = -k, \ -k+1, \ \ldots, \ k-1, \ k), \tag{151}$$

где коэффициенты $d_{mm'}^{(m'')}$ не содержат уже u_k и v_k. Заметим, что в выражении (146) переменные u_1 и v_1 входят только или в соединении с множителем x_1, или в первом из множителей (146), причем этот первый множитель дает сумму показателей у u_1 и v_1, равную l. Принимая во внимание, что $y_{m''}$ содержит $x_1^{k+m''}$, мы можем утверждать, что в слагаемых суммы (151) сумма показателей у u_1 и v_1 будет $k + m'' + l$ или, в силу (148), эта сумма показателей будет $j + j' + m'$. Но U_m содержит u_1^{j+m} и $V_{m'}$ содержит $v_1^{j'+m'}$, и отсюда непосредственно следует, что каждое из выражений (151) содержит лишь те произведения $U_m V_{m'}$, для которых $m + m' = m''$. Покажем теперь, что линейные комбинации (151) величин $U_m V_{m'}$ как раз и дают линейное представление группы вращения, эквивалентное D_k.

Предварительно напомним определение контраградиентного преобразования. Если имеются два линейных преобразования

$$(x'_1, \ \ldots, \ x'_n) = A(x_1, \ \ldots, \ x_n) \ \text{и} \ (y'_1, \ \ldots, \ y'_n) = B(y_1, \ \ldots, \ y_n),$$

то для выполнения равенства

$$x'_1 y'_1 + \ldots + x'_n y'_n = x_1 y_1 + \ldots + x_n y_n$$

необходимо и достаточно, чтобы B было контраградиентно с A, т. е. $B = A^{(*)-1}$ (ср. [21] и [40]).

Положим, что переменные (u_1, u_2) и (v_1, v_2) одновременно подвергаются некоторому унитарному преобразованию A с определителем $(+1)$. Положим, что переменные x_1 и x_2 подвергаются при этом преобразованию $A^{(*)-1}$, контраградиентному преобразованию A. Из определения контраградиентного преобразования следует, что при этом суммы

$$u_1x_1 + u_2x_2 \text{ и } v_1x_1 + v_2x_2$$

остаются неизменными. Кроме того, как мы показали выше, и первый множитель в выражении (146) остается неизменным при указанном преобразовании наших переменных. Таким образом, и вся сумма L останется неизменной, т. е., иначе говоря, в силу (150), переменные $c_{m''}$ испытают преобразование B, контраградиентное тому преобразованию C, которое испытают переменные $y_{m''}$.

Введем новые переменные

$$z_{m''} = \frac{u_1^{k+m''}u_2^{k-m''}}{\sqrt{(k+m'')!\,(k-m'')!}} \qquad (m'' = -k, \ -k+1, \ \ldots, \ k-1, \ k).$$

Применяя формулу бинома Ньютона, можем написать:

$$(u_1x_1 + u_2x_2)^{2k} = (2k)! \sum_{m''=-k}^{+k} z_{m''}y_{m''}.$$

Левая часть последнего выражения остается инвариантной при наших преобразованиях, и, следовательно, то же можно сказать и о правой части, т. е. переменные $z_{m''}$ подвергаются тому же самому преобразованию B, контраградиентному преобразованию C, что и переменные $c_{m''}$. Но мы знаем, что переменные $z_{m''}$ как раз и дают нам линейное представление D_k группы вращения, если (u_1, u_2) суть объекты унитарной группы с определителем $(+1)$. Следовательно, наше утверждение доказано.

Мы можем таким образом из переменных (144), которые толкуем как составляющие некоторого вектора в пространстве с $(2j+1)(2j'+1)$ измерениями, составить линейные комбинации числом $(2k+1)$, которые дают линейное представление D_k группы вращения. Принимая во внимание формулу (148) и неравенства (147), мы видим, что числу k можем придавать следующие значения:

$$k = j+j', \ j+j'-1, \ \ldots, \ |j-j'|. \tag{152}$$

Подсчитаем теперь, сколько всего линейных комбинаций из величин (144) нам придется составить. Для определенности будем предполагать, что $j \geqslant j'$. Упомянутое общее число линейных комбинаций будет:

$$(2j+2j'+1) + (2j+2j'-1) + \ldots + (2j-2j'+1).$$

Это есть сумма арифметической прогрессии с числом членов

$$\frac{(2j+2j'+1)-(2j-2j'+1)}{2}+1=2j'+1,$$

и общее число линейных комбинаций будет $(2j+1)(2j'+1)$, т. е. будет равно числу величин (144). Тот же результат получился бы и при условии $j<j'$. Полагая для краткости

$$(2j+1)(2j'+1)=r,$$

обозначим упомянутые выше линейные комбинации величин (144) следующим образом:

$$w_1,\ w_2,\ \ldots,\ w_r, \tag{153}$$

причем будем считать, что эти линейные комбинации расположены в том порядке, который дает линейные представления D_k, где k имеет значения (152). В результате некоторого унитарного преобразования с определителем $(+1)$ над переменными $(u_1,\ u_2)$ и $(v_1,\ v_2)$ мы получим новые значения $U'_m V'_m$ переменных (144) и новые значения $w'_s\,(s=1,\ 2,\ \ldots,\ r)$ переменных (153), причем w'_s выражаются через w_s при помощи квазидиагональной матрицы

$$[D_{j+j'},\ D_{j+j'-1},\ \ldots, D_{|j-j'|}], \tag{154}$$

и каждое D_k соответствует тому унитарному преобразованию, которому мы подвергли $(u_1,\ u_2)$ и $(v_1,\ v_2)$. Дальше покажем, что линейные формы (153) величин (144) между собой линейно-независимы. Пусть T — матрица того линейного преобразования, при помощи которого w_s выражаются через переменные (144). Прямое произведение $D_j\times D_{j'}$ есть матрица линейного преобразования для переменных (144), и мы имеем в силу предыдущего:

$$(D_{j+j'},\ D_{j+j'-1},\ \ldots,\ D_{|j-j'|}]=T\,(D_j\times D_{j'})\,T^{-1}, \tag{155}$$

что и дает разложение прямого произведения на неприводимые части. Предыдущую формулу обычно записывают следующим образом:

$$D_j\times D_{j'}=D_{j+j'}+D_{j+j'-1}+\ldots+D_{|j-j'|}. \tag{156}$$

Напомним, что каждое D_k определяется унитарным преобразованием и полностью записывается так: $D_k\left\{\begin{matrix} a & b\\ -\bar b & \bar a\end{matrix}\right\}$. Предыдущий результат обобщается и на случай нескольких сомножителей. Так, например, мы можем написать:

$$\begin{aligned}D_1\times D_1\times D_1&=(D_2+D_1+D_0)\times D_1=\\&=D_3+D_2+D_1+D_2+D_1+D_0+D_1=\\&=D_3+2D_2+3D_1+D_0.\end{aligned}$$

Сама матрица D_1 есть матрица третьего порядка [68]. Прямое произведение $D_1 \times D_1$ будет матрицей девятого порядка, и, наконец, прямое произведение $D_1 \times D_1 \times D_1$ будет матрицей двадцать седьмого порядка. Предыдущая формула показывает, что эта матрица эквивалентна при всяком выборе унитарного преобразования квазидиагональной матрице

$$[D_3, D_2, D_2, D_1, D_1, D_1, D_0].$$

Порядок этой последней матрицы равен [68]:

$$(2 \cdot 3 + 1) + 2(2 \cdot 2 + 1) + 3(2 \cdot 1 + 1) + (2 \cdot 0 + 1) = 27.$$

Обратимся теперь к доказательству линейной независимости w_s, как линейных форм от величин (144). Величины w_s в прежних обозначениях суть величины $c_{m''}$, но только надо принять во внимание, что при построении $c_{m''}$ мы можем брать различные значения k или, что то же, различные значения l, так что правильнее было бы писать $c_{m''}^{(l)}$. Как мы видели выше, всякое $c_{m''}^{(l)}$ выражается только через те величины $U_m V_{m'}$, для которых $m + m' = m''$. Отсюда непосредственно вытекает, что линейно-зависимыми могут оказаться лишь $c_{m''}^{(l)}$ при различных l, но одинаковых m''. Открывая в выражении (146) скобки в двух последних множителях и собирая члены при $x_1^{k+m''} x_2^{k-m''}$, где k определяется формулой (148), мы и получим с точностью до постоянного множителя $c_{m''}^{(l)}$, выраженные через u_k и v_k. Они будут, очевидно, произведениями $(u_1 v_2 - u_2 v_1)^l$ на некоторый полином от u_1, u_2, v_1 и v_2 с целыми положительными коэффициентами. Нетрудно видеть, что такие выражения не могут быть при различных l линейно-зависимыми. Предположим, например, что мы имеем линейную зависимость вида:

$$\alpha_1 c_{m''}^{(l_1)} + \alpha_2 c_{m''}^{(l_2)} + \alpha_3 c_{m''}^{(l_3)} = 0,$$

где $l_1 < l_2 < l_3$ и α_k — некоторые постоянные, отличные от нуля. Написанное соотношение должно выполняться тождественно при всяких u_1, u_2, v_1 и v_2. Положим, например, $u_2 = v_1 = v_2 = 1$. В силу сказанного выше о форме выражений $c_{m''}^{(l)}$, мы получим соотношение вида

$$\alpha_1 (u_1 - 1)^{l_1} p_1(u_1) + \alpha_2 (u_1 - 1)^{l_2} p_2(u_1) + \alpha_3 (u_1 - 1)^{l_3} p_3(u_1) = 0,$$

где $p_k(u_1)$ — полиномы от u_1 с целыми положительными коэффициентами. Деля предыдущее соотношение на $(u_1 - 1)^{l_1}$ и полагая затем $u_1 = 1$, получим $\alpha_1 = 0$, что противоречит сказанному выше и доказывает таким образом невозможность линейных соотношений.

Раскрывая скобки в выражении (146), мы могли бы, конечно, и фактически построить выражения w_s через переменные (144).

76. Свойство ортогональности. Матрицы, образующие неэквивалентные унитарные неприводимые представления, обладают некоторым свойством, которое называется обычно свойством *ортогональности*. Им часто пользуются при применении теории групп к физике. Сначала сформулируем это свойство.

Пусть имеется конечная группа G порядка m с элементами

$$G_1, G_2, \ldots, G_m$$

и пусть

$$A^{(1)}, \ldots, A^{(m)} \text{ и } B^{(1)}, \ldots, B^{(m)}$$

— две системы матриц, дающих линейное представление группы G. Обозначая элементы этих матриц малыми буквами с двумя значками снизу и считая, что указанные два линейных представления неэквивалентные, неприводимые представления и состоят из унитарных матриц, мы будем иметь следующие равенства:

$$\sum_{s=1}^{m} a_{ij}^{(s)} \overline{b_{kl}^{(s)}} = 0, \tag{157}$$

которые имеют место при любых нижних значках. Аналогичные равенства имеют место и для одного неприводимого, унитарного представления. Пусть порядок матриц $A^{(s)}$, дающих унитарное, неприводимое представление, равен p. Имеют место следующие формулы:

$$\sum_{s=1}^{m} a_{ij}^{(s)} \overline{a_{kl}^{(s)}} = \frac{m}{p} \delta_{ik}\delta_{jl}, \tag{158}$$

т. е. сумма, стоящая слева, равна нулю, если пары чисел (i, j) и (k, l) различны, и равна $\frac{m}{p}$, если эти пары одинаковы.

Доказательство ортогональности основано на теореме III из [66]. Предварительно напомним понятие умножения для случая неквадратных прямоугольных матриц. Пусть имеются две матрицы C и D с элементами

$$\{D\}_{ik} \begin{pmatrix} i = 1, 2, \ldots, n_1 \\ k = 1, 2, \ldots, n_2 \end{pmatrix} \text{ и } \{C\}_{jl} \begin{pmatrix} j = 1, 2, \ldots, n_2 \\ l = 1, 2, \ldots, n_3 \end{pmatrix},$$

причем число n_2 столбцов матрицы D совпадает с числом строк матрицы C. Элементы произведения DC определим обычной формулой

$$\{DC\}_{ik} = \sum_{s=1}^{n_2} \{D\}_{is} \{C\}_{sk}.$$

Новая матрица DC будет иметь, очевидно, n_1 строк и n_3 столбцов.

Формулируем теперь основную теорему.

Теорема. *Если унитарные матрицы $A^{(s)}$ порядка p и унитарные матрицы $B^{(s)}$ порядка q дают неэквивалентные неприводимые представления группы G и если некоторая прямоугольная матрица C с p строками и q столбцами удовлетворяет при всяком s условиям*

$$A^{(s)}C = CB^{(s)} \quad (s = 1, 2, \ldots, m), \tag{159}$$

то C — нулевая матрица, т. е. *все ее элементы равны нулю.*

Рассмотрим сначала случай $p = q$, когда C есть также квадратная матрица. Если определитель C отличен от нуля, то существует C^{-1}, и из (159) следует

$$A^{(s)} = CB^{(s)}C^{-1},$$

т. е. наши два представления эквивалентны, что противоречит условию теоремы. Итак, определитель C должен равняться нулю. Положим, что не все элементы C равны нулю, и обозначим эти элементы через c_{ik}. Как известно, линейные формы

$$c_{i1}x_1+\ldots+c_{ip}x_p \quad (l=1, 2, \ldots, p)$$

определяют при произвольных x_s подпространство, число измерений которого равно рангу матрицы C [**14**], т. е. в данном случае это будет подпространство с числом измерений $\geqslant 1$ и $<p$. Иначе говоря, это будет действительно не полное пространство p измерений, а некоторое подпространство R. Напишем (159) как некоторое линейное преобразование над вектором с составляющими $(x_1, \ldots, x_p)$:

$$A^{(s)}C(x_1, \ldots, x_p)=CB^{(s)}(x_1, \ldots, x_p) \quad (s=1, 2, \ldots, m).$$

Слева $C(x_1, \ldots, x_p)$ есть произвольный вектор из R, а вся правая часть, представляющая собой линейное преобразование C над вектором $B^{(s)}(x_1, \ldots, x_p)$, также принадлежит R. Иначе говоря, преобразование $A^{(s)}$ над любым вектором из R дает опять вектор из R. При этом, как мы знаем [**66**], $A^{(s)}$ дают приводимое представление, что противоречит условию теоремы.

Доказательство это остается в силе и при $p>q$. Тогда ранг матрицы C во всяком случае меньше p, и линейные формы

$$c_{i1}x_1+\ldots+c_{iq}x_q \quad (i=1, 2, \ldots, p)$$

определяют в пространстве p измерений некоторое подпространство R, с числом измерений $<p$, так что предыдущее доказательство остается справедливым. Положим, наконец, что $p<q$, и в условиях (159) перейдем к транспонированным матрицам. Это даст нам

$$B^{(s)*}C^*=C^*A^{(s)*}.$$

В данном случае порядок q матриц $B^{(s)*}$ выше порядка p матриц $A^{(s)*}$, и, как и выше, мы заключаем отсюда, что унитарные матрицы $B^{(s)*}$ оставляют инвариантным некоторое пространство, а следовательно, мы можем соответствующим подбором основных ортов привести их к квазидиагональной форме. При этом и матрицы $B^{(s)}$ будут квазидиагональными, что противоречит условию теоремы. Таким образом, теорема доказана.

Мы могли бы не упоминать в условиях теоремы о том, что матрицы $A^{(s)}$ и $B^{(s)}$ унитарны. Как известно, переходя к подобным представлениям, мы можем всегда считать $A^{(s)}$ и $B^{(s)}$ унитарными, причем переход к подобным представлениям в соотношениях (159) вводит вместо C новую матрицу C_1, связанную с C соотношением вида:

$$C=D_1C_1D_2,$$

если C_1 есть нулевая матрица, то такой же будет и C.

Обратимся теперь к доказательству формул (157). Вместо $A^{(s)}$ и $B^{(s)}$ введем обозначения $A(G_s)$ и $B(G_s)$, где G_s — тот элемент группы G, которому соответствуют матрицы $A^{(s)}$ и $B^{(s)}$. Пусть X — любая матрица, имеющая p строк и q столбцов. Введем матрицу

$$C=\sum_{s=1}^{m} A(G_s)XB(G_s)^{-1} \tag{160}$$

и покажем, что она удовлетворяет соотношениям (159).

Пусть G_t — какой-либо фиксированный элемент группы G. Мы имеем:

$$A(G_t)C=\sum_{s=1}^{m} A(G_t)A(G_s)XB(G_s)^{-1}.$$

Но в силу определения линейного представления

$$A(G_t)A(G_s)=A(G_tG_s) \text{ и } B(G_t)B(G_s)=B(G_tG_s),$$

и отсюда

$$A(G_t)C=\sum_{s=1}^{m} A(G_tG_s)XB(G_tG_s)^{-1}B(G_t).$$

Элемент G_s пробегает всю группу, то же самое можно сказать и о произведении G_tG_s, так что предыдущую формулу можно записать в виде:

$$A(G_t)C=CB(G_t),$$

т. е. матрица C, определяемая формулой (160), действительно удовлетворяет соотношениям (159), и, следовательно, эта матрица C есть нулевая матрица Итак, при любом выборе матрицы X имеем:

$$\sum_{s=1}^{m} A(G_s)XB(G_s)^{-1}=0.$$

Положим, что в матрице X некоторый фиксированный элемент $\{X\}_{jl}$ равен единице, а остальные нулю. Написанная формула дает нам при этом

$$\sum_{s=1}^{m} \{A(G_s)\}_{ij}\{B(G_s)^{-1}\}_{lk}=0.$$

В силу унитарности матрица $B(G_s)$ получается из $B(G_s)^{-1}$ заменой строк столбцами и всех элементов сопряженными, так что предыдущая формула при прежних обозначениях дает

$$\sum_{s=1}^{m} a_{ij}^{(s)}\overline{b_{kl}^{(s)}}=0,$$

что и совпадает с (157).

Точно так же строя матрицу

$$D=\sum_{s=1}^{m} A(G_s)XA(G_s)^{-1},$$

где X — любая квадратная матрица порядка p, мы можем показать, что

$$A(G_s)D=DA(G_s) \quad (s=1, 2, \ldots, m),$$

и в силу теоремы III из [66] можно утверждать, что D кратно единичной матрице, или

$$\sum_{s=1}^{m} A(G_s)XA(G_s)^{-1}=cI,$$

где число c зависит от выбора X. Положим опять $\{X\}_{jl}=1$, а остальные элементы X равными нулю, и обозначим через c_{jl} соответствующее значение числа c. Мы можем написать:

$$\sum_{s=1}^{m} \{A(G_s)\}_{ij}\{A(G_s)^{-1}\}_{lk}=c_{jl}\delta_{ik}. \tag{161}$$

Для определения c_{jl} положим $i=k$ и будем суммировать по i от 1 до p

$$pc_{jl}=\sum_{s=1}^{m}\sum_{i=1}^{p}\{A(G_s)^{-1}\}_{li}\{A(G_s)\}_{ij}=\sum_{s=1}^{m}\{I\}_{lj}.$$

Если $l=j$, то правая часть равна m, а при $l\neq j$ она равна нулю. Отсюда $c_{jl}=\frac{m}{p}\delta_{jl}$ и, следовательно, формула (161) переписывается в виде:

$$\sum_{s=1}^{m}\{A(G_s)\}_{ij}\{A(G_s)^{-1}\}_{lk}=\frac{m}{p}\delta_{ik}\delta_{jl}, \tag{162}$$

что и совпадает с (158), если воспользоваться унитарностью матриц $A(G_s)$.

Нетрудно видеть, что соотношение (157) имеет место не только для унитарных представлений группы, но и для любых неэквивалентных, неприводимых представлений. Пусть $A'(G_s)$ и $B'(G_s)$ — два таких представления порядков p и q, а $A(G_s)$ и $B(G_s)$ — эквивалентные им унитарные представления, так что

$$A(G_s)=C_1A'(G_s)C_1^{-1};\quad B(G_s)=C_2B'(G_s)C_2^{-1},$$

где C_1 и C_2 — определенные матрицы, не зависящие от s. В силу унитарности $B(G_s)$ имеем:

$$B(G_s)^{-1}=\overline{B(G_s)^*}=\overline{(C_2^{-1})^*}\,\overline{B'(G_s)^*}\,\overline{C_2^*},$$

и формула (157) может быть записана в виде:

$$\sum_{s=1}^{m}C_1A'(G_s)C_1^{-1}X\,\overline{(C_2^{-1})^*}\,\overline{B'(G_s)^*}\,\overline{C_2^*}=0,$$

откуда, умножая слева на C_1^{-1} и справа на $\overline{(C_2^*)^{-1}}$ и вводя произвольную матрицу $Y=C_1^{-1}X\,\overline{(C_2^{-1})^*}$, имеющую p строк и q столбцов:

$$\sum_{s=1}^{m}A'(G_s)\,Y\,\overline{B'(G_s)^*}=0,$$

и пользуясь произвольностью Y, как и выше, получим:

$$\sum_{s=1}^{m}a_{ij}^{(s)}\overline{b_{kl}^{(s)}}=0.$$

Отметим также, что формула (162) имеет место для любого, а не только для унитарного представления, что вытекает из ее доказательства и из того, что в формулировке приведенной выше теоремы не нужно упоминать об унитарности матриц $A^{(s)}$ и $B^{(s)}$.

77. Характеры. Пусть, как и выше, $A(G_s)$ и $B(G_s)$ — два неэквивалентных, неприводимых представления порядков p и q группы G с элементами $G_1, G_2, \ldots, G_m$. Обозначим через $X(G_s)$ и $X'(G_s)$ — следы матриц в этих представлениях, т. е. сумму их диагональных элементов:

$$X(G_s)=\sum_{i=1}^{p}\{A(G_s)\}_{ii},\quad X'(G_s)=\sum_{k=1}^{q}\{B(G_s)\}_{kk}.$$

Эти числа называются *характерами* упомянутых представлений. У эквивалентных представлений характеры, очевидно, одинаковы [27], и мы можем считать, что упомянутые представления суть унитарные представления. Формула ортогональности дает:

$$\sum_{s=1}^{m} \{A(G_s)\}_{ii} \overline{\{B(G_s)\}_{kk}} = 0,$$

откуда, суммируя по i и по k, получаем формулу ортогональности для характеров:

$$\sum_{s=1}^{m} X(G_s) \overline{X'(G_s)} = 0. \tag{163}$$

Точно так же формула (158) дает:

$$\sum_{s=1}^{m} \{A(G_s)\}_{ii} \overline{\{A(G_s)_{kk}} = \frac{m}{p} \delta_{ik}$$

и, суммируя по i и k, получим

$$\sum_{s=1}^{m} X(G_s) \overline{X(G_s)} = m. \tag{164}$$

Пользуясь этими формулами, докажем некоторые теоремы.

Теорема 1. *Необходимым и достаточным условием эквивалентности двух неприводимых представлений является совпадение всех их характеров.*

Мы уже упоминали о том, что характеры у эквивалентных (приводимых или неприводимых) представлений одинаковы, и тем самым необходимость условия установлена. Положим теперь наоборот, что дано совпадение системы характеров двух неприводимых представлений, т. е. $X(G_s) = X'(G_s)$ $(s = 1, 2, \ldots, m)$, и докажем эквивалентность представлений. В силу (164) имеем:

$$\sum_{s=1}^{m} X(G_s) \overline{X'(G_s)} = m,$$

откуда вытекает эквивалентность представлений, ибо если бы они были не эквивалентны, то мы должны были иметь равенство (163). Отметим, что порядок матриц в эквивалентных представлениях должен быть, очевидно, одинаковым. В соответствии с каждым неприводимым представлением введем в m-мерном комплексном пространстве R_m векторы с составляющими:

$$\frac{1}{\sqrt{m}} X(G_1), \quad \frac{1}{\sqrt{m}} X(G_2), \ \ldots, \quad \frac{1}{\sqrt{m}} X(G_m).$$

В силу (164) эти векторы нормированы и, в силу (163), векторы, соответствующие неэквивалентным представлениям, взаимно ортогональны Отсюда следует, что *не может существовать больше чем m неэквивалентных неприводимых представлений группы G порядка m.* В дальнейшем мы уточним общее число неэквивалентных неприводимых представлений группы. Обозначим пока это число буквою l Пусть $\omega^{(i)}$ $(i = 1, 2, \ldots, l)$ — эти неэквивалентные неприводимые представления и

$$X^{(i)}(G_1), \quad X^{(i)}(G_2), \ \ldots, \ X^{(i)}(G_m) \qquad (i = 1, 2, \ldots, l)$$

— характеры этих представлений. Пусть имеется некоторое представление ω с характерами:

$$X(G_1),\ X(G_2),\ \ldots,\ X(G_m).$$

В результате приведения представления ω оно изобразится квазидиагональными матрицами, составленными из матриц представлений $\omega^{(i)}$. Для характеров мы будем иметь таким образом:

$$X(G_s) = \sum_{i=1}^{l} a_i X^{(i)}(G_s), \tag{165}$$

где a_i — целые числа, не меньшие нуля, которые показывают, сколько раз представление $\omega^{(i)}$ входит в состав представления ω после его приведения.

Можно указать формулы для определения коэффициентов a_i по характерам представления ω. Пусть k — одно из чисел $1, 2, \ldots, l$. Умножим обе части (165) на $\overline{X^{(k)}(G_s)}$ и просуммируем по s. Пользуясь (163) и (164), получим:

$$\sum_{s=1}^{m} X(G_s)\overline{X^{(k)}(G_s)} = a_k m,$$

откуда

$$a_k = \frac{1}{m}\sum_{s=1}^{m} X(G_s)\overline{X^{(k)}(G_s)}. \tag{166}$$

Эта формула дает для всякого a_k определенное значение, откуда вытекает следующая теорема:

Теорема 2. *Всякое приводимое представление распадается на единственную совокупность неприводимых представлений.*

Пользуясь формулой (166), нетрудно обобщить теорему 1 и на случай любых, а не только неприводимых, представлений.

Теорема 3 *Необходимым и достаточным условием эквивалентности двух представлений является совпадение всех их характеров.*

Необходимость условия отмечалась и при доказательстве теоремы 1. Наоборот, если характеры $X(G_s)$ двух представлений одинаковы, то, согласно формулам (166), мы получим одинаковые значения для чисел a_k, и оба представления приводятся, следовательно, к квазидиагональной матрице, состоящей из одинаковых неприводимых представлений. При этом, переходя, если нужно, к эквивалентным представлениям, можем считать, что указанные неприводимые представления расположены в квазидиагональной матрице в одинаковом порядке, ибо одна и та же перестановка строк и столбцов равносильна переходу к эквивалентному представлению.

Тем самым представления с одинаковыми характерами приводятся к одним и тем же квазидиагональным матрицам, т. е. они эквивалентны.

Перейдем теперь к исследованию числа l всех неприводимых, не эквивалентных между собой представлений группы G. Элементы этой группы распределяются по классам. В один и тот же класс входят те элементы, которые получаются из одного из них G_t при помощи формул:

$$G_s G_t G_s^{-1} \qquad (s = 1, 2, \ldots, m).$$

Всем этим элементам соответствуют в любом представлении подобные матрицы, имеющие одинаковый след. Пусть r — число классов в группе G. В силу сказанного выше, всякое линейное представление группы G имеет не больше чем r различных характеров, причем каждый характер соответствует не отдельному элементу, а всем элементам, входящим в некоторый класс. Пусть класс C_1 состоит из g_1 элементов, класс C_2 — из g_2 элементов и, наконец, класс C_r — из g_r элементов. Слагаемые суммы (163) будут одинаковыми для элементов одного и того же класса и, обозначая через $X(C_k)$, $X'(C_k)$ — характеры, соответствующие элементам, входящим в класс C_k, для двух неэквивалентных неприводимых представлений можем переписать (163) в виде:

$$\sum_{k=1}^{r} X(C_k)\overline{X'(C_k)}\,g_k = 0$$

и (164) в виде:

$$\sum_{k=1}^{r} X(C_k)\overline{X(C_k)}\,g_k = m.$$

Таким образом, для характеров $X^{(i)}(C_k)$ неэквивалентных, неприводимых представлений $\omega^{(i)}$ $(i = 1, 2, \ldots, l)$ будем иметь:

$$\begin{gathered}\sum_{k=1}^{r} X^{(i_1)}(C_k)\overline{X^{(i_2)}(C_k)}\,g_k = 0 \quad \text{при } i_1 \neq i_2,\\ \sum_{k=1}^{r} X^{(i)}(C_k)\overline{X^{(i)}(C_k)}\,g_k = m.\end{gathered} \tag{167}$$

Введем в пространстве R_r, число измерений которого равно r, l векторов с составляющими:

$$\sqrt{\frac{g_1}{m}}X^{(i)}(C_1), \quad \sqrt{\frac{g_2}{m}}X^{(i)}(C_2), \ \ldots, \quad \sqrt{\frac{g_r}{m}}X^{(i)}(C_r) \qquad (i = 1, 2, \ldots, l).$$

Предыдущие равенства показывают, что эти векторы попарно ортогональны и нормированы, а потому и линейно-независимы. Отсюда вытекает, что их число l не больше числа измерений, т. е. $l \leqslant r$. Мы получили теорему:

Теорема 4. *Число неэквивалентных, неприводимых представлений группы не больше числа классов группы.*

В следующем номере покажем, что всегда $l = r$. Поскольку мы только что показали, что $l \leqslant r$, для доказательства равенства $l = r$ достаточно доказать неравенство $l \geqslant r$. Доказательство этого неравенства связано с введением некоторых новых понятий и соотношений между характерами, которые представляют интерес и сами по себе.

Установим еще одно соотношение между характерами любого неприводимого представления. Положим, что класс C_k состоит из элементов $G_1^{(k)}$, $G_2^{(k)}, \ldots, G_{g_k}^{(k)}$. Если G_s — какой-либо элемент группы, то элементы $G_s G_i^{(k)} G_s^{-1}$ $(i = 1, 2, \ldots, g_k)$ дадут опять все элементы из класса C_k, но уже в другом порядке. Отсюда следует, что если мы возьмем совокупность всех произведений элементов, входящих в некоторые классы C_p и C_q:

$$G_u^{(p)} G_v^{(q)} \quad (u = 1, 2, \ldots, g_p; \quad v = 1, 2, \ldots, g_q), \tag{168}$$

то совокупность элементов

$$G_s G_u^{(p)} G_v^{(q)} G_s^{-1} = (G_s G_u^{(p)} G_s^{-1})(G_s G_v^{(q)} G_s^{-1})$$

будет той же самой. Отсюда следует, что совокупность элементов (168) обладает тем свойством, что если некоторый элемент принадлежит этой совокупности, то этой же совокупности принадлежит целиком весь класс, содержащий этот элемент, причем каждый элемент этого класса входит в совокупность элементов (168) одинаковое число раз *Обозначим через a_{pqk} целое число, не меньшее нуля, которое показывает, сколько раз элементы класса C_k входят в совокупность элементов* (168). Иначе это записывают, чисто условно, следующим образом:

$$C_pC_q = \sum_{k=1}^{r} a_{pqk}C_k \tag{169}$$

или

$$(G_1^{(p)} + G_2^{(p)} + \ldots + G_{g_p}^{(p)})\,(G_1^{(q)} + G_2^{(q)} + \ldots + G_{g_q}^{(q)}) =$$
$$= \sum_{k=1}^{r} a_{pqk}\,(G_1^{(k)} + G_2^{(k)} + \ldots + G_{g_k}^{(k)}). \tag{170}$$

Пусть $A(G_s)$ — матрицы порядка n некоторого неприводимого линейного представления группы G. Образуем сумму матриц, соответствующих элементам класса C_k, и обозначим эту матрицу через $A(C_k)$:

$$A(C_k) = \sum_{j=1}^{g_k} A(G_j^{(k)}).$$

Принимая во внимание, что элементы $G_sG_i^{(k)}G_s^{-1}$ при $i = 1, 2, \ldots, g_k$ и любом G_s из G дают всю совокупность элементов класса C_k, мы видим, что матрица $A(C_k)$ коммутирует со всеми матрицами $A(G_s)$. Отсюда следует, что эта матрица $A(G_k)$ кратна единичной матрице [66], так что можем написать:

$$A(C_k) = b_kI \qquad (k = 1, 2, \ldots, r), \tag{171}$$

где b_k — некоторые числа. Принимая во внимание опрецеление чисел a_{pqk}, т. е. символическую формулу (170), получаем следующее соотношение между числами b_k:

$$b_pb_q = \sum_{k=1}^{r} a_{pqk}b_k \tag{172}$$

След матрицы $A(C_k)$ равен сумме следов матриц $A(G_i^{(k)})$ $(i = 1, 2, \ldots, g_k)$, т. е. равен $g_kX(C_k)$. С другой стороны, из (171) следует, что след $A(C_k)$ равен nb_k, т. е. $nb_k = g_kX(C_k)$, откуда

$$b_k = \frac{g_k}{n}X(C_k),$$

и соотношение (172) приводит нас к следующей теореме.

Теорема 5. *Между характерами любого неприводимого представления, образованного матрицами порядка n, имеют место соотношения:*

$$g_pX(C_p)\,g_qX(C_q) = n\sum_{k=1}^{r} a_{pqk}g_kX(C_k). \tag{173}$$

Отметим, что среди классов C_k имеется класс, состоящий только из единичного элемента E группы G. В любом линейном представлении ему соответствует единичная матрица, след которой равен ее порядку n. Этот класс мы будем всегда обозначать через C_1, так что $X(C_1)=n$, и предыдущую формулу можно переписать в виде:

$$g_pX(C_p)\,g_qX(C_q)=X(C_1)\sum_{k=1}^{r}a_{pqk}g_kX(C_k). \tag{174}$$

Определим теперь значения постоянных a_{pq1}. Каждому классу C_p соответствует некоторый класс $C_{p'}$, состоящий из элементов, обратных тем, которые входят в класс C_p. Это следует непосредственно из определения классов и из того, что формула $G_sG_tG_s^{-1}=G_u$ приводит к формуле $G_sG_t^{-1}G_s^{-1}=G_u^{-1}$.

Класс $C_{p'}$ может и совпадать с C_p, т. е. может случиться, что $p'=p$. Во всяком случае классы C_p и $C_{p'}$ содержат одинаковое число элементов, т. е. $g_{p'}=g_p$. Если в формуле (173) или (174) взять $q=p'$, то в правую часть класс C_1 будет входить g_p раз, а при $q\neq p'$, правая часть не содержит C_1, т. е.

$$a_{pq1}=\begin{cases}0 & \text{при } q\neq p',\\ g_p & \text{при } q=p'.\end{cases} \tag{175}$$

78. Регулярное представление группы. Мы уже указывали прием представления любой конечной группы при помощи групп перестановок. Всякую группу перестановок мы можем изобразить как группу преобразований.

Действительно, если имеется, например, перестановка

$$\begin{matrix}1, & 2, & 3, & 4\\ 2, & 4, & 3, & 1,\end{matrix}$$

то ее можно записать в виде линейного преобразования, при котором x_1 переходит в y_2, x_2 в y_4, x_3 в y_3 и x_4 в y_1:

$$\begin{aligned}
y_1&=0x_1+0x_2+0x_3+x_4\\
y_2&=x_1\;\;+0x_2+0x_3+0x_4\\
y_3&=0x_1+0x_2+x_3\;\;+0x_4\\
y_4&=0x_1+x_2\;\;+0x_3+0x_4.
\end{aligned}$$

Рассмотрим следующее представление группы G группой перестановок. Умножаем элементы $G_1, G_2, \ldots, G_m$ справа на некоторый элемент G_s. Это приводит к некоторой перестановке элементов, т. е. в силу сказанного выше, к некоторой матрице P_s, которая и считается соответствующей элементу G_s. Это представление называется обычно *регулярным представлением группы* G. Один из элементов G_k есть единичный элемент группы, который мы, как всегда, будем обозначать буквою E. Этому элементу соответствует единичная матрица P_s, и, следовательно, след этой матрицы равен m, т. е. $X(E)=m$. При умножении элементов $G_1, G_2, \ldots, G_m$ на любой элемент G_s ни один элемент G_k не остается на месте, т. е. в соответствующей матрице все диагональные элементы равны нулю, и в регулярном представлении $X(G_s)=0$ при $G_s\neq E$.

Положим, что при приведении регулярного представления оно содержит представление $\omega^{(k)}$, о котором мы говорили выше, h_k раз. Мы имеем при этом, в силу сказанного выше:

$$\sum_{t=1}^{l}h_tX^{(t)}(G_s)=\begin{cases}0 & \text{при } G_s\neq E,\\ m & \text{при } G_s=E.\end{cases} \tag{176}$$

Умножая обе части этого равенства на $\overline{X^{(k)}(G_s)}$ и суммируя по s, получим в силу (163) и (164):

$$h_k m = m X^{(k)}(E),$$

но $X^{(k)}(E)$ равно порядку матриц в представлении $\omega^{(k)}$, который мы обозначим через n_k, и $\overline{X^{(k)}(E)} = X^{(k)}(E) = h_k$, откуда $h_k = n_k$, и формула (176) может быть переписана в виде:

$$\sum_{t=}^{l} X^{(t)}(E)\, X^{(t)}(G_s) = \sum_{t=1}^{l} n_t X^{(t)}(G_s) = \begin{cases} 0 & \text{при} \quad G_s \neq E. \\ m & \text{при} \quad G_s = E. \end{cases} \tag{177}$$

Мы приходим таким образом к следующей теореме:

Теорема 6. *Регулярное представление содержит каждое неприводимое представление $\omega^{(k)}$ число раз, равное порядку матриц n_k в представлении $\omega^{(k)}$, и для характеров представлений $\omega^{(k)}$ имеет место формула* (177).

Напишем теперь формулу (174) для представлений $\omega^{(t)}$:

$$g_p X^{(t)}(C_p)\, g_q X^{(t)}(C_q) = X^{(t)}(C_1) \sum_{k=1}^{r} a_{pqk} X^{(t)}(C_k)\, g_k,$$

и просуммируем по t от $t = 1$ до $t = l$:

$$g_p g_q \sum_{t=1}^{l} X^{(t)}(C_p)\, X^{(t)}(C_q) = \sum_{k=1}^{r} a_{pqk} \sum_{t=1}^{l} X^{(t)}(C_1)\, X^{(t)}(C_k)\, g_k.$$

Принимая во внимание (177), получим:

$$g_p g_q \sum_{t=1}^{l} X^{(t)}(C_p)\, X^{(t)}(C_q) = a_{pq1} m,$$

т. е. в силу (175):

$$\sum_{t=1}^{l} X^{(t)}(C_p)\, X^{(t)}(C_q) = \begin{cases} 0 & \text{при} \quad q \neq p', \\ \dfrac{m}{g_{p'}} & \text{при} \quad q = p'. \end{cases} \tag{178}$$

Составим l линейных однородных уравнений относительно $x_1, x_2, \ldots, x_r$:

$$\sum_{q=1}^{r} x_q X^{(k)}(C_q) = 0 \quad (k = 1, 2, \ldots, l) \tag{179}$$

и покажем, что полученная система имеет только нулевое решение.

Действительно, умножая обе части (179) на $X^{(k)}(C_p)$ и суммируя по k, получим $x_{p'} = 0$, причем p' может быть любым числом из ряда $1, 2, \ldots, r$. Поскольку система (179) имеет только нулевое решение, число уравнений в этой системе не меньше числа неизвестных, т. е. $l \geqslant r$. Раньше мы доказали неравенство $l \leqslant r$, откуда следует $l = r$, т. е.:

Теорема 7. *Общее число неэквивалентных, неприводимых представлений конечной группы G равно числу классов этой группы.*

Выясним еще одно следствие теоремы 6. Регулярное представление группы G состоит из матриц порядка m. С другой стороны, в силу теоремы 6, оно содержит n_k раз каждое представление $\omega^{(k)}$, состоящее из матриц порядка n_k.

Отсюда следует равенство

$$\sum_{k=1}^{r} n_k^2 = m, \tag{180}$$

что можно формулировать следующим образом:

Теорема 8. *Сумма квадратов порядков неэквивалентных, неприводимых представлений $\omega^{(k)}$ равна порядку группы G.*

79. Примеры представления конечных групп. 1. Рассмотрим абелеву группу G, состоящую из элементов $A_2^k A_1^i$, где $i=0, 1, 2, \ldots, m-1$; $k=0, 1, 2, \ldots, n-1$, причем элементы A_1 и A_2 коммутируют, $A_1^m = E$; $A_2^n = E$, и при $i=k=0$ надо считать $A_2^0 A_1^0 = E$. Каждый отдельный элемент G образует класс, и все неприводимые представления группы имеют первый порядок. Пусть α и β — какие-либо значения корней степени m и n из единицы. Элементу $A_2^k A_1^i$ мы сопоставляем число $\beta^k \alpha^i$ и таким образом, как нетрудно проверить, получаем представление группы. Придавая α и β всевозможные значения упомянутых корней из единицы, получим всего mn различных представлений первого порядка. Общее число классов (т. е. элементов) также равно mn, и тем самым построены все неэквивалентные, неприводимые представления. Аналогично строятся представления и в том случае, когда число „производящих элементов“ (т. е. элементов A_s) не два, а больше.

2. Перейдем к группе диэдра порядка n. Она состоит из $2n$ элементов

$$E,\ A^i,\ T,\ TA^i \quad (i=1, 2, \ldots, n-1),$$

причем:

$$A^n = E; \quad T^2 = E; \quad TAT^{-1} = A^{-1} \quad (T^{-1} = T). \tag{181}$$

Последние из написанных соотношений непосредственно очевидны из геометрического смысла вращений A и T. Из этого соотношения непосредственно следует соотношение $TA^iT^{-1} = A^{-i}$. Положим сначала, что $n = 2m+1$ есть нечетное число. Группа будет при этом состоять из $(m+2)$ классов. Один из них содержит E, m классов содержат по два элемента A^s и A^{-s} $(s = 1, 2, \ldots, m)$, причем $A^{-s} = A^{2l+1-s}$, и один класс содержит все элементы вида T и TA^i. Все это легко проверяется при помощи указанных соотношений.

Имеются два представления первого порядка. В одном из них каждому элементу сопоставляется число 1. В другом элементу A сопоставляется 1 и элементу T число (-1). Пусть, далее, $\varepsilon = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$. Можно построить m представлений второго порядка, сопоставляя элементам A и T следующие матрицы:

$$A \to \begin{Vmatrix} \varepsilon^s & 0 \\ 0 & \varepsilon^{-s} \end{Vmatrix}; \quad T \to \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \qquad (s = 1, 2, \ldots, m). \tag{182}$$

Написанные матрицы удовлетворяют соотношениям (181) и тем самым дают представление группы, ибо всякое соотношение между элементами A и T является следствием соотношений (181). Неприводимость каждого из представлений вытекает из того, что в противном случае представление привелось бы к двум представлениям первого порядка, и матрицы нашего представления должны были бы коммутировать, чего нет на самом деле ни при каком s, в чем легко убедиться.

Неэквивалентность представлений (182) при разных s вытекает из того, что при разных s матрицы, соответствующие элементу A, имеют различные наборы характеристических чисел ε^s и ε^{-s}. Таким образом построены все

$(m+2)$ неэквивалентных, неприводимых представлений. Формула (180) в рассматриваемом примере сводится к следующей:

$$2\cdot 1^2+m\cdot 2^2=4m+2=2n.$$

При четном $n=2m$ представление (182), соответствующее значению $s=m$, имеет вид:

$$A\to\begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix}; \qquad T\to\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

и распадается на два представления первого порядка:

$$A\to(-1);\quad T\to(+1)\quad \text{и}\quad A\to(-1);\quad T\to(-1).$$

Чтобы получить это, достаточно применить такую матрицу S, что STS^{-1} приводится к диагональному виду, причем характеристические числа матрицы T равны, очевидно, ± 1. Таким образом при $n=2m$ имеются четыре представления первого порядка и $(m-1)$ представлений второго порядка. Формула (180) принимает вид:

$$4\cdot 1^2+(m-1)\,2^2=4m=2n.$$

3. Рассмотрим представления группы тетраэдра или, что то же, изоморфной ей знакопеременной группы при $n=4$ [**59**]. Группа состоит из четырех классов и порядок ее равен двенадцати. Она должна иметь четыре неэквивалентных, неприводимых представления. Порядки этих представлений должны удовлетворять равенству

$$n_1^2+n_2^2+n_3^2+n_4^2=12.$$

Это уравнение имеет с точностью до порядка слагаемых в левой части единственное решение в целых положительных числах, а именно

$$n_1=n_2=n_3=1;\quad n_4=3,$$

т. е. группа имеет три представления первого порядка и одно — третьего порядка. В представлениях первого порядка элементам, входящим в один и тот же класс, соответствует одно и то же число. Нетрудно показать, что в трех представлениях первого порядка классам соответствуют следующие числа:

$$\begin{aligned}
&\mathrm{I}\to 1;\quad \mathrm{II}\to 1;\quad \mathrm{III}\to 1;\quad \mathrm{IV}\to 1\\
&\mathrm{I}\to 1;\quad \mathrm{II}\to 1;\quad \mathrm{III}\to \varepsilon;\quad \mathrm{IV}\to \varepsilon^2\\
&\mathrm{I}\to 1;\quad \mathrm{II}\to 1;\quad \mathrm{III}\to \varepsilon^2;\quad \mathrm{IV}\to \varepsilon,
\end{aligned}$$

где

$$\varepsilon=\cos\frac{2\pi}{3}+i\sin\frac{2\pi}{3}.$$

Неприводимое представление третьего порядка дает сама группа тетраэдра, т. е. группа тех вращений пространства (матрицы третьего порядка), при которых тетраэдр переходит в себя. Если бы это представление оказалось приводимым, то оно должно было бы привести к трем представлениям первого порядка, что невозможно хотя бы потому, что группа тетраэдра не есть абелева группа. Изложенная в последних номерах теория касается конечных групп. Для того чтобы перенести ее на группу вращения, мы должны более подробно остановиться на рассмотрении бесконечных групп, зависящих от параметров. Прежде чем переходить к общему рассмотрению таких групп, мы изложим вопрос о линейных представлениях группы Лоренца. Эти представления, наряду с представлениями группы вращения, будут служить для нас основными примерами бесконечных групп, зависящих от параметров.

80. Представления линейной группы с двумя переменными. В [68] мы построили линейные представления унитарной группы с двумя переменными, что привело нас к линейным представлениям групп вращения. Аналогично можно построить представления линейной группы с двумя переменными и с определителем, равным единице:

$$\begin{aligned} x_1' &= a x_1 + b x_2, \\ x_2' &= c x_1 + d x_2, \end{aligned} \quad ad - bc = 1. \tag{183}$$

В силу сказанного в [64] это приведет нас к однозначным и двузначным представлениям группы положительных преобразований Лоренца. Результаты окажутся существенно отличными от результатов из [68].

Одним из возможных линейных представлений унитарной группы (93) является представление этой группы самой собой, т. е. линейное представление, при котором каждому преобразованию (93) соответствует это же преобразование. Легко видеть, что другим линейным представлением является следующее: каждому преобразованию (93) приводится в соответствие преобразование с комплексными сопряженными коэффициентами:

$$y_1' = \bar{a} y_1 + \bar{b} y_2; \qquad y_2' = -b y_1 + a y_2.$$

Но это представление эквивалентно предыдущему, что непосредственно следует из легко проверяемой формулы:

$$\left\| \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \right\| \left\| \begin{matrix} a & b \\ -\bar{b} & \bar{a} \end{matrix} \right\| = \left\| \begin{matrix} \bar{a} & \bar{b} \\ -b & a \end{matrix} \right\| \left\| \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \right\|.$$

Для группы (183) сопряженное представление

$$y_1' = \bar{a} y_1 + \bar{b} y_2; \qquad y_2' = \bar{c} y_1 + \bar{d} y_2 \tag{184}$$

не эквивалентно самой группе (183). Чтобы убедиться в этом, достаточно рассмотреть случай $b = c = 0$. При этом матрица преобразования (183) имеет характеристические числа a и d, а матрица (184) — характеристические числа $\bar{a}$ и $\bar{d}$. Очевидно, можно выбрать комплексные числа a и d, удовлетворяющие условию $ad = 1$ так, что совокупность чисел $\bar{a}$ и $\bar{d}$ будет отлична от совокупности чисел a и d, и, следовательно, соответствующие преобразования не могут быть подобными. Таким образом в данном случае мы уже имеем два неэквивалентных представления второго порядка — саму группу (183) и группу (184). О неприводимости представлений будет сказано ниже.

Мы можем далее построить представления группы (183) совершенно так же, как это было сделано в [68]. В формулах (99) надо только заменить $\bar{a}$ на d и $\bar{b}$ на $(-c)$. Это приведет к следующему представлению порядка $(2j+1)$, где j — целое неотрицательное число или половина нечетного числа:

$$D_j\begin{Bmatrix} a, & b \\ c, & d \end{Bmatrix}_{ls} = \sum_k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times a^{j+l-k} b^k c^{k+s-l} d^{j-k-s} \left(j = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots\right). \quad (185)$$

Здесь l и s пробегают следующий ряд значений:

$$l \text{ и } s = -j,\quad -j+1,\ \ldots,\ j-1,\ j,$$

и суммирование по k определяется неравенствами:

$$k \geqslant 0;\quad k \geqslant l-s;\quad k \leqslant j-s;\quad k \leqslant j+l.$$

В формулах (185) надо считать $0! = 1$ и $0^0 = 1$. При $j = 0$ получается тождественное представление единицей. Кроме представлений (185), мы можем написать непосредственно другие представления, заменив в правых частях (185) числа a, b, c и d сопряженными. Соответствующие представления обозначим следующим образом:

$$\overline{D}_{j'}\begin{Bmatrix} a & b \\ c & d \end{Bmatrix} \left(j' = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots\right). \quad (186)$$

Мы можем составить теперь композицию представлений (185) и (186) [**73**], в результате чего получится новое представление порядка $(2j+1)(2j'+1)$. Обозначим его следующим образом:

$$E_{j,\,j'}\begin{Bmatrix} a & b \\ c & d \end{Bmatrix}. \quad (187)$$

Пользуясь формулами (185), легко выписать элементы матриц, соответствующих этому представлению. Возьмем два различных представления (187), но так, чтобы порядок их был одинаковым:

$$E_{p,\,q}\begin{Bmatrix} a & b \\ c & d \end{Bmatrix} \text{ и } E_{p_1,q_1}\begin{Bmatrix} a & b \\ c & d \end{Bmatrix};\ (2p+1)(2q+1) = (2p_1+1)(2q_1+1).$$

Покажем, что такие два представления не эквивалентны. Положим $b = c = 0$. При этом матрица (185) приведется к диагональной матрице с диагональными элементами:

$$D_j\begin{Bmatrix} a & 0 \\ 0 & d \end{Bmatrix}_{ll} = a^{j+l} d^{j-l} \quad (l = -j,\ -j+1,\ \ldots,\ j-1,\ j).$$

Прямое произведение двух диагональных матриц есть диагональная матрица, и, следовательно, матрицы $E_{p,q}$ и E_{p_1,q_1} при $b=c=0$ имеют следующие характеристические числа:

$$E_{p,q}: a^{p+l}d^{p-l}(\bar{a})^{q+m}(\bar{d})^{q-m} \begin{pmatrix} l=-p, & -p+1, & \ldots, & p-1, & p \\ m=-q, & -q+1, & \ldots, & q-1, & q \end{pmatrix}$$

$$E_{p_1,q_1}: a^{p_1+l_1}d^{p_1-l_1}(\bar{a})^{q_1+m_1}(\bar{d})^{q_1-m_1}$$

$$\begin{pmatrix} l_1=-p_1, & -p_1+1, & \ldots, & p_1-1, & p_1 \\ m_1=-q_1, & -q_1+1, & \ldots, & q_1-1, & q_1 \end{pmatrix}$$

или, принимая во внимание, что $ad=1$:

$$E_{p,q}: a^{2l}(\bar{a})^{2m}; \quad E_{p_1,q_1}: a^{2l_1}(\bar{a})^{2m_1}.$$

В качестве a мы можем взять любое комплексное число, отличное от нуля, и можно, очевидно, выбрать это число так, что совокупность характеристических чисел матрицы $E_{p,q}$ будет отлична от совокупности характеристических чисел матрицы E_{p_1,q_1}, что и доказывает неэквивалентность представлений (187) при различных выборах j и j'. Отметим, что при $j'=0$ представление (187) совпадает с представлением (185), а при $j=0$ оно совпадает с тем представлением, которое получается из (185) при $j=j'$ и при замене a, b, c и d сопряженными величинами. Отметим одну особенность представлений (187). Эти представления не эквивалентны унитарным представлениям. Если бы они были эквивалентны некоторым унитарным представлениям, то все характеристические числа любой матрицы представления должны иметь модуль, равный единице, а выше мы видели, что у представления $E_{p,q}$ эти характеристические числа при $b=c=0$ равны $a^{2l}(\bar{a})^{2m}$ и могут быть по модулю, очевидно, отличными от единицы. Исключением является лишь представление $E_{0,0}$, которое является тривиальным тождественным представлением, при котором каждому элементу группы (183) соответствует единица.

В [**66**] мы видели, что если некоторое представление, не обязательно эквивалентное унитарному, приводимо, т. е. эквивалентно некоторому представлению с квазидиагональными матрицами одной и той же структуры, то обязательно существует матрица, отличная от кратной единичной матрицы и коммутирующая со всеми матрицами представления. Таким образом, для того, чтобы доказать, что любое представление (187) не является приводимым, достаточно показать, что если некоторая матрица коммутирует со всеми матрицами представления (187), то эта матрица кратна единичной матрице. Это можно сделать совершенно так же, как и в [**68**]. Итак, представления (187) попарно неэквивалентны и каждое из них является неприводимым представлением. Часто пользуются определением приводимости, отличным от того, которое мы привели в [**58**], а именно представление называют приводимым, если все его линейные преобразования (пусть их порядок равен n) оставляют инвариантным некоторое подпространство L_k, причем $0 < k < n$.

Мы видели [**58**], что если приводимое в этом смысле представление состоит из унитарных матриц, то оно приводимо и в смысле определения из [**65**], т. е. эквивалентно некоторому квазидиагональному представлению. Если представление не унитарно, то из инвариантности некоторого подпространства не следует приводимость в смысле определения из [**65**]. Можно показать, что всякое представление (187) группы не только неприводимо в том смысле, как это мы доказали, но что оно не оставляет инвариантным никакое подпространство. Кроме того, можно показать, что всякое линейное представление группы (183) или эквивалентно одному из представлений (187), или эквивалентно представлению, имеющему приведенную формулу и состоящему из нескольких представлений (187).

В [**73**] мы видели, что композиция двух линейных представлений группы равносильна перемножению объектов тех линейных представлений, которые входят в эту композицию. Принимая это во внимание, мы можем утверждать, что для представлений (187) объектами представлений являются выражения:

$$\eta_{kk'} = \frac{x_1^{j+k} x_2^{j-k}}{\sqrt{(j+k)!\,(j-k)!}} \cdot \frac{y_1^{j'+k'} y_2^{j'-k'}}{\sqrt{(j'+k')!\,(j'-k')!}}$$

$$\begin{pmatrix} k = j, & j-1, & \ldots, & -j+1, & -j \\ k' = j', & j'-1, & \ldots, & -j'+1, & -j' \end{pmatrix},$$

причем x_1 и x_2 испытывают преобразование (183), а y_1, y_2 — преобразование (184).

Мы говорили до сих пор о линейных представлениях группы, состоящей из положительных преобразований Лоренца [**64**]. Эти положительные преобразования составляют лишь часть преобразований Лоренца с определителем, равным единице. Кроме того, имеются преобразования Лоренца и с определителем (-1). Исследование структуры этих более общих множеств преобразований и расширение линейных представлений группы положительных преобразований Лоренца на случай полной группы Лоренца представляет некоторые особенности по сравнению с группой ортогональных преобразований в трехмерном пространстве. Отметим, что при определении полной группы Лоренца мы можем поставить требование неизменности в направлении отсчета времени. При этом к рассмотренной группе Лоренца надо добавить отражение:

$$x_1' = -x_1; \quad x_2' = -x_2; \quad x_3' = -x_3; \quad x_4' = x_4.$$

Исследование всех указанных вопросов можно найти, например, в книге Картана „Теория спиноров“ (Москва, 1947 г.) и в книге Ван-дер-Вардена „Метод теории групп в квантовой механике“.

81. Теорема о простоте группы Лоренца. Пользуясь методом, аналогичным тому, который мы применили в [**70**], докажем сейчас, что группа Лоренца есть простая группа. Для этого достаточно показать, что у группы G, состоящей из преобразований (183), нет нормальных делителей, отличных от нормального делителя, состоящего из матриц E и $(-E)$. Пусть имеем такой нормальный делитель H_1, содержащий матрицу

$$A = \begin{Vmatrix} a & b \\ c & d \end{Vmatrix} \qquad (ad - bc = 1),$$

отличную от E и $(-E)$. Надо доказать, что H_1 совпадает с G. Если H_1 содержит некоторую матрицу B, то подгруппа H_1 содержит и все матрицы $U^{-1}BU$, где U — любая матрица из G. Принимая во внимание основной результат о приведении матриц к каноническому виду, а также тот факт, что определитель матрицы U, приводящей какую-либо матрицу к каноническому виду, можно всегда считать равным единице [**27**], мы видим, что достаточно показать, что H_1 содержит, во-первых, матрицы с любыми допустимыми различными характеристическими числами t и t^{-1}, где t — любое комплексное число, отличное от нуля и $(+1)$. Отметим, что произведение характеристических чисел матриц группы G должно равняться единице. Далее, H_1 должно содержать матрицы E и $(-E)$, и, кроме того, учитывая случай равных характеристических чисел и двойного элементарного делителя, мы должны еще показать, что H_1 содержит матрицы:

$$\begin{Vmatrix} 1 & 0 \\ 1 & 1 \end{Vmatrix} \quad \text{и} \quad \begin{Vmatrix} -1 & 0 \\ 1 & -1 \end{Vmatrix}. \tag{188}$$

Возьмем переменную матрицу группы G:

$$X = \begin{Vmatrix} x & y \\ z & x \end{Vmatrix} \qquad (x^2 - yz = 1)$$

и составим матрицу:

$$Y = A(XA^{-1}X^{-1}),$$

которая, как и в [**70**], должна входить в H_1. Для следа s матрицы Y получаем выражение:

$$s = 2 + bz^2 + cy^2 - [(a-d)^2 + 2bc]\,yz.$$

Поскольку A отлична от E и $(-E)$, мы не будем иметь одновременно: $b = c = 0$ и $a = d$. Отсюда ясно, что s не есть постоянная, и, меняя z и y, мы можем придавать s любые комплексные значения. Характеристические числа матрицы Y определяются из квадратного уравнения

$$\lambda^2 - s\lambda + 1 = 0.$$

Таким образом мы можем получать для этих корней любые значения t и t^{-1}, следовательно, H_1 содержит все матрицы с различными характеристическими числами и определителем единица. Далее, H_1, очевидно, содержит E, а также $(-E)$, которое можно представить в виде произведения:

$$-E=[t, t^{-1}]\cdot[-t^{-1}, -t],$$

каждый из множителей которого входит в H_1. Матрицы (188) легко представить в виде произведения двух матриц с определителем единица и с различными характеристическими числами, откуда следует, что H_1 содержит и эти матрицы. Действительно:

$$\left\|\begin{matrix}1 & 0\\ 1 & 1\end{matrix}\right\| = \left\|\begin{matrix}\frac{1}{\beta} & 0\\ 0 & \beta\end{matrix}\right\| \cdot \left\|\begin{matrix}\beta & 0\\ \frac{1}{\beta} & \frac{1}{\beta}\end{matrix}\right\|,$$

$$\left\|\begin{matrix}-1 & 0\\ 1 & -1\end{matrix}\right\| = \left\|\begin{matrix}\frac{1}{\beta} & 0\\ 0 & \beta\end{matrix}\right\| \cdot \left\|\begin{matrix}-\beta & 0\\ \frac{1}{\beta} & -\frac{1}{\beta}\end{matrix}\right\| \qquad (\beta \neq 0 \text{ и } \pm 1).$$

Таким образом доказано, что H_1 должно совпадать с G, т. е. G не имеет нормальных делителей, отличных от нормального делителя, состоящего из E и $(-E)$, и тем самым доказано, что группа положительных преобразований Лоренца есть простая группа. Отсюда следует, как и в [**70**], что эта группа не может иметь гомоморфных (не изоморфных) представлений.

§ 7. НЕПРЕРЫВНЫЕ ГРУППЫ

82. Непрерывные группы. Структурные постоянные. Группы вращений трехмерного пространства и группа положительных преобразований Лоренца представляют собою примеры бесконечных групп, элементы которых зависят от параметров, которые могут меняться непрерывным образом. Для группы вращения роль параметров могли играть, например, углы Эйлера. В рассматриваемых случаях группы состоят из линейных преобразований, и зависимость группы от параметров сводится к тому, что элементы матриц, которыми определяются упомянутые линейные преобразования, зависят от этих параметров. В дальнейшем мы будем рассматривать группы линейных преобразований.

Положим, что элементы a_{ik} матриц линейных преобразований, составляющих некоторую группу G, суть функции r вещественных параметров $\alpha_1, \alpha_2, \ldots, \alpha_r$, причем выполнены условия, которые мы сейчас укажем. Положим, что a_{ik} суть однозначные функции параметров a_s при всех значениях этих параметров, достаточно близких к нулю, и что нулевым значениям параметров $\alpha_1=\alpha_2=\ldots=\alpha_r=0$.

соответствует единичный элемент группы G, который характеризуется условиями: $a_{ik}=0$ при $i \neq k$ и $a_{ii}=1$. Положим далее, что всякому элементу группы G, достаточно близкому к единичному элементу, соответствуют определенные значения параметров α_s, достаточно близкие к нулю. Близость элемента группы к единичному элементу определяется тем, что элементы a_{ik} соответствующих матриц близки к нулю при $i \neq k$ и к единице при $i=k$. Таким образом, при сделанных предположениях, мы будем иметь биоднозначное соответствие между элементами группы G, находящимися в определенной окрестности единичного элемента, и некоторой окрестностью начала координат r-мерного вещественного пространства T_r. В дальнейшем мы будем иметь не только такое локальное биоднозначное соответствие, но биоднозначное соответствие в целом, при котором каждому элементу группы G соответствует определенная точка пространства T_r, принадлежащая некоторой области V этого пространства, содержащей начало внутри себя, и, наоборот, любой точке из V отвечает определенный элемент группы G Пока нам потребуется лишь указанное выше локальное соответствие. Символами G_α, G_β, G_γ и т. д. мы будем обозначать те элементы группы G, которые соответствуют значениям параметров α_s, β_s, γ_s $(s=1, 2, \ldots, r)$. При локальной точке зрения надо считать, что параметры достаточно близки к нулю, а элементы группы — к единичному элементу.

Рассмотрим произведение каких-либо элементов группы

$$G_\beta G_\alpha = G_\gamma.$$

Параметры γ_s, характеризующие элемент G_γ, полученный в результате указанного умножения, суть однозначные функции параметров α_s и β_s:

$$\gamma_s = \varphi_s(\beta_1, \beta_2, \ldots, \beta_r; \alpha_1, \alpha_2, \ldots, \alpha_r). \tag{189}$$

Мы предполагаем, что это суть непрерывные функции, имеющие непрерывные производные до четвертого порядка при всех α_s и β_s, достаточно близких к нулю.

Из того, что нулевым значениям параметров соответствует единичный элемент группы, непосредственно следуют равенства

$$\left.\begin{aligned} &\varphi_s(\beta_1, \beta_2, \ldots, \beta_r; \ 0, 0, \ldots, 0) = \beta_s; \\ &\varphi_s(0, 0, \ldots, 0; \ \alpha_1, \alpha_2, \ldots, \alpha_r) = \alpha_s \end{aligned}\right\} \quad (s=1, 2, \ldots, r), \tag{190}$$

откуда

$$\left.\begin{aligned} &\frac{\partial \varphi_i}{\partial \beta_k} = \delta_{ik} \quad \text{при} \quad \alpha_s = 0; \\ &\frac{\partial \varphi_i}{\partial \alpha_k} = \delta_{ik} \quad \text{при} \quad \beta_s = 0 \end{aligned}\right\} \quad (s=1, 2, \ldots, r). \tag{191}$$

Параметры $\tilde{\alpha}_s$, соответствующие обратному элементу G_α^{-1}, определяются, очевидно, из соотношений:

$$\varphi_s(\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_r; \ \alpha_1, \alpha_2, \ldots, \alpha_r) = 0 \quad (s = 1, 2, \ldots, r), \tag{192}$$

причем написанные равенства имеют место, если положить все α_s и все $\tilde{\alpha}_s$ равными нулю. Функциональный определитель от левых частей уравнений (192) по $\tilde{\alpha}_s$ равен, в силу (191), единице при α_s и $\tilde{\alpha}_s$, равных нулю. Таким образом, в силу теоремы о неявных функциях, уравнения (192) определяют $\tilde{\alpha}_s$ как непрерывные функции при всех α_s, достаточно близких к нулю, причем $\tilde{\alpha}_s$ обращаются в нуль при $\alpha_s = 0$. Разложим функции (189) по степеням α_s и β_s, пользуясь формулой Маклорена, причем доведем разложение до членов третьего порядка. Принимая во внимание формулы (190) и (191), получим:

$$\gamma_s = \alpha_s + \beta_s + \\ + \sum_{i,k} a_{i,k}^{(s)} \alpha_i \beta_k + \sum_{i,k,l} a_{i,k,l}^{(s)} \alpha_i \alpha_k \beta_l + \sum_{i,k,l} b_{i,k,l}^{(s)} \alpha_i \beta_k \beta_l + \varepsilon^{(s)}, \tag{193}$$

где $a_{i,k}^{(s)}$, $a_{i,k,l}^{(s)}$ и $b_{i,k,l}^{(s)}$ — численные коэффициенты, $\varepsilon^{(s)}$ — не ниже четвертого порядка малости относительно α_s и β_s, и суммирование по i, k и l производится от 1 до r. Числа

$$C_{ik}^{(s)} = a_{ik}^{(s)} - a_{ki}^{(s)} \quad (s, i, k = 1, 2, \ldots, r) \tag{194}$$

называются *структурными постоянными группы* G при принятом выборе параметров α_s.

Если вместо α_s ввести другие параметры α'_s:

$$\alpha_s = \omega_s(\alpha'_1, \alpha'_2, \ldots, \alpha'_r) \quad (s = 1, 2, \ldots, r)$$

так, что $\omega_s(0, 0, \ldots, 0) = 0$, написанные равенства однозначно разрешимы относительно α'_s при всех α_s, достаточно близких к нулю, и функции ω_s имеют достаточное число производных, то структурные постоянные в новых параметрах α'_s будут уже другими.

Из определения (194) непосредственно следует:

$$C_{ki}^{(s)} = -C_{ik}^{(s)}. \tag{194$_1$}$$

Кроме того, пользуясь (192) и законом ассоциативности при перемножении элементов группы G, можно доказать еще следующее соотношение между структурными постоянными:

$$\sum_{s=1}^{r} (C_{is}^{(t)} C_{jk}^{(s)} + C_{js}^{(t)} C_{ki}^{(s)} + C_{ks}^{(t)} C_{ij}^{(s)}) = 0 \ (i, j, k, t = 1, 2, \ldots, r). \tag{194$_2$}$$

Мы не будем пользоваться этим соотношением и не приводим его доказательства.

Вернемся к формулам (193). При α_s и β_s, достаточно близких к нулю, и величины γ_s будут близкими к нулю. Принимая во внимание формулы (191) и теорему о неявных функциях, можно утверждать, что уравнения (193) в некоторой окрестности начала координат пространства T_r разрешимы относительно β_s:

$$\beta_s = \psi_s(\gamma_1, \gamma_2, \ldots, \gamma_r; \quad \alpha_1, \alpha_2, \ldots, \alpha_r) \quad (s = 1, 2, \ldots, r). \tag{195}$$

Отметим при этом, что условия: $\beta_s = 0$ $(s = 1, 2, \ldots, r)$ равносильны условиям: $\gamma_s = \alpha_s$ $(s = 1, 2, \ldots, r)$. Пользуясь формулами (193) и (195), составим две квадратные матрицы $S(\alpha_s)$ и $T(\alpha_s)$ порядка r с элементами $S_{ik}(\alpha_s)$ и $T_{ik}(\alpha_s)$, зависящими от параметров α_s:

$$S_{ik}(\alpha_s) = \left(\frac{\partial\gamma_i}{\partial\beta_k}\right)_{\beta_s=0}; \qquad T_{ik}(\alpha_s) = \left(\frac{\partial\beta_i}{\partial\gamma_k}\right)_{\gamma_s=\alpha_s}$$

$$(s, i, k = 1, 2, \ldots, r). \tag{196}$$

Принимая во внимание правило дифференцирования сложных функций и вычисляя производную от γ_i по γ_k или производную от β_i по β_k, получим:

$$S(\alpha_s)\,T(\alpha_s) = E \quad \text{и} \quad T(\alpha_s)\,S(\alpha_s) = E, \tag{197}$$

где E — единичная матрица порядка r. Из формул (191) следует, что $S(\alpha_s)$ при нулевых значениях $\alpha_s = 0$ обращается в единичную матрицу. Из (197) при этом следует, что и $T(\alpha_s)$ обладает этими же свойствами. Нетрудно выразить структурные постоянные $C_{ik}^{(p)}$ через элементы упомянутых матриц, а именно:

$$C_{ik}^{(p)} = \left(\frac{\partial S_{pk}(\alpha_s)}{\partial\alpha_i} - \frac{\partial S_{pi}(\alpha_s)}{\partial\alpha_k}\right)_{\alpha_s=0} \tag{198}$$

или

$$C_{ik}^{(p)} = \left(\frac{\partial T_{pi}(\alpha_s)}{\partial\alpha_k} - \frac{\partial T_{pk}(\alpha_s)}{\partial\alpha_i}\right)_{\alpha_s=0}. \tag{199}$$

Действительно, в силу (193) и (196), получим:

$$a_{ik}^{(p)} = \left(\frac{\partial\gamma_p}{\partial\alpha_i\partial\beta_k}\right)_{\alpha_s=\beta_s=0} = \left(\frac{\partial S_{pk}(\alpha_s)}{\partial\alpha_i}\right)_{\alpha_s=0}, \tag{200}$$

и, переставляя значки i и k, можем написать:

$$a_{ki}^{(p)} = \left(\frac{\partial S_{pi}(\alpha_s)}{\partial\alpha_k}\right)_{\alpha_s=0}, \tag{201}$$

откуда и следует непосредственно формула (198). Далее, принимая во внимание (197), имеем:

$$\sum_{j=1}^{r} S_{pj}(\alpha_s)\,T_{jk}(\alpha_s) = \delta_{pk}.$$

Дифференцируем обе части по α_l и полагаем затем все α_s равными нулю. Принимая во внимание, что матрицы $S(\alpha_s)$ и $T(\alpha_s)$ обращаются в единичную матрицу при $\alpha_s=0$ $(s=1, 2, \ldots, r)$, получим:

$$\left(\frac{\partial S_{pk}(\alpha_s)}{\partial \alpha_i}\right)_{\alpha_s=0}+\left(\frac{\partial T_{pk}(\alpha_s)}{\partial \alpha_i}\right)_{\alpha_s}=0,$$

т. е., в силу (201):

$$a_{ik}^{(p)}=-\left(\frac{\partial T_{pk}(\alpha_s)}{\partial \alpha_i}\right)_{\alpha_s=0},$$

откуда, как и выше, следует формула (199). Формулы (193) определяют основную групповую операцию, которая по параметрам α_s и β_s элементов G_α и G_β группы G дает параметры γ_s, соответствующие произведению $G_\beta G_\alpha$. Из выражений (193) видно, что при α_s и β_s, близких к нулю, в первом приближении групповая операция сводится к следующей: $\gamma_s=\alpha_s+\beta_s$, так что в первом приближении группа оказывается абелевой. Если группа в точности есть абелева, то:

$$\varphi_s(\beta_1\beta_2, \ldots, \beta_r;\quad \alpha_1\alpha_2, \ldots, \alpha_r)=\varphi_s(\alpha_1\alpha_2, \ldots, \alpha_r;\quad \beta_1\beta_2, \ldots, \beta_r)$$
$$(s=1, 2, \ldots, r)$$

и в разложениях (193) $a_{ki}^{(s)}=a_{ik}^{(s)}$, т. е. у абелевой группы все структурные постоянные равны нулю. Для общих групп уже члены второго измерения в разложениях (193) дают уклонение от коммутативности, что и характеризуется наличием структурных постоянных, отличных от нуля. Пользуясь разложением (193), нетрудно получить и разложение параметров $\tilde{\alpha}_s$, соответствующих элементу G_α^{-1}. Для этого в формулах (193) надо положить $\gamma_s=0$ и заменить β_s на $\tilde{\alpha}_s$. Применяя обычные правила дифференцирования неявных функций, получим:

$$\tilde{\alpha}_s=-\alpha_s+\sum_{i,k} a_{ik}^{(s)}\alpha_i\alpha_k+\varepsilon_1^{(s)},$$

где $\varepsilon_1^{(s)}$, по крайней мере, третьего порядка малости относительно $\alpha_1, \alpha_2, \ldots, \alpha_r$.

83. Бесконечно малые преобразования. Пусть, как и выше, имеется непрерывная группа G линейных преобразований порядка n, определяемая параметрами α_s $(s=1, 2, \ldots, r)$. Будем, как и выше, обозначать символом G_α — матрицу преобразования, соответствующего параметрам α_s, так что линейное преобразование имеет вид:

$$\boldsymbol{x}=G_\alpha \boldsymbol{u}, \tag{202}$$

где $\boldsymbol{u}$ — любой вектор n-мерного комплексного пространства R_n и $\boldsymbol{x}$—преобразованный вектор. Введем операцию дифференцирования матрицы, а именно: если элементы некоторой матрицы A суть дифференцируемые функции некоторого параметра t, то производной от матрицы A по параметру t назовем матрицу, элементы которой получаются дифференцированием элементов матрицы A по t, т. е.

$$\left\{\frac{dA}{dt}\right\}_{ik} = \frac{d\,\{A\}_{ik}}{dt}.$$

Если элементы A зависят от нескольких переменных, то мы будем иметь частные производные.

Совершенно так же, если составляющие некоторого вектора $\boldsymbol{z}(z_1, z_2, \ldots, z_n)$ из пространства R_n суть дифференцируемые функции t, то вектор $\frac{d\boldsymbol{z}}{dt}$ определяется как вектор с составляющими $\frac{dz_i}{dt}$, т. е. дифференцирование вектора сводится к дифференцированию его составляющих [**II, 119**].

Введем теперь так называемые *бесконечно малые преобразования группы* G:

$$I_k = \left(\frac{\partial G_\alpha}{\partial \alpha_k}\right)_{\alpha_s = 0} \quad (k = 1, 2, \ldots, r). \tag{203}$$

Символ I_k обозначает, очевидно, некоторую матрицу порядка n с численными элементами.

Обратимся теперь к формуле (202) и положим, что $\boldsymbol{u}$ есть фиксированный вектор, т. е. его составляющие не зависят от α_s. Преобразованный вектор уже зависит, вообще говоря, от этих параметров, и мы выведем сейчас основные дифференциальные уравнения для этого вектора. Для этого применим к обеим частям (202) линейную операцию, определяемую матрицей G_β:

$$G_\beta \boldsymbol{x} = G_\gamma \boldsymbol{u},$$

где $G_\gamma = G_\beta G_\alpha$, и параметры γ_s определяются через α_s и β_s согласно основной групповой операции (193). Дифференцируем обе части последней формулы по β_p и полагаем затем $\beta_s = 0$, т. е. $\gamma_s = \alpha_s$. Пользуясь определением (203), получим:

$$I_p \boldsymbol{x} = \sum_{j=1}^{r} \left[\frac{\partial\,(G_\gamma \boldsymbol{u})}{\partial \gamma_j}\right]_{\gamma_s = \alpha_s} \left(\frac{\partial \gamma_j}{\partial \beta_p}\right)_{\beta_s = 0}.$$

Первый сомножитель под знаком суммы равен, очевидно, производной от правой части (202) по α_j, и, принимая во внимание обозначение (196), можем переписать последнюю формулу в виде:

$$I_p \boldsymbol{x} = \sum_{j=1}^{r} S_{jp}(\alpha_s) \frac{\partial \boldsymbol{x}}{\partial \alpha_j} \qquad (p = 1, 2, \ldots, r).$$

Если ввести векторы:

$$X\left(\frac{\partial \boldsymbol{x}}{\partial \alpha_1}, \frac{\partial \boldsymbol{x}}{\partial \alpha_2}, \ldots, \frac{\partial \boldsymbol{x}}{\partial \alpha_r}\right) \quad \text{и} \quad Y(I_1 \boldsymbol{x}, I_2 \boldsymbol{x}, \ldots, I_r \boldsymbol{x}),$$

то предыдущие формулы можно записать в виде линейного преобразования

$$Y = S^*(\alpha_s) X,$$

где $S^*(\alpha_s)$ — обычное обозначение транспонированной матрицы. Умножая слева на $T^*(a_s)$ и принимая во внимание (197), получим:

$$X = T^*(\alpha_s) Y,$$

или, в раскрытом виде:

$$\frac{\partial x}{\partial \alpha_p} = \sum_{j=1}^{r} T_{jp}(\alpha_s) I_j x \qquad (p = 1, 2, \ldots, r). \tag{204}$$

Для составляющей x_k вектора x, определяемого формулой (202), мы имеем:

$$\frac{\partial x_k}{\partial \alpha_p} = \sum_{j=1}^{r} T_{jp}(\alpha_s) \sum_{t=1}^{n} \{I_j\}_{kt} x_t \quad \begin{pmatrix} k = 1, 2, \ldots, n \\ p = 1, 2, \ldots, r \end{pmatrix}, \tag{205}$$

где $\{I_j\}_{kt}$ — составляющие матрицы I_j. К уравнению (204) для x мы должны добавить начальное условие, непосредственно вытекающее из формулы (202):

$$x\,|_{\alpha_s = 0} = u, \tag{206}$$

где u — произвольный заданный вектор. Отметим, что величины $T_{jp}(\alpha_s)$, входящие в коэффициенты уравнения (204), определяются непосредственно по групповой операции (193). Уравнение (204) приведет нас к некоторым соотношениям между I_j. Для этого достаточно написать, что вторая производная от x по α_p и α_q не зависит от порядка дифференцирования.

Из (204) следует:

$$\frac{\partial^2 x}{\partial \alpha_p \partial \alpha_q} = \sum_{j=1}^{r} \left(\frac{\partial T_{jp}(\alpha_s)}{\partial \alpha_q} I_j x + T_{jp}(\alpha_s) I_j \frac{\partial x}{\partial \alpha_q} \right)$$

или, заменяя величину $\dfrac{\partial x}{\partial \alpha_q}$ ее выражением из (204) при $p = q$, получим:

$$\frac{\partial^2 x}{\partial \alpha_p \partial \alpha_q} = \sum_{j=1}^{r} \frac{\partial T_{jp}(\alpha_s)}{\partial \alpha_q} I_j x + \sum_{j=1}^{r} \sum_{k=1}^{r} T_{jp}(\alpha_s) T_{kq}(\alpha_s) I_j I_k x.$$

Переставляя справа p и q и приравнивая полученную правую часть написанной выше, будем иметь следующее следствие системы (205):

$$\left[\sum_{j=1}^{r} \left(\frac{\partial T_{jp}(\alpha_s)}{\partial \alpha_q} - \frac{\partial T_{jq}(\alpha_s)}{\partial \alpha_p} \right) I_j + \right.$$

$$\left. + \sum_{j=1}^{r} \sum_{k=1}^{r} (T_{jp}(\alpha_s) T_{kq}(\alpha_s) - T_{jq}(\alpha_s) T_{kp}(\alpha_s)) I_j I_k \right] x = 0. \tag{207}$$

Положим в этом соотношении все α_s равными нулю. Принимая во внимание формулы (199) и тот факт, что $T(\alpha_s)=E$, если все α_s равны нулю, получим:

$$\left[\sum_{j=1}^{r} C_{pq}^{(j)} I_j + (I_p I_q - I_q I_p)\right] u = 0,$$

откуда следуют, в силу произвольности вектора u, следующие соотношения между бесконечно малыми преобразованиями:

$$I_q I_p - I_p I_q = \sum_{j=1}^{r} C_{pq}^{(j)} I_j \qquad (p,\ q=1,\ 2,\ \ldots,\ r). \tag{208}$$

Мы определили I_j и доказали соотношения (208), исходя от заданной непрерывной группы G и пользуясь уравнением (204). Покажем, что это уравнение или, что то же, система (205), имеет единственное решение при заданном начальном условии (206). Пусть имеются два таких решения. В силу линейности уравнения (204) их разность также должна удовлетворять уравнению и должна обращаться в нулевой вектор при $\alpha_s=0$. Таким образом, надо показать, что решение x уравнения (204) при нулевом начальном условии равно тождественно нулю. Для простоты письма будем считать $r=3$. Пусть $x(\alpha_1, \alpha_2, \alpha_3)$ — упомянутое решение. Напишем уравнение (204) при $p=1$ и в правой части положим $\alpha_2=\alpha_3=0$. Получится обыкновенное дифференциальное уравнение с независимой переменной α_1 и нулевым начальным условием. В силу известной теоремы единственности [**II,50**], оно тождественно равно нулю, т. е. $x(\alpha_1\ 0,\ 0)\equiv 0$. Напишем теперь уравнение (204) при $p=2$ и в правой части положим $\alpha_3=0$. Это обыкновенное дифференциальное уравнение с независимой переменной α_2 имеет, как мы только что показали, нулевое начальное условие: $x(\alpha_1,\ \alpha_2,\ 0)=0$ при $\alpha_2=0$, а потому, в силу теоремы единственности $x(\alpha_1,\ \alpha_2,\ 0)=0$. Напишем теперь уравнение (204) при $p=3$. Это обыкновенное дифференциальное уравнение имеет нулевое начальное условие: $x(\alpha_1,\ \alpha_2,\ \alpha_3)=0$ при $\alpha_3=0$, и, следовательно, $x(\alpha_1,\ \alpha_2,\ \alpha_3)\equiv 0$, что мы и хотели доказать.

Таким образом уравнение (204) может приводить только к одному конечному преобразованию (202) при заданных бесконечно малых преобразованиях I_j и заданных $T_{jp}(\alpha_s)$, которые определяются групповой операцией (193). Иначе говоря, бесконечно малые преобразования определяют группу. Это будет нам существенно в дальнейшем. Доказательство существования решения уравнения (204) основано на одной общей теореме из уравнений с частными производными, которая применительно к уравнению (204) формулируется следующим образом: для того чтобы уравнение (204) имело решение при любом начальном условии (206), необходимо и достаточно, чтобы квадратная скобка, входящая в формулу (207), при любом выборе p и q была равна нулю тождественно относительно α_s. Дальше мы не будем пользоваться этой теоремой существования.

84. Группа вращения. Рассмотрим в качестве примера группу вращения пространства вокруг начала координат. Соответствующие матрицы третьего порядка зависят от трех параметров. Роль этих параметров могут играть, например, три угла Эйлера. Мы введем сейчас другие параметры α_1, α_2, α_3, в которых и будем производить дальше все вычисления. Всякое вращение мы можем рассматривать как вращение против часовой стрелки вокруг некоторой направленной оси l, выходящей из начала координат, на угол, не превышающий π. При этом два вращения на угол π относительно противоположно направленных осей приводят к одному и тому же конечному положению. Мы можем таким образом изобразить всякое вращение вектором из начала, направленным по оси вращения и по длине равным углу вращения. Проекции этого вектора $(\alpha_1, \alpha_2, \alpha_3)$ на координатные оси и будут служить нам параметрами.

Если мы возьмем сферу V с центром в начале и радиусом π и отождествим концы любого из ее диаметров, то между точками $(\alpha_1, \alpha_2, \alpha_3)$ сферы V и элементами группы вращения будет установлено биоднозначное соответствие. В данном случае оно будет иметь место не только в окрестности начала координат и единичного элемента группы, но и для всей группы, если взять всю сферу V. Можно выразить все матрицы, входящие в группу вращения, через параметры α_1, α_2, α_3 и убедиться в непрерывности и существовании производных, о чем мы говорили выше.

Мы не будем выводить формулу (193) для основной групповой операции в рассматриваемом случае, а определим структурные постоянные, вычисляя непосредственно матрицы бесконечно малых преобразований.

Для вычисления I_1 мы можем считать $\alpha_2 = \alpha_3 = 0$, продифференцировать матрицу преобразования по α_1 и положить затем $\alpha_1 = 0$. Но при $\alpha_2 = \alpha_3 = 0$ мы имеем вращение вокруг оси X на угол α_1, что приводит к формулам:

$$
\begin{aligned}
x_1' &= x_1,\\
x_2' &= \quad x_2 \cos\alpha_1 - x_3 \sin\alpha_1,\\
x_3' &= \quad x_2 \sin\alpha_1 + x_3 \cos\alpha_1.
\end{aligned}
$$

Дифференцируя матрицу этого преобразования по α_1 и полагая затем $\alpha_1 = 0$, получим:

$$
I_1 = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{Vmatrix}. \tag{209}
$$

Совершенно аналогично

$$I_2 = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{Vmatrix}; \qquad I_3 = \begin{Vmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}. \tag{210}$$

После этого мы можем непосредственно вычислить левые части соотношений (208) и тем самым определить структурные постоянные. Это элементарное вычисление приводит к трем следующим соотношениям:

$$I_1 I_2 - I_2 I_1 = I_3; \; I_2 I_3 - I_3 I_2 = I_1; \; I_3 I_1 - I_1 I_3 = I_2. \tag{211}$$

Если правую часть формулы (202) разложим по степеням α_s, ограничиваясь членами первого порядка, то получим с этой точностью:

$$\boldsymbol{x} \doteq \boldsymbol{u} + (\alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3)\, \boldsymbol{u}.$$

Таким образом вектор $\boldsymbol{u}$ в результате указанного преобразования испытывает следующее изменение:

$$\delta \boldsymbol{u} \doteq \alpha_1 I_1 \boldsymbol{u} + \alpha_2 I_2 \boldsymbol{u} + \alpha_3 I_3 \boldsymbol{u}.$$

Каждое слагаемое справа дает изменение $\boldsymbol{u}$ при малом вращении вокруг одной из осей координат. Так, например, мы получаем следующие изменения составляющих (u_1, u_2, u_3) вектора $\boldsymbol{u}$ при повороте на малый угол α_1 вокруг оси X:

$$\delta u_1 \doteq 0; \; \delta u_2 \doteq -u_3 \alpha_1; \; \delta u_3 \doteq u_2 \alpha_1.$$

При этом, как и выше, мы ограничиваемся лишь членами первого порядка относительно α_1.

85. Бесконечно малые преобразования и представления группы вращения. Сейчас выясним связь сказанного выше о бесконечно малых преобразованиях с представлением группы вращения. Мы будем подразумевать биоднозначное представление в окрестности тождественного преобразования матрицами $F(\alpha_1, \alpha_2, \alpha_3)$ порядка n, причем элементы матрицы считаются непрерывными и дифференцируемыми функциями параметров $\alpha_1, \alpha_2, \alpha_3$. Каждое вращение D может быть получено в виде произведения конечного числа вращений из упомянутой окрестности, и произведение соответствующих матриц представления дает представления и для D. Но таким образом в целом может получиться и многозначное представление группы вращения, поскольку при непрерывном изменении параметров вращения можно, возвращаясь к исходному вращению, получить для него новое представление. Это, например, было раньше при двузначном представлении полной группы вращения [**69**].

Для матриц $F(\alpha_1, \alpha_2, \alpha_3)$ мы имеем ту же групповую операцию, что и для самих вращений, а следовательно, и те же структурные постоянные. Для группы G' матриц $F(\alpha_1, \alpha_2, \alpha_3)$ можно составить бесконечно малые преобразования I_k. Это будут некоторые матрицы порядка n, связанные соотношениями (211). Если удастся найти матрицы I_k, то можно написать для вектора $\boldsymbol{x}$ из R_n дифференциальные уравнения (204), ибо $T_{jp}(\alpha_s)$ определяются лишь групповой операцией. Эти уравнения могут иметь при заданном начальном условии (206) только одно решение; только это решение, очевидно, и может быть тем преобразованием

$$\boldsymbol{x} = F(\alpha_1, \alpha_2, \alpha_3)\boldsymbol{u},$$

которое дает представление группы вращения в окрестности тождественного преобразования.

В рассматриваемом случае $r=3$, и если в уравнении (204) перейти к составляющим вектора $\boldsymbol{x}$, то получатся $3n$ уравнений для n составляющих вектора

$$\boldsymbol{x}(x_1, x_2, \ldots, x_n).$$

В дальнейшем для нас будет важно лишь то, что уравнение (204) не может иметь более одного решения при заданном начальном условии (206). Как уже говорилось выше, это можно формулировать так: *представление группы вращения вполне определяется своими бесконечно малыми преобразованиями* I_1, I_2, I_3.

Таким образом все сводится к определению бесконечно малых преобразований представления, к чему мы и переходим. Вместо искомых матриц I_1, I_2, I_3 вводим новые матрицы:

$$A_1 = -I_2 + iI_1;\ A_2 = I_2 + iI_1;\ A_3 = iI_3. \tag{212}$$

Легко проверить, что для них, вместо (211), получаются следующие соотношения:

$$\left.\begin{aligned} A_3A_1 - A_1A_3 &= A_1, \\ A_3A_2 - A_2A_3 &= -A_2, \\ A_1A_2 - A_2A_1 &= 2A_3. \end{aligned}\right\} \tag{213}$$

В представлении матрицами $F(\alpha_1, \alpha_2, \alpha_3)$ должно заключаться, в частности, и представление абелевой группы вращения вокруг оси Z, элементам которой соответствуют матрицы $F(0, 0, \alpha_3)$. При помощи соответствующего выбора ортов все эти матрицы одновременно преобразуются к диагональной форме, ибо неприводимые представления абелевой группы суть представления первого порядка. Для векторов, которые при этом играют роль ортов, преобразование $F(0, 0, \alpha_3)$ будет иметь вид [**69**]:

$$F(0, 0, \alpha_3)\boldsymbol{v} = e^{l\alpha_3}\boldsymbol{v}$$

или, полагая $l = -im$ и обозначая $\boldsymbol{v}$ через $\boldsymbol{v}_m$:

$$F(0, 0, \alpha_3)\boldsymbol{v}_m = e^{-im\alpha_3}\boldsymbol{v}_m.$$

Поскольку мы поставили условие однозначности представления только в окрестности $\alpha_s = 0$ ($s = 1, 2, 3$), мы не должны считать m целым числом. Отсюда получаем, основываясь на определении I_3:

$$A_3 v_m = i I_3 v_m = i \left[\frac{\partial}{\partial \alpha_3} F(0, 0, \alpha_3) v_m \right] = i \left(\frac{\partial}{\partial \alpha_3} e^{-im\alpha_3} v_m \right) = m v_m.$$

Итак,

$$A_3 v_m = m v_m, \tag{214}$$

т. е. v_m есть собственный вектор оператора A_3, соответствующий собственному значению m. Если таких собственных векторов несколько, то через v_m обозначаем один из них.

Докажем теперь следующую лемму:

Лемма. *Если некоторый вектор v есть собственный вектор оператора A_3, соответствующий собственному значению a, то вектор $A_1 v$, если он отличен от нулевого, есть также собственный вектор A_3, соответствующий собственному значению $(a+1)$, и аналогично $A_2 v$ есть собственный вектор A_3, соответствующий собственному значению $(a-1)$.*

По условию леммы $A_3 v = a v$, мы получаем в силу (213):

$$A_3(A_1 v) = (A_1 A_3 + A_1) v = A_1 (A_3 v) + A_1 v = \\ = A_1 (a v) + A_1 v = (a+1) A_1 v$$

и совершенно аналогично $A_3(A_2 v) = (a-1) A_2 v$.

Число различных собственных значений A_3 не больше n. Среди этих значений имеется одно или несколько с наибольшей вещественной частью. Обозначим это собственное значение или одно из них, если их несколько, через j, и пусть v_j — соответствующий собственный вектор (один из них, если их несколько). В силу леммы вектор $A_1 v_j$ должен был бы относиться к собственному значению $(j+1)$, но, согласно определению j, такого собственного значения у A_3 нет и, следовательно, мы должны иметь:

$$A_1 v_j = 0. \tag{215}$$

В силу доказанной выше леммы векторы:

$$v_{j-1} = A_2 v_j; \quad v_{j-2} = A_2 v_{j-1}; \ldots, \tag{216}$$

если они отличны от нуля, относятся к собственным значениям $(j-1)$, $(j-2)$, ... оператора A_3. Последовательность векторов (216) должна, конечно, привести к нулевому вектору, поскольку число различных собственных значений у A_3 не больше n. Докажем теперь формулу

$$A_1 v_k = \rho_k v_{k+1} \qquad (k = j, j-1, j-2, \ldots), \tag{217}$$

где ρ_k — некоторые целые числа. В силу (215) она верна при $k = j$, причем $\rho_j = 0$, а за v_{j+1} можно взять, например, нулевой вектор. Положим теперь, что формула (217) верна при некотором из указанных

k, и докажем ее для значения $(k-1)$, на единицу меньшего. Имеем в силу (213), (216) и (217):

$$A_1 v_{k-1} = A_1(A_2 v_k) = (A_2 A_1 + 2A_3) v_k =$$
$$= A_2(A_1 v_k) + 2A_3 v_k = A_2(\rho_k v_{k+1}) + 2k v_k = (\rho_k + 2k) v_k.$$

Отметим, что при $k=j$ мы не пользуемся при этом формулой

$$A_2 v_{k+1} = v_k,$$

ибо $\rho_k = 0$ при $k=j$. Таким образом формула (217) доказана, и числа ρ_k определяются из соотношений:

$$\rho_{k-1} = \rho_k + 2k; \quad \rho_j = 0 \qquad (k=j,\ j-1,\ \ldots).$$

Проводя последовательные вычисления, получаем:

$$\rho_k = j(j+1) - k(k+1),$$

т. е.

$$A_1 v_k = [j(j+1) - k(k+1)] v_{k+1} \qquad (k=j,\ j-1,\ \ldots). \tag{218}$$

Пользуясь этим равенством, определим значок s первого из векторов (216), равного нулю, т. е. $v_s = 0$ и вектор v_{s+1} отличен от нулевого. Из (217) при этом следует $\rho_s = 0$, т. е.

$$j(j+1) - s(s+1) = 0.$$

Это квадратное относительно s уравнение имеет корни $s=j$ и $s=-(j+1)$. Значение $s=j$ не годится, ибо вектор v_j отличен от нулевого и не входит в последовательность (216). Таким образом в последовательности (216) векторы

$$v_j,\ v_{j-1},\ \ldots,\ v_{-j+1},\ v_{-j} \tag{219}$$

отличны от нулевого, и $A_2 v_{-j} = 0$. Число этих векторов равно $(2j+1)$, откуда видно, что j есть или целое неотрицательное число, или половина нечетного положительного числа. Если $2j+1=n$, то мы можем принять векторы (219) за основные орты в пространстве R_n, если же $2j+1<n$, то они образуют в R_n некоторое подпространство L_{2j+1}. Положим, что мы имеем последнее. Каждый вектор v_k из последовательности (219) удовлетворяет уравнению:

$$A_3 v_k = k v_k \qquad (k=j,\ j-1,\ \ldots,\ -j+1,\ -j).$$

Далее мы имеем $A_2 v_k = v_{k-1}$, причем $v_{j-1} = 0$, и формулу (218). Тем самым операторы A_1, A_2 и A_3 переводят подпространство L_{2j+1} в себя, и данные формулы вполне определяют указанные операторы в подпространстве L_{2j+1}. Больше того, из формул (216) и (218) непосредственно следует, что никакое подпространство L_k, лежащее в L_{2j+1} и для которого $0<k<2j+1$, не остается инвариантным при применении операторов A_1, A_2, A_3. Определив A_k, мы можем для подпространства L_{2j+1} построить уравнения (204), которым должен удовлетворять вектор

$$x = F_j(\alpha_1, \alpha_2, \alpha_3)\, u \tag{220}$$

искомого представления в подпространстве L_{2j+1}. Это представление не может оставлять инвариантным никакое подпространство L_k, входящее в L_{2j+1}, т. е. представление неприводимо в L_{2j+1}, ибо если бы это было так, то и всякое A_s, в силу их определения, должно было бы оставлять инвариантным L_k, что, как мы только что видели, не имеет места. Если $2j+1=n$, то приведенное рассуждение относится ко всему R_n. При $2j+1<n$ мы отделили от общего представления в R_n неприводимое в указанном смысле представление порядка $(2j+1)$, т. е. оно не оставляет инвариантным никакое подпространство L_k при $0<k<2j+1$. Из наших рассуждений непосредственно вытекает, что существует только одно с точностью до подобных представлений неприводимое представление данного порядка. Но в [**69**] мы уже построили унитарные неприводимые представления любого порядка.

Тем самым они исчерпывают всевозможные неприводимые представления, и указанные выше представления, основанные на построенных в L_{2j+1} операторах A_s, должны быть им подобны.

Векторы (219) можно умножать на произвольные численные множители, отличные от нуля. При этом в соотношениях (216) и (218) также появятся численные множители. Указанные множители можно подобрать так, чтобы иметь окончательно следующие соотношения:

$$\left.\begin{aligned} A_1 v_k &= \sqrt{j(j+1)-k(k+1)}\, v_{k+1}, \\ A_2 v_k &= \sqrt{j(j+1)-k(k-1)}\, v_{k-1}, \\ A_3 v_k &= k v_k, \end{aligned}\right\} \tag{221}$$

причем $v_{j+1}=0$ и $v_{-j-1}=0$.

При таком выборе множителей получатся те представления, которые мы построили в [**69**], исходя от величины

$$\eta_l = \frac{x_1^{j+l} x_2^{j-l}}{\sqrt{(j+l)!\,(j-l)!}}. \tag{222}$$

Указанное выше построение дает возможность из любого представления выделять его неприводимые части. Все сводится к отысканию собственных векторов v_j оператора A_3 с наибольшим собственным значением и построению (216).

86. Представления группы Лоренца. Рассмотрим группу линейных преобразований с определителем единица:

$$\begin{aligned} x_1' &= a x_1 + b x_2 \\ x_2' &= c x_1 + d x_2. \end{aligned} \quad (ad - bc = 1) \tag{223}$$

Матрица преобразования содержит четыре комплексных коэффициента, между которыми имеется одно соотношение. Произвольными остаются три комплексные величины, что сводится к шести вещественным параметрам. Введем эти параметры, принимая следующее обозначение для матрицы преобразования:

$$A=\left\|\begin{matrix} 1+\alpha_1+i\alpha_2 & \alpha_3+i\alpha_4 \\ \alpha_5+i\alpha_6 & d(\alpha_s) \end{matrix}\right\|, \tag{224}$$

где

$$d(\alpha_s)=\frac{1+(\alpha_3+i\alpha_4)(\alpha_5+i\alpha_6)}{1+\alpha_1+i\alpha_2}.$$

Мы получаем шесть бесконечно малых преобразований I'_k, которые легко построить. Например, для построения I'_1 надо в матрице A положить все α_s, кроме α_1, равными нулю, продифференцировать матрицу по α_1 и затем положить $\alpha_1=0$. Таким образом, будем иметь:

$$I'_1=\left\|\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}\right\|; \quad I'_2=\left\|\begin{matrix} i & 0 \\ 0 & -i \end{matrix}\right\|; \quad I'_3=\left\|\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}\right\|;$$

$$I'_4=\left\|\begin{matrix} 0 & i \\ 0 & 0 \end{matrix}\right\|; \quad I'_5=\left\|\begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}\right\|; \quad I'_6=\left\|\begin{matrix} 0 & 0 \\ i & 0 \end{matrix}\right\|.$$

Структурные постоянные $C^{(j)}_{pq}$, входящие в соотношения (208), по самому их определению должны быть вещественными и могут быть найдены из этих соотношений:

$$I'_p I_q - I'_q I'_p = \sum_{j=1}^{6} C^{(j)}_{pq} I'_j \qquad (p<q;\ p,\ q=1,\ 2,\ \ldots,\ 6).$$

При этом надо отметить, что между матрицами I'_j не существует линейного соотношения (кроме тривиального) с вещественными коэффициентами, поэтому можно получить следующие пятнадцать равенств:

$$\begin{array}{lll}
I'_1I'_3-I'_3I'_1=2I'_3, & I'_1I'_4-I'_4I'_1=2I'_4, & I'_2I'_3-I'_3I'_2=2I'_4, \\
I'_1I'_5-I'_5I'_1=-2I'_5, & I'_1I'_6-I'_6I'_1=-2I'_6, & I'_2I'_5-I'_5I'_2=-2I'_6, \\
I'_3I'_5-I'_5I'_3=I'_1, & I'_3I'_6-I'_6I'_3=I'_2, & I'_4I'_5-I'_5I'_4=I'_2. \\
I'_2I'_4-I'_4I'_2=-2I'_3, & I'_1I'_2-I'_2I'_1=0, & \\
I'_2I'_6-I'_6I'_2=2I'_5, & I'_3I'_4-I'_4I'_3=0, & \\
I'_4I'_6-I'_6I'_4=-I'_1, & I'_5I'_6-I'_6I'_5=0. &
\end{array}$$

Если I_k $(k=1, 2, \ldots, 6)$ — бесконечно малые преобразования для любого представления исследуемой группы, то между ними также имеют место пятнадцать соотношений

$$I_p I_q - I_q I_p = \sum_{j=1}^{6} C_{pq}^{(j)} I_j$$

с теми же коэффициентами $C_{pq}^{(j)}$. Если мы введем обозначения

$$\begin{aligned} &I_3 + iI_4 = 2A_1; \quad I_5 + iI_6 = 2A_2; \quad I_1 + iI_2 = 4A_3; \\ &I_3 - iI_4 = 2B_1; \quad I_5 - iI_6 = 2B_2; \quad I_1 - iI_2 = 4B_3, \end{aligned} \tag{225}$$

то упомянутые пятнадцать соотношений записываются в следующем виде:

$$A_p B_q - B_q A_p = 0 \quad (p, q = 1, 2, 3) \tag{226}$$

и, кроме того,

$$\begin{aligned} A_3A_1 - A_1A_3 &= A_1, \\ A_3A_2 - A_2A_3 &= -A_2, \\ A_1A_2 - A_2A_1 &= 2A_3, \end{aligned} \tag{227}$$

$$\begin{aligned} B_3B_1 - B_1B_3 &= B_1, \\ B_3B_2 - B_2B_3 &= -B_2, \\ B_1B_2 - B_2B_1 &= 2B_3. \end{aligned} \tag{228}$$

Отметим, что соотношения (226) и (227) выполняются тривиально, если взять матрицы I'_k, ибо при этом $A'_k = 0$ $(k=1, 2, 3)$. Соотношения (227) совпадают с соотношениями (213) и рассуждения предыдущего номера остаются по существу в силе. Мы применяем написанные соотношения для бесконечно малых преобразований любого линейного представления группы (223). Если $\boldsymbol{v}_j$ — собственный вектор оператора A_3, относящийся к наибольшему собственному значению, то имеется $(2j+1)$ собственных векторов $\boldsymbol{v}_k$ $(k=j, j-1, \ldots, -j+1, -j)$ оператора A_3, которые преобразуются операторами A_1, A_2, A_3 согласно формулам (221), причем $\boldsymbol{v}_{j+1}=0$ и $\boldsymbol{v}_{-j-1}=0$. Пусть $L^{(j)}$ — подпространство, образованное всеми собственными векторами оператора A_3, относящимися к собственному значению j. Покажем, что если вектор $\boldsymbol{v}$ принадлежит $L^{(j)}$, то и векторы $B_q\boldsymbol{v}$ $(q=1, 2, 3)$ принадлежат $L^{(j)}$. Действительно, в силу (226):

$$A_3(B_q\boldsymbol{v}) = B_q(A_3\boldsymbol{v}) = B_q(j\boldsymbol{v}) = jB_q\boldsymbol{v},$$

откуда и следует, что $B_q\boldsymbol{v}$ есть собственный вектор A_3, соответствующий собственному значению j (или нулевой вектор), т. е. $B_q\boldsymbol{v}$ входит в $L^{(j)}$. В $L^{(j)}$ мы можем повторить наши рассуждения из [85], заменяя операторы A_k операторами B_k. Следовательно, можно построить в $L^{(j)}$ ряд векторов $\boldsymbol{v}_{jk'}$ $(k'=j', j'-1, \ldots, -j+1, -j')$, которые преобразуются согласно формулам (221) при замене j на j' и A_k на B_k. Каждый вектор $\boldsymbol{v}_{jk'}$ при повторном применении операции A_2 дает $(2j+1)$ векторов $\boldsymbol{v}_{kk'}$ $(k=j, j-1, \ldots, -j+1, -j)$.

Таким образом окончательно получается $(2j+1)(2j'+1)$ векторов $\boldsymbol{v}_{kk'}$, для которых имеют место соотношения:

$$\left.\begin{aligned}
A_1\boldsymbol{v}_{kk'} &= \sqrt{j(j+1)-k(k+1)}\,\boldsymbol{v}_{k+1,\,k'},\\
A_2\boldsymbol{v}_{kk'} &= \sqrt{j(j+1)-k(k-1)}\,\boldsymbol{v}_{k-1,\,k'},\\
A_3\boldsymbol{v}_{kk'} &= k\boldsymbol{v}_{kk'},\\
B_1\boldsymbol{v}_{kk'} &= \sqrt{j(j+1)-k'(k'+1)}\,\boldsymbol{v}_{k,\,k'+1},\\
B_2\boldsymbol{v}_{kk'} &= \sqrt{j(j+1)-k'(k'-1)}\,\boldsymbol{v}_{k,\,k'-1},\\
B_3\boldsymbol{v}_{kk'} &= k'\boldsymbol{v}_{kk'}.
\end{aligned}\right\} \tag{229}$$

Эти формулы определяют в пространстве с числом измерений $(2j+1)(2j'+1)$ операторы A_p и B_q, и согласно формулам (225) определяются операторы I_k, после чего уравнение (204) может приводить лишь к одному линейному представлению группы. Это есть то представление, которое мы строили в [**80**].

В последних номерах мы следовали изложению, приведенному в книге Ван-дер-Вардена „Метод теории групп в квантовой механике“.

87. Вспомогательные формулы. Вернемся к формулам из [**82**]. Мы имеем:

$$G_\beta G_\alpha = G_\gamma, \tag{230}$$

причем γ_s выражается через α_s и β_s согласно формулам (189) или (192), которые определяют основную групповую операцию. Составим матрицу, зависящую от переменных α_s и β_s, т. е. элементов группы G_α и G_β. Обозначим эту матрицу символом $S(G_\beta, G_\alpha)$, и элементы ее определим следующими формулами:

$$S_{ik}(G_\beta, G_\alpha) = \frac{\partial\gamma_i}{\partial\beta_k} \qquad (i, k = 1, 2, \ldots, r). \tag{231}$$

Мы уже рассматривали эту матрицу в [**82**] при $\beta_s = 0$, т. е. при $G_\beta = E$, где E — единичный элемент группы. Изучим свойства этой матрицы. Из ее определения непосредственно следует:

$$S(G_\beta, E) = I. \tag{232}$$

Докажем еще формулу:

$$S(G_\beta, G_\alpha)\cdot S(E, G_\beta) = S(E, G_\beta G_\alpha). \tag{233}$$

Положим $G_\alpha = G_{\alpha''}G_{\alpha'}$, так что

$$G_\gamma = G_\beta G_\alpha = (G_\beta G_{\alpha''})\,G_{\alpha'} = G_\delta G_{\alpha'} \qquad (G_\delta = G_\beta G_{\alpha''}).$$

Применяем правило дифференцирования сложных функций

$$\frac{\partial\gamma_i}{\partial\beta_k} = \sum_{s=1}^{r}\frac{\partial\gamma_i}{\partial\delta_s}\cdot\frac{\partial\delta_s}{\partial\beta_k} = \sum_{s=1}^{r} S_{is}(G_\delta, G_{\alpha'})\,S_{sk}(G_\beta, G_{\alpha''}),$$

откуда

$$S(G_\beta, G_{\alpha''}G_{\alpha'}) = S(G_\delta, G_{\alpha'})\,S(G_\beta, G_{\alpha''}).$$

Положив в этом равенстве $G_\beta = E$, $G_{\alpha''} = G_\beta$ и $G_{\alpha'} = G_\alpha$, получим равенство (233). При $G_\alpha = G_\beta^{-1}$ получим выражение матрицы, обратной матрице $S(E, G_\beta)$:

$$S^{-1}(E, G_\beta) = S(G_\beta, G_\beta^{-1}). \tag{234}$$

Матрица $S(E, G_\beta)$ в обозначениях из [82] будет $S(\beta_s)$ и обратная матрица будет $T(\beta_s)$. Сейчас мы их будем обозначать символами $S(G_\beta)$ и $T(G_\beta)$:

$$S(E, G_\beta) = S(G_\beta); \qquad S^{-1}(E, G_\beta) = T(G_\beta). \tag{235}$$

Мы имеем:

$$S(G_\beta)\, T(G_\beta) = T(G_\beta)\, S(G_\beta) = E. \tag{236}$$

Формула (233) дает:

$$S(G_\beta, G_\alpha) = S(E, G_\gamma)\, S^{-1}(E, G_\beta) = S(G_\gamma)\, S^{-1}(G_\beta), \tag{237}$$

и соотношение (231) может быть записано в виде:

$$\frac{\partial \gamma_i}{\partial \beta_k} = \sum_{s=1}^{r} S_{is}(G_\gamma)\, T_{sk}(G_\beta). \tag{238}$$

Умножая обе части на $T_{mi}(G_\gamma)$ и суммируя по i, получим, в силу (236):

$$\sum_{i=1}^{r} T_{mi}(G_\gamma) \frac{\partial \gamma_i}{\partial \beta_k} = T_{mk}(G_\beta). \tag{239}$$

Дифференцируем (238) по β_l:

$$\frac{\partial^2 \gamma_i}{\partial \beta_k\, \partial \beta_l} = \sum_{s,\,p=1}^{r} \frac{\partial S_{is}(G_\gamma)}{\partial \gamma_p} \frac{\partial \gamma_p}{\partial \beta_l} T_{sk}(G_\beta) + \sum_{s=1}^{r} S_{is}(G_\gamma) \frac{\partial T_{sk}(G_\beta)}{\partial \beta_l},$$

откуда, выражая $\frac{\partial \gamma_p}{\partial \beta_l}$ согласно формуле (238):

$$\frac{\partial^2 \gamma_i}{\partial \beta_k\, \partial \beta_l} = \sum_{s,\,p,\,q=1}^{r} \frac{\partial S_{is}(G_\gamma)}{\partial \gamma_p} S_{pq}(G_\gamma)\, T_{ql}(G_\beta)\, T_{sk}(G_\beta) + \sum_{s=1}^{r} S_{is}(G_\gamma) \frac{\partial T_{sk}(G_\beta)}{\partial \beta_l}.$$

Переставляя в правой части k и l, пользуясь независимостью левой части от порядка дифференцирования и переставляя переменные суммирования s и q, получим:

$$\sum_{s,\,p,\,q=1}^{r} \left[\frac{\partial S_{is}(G_\gamma)}{\partial \gamma_p} S_{pq}(G_\gamma) - \frac{\partial S_{iq}(G_\gamma)}{\partial \gamma_p} S_{ps}(G_\gamma)\right] T_{ql}(G_\beta)\, T_{sk}(G_\beta) =$$

$$= -\sum_{s=1}^{r} S_{is}(G_\gamma) \left[\frac{\partial T_{sk}(G_\beta)}{\partial \beta_l} - \frac{\partial T_{sl}(G_\beta)}{\partial \beta_k}\right].$$

Умножим обе части на произведение $S_{lf}(G_\beta)\,S_{kg}(G_\beta)\,T_{hi}(G_\gamma)$ и просуммируем по i, k и l от 1 до r. Принимая во внимание (236), получим равносильную систему равенств:

$$\sum_{i,\,p=1}^{r}\left[\frac{\partial S_{ig}(G_\gamma)}{\partial\gamma_p}S_{pf}(G_\gamma)-\frac{\partial S_{if}(G_\gamma)}{\partial\gamma_p}S_{pg}(G_\gamma)\right]T_{hi}(G_\gamma)=$$
$$=-\sum_{k,\,l=1}^{r}S_{lf}(G_\beta)\,S_{kg}(G_\beta)\left[\frac{\partial T_{hk}(G_\beta)}{\partial\beta_l}-\frac{\partial T_{hl}(G_\beta)}{\partial\beta_k}\right].$$

От этих равенств легко перейти к предыдущим, умножая обе части на произведение $T_{fl_1}(G_\beta)\,T_{gk_1}(G_\beta)\,S_{i_1h}(G_\gamma)$ и суммируя по f, g и h. В последнем равенстве левая часть зависит только от γ_s, а правая — только от β_s. Таким образом, ввиду произвольности G_α в формуле (230) и тем самым независимости β_s и γ_s обе части последней формулы должны равняться одной и той же постоянной, и, в частности:

$$\sum_{k,\,l=1}^{r}S_{lf}(G_\beta)\,S_{kg}(G_\beta)\left[\frac{\partial T_{hk}(G_\beta)}{\partial\beta_l}-\frac{\partial T_{hl}(G_\beta)}{\partial\beta_k}\right]=C_{fg}^{(h)}.$$

Меняя значки, можем написать:

$$\sum_{s,\,t=1}^{r}S_{ti}(G_\alpha)\,S_{sk}(G_\alpha)\left[\frac{\partial T_{ps}(G_\alpha)}{\partial\alpha_t}-\frac{\partial T_{pt}(G_\alpha)}{\partial\alpha_s}\right]=-C_{ik}^{(p)}. \tag{240}$$

Если положим в этом тождестве $G_\alpha=E$, т. е. $\alpha_s=0$ $(s=1,\ldots,r)$, и примем во внимание, что $S(E)=E$, то получим:

$$-C_{ik}^{(p)}=\left[\frac{\partial T_{pk}(G_\alpha)}{\partial\alpha_i}-\frac{\partial T_{pi}(G_\alpha)}{\partial\alpha_k}\right].$$

Сравнивая с формулой (199) из [82], мы видим, что $C_{ik}^{(p)}$ суть структурные постоянные, определенные нами выше. Умножая обе части (240) на $T_{jl}(G_\alpha)\,T_{km}(G_\alpha)$ и суммируя по i и k, получим в силу (236):

$$\frac{\partial T_{pm}(G_\alpha)}{\partial\alpha_l}-\frac{\partial T_{pl}(G_\alpha)}{\partial\alpha_m}=-\sum_{i,\,k=1}^{r}C_{ik}^{(p)}T_{il}(G_\alpha)\,T_{km}(G_\alpha). \tag{241}$$

Вернемся к формулам (207) и (208). Формула (208) получается, как мы видели, путем приравнивания нулю квадратной скобки формулы (207) при $\alpha_s=0$ $(s=1,\ldots,r)$. Пользуясь (241), легко показать, что *из* (208) *вытекает, что квадратная скобка формулы* (207) *равна нулю и при любых* α_s.

Второе слагаемое этой скобки представим в виде:

$$\sum_{j,\,k=1}^{r}T_{jp}T_{kq}I_jI_k-\sum_{j,\,k=1}^{r}T_{jq}T_{kp}I_jI_k,$$

причем мы не выписываем аргумента G_α у T. Заменяя у вычитаемого j на k и k на j, получим:

$$\sum_{j,\ k=1}^{r} T_{jp}T_{kq}(I_jI_k - I_kI_j) = \sum_{j,\ k,\ s=1}^{r} T_{jp}T_{kq}C_{jk}^{(s)}I_s.$$

Преобразуя первое слагаемое скобки формулы (207):

$$\sum_{j=1}^{r}\left(\frac{\partial T_{jp}}{\partial \alpha_q} - \frac{\partial T_{jq}}{\partial \alpha_p}\right) I_j$$

согласно (241), получим непосредственно тот же результат, но с обратным знаком. Наряду с матрицей $S(G_\beta, G_\alpha)$ рассмотрим матрицу $S'(G_\beta, G_\alpha)$, элементы которой определяются формулами:

$$\frac{\partial \gamma_i}{\partial \alpha_k} = S'(G_\beta,\ G_\alpha). \tag{242}$$

Совершенно так же, как и выше, можно доказать формулы:

$$\left.\begin{aligned} S'(E,\ G_\alpha) &= I, \\ S'(G_\beta G_\alpha,\ E) &= S'(G_\beta,\ G_\alpha)\, S'(G_\alpha,\ E), \\ S'^{-1}(G_\alpha,\ E) &= S'(G_\alpha^{-1},\ G_\alpha), \end{aligned}\right\} \tag{243}$$

которые нам понадобятся в дальнейшем.

88. Построение группы по структурным постоянным. В настоящем номере мы в общих чертах коснемся вопроса о построении групповой операции и группы линейных преобразований по заданным структурным постоянным $C_{ik}^{(p)}$, которые удовлетворяют соотношениям (194_1) и (194_2). Это построение основано на одной теореме из теории уравнений с частными производными, о которой мы упоминали выше. Сформулируем сейчас эту теорему.

Пусть имеется следующая система дифференциальных уравнений с частными производными:

$$\frac{\partial z_i}{\partial x_k} = X_{ik}(x_1,\ \ldots,\ x_n;\ z_1,\ \ldots,\ z_m) \tag{244}$$

$$(i = 1,\ 2,\ \ldots,\ m;\ k = 1,\ 2,\ \ldots,\ n).$$

Напишем, пользуясь этой системой, условие того, что

$$\frac{\partial^2 z_i}{\partial x_k\, \partial x_l} = \frac{\partial^2 z_i}{\partial x_l\, \partial x_k}.$$

Оно имеет, очевидно, вид:

$$\frac{\partial X_{ik}}{\partial x_l} + \sum_{s=1}^{m} \frac{\partial X_{ik}}{\partial z_s} \cdot \frac{\partial z_s}{\partial x_l} = \frac{\partial X_{il}}{\partial x_k} + \sum_{s=1}^{m} \frac{\partial X_{il}}{\partial z_s} \cdot \frac{\partial z_s}{\partial x_k},$$

или, заменяя $\frac{\partial z_s}{\partial x_l}$ и $\frac{\partial z_s}{\partial x_k}$ правой частью системы (244), получим:

$$\frac{\partial X_{ik}}{\partial x_l}+\sum_{s=1}^{m}\frac{\partial X_{ik}}{\partial z_s}\cdot X_{sl}=\frac{\partial X_{il}}{\partial x_k}+\sum_{s=1}^{m}\frac{\partial X_{il}}{\partial z_s}X_{sk} \qquad (k\neq l). \tag{245}$$

Это равенство является соотношением между переменными x_k, z_i.

Теорема. *Если функции X_{ik} в окрестности значений $x_k=x_k^{(0)}$, $z_i=z_i^{(0)}$ (и при этих значениях) непрерывны и имеют непрерывные частные производные, которые входят в соотношения* (245), *и все эти последние соотношения выполняются тождественно относительно x_k, z_i, то система* (244) *при начальных условиях*

$$z_i\Big|_{x_k=x_k^{(0)}}=z_i^{(0)}$$

имеет решение и притом единственное.

Тождественное выполнение всех соотношений (245) при наличии указанных условий непрерывности называется обычно условием полной интегрируемости системы (244). Опишем теперь схему построений групповой операции и группы линейных преобразований по заданным структурным постоянным.

Итак, пусть заданы постоянные $C_{ik}^{(p)}$, где i, k, $p=1, 2, \ldots, r$, причем эти постоянные удовлетворяют соотношениям (194_1) и (194_2).

Если решить систему (241) относительно частных производных, то можно проверить, что упомянутые соотношения являются условиями полной интегрируемости системы (241). Таким образом, существует единственная матрица $T(G_\alpha)$ с элементами $T_{pq}(G_\alpha)$ $(p, q=1, 2, \ldots, r)$, которая обращается в единичную матрицу при $G_\alpha=E$, т. е. при $\alpha_s=0$ $(s=1, 2, \ldots, r)$, и удовлетворяет системе (241). Имея $T(G_\alpha)$, мы можем построить и обратную матрицу $S(G_\alpha)=T^{-1}(G_\alpha)$. Для построения групповой операции обращаемся к системе (238). Правые части этих уравнений — известные функции β_s и γ_s $(s=1, 2, \ldots, r)$. Можно проверить, что система (241) выражает условия полной интегрируемости системы (238). Следовательно, существует единственное решение системы (238), которое удовлетворяет начальным условиям

$$\gamma_i\Big|_{\beta_k=0}=\alpha_i.$$

Построенное решение и дает групповую операцию. Начальные условия выражают тот факт, что элемент G_γ, определяемый формулой (230), обращается в G_α при $\beta_s=0$ $(s=1, \ldots, r)$. Переходим теперь к построению группы линейных преобразований, т. е. группы матриц заданного порядка по структурным постоянным, причем, как указано выше, мы имеем уже матрицу $T(G_\alpha)$. Как мы показали в [**83**], условия полной интегрируемости (204) или системы (205) сводятся к тождественному равенству нулю квадратной скобки уравнения (207) при любом выборе значков, а эти последние условия выполнены, как мы показали в [**87**], если матрицы I_s удовлетворяют соотношениям (208). Таким образом, решение задачи должно начинаться с построения матриц I_s заданного порядка, удовлетворяющих соотношениям (208). Это сложная алгебраическая задача. Имея матрицы I_s, мы можем уже утверждать, что система (205) имеет единственное решение, удовлетворяющее начальным условиям (206). Это решение и дает группу матриц с заданными структурными постоянными $C_{ik}^{(p)}$.

Можно показать, что интегрирование системы (241) при начальных условиях $T(E)=I$ сводится к интегрированию системы обыкновенных линейных дифференциальных уравнений с постоянными коэффициентами. Сформулируем соответствующий результат. Построим систему линейных обыкновенных дифференциальных уравнений с постоянными коэффициентами:

$$\frac{dw_{ik}(t)}{dt}=\delta_{ik}+\sum_{p,\ q=1}^{r} C_{pq}^{(i)}\alpha_p w_{qk}(t),$$

где $\delta_{ik}=0$ при $i\neq k$, $\delta_{ii}=1$ и $\alpha_1, \alpha_2, \ldots, \alpha_r$ считаются заданными постоянными. При этом функции $T_{ik}(\alpha_s)=w_{ik}(t)$ удовлетворяют системе (241) и начальному условию $T(E)=I$. Подробное рассмотрение вопроса о построении непрерывной группы по заданным структурным постоянным так же, как и исследование других вопросов теории непрерывных групп, можно найти в книге Л. С. Понтрягина „Непрерывные группы".

89. Интегрирование на группе. В [76, 77] мы доказывали ряд соотношений, которые содержат суммы некоторых величин, зависящих от элементов группы, причем суммирование распространялось на все элементы группы. В случае непрерывной группы суммирование заменяется интегрированием по параметрам, определяющим элементы группы. Положим, что непрерывная группа G такова, что при некотором выборе параметров, этой группе в вещественном r-мерном пространстве T_r определяемой параметрами $\alpha_1, \alpha_2, \ldots, \alpha_r$, соответствует ограниченная замкнутая область V (область вместе с ее границей), так что всякому элементу из G соответствует определенная точка V и наоборот. Внутри области V функции $\varphi_j(\beta_1, \ldots, \beta_r; \alpha_1, \ldots, \alpha_r)$, определяющие групповую операцию, считаются непрерывными и достаточное число раз дифференцируемыми. Кроме того, эти функции и их производные считаются непрерывными вплоть до границы V. Зависимость параметров $\tilde{\alpha}_s$, соответствующих элементу G_α^{-1}, от параметров α_s также считается непрерывной. Группы с такими свойствами называются обычно *компактными*. Для определения интегрирования на группе рассмотрим определитель матрицы $S'(G_\beta, G_\alpha)$ [87] и введем для него следующее обозначение:

$$\Delta'(G_\beta, G_\alpha)=\left|\frac{\partial\gamma_i}{\partial\alpha_k}\right|_1^r. \tag{246}$$

Из (243) непосредственно следует:

$$\Delta'(E, G_\alpha)=1, \tag{247$_1$}$$

$$\Delta'(G_\beta G_\alpha, E)=\Delta'(G_\beta, G_\alpha)\cdot\Delta'(G_\alpha, E). \tag{247$_2$}$$

Обозначая $\delta'(G_\beta)=\Delta'(G_\beta, E)$, можем написать:

$$\Delta'(G_\beta, G_\alpha)=\frac{\delta'(G_\beta G_\alpha)}{\delta'(G_\alpha)}. \tag{248}$$

Отсюда, принимая во внимание, что $\delta'(E)=\Delta'(E, E)=1$, получаем:

$$\Delta'(G_\alpha^{-1}, G_\alpha)=\frac{1}{\delta'(G_\alpha)}. \tag{249}$$

Введем еще одно обозначение:

$$u'(G_\alpha)=\Delta'(G_\alpha^{-1}, G_\alpha). \tag{250}$$

В силу сделанных выше предположений и $u'(G_\alpha)$ есть непрерывная функция в замкнутой области V. Она не обращается в нуль, ибо

$$\frac{1}{u'(G_\alpha)} = \delta'(G_\alpha) = \Delta'(G_\alpha, E)$$

есть также непрерывная функция. Принимая во внимание, что $u'(E) = 1$, можем утверждать, что $u'(G_\alpha)$ и $\delta(G_\alpha)$ — положительные функции. То же можно утверждать, в силу (248), и относительно $\Delta'(G_\beta, G_\alpha)$.

Пусть $f(G_\alpha) = f(\alpha_1, \ldots, \alpha_r)$ — любая непрерывная в замкнутой области V функция.

Определим интеграл от этой функции на группе G формулой:

$$\int_G f(G_\alpha)\, dG_\alpha = \int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r, \tag{251}$$

где интеграл, стоящий справа, есть обычный интеграл по области V. Докажем, что такой интеграл обладает следующим свойством левой инвариантности:

$$\int_G f(G_\alpha)\, dG_\alpha = \int_G f(G_\beta G_\alpha)\, dG_\alpha, \tag{252}$$

или в координатах

$$\int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha \ldots d\alpha_r = \int_V f(\gamma_1, \ldots, \gamma_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r, \tag{253}$$

где G_β — любой фиксированный элемент группы G. Для доказательства заменим в интеграле, стоящем в левой части, переменный элемент G_α переменным элементом G_δ, полагая $G_\alpha = G_\beta G_\delta$, причем область изменения параметров $\delta_1, \ldots, \delta_r$ по-прежнему V. Определитель преобразования будет

$$\left|\frac{\partial \alpha_i}{\partial \alpha_k}\right|_1^r = \Delta'(G_\beta, G_\delta) = \frac{\delta'(G_\beta G_\delta)}{\delta'(G_\delta)} = \frac{u'(G_\delta)}{u'(G_\beta G_\delta)} = \frac{u'(G_\delta)}{u'(G_\alpha)},$$

и мы получим:

$$\int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha) \frac{u'(G_\delta)}{u'(G_\alpha)}\, d\delta_1 \ldots d\delta_r =$$
$$= \int_G f(G_\beta G_\delta)\, dG_\delta.$$

Это и совпадает с (253). Замена в правой части G_α на G_δ несущественна.

Аналогично строится правоинвариантный интеграл. Введем определитель матрицы $S(G_\beta, G_\alpha)$:

$$\Delta(G_\beta, G_\alpha) = \left|\frac{\partial \gamma_i}{\partial \beta_k}\right|_1^r. \tag{254}$$

Как и выше, имеем:

$$\left.\begin{aligned} \Delta(G_\alpha, E) &= 1, \\ \Delta(E, G_\beta G_\alpha) &= \Delta(G_\beta, G_\alpha)\, \Delta(E, G_\beta), \\ \Delta(G_\beta, G_\alpha) &= \frac{\delta(G_\beta, G_\alpha)}{\delta(G_\beta)}, \end{aligned}\right\} \tag{255}$$

где $\delta(G_\alpha) = \Delta(E, G_\alpha)$. Вводится положительная функция

$$u(G_\alpha) = \frac{1}{\delta(G_\alpha)} = \Delta(G_\alpha, G_\alpha^{-1}), \tag{256}$$

и интеграл определяется формулой:

$$\int\limits_V f(\alpha_1, \ldots, \alpha_r)\, u(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \int\limits_G f(G_\alpha)\, \widetilde{dG}_\alpha. \tag{257}$$

Волна сверху знака дифференциала отличает этот интеграл от интеграла (251). При этом имеет место свойство правой инвариантности:

$$\int\limits_G f(G_\alpha)\, \widetilde{dG}_\alpha = \int\limits_G f(G_\alpha G_\beta)\, \widetilde{dG}_\alpha. \tag{258}$$

Докажем теперь, что замена G_α на G_α^{-1} под знаком подынтегральной функции влечет преобразование левоинвариантного интеграла в правоинвариантный и обратно. Дифференцируем равенство $G_\lambda = G_\alpha G_\beta$, записанное в параметрах, по α_s, причем во всех дальнейших формулах мы считаем $G_\beta = G_\alpha^{-1}$:

$$\frac{\partial \lambda_i}{\partial \alpha_k} + \sum_{s=1}^{r} \frac{\partial \lambda_i}{\partial \beta_s} \cdot \frac{\partial \beta_s}{\partial \alpha_k} = 0, \text{ откуда } \left|\frac{\partial \lambda_i}{\partial \alpha_k}\right|_1^r = (-1)^r \left|\frac{\partial \lambda_i}{\partial \beta_k}\right|_1^r \left|\frac{\partial \beta_i}{\partial \alpha_k}\right|_1^r,$$

и, принимая во внимание (246) и (245), получаем:

$$\left|\frac{\partial \beta_i}{\partial \alpha_k}\right|_1^r = (-1)^r \frac{\Delta(G_\alpha, G_\alpha^{-1})}{\Delta'(G_\alpha, G_\alpha^{-1})} = (-1)^r \frac{u(G_\alpha)}{u'(G_\alpha^{-1})}. \tag{259}$$

Можно дать другое представление этого определителя. Из равенства

$$\left|\frac{\partial \beta_i}{\partial \alpha_k}\right|_1^r \left|\frac{\partial \alpha_i}{\partial \beta_k}\right|_1^r = 1$$

следует

$$\left|\frac{\partial \alpha_i}{\partial \beta_k}\right|_1^r = (-1)^r \frac{u'(G_\alpha^{-1})}{u(G_\alpha)} \tag{260}$$

или, меняя местами G_α и G_α^{-1}:

$$\left|\frac{\partial \beta_i}{\partial \alpha_k}\right| = (-1)^r \frac{u'(G_\alpha)}{u(G_\alpha^{-1})}. \tag{261}$$

Обращаемся теперь к интегралам. Совершая обычным образом замену переменных интегрирования, получим, пользуясь (260):

$$\int\limits_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \int\limits_V f(\tilde{\beta}_1, \ldots, \tilde{\beta}_r)\, u'(G_\alpha) \left\| \frac{\partial \alpha_i}{\partial \beta_k} \right|_1^r \right| d\beta_1 \ldots d\beta_r =$$
$$= \int\limits_V f(\tilde{\beta}_1, \ldots, \tilde{\beta}_r)\, u'(G_\alpha) \frac{u(G_\beta)}{u'(G_\alpha)}\, d\beta_1 \ldots d\beta_r.$$

Сокращая на $u'(G_\alpha)$ и заменяя переменный элемент G_β переменным элементом G_α, получим:

$$\int\limits_V f(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_r)\, u(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \int\limits_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r. \tag{262}$$

Совершенно так же, принимая во внимание формулу (261), получим:

$$\int\limits_V f(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \int\limits_V f(\alpha_1, \ldots, \alpha_r)\, u(G_\alpha)\, d\alpha_1 \ldots d\alpha_r. \tag{263}$$

До сих пор мы не использовали компактности группы. Область V может быть и бесконечной. Но при этом надо предполагать функцию $f(\alpha_1, \ldots, \alpha_r)$ такой, что все написанные интегралы имеют смысл. Сейчас, пользуясь компактностью, мы докажем, что $u(G_\alpha) = u'(G_\alpha)$. Для этого рассмотрим определитель

$$D(G_\beta, G_\alpha) = \left| \frac{\partial \mu_i}{\partial \beta_k} \right|_1^r, \tag{264}$$

где $G_\mu = G_\alpha^{-1} G_\beta G_\alpha$, и докажем формулу

$$D(G_\beta, G_{\alpha''} G_{\alpha'}) = D(G_{\alpha'}^{-1} G_\beta G_{\alpha''}, G_{\alpha'}) D(G_\beta, G_{\alpha''}). \tag{265}$$

Мы можем написать:

$$G_\mu = (G_{\alpha''} G_{\alpha'})^{-1} G_\beta (G_{\alpha''} G_{\alpha'}) = G_{\alpha'}^{-1} G_\nu G_{\alpha'},$$

где $G_\nu = G_{\alpha''}^{-1} G_\beta G_{\alpha''}$, а потому

$$\left| \frac{\partial \mu_i}{\partial \beta_k} \right|_1^r = \left| \frac{\partial \mu_i}{\partial \nu_k} \right|_1^r \cdot \left| \frac{\partial \nu_i}{\partial \beta_k} \right|_1^r = D(G_\nu, G_{\alpha'}) D(G_\beta, G_{\alpha''}),$$

откуда и следует (265). Положив в этой формуле $G_\beta = E$, получим

$$D(E, G_{\alpha''} G_{\alpha'}) = D(E, G_{\alpha'}) D(E, G_{\alpha''}). \tag{266}$$

Если мы введем численную функцию элемента:

$$\eta(G_\alpha) = D(E, G_\alpha), \tag{267}$$

то, в силу (266), можем написать:

$$\eta(G_{\alpha''} G_{\alpha'}) = \eta(G_{\alpha''}) \eta(G_{\alpha'}), \tag{268}$$

т. е. перемножение элементов означает умножение соответствующих значений функции $\eta(G_\alpha)$. Мы имеем, очевидно:

$$\eta(E) = 1 \text{ и } \eta(G_\alpha) \eta(G_\alpha^{-1}) = 1, \tag{269}$$

и функция $\eta(G_\alpha)$ непрерывна и положительна в замкнутой области V.

Используя компактность группы, докажем сейчас, что $\eta(G_\alpha) = 1$ для любого элемента G_α. Положим, что для некоторого элемента G_α мы имеем $\eta(G_\alpha) \neq 1$. Если, например, $\eta(G_\alpha) < 1$, то в силу (269): $\eta(G_\alpha^{-1}) > 0$, и мы можем считать всегда, что $\eta(G_\alpha) > 1$. При этом

$$\eta(G_\alpha^n) = [\eta(G_\alpha)]^n \to \infty \text{ при } n \to \infty.$$

Это противоречит тому, что непрерывная в замкнутой области V функция $\eta(G_\alpha)$ должна быть ограниченной. Переходим теперь к установлению связи между $u(G_\alpha)$ и $u'(G_\alpha)$. Пусть

$$G_\gamma = G_\beta G_\alpha = G_\alpha^{-1} (G_\alpha G_\beta) G_\alpha = G_\alpha^{-1} G_\rho G_\alpha \quad (G_\rho = G_\alpha G_\beta).$$

Мы имеем:

$$\left| \frac{\partial \gamma_i}{\partial \beta_k} \right|_1^r = \Delta(G_\beta, G_\alpha).$$

Но, с другой стороны:

$$\left| \frac{\partial \gamma_i}{\partial \beta_k} \right|_1^r = \left| \frac{\partial \gamma_i}{\partial \rho_k} \right|_1^r \left| \frac{\partial \rho_i}{\partial \beta_k} \right|_1^r = D(G_\rho, G_\alpha) \Delta'(G_\alpha, G_\beta),$$

т. е.

$$\Delta(G_\beta, G_\alpha) = D(G_\alpha G_\beta, G_\alpha) \cdot \Delta'(G_\alpha, G_\beta).$$

Полагая $G_\beta = G_\alpha^{-1}$, получим:

$$\Delta(G_\alpha^{-1}, G_\alpha) = D(E, G_\alpha) \cdot \Delta'(G_\alpha, G_\alpha^{-1}),$$

т. е.

$$u(G_\alpha^{-1}) = \eta(G_\alpha)\, u'(G_\alpha^{-1}) \quad \text{или} \quad u(G_\alpha^{-1}) = u'(G_\alpha^{-1})$$

при любом G_α, ибо $\eta(G_\alpha) = 1$. Таким образом для компактных групп левоинвариантный интеграл (251) совпадает с правоинвариантным интегралом (257). Кроме того, из (262) или (263) следует, что этот интеграл совпадает также с интегралом

$$\int\limits_V f(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r.$$

Для некомпактных групп левоинвариантный интеграл может быть отличным от правоинвариантного. В качестве примера рассмотрим группу линейных преобразований вида:

$$z' = e^{\alpha_1} z + \alpha_2,$$

где α_1 и α_2 меняются от $(-\infty)$ до $(+\infty)$. В данном случае $r = 2$ и V есть вся плоскость. Композиция двух преобразований дает:

$$z' = e^{\alpha_1} z + \alpha_2; \quad z'' = e^{\beta_1} z' + \beta_2; \qquad \gamma_1 = \varphi_1(\beta_1, \beta_2; \alpha_1, \alpha_2) = \beta_1 + \alpha_1;$$

т. е.

$$z'' = e^{\beta_1 + \alpha_1} z + (e^{\beta_1}\alpha_2 + \beta_2), \qquad \gamma_2 = \varphi_2(\beta_1, \beta_2; \alpha_1, \alpha_2) = e^{\alpha_1}\beta_2 + \alpha_2.$$

Единичному элементу соответствуют параметры $\alpha_1 = \alpha_2 = 0$. Элемент G_α^{-1} имеет параметры $\tilde{\alpha}_1 = -\alpha_1$, $\tilde{\alpha}_2 = -\alpha_2 e^{-\alpha_1}$. Вычисляем функциональные определители:

$$\Delta'(G_\beta, G_\alpha) = \begin{vmatrix} 1, & 0 \\ 0, & e^{\alpha_1} \end{vmatrix} = e^{\alpha_1}; \quad \delta'(G_\alpha) = e^{\alpha_1}; \quad u'(G_\alpha) = e^{-\alpha_1};$$

$$\Delta(G_\beta, G_\alpha) = \begin{vmatrix} 1, & e^{\alpha_1}\beta_2 \\ 0, & 1 \end{vmatrix} = 1; \quad \delta(G_\alpha) = u(G_\alpha) = 1.$$

Левоинвариантный интеграл имеет вид:

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(\alpha_1, \alpha_2)\, e^{-\alpha_1}\, d\alpha_1\, d\alpha_2$$

и правоинвариантный интеграл:

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(\alpha_1, \alpha_2)\, d\alpha_1\, d\alpha_2.$$

Отметим, что при доказательстве равенства правоинвариантного и левоинвариантного интегралов, т. е. равенства $u(G_\alpha) = u'(G_\alpha)$, можно требование компактности группы заменить другим требованием.

Пусть G' — подгруппа, состоящая из элементов G, которые имеют вид:

$$G_\alpha G_\beta G_\alpha^{-1} G_\beta^{-1} \tag{270}$$

или получаются из таких элементов путем их перемножения, причем G_α и G_β — любые элементы G.

Нетрудно видеть, что если некоторый элемент G_γ содержится среди элементов (270), то и обратный элемент G_γ^{-1} содержится среди элементов (270).

Точно так же и элемент $G_\delta^{-1}G_\gamma G_\delta$, при любом выборе G_δ из G, содержится среди элементов (270). Иначе говорят, что подгруппа G' порождается элементами (270). Из сказанного выше следует, что G' является нормальным делителем G. Подгруппа G' приводится к единичному элементу в том и только в том случае, если все элементы (270) суть единичные элементы, т. е. в том и только в том случае, когда G есть абелева группа. Подгруппа G' может и совпадать с G. В частности, это будет иметь место, если G есть не абелева, простая группа. Подгруппа G' называется обычно *коммутантом* группы G.

Из определения (268) и (269) следует, что $\eta(G_\alpha G_\beta G_\alpha^{-1}G_\beta^{-1})=1$, что $\eta(G_\gamma)=1$ для всех G_γ из G' и что $\eta(G_\alpha)$ имеет одинаковое значение для всех элементов, принадлежащих одной и той же совокупности по группе G', т. е. что функция $\eta(G_\alpha)$ имеет определенное значение для всякого элемента дополнительной к G' группы. Если G' совпадает с G, то $\eta(G_\alpha)=1$ при любом G_α из G. То же имеет место, если упомянутая выше дополнительная группа компактна. Но раз

$$\eta(G_\alpha)=1, \text{ то } u(G_\alpha)=u'(G_\alpha).$$

90. Свойство ортогональности. Примеры. Свойство правой и левой инвариантности интеграла аналогично в случае конечных групп тому свойству, что при переменном элементе G_s и фиксированном элементе G_t произведение G_sG_t или G_tG_s пробегает по одному разу все элементы группы. Этим свойством мы пользовались при доказательстве того, что всякое представление группы эквивалентно унитарному представлению и при доказательстве свойств ортогональности. Пользуясь инвариантным интегралом, можно доказать аналогичные предложения и для компактных групп. Если $A(G_\alpha)$ — унитарные матрицы, дающие неприводимое линейное представление компактной группы G, и $B(G_\sigma)$ — унитарные матрицы, дающие неэквивалентное неприводимое представление, то, обозначая, как всегда, двумя значками снизу элементы матриц, мы будем иметь следующие формулы, выражающие ортогональность неэквивалентных неприводимых унитарных представлений:

$$\int\limits_V \{A(G_\alpha)\}_{ij}\overline{\{B(G_\alpha)\}}_{kl}\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=0. \tag{271}$$

Для одного неприводимого представления получим

$$\int\limits_V \{A(G_\alpha)\}_{ij}\overline{\{A(G_\alpha)\}}_{kl}\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=\frac{\delta_{ik}\delta_{jl}}{p}\int\limits_V u(G_\alpha)\,d\alpha_1\ldots d\alpha_r, \tag{272}$$

где p — порядок матриц. Точно так же для характеров:

$$X(G_\alpha)=\sum_{i=1}^{p}\{A(G_\sigma)\}_{ii};\quad X'(G_\alpha)=\sum_{i=1}^{q}\{B(G_\alpha)\}_{ii},$$

где p и q — порядок матриц $A(G_\alpha)$ и $B(G_\alpha)$, имеют место следующие свойства:

$$\int\limits_V X(G_\alpha)\overline{X'(G_\alpha)}\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=0, \tag{273}$$

$$\int\limits_V X(G_\alpha)\overline{X(G_\sigma)}\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=\int\limits_V u(G_\alpha)\,d\alpha_1\ldots d\alpha_r. \tag{274}$$

1. Перейдем к рассмотрению примеров. Пусть G — абелева группа вращения плоскости вокруг начала. Для нее $r=1$ и единственный параметр α дает величину угла поворота. Будем считать, что α принадлежит промежуток $(0, 2\pi)$, причем концы этого промежутка отождествляются. Последовательные вращения на углы α и β приводят к вращению на угол $\beta+\alpha$, причем, вычитая, если надо, 2π, мы должны привести эти суммы к указанному промежутку. В рассматриваемом случае функциональные определители $\Delta(G_\beta, G_\alpha)$ и $\Delta'(G_\beta, G_\alpha)$ сводятся к производной от $\beta+\alpha$ по β или α, равной единице, так что $u(G_\alpha)=u'(G_\alpha)=1$. Мы знаем, что группа G имеет неприводимые унитарные представления первого порядка $e^{im\alpha}$ $(m=0, \pm 1, \pm 2, \ldots)$, и формулы (357) и (358) дают известные формулы:

$$\int_0^{2\pi} e^{im_1\alpha}\overline{e^{im_2\alpha}}\,d\alpha=\int_0^{2\pi} e^{i(m_1-m_2)\alpha}\,d\alpha=\begin{cases}0 & \text{при } m_1\neq m_2,\\ 2\pi & \text{при } m_1=m_2.\end{cases} \tag{275}$$

Отметим, что в силу необходимости приведения суммы $\beta+\alpha$ к промежутку $(0, 2\pi)$, мы имеем некоторую особенность в непрерывности и определении производных этой суммы в тех случаях, когда для α и β, лежащих внутри промежутка $(0, 2\pi)$, сумма оказывается равной 2π.

2. Рассмотрим группу вращения трехмерного пространства и применим параметры, несколько отличные от тех, о которых мы говорили в [84]. Пусть пространство повернулось на угол ω вокруг оси, которая образует углы α, β и γ с осями координат X, Y и Z.

Введем четыре параметра:

$$a_0=\cos\frac{1}{2}\omega;\ a_1=\cos\alpha\sin\frac{1}{2}\omega;\ a_2=\cos\beta\sin\frac{1}{2}\omega;\ a_3=\cos\gamma\sin\frac{1}{2}\omega. \tag{276}$$

Они связаны соотношением

$$a_0^2+a_1^2+a_2^2+a_3^2=1. \tag{277}$$

Единичному преобразованию соответствуют значения $a_0=1$, $a_1=a_2=a_3=0$. Мы можем принять a_1, a_2, a_3 за параметры. При этом a_0 считается их функцией.

Если произвести последовательно вращения, определяемые параметрами (a_0, a_1, a_2, a_3) и (b_0, b_1, b_2, b_3), то параметры результирующего вращения (c_0, c_1, c_2, c_3) определятся, как нетрудно проверить, формулами:

$$\begin{aligned} c_0&=a_0b_0-a_1b_1-a_2b_2-a_3b_3, & c_2&=a_0b_2-a_1b_3+a_2b_0+a_3b_1,\\ c_1&=a_0b_1+a_1b_0+a_2b_3-a_3b_2, & c_3&=a_0b_3+a_1b_2-a_2b_1+a_3b_0. \end{aligned} \tag{278}$$

Считая a_0 функцией a_1, a_2, a_3, получим, в силу (277):

$$a_0\frac{\partial a_0}{\partial a_j}+a_j=0,\quad (j=1, 2, 3),$$

откуда $\frac{\partial a_0}{\partial a_j}=0$ для E. Пользуясь этим, мы можем легко составить функциональный определитель при $b_0=1$, $b_1=b_2=b_3=0$:

$$\frac{D(c_1, c_2, c_3)}{D(b_1, b_2, b_3)}=\begin{vmatrix} a_0 & -a_3 & a_2\\ a_3 & a_0 & -a_1\\ -a_2 & a_1 & a_0\end{vmatrix}=a_0(a_0^2+a_1^2+a_2^2+a_3^2)=$$

$$=a_0=\sqrt{1-a_1^2-a_2^2-a_3^2}.$$

Инвариантный интеграл имеет вид:

$$\int\limits_V f(a_1, a_2, a_3) \frac{1}{\sqrt{1-a_1^2-a_2^2-a_3^2}} da_1\, da_2\, da_3. \tag{279}$$

Область V есть сфера с центром в начале координат и радиусом единица. Отметим, что формулы (278) непосредственно получаются из правила умножения кватернионов

$$c_0+c_1 i+c_2 j+c_3 k=(a_0+a_1 i+a_2 j+a_3 k)(b_0+b_1 i+b_2 j+b_3 k),$$

причем единицы i, j и k подчиняются следующему закону умножения:

$$i^2=j^2=k^2=-1;\quad ij=-ji=k;\quad jk=-kj=i;\quad ki=-ik=j.$$

Нетрудно установить связь между параметрами (a_0, a_1, a_2, a_3) и углами Эйлера α, β, γ. Приведем соответствующие формулы:

$$a_0=\cos\frac{1}{2}\beta\cos\frac{1}{2}(\alpha+\gamma);\qquad a_2=\sin\frac{1}{2}\beta\sin\frac{1}{2}(\gamma-\alpha);$$
$$a_1=\sin\frac{1}{2}\beta\cos\frac{1}{2}(\gamma-\alpha);\qquad a_3=\cos\frac{1}{2}\beta\sin\frac{1}{2}(\alpha+\gamma).$$

Инвариантный интеграл в параметрах (α, β, γ) записывается при этом в виде:

$$\int\limits_V f(\alpha, \beta, \gamma)\sin\beta\sin^2\frac{1}{2}(\alpha-\gamma)\, d\alpha\, d\beta\, d\gamma, \tag{280}$$

причем $0\leqslant\alpha<2\pi$, $0\leqslant\beta<\pi$, $0\leqslant\gamma<2\pi$. Отметим, что в интеграле (279) функция

$$\frac{1}{a_0}=\frac{1}{\sqrt{1-a_1^2-a_2^2-a_3^2}}$$

обращается в бесконечность, если $\omega=\pi$. Это связано с тем, что в формулах (276) для a_1, a_2, a_3 вместо ω стоит $\sin\frac{1}{2}\omega$. Отметим в связи с этим, что те свойства, о которых говорилось в [89] в связи с определением компактности, должны быть выполнены лишь при некотором выборе параметров. При перемене параметров эти свойства могут уже потеряться. Кроме того, для группы вращения трехмерного пространства будет иметь место особенность в непрерывности и определении производных, о которых мы уже упоминали в конце первого примера в связи с группой вращения плоскости вокруг начала.

Отметим еще, что совпадение инвариантных интегралов для группы вращения пространства может быть непосредственно связано с тем, что это — простая не абелева группа.

3. Легко непосредственным вычислением проверить совпадение левоинвариантного и правоинвариантного интеграла для группы Лоренца, которая, как мы видели, гомоморфна группе линейных преобразований с определителем единица:

$$\left.\begin{aligned} x_1' &= a_0x_1+a_1x_2,\\ x_2' &= a_2x_1+a_3x_2 \end{aligned}\right\}\quad (a_0a_3-a_1a_2=1). \tag{281}$$

Единичному элементу соответствуют значения $a_0=a_3=1$, $a_1=a_2=0$. Можно считать a_0 функцией a_1, a_2, a_3 и принять за параметры вещественные и мнимые части величин a_1, a_2 и a_3-1. Групповая операция сводится к умножению матриц второго порядка, и мы имеем:

$$c_0=b_0a_0+b_1a_2,\quad c_1=b_0a_1+b_1a_3,\quad c_2=b_2a_0+b_3a_2,\quad c_3=b_2a_1+b_3a_3. \tag{282}$$

Если положить $a_k = \alpha'_k + i\alpha''_k$ ($k = 0, 1, 2, 3$), то параметрами группы будут $\alpha'_1, \alpha'_2, \alpha'_3, \alpha''_1, \alpha''_2, \alpha''_3$. Полагая далее $b_k = \beta'_k + i\beta''_k$ и $c_k = \gamma'_k + i\gamma''_k$, мы для определения инвариантного интеграла должны вычислить функциональные определители:

$$\frac{D(\gamma'_1, \gamma'_2, \gamma'_3, \gamma''_1, \gamma''_2, \gamma''_3)}{D(\beta'_1, \beta'_2, \beta'_3, \beta''_1, \beta''_2, \beta''_3)} \text{ при } \beta'_1 = \beta'_2 = \beta''_1 = \beta''_2 = \beta''_3 = 0; \quad \beta'_s = 1$$

или

$$\frac{D(\gamma'_1, \gamma'_2, \gamma'_3, \gamma''_1, \gamma''_2, \gamma''_3)}{D(\alpha'_1, \alpha'_2, \alpha'_3, \alpha''_1, \alpha''_2, \alpha''_3)} \text{ при } \alpha'_1 = \alpha'_2 = \alpha''_1 = \alpha''_2 = \alpha''_3 = 0; \ \alpha'_3 = 1,$$

причем несуществен тот факт, что $\alpha'_3 = 1$, а не нулю для тождественного преобразования группы. В обоих случаях мы получаем один и тот же инвариантный интеграл:

$$\int\limits_V f(\alpha'_1, \alpha'_2, \alpha'_3, \alpha''_1, \alpha''_2, \alpha''_3) \frac{1}{\alpha'^2_3 + \alpha''^2_3} d\alpha'_1\, d\alpha'_2\, d\alpha'_3, d\alpha''_1\, d\alpha''_2\, d\alpha''_3. \tag{283}$$

Область V есть полное шестимерное пространство. Совпадение инвариантных интегралов связано с тем, что для группы (281) подгруппа G', образованная производящими элементами $G_\alpha G_\beta G_\alpha^{-1} G_\beta^{-1}$, о которой мы говорили в [**89**], совпадает с самой группой. Нетрудно показать, действительно, что G' не сводится к тождественному преобразованию или к нормальному делителю, образованному элементами E и $(-E)$. Фактическое вычисление плотности в инвариантном интеграле (283) можно просто провести на основании леммы, в которой используется понятие аналитической функции нескольких комплексных переменных (см. главу IV второй части этого тома).

Лемма. *Пусть $w_s = u_s + iv_s$ ($x = 1, 2, \ldots, k$) суть аналитические функции комплексных переменных $z_s = x_s + iy_s$ ($s = 1, 2, \ldots, k$). При этом функциональный определитель функций $(u_1, v_1, \ldots, u_k, v_k)$ по переменным $(x_1, y_1, \ldots, x_k, y_k)$ равен квадрату модуля функционального определителя функций по $(w_1, \ldots, w_k)$ по переменным $(z_1, \ldots, z_k)$.*

Мы имеем (см. главы I и IV второй части этого тома)

$$\frac{\partial u_i}{\partial x_k} = \frac{\partial v_i}{\partial y_k}; \qquad \frac{\partial v_i}{\partial x_k} = -\frac{\partial u_i}{\partial y_k},$$

и можем написать:

$$\frac{D(u_1, v_1, \ldots, u_k, v_k)}{D(x_1, y_1, \ldots, x_k, y_k)} = \begin{vmatrix} a_{11} & b_{11} & a_{12} & b_{12} & \ldots & a_{1k} & b_{1k} \\ -b_{11} & a_{11} & -b_{12} & a_{12} & \ldots & -b_{1k} & a_{1k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & b_{k1} & a_{k2} & b_{k2} & \ldots & a_{kk} & b_{kk} \\ -b_{k1} & a_{k1} & -b_{k2} & a_{k2} & \ldots & -b_{kk} & a_{kk} \end{vmatrix},$$

где

$$a_{ik} = \frac{\partial u_i}{\partial x_k}; \quad b_{ik} = \frac{\partial v_i}{\partial x_k}.$$

Добавляя к каждому нечетному столбцу четный, умноженный на i, получим определитель:

$$\begin{vmatrix} c_{11} & b_{11} & c_{12} & b_{12} & \ldots & c_{1k} & b_{1k} \\ ic_{11} & a_{11} & ic_{12} & a_{12} & \ldots & ic_{1k} & a_{1k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & b_{k1} & c_{k2} & b_{k2} & \ldots & c_{kk} & b_{kk} \\ ic_{k1} & a_{k1} & ic_{k2} & a_{k2} & \ldots & ic_{kk} & a_{kk} \end{vmatrix} \quad (c_{ik} = a_{ik} + ib_{ik}).$$

Далее, отнимая от каждой четной стороны строки нечетную, умноженную на i, получим:

$$\begin{vmatrix} c_{11} & b_{11} & c_{12} & b_{12} & \dots & c_{1k} & b_{1k} \\ 0 & \overline{c_{11}} & 0 & \overline{c_{12}} & \dots & 0 & \overline{c_{1k}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & b_{k1} & c_{k2} & b_{k2} & \dots & c_{kk} & b_{kk} \\ 0 & \overline{c_{k1}} & 0 & \overline{c_{k2}} & \dots & 0 & \overline{c_{kk}} \end{vmatrix}.$$

Выносяя налево все нечетные столбцы и наверх все нечетные строки, будем иметь:

$$\begin{vmatrix} c_{11} & c_{12} & \dots & c_{1k} & b_{11} & b_{12} & \dots & b_{1k} \\ c_{21} & c_{22} & \dots & c_{2k} & b_{21} & b_{22} & \dots & b_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & c_{k2} & \dots & c_{kk} & b_{k1} & b_{k2} & \dots & b_{kk} \\ 0 & 0 & \dots & 0 & \overline{c_{11}} & \overline{c_{12}} & \dots & \overline{c_{1k}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & \overline{c_{k1}} & \overline{c_{k2}} & \dots & \overline{c_{kk}} \end{vmatrix},$$

откуда и следует

$$\frac{D(u_1, v_1, \dots, u_{k1}, v_{k1})}{D(x_1, y_1, \dots, x_k, y_k)} = \begin{vmatrix} c_{11} & \dots & c_{1k} \\ \cdot & \cdot & \cdot \\ c_{k1} & \dots & c_{kk} \end{vmatrix} \begin{vmatrix} \overline{c_{11}} & \dots & \overline{c_{1k}} \\ \cdot & \cdot & \cdot \\ \overline{c_{k1}} & \dots & \overline{c_{kk}} \end{vmatrix} = \left| \frac{D(w_1, \dots, w_k)}{D(z_1, \dots, z_k)} \right|^2.$$

Переходим к вычислению функции $u(G_\alpha)$ в инвариантном интеграле. Для этого, согласно лемме, надо вычислить функциональный определитель:

$$\frac{D(c_1, c_2, c_3)}{D(b_1, b_2, b_3)} \text{ при } b_0 = b_3 = 1; \quad b_1 = b_2 = 0 \tag{284}$$

или

$$\frac{D(c_1, c_2, c_3)}{D(a_1, a_2, a_3)} \text{ при } a_0 = a_3 = 1; \quad a_1 = a_2 = 0. \tag{285}$$

Из соотношения $a_0 a_3 - a_1 a_2 = 0$ следует:

$$-a_2 + a_3 \frac{\partial a_0}{\partial a_1} = 0; \quad -a_1 + a_3 \frac{\partial a_0}{\partial a_2} = 0; \quad a_0 + a_3 \frac{\partial a_0}{\partial a_3} = 0.$$

Далее имеем:

$$\begin{aligned} &\frac{\partial c_1}{\partial a_1} = b_0; && \frac{\partial c_1}{\partial a_2} = 0; && \frac{\partial c_1}{\partial a_3} = b_1; \\ &\frac{\partial c_2}{\partial a_1} = b_2 \frac{\partial a_0}{\partial a_1}; && \frac{\partial c_2}{\partial a_2} = b_2 \frac{\partial a_0}{\partial a_2} + b_3; && \frac{\partial c_2}{\partial a_3} = b_2 \frac{\partial a_0}{\partial a_3}; \\ &\frac{\partial c_3}{\partial a_1} = b_2; && \frac{\partial c_3}{\partial a_2} = 0; && \frac{\partial c_3}{\partial a_3} = b_3, \end{aligned}$$

откуда и следует (283). Тот же результат мы получим и на основании формулы (284)

АЛФАВИТНЫЙ УКАЗАТЕЛЬ

Владимир Иванович Смирнов
Курс высшей математики, том III, часть I
М., 1974 г., 324 стр. с илл.
Редактор *И. Е. Морозова*
Техн. редактор *А. П. Колесникова*
Корректоры *О. А. Бутусова* и *Т. Д. Доверман*

Печать с матриц. Подписано к печати 12/VI 1974 г. Бумага 60×90[1]/16, тип. № 3. Физ. печ. л. 20,25. Условн. печ. л. 20,25. Уч.-изд. л. 19,34. Тираж 50 000 экз. Цена книги 84 коп. Заказ № 223.

Издательство «Наука»
Главная редакция
физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Отпечатано в ордена Трудового Красного Знамени Ленинградской типографии № 2 им. Евгении Соколовой Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств, полиграфии и книжной торговли, 198052, Ленинград, Л-52, Измайловский пр., 29, с матриц ордена Трудового Красного Знамени Ленинградской типографии № 1 «Печатный Двор» имени А. М. Горького Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств полиграфии и книжной торговли Ленинград, Гатчинская ул., 26.