Таким образом, неравенство (2) будет выполняться, если $x$ удовлетворяет, помимо условия $|x|>1$, также неравенству

$$kA\frac{|x|^n}{|x|-1}\leqslant|a_0x^n|=|a_0||x|^n,$$

т. е. если

$$|x|\geqslant\frac{kA}{|a_0|}+1. \tag{3}$$

Так как правая часть неравенства (3) больше 1, то можно утверждать, что для значений $x$, удовлетворяющих этому неравенству, имеет место неравенство (2), что доказывает лемму.

Лемма о возрастании модуля многочлена. *Для всякого многочлена $f(x)$ с комплексными коэффициентами, степень которого не меньше единицы, и всякого положительного действительного числа $M$, сколь угодно большого, можно подобрать такое положительное действительное число $N$, что при $|x|>N$ будет $|f(x)|>M$.*

Пусть

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n.$$

По формуле (11) § 18

$$|f(x)|=|a_0x^n+(a_1x^{n-1}+\ldots+a_n)|\geqslant$$
$$\geqslant|a_0x^n|-|a_1x^{n-1}+\ldots+a_n|. \tag{4}$$

Применим лемму о модуле старшего члена, положив $k=2$: существует такое число $N_1$, что при $|x|>N_1$ будет

$$|a_0x^n|>2|a_1x^{n-1}+\ldots+a_n|.$$

Отсюда

$$|a_1x^{n-1}+\ldots+a_n|<\frac{1}{2}|a_0x^n|,$$

т. е., по (4),

$$|f(x)|>|a_0x^n|-\frac{1}{2}|a_0x^n|=\frac{1}{2}|a_0x^n|.$$

Правая часть этого неравенства будет больше $M$ при

$$|x|>N_2=\sqrt[n]{\frac{2M}{|a_0|}}.$$

Таким образом, при $|x|>N=\max(N_1,\ N_2)$ будет $|f(x)|>M$.

Смысл этой леммы может быть выяснен при помощи следующей геометрической иллюстрации, которая в настоящем параграфе будет неоднократно использоваться. Предположим, что в каждой точке $x_0$ комплексной плоскости к этой плоскости восставлен перпендикуляр, длина которого (при заданной единице масштаба) равна модулю значения многочлена $f(x)$ в этой точке, т. е. равна $|f(x_0)|$. Концы перпендикуляров будут составлять ввиду доказанной выше

непрерывности модуля многочлена некоторую непрерывную кривую поверхность, расположенную над комплексной плоскостью. Лемма о возрастании модуля многочлена показывает, что эта поверхность при возрастании $|x_0|$ все больше и больше удаляется от комплексной плоскости, хотя, понятно, это удаление вовсе не является монотонным. Рис. 8 схематически изображает линию пересечения этой поверхности с плоскостью, перпендикулярной к комплексной плоскости и проходящей через точку $O$.

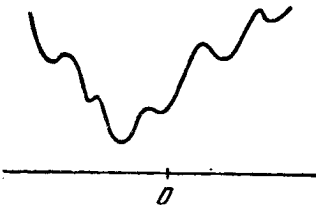

Рис. 8.

Основную роль в доказательстве играет следующая лемма:

Лемма Даламбера. *Если при $x=x_0$ многочлен $f(x)$ степени $n$, $n\geqslant 1$, не обращается в нуль, $f(x_0)\neq 0$ и поэтому $|f(x_0)|>0$, то можно найти такое приращение $h$, в общем случае комплексное, что*

$$|f(x_0+h)|<|f(x_0)|.$$

По формуле Тэйлора, если приращение $h$ пока произвольно, будет

$$f(x_0+h)=f(x_0)+hf'(x_0)+\frac{h^2}{2!}f''(x_0)+\ldots+\frac{h^n}{n!}f^{(n)}(x_0).$$

По условию, $x_0$ не является корнем для $f(x)$. Случайно, однако, это число может оказаться корнем для $f'(x)$, а также, быть может, для некоторых из дальнейших производных. Пусть $k$-я производная ($k\geqslant 1$) будет первой, не имеющей $x_0$ своим корнем, т. е.

$$f'(x_0)=f''(x_0)=\ldots=f^{(k-1)}(x_0)=0,\quad f^{(k)}(x_0)\neq 0.$$

Такое $k$ существует, так как, если $a_0$ есть старший коэффициент многочлена $f(x)$, то

$$f^{(n)}(x_0)=n!a_0\neq 0.$$

Таким образом,

$$f(x_0+h)=f(x_0)+\frac{h^k}{k!}f^{(k)}(x_0)+$$
$$+\frac{h^{k+1}}{(k+1)!}f^{(k+1)}(x_0)+\ldots+\frac{h^n}{n!}f^{(n)}(x_0).$$

Некоторые из чисел $f^{(k+1)}(x_0),\ \ldots,\ f^{(n-1)}(x_0)$ также могут равняться нулю, но это для нас несущественно.

Деля обе части этого равенства на $f(x_0)$, отличное, по условию, от нуля, и вводя обозначение

$$c_j=\frac{f^{(j)}(x_0)}{j!\,f(x_0)},\quad j=k,\ k+1,\ \ldots,\ n,$$

мы получим:

$$\frac{f(x_0+h)}{f(x_0)}=1+c_kh^k+c_{k+1}h^{k+1}+\ldots+c_nh^n$$

или, ввиду $c_k\neq 0$,

$$\frac{f(x_0+h)}{f(x_0)}=(1+c_kh^k)+c_kh^k\left(\frac{c_{k+1}}{c_k}h+\ldots+\frac{c_n}{c_k}h^{n-k}\right).$$

Переходя к модулям, получим:

$$\left|\frac{f(x_0+h)}{f(x_0)}\right|\leqslant|1+c_kh^k|+|c_kh^k|\left|\frac{c_{k+1}}{c_k}h+\ldots+\frac{c_n}{c_k}h^{n-k}\right|. \qquad (5)$$

До этого момента мы не делали никаких предположений о приращении $h$. Теперь мы будем выбирать $h$, причем будем отдельно выбирать его модуль и его аргумент. Модуль $h$ будет выбираться следующим образом. Так как

$$\frac{c_{k+1}}{c_k}h+\ldots+\frac{c_n}{c_k}h^{n-k}$$

является многочленом от $h$ без свободного члена, то, по лемме 1 $\left(\text{полагая } \varepsilon=\frac{1}{2}\right)$, можно найти такое $\delta_1$, что при $|h|<\delta_1$ будет

$$\left|\frac{c_{k+1}}{c_k}h+\ldots+\frac{c_n}{c_k}h^{n-k}\right|<\frac{1}{2}. \qquad (6)$$

С другой стороны, при

$$|h|<\delta_2=\sqrt[k]{|c_k|^{-1}}$$

будет

$$|c_kh^k|<1. \qquad (7)$$

Положим, что модуль $h$ выбран в соответствии с неравенством

$$|h|<\min(\delta_1,\ \delta_2). \qquad (8)$$

Тогда, ввиду (6), неравенство (5) превращается в строгое неравенство

$$\left|\frac{f(x_0+h)}{f(x_0)}\right|<|1+c_kh^k|+\frac{1}{2}|c_kh^k|; \qquad (9)$$

условием (7) мы воспользуемся лишь позже.

Для выбора аргумента $h$ потребуем, чтобы число $c_kh^k$ было отрицательным действительным числом. Иными словами,

$$\arg(c_kh^k)=\arg c_k+k\arg h=\pi,$$

откуда

$$\arg h=\frac{\pi-\arg c_k}{k}. \qquad (10)$$

При этом выборе $h$ число $c_k h^k$ будет отличаться знаком от своей абсолютной величины,

$$c_k h^k = -|c_k h^k|,$$

а поэтому, используя неравенство (7),

$$|1 + c_k h^k| = |1 - |c_k h^k|| = 1 - |c_k h^k|.$$

Таким образом, при выборе $h$ на основании условий (8) и (10) неравенство (9) принимает вид

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| < 1 - |c_k h^k| + \frac{1}{2}|c_k h^k| = 1 - \frac{1}{2}|c_k h^k|,$$

т. е. тем более

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| = \frac{|f(x_0+h)|}{|f(x_0)|} < 1,$$

откуда следует

$$|f(x_0+h)| < |f(x_0)|,$$

что доказывает лемму Даламбера.

При помощи той геометрической иллюстрации, которая была дана выше, можно следующим образом пояснить лемму Даламбера. Дано, что $|f(x_0)| > 0$. Это значит, что длина перпендикуляра, восставленного к комплексной плоскости в точке $x_0$, отлична от нуля. Тогда, по лемме Даламбера, можно найти такую точку $x_1 = x_0 + h$, что $|f(x_1)| < |f(x_0)|$, т. е. перпендикуляр в точке $x_1$ будет более коротким, чем в точке $x_0$, и, следовательно, поверхность, образованная концами перпендикуляров, будет в этой новой точке несколько ближе к комплексной плоскости. Как показывает доказательство леммы, модуль $h$ можно считать сколь угодно малым, т. е. точку $x_1$ можно выбрать как угодно близко к точке $x_0$; мы не будем, однако, пользоваться в дальнейшем этим замечанием.

Корнями многочлена $f(x)$ будут служить, очевидно, те комплексные числа (т. е. те точки комплексной плоскости), в которых поверхность, образованная концами перпендикуляров, коснется этой плоскости. Опираясь лишь на лемму Даламбера, нельзя доказать существование таких точек. В самом деле, пользуясь этой леммой, можно найти такую бесконечную последовательность точек $x_0, x_1, x_2, \ldots$, что

$$|f(x_0)| > |f(x_1)| > |f(x_2)| > \ldots, \tag{11}$$

Отсюда не следует, однако, существование такой точки $\bar{x}$, что $f(\bar{x}) = 0$, тем более, что убывающая последовательность положительных действительных чисел (11) вовсе не обязана стремиться к нулю.

Дальнейшие рассмотрения будут основаны на одной теореме из теории функций комплексного переменного, обобщающей теорему Вейерштрасса, известную читателю из курса математического анализа. Она относится к действительным функциям комплексного перемен-

ного, т. е. к функциям комплексного переменного, принимающим лишь действительные значения; примером таких функций служит модуль многочлена. В формулировке этой теоремы мы будем говорить для простоты *о замкнутом круге* $E$, понимая под этим круг на комплексной плоскости, к которому присоединены все точки его границы.

*Если действительная функция $g(x)$ комплексного переменного $x$ непрерывна во всех точках замкнутого круга $E$, то существует в круге $E$ такая точка $x_0$, что для всех $x$ из $E$ имеет место неравенство $g(x) \geqslant g(x_0)$. Точка $x_0$ является, следовательно, точкой минимума для $g(x)$ в круге $E$.*

Доказательство этой теоремы можно найти во всех курсах теории функций комплексного переменного, и мы его не приводим

Ограничиваясь случаем, когда функция $g(x)$ неотрицательна во всех точках круга $E$,— только этот случай представляет для нас интерес,— поясним геометрически эту теорему при помощи той иллюстрации, которая уже использована выше. В каждой точке $x_0$ круга $E$ проводим перпендикуляр длины $g(x_0)$. Концы этих перпендикуляров составляют кусок непрерывной кривой поверхности, причем благодаря замкнутости круга $E$ существование точек минимума для этого куска поверхности делается геометрически достаточно ясным. Эта иллюстрация не заменяет, конечно, доказательства теоремы.

Теперь мы можем перейти к непосредственному доказательству основной теоремы. Пусть дан многочлен $f(x)$ степени $n$, $n \geqslant 1$. Если его свободный член есть $a_n$, то, очевидно, $f(0) = a_n$. Применим к нашему многочлену лемму о возрастании модуля многочлена, полагая $M = |f(0)| = |a_n|$. Существует, следовательно, такое $N$, что при $|x| > N$ будет $|f(x)| > |f(0)|$. Очевидно, далее, что указанное выше обобщение теоремы Вейерштрасса применимо к функции $|f(x)|$ при любом выборе замкнутого круга $E$. В качестве $E$ мы возьмем замкнутый круг, ограниченный окружностью радиуса $N$ с центром в точке 0. Пусть точка $x_0$ будет точкой минимума для $|f(x)|$ в круге $E$, откуда, в частности, следует $|f(x_0)| \leqslant |f(0)|$.

Легко видеть, что $x_0$ *на самом деле будет служить точкой минимума для $|f(x)|$ на всей комплексной плоскости*: если точка $x'$ лежит вне $E$, то $|x'| > N$, и поэтому

$$|f(x')| > |f(0)| \geqslant |f(x_0)|.$$

Отсюда следует, наконец, что $f(x_0) = 0$, т. е. что $x_0$ *служит корнем для $f(x)$*; если бы было $f(x_0) \neq 0$, то, по лемме Даламбера, существовала бы такая точка $x_1$, что $|f(x_1)| < |f(x_0)|$; это противоречит, однако, только что установленному свойству точки $x_0$.

Заметим, что еще одно доказательство основной теоремы будет приведено в § 55.

## § 24. Следствия из основной теоремы

Пусть дан многочлен $n$-й степени, $n \geqslant 1$,

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \tag{1}$$

с любыми комплексными коэффициентами. Мы снова рассматриваем его как формально-алгебраическое выражение, вполне определяемое набором своих коэффициентов. Основная теорема о существовании корня, доказанная в предшествующем параграфе, позволяет утверждать существование для $f(x)$ корня $\alpha_1$, комплексного или действительного. Поэтому многочлен $f(x)$ обладает разложением

$$f(x) = (x - \alpha_1)\,\varphi(x).$$

Коэффициенты многочлена $\varphi(x)$ снова являются действительными или комплексными числами, и поэтому $\varphi(x)$ обладает корнем $\alpha_2$, откуда

$$f(x) = (x - \alpha_1)(x - \alpha_2)\,\psi(x).$$

Продолжая так далее, мы придем после конечного числа шагов к разложению многочлена $n$-й степени $f(x)$ в произведение $n$ линейных множителей,

$$f(x) = a_0(x - \alpha_1)(x - \alpha_2)\ \ldots\ (x - \alpha_n). \tag{2}$$

Коэффициент $a_0$ появился по следующей причине: если бы справа в выражении (2) стоял некоторый коэффициент $b$, то после раскрытия скобок старший член многочлена $f(x)$ имел бы вид $bx^n$, хотя на самом деле, ввиду (1), им является член $a_0x^n$. Поэтому $b = a_0$.

*Разложение* (2) *является для многочлена* $f(x)$ *единственным с точностью до порядка сомножителей разложением такого типа.*

Пусть, в самом деле, имеется еще разложение

$$f(x) = a_0(x - \beta_1)(x - \beta_2)\ \ldots\ (x - \beta_n). \tag{3}$$

Из (2) и (3) следует равенство

$$(x - \alpha_1)(x - \alpha_2)\ \ldots\ (x - \alpha_n) = (x - \beta_1)(x - \beta_2)\ \ldots\ (x - \beta_n). \tag{4}$$

Если бы корень $\alpha_i$ был отличен от всех $\beta_j$, $j = 1, 2, \ldots, n$, то, подставляя $\alpha_i$ вместо неизвестного в (4), мы получили бы слева нуль, а справа число, отличное от нуля. Таким образом, *всякий корень* $\alpha_i$ *равен некоторому корню* $\beta_j$ *и обратно.*

Отсюда еще не вытекает совпадение разложений (2) и (3). Действительно, среди корней $\alpha_i$, $i = 1, 2, \ldots, n$, могут быть равные между собой. Пусть, например, $s$ этих корней равны $\alpha_1$ и пусть, с другой стороны, среди корней $\beta_j$, $j = 1, 2, \ldots, n$, содержится $t$ равных корню $\alpha_1$. Нужно показать, что $s = t$.

Так как степень произведения многочленов равна сумме степеней сомножителей, то произведение двух многочленов, отличных от нуля, не может равняться нулю. Отсюда вытекает, что если два произведения многочленов равны друг другу, *то обе части равенства можно сократить на общий множитель*: если

$$f(x)\,\varphi(x)=g(x)\,\varphi(x)$$

и $\varphi(x)\neq 0$, то из

$$[f(x)-g(x)]\,\varphi(x)=0$$

следует

$$f(x)-g(x)=0,$$

т. е.

$$f(x)=g(x).$$

Применим это к равенству (4). Если, например, $s>t$, то, сокращая обе части равенства (4) на множитель $(x-\alpha_1)^t$, мы придем к равенству, левая часть которого еще содержит множитель $x-\alpha_1$, а правая его не содержит. Выше показано, однако, что это приводит к противоречию. Таким образом, единственность разложения (2) для многочлена $f(x)$ доказана.

Объединяя вместе одинаковые множители, разложение (2) можно переписать в виде

$$f(x)=a_0(x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\ \ldots\ (x-\alpha_l)^{k_l}, \tag{5}$$

где

$$k_1+k_2+\ldots+k_l=n.$$

При этом предполагается, что среди корней $\alpha_1, \alpha_2, \ldots, \alpha_l$ уже нет равных.

Докажем, что *число $k_i$ из* (5), $i=1, 2, \ldots, l$, *является кратностью корня $\alpha_i$ в многочлене $f(x)$*. Действительно, если эта кратность равна $s_i$, то $k_i\leqslant s_i$. Пусть, однако, $k_i<s_i$. В силу определения кратности корня для $f(x)$ существует разложение

$$f(x)=(x-\alpha_i)^{s_i}\,\varphi(x).$$

Заменяя в этом разложении множитель $\varphi(x)$ его разложением на линейные множители, мы получили бы для $f(x)$ разложение на линейные множители, заведомо отличное от разложения (2), т. е. пришли бы к противоречию с доказанной выше единственностью этого разложения.

Мы доказали, таким образом, следующий важный результат:

*Всякий многочлен $f(x)$ степени $n$, $n\geqslant 1$, с любыми числовыми коэффициентами имеет $n$ корней, если каждый из корней считать столько раз, какова его кратность.*

Заметим, что наша теорема справедлива и при $n=0$, так как многочлен нулевой степени не имеет, понятно, корней. Эта теорема неприменима лишь к многочлену 0, не имеющему степени и равному нулю при любом значении $x$. Этим последним замечанием мы воспользуемся при доказательстве следующей теоремы:

*Если многочлены $f(x)$ и $g(x)$, степени которых не превосходят $n$, имеют равные значения более чем при $n$ различных значениях неизвестного, то $f(x)=g(x)$.*

Действительно, многочлен $f(x)-g(x)$ имеет при наших предположениях более чем $n$ корней, а так как его степень не превосходит $n$, то должно иметь место равенство $f(x)-g(x)=0$.

Таким образом, учитывая, что различных чисел бесконечно много, можно утверждать, что *для любых двух различных многочленов $f(x)$ и $g(x)$ найдутся такие значения с неизвестного $x$, что $f(c)\neq g(c)$.* Такие $c$ можно найти не только среди комплексных чисел, но и среди действительных, среди рациональных и даже среди целых чисел.

Таким образом, два многочлена с числовыми коэффициентами, имеющие хотя бы при одной степени неизвестного $x$ различные коэффициенты, будут различными комплексными функциями комплексного переменного $x$. Этим доказана, наконец, *равносильность для многочленов с числовыми коэффициентами двух указанных в § 20 определений равенства многочленов — алгебраического и теоретико-функционального.*

Теорема, доказанная выше, позволяет утверждать, что *многочлен, степень которого не больше $n$, вполне определяется своими значениями при любых различных значениях неизвестного, число которых больше $n$.* Можно ли эти значения многочлена задавать произвольно? Если предположить, что задаются значения многочлена при $n+1$ различных значениях неизвестного, то ответ будет положительным: *всегда существует многочлен не более чем $n$-й степени, принимающий наперед заданные значения при $n+1$ заданных различных значениях неизвестного.*

В самом деле, пусть нужно построить многочлен не более чем $n$-й степени, который при значениях неизвестного $a_1, a_2, \ldots, a_{n+1}$, предполагаемых различными, принимает соответственно значения $c_1, c_2, \ldots, c_{n+1}$. Этим многочленом будет:

$$f(x)=\sum_{i=1}^{n+1}\frac{c_i(x-a_1)\ldots(x-a_{i-1})(x-a_{i+1})\ldots(x-a_{n+1})}{(a_i-a_1)\ldots(a_i-a_{i-1})(a_i-a_{i+1})\ldots(a_i-a_{n+1})}. \qquad (6)$$

Действительно, его степень не больше $n$, а значение $f(a_i)$ равно $c_i$.

Формула (6) называется *интерполяционной формулой Лагранжа.* Название «интерполяционная» связано с тем, что по этой формуле, зная значения многочлена в $n+1$ точке, можно вычислять его значения во всех других точках.

**Формулы Вьета.** Пусть дан многочлен $f(x)$ степени $n$ со старшим коэффициентом 1,

$$f(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n, \qquad (7)$$

и пусть $\alpha_1, \alpha_2, \ldots, \alpha_n$ — его корни[1]). Тогда $f(x)$ обладает следующим разложением:

$$f(x) = (x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n).$$

Перемножая скобки, стоящие справа, а затем приводя подобные члены и сравнивая полученные коэффициенты с коэффициентами из (7), мы получим следующие равенства, называемые *формулами Вьета* и выражающие коэффициенты многочлена через его корни:

$$\begin{aligned}
a_1 &= -(\alpha_1 + \alpha_2 + \ldots + \alpha_n),\\
a_2 &= \alpha_1\alpha_2 + \alpha_1\alpha_3 + \ldots + \alpha_1\alpha_n + \alpha_2\alpha_3 + \ldots + \alpha_{n-1}\alpha_n,\\
a_3 &= -(\alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2\alpha_4 + \ldots + \alpha_{n-2}\alpha_{n-1}\alpha_n),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{n-1} &= (-1)^{n-1}(\alpha_1\alpha_2\ldots\alpha_{n-1} + \alpha_1\alpha_2\ldots\alpha_{n-2}\alpha_n + \ldots + \alpha_2\alpha_3\ldots\alpha_n),\\
a_n &= (-1)^n\alpha_1\alpha_2\ldots\alpha_n.
\end{aligned}$$

Таким образом, в правой части $k$-го равенства, $k = 1, 2, \ldots, n$, стоит сумма всевозможных произведений по $k$ корней, взятая со знаком плюс или минус, в зависимости от четности или нечетности $k$.

При $n = 2$ эти формулы превращаются в известную из элементарной алгебры связь между корнями и коэффициентами квадратного многочлена. При $n = 3$, т. е. для кубичного многочлена, эти формулы принимают вид

$$a_1 = -(\alpha_1 + \alpha_2 + \alpha_3), \quad a_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3, \quad a_3 = -\alpha_1\alpha_2\alpha_3.$$

Формулы Вьета облегчают написание многочлена по заданным его корням. Так, найдем многочлен $f(x)$ четвертой степени, имеющий простыми корнями числа 5 и $-2$ и двукратным корнем число 3. Мы получим:

$$\begin{aligned}
a_1 &= -(5 - 2 + 3 + 3) = -9,\\
a_2 &= 5\cdot(-2) + 5\cdot 3 + 5\cdot 3 + (-2)\cdot 3 + (-2)\cdot 3 + 3\cdot 3 = 17,\\
a_3 &= -[5\cdot(-2)\cdot 3 + 5\cdot(-2)\cdot 3 + 5\cdot 3\cdot 3 + (-2)\cdot 3\cdot 3] = 33,\\
a_4 &= 5\cdot(-2)\cdot 3\cdot 3 = -90,
\end{aligned}$$

а поэтому

$$f(x) = x^4 - 9x^3 + 17x^2 + 33x - 90.$$

Если старший коэффициент $a_0$ многочлена $f(x)$ отличен от 1, то для применения формул Вьета необходимо сначала разделить все коэффициенты на $a_0$, что не влияет на корни многочлена. Таким образом, в этом случае формулы Вьета дают выражение для отношений всех коэффициентов к старшему.

**Многочлены с действительными коэффициентами.** Сейчас будут выведены некоторые следствия из основной теоремы алгебры комплексных чисел, относящиеся к многочленам с действительными

---

[1]) Каждый кратный корень взят здесь соответствующее число раз.

коэффициентами. По существу, именно на этих следствиях основано то исключительно большое значение основной теоремы, о котором говорилось раньше.

Пусть многочлен с действительными коэффициентами

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

имеет комплексный корень $\alpha$, т. е.

$$a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_{n-1}\alpha + a_n = 0.$$

Мы знаем, что последнее равенство не нарушится, если в нем все числа заменить на сопряженные. Однако все коэффициенты $a_0, a_1, \ldots$ $\ldots, a_{n-1}, a_n$, а также число 0, стоящее справа, будучи действительными, останутся при этой замене без изменения, и мы приходим к равенству

$$a_0\overline{\alpha}^n + a_1\overline{\alpha}^{n-1} + \ldots + a_{n-1}\overline{\alpha} + a_n = 0,$$

т. е.

$$f(\overline{\alpha}) = 0.$$

Таким образом, *если комплексное (но не действительное) число $\alpha$ служит корнем многочлена $f(x)$ с действительными коэффициентами, то корнем для $f(x)$ будет и сопряженное число $\overline{\alpha}$.*

Многочлен $f(x)$ будет делиться, следовательно, на квадратный трехчлен

$$\varphi(x) = (x-\alpha)(x-\overline{\alpha}) = x^2 - (\alpha+\overline{\alpha})x + \alpha\overline{\alpha}, \tag{8}$$

коэффициенты которого, как мы знаем из § 18, действительны. Пользуясь этим, докажем, что *корни $\alpha$ и $\overline{\alpha}$ имеют в многочлене $f(x)$ одну и ту же кратность.*

Пусть, в самом деле, эти корни имеют соответственно кратности $k$ и $l$ и пусть, например, $k > l$. Тогда $f(x)$ делится на $l$-ю степень многочлена $\varphi(x)$,

$$f(x) = \varphi^l(x)\,q(x).$$

Многочлен $q(x)$, как частное двух многочленов с действительными коэффициентами, также имеет действительные коэффициенты, но, в противоречие с доказанным выше, он имеет число $\alpha$ своим $(k-l)$-кратным корнем, тогда как число $\overline{\alpha}$ не является для него корнем. Отсюда следует, что $k = l$.

Таким образом, теперь можно сказать, что *комплексные корни всякого многочлена с действительными коэффициентами попарно сопряжены.* Отсюда и из доказанной выше единственности разложений вида (2) вытекает следующий окончательный результат:

*Всякий многочлен $f(x)$ с действительными коэффициентами представим, притом единственным способом (с точностью до порядка множителей), в виде произведения своего старшего коэффициента $a_0$ и нескольких многочленов с действительными коэф-*

*фициентами, линейных вида $x-\alpha$, соответствующих его действительным корням, и квадратных вида* (8), *соответствующих парам сопряженных комплексных корней.*

Для дальнейшего полезно подчеркнуть, что среди многочленов с действительными коэффициентами и со старшим коэффициентом 1, неразложимыми на множители меньшей степени или, как мы будем говорить, *неприводимыми*, являются лишь линейные многочлены вида $x-\alpha$ и квадратные многочлены вида (8).

## § 25*. Рациональные дроби

В курсе математического анализа изучаются, помимо целых рациональных функций, названных нами многочленами, также *дробно-рациональные функции*; это будут частные $\frac{f(x)}{g(x)}$ двух целых рациональных функций, где $g(x)\neq 0$. Над этими функциями производятся алгебраические операции по таким же законам, как над рациональными числами, т. е. как над дробями с целыми числителями и знаменателями. Равенство двух дробно-рациональных функций или, как мы будем дальше говорить, *рациональных дробей* также понимается в том же смысле, что и равенство дробей в элементарной арифметике. Для определенности мы будем рассматривать рациональные дроби с действительными коэффициентами; читатель без труда заметит, что все содержание настоящего параграфа может быть почти дословно перенесено на случай рациональных дробей с комплексными коэффициентами.

Рациональная дробь называется *несократимой*, если ее числитель взаимно прост со знаменателем.

*Всякая рациональная дробь равна некоторой несократимой дроби, определяемой однозначно с точностью до множителя нулевой степени, общего для числителя и знаменателя.*

Действительно, всякую рациональную дробь можно сократить на наибольший общий делитель ее числителя и знаменателя, после чего будет получена равная ей несократимая дробь. Если, далее, равны друг другу несократимые дроби $\frac{f(x)}{g(x)}$ и $\frac{\varphi(x)}{\psi(x)}$, т. е.

$$f(x)\,\psi(x)=g(x)\,\varphi(x), \tag{1}$$

то из взаимной простоты $f(x)$ и $g(x)$ следует, по свойству б) из § 21, что $\varphi(x)$ делится на $f(x)$, а из взаимной простоты $\varphi(x)$ и $\psi(x)$ следует, что $f(x)$ делится на $\varphi(x)$. Таким образом, $f(x)=c\varphi(x)$, а тогда из (1) следует $g(x)=c\psi(x)$.

Рациональная дробь называется *правильной*, если степень числителя меньше степени знаменателя. Если к числу правильных дробей мы условимся причислять многочлен 0, то справедлива следующая теорема:

*Всякая рациональная дробь представима, притом единственным способом, в виде суммы многочлена и правильной дроби.*

Действительно, если дана рациональная дробь $\frac{f(x)}{g(x)}$ и если, деля числитель на знаменатель, мы получим равенство

$$f(x)=g(x)\,q(x)+r(x),$$

где степень $r(x)$ меньше степени $g(x)$, то, как легко проверить,

$$\frac{f(x)}{g(x)}=q(x)+\frac{r(x)}{g(x)}.$$

Если имеет место также равенство

$$\frac{f(x)}{g(x)}=\bar{q}(x)+\frac{\varphi(x)}{\psi(x)},$$

где степень $\varphi(x)$ меньше степени $\psi(x)$, то мы получаем равенство

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

Так как слева стоит многочлен, а справа, как легко видеть, правильная дробь, то мы получим $q(x)-\bar{q}(x)=0$ и

$$\frac{\varphi(x)}{\psi(x)}-\frac{r(x)}{g(x)}=0.$$

Правильные рациональные дроби могут быть подвергнуты дальнейшему изучению. При этом напомним, что, как отмечено в конце предшествующего параграфа, неприводимыми действительными многочленами являются многочлены вида $x-\alpha$, где число $\alpha$ действительное, и многочлены вида $x^2-(\beta+\bar{\beta})\,x+\beta\bar{\beta}$, где $\beta$ и $\bar{\beta}$ — пара сопряженных комплексных чисел. Как легко проверить, в комплексном случае аналогичную роль играют многочлены вида $x-\alpha$, где $\alpha$ — любое комплексное число.

Правильная рациональная дробь $\frac{f(x)}{g(x)}$ называется *простейшей*, если ее знаменатель $g(x)$ является степенью неприводимого многочлена $p(x)$,

$$g(x)=p^k(x),\quad k\geqslant 1,$$

а степень числителя $f(x)$ меньше степени $p(x)$.

Справедлива следующая основная теорема:

*Всякая правильная рациональная дробь разлагается в сумму простейших дробей.*

Доказательство. Рассмотрим сначала правильную рациональную дробь $\frac{f(x)}{g(x)\,h(x)}$, где многочлены $g(x)$ и $h(x)$ взаимно просты,

$$(g(x),\ h(x))=1.$$

Существуют, следовательно, ввиду § 21, такие многочлены $\bar{u}(x)$ и $\bar{v}(x)$, что

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

Отсюда

$$g(x)[\bar{u}(x)f(x)]+h(x)[\bar{v}(x)f(x)]=f(x). \tag{2}$$

Пусть, деля произведение $\bar{u}(x)f(x)$ на $h(x)$, мы получим остаток $u(x)$, степень которого меньше степени $h(x)$. Тогда равенство (2) можно будет переписать в виде

$$g(x)u(x)+h(x)v(x)=f(x), \tag{3}$$

где $v(x)$ — многочлен, выражение которого могло бы быть без труда написано. Так как степень произведения $g(x)u(x)$ меньше степени произведения $g(x)h(x)$ и это же, по условию, верно для многочлена $f(x)$, то и произведение $h(x)v(x)$ имеет степень меньшую, чем $g(x)h(x)$, а поэтому степень $v(x)$ меньше степени $g(x)$. Из (3) вытекает теперь равенство

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

в правой части которого стоит сумма правильных дробей.

Если хотя бы один из знаменателей $g(x)$, $h(x)$ разлагается в произведение взаимно простых множителей, то можно выполнить дальнейшее разложение. Продолжая так далее, мы получим, что *всякая правильная дробь разлагается в сумму нескольких правильных дробей, каждая из которых имеет знаменателем степень некоторого неприводимого многочлена.* Точнее, если дана правильная дробь $\frac{f(x)}{g(x)}$, знаменатель которой имеет разложение на неприводимые множители

$$g(x)=p_1^{k_1}(x)p_2^{k_2}(x)\ldots p_l^{k_l}(x)$$

(всегда можно считать, конечно, что старший коэффициент знаменателя рациональной дроби равен единице), причем $p_i(x)\neq p_j(x)$ при $i\neq j$, то

$$\frac{f(x)}{g(x)}=\frac{u_1(x)}{p_1^{k_1}(x)}+\frac{u_2(x)}{p_2^{k_2}(x)}+\ldots+\frac{u_l(x)}{p_l^{k_l}(x)};$$

все слагаемые в правой части этого равенства являются правильными дробями.

Нам остается рассмотреть правильную дробь вида $\frac{u(x)}{p^k(x)}$, где $p(x)$ — неприводимый многочлен. Применяя алгоритм деления с остатком, разделим $u(x)$ на $p^{k-1}(x)$, полученный остаток разделим на $p^{k-2}(x)$ и т. д.

Мы придем к следующим равенствам:

$$u(x) = p^{k-1}(x)\, s_1(x) + u_1(x),$$
$$u_1(x) = p^{k-2}(x)\, s_2(x) + u_2(x),$$
$$\cdots\cdots\cdots\cdots$$
$$u_{k-2}(x) = p(x)\, s_{k-1}(x) + u_{k-1}(x).$$

Так как при этом степень $u(x)$, по условию, меньше степени $p^k(x)$, а степень каждого из остатков $u_i(x)$, $i = 1, 2, \ldots, k-1$, меньше степени соответствующего делителя $p^{k-i}(x)$, то степени всех частных $s_1(x), s_2(x), \ldots, s_{k-1}(x)$ будут строго меньше степени многочлена $p(x)$. Степень последнего остатка $u_{k-1}(x)$ также меньше степени $p(x)$. Из полученных равенств следует:

$$u(x) = p^{k-1}(x)\, s_1(x) + p^{k-2}(x)\, s_2(x) + \ldots + p(x)\, s_{k-1}(x) + u_{k-1}(x).$$

Отсюда мы приходим к искомому представлению рациональной дроби $\frac{u(x)}{p^k(x)}$ в виде суммы простейших дробей:

$$\frac{u(x)}{p^k(x)} = \frac{u_{k-1}(x)}{p^k(x)} + \frac{s_{k-1}(x)}{p^{k-1}(x)} + \ldots + \frac{s_2(x)}{p^2(x)} + \frac{s_1(x)}{p(x)}.$$

Основная теорема доказана. Ее можно дополнить следующей теоремой единственности:

*Всякая правильная рациональная дробь обладает единственным разложением в сумму простейших дробей.*

Пусть, в самом деле, некоторая правильная дробь может быть двумя способами представлена в виде суммы простейших дробей. Вычитая одно из этих представлений из другого и приводя подобные члены, мы получим сумму простейших дробей, тождественно равную нулю. Пусть знаменатели простейших дробей, составляющих эту сумму, будут некоторыми степенями различных неприводимых многочленов $p_1(x), p_2(x), \ldots, p_s(x)$ и пусть наивысшая степень многочлена $p_i(x)$, $i = 1, 2, \ldots, s$, являющаяся одним из этих знаменателей, будет $p_i^{k_i}(x)$. Умножим обе части рассматриваемого равенства на произведение $p_1^{k_1-1}(x)\, p_2^{k_2}(x) \ldots p_s^{k_s}(x)$. Все слагаемые нашей суммы, кроме одного, превратятся при этом в многочлены. Что же касается слагаемого $\frac{u(x)}{p_1^{k_1}(x)}$, то оно превратится в дробь, знаменателем которой служит $p_1(x)$, а числителем — произведение $u(x)\, p_2^{k_2}(x) \ldots p_s^{k_s}(x)$. Числитель не делится нацело на знаменатель, так как многочлен $p_1(x)$ неприводим, а все множители числителя с ним взаимно просты. Выполняя деление с остатком, мы в результате получим, что равна нулю сумма многочлена и отличной от нуля правильной дроби, что, однако, невозможно.

Пример. Разложить в сумму простейших дробей действительную правильную дробь $\frac{f(x)}{g(x)}$, где

$$f(x)=2x^4-10x^3+7x^2+4x+3,$$
$$g(x)=x^5-2x^3+2x^2-3x+2.$$

Легко проверяется, что

$$g(x)=(x+2)(x-1)^2(x^2+1),$$

причем каждый из многочленов $x+2$, $x-1$, $x^2+1$ неприводим. Из изложенной выше теории вытекает, что искомое разложение должно иметь вид

$$\frac{f(x)}{g(x)}=\frac{A}{x+2}+\frac{B}{(x-1)^2}+\frac{C}{x-1}+\frac{Dx+E}{x^2+1}, \tag{4}$$

где числа $A$, $B$, $C$, $D$ и $E$ еще должны быть разысканы.

Из (4) вытекает равенство

$$f(x)=A(x-1)^2(x^2+1)+B(x+2)(x^2+1)+C(x+2)(x-1)(x^2+1)+$$
$$+Dx(x+2)(x-1)^2+E(x+2)(x-1)^2. \tag{5}$$

Приравнивая коэффициенты при одинаковых степенях неизвестного $x$ из обеих частей равенства (5), мы получили бы систему пяти линейных уравнений относительно пяти неизвестных $A$, $B$, $C$, $D$, $E$, причем, как вытекает из доказанного выше, эта система обладает решением и притом единственным. Мы пойдем, однако, иным путем.

Полагая в равенстве (5) $x=-2$, мы придем к равенству $45A=135$, откуда

$$A=3. \tag{6}$$

Полагая, далее, в (5) $x=1$, мы получим $6B=6$, т. е.

$$B=1. \tag{7}$$

После этого положим в равенстве (5) последовательно $x=0$ и $x=-1$. Используя (6) и (7), мы получим уравнения

$$\left.\begin{aligned}-2C+2E&=-2,\\ -4C-4D+4E&=-8.\end{aligned}\right\} \tag{8}$$

Отсюда

$$D=1. \tag{9}$$

Положим, наконец, в равенстве (5) $x=2$. Используя (6), (7) и (9), мы придем к уравнению

$$20C+4E=-52,$$

которое вместе с первым из уравнений (8) дает

$$C=-2, \quad E=-3.$$

Таким образом,

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

ГЛАВА ШЕСТАЯ

# КВАДРАТИЧНЫЕ ФОРМЫ

## § 26. Приведение квадратичной формы к каноническому виду

Истоки теории квадратичных форм лежат в аналитической геометрии, а именно в теории кривых (и поверхностей) второго порядка. Известно, что уравнение центральной кривой второго порядка на плоскости, после перенесения начала прямоугольных координат в центр этой кривой, имеет вид

$$Ax^2+2Bxy+Cy^2=D. \tag{1}$$

Известно, далее, что можно совершить такой поворот осей координат на некоторый угол $\alpha$, т. е. такой переход от координат $x$, $y$ к координатам $x'$, $y'$:

$$\left.\begin{aligned} x&=x'\cos\alpha-y'\sin\alpha,\\ y&=x'\sin\alpha+y'\cos\alpha,\end{aligned}\right\} \tag{2}$$

что в новых координатах уравнение нашей кривой будет иметь «канонический» вид

$$A'x'^2+C'y'^2=D; \tag{3}$$

в этом уравнении коэффициент при произведении неизвестных $x'y'$ равен, следовательно, нулю. Преобразование координат (2) можно толковать, очевидно, как линейное преобразование неизвестных (см. § 13), притом невырожденное, так как определитель из его коэффициентов равен единице. Это преобразование применяется к левой части уравнения (1), и поэтому можно сказать, что левая часть уравнения (1) невырожденным линейным преобразованием (2) превращается в левую часть уравнения (3).

Многочисленные приложения потребовали построения аналогичной теории для случая, когда число неизвестных вместо двух равно любому $n$, а коэффициенты являются или действительными, или же любыми комплексными числами.

Обобщая выражение, стоящее в левой части уравнения (1), мы приходим к следующему понятию.

*Квадратичной формой* $f$ от $n$ неизвестных $x_1, x_2, \ldots, x_n$ называется сумма, каждый член которой является или квадратом одного из этих неизвестных, или произведением двух разных неизвестных. Квадратичная форма называется *действительной* или *комплексной* в зависимости от того, являются ли ее коэффициенты действительными или же могут быть любыми комплексными числами.

Считая, что в квадратичной форме $f$ уже сделано приведение подобных членов, введем следующие обозначения для коэффициентов этой формы: коэффициент при $x_i^2$ обозначим через $a_{ii}$, а коэффициент при произведении $x_ix_j$ для $i \neq j$ — через $2a_{ij}$ (сравнить с (1)!). Так как, однако, $x_ix_j = x_jx_i$, то коэффициент при этом произведении мог бы быть обозначен и через $2a_{ji}$, т. е. введенные нами обозначения предполагают справедливость равенства

$$a_{ji} = a_{ij}. \tag{4}$$

Член $2a_{ij}x_ix_j$ можно записать теперь в виде

$$2a_{ij}x_ix_j = a_{ij}x_ix_j + a_{ji}x_jx_i,$$

а всю квадратичную форму $f$ — в виде суммы всевозможных членов $a_{ij}x_ix_j$, где $i$ и $j$ уже независимо друг от друга принимают значения от 1 до $n$:

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j; \tag{5}$$

в частности, при $i = j$ получается член $a_{ii}x_i^2$.

Из коэффициентов $a_{ij}$ можно составить, очевидно, квадратную матрицу $A = (a_{ij})$ порядка $n$; она называется *матрицей квадратичной формы* $f$, а ее ранг $r$ — *рангом* этой квадратичной формы. Если, в частности, $r = n$, т. е. матрица — невырожденная, то и квадратичная форма $f$ называется *невырожденной*. Ввиду равенства (4) элементы матрицы $A$, симметричные относительно главной диагонали, равны между собой, т. е. матрица $A$ — *симметрическая*. Обратно, для любой симметрической матрицы $A$ $n$-го порядка можно указать вполне определенную квадратичную форму (5) от $n$ неизвестных, имеющую элементы матрицы $A$ своими коэффициентами.

Квадратичную форму (5) можно записать в ином виде, используя введенное в § 14 умножение прямоугольных матриц. Условимся сначала о следующем обозначении: если дана квадратная или вообще прямоугольная матрица $A$, то через $A'$ будет обозначаться матрица, полученная из матрицы $A$ транспонированием. Если матрицы $A$ и $B$ таковы, что их произведение определено, то имеет место равенство:

$$(AB)' = B'A', \tag{6}$$

т. е. *матрица, полученная транспонированием произведения, равна произведению матриц, получающихся транспонированием сомножителей, притом взятых в обратном порядке.*

В самом деле, если произведение $AB$ определено, то будет определено, как легко проверить, и произведение $B'A'$: число столбцов матрицы $B'$ равно числу строк матрицы $A'$. Элемент матрицы $(AB)'$, стоящий в ее $i$-й строке и $j$-м столбце, в матрице $AB$ расположен в $j$-й строке и $i$-м столбце. Он равен поэтому сумме произведений соответственных элементов $j$-й строки матрицы $A$ и $i$-го столбца матрицы $B$, т. е. равен сумме произведений соответственных элементов $j$-го столбца матрицы $A'$ и $i$-й строки матрицы $B'$. Этим равенство (6) доказано.

Заметим, что *матрица $A$ тогда и только тогда будет симметрической, если она совпадает со своей транспонированной*, т. е. если

$$A' = A.$$

Обозначим теперь через $X$ столбец, составленный из неизвестных,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}.$$

$X$ является матрицей, имеющей $n$ строк и один столбец. Транспонируя эту матрицу, получим матрицу

$$X' = (x_1, \ x_2, \ \ldots, \ x_n),$$

составленную из одной строки.

Квадратичная форма (5) с матрицей $A = (a_{ij})$ может быть записана теперь в виде следующего произведения:

$$f = X'AX. \tag{7}$$

Действительно, произведение $AX$ будет матрицей, состоящей из одного столбца:

$$AX = \begin{Bmatrix} \sum\limits_{j=1}^{n} a_{1j}x_j \\ \sum\limits_{j=1}^{n} a_{2j}x_j \\ \cdot \\ \cdot \\ \cdot \\ \sum\limits_{j=1}^{n} a_{nj}x_j \end{Bmatrix}.$$

Умножая эту матрицу слева на матрицу $X'$, мы получим «матрицу», состоящую из одной строки и одного столбца, а именно правую часть равенства (5).

Что произойдет с квадратичной формой $f$, если входящие в нее неизвестные $x_1, x_2, \ldots, x_n$ будут подвергнуты линейному преобразованию

$$x_i = \sum_{k=1}^{n} q_{ik} y_k, \quad i = 1, 2, \ldots, n, \tag{8}$$

с матрицей $Q = (q_{ik})$? Будем считать при этом, что если форма $f$ действительная, то и элементы матрицы $Q$ должны быть действительными. Обозначая через $Y$ столбец из неизвестных $y_1, y_2, \ldots, y_n$, запишем линейное преобразование (8) в виде матричного равенства:

$$X = QY. \tag{9}$$

Отсюда по (6)

$$X' = Y'Q'. \tag{10}$$

Подставляя (9) и (10) в запись (7) формы $f$, получаем:

$$f = Y'(Q'AQ)Y,$$

или

$$f = Y'BY,$$

где

$$B = Q'AQ.$$

Матрица $B$ будет симметрической, так как ввиду равенства (6), справедливого, очевидно, для любого числа множителей, и равенства $A' = A$, равносильного симметричности матрицы $A$, имеем:

$$B' = Q'A'Q = Q'AQ = B.$$

Таким образом, доказана следующая теорема:

*Квадратичная форма от $n$ неизвестных, имеющая матрицу $A$, после выполнения линейного преобразования неизвестных с матрицей $Q$ превращается в квадратичную форму от новых неизвестных, причем матрицей этой формы служит произведение $Q'AQ$.*

Предположим теперь, что мы выполняем невырожденное линейное преобразование, т. е. $Q$, а поэтому и $Q'$ — матрицы невырожденные. Произведение $Q'AQ$ получается в этом случае умножением матрицы $A$ на невырожденные матрицы и поэтому, как следует из результатов § 14, ранг этого произведения равен рангу матрицы $A$. Таким образом, *ранг квадратичной формы не меняется при выполнении невырожденного линейного преобразования.*

Рассмотрим теперь, по аналогии с указанной в начале параграфа геометрической задачей приведения уравнения центральной кривой второго порядка к каноническому виду (3), вопрос о приведении произвольной квадратичной формы некоторым невырожденным линейным преобразованием к виду суммы квадратов неизвестных, т. е. к такому виду, когда все коэффициенты при произведениях различных неизвестных равны нулю; этот специальный вид квадратичной формы называется *каноническим*. Предположим сначала, что

квадратичная форма $f$ от $n$ неизвестных $x_1, x_2, \ldots, x_n$ уже приведена невырожденным линейным преобразованием к каноническому виду

$$f = b_1 y_1^2 + b_2 y_2^2 + \ldots + b_n y_n^2, \tag{11}$$

где $y_1, y_2, \ldots, y_n$ — новые неизвестные. Некоторые из коэффициентов $b_1, b_2, \ldots, b_n$ могут, конечно, быть нулями. Докажем, что *число отличных от нуля коэффициентов в* (11) *непременно равно рангу $r$ формы $f$.*

В самом деле, так как мы пришли к (11) при помощи невырожденного преобразования, то квадратичная форма, стоящая в правой части равенства (11), также должна быть ранга $r$. Однако матрица этой квадратичной формы имеет диагональный вид

$$\begin{pmatrix} b_1 & & & 0 \\ & b_2 & & \\ & & \ddots & \\ 0 & & & b_n \end{pmatrix},$$

и требование, чтобы эта матрица имела ранг $r$, равносильно предположению, что на ее главной диагонали стоит ровно $r$ отличных от нуля элементов.

Перейдем к доказательству следующей основной теоремы о квадратичных формах.

*Всякая квадратичная форма может быть приведена некоторым невырожденным линейным преобразованием к каноническому виду. Если при этом рассматривается действительная квадратичная форма, то все коэффициенты указанного линейного преобразования можно считать действительными.*

Эта теорема верна для случая квадратичных форм от одного неизвестного, так как всякая такая форма имеет вид $ax^2$, являющийся каноническим. Мы можем, следовательно, вести доказательство индукцией по числу неизвестных, т. е. доказывать теорему для квадратичных форм от $n$ неизвестных, считая ее уже доказанной для форм с меньшим числом неизвестных.

Пусть дана квадратичная форма

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{12}$$

от $n$ неизвестных $x_1, x_2, \ldots, x_n$. Мы постараемся найти такое невырожденное линейное преобразование, которое выделило бы из $f$ квадрат одного из неизвестных, т. е. привело бы $f$ к виду суммы этого квадрата и некоторой квадратичной формы от остальных неизвестных. Эта цель легко достигается в том случае, если среди коэффициентов $a_{11}, a_{22}, \ldots, a_{nn}$, стоящих в матрице формы $f$ на главной диагонали, есть отличные от нуля, т. е. если в (12) входит

с отличным от нуля коэффициентом квадрат хотя бы одного из неизвестных $x_i$.

Пусть, например, $a_{11} \neq 0$. Тогда, как легко проверить, выражение $a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2$, являющееся квадратичной формой, содержит такие же члены с неизвестным $x_1$, как и наша форма $f$, а поэтому разность

$$f - a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2 = g$$

будет квадратичной формой, содержащей лишь неизвестные $x_2, \ldots, x_n$, но не $x_1$. Отсюда

$$f = a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2 + g.$$

Если мы введем обозначения

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \quad y_i = x_i \quad \text{при} \quad i = 2, 3, \ldots, n, \tag{13}$$

то получим

$$f = a_{11}^{-1}y_1^2 + g, \tag{14}$$

где $g$ будет теперь квадратичной формой от неизвестных $y_2, y_3, \ldots, y_n$. Выражение (14) есть искомое выражение для формы $f$, так как оно получено из (12) невырожденным линейным преобразованием, а именно преобразованием, обратным линейному преобразованию (13), которое имеет своим определителем $a_{11}$ и поэтому не вырождено.

Если же имеют место равенства $a_{11} = a_{22} = \ldots = a_{nn} = 0$, то предварительно нужно совершить вспомогательное линейное преобразование, приводящее к появлению в нашей форме $f$ квадратов неизвестных. Так как среди коэффициентов в записи (12) этой формы должны быть отличные от нуля,— иначе нечего было бы доказывать,— то пусть, например, $a_{12} \neq 0$, т. е. $f$ является суммой члена $2a_{12}x_1x_2$ и членов, в каждый из которых входит хотя бы одно из неизвестных $x_3, \ldots, x_n$.

Совершим теперь линейное преобразование

$$x_1 = z_1 - z_2, \quad x_2 = z_1 + z_2, \quad x_i = z_i \quad \text{при } i = 3, \ldots, n. \tag{15}$$

Оно будет невырожденным, так как имеет определитель

$$\begin{vmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} = 2 \neq 0.$$

В результате этого преобразования член $2a_{12}x_1x_2$ нашей формы примет вид

$$2a_{12}x_1x_2 = 2a_{12}(z_1 - z_2)(z_1 + z_2) = 2a_{12}z_1^2 - 2a_{12}z_2^2,$$

т. е. в форме $f$ появятся, с отличными от нуля коэффициентами, квадраты сразу двух неизвестных, причем они не могут сократиться ни с одним из остальных членов, так как в каждый из этих последних входит хотя бы одно из неизвестных $z_3, \ldots, z_n$. Теперь мы находимся в условиях уже рассмотренного выше случая, т. е. еще одним невырожденным линейным преобразованием можем привести форму $f$ к виду (14).

Для окончания доказательства остается отметить, что квадратичная форма $g$ зависит от меньшего, чем $n$, числа неизвестных и поэтому, по предположению индукции, некоторым невырожденным преобразованием неизвестных $y_2, y_3, \ldots, y_n$ приводится к каноническому виду. Это преобразование, рассматриваемое как (невырожденное, как легко видеть) преобразование всех $n$ неизвестных, при котором $y_1$ остается без изменения, приводит, следовательно, (14) к каноническому виду. Таким образом, квадратичная форма $f$ двумя или тремя невырожденными линейными преобразованиями, которые можно заменить одним невырожденным преобразованием — их произведением, приводится к виду суммы квадратов неизвестных с некоторыми коэффициентами. Число этих квадратов равно, как мы знаем, рангу формы $r$. Если, сверх того, квадратичная форма $f$ действительная, то коэффициенты как в каноническом виде формы $f$, так и в линейном преобразовании, приводящем $f$ к этому виду, будут действительными; в самом деле, и линейное преобразование, обратное (13), и линейное преобразование (15) имеют действительные коэффициенты.

Доказательство основной теоремы закончено. Метод, использованный в этом доказательстве, может быть применен в конкретных примерах для действительного приведения квадратичной формы к каноническому виду. Нужно лишь вместо индукции, которую мы использовали в доказательстве, последовательно выделять изложенным выше методом квадраты неизвестных.

Пример. Привести к каноническому виду квадратичную форму

$$f = 2x_1x_2 - 6x_2x_3 + 2x_3x_1. \tag{16}$$

Ввиду отсутствия в этой форме квадратов неизвестных мы выполним сначала невырожденное линейное преобразование

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_3 = y_3$$

с матрицей

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

после чего получим:

$$f = 2y_1^2 - 2y_2^2 - 4y_1y_3 - 8y_2y_3.$$

Теперь коэффициент при $y_1^2$ отличен от нуля, и поэтому из нашей формы можно выделить квадрат одного неизвестного. Полагая

$$z_1 = 2y_1 - 2y_3, \quad z_2 = y_2, \quad z_3 = y_3,$$

т. е. совершая линейное преобразование, для которого обратное будет иметь матрицу

$$B=\begin{pmatrix} \frac{1}{2} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

мы приведем $f$ к виду

$$f=\frac{1}{2}z_1^2-2z_2^2-2z_3^2-8z_2z_3.$$

Пока выделился лишь квадрат неизвестного $z_1$, так как форма еще содержит произведение двух других неизвестных. Используя неравенство нулю коэффициента при $z_2^2$, еще раз применим изложенный выше метод. Совершая линейное преобразование

$$t_1=z_1,\quad t_2=-2z_2-4z_3,\quad t_3=z_3,$$

для которого обратное имеет матрицу

$$C=\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix},$$

мы приведем, наконец, форму $f$ к каноническому виду

$$f=\frac{1}{2}t_1^2-\frac{1}{2}t_2^2+6t_3^2. \tag{17}$$

Линейное преобразование, приводящее (16) сразу к виду (17), будет иметь своей матрицей произведение

$$ABC=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 3 \\ \frac{1}{2} & -\frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Можно и непосредственной подстановкой проверить, что невырожденное $\left(\text{так как определитель равен } -\frac{1}{2}\right)$ линейное преобразование

$$\begin{aligned} x_1&=\frac{1}{2}t_1+\frac{1}{2}t_2+3t_3, \\ x_2&=\frac{1}{2}t_1-\frac{1}{2}t_2-t_3, \\ x_3&=\qquad\qquad\quad t_3 \end{aligned}$$

превращает (16) в (17).

Теория приведения квадратичной формы к каноническому виду построена по аналогии с геометрической теорией центральных кривых второго порядка, но не может считаться обобщением этой последней теории. В самом деле, в нашей теории допускается использование

любых невырожденных линейных преобразований, в то время как приведение кривой второго порядка к каноническому виду достигается применением линейных преобразований весьма специального вида (2), являющихся вращениями плоскости. Эта геометрическая теория может быть, однако, обобщена на случай квадратичных форм от $n$ неизвестных с действительными коэффициентами. Изложение этого обобщения, называемого приведением квадратичных форм к главным осям, будет дано в гл. 8.

## § 27. Закон инерции

Канонический вид, к которому приводится данная квадратичная форма, вовсе не является для нее однозначно определенным: всякая квадратичная форма может быть приведена к каноническому виду многими различными способами. Так, рассмотренная в предшествующем параграфе квадратичная форма $f=2x_1x_2-6x_2x_3+2x_3x_1$ невырожденным линейным преобразованием

$$\begin{aligned} x_1 &= t_1+3t_2+2t_3, \\ x_2 &= t_1-t_2-2t_3, \\ x_3 &= \phantom{t_1-{}} t_2 \end{aligned}$$

приводится к каноническому виду

$$f=2t_1^2+6t_2^2-8t_3^2,$$

отличному от полученного ранее.

Возникает вопрос, что общего у тех различных канонических квадратичных форм, к которым приводится данная форма $f$? Этот вопрос тесно связан, как мы увидим, с таким вопросом: при каком условии одна из двух данных квадратичных форм может быть переведена в другую невырожденным линейным преобразованием? Ответ на эти вопросы зависит, однако, от того, рассматриваются ли комплексные или действительные квадратичные формы.

Предположим сначала, что рассматриваются произвольные комплексные квадратичные формы и, вместе с тем, допускается употребление невырожденных линейных преобразований также с произвольными комплексными коэффициентами. Мы знаем, что всякая квадратичная форма $f$ от $n$ неизвестных, имеющая ранг $r$, приводится к каноническому виду

$$f=c_1y_1^2+c_2y_2^2+\ldots+c_ry_r^2,$$

где все коэффициенты $c_1, c_2, \ldots, c_r$ отличны от нуля. Пользуясь тем, что из всякого комплексного числа извлекается квадратный корень, выполним следующее невырожденное линейное преобразование:

$$z_i=\sqrt{c_i}\,y_i \quad \text{при} \quad i=1, 2, \ldots, r; \quad z_j=y_j \quad \text{при} \quad j=r+1, \ldots, n.$$

Оно приводит форму $f$ к виду

$$f = z_1^2 + z_2^2 + \ldots + z_r^2, \tag{1}$$

называемому ***нормальным***; это — просто сумма квадратов $r$ неизвестных с коэффициентами, равными единице.

Нормальный вид зависит лишь от ранга $r$ формы $f$, т. е. все квадратичные формы ранга $r$ приводятся к одному и тому же нормальному виду (1). Если, следовательно, формы $f$ и $g$ от $n$ неизвестных имеют одинаковый ранг $r$, то можно перевести $f$ в (1), а затем (1) в $g$, т. е. существует невырожденное линейное преобразование, переводящее $f$ в $g$. Так как, с другой стороны, никакое невырожденное линейное преобразование не изменяет ранга формы, то мы приходим к следующему результату:

***Две комплексные квадратичные формы от n неизвестных тогда и только тогда переводятся друг в друга невырожденными линейными преобразованиями с комплексными коэффициентами, если эти формы имеют один и тот же ранг.***

Из этой теоремы без труда вытекает, что ***каноническим видом комплексной квадратичной формы ранга r может служить всякая сумма квадратов r неизвестных с любыми отличными от нуля комплексными коэффициентами.***

Положение несколько более сложно в том случае, если рассматриваются действительные квадратичные формы и, что особенно важно, допускаются лишь линейные преобразования с действительными коэффициентами. В этом случае уже не всякую форму можно привести к виду (1), так как это могло бы потребовать извлечения квадратного корня из отрицательного числа. Если, однако, мы назовем теперь ***нормальным видом*** квадратичной формы сумму квадратов нескольких неизвестных с коэффициентами $+1$ или $-1$, то легко показать, что ***всякую действительную квадратичную форму f можно привести невырожденным линейным преобразованием с действительными коэффициентами к нормальному виду.***

В самом деле, форма $f$ ранга $r$ от $n$ неизвестных приводится к каноническому виду, который можно записать следующим образом (меняя, если нужно, нумерацию неизвестных):

$$f = c_1 y_1^2 + \ldots + c_k y_k^2 - c_{k+1} y_{k+1}^2 - \ldots - c_r y_r^2, \quad 0 \leqslant k \leqslant r,$$

где все числа $c_1, \ldots, c_k; c_{k+1}, \ldots, c_r$ отличны от нуля и положительны. Тогда невырожденное линейное преобразование с действительными коэффициентами

$$z_i = \sqrt{c_i}\, y_i \quad \text{при } i = 1, 2, \ldots, r, \quad z_j = y_j \quad \text{при } j = r+1, \ldots, n,$$

приводит $f$ к нормальному виду,

$$f = z_1^2 + \ldots + z_k^2 - z_{k+1}^2 - \ldots - z_r^2.$$

Общее число входящих сюда квадратов будет равно рангу формы.

Действительная квадратичная форма может быть приведена к нормальному виду многими различными преобразованиями, однако с точностью до нумерации неизвестных она приводится лишь к одному нормальному виду. Это показывает следующая важная теорема, называемая *законом инерции действительных квадратичных форм*:

***Число положительных и число отрицательных квадратов в нормальном виде, к которому приводится данная квадратичная форма с действительными коэффициентами действительным невырожденным линейным преобразованием, не зависят от выбора этого преобразования.***

Пусть, в самом деле, квадратичная форма $f$ ранга $r$ от $n$ неизвестных $x_1, x_2, \ldots, x_n$ двумя способами приведена к нормальному виду:

$$\begin{aligned} f &= y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 = \\ &= z_1^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2. \end{aligned} \tag{2}$$

Так как переход от неизвестных $x_1, x_2, \ldots, x_n$ к неизвестным $y_1, y_2, \ldots, y_n$ был невырожденным линейным преобразованием, то, обратно, вторые неизвестные также будут линейно выражаться через первые с отличным от нуля определителем:

$$y_i = \sum_{s=1}^{n} a_{is} x_s, \qquad i = 1, 2, \ldots, n. \tag{3}$$

Аналогично

$$z_j = \sum_{t=1}^{n} b_{jt} x_t, \qquad j = 1, 2, \ldots, n, \tag{4}$$

причем определитель из коэффициентов снова отличен от нуля. Коэффициенты же как в (3), так и в (4) — действительные числа.

Предположим теперь, что $k < l$, и напишем систему равенств

$$y_1 = 0, \ \ldots, \ y_k = 0, \quad z_{l+1} = 0, \ \ldots, \ z_r = 0, \ \ldots, \ z_n = 0. \tag{5}$$

Если левые части этих равенств будут заменены их выражениями из (3) и (4), мы получим систему $n - l + k$ линейных однородных уравнений с $n$ неизвестными $x_1, x_2, \ldots, x_n$. Число уравнений в этой системе меньше числа неизвестных, поэтому, как мы знаем из § 1, наша система обладает ненулевым действительным решением $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Заменим теперь в равенстве (2) все $y$ и все $z$ их выражениями (3) и (4), а затем подставим вместо неизвестных числа $\alpha_1, \alpha_2, \ldots, \alpha_n$. Если для краткости через $y_i(\alpha)$ и $z_j(\alpha)$ будут обозначены значения неизвестных $y_i$ и $z_j$, получающиеся после такой подстановки, то (2) превращается, ввиду (5), в равенство

$$-y_{k+1}^2(\alpha) - \ldots - y_r^2(\alpha) = z_1^2(\alpha) + \ldots + z_l^2(\alpha). \tag{6}$$

Так как все коэффициенты в (3) и (4) действительные, то все квадраты, входящие в равенство (6), положительны, а поэтому (6) влечет за собой равенство нулю всех этих квадратов; отсюда следуют равенства

$$z_1(\alpha)=0, \ \ldots, \ z_l(\alpha)=0. \tag{7}$$

С другой стороны, по самому выбору чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$

$$z_{l+1}(\alpha)=0, \ \ldots, \ z_r(\alpha)=0, \ \ldots, \ z_n(\alpha)=0. \tag{8}$$

Таким образом, система $n$ линейных однородных уравнений

$$z_i=0, \quad i=1, 2, \ldots, n,$$

с $n$ неизвестными $x_1, x_2, \ldots, x_n$ обладает, ввиду (7) и (8), ненулевым решением $\alpha_1, \alpha_2, \ldots, \alpha_n$, т. е. определитель этой системы должен быть равен нулю. Это противоречит, однако, тому, что преобразование (4) предполагалось невырожденным. К такому же противоречию мы придем при $l<k$. Отсюда следует равенство $k=l$, доказывающее теорему.

Число положительных квадратов в той нормальной форме, к которой приводится данная действительная квадратичная форма $f$, называется *положительным индексом инерции* этой формы, число отрицательных квадратов — *отрицательным индексом инерции*, а разность между положительным и отрицательным индексами инерции — *сигнатурой* формы $f$. Понятно, что при заданном ранге формы задание любого из определенных сейчас трех чисел вполне определяет два других, и поэтому в дальнейших формулировках можно будет говорить о любом из этих трех чисел.

Докажем теперь следующую теорему:

*Две квадратичные формы от $n$ неизвестных с действительными коэффициентами тогда и только тогда переводятся друг в друга невырожденными действительными линейными преобразованиями, если эти формы имеют одинаковые ранги и одинаковые сигнатуры.*

В самом деле, пусть форма $f$ переводится в форму $g$ невырожденным действительным преобразованием. Мы знаем, что это преобразование не меняет ранга формы. Оно не может менять и сигнатуры, так как в противном случае $f$ и $g$ приводились бы к различным нормальным видам, а тогда форма $f$ приводилась бы, в противоречие с законом инерции, к этим обоим нормальным видам. Обратно, если формы $f$ и $g$ имеют одинаковые ранги и одинаковые сигнатуры, то они приводятся к одному и тому же нормальному виду и поэтому могут быть переведены друг в друга.

Если дана квадратичная форма $g$ в каноническом виде,

$$g=b_1y_1^2+b_2y_2^2+\ldots+b_ry_r^2, \tag{9}$$

с не равными нулю действительными коэффициентами, то ранг этой формы равен, очевидно, $r$. Легко видеть, далее, употребляя уже применявшийся выше способ приведения такой формы к нормальному виду, что положительный индекс инерции формы $g$ равен числу положительных коэффициентов в правой части равенства (9). Отсюда и из предшествующей теоремы вытекает такой результат:

*Квадратичная форма $f$ тогда и только тогда будет иметь форму* (9) *своим каноническим видом, если ранг формы $f$ равен $r$, а положительный индекс инерции этой формы совпадает с числом положительных коэффициентов в* (9).

**Распадающиеся квадратичные формы.** Перемножая любые две линейные формы от $n$ неизвестных,

$$\varphi = a_1x_1 + a_2x_2 + \ldots + a_nx_n, \quad \psi = b_1x_1 + b_2x_2 + \ldots + b_nx_n,$$

мы получим, очевидно, некоторую квадратичную форму. Не всякая квадратичная форма может быть представлена в виде произведения двух линейных форм, и мы хотим вывести условия, при которых это имеет место, т. е. при которых квадратичная форма является *распадающейся*.

*Комплексная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ распадается тогда и только тогда, если ее ранг меньше или равен двум. Действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ распадается тогда и только тогда, если или ее ранг не больше единицы, или же он равен двум, а сигнатура равна нулю.*

Рассмотрим сначала произведение линейных форм $\varphi$ и $\psi$. Если хотя бы одна из этих форм нулевая, то их произведение будет квадратичной формой с нулевыми коэффициентами, т. е. оно имеет ранг 0. Если линейные формы $\varphi$ и $\psi$ пропорциональны,

$$\psi = c\varphi,$$

причем $c \neq 0$ и форма $\varphi$ ненулевая, то пусть, например, коэффициент $a_1$ отличен от нуля. Тогда невырожденное линейное преобразование

$$y_1 = a_1x_1 + \ldots + a_nx_n, \quad y_i = x_i \quad \text{при} \quad i = 2, 3, \ldots, n$$

приводит квадратичную форму $\varphi\psi$ к виду

$$\varphi\psi = cy_1^2.$$

Справа стоит квадратичная форма ранга 1, а поэтому и квадратичная форма $\varphi\psi$ имеет ранг 1. Если же, наконец, линейные формы $\varphi$ и $\psi$ не являются пропорциональными, то пусть, например,

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0.$$

Тогда линейное преобразование

$$y_1 = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$
$$y_2 = b_1x_1 + b_2x_2 + \ldots + b_nx_n,$$
$$y_i = x_i \text{ при } i = 3, 4, \ldots, n$$

будет невырожденным; оно приводит квадратичную форму $\varphi\psi$ к виду

$$\varphi\psi = y_1y_2.$$

Справа стоит квадратичная форма ранга 2, имеющая в случае действительных коэффициентов сигнатуру 0.

Перейдем к доказательству обратного утверждения. Квадратичная форма ранга 0 может, конечно, рассматриваться как произведение двух линейных форм, одна из которых нулевая. Далее, квадратичная форма $f(x_1, x_2, \ldots, x_n)$ ранга 1 невырожденным линейным преобразованием приводится к виду

$$f = cy_1^2, \quad c \neq 0,$$

т. е. к виду

$$f = (cy_1)\, y_1.$$

Выражая $y_1$ линейно через $x_1, x_2, \ldots, x_n$, мы получим представление формы $f$ в виде произведения двух линейных форм. Наконец, действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ ранга 2 и сигнатуры 0 приводится невырожденным линейным преобразованием к виду

$$f = y_1^2 - y_2^2;$$

к этому же виду может быть приведена любая комплексная квадратичная форма ранга 2. Однако

$$y_1^2 - y_2^2 = (y_1 - y_2)(y_1 + y_2),$$

но справа, после замены $y_1$ и $y_2$ их линейными выражениями через $x_1, x_2, \ldots, x_n$, будет стоять произведение двух линейных форм. Теорема доказана.

## § 28. Положительно определенные формы

Квадратичная форма $f$ от $n$ неизвестных с действительными коэффициентами называется *положительно определенной*, если она приводится к нормальному виду, состоящему из $n$ положительных квадратов, т. е. если и ранг, и положительный индекс инерции этой формы равны числу неизвестных.

Следующая теорема дает возможность охарактеризовать положительно определенные формы, не приводя их к нормальному или каноническому виду.

*Квадратичная форма $f$ от $n$ неизвестных $x_1, x_2, \ldots, x_n$ с действительными коэффициентами тогда и только тогда будет положительно определенной, если при всяких действительных значениях этих неизвестных, хотя бы одно из которых отлично от нуля, эта форма получает положительные значения.*

Доказательство. Пусть форма $f$ положительно определенная, т. е. приводится к нормальному виду

$$f = y_1^2 + y_2^2 + \ldots + y_n^2, \tag{1}$$

причем

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \ldots, n, \tag{2}$$

с отличным от нуля определителем из действительных коэффициентов $a_{ij}$. Если мы хотим подставить в $f$ произвольные действительные значения неизвестных $x_1, x_2, \ldots, x_n$, хотя бы одно из которых отлично от нуля, то можно подставить их сначала в (2), а затем значения, полученные для всех $y_i$,— в (1). Заметим, что значения, полученные для $y_1, y_2, \ldots, y_n$ из (2), не могут все сразу равняться нулю, так как иначе мы получили бы, что система линейных однородных уравнений

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = 1, 2, \ldots, n,$$

обладает ненулевым решением, хотя ее определитель отличен от нуля. Подставляя найденные для $y_1, y_2, \ldots, y_n$ значения в (1), мы получим значение формы $f$, равное сумме квадратов $n$ действительных чисел, которые не все равны нулю; это значение будет, следовательно, строго положительным.

Обратно, пусть форма $f$ не является положительно определенной, т. е. или ее ранг, или положительный индекс инерции меньше $n$. Это означает, что в нормальном виде этой формы, к которому она приводится, скажем, невырожденным линейным преобразованием (2), квадрат хотя бы одного из новых неизвестных, например $y_n$, или отсутствует совсем, или же содержится со знаком минус. Покажем, что в этом случае можно подобрать такие действительные значения для неизвестных $x_1, x_2, \ldots, x_n$, которые не все равны нулю, что значение формы $f$ при этих значениях неизвестных равно нулю или даже отрицательно. Такими будут, например, те значения для $x_1, x_2, \ldots, x_n$, которые мы получим, решая по правилу Крамера систему линейных уравнений, получающихся из (2) при $y_1 = y_2 = \ldots = y_{n-1} = 0$, $y_n = 1$. Действительно, при этих значениях неизвестных $x_1, x_2, \ldots, x_n$ форма $f$ равна нулю, если $y_n^2$ не входит в нормальный вид этой формы, и равна $-1$, если $y_n^2$ входит в нормальный вид со знаком минус.

Теорема, сейчас доказанная, используется всюду, где применяются положительно определенные квадратичные формы. С ее помощью нельзя, однако, по коэффициентам формы установить, будет ли эта форма положительно определенной. Для этой цели служит другая теорема, которую мы сформулируем и докажем после того, как введем одно вспомогательное понятие.

Пусть дана квадратичная форма $f$ от $n$ неизвестных с матрицей $A=(a_{ij})$. Миноры порядка 1, 2, ..., $n$ этой матрицы, расположенные в ее левом верхнем углу, т. е. миноры

$$a_{11},\ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \ldots,\ \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix},\ \ldots,\ \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

из которых последний совпадает, очевидно, с определителем матрицы $A$, называются *главными минорами* формы $f$.

Справедлива следующая теорема:

***Квадратичная форма $f$ от $n$ неизвестных с действительными коэффициентами тогда и только тогда будет положительно определенной, если все ее главные миноры строго положительны.***

Доказательство. При $n=1$ теорема верна, так как форма имеет в этом случае вид $ax^2$ и поэтому положительно определена тогда и только тогда, если $a>0$. Будем поэтому доказывать теорему для случая $n$ неизвестных, предполагая, что для квадратичных форм от $n-1$ неизвестных она уже доказана.

Сделаем сначала следующее замечание:

Если квадратичная форма $f$ с действительными коэффициентами, составляющими матрицу $A$, подвергается невырожденному линейному преобразованию с действительной матрицей $Q$, то *знак определителя формы (т. е. определителя ее матрицы) не меняется.*

Действительно, после преобразования мы получаем квадратичную форму с матрицей $Q'AQ$, однако, ввиду $|Q'|=|Q|$,

$$|Q'AQ|=|Q'|\cdot|A|\cdot|Q|=|A|\cdot|Q|^2,$$

т. е. определитель $|A|$ умножается на положительное число.

Пусть теперь дана квадратичная форма

$$f=\sum_{i,\,j=1}^{n} a_{ij}x_ix_j.$$

Ее можно записать в виде

$$f=\varphi(x_1,\ x_2,\ \ldots,\ x_{n-1})+2\sum_{i=1}^{n-1} a_{in}x_ix_n+a_{nn}x_n^2, \tag{3}$$

где $\varphi$ будет квадратичной формой от $n-1$ неизвестных, составленной из тех членов формы $f$, в которые не входит неизвестное $x_n$.

Главные миноры формы $\varphi$ совпадают, очевидно, со всеми, кроме последнего, главными минорами формы $f$.

Пусть форма $f$ положительно определена. Форма $\varphi$ также будет в этом случае положительно определенной: если бы существовали такие значения неизвестных $x_1, x_2, \ldots, x_{n-1}$, не все равные нулю, при которых форма $\varphi$ получает не строго положительное значение, то, полагая дополнительно $x_n=0$, мы получили бы, ввиду (3), также не строго положительное значение формы $f$, хотя не все значения неизвестных $x_1, x_2, \ldots, x_{n-1}, x_n$ равны нулю. Поэтому, по индуктивному предположению, все главные миноры формы $\varphi$, т. е. все главные миноры формы $f$, кроме последнего, строго положительны. Что же касается последнего главного минора формы $f$, т. е. определителя самой матрицы $A$, то его положительность вытекает из следующих соображений: форма $f$, ввиду ее положительной определенности, невырожденным линейным преобразованием приводится к нормальному виду, состоящему из $n$ положительных квадратов. Определитель этого нормального вида строго положителен, а поэтому ввиду сделанного выше замечания положителен и определитель самой формы $f$.

Пусть теперь строго положительны все главные миноры формы $f$. Отсюда вытекает положительность всех главных миноров формы $\varphi$, т. е., по индуктивному предположению, положительная определенность этой формы. Существует, следовательно, такое невырожденное линейное преобразование неизвестных $x_1, x_2, \ldots, x_{n-1}$, которое приводит форму $\varphi$ к виду суммы $n-1$ положительных квадратов от новых неизвестных $y_1, y_2, \ldots, y_{n-1}$. Это линейное преобразование можно дополнить до (невырожденного) линейного преобразования всех неизвестных $x_1, x_2, \ldots, x_n$, полагая $x_n=y_n$. Ввиду (3) форма $f$ приводится указанным преобразованием к виду

$$f=\sum_{i=1}^{n-1} y_i^2+2\sum_{i=1}^{n-1} b_{in}y_iy_n+b_{nn}y_n^2; \tag{4}$$

точные выражения коэффициентов $b_{in}$ для нас несущественны. Так как

$$y_i^2+2b_{in}y_iy_n=(y_i+b_{in}y_n)^2-b_{in}^2y_n^2,$$

то невырожденное линейное преобразование

$$\begin{aligned} z_i&=y_i+b_{in}y_n, \quad i=1, 2, \ldots, n-1,\\ z_n&=\qquad\quad y_n \end{aligned}$$

приводит, ввиду (4), форму $f$ к каноническому виду

$$f=\sum_{i=1}^{n-1} z_i^2+cz_n^2. \tag{5}$$

Для доказательства положительной определенности формы $f$ остается доказать положительность числа $c$. Определитель формы, стоящей в правой части равенства (5), равен $c$. Этот определитель должен, однако, быть положительным, так как правая часть равенства (5) получена из формы $f$ двумя невырожденными линейными преобразованиями, а определитель формы $f$ был, как последний из главных миноров этой формы, положительным.

Доказательство теоремы закончено.

Примеры 1. Квадратичная форма

$$f=5x_1^2+x_2^2+5x_3^2+4x_1x_2-8x_1x_3-4x_2x_3$$

положительно определена, так как ее главные миноры

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

положительны.

2. Квадратичная форма

$$f=3x_1^2+x_2^2+5x_3^2+4x_1x_2-8x_1x_3-4x_2x_3$$

не будет положительно определенной, так как ее второй главный минор отрицателен:

$$\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}=-1.$$

Заметим, что по аналогии с положительно определенными квадратичными формами можно ввести *отрицательно определенные формы*, т. е. такие невырожденные квадратичные формы с действительными коэффициентами, нормальный вид которых содержит лишь отрицательные квадраты неизвестных. Вырожденные квадратичные формы, нормальный вид которых состоит из квадратов одного знака, называются иногда *полуопределенными*. Наконец, *неопределенными* будут такие квадратичные формы, нормальный вид которых содержит как положительные, так и отрицательные квадраты неизвестных.

ГЛАВА СЕДЬМАЯ

# ЛИНЕЙНЫЕ ПРОСТРАНСТВА

## § 29. Определение линейного пространства. Изоморфизм

Определение $n$-мерного векторного пространства, данное в § 8, начиналось с определения $n$-мерного вектора как упорядоченной системы $n$ чисел. Для $n$-мерных векторов были введены затем сложение и умножение на числа, что и привело к понятию $n$-мерного векторного пространства. Первыми примерами векторных пространств являются совокупности векторов-отрезков, выходящих из начала координат на плоскости или в трехмерном пространстве. Однако, встречаясь с этими примерами в курсе геометрии, мы не всегда считаем необходимым задавать векторы их компонентами в некоторой фиксированной системе координат, так как и сложение векторов и их умножение на скаляр определяются геометрически, независимо от выбора системы координат. Именно, сложение векторов на плоскости или в пространстве производится по правилу параллелограмма, а умножение вектора на число $\alpha$ означает растяжение этого вектора в $\alpha$ раз (с изменением направления вектора на противоположное, если $\alpha$ отрицательно). Целесообразно и в общем случае дать «бескоординатное» определение векторного пространства, т. е. определение, не требующее задания векторов упорядоченными системами чисел. Сейчас будет дано такое определение. Это определение является аксиоматическим: в нем ничего не будет сказано о свойствах отдельного вектора, но будут перечислены те свойства, которыми должны обладать операции над векторами.

Пусть дано множество $V$; его элементы будут обозначаться малыми латинскими буквами: $a$, $b$, $c$, ...[1]). Пусть, далее, в множестве $V$ определены *операция сложения*, ставящая в соответствие всякой паре элементов $a$, $b$ из $V$ однозначно определенный элемент $a+b$ из $V$, называемый их *суммой*, и *операция умножения на действительное число*, причем *произведение* $\alpha a$ *элемента* $a$ на число $\alpha$ однозначно определено и принадлежит к $V$.

---

[1]) В отличие от того, что было принято в гл. 2, мы будем в настоящей и следующей главах обозначать векторы малыми латинскими буквами, а числа — малыми греческими буквами.

Элементы множества $V$ будут называться *векторами*, а само $V$—*действительным линейным* (или *векторным*, или *аффинным*) *пространством*, если указанные операции обладают следующими свойствами I—VIII:

I. Сложение коммутативно, $a+b=b+a$.

II. Сложение ассоциативно, $(a+b)+c=a+(b+c)$.

III. В $V$ существует *нулевой элемент* 0, удовлетворяющий условию: $a+0=a$ для всех $a$ из $V$.

Легко доказать, используя I, *единственность нулевого элемента:* если $0_1$ и $0_2$—два нулевых элемента, то

$$0_1+0_2=0_1,$$
$$0_1+0_2=0_2+0_1=0_2,$$

откуда $0_1=0_2$.

IV. Для всякого элемента $a$ в $V$ существует *противоположный элемент* $-a$, удовлетворяющий условию: $a+(-a)=0$.

Легко проверяется, ввиду II и I, *единственность противоположного элемента:* если $(-a)_1$ и $(-a)_2$—два противоположных элемента для $a$, то

$$(-a)_1+[a+(-a)_2]=(-a_1)+0=(-a)_1,$$
$$[(-a)_1+a]+(-a)_2=0+(-a)_2=(-a)_2,$$

откуда $(-a)_1=(-a)_2$.

Из аксиом I—IV выводится *существование и единственность разности* $a-b$, т. е. такого элемента, который удовлетворяет уравнению

$$b+x=a. \tag{1}$$

Действительно, можно положить

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

так как

$$b+[a+(-b)]=[b+(-b)]+a=0+a=a.$$

Если же существует еще такой элемент $c$, который удовлетворяет уравнению (1), т. е.

$$b+c=a,$$

то, прибавляя к обеим частям этого равенства элемент $-b$, получаем, что

$$c=a+(-b).$$

Дальнейшие аксиомы V—VIII (ср. § 8) связывают умножение на число со сложением и с операциями над числами. Именно, для любых элементов $a$, $b$ из $V$, для любых действительных чисел $\alpha$, $\beta$

и для действительного числа 1 должны иметь место равенства:

V. $$\alpha(a+b)=\alpha a+\alpha b;$$

VI. $$(\alpha+\beta)a=\alpha a+\beta a;$$

VII. $$(\alpha\beta)a=\alpha(\beta a);$$

VIII. $$1\cdot a=a.$$

Укажем некоторые простейшие следствия из этих аксиом.

[1]. $$\alpha\cdot 0=0.$$

Действительно, для некоторого $a$ из $V$

$$\alpha a=\alpha(a+0)=\alpha a+\alpha\cdot 0,$$

т. е.

$$\alpha\cdot 0=\alpha a-\alpha a=\alpha a+[-(\alpha a)]=0.$$

[2]. $$0\cdot a=0,$$

где слева стоит число нуль, а справа — нулевой элемент из $V$.

Для доказательства возьмем любое число $\alpha$. Тогда

$$\alpha a=(\alpha+0)a=\alpha a+0\cdot a,$$

откуда

$$0\cdot a=\alpha a-\alpha a=0.$$

[3]. Если $\alpha a=0$, то или $\alpha=0$, или $a=0$.

Действительно, если $\alpha\neq 0$, т. е. число $\alpha^{-1}$ существует, то

$$a=1\cdot a=(\alpha^{-1}\alpha)a=\alpha^{-1}(\alpha a)=\alpha^{-1}\cdot 0=0.$$

[4]. $$\alpha(-a)=-\alpha a.$$

В самом деле,

$$\alpha a+\alpha(-a)=\alpha[a+(-a)]=\alpha\cdot 0=0,$$

т. е. элемент $\alpha(-a)$ противоположен элементу $\alpha a$.

[5]. $$(-\alpha)a=-\alpha a.$$

Действительно,

$$\alpha a+(-\alpha)a=[\alpha+(-\alpha)]a=0\cdot a=0,$$

т. е. элемент $(-\alpha)a$ противоположен элементу $\alpha a$.

[6]. $$\alpha(a-b)=\alpha a-\alpha b.$$

Действительно, по [4],

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

[7]. $$(\alpha-\beta)a=\alpha a-\beta a.$$

В самом деле,

$$(\alpha-\beta)a=[\alpha+(-\beta)]a=\alpha a+(-\beta)a=\alpha a+(-\beta a)=\alpha a-\beta a.$$

Заметим, что перечисленными выше аксиомами и следствиями из них мы будем пользоваться дальше без специальных оговорок.

Выше дано определение действительного линейного пространства. Если бы мы предположили, что в множестве $V$ определено умножение не только на действительные, но и на любые комплексные числа, то, сохраняя те же аксиомы I—VIII, получили бы определение *комплексного линейного пространства*. Для определенности ниже рассматриваются действительные линейные пространства, однако все, что будет сказано в настоящей главе, переносится дословно на случай комплексных линейных пространств.

Примеры действительных линейных пространств могут быть легко указаны. Ими будут, прежде всего, те $n$-мерные действительные векторные пространства, составленные из векторов-строк, которые изучались в гл. 2. Линейными пространствами будут и множества векторов-отрезков, выходящих из начала координат на плоскости или в трехмерном пространстве, если операции сложения и умножения на число понимать в том геометрическом смысле, который был указан в начале параграфа.

Существуют также примеры линейных пространств, так сказать «бесконечномерных». Рассмотрим всевозможные последовательности действительных чисел; они имеют вид

$$a=(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n,\ \ldots).$$

Операции над последовательностями будем производить покомпонентно: если

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

то

$$a+b=(\alpha_1+\beta_1,\ \alpha_2+\beta_2,\ \ldots,\ \alpha_n+\beta_n,\ \ldots);$$

с другой стороны, для любого действительного числа $\gamma$

$$\gamma a=(\gamma\alpha_1,\ \gamma\alpha_2,\ \ldots,\ \gamma\alpha_n,\ \ldots).$$

Все аксиомы I—VIII выполняются, т. е. мы получаем действительное линейное пространство.

Примером бесконечномерного пространства будет также множество всевозможных действительных функций действительного переменного, если сложение функций и их умножение на действительное число понимать так, как это принято в теории функций, т. е. как сложение или умножение на число значений функций при каждом значении независимого переменного.

**Изоморфизм.** Нашей ближайшей целью будет выделение среди всех линейных пространств тех, которые естественно назвать конечномерными. Введем сначала одно общее понятие.

В определении линейного пространства говорилось о свойствах операций над векторами, но ничего не говорилось о свойствах

самих векторов. Ввиду этого может случиться, что хотя векторы некоторых двух данных линейных пространств по своей природе совершенно различны, однако с точки зрения свойств операций эти два пространства неразличимы. Точное определение таково:

Два действительных линейных пространства $V$ и $V'$ называются *изоморфными*, если между их векторами установлено взаимно однозначное соответствие — всякому вектору $a$ из $V$ сопоставлен вектор $a'$ из $V'$, образ вектора $a$, причем различные векторы из $V$ обладают различными образами и всякий вектор из $V'$ служит образом некоторого вектора из $V$, — и если при этом соответствии образом суммы двух векторов служит сумма образов этих векторов,

$$(a+b)'=a'+b', \tag{2}$$

а образом произведения вектора на число служит произведение образа этого вектора на то же число,

$$(\alpha a)'=\alpha a'. \tag{3}$$

Отметим, что взаимно однозначное соответствие между пространствами $V$ и $V'$, удовлетворяющее условиям (2) и (3), называется *изоморфным соответствием*.

Так, пространство векторов-отрезков на плоскости, выходящих из начала координат, изоморфно двумерному векторному пространству, составленному из упорядоченных пар действительных чисел: мы получим изоморфное соответствие между этими пространствами, если на плоскости фиксируем некоторую систему координат и всякому вектору-отрезку сопоставим упорядоченную пару его координат.

Докажем следующее свойство изоморфизма линейных пространств: *образом нуля пространства $V$ при изоморфном соответствии между пространствами $V$ и $V'$ служит нуль пространства $V'$.*

Пусть, в самом деле, $a$ будет некоторый вектор из $V$, $a'$ — его образ в $V'$. Тогда, ввиду (2),

$$a'=(a+0)'=a'+0',$$

т. е. $0'$ будет нулем пространства $V'$.

## § 30. Конечномерные пространства. Базы

Как читатель без труда может проверить, те два определения линейной зависимости векторов-строк, которые были даны в § 9, равно как и доказательство эквивалентности этих определений, используют лишь операции над векторами и поэтому могут быть перенесены на случай любых линейных пространств. В аксиоматически определенных линейных пространствах можно говорить, следовательно, о линейно независимых системах векторов, о макси-

мальных линейно независимых системах, если такие существуют, и т. д.

*Если линейные пространства $V$ и $V'$ изоморфны, то система векторов $a_1, a_2, \ldots, a_k$ из $V$ тогда и только тогда линейно зависима, если линейно зависима система их образов $a_1', a_2', \ldots, a_k'$ в $V'$.*

Заметим, что если соответствие $a \to a'$ (для всех $a$ из $V$) является изоморфным соответствием между $V$ и $V'$, то и обратное соответствие $a' \to a$ также будет изоморфным. Поэтому достаточно рассмотреть случай, когда линейно зависима система $a_1, a_2, \ldots, a_k$. Пусть существуют такие числа $\alpha_1, \alpha_2, \ldots, \alpha_k$, не все равные нулю, что

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

Образом правой части этого равенства при рассматриваемом изоморфизме служит, как мы знаем, нуль $0'$ пространства $V'$. Беря образ левой части и применяя несколько раз (2) и (3), получаем

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

т. е. система $a_1', a_2', \ldots, a_k'$ также оказалась линейно зависимой.

**Конечномерные пространства.** Линейное пространство $V$ называется *конечномерным*, если в нем можно найти конечную максимальную линейно независимую систему векторов; всякая такая система векторов будет называться *базой* пространства $V$.

Конечномерное линейное пространство может обладать многими различными базами. Так, в пространстве векторов-отрезков на плоскости базой служит любая пара векторов, отличных от нуля и не лежащих на одной прямой. Заметим, что наше определение конечномерного пространства не дает пока ответа на вопрос, могут ли в этом пространстве существовать базы, состоящие из разного числа векторов. Больше того, можно было бы допустить даже, что в некоторых конечномерных пространствах существуют базы со сколь угодно большим числом векторов. Сейчас мы приступим к выяснению того, каково же положение на самом деле.

Пусть линейное пространство $V$ обладает базой

$$e_1, e_2, \ldots, e_n, \tag{1}$$

состоящей из $n$ векторов. Если $a$ — произвольный вектор из $V$, то из максимальности линейно независимой системы (1) следует, что $a$ линейно выражается через эту систему,

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n. \tag{2}$$

С другой стороны, ввиду линейной независимости системы (1) выражение (2) будет для вектора $a$ единственным: если

$$a = \alpha_1' e_1 + \alpha_2' e_2 + \ldots + \alpha_n' e_n,$$

то

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

откуда

$$\alpha_i=\alpha_i', \quad i=1,\ 2,\ \ldots,\ n.$$

Таким образом, вектору $a$ однозначно соответствует строка

$$(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n) \tag{3}$$

коэффициентов его выражения (2) через базу (1) или, как мы будем говорить, *строка его координат в базе* (1). Обратно, всякая строка вида (3), т. е. всякий $n$-мерный вектор в смысле гл. 2, служит строкой координат в базе (1) для некоторого вектора пространства $V$, а именно для вектора, записывающегося через базу (1) в виде (2).

Мы получили, следовательно, взаимно однозначное соответствие между всеми векторами пространства $V$ и всеми векторами $n$-мерного векторного пространства строк. Покажем, что это соответствие, зависящее, понятно, от выбора базы (1), является изоморфным.

Возьмем в пространстве $V$, помимо вектора $a$, выражающегося через базу (1) в виде (2), также вектор $b$, выражение которого через базу (1) будет

$$b=\beta_1e_1+\beta_2e_2+\ldots+\beta_ne_n.$$

Тогда

$$a+b=(\alpha_1+\beta_1)\,e_1+(\alpha_2+\beta_2)\,e_2+\ldots+(\alpha_n+\beta_n)\,e_n,$$

т. е. *сумме векторов $a$ и $b$ соответствует сумма строк их координат в базе* (1). С другой стороны,

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

т. е. *произведению вектора $a$ на число $\gamma$ соответствует произведение строки его координат в базе* (1) *на это же число* $\gamma$.

Этим доказана следующая теорема:

*Всякое линейное пространство, обладающее базой из $n$ векторов, изоморфно $n$-мерному векторному пространству строк.*

Как мы знаем, при изоморфном соответствии между линейными пространствами линейно зависимая система векторов переходит в линейно зависимую и обратно, а поэтому линейно независимая переходит в линейно независимую. Отсюда следует, что *при изоморфном соответствии база переходит в базу.*

В самом деле, пусть база $e_1,\ e_2,\ \ldots,\ e_n$ пространства $V$ переходит при изоморфном соответствии между пространствами $V$ и $V'$ в систему векторов $e_1',\ e_2',\ \ldots,\ e_n'$ пространства $V'$, которая хотя и линейно независима, но не является максимальной. В $V'$ можно найти, следовательно, такой вектор $f'$, что система $e_1',\ e_2',\ \ldots,\ e_n',\ f'$ остается линейно независимой. Вектор $f'$ служит, однако, образом при рассматриваемом изоморфизме для некоторого вектора $f$ из $V$.

Мы получаем, что система векторов $e_1, e_2, \ldots, e_n, f$ должна быть линейно независимой в противоречие с определением базы.

Мы знаем, далее (см. § 9), что в $n$-мерном векторном пространстве строк все максимальные линейно независимые системы состоят из $n$ векторов, что всякая система из $n+1$ вектора линейно зависима и что всякая линейно независимая система векторов содержится в некоторой максимальной линейно независимой системе. Используя установленные выше свойства изоморфных соответствий, мы приходим к следующим результатам:

*Все базы конечномерного линейного пространства $V$ состоят из одного и того же числа векторов.* Если это число равно $n$, то $V$ будет называться *$n$-мерным линейным пространством*, а число $n$ — *размерностью* этого пространства.

*Всякая система из $n+1$ вектора $n$-мерного линейного пространства линейно зависима.*

*Всякая линейно независимая система векторов $n$-мерного линейного пространства содержится в некоторой базе этого пространства.*

Теперь легко проверить, что указанные выше примеры действительных линейных пространств — пространство последовательностей и пространство функций — не являются конечномерными пространствами: в каждом из этих пространств читатель без труда найдет линейно независимые системы, состоящие из сколь угодно большого числа векторов.

**Связь между базами.** Объектом изучения являются для нас конечномерные линейные пространства. Понятно, что, изучая $n$-мерные линейные пространства, мы по существу изучаем то $n$-мерное векторное пространство строк, которое было введено еще в гл. 2. Однако раньше в этом пространстве была выделена одна база — а именно база, составленная из единичных векторов, т. е. векторов, у которых одна координата равна единице, а все остальные координаты равны нулю, — и все векторы пространства задавались строками их координат в этой базе; теперь же все базы пространства являются для нас равноправными.

Посмотрим, как много баз можно найти в $n$-мерном линейном пространстве и как эти базы связаны друг с другом.

Пусть в $n$-мерном линейном пространстве $V$ заданы базы

$$e_1, \quad e_2, \ldots, e_n \tag{4}$$

и

$$e_1', \quad e_2', \ldots, e_n'. \tag{5}$$

Каждый вектор базы (5), как и всякий вектор пространства $V$, однозначно записывается через базу (4),

$$e_i' = \sum_{j=1}^{n} \tau_{ij} e_j, \quad i = 1, 2, \ldots, n. \tag{6}$$

Матрица

$$T=\begin{pmatrix}\tau_{11} & \dots & \tau_{1n}\\ \cdot & \cdot & \cdot\\ \tau_{n1} & \dots & \tau_{nn}\end{pmatrix},$$

строки которой являются строками координат векторов (5) в базе (4), называется *матрицей перехода* от базы (4) к базе (5).

Связь между базами (4) и (5) и матрицей перехода $T$ можно записать, ввиду (6), в виде матричного равенства

$$\begin{pmatrix}e_1'\\ e_2'\\ \cdot\\ \cdot\\ \cdot\\ e_n'\end{pmatrix}=\begin{pmatrix}\tau_{11} & \tau_{12} & \dots & \tau_{1n}\\ \tau_{21} & \tau_{22} & \dots & \tau_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ \tau_{n1} & \tau_{n2} & \dots & \tau_{nn}\end{pmatrix}\begin{pmatrix}e_1\\ e_2\\ \cdot\\ \cdot\\ \cdot\\ e_n\end{pmatrix} \tag{7}$$

или, обозначая базы (4) и (5), записанные в столбец, соответственно через $e$ и $e'$, в виде

$$e'=Te.$$

С другой стороны, если $T'$ — матрица перехода от базы (5) к базе (4), то

$$e=T'e'.$$

Отсюда

$$e=(T'T)\,e,$$
$$e'=(TT')\,e',$$

т. е., ввиду линейной независимости баз $e$ и $e'$,

$$T'T=TT'=E,$$

откуда

$$T'=T^{-1}.$$

Этим доказано, что *матрица перехода от одной базы к другой всегда является невырожденной матрицей.*

*Всякая невырожденная квадратная матрица порядка $n$ с действительными элементами служит матрицей перехода от данной базы $n$-мерного действительного линейного пространства к некоторой другой базе.*

Пусть, в самом деле, дана база (4) и невырожденная матрица $T$ порядка $n$. Возьмем в качестве (5) систему векторов, для которых строки матрицы $T$ служат строками координат в базе (4); имеет место, следовательно, равенство (7). Векторы (5) линейно независимы — линейная зависимость между ними влекла бы за собой линейную зависимость строк матрицы $T$ в противоречие с ее невырожденностью. Поэтому система (5), как линейно независимая

система, состоящая из $n$ векторов, является базой нашего пространства, а матрица $T$ служит матрицей перехода от базы (4) к базе (5).

Мы видим, что в $n$-мерном линейном пространстве можно найти столь же много различных баз, как много существует различных невырожденных квадратных матриц порядка $n$. Правда, при этом две базы, состоящие из одних и тех же векторов, но записанных в различном порядке, считаются различными.

**Преобразование координат вектора.** Пусть в $n$-мерном линейном пространстве даны базы (4) и (5) с матрицей перехода $T=(\tau_{ij})$,

$$e' = Te.$$

Найдем связь между строками координат произвольного вектора $a$ в этих базах.

Пусть

$$a = \sum_{j=1}^{n} \alpha_j e_j, \tag{8}$$

$$a = \sum_{i=1}^{n} \alpha_i' e_i'.$$

Используя (6), получаем:

$$a = \sum_{i=1}^{n} \alpha_i' \left( \sum_{j=1}^{n} \tau_{ij} e_j \right) = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} \alpha_i' \tau_{ij} \right) e_j.$$

Сравнивая с (8) и используя единственность записи вектора через базу, получаем:

$$\alpha_j = \sum_{i=1}^{n} \alpha_i' \tau_{ij}, \qquad j = 1, 2, \ldots, n,$$

т. е. имеет место матричное равенство

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1', \alpha_2', \ldots, \alpha_n')\, T.$$

Таким образом, *строка координат вектора $a$ в базе $e$ равна строке координат этого вектора в базе $e'$, умноженной справа на матрицу перехода от базы $e$ к базе $e'$.*

Отсюда следует, понятно, равенство

$$(\alpha_1', \alpha_2', \ldots, \alpha_n') = (\alpha_1, \alpha_2, \ldots, \alpha_n)\, T^{-1}.$$

Пример. Рассмотрим трехмерное действительное линейное пространство с базой

$$e_1, e_2, e_3. \tag{9}$$

Векторы

$$\left.\begin{aligned} e_1' &= 5e_1 - e_2 - 2e_3, \\ e_2' &= 2e_1 + 3e_2, \\ e_3' &= -2e_1 + e_2 + e_3 \end{aligned}\right\} \tag{10}$$

также составляют базу в этом пространстве, причем матрицей перехода от (9) к (10) служит матрица

$$T=\begin{pmatrix} 5 & -1 & -2 \\ 2 & 3 & 0 \\ -2 & 1 & 1 \end{pmatrix},$$

откуда

$$T^{-1}=\begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix}.$$

Вектор

$$a=e_1+4e_2-e_3$$

имеет поэтому в базе (10) строку координат

$$(\alpha_1', \alpha_2', \alpha_3')=(1,\ 4,\ -1)\begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix}=(-13,\ 6,\ -27),$$

т. е.

$$a=-13e_1'+6e_2'-27e_3'.$$

## § 31. Линейные преобразования

В гл. 3 мы уже встречались с понятием линейного преобразования неизвестных. Понятие, которое будет сейчас введено, носит такое же название, но имеет иной характер. Впрочем, некоторые связи между этими двумя одноименными понятиями могли бы быть указаны.

Пусть дано $n$-мерное действительное линейное пространство, которое обозначим через $V_n$. Рассмотрим *преобразование* этого пространства, т. е. отображение, переводящее каждый вектор $a$ пространства $V_n$ в некоторый вектор $a'$ этого же пространства. Вектор $a'$ называется *образом* вектора $a$ при рассматриваемом преобразовании.

Если преобразование обозначено через $\varphi$, то образ вектора $a$ условимся записывать не через $\varphi(a)$ или $\varphi a$, что читателю было бы привычнее, а через $a\varphi$. Таким образом,

$$a'=a\varphi.$$

Преобразование $\varphi$ линейного пространства $V_n$ называется *линейным преобразованием* этого пространства, если сумму любых двух векторов $a$, $b$ оно переводит в сумму образов этих векторов,

$$(a+b)\varphi=a\varphi+b\varphi, \tag{1}$$

а произведение любого вектора $a$ на любое число $\alpha$ переводит в произведение образа вектора $a$ на это же число $\alpha$,

$$(\alpha a)\varphi=\alpha(a\varphi). \tag{2}$$

Из этого определения немедленно вытекает, что *линейное преобразование линейного пространства переводит любую линейную*

*комбинацию данных векторов* $a_1, a_2, \ldots, a_k$ *в линейную комбинацию (с теми же коэффициентами) образов этих векторов,*

$$(\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k)\varphi = \alpha_1 (a_1\varphi) + \alpha_2 (a_2\varphi) + \ldots + \alpha_k (a_k\varphi). \quad (3)$$

Докажем следующее утверждение:

*При любом линейном преобразовании* $\varphi$ *линейного пространства* $V_n$ *нулевой вектор* 0 *остается неподвижным,*

$$0\varphi = 0,$$

*а образом вектора, противоположного для данного вектора* $a$, *служит вектор, противоположный для образа вектора* $a$,

$$(-a)\varphi = -a\varphi.$$

В самом деле, если $b$ — произвольный вектор, то, ввиду (2),

$$0\varphi = (0 \cdot b)\varphi = 0 \cdot (b\varphi) = 0.$$

С другой стороны,

$$(-a)\varphi = [(-1)a]\varphi = (-1)(a\varphi) = -a\varphi.$$

Понятие линейного преобразования линейного пространства возникло как обобщение известного из курса аналитической геометрии понятия аффинного преобразования плоскости или трехмерного пространства; действительно, условия (1) и (2) для аффинных преобразований выполняются. Эти условия выполняются и для проекций векторов на плоскости или в трехмерном пространстве на некоторую прямую (или на некоторую плоскость). Таким образом, например, в двумерном линейном пространстве векторов-отрезков, выходящих из начала координат плоскости, преобразование, переводящее всякий вектор в его проекцию на некоторую ось, проходящую через начало координат, будет линейным преобразованием.

Примерами линейных преобразований в произвольном пространстве $V_n$ служат *тождественное преобразование* $\varepsilon$, оставляющее всякий вектор $a$ на месте,

$$a\varepsilon = a,$$

и *нулевое преобразование* $\omega$, отображающее всякий вектор $a$ в нуль,

$$a\omega = 0.$$

Сейчас будет получено некоторое обозрение всех линейных преобразований линейного пространства $V_n$. Пусть

$$e_1, e_2, \ldots, e_n \quad (4)$$

— база этого пространства; как и раньше, базу (4), расположенную в столбец, будем обозначать через $e$. Так как всякий вектор $a$ пространства $V_n$ однозначно представляется в виде линейной комбинации

векторов базы (4), то, ввиду (3), образ вектора $a$ с теми же коэффициентами выражается через образы векторов (4). Иными словами, *всякое линейное преобразование* $\varphi$ *пространства* $V_n$ *однозначно определяется заданием образов* $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ *всех векторов фиксированной базы* (4).

*Какова бы ни была упорядоченная система из* $n$ *векторов пространства* $V_n$,

$$c_1, c_2, \ldots, c_n, \tag{5}$$

*существует, притом единственное, такое линейное преобразование* $\varphi$ *этого пространства, что* (5) *служит системой образов векторов базы* (4) *при этом преобразовании,*

$$e_i\varphi = c_i, \quad i = 1, 2, \ldots, n. \tag{6}$$

Единственность преобразования $\varphi$ уже доказана выше и нужно доказать лишь его существование. Определим преобразование $\varphi$ следующим образом: если $a$ — произвольный вектор пространства и

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

— его запись в базе (4), то положим

$$a\varphi = \sum_{i=1}^{n} \alpha_i c_i. \tag{7}$$

Докажем линейность этого преобразования. Если

$$b = \sum_{i=1}^{n} \beta_i e_i$$

— любой другой вектор пространства, то

$$(a+b)\varphi = \left[\sum_{i=1}^{n} (\alpha_i + \beta_i) e_i\right]\varphi = \sum_{i=1}^{n} (\alpha_i + \beta_i) c_i =$$
$$= \sum_{i=1}^{n} \alpha_i c_i + \sum_{i=1}^{n} \beta_i c_i = a\varphi + b\varphi.$$

Если же $\gamma$ — любое число, то

$$(\gamma a)\varphi = \left[\sum_{i=1}^{n} (\gamma\alpha_i) e_i\right]\varphi = \sum_{i=1}^{n} (\gamma\alpha_i) c_i = \gamma \sum_{i=1}^{n} \alpha_i c_i = \gamma (a\varphi).$$

Что же касается справедливости равенств (6), то она вытекает из определения (7) преобразования $\varphi$, так как все координаты вектора $e_i$ в базе (4) равны нулю, кроме $i$-й координаты, равной единице.

*Нами установлено, следовательно, взаимно однозначное соответствие между всеми линейными преобразованиями линейного пространства* $V_n$ *и всеми упорядоченными системами* (5) *из* $n$ *векторов этого пространства.*

Всякий вектор $c_i$ обладает, однако, определенной записью в базе (4),

$$c_i = \sum_{j=1}^{n} \alpha_{ij} e_j, \quad i = 1, 2, \ldots, n. \tag{8}$$

Из координат вектора $c_i$ в базе (4) можно составить квадратную матрицу

$$A = (\alpha_{ij}), \tag{9}$$

беря в качестве ее $i$-й строки строку координат вектора $c_i$, $i = 1, 2, \ldots, n$. Так как система (5) была произвольной, то матрица $A$ будет произвольной квадратной матрицей порядка $n$ с действительными элементами.

*Мы имеем, таким образом, взаимно однозначное соответствие между всеми линейными преобразованиями пространства $V_n$ и всеми квадратными матрицами порядка $n$*; это соответствие зависит, конечно, от выбора базы (4).

Будем говорить, что матрица $A$ *задает* линейное преобразование $\varphi$ в базе (4), или, короче, что $A$ есть *матрица линейного преобразования* $\varphi$ в базе (4). Если через $e\varphi$ мы обозначим столбец, составленный из образов векторов базы (4), то из (6), (8) и (9) вытекает следующее матричное равенство, полностью описывающее связи, существующие между линейным преобразованием $\varphi$, базой $e$ и матрицей $A$, задающей это линейное преобразование в этой базе:

$$e\varphi = Ae. \tag{10}$$

Покажем, как, зная матрицу $A$ линейного преобразования $\varphi$ в базе (4), по координатам вектора $a$ в этой базе найти координаты его образа $a\varphi$. Если

$$a = \sum_{i=1}^{n} \alpha_i e_i,$$

то

$$a\varphi = \sum_{i=1}^{n} \alpha_i (e_i\varphi),$$

что равносильно матричному равенству

$$a\varphi = (\alpha_1, \alpha_2, \ldots, \alpha_n)(e\varphi).$$

Используя (10) и учитывая, что ассоциативность умножения матриц легко проверяется и в том случае, когда одна из матриц является столбцом, составленным из векторов, мы получаем:

$$a\varphi = [(\alpha_1, \alpha_2, \ldots, \alpha_n) A] e.$$

Отсюда следует, что *строка координат вектора $a\varphi$ равна строке координат вектора $a$, умноженной справа на матрицу $A$ линейного преобразования $\varphi$, все в базе* (4).

Пример. Пусть в базе $e_1$, $e_2$, $e_3$ трехмерного линейного пространства линейное преобразование $\varphi$ задается матрицей

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

Если

$$a=5e_1+e_2-2e_3,$$

то

$$(5,\ 1,\ -2)\begin{pmatrix} -2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & -4 & 1 \end{pmatrix}=(-9,\ 16,\ 0),$$

т. е.

$$a\varphi=-9e_1+16e_2.$$

**Связь между матрицами линейного преобразования в разных базах.** Само собою разумеется, что матрица, задающая линейное преобразование, зависит от выбора базы. Покажем, какова связь между матрицами, задающими в разных базах одно и то же линейное преобразование.

Пусть даны базы $e$ и $e'$ с матрицей перехода $T$,

$$e'=Te, \tag{11}$$

и пусть линейное преобразование $\varphi$ задается в этих базах соответственно матрицами $A$ и $A'$,

$$e\varphi=Ae, \quad e'\varphi=A'e'. \tag{12}$$

Второе из равенств (12) приводит, ввиду (11), к равенству

$$(Te)\,\varphi=A'\,(Te).$$

Однако

$$(Te)\,\varphi=T\,(e\varphi).$$

Действительно, если $(\tau_{i1},\ \tau_{i2},\ \ldots,\ \tau_{in})$ — $i$-я строка матрицы $T$, то

$$(\tau_{i1}e_1+\tau_{i2}e_2+\ldots+\tau_{in}e_n)\,\varphi=\tau_{i1}\,(e_1\varphi)+\tau_{i2}\,(e_2\varphi)+\ldots+\tau_{in}\,(e_n\varphi).$$

Таким образом, ввиду (12),

$$(Te)\,\varphi=T\,(e\varphi)=T\,(Ae)=(TA)\,e,$$
$$A'\,(Te)=(A'T)\,e,$$

т. е.

$$(TA)\,e=(A'T)\,e.$$

Если хотя бы для одного $i$, $1\leqslant i\leqslant n$, $i$-я строка матрицы $TA$ будет отлична от $i$-й строки матрицы $A'T$, то две различные линейные комбинации векторов $e_1$, $e_2$, $\ldots$, $e_n$ окажутся равными друг другу, что противоречит линейной независимости базы $e$. Таким образом,

$$TA=A'T,$$

откуда, ввиду невырожденности матрицы перехода $T$,

$$A' = TAT^{-1}, \quad A = T^{-1}A'T. \tag{13}$$

Заметим, что квадратные матрицы $B$ и $C$ называются ***подобными***, если они связаны равенством

$$C = Q^{-1}BQ,$$

где $Q$ — некоторая невырожденная матрица. При этом говорят, что матрица $C$ получена из матрицы $B$ *трансформированием* матрицей $Q$.

Доказанные выше равенства (13) можно сформулировать, таким образом, в виде следующей важной теоремы:

*Матрицы, задающие одно и то же линейное преобразование в разных базах, подобны между собой. При этом матрица линейного преобразования $\varphi$ в базе $e'$ получается трансформированием матрицы этого преобразования в базе $e$ матрицей перехода от базы $e'$ к базе $e$.*

Подчеркнем, что если матрица $A$ задает линейное преобразование $\varphi$ в базе $e$, то любая матрица $B$, подобная матрице $A$,

$$B = Q^{-1}AQ,$$

также задает преобразование $\varphi$ в некоторой базе, а именно в базе, получающейся из базы $e$ при помощи матрицы перехода $Q^{-1}$.

**Операции над линейными преобразованиями.** Сопоставляя каждому линейному преобразованию пространства $V_n$ его матрицу в фиксированной базе, мы получаем, как доказано, взаимно однозначное соответствие между всеми линейными преобразованиями и всеми квадратными матрицами порядка $n$. Естественно ожидать, что операциям сложения и умножения матриц, а также умножения матрицы на число, будут соответствовать аналогичные операции над линейными преобразованиями.

Пусть в пространстве $V_n$ даны линейные преобразования $\varphi$ и $\psi$. Назовем *суммой* этих преобразований преобразование $\varphi + \psi$, определяемое равенством

$$a(\varphi + \psi) = a\varphi + a\psi; \tag{14}$$

оно переводит, следовательно, любой вектор $a$ в сумму его образов при преобразованиях $\varphi$ и $\psi$.

*Преобразование $\varphi + \psi$ является линейным.* Действительно, для любых векторов $a$ и $b$ и любого числа $\alpha$

$$(a+b)(\varphi+\psi) = (a+b)\varphi + (a+b)\psi = \\ = a\varphi + b\varphi + a\psi + b\psi = a(\varphi+\psi) + b(\varphi+\psi);$$

$$(\alpha a)(\varphi+\psi) = (\alpha a)\varphi + (\alpha a)\psi = \alpha(a\varphi) + \alpha(a\psi) = \\ = \alpha(a\varphi + a\psi) = \alpha[a(\varphi+\psi)].$$

С другой стороны, назовем *произведением* линейных преобразований $\varphi$ и $\psi$ преобразование $\varphi\psi$, определяемое равенством

$$a(\varphi\psi) = (a\varphi)\psi, \tag{15}$$

т. е. получающееся в результате последовательного выполнения преобразований $\varphi$ и $\psi$.

*Преобразование* $\varphi\psi$ *является линейным*:

$$(a+b)(\varphi\psi) = [(a+b)\varphi]\psi = (a\varphi + b\varphi)\psi = \\ = (a\varphi)\psi + (b\varphi)\psi = a(\varphi\psi) + b(\varphi\psi);$$

$$(\alpha a)(\varphi\psi) = [(\alpha a)\varphi]\psi = [\alpha(a\varphi)]\psi = \alpha[(a\varphi)\psi] = \alpha[a(\varphi\psi)].$$

Назовем, наконец, *произведением линейного преобразования* $\varphi$ *на число* $\varkappa$ преобразование $\varkappa\varphi$, определяемое равенством

$$a(\varkappa\varphi) = \varkappa(a\varphi); \tag{16}$$

образы при преобразовании $\varphi$ всех векторов умножаются, следовательно, на число $\varkappa$.

*Преобразование* $\varkappa\varphi$ *является линейным*:

$$(a+b)(\varkappa\varphi) = \varkappa[(a+b)\varphi] = \varkappa(a\varphi + b\varphi) = \\ = \varkappa(a\varphi) + \varkappa(b\varphi) = a(\varkappa\varphi) + b(\varkappa\varphi);$$

$$(\alpha a)(\varkappa\varphi) = \varkappa[(\alpha a)\varphi] = \varkappa[\alpha(a\varphi)] = \alpha[\varkappa(a\varphi)] = \alpha[a(\varkappa\varphi)].$$

Пусть в базе $e_1, e_2, \ldots, e_n$ преобразования $\varphi$ и $\psi$ задаются соответственно матрицами $A = (\alpha_{ij})$ и $B = (\beta_{ij})$,

$$e\varphi = Ae, \quad e\psi = Be.$$

Тогда, ввиду (14),

$$e_i(\varphi + \psi) = e_i\varphi + e_i\psi = \sum_{j=1}^{n} \alpha_{ij} e_j + \sum_{j=1}^{n} \beta_{ij} e_j = \sum_{j=1}^{n} (\alpha_{ij} + \beta_{ij}) e_j,$$

т. е.

$$e(\varphi + \psi) = (A + B)e.$$

Таким образом, *матрица суммы линейных преобразований в любой базе равна сумме матриц этих преобразований в той же базе.*

С другой стороны, ввиду (15),

$$e_i(\varphi\psi) = (e_i\varphi)\psi = \left(\sum_{j=1}^{n} \alpha_{ij} e_j\right)\psi = \sum_{j=1}^{n} \alpha_{ij}(e_j\psi) = \\ = \sum_{j=1}^{n} \alpha_{ij}\left(\sum_{k=1}^{n} \beta_{jk} e_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} \alpha_{ij}\beta_{jk}\right) e_k,$$

т. е.

$$e(\varphi\psi) = (AB)e.$$