А. Л. Городенцев*

А Л Г Е Б Р А

1-й курс

Независимый Московский Университет
2020/21 уч. год

* ВШЭ, ИТЭФ, НМУ, e-mail:gorod@itep.ru , http://gorod.bogomolov-lab.ru/

Оглавление

§1. Множества и отображения

1.1. Множества. Мы не будем заниматься основаниями теории множеств, полагаясь на школьное интуитивное представление о множестве как «абстрактной совокупности элементов произвольной природы». Элементы множества мы часто будем называть *точками*. Все точки в любом множестве, по определению, различны.

Множество X задано, как только про любой объект можно сказать, является он элементом множества X или нет. Принадлежность точки x множеству X записывается как $x \in X$. Два множества *равны*, если они состоят из одних и тех же элементов. Существует единственное множество, не содержащее ни одного элемента. Оно называется *пустым* и обозначается $\varnothing$. Если множество X конечно, то мы обозначаем через $|X|$ количество точек в нём.

Множество X называется *подмножеством* множества Y, если каждый его элемент $x \in X$ лежит также и в Y. В этом случае пишут $X \subset Y$. Отметим, что пустое множество является подмножеством любого множества и всякое множество является подмножеством самого себя. Подмножества, отличные от всего множества, называются *собственными подмножествами*. В частности, пустое подмножество собственное.

Упражнение 1.1. Сколько всего подмножеств (включая пустое и несобственное) имеется у множества, состоящего из n элементов?

Для заданных множеств X, Y их *объединение* $X \cup Y$ состоит из всех элементов, принадлежащих хотя бы одному из множеств X, Y; *пересечение* $X \cap Y$ состоит из всех элементов, принадлежащих одновременно каждому из множеств X, Y; *разность* $X \setminus Y$ состоит из всех элементов множества X, которые не содержатся в Y.

Упражнение 1.2. Проверьте, что операция пересечения выражается через разность по формуле $X \cap Y = X \setminus (X \setminus Y)$. Можно ли выразить разность через пересечение и объединение?

Если множество X является объединением непересекающихся подмножеств Y и Z, то говорят, что X является *дизъюнктным объединением* Y и Z и пишут $X = Y \sqcup Z$.

Множество $X \times Y$, элементами которого являются, по определению, всевозможные пары (x, y) с $x \in X$, $y \in Y$, называется *декартовым* (или *прямым*) *произведением* множеств X и Y.

1.2. Отображения. Отображение $f : X \to Y$ из множества X в множество Y есть правило, однозначно сопоставляющее каждой точке $x \in X$ некоторую точку $y = f(x) \in Y$, которая называется *образом* точки x при отображении f. Множество всех таких точек $x \in X$, образ которых равен заданной точке $y \in Y$, называется *полным прообразом* точки y (или *слоем* отображения f над y) и обозначается

$$f^{-1}(y) \stackrel{\text{def}}{=} \{x \in X \mid f(x) = y\}.$$

Полные прообразы различных точек не пересекаются и могут быть как пустыми, так и состоять из многих точек. Множество всех $y \in Y$, имеющих непустой прообраз, называется *образом отображения* $f : X \to Y$ и обозначается

$$\operatorname{im}(f) \stackrel{\text{def}}{=} \{y \in Y \mid f^{-1}(y) \neq \varnothing\} = \{y \in Y \mid \exists\, x \in X : f(x) = y\}.$$

Два отображения $f : X \to Y$ и $g : X \to Y$ *равны*, если $f(x) = g(x)$ для всех $x \in X$. Множество всех отображений из множества X в множество Y обозначается $\operatorname{Hom}(X, Y)$.

Отображение $f : X \to Y$ называется *наложением* (а также *сюръекцией* или *эпиморфизмом*), если $\operatorname{im}(f) = Y$, т. е. когда прообраз каждой точки $y \in Y$ не пуст. Мы будем изображать сюръективные отображения стрелками $X \twoheadrightarrow Y$. Отображение f называется *вложением* (а также

инъекцией, или *мономорфизмом*), если $f(x_1) \neq f(x_2)$ при $x_1 \neq x_2$, т. е. когда прообраз каждой точки $y \in Y$ содержит не более одного элемента. Инъективные отображения изображаются стрелками $X \hookrightarrow Y$.

Упражнение 1.3. Перечислите все отображения $\{0, 1, 2\} \to \{0, 1\}$ и $\{0, 1\} \to \{0, 1, 2\}$. Сколько среди них вложений и сколько наложений?

Отображение $f : X \to Y$, которое является одновременно и вложением и наложением, называется *взаимно однозначным* (а также *биекцией* или *изоморфизмом*). Биективность отображения f означает, что для каждого $y \in Y$ существует единственный $x \in X$, такой что $f(x) = y$. Мы будем обозначать биекции стрелками $X \xrightarrow{\sim} Y$.

Упражнение 1.4. Из отображений: а) $\mathbb{N} \to \mathbb{N} : x \mapsto x^2$ б) $\mathbb{Z} \to \mathbb{Z} : x \mapsto x^2$ в) $\mathbb{Z} \to \mathbb{Z} : x \mapsto 7x$ г) $\mathbb{Q} \to \mathbb{Q} : x \mapsto 7x$ выделите все инъекции, все сюръекции и все биекции.

Отображения $X \to X$ из множества X в себя обычно называют *эндоморфизмами* множества X. Множество всех эндоморфизмов обозначается $\mathrm{End}(X) \stackrel{\text{def}}{=} \mathrm{Hom}(X, X)$.

Упражнение 1.5 (принцип Дирихле). Покажите, что следующие три условия на множество X равносильны: а) X бесконечно б) существует вложение $X \hookrightarrow X$, не являющееся наложением в) существует наложение $X \twoheadrightarrow X$, не являющееся вложением.

Взаимно однозначные эндоморфизмы $X \xrightarrow{\sim} X$ называются *автоморфизмами* X. Множество всех автоморфизмов обозначается через $\mathrm{Aut}(X)$. Автоморфизмы можно воспринимать как *перестановки* элементов множества X. У всякого множества X имеется *тождественный автоморфизм* $\mathrm{Id}_X : X \to X$, который переводит каждый элемент в самого себя: $\forall\, x \in X\ \mathrm{Id}_X(x) = x$.

Упражнение 1.6. Счётно[1] ли множество $\mathrm{Aut}(\mathbb{N})$?

Пример 1.1 (запись отображений словами)
Рассмотрим множества $X = \{1, 2, \ldots, n\}$ и $Y = \{1, 2, \ldots, m\}$, сопоставим каждому отображению $f : X \to Y$ последовательность его значений:

$$w(f) \stackrel{\text{def}}{=} \big(f(x_1), f(x_2), \ldots, f(x_n)\big) \tag{1-1}$$

и будем воспринимать её как n-буквенное слово, написанное при помощи m-буквенного алфавита Y. Так, отображениям $f : \{1, 2\} \to \{1, 2, 3\}$ и $g : \{1, 2, 3\} \to \{1, 2, 3\}$, действующим по правилам $f(1) = 3, f(2) = 2$ и $g(1) = 1, g(2) = 2, g(3) = 2$, сопоставятся слова $w(f) = (3, 2)$ и $w(g) = (1, 2, 2)$, составленные из букв алфавита $\{1, 2, 3\}$. Запись отображения словом задаёт биекцию

$$w : \ \mathrm{Hom}(X, Y) \xrightarrow{\sim} \{\text{слова из } |X| \text{ букв в алфавите } Y\}\,, \quad f \mapsto w(f)\,. \tag{1-2}$$

Инъективные отображения записываются при этом словами, в которых нет повторяющихся букв, а сюръективные отображения — словами, в которых используются все без исключения буквы алфавита Y. Взаимно однозначным отображениям отвечают слова, в которых задействованы все буквы алфавита Y, причём каждая — ровно по одному разу.

1.3. Слои отображений. Задание отображения $f : X \to Y$ равносильно разбиению X в дизъюнктное объединение непустых подмножеств $f^{-1}(y)$, занумерованных точками $y \in \mathrm{im}(f)$:

$$X = \bigsqcup_{y \in \mathrm{im}(f)} f^{-1}(y)\,. \tag{1-3}$$

[1] Множество M называется *счётным* если существует биекция $\mathbb{N} \xrightarrow{\sim} M$.

Такой взгляд на отображения часто оказывается полезным при подсчёте числа элементов в том или ином множестве. Например, когда все непустые слои отображения $f\colon X \to Y$ состоят из одного и того же числа точек $m = |f^{-1}(y)|$, число элементов в образе отображения f связано с числом элементов в множестве X соотношением

$$|X| = m \cdot |\operatorname{im} f|\,, \tag{1-4}$$

которое при всей своей простоте имеет много разнообразных применений.

Предложение 1.1
Если множества X и Y конечны, то $|\operatorname{Hom}(X,Y)| = |Y|^{|X|}$.

Доказательство. Зафиксируем какую-нибудь точку $x \in X$ и рассмотрим *отображение вычисления*[1], сопоставляющее каждому отображению $f\colon X \to Y$ его значение в точке x:

$$\operatorname{ev}_x\colon \operatorname{Hom}(X,Y) \to Y\,, \quad f \mapsto f(x)\,. \tag{1-5}$$

Для каждого $y \in Y$ слой $\operatorname{ev}_x^{-1}(y)$ отображения (1-5) над точкой y состоит из всех отображений $f\colon X \to Y$, у которых $f(x) = y$. Сопоставляя такому отображению f его ограничение на подмножество $X \smallsetminus \{x\}$, мы получаем биекцию $\operatorname{ev}_x^{-1}(y) \xrightarrow{\sim} \operatorname{Hom}(X \smallsetminus \{x\}, Y)$. Таким образом, все слои отображения (1-5) состоят из одинакового числа элементов, равного количеству всех отображений из $(n-1)$-элементного множества $X \smallsetminus \{x\}$ в Y, откуда по формуле (1-4) $|\operatorname{Hom}(X,Y)| = |\operatorname{Hom}(X \smallsetminus \{x\}, Y)| \cdot |Y|$, т. е. при добавлении к X одной точки число отображений $X \to Y$ увеличивается в $|Y|$ раз. □

Замечание 1.1. Имея в виду предл. 1.1, множество $\operatorname{Hom}(X,Y)$, состоящее из всех отображений $X \to Y$, часто обозначают через Y^X. В терминах прим. 1.1 отображение $X \to Y$ представляет собою слово, в котором места расположения букв отвечают точкам множества X, а сами буквы независимо выбираются из алфавита Y. Отображение вычисления ev_x сопоставляет слову его x-ю букву.

Замечание 1.2. В доказательстве предл. 1.1 мы молчаливо предполагали, что оба множества непусты. Если $X = \varnothing$, то для любого множества Y множество $\operatorname{Hom}(\varnothing, Y)$ по определению состоит из единственного элемента — вложения $\varnothing$ в Y в качестве пустого подмножества или, что то же самое, пустого слова в алфавите Y. Отображение вычисления (1-5) в этом случае не определено, но предл. 1.1 остаётся в силе: $1 = |Y|^0$. В частности, множество $\operatorname{Hom}(\varnothing, \varnothing)$ тоже состоит из одного элемента[2] — тождественного автоморфизма $\operatorname{Id}_\varnothing$. Если же $Y = \varnothing$, $X \neq \varnothing$, то $\operatorname{Hom}(X, \varnothing) = \varnothing$, что тоже согласуется с предл. 1.1: $0^{|X|} = 0$ при $|X| > 0$.

Предложение 1.2
Если $|X| = n$, то $|\operatorname{Aut}(X)| = n! \stackrel{\text{def}}{=} n \cdot (n-1) \cdot (n-2) \cdot \,\cdots\, \cdot 1$.

Доказательство. Положим $Y = X$ в доказательстве предл. 1.1 и ограничим отображение вычисления (1-5) на подмножество биекций $\operatorname{Aut}(X) \subset \operatorname{Hom}(X,X)$. Получим отображение

$$\operatorname{ev}_x\colon \operatorname{Aut}(X) \to X\,, \quad f \mapsto f(x)\,.$$

[1] Обозначение «ev» является сокращением слова *evaluation*.
[2] Т. е. 0^0 в этом контексте считается равным 1.

Его слой $\mathrm{ev}_x^{-1}(x')$ над произвольной точкой $x' \in X$ состоит из всех биекций $X \xrightarrow{\sim} X$, переводящих x в x'.

Упражнение 1.7. Постройте взаимно однозначное отображение между биекциями $X \xrightarrow{\sim} X$, переводящими x в x', и биекциями $X \xrightarrow{\sim} X$, оставляющими точку x на месте.

Таким образом, слои $\mathrm{ev}_x^{-1}(x')$ над всеми точками $x' \in X$ непусты и состоят из одного и того же числа элементов, равного количеству автоморфизмов $(n-1)$-элементного множества $X \smallsetminus \{x\}$. По формуле (1-4), $|\operatorname{Aut}(X)| = |\operatorname{Aut}(X \smallsetminus \{x\})| \cdot |X|$, т. е. при добавлении n-той точки к $(n-1)$-элементному множеству количество автоморфизмов увеличивается в n раз. Поэтому $|\operatorname{Aut}(X)| = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 1 = n!$. □

Замечание 1.3. Число $n! = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 1$ называется *n-факториал*. Так как $|\operatorname{Aut}(\varnothing)| = |\{\mathrm{Id}_\varnothing\}| = 1$, мы полагаем $0! \stackrel{\text{def}}{=} 1$.

Замечание 1.4. В терминах прим. 1.1 автоморфизм n-элементного множества X представляет собою n-буквенное слово без повторяющихся букв в алфавите X, т. е. перестановку элементов множества X, и предл. 1.2 утверждает, что имеется ровно $n!$ различных слов, которые можно получить, переставляя буквы в заданном n-буквенном слове без повторяющихся букв.

Пример 1.2 (мультиномиальные коэффициенты)
При раскрытии скобок в выражении $(a_1 + a_2 + \cdots + a_m)^n$ получится сумма одночленов вида $a_1^{k_1} a_2^{k_2} \cdots a_m^{k_m}$, где каждый показатель k_i заключен в пределах $0 \leqslant k_i \leqslant n$, а общая степень $k_1 + k_2 + \cdots + k_m = n$. Коэффициент, возникающий при таком одночлене после приведения подобных слагаемых, называется *мультиномиальным коэффициентом* и обозначается $\binom{n}{k_1 \ldots k_m}$. Таким образом,

$$(a_1 + a_2 + \cdots + a_m)^n = \sum_{\substack{k_1+k_2+\cdots+k_m=n \\ \forall i\ 0 \leqslant k_i \leqslant n}} \binom{n}{k_1 \ldots k_m} \cdot a_1^{k_1} a_2^{k_2} \cdots a_m^{k_m}\,, \tag{1-6}$$

Чтобы явно выразить $\binom{n}{k_1 \ldots k_m}$ через $k_1, \ldots, k_m$, заметим, что раскрытие n скобок

$$(a_1 + a_2 + \cdots + a_m)(a_1 + a_2 + \cdots + a_m) \cdots (a_1 + a_2 + \cdots + a_m)$$

заключается в выборе внутри каждой из скобок какой-нибудь одной буквы и выписывании их слева направо друг за другом в одно n-буквенное слово. Это надо сделать всеми возможными способами и сложить все полученные слова. Подобные слагаемые, вносящие вклад в коэффициент при $a_1^{k_1} a_2^{k_2} \cdots a_m^{k_m}$, суть слова, состоящие ровно из k_1 букв a_1, k_2 букв a_2, … , k_m букв a_m. Количество таких слов легко подсчитать по формуле (1-4). А именно, сделаем на время k_1 букв a_1 попарно разными, снабдив каждую из них дополнительным верхним индексом; аналогично поступим с k_2 буквами a_2, k_3 буквами a_3 и т. д. В результате получится набор из $n = k_1 + k_2 + \cdots + k_m$ попарно различных букв:

$$\underbrace{a_1^{(1)}, a_1^{(2)}, \ldots, a_1^{(k_1)}}_{k_1 \text{ меченых букв } a_1}, \underbrace{a_2^{(1)}, a_2^{(2)}, \ldots, a_2^{(k_2)}}_{k_2 \text{ меченых букв } a_2}, \ \ldots\ \ldots\ \ldots, \underbrace{a_m^{(1)}, a_m^{(2)}, \ldots, a_m^{(k_m)}}_{k_m \text{ меченых букв } a_m}.$$

Обозначим через X множество всех n-буквенных слов, которые можно написать этими n различными буквами, используя каждую букву ровно по одному разу. Как мы уже знаем, $|X| = n!$.

В качестве Y возьмём интересующее нас множество слов из k_1 одинаковых букв a_1, k_2 одинаковых букв a_2, и т. д. и рассмотрим отображение $f : X \to Y$, стирающее верхние индексы у всех букв. Оно эпиморфно, и полный прообраз каждого слова $y \in Y$ состоит из $k_1! \cdot k_2! \cdot \ \cdots \ \cdot k_m!$ слов, которые получаются из y всевозможными расстановками k_1 верхних индексов у букв a_1, k_2 верхних индексов у букв a_2, и т. д. По формуле (1-4)

$$\binom{n}{k_1 \dots k_m} = \frac{n!}{k_1! \cdot k_2! \cdot \ \cdots \ \cdot k_m!} \,. \tag{1-7}$$

Тем самым, разложение (1-6) имеет вид

$$(a_1 + a_2 + \ \cdots \ + a_m)^n = \sum_{\substack{k_1 + \ \cdots \ + k_m = n \\ \forall i \ 0 \leqslant k_i \leqslant n}} \frac{n! \cdot \ a_1^{k_1} a_2^{k_2} \cdots a_m^{k_m}}{k_1! \cdot k_2! \cdot \ \cdots \ \cdot k_m!} \,. \tag{1-8}$$

Упражнение 1.8. Сколько всего слагаемых в правой части формулы (1-8) ?

В частности, при $m = 2$ мы получаем известную формулу для раскрытия бинома с натуральным показателем[1]:

$$(a + b)^n = \sum_{k=0}^{n} \frac{n! \cdot \ a^k b^{n-k}}{k!(n-k)!} \,. \tag{1-9}$$

При $m = 2$ мультиномиальный коэффициент $\binom{n}{k,n-k}$ принято обозначать через $\binom{n}{k}$ или C_n^k и называть k-тым *биномиальным коэффициентом* степени n или *числом сочетаний* из n по k. Он равен

$$\binom{n}{k} = C_n^k = \frac{n \cdot (n-1) \cdot \ \cdots \ \cdot (n-k+1)}{k \cdot (k-1) \cdot \ \cdots \ \cdot 1}$$

(сверху и снизу стоит по k последовательно убывающих сомножителей).

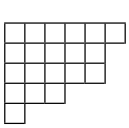
Пример 1.3 (диаграммы Юнга)
Разбиение конечного множества $X = \{1, 2, \dots, n\}$ в объединение непересекающихся подмножеств

$$X = X_1 \sqcup X_2 \sqcup \ \dots \ \sqcup X_k \tag{1-10}$$

можно кодировать следующим образом. Занумеруем подмножества в порядке нестрогого убывания их размера и обозначим количество элементов в i-том подмножестве через $\lambda_i = |X_i|$. Получим невозрастающую последовательность чисел

$$\lambda = (\lambda_1, \dots, \lambda_k), \quad \lambda_1 \geqslant \lambda_2 \geqslant \ \cdots \ \geqslant \lambda_k \,,$$

которая называется *формой* разбиения (1-10). Форму разбиения удобно представлять себе в виде *диаграммы Юнга* — картинки вида

 , (1-11)

[1] Это частный случай *формулы Ньютона*, которую в полной общности мы обсудим чуть позже, когда будем заниматься степенными рядами.

составленной из выровненных по левому краю горизонтальных клетчатых полосок, занумерованных сверху вниз, так что в i-той сверху полоске λ_i клеток. Общее число клеток в диаграмме λ называется её *весом* и обозначается $|\lambda|$, а количество строк называется *длиной* и обозначается $\ell(\lambda)$. Так, диаграмма Юнга (1-11) отвечает разбиению формы $\lambda = (6,5,5,3,1)$, имеет вес $|\lambda| = 20$ и длину $\ell(\lambda) = 5$.

Упражнение 1.9. Подсчитайте количество всех диаграмм Юнга, умещающихся в прямоугольнике размером $k \times n$ клеток (включая пустую диаграмму и сам прямоугольник).

Будем называть *заполнением* диаграммы λ множеством X из $|X| = |\lambda|$ элементов произвольную расстановку этих элементов в клетки диаграммы по одному элементу в каждую клетку. Таким образом, каждая диаграмма λ веса n имеет $n!$ различных заполнений заданным n-элементным множеством X.

Объединяя элементы, стоящие в i-той строке диаграммы в одно подмножество X_i, мы получаем разбиение множества X в дизъюнктное объединение k непересекающихся подмножеств $X_1, \dots, X_k$. Поскольку любое разбиение (1-10) заданной формы λ можно получить таким образом, возникает сюрьективное отображение из множества заполнений диаграммы λ в множество разбиений множества X формы λ. Покажем, что все слои этого отображения состоят из одного и того же числа элементов. Два заполнения приводят к одинаковым разбиениям тогда и только тогда, когда они получаются друг из друга перестановками элементов внутри строк и перестановками строк одинаковой длины между собою как единого целого. Если обозначить через $m_i = m_i(\lambda)$ число строк длины[1] i в диаграмме λ, то перестановок первого типа будет $\prod \lambda_i! = \prod_{i=1}^{n} (i!)^{m_i}$ штук, а второго типа — $\prod_{i=1}^{n} m_i!$ штук. Так как все эти перестановки действуют независимо друг от друга, каждый слой нашего отображения состоит из $\prod_{i=1}^{n} (i!)^{m_i} m_i!$ элементов. Из формулы (1-4) вытекает

Предложение 1.3
Число разбиений n-элементного множества X в дизъюнктное объединение m_1 1-элементных, m_2 2-элементных, … , m_n n-элементных подмножеств равно

$$\frac{n!}{\prod_{i=1}^{n} m_i! \cdot (i!)^{m_i}}. \tag{1-12}$$

□

1.4. Классы эквивалентности. Альтернативный способ разбить заданное множество X в дизъюнктное объединение подмножеств состоит в том, чтобы объявить элементы, входящие в одно подмножество такого разбиения «эквивалентными». Формализуется это так. Назовём *бинарным отношением* на множестве X любое подмножество $R \subset X \times X = \{(x_1, x_2) \mid x_1, x_2 \in X\}$. Принадлежность пары (x_1, x_2) отношению R обычно записывают как $x_1 \underset{R}{\sim} x_2$.

[1] Отметим, что многие $m_i = 0$, поскольку $|\lambda| = n = m_1 + 2m_2 + \cdots + nm_n$

Например, на множестве целых чисел $X = \mathbb{Z}$ имеются бинарные отношения

$$\text{равенство} \qquad x_1 \underset{R}{\sim} x_2 \overset{\text{def}}{\iff} x_1 = x_2 \tag{1-13}$$

$$\text{неравенство} \qquad x_1 \underset{R}{\sim} x_2 \overset{\text{def}}{\iff} x_1 \leqslant x_2 \tag{1-14}$$

$$\text{делимость} \qquad x_1 \underset{R}{\sim} x_2 \overset{\text{def}}{\iff} x_1 | x_2 \tag{1-15}$$

$$\text{сравнимость по модулю } n \qquad x_1 \underset{R}{\sim} x_2 \overset{\text{def}}{\iff} x_1 \equiv x_2 \pmod{n} \tag{1-16}$$

(последнее условие $x_1 \equiv x_2 \pmod{n}$ читается как «x_1 сравнимо с x_2 по модулю n» и по определению означает, что x_1 и x_2 имеют одинаковые остатки от деления на n).

Определение 1.1
Бинарное отношение $\underset{R}{\sim}$ называется *эквивалентностью*, если оно обладает следующими тремя свойствами:

$$\begin{aligned} &\textit{рефлексивность}: && \forall\, x \in X \; x \underset{R}{\sim} x \\ &\textit{транзитивность}: && \forall\, x_1, x_2, x_3 \in X \text{ из } x_1 \underset{R}{\sim} x_2 \text{ и } x_2 \underset{R}{\sim} x_3 \text{ вытекает } x_1 \underset{R}{\sim} x_3 \\ &\textit{симметричность}: && \forall\, x_1, x_2 \in X \; x_1 \underset{R}{\sim} x_2 \iff x_2 \underset{R}{\sim} x_1. \end{aligned}$$

Среди перечисленных выше бинарных отношений на множестве $\mathbb{Z}$ отношения (1-13) и (1-16) являются эквивалентностями, а (1-14) и (1-15) не являются (они не симметричны).

Если множество X разбито в объединение непересекающихся подмножеств, то отношение $x_1 \sim x_2$, означающее, что x_1 и x_2 лежат в одном и том же подмножестве этого разбиения, очевидно, является эквивалентностью.

Наоборот, пусть на множестве X задано какое-нибудь отношение эквивалентности R. Рассмотрим для каждого $x \in X$ подмножество в X, состоящее из всех элементов, эквивалентных x. Оно называется *классом эквивалентности* элемента x и обозначается

$$[x]_R = \{z \in X \mid x \underset{R}{\sim} z\} = \{z \in X \mid z \underset{R}{\sim} x\}$$

(второе равенство выполняется благодаря симметричности отношения R). Два класса $[x]_R$ и $[y]_R$ либо вообще не пересекаются, либо полностью совпадают. В самом деле, если существует элемент z, эквивалентный и x и y, то в силу симметричности и транзитивности отношения $\underset{R}{\sim}$ элементы x и y будут эквивалентны между собой, а значит, любой элемент, эквивалентный x, будет эквивалентен также и y, и наоборот. Таким образом, множество X распадается в дизъюнктное объединение различных классов эквивалентности.

Множество классов эквивалентности по отношению $R \subset X \times X$ обозначается X/R и называется *фактором* множества X по отношению R. Сюръективное отображение

$$f : X \twoheadrightarrow X/R\,, \quad x \mapsto [x]_R\,, \tag{1-17}$$

сопоставляющее каждому элементу $x \in X$ его класс эквивалентности $[x]_R \in X/R$, называется *отображением факторизации*. Слои этого отображения суть классы эквивалентных элементов. Наоборот, любое сюръективное отображение $f : X \twoheadrightarrow Y$ является отображением факторизации по отношению эквивалентности $x_1 \sim x_2$, означающему, что $f(x_1) = f(x_2)$.

Пример 1.4 (классы вычетов)
Фиксируем ненулевое целое число $n \in \mathbb{Z}$. Фактор множества целых чисел $\mathbb{Z}$ по отношению сравнимости по модулю n из (1-16) обозначается $\mathbb{Z}/(n)$. Мы будем записывать его элементы символами $[z]_n$, где $z \in \mathbb{Z}$, и опускать индекс n, когда понятно чему он равен. Класс эквивалентности

$$[z]_n \stackrel{\text{def}}{=} \{x \in \mathbb{Z} \mid (z-x) \vdots n\} \tag{1-18}$$

называется *классом вычетов по модулю* n. Отображение факторизации

$$\mathbb{Z} \twoheadrightarrow \mathbb{Z}/(n), \quad z \mapsto [z]_n$$

называется *приведением по модулю* n. Множество $\mathbb{Z}/(n)$ состоит из n различных классов

$$[0]_n, \ [1]_n, \ \dots, \ [n-1]_n.$$

При желании их можно воспринимать как остатки от деления на n, но в практических вычислениях удобнее работать с ними именно как с *подмножествами* в $\mathbb{Z}$, поскольку возможность по-разному записывать один и тот же класс часто упрощает вычисления. Например, остаток от деления 12^{100} на 13 можно искать как

$$\left[12^{100}\right]_{13} = [12]_{13}^{100} = [-1]_{13}^{100} = \left[(-1)^{100}\right]_{13} = [1]_{13}. \tag{1-19}$$

Упражнение 1.10. Докажите правомочность этого вычисления: проверьте, что классы вычетов $[x+y]_n$ и $[xy]_n$ не зависят от выбора чисел $x \in [x]_n$ и $y \in [y]_n$, т. е. правила

$$[x]_n + [y]_n \stackrel{\text{def}}{=} [x+y]_n \tag{1-20}$$

$$[x]_n \cdot [y]_n \stackrel{\text{def}}{=} [xy]_n \tag{1-21}$$

корректно определяют на множестве $\mathbb{Z}/(n)$ операции сложения и умножения[1].

1.4.1. Неявное задание эквивалентности. Для любого семейства отношений эквивалентности $R_\nu \subset X \times X$ пересечение $\bigcap_\nu R_\nu \subset X \times X$ также является отношением эквивалентности. В самом деле, если каждое из множеств $R_\nu \subset X \times X$ содержит диагональ

$$\Delta = \{(x,x) \mid x \in X\} \subset X \times X,$$

переходит в себя при симметрии $(x,y) \leftrightarrows (y,x)$ и вместе с каждой парой точек вида (x,y), (y,z) содержит также и точку (x,z), то этими свойствами обладает и пересечение $\bigcap_\nu R_\nu$ всех этих множеств. Поэтому для любого подмножества $R \subset X \times X$ существует *наименьшее по включению* отношение эквивалентности $\overline{R}$, содержащее R, а именно, пересечение всех содержащих R отношений эквивалентности. Отношение $\overline{R}$ называется эквивалентностью, *порождённой* отношением R.

Упражнение 1.11. Проверьте, что $(x,y) \in \overline{R}$ если и только если в X существует такая конечная последовательность точек $x = z_0, z_1, z_2, \dots, z_n = y$, что $(x_{i-1}, x_i) \in R$ или $(x_i, x_{i-1}) \in R$ при каждом $i = 1, 2, \dots, n$.

К сожалению, по данному подмножеству $R \subset X \times X$ не всегда легко судить о том, как устроена порождённая им эквивалентность $\overline{R}$. Даже выяснить, не окажутся ли в результате все точки эквивалентными друг другу может быть не просто.

[1] Именно такое умножение $[12]^{100} = \underbrace{[12] \cdot [12] \cdot \dots \cdot [12]}_{100} = \left[12^{100}\right]$ было использовано в (1-19).

Пример 1.5 (дроби)
Множество рациональных чисел $\mathbb{Q}$ обычно определяют как множество дробей a/b с $a, b \in \mathbb{Z}$ и $b \neq 0$. При этом под *дробью* понимается класс эквивалентности упорядоченных пар (a, b), где $a \in \mathbb{Z}$, $b \in \mathbb{Z} \smallsetminus 0$, для минимальной эквивалентности, содержащей все отождествления

$$(a, b) \sim (ac, bc) \quad \text{с произвольными } c \in \mathbb{Z} \smallsetminus \{0\}\,. \tag{1-22}$$

Отношения (1-22) выражают собою равенства дробей $a/b = (ac)/(bc)$, но сами по себе не образуют эквивалентности. Например, при $a_1 b_2 = a_2 b_1$ в двухшаговой цепочке отождествлений $(a_1, b_1) \sim (a_1 b_2, b_1 b_2) = (a_2 b_1, b_1 b_2) \sim (a_2, b_2)$ самый левый и самый правый элементы могут не отождествляться напрямую по правилу (1-22), как, например, $3/6$ и $5/10$. Поэтому эквивалентность, порождённая отождествлениями (1-22), обязана содержать все отождествления

$$(a_1, b_1) \sim (a_2, b_2) \quad \text{при} \quad a_1 b_2 = a_2 b_1\,. \tag{1-23}$$

Оказывается, что к этим отношениям больше уже ничего добавлять не надо.

Упражнение 1.12. Проверьте, что набор отношений (1-23) рефлексивен, симметричен и транзитивен.

Тем самым, он является минимальным отношением эквивалентности, содержащим все отождествления (1-22). Отметим, что если в отношениях (1-22) разрешить нулевые c, то все пары (a, b) окажутся эквивалентны паре $(0, 0)$.

1.5. Композиции отображений. Отображение $X \to Z$, получающееся в результате последовательного выполнения двух отображений $f\colon X \to Y$ и $g\colon Y \to Z$ называется *композицией* отображений g и f и обозначается $g \circ f$ или просто gf. Таким образом, композиция gf определена если и только если образ f содержится в множестве, на котором определено отображение g, и $gf\colon X \to Z\,,\ x \mapsto g(f(x))$.

Хотя композицию и принято записывать точно так же, как умножение чисел, единственным общим свойством этих операций является их *ассоциативность* или *сочетательный закон*: композиция трёх последовательных отображений

$$X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} T\,,$$

как и произведение трёх чисел, не зависит от того, в каком порядке перемножаются последовательные пары элементов, т. е. $(hg)f = h(gf)$, если хотя бы одна из двух частей этого равенства определена. Действительно, в этом случае вторая часть тоже определена, и обе части действуют на каждую точку $x \in X$ по правилу $x \mapsto h(g(f(x)))$.

В остальном алгебраические свойства композиции весьма далеки от привычных свойств умножения чисел. Если композиция fg определена, то противоположная композиция gf часто бывает не определена. Даже если $f, g\colon X \to X$ являются эндоморфизмами одного и того же множества X, так что обе композиции fg и gf определены, равенство $fg = gf$ может не выполняться.

Упражнение 1.13. Рассмотрим на плоскости пару различных прямых ℓ_1, ℓ_2, пересекающихся в точке O, и обозначим через σ_1 и σ_2 осевые симметрии относительно этих прямых. Явно опишите движения плоскости, задаваемые композициями $\sigma_1\sigma_2$ и $\sigma_2\sigma_1$. При каком условии на прямые выполняется равенство $\sigma_1\sigma_2 = \sigma_2\sigma_1$?

Общие множители тоже бывает нельзя сокращать, т. е. ни равенство $fg = fh$, ни равенство $gf = hf$, вообще говоря, не влекут равенства $g = h$.

Пример 1.6 (эндоморфизмы двухэлементного множества)
Двухэлементное множество $X = \{1, 2\}$ имеет ровно четыре эндоморфизма. Если кодировать отображение $f : X \to X$ двубуквенным словом $(f(1), f(2))$, как в прим. 1.1, то эти четыре эндоморфизма запишутся словами $(1,1)$, $(1,2) = \mathrm{Id}_X$, $(2,1)$ и $(2,2)$. Все композиции между ними определены, и таблица композиций gf имеет вид:

$g \diagdown f$	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
$(1,1)$	$(1,1)$	$(1,1)$	$(1,1)$	$(1,1)$
$(1,2)$	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
$(2,1)$	$(2,2)$	$(2,1)$	$(1,2)$	$(1,1)$
$(2,2)$	$(2,2)$	$(2,2)$	$(2,2)$	$(2,2)$

(1-24)

Обратите внимание на то, что $(2,2) \circ (1,1) \neq (1,1) \circ (2,2)$ и что $(1,1) \circ (1,2) = (1,1) \circ (2,1)$, хотя $(1,2) \neq (2,1)$, и $(1,1) \circ (2,2) = (2,1) \circ (2,2)$, хотя $(1,1) \neq (2,1)$.

Лемма 1.1 (левые обратные отображения)
Если $X \neq \varnothing$, то следующие условия на отображение $f : X \to Y$ эквивалентны:

1) f инъективно
2) существует такое отображение $g : Y \to X$, что $gf = \mathrm{Id}_X$
3) для любых отображений $g_1, g_2 : Z \to X$ из равенства $fg_1 = fg_2$ вытекает равенство $g_1 = g_2$.

Доказательство. Импликация (1) ⇒ (2): для точек $y = f(x) \in \operatorname{im} f$ положим $g(y) = x$, а в точках $y \notin \operatorname{im} f$ зададим g как угодно[1]. Импликация (2) ⇒ (3): если $fg_1 = fg_2$, то умножая обе части слева на любое такое отображение $g : Y \to X$, что $gf = \mathrm{Id}_X$, получаем $g_1 = g_2$. Импликация (3) ⇒ (1) доказывается от противного: если $f(x_1) = f(x_2)$ для каких-то $x_1 \neq x_2$, то пусть $g_1 = \mathrm{Id}_X$, а $g_2 : X \rightrightarrows X$ переставляет между собою точки x_1 и x_2, а все остальные точки оставляет на месте. Тогда $g_1 \neq g_2$, но $fg_1 = fg_2$. □

Определение 1.2
Отображение $f : X \to Y$, удовлетворяющее лем. 1.1, называется *обратимым слева*, и любое отображение $g : Y \to X$, такое что $gf = \mathrm{Id}_X$, называется *левым обратным* к f.

Упражнение 1.14. В условиях лем. 1.1 убедитесь, что вложение f тогда и только тогда имеет несколько различных левых обратных, когда оно не сюръективно.

1.5.1. Правое обратное отображение и аксиома выбора. Чувство гармонии заставляет думать, что у лем. 1.1 должна быть симметричная «правая» версия. А именно, хочется ожидать, что следующие три свойства отображения $f : X \to Y$ эквивалентны друг другу:

1) f сюръективно
2) существует такое отображение $g : Y \to X$, что $fg = \mathrm{Id}_Y$
3) для любых отображений $g_1, g_2 : Y \to Z$ из равенства $g_1 f = g_2 f$ вытекает равенство $g_1 = g_2$.

[1] Например, отобразим их все в одну и ту же произвольно выбранную точку $x \in X$.

Отображение f, удовлетворяющее свойству (2), называются *обратимым справа*, и всякое отображение $g : Y \to X$, такое что $fg = \mathrm{Id}_Y$, называется *правым обратным* к f или *сечением* эпиморфизма f. Второе название связано с тем, что если отображение f сюръективно, то отображение g, удовлетворяющее свойству (2), переводит каждую точку $y \in Y$ в точку $g(y) \in f^{-1}(y)$, лежащую в слое отображения f над точкой y. В строгой теории множеств, углубления в которую мы пытаемся избежать, импликация (1) ⇒ (2) постулируется в качестве одной из аксиом. Эта аксиома называется *аксиомой выбора* и утверждает, что в каждом слое любого сюръективного отображения можно выбрать по элементу[1].

Итак, импликация (1) ⇒ (2) является частью строго определения понятия «множество». Доказательство импликации (2) ⇒ (3) полностью симметрично доказательству аналогичной импликации из лем. 1.1: применяя отображения, стоящие в обеих частях равенства $g_1 f = g_2 f$, вслед за таким отображением $g : Y \to X$, что $fg = \mathrm{Id}_Y$, получаем равенство $g_1 = g_2$. Импликация (3) ⇒ (1), как и в лем. 1.1, доказывается от противного: если $y \notin \operatorname{im} f$, то свойство (3) не выполняется для отображения $g_1 = \mathrm{Id}_Y$ и любого отображения $g_2 : Y \to Y$, переводящего точку y в какую-нибудь точку из $\operatorname{im} f$ и оставляющего на месте все остальные точки. Таким образом, перечисленные выше свойства (1) – (3) действительно эквивалентны друг другу.

1.5.2. Обратимые отображения. Если отображение $g : X \xrightarrow{\sim} Y$ биективно, то прообраз $g^{-1}(y) \subset X$ каждой точки $y \in Y$ состоит ровно из одной точки. В этом случае правило $y \mapsto g^{-1}(y)$ определяет отображение $g^{-1} : Y \to X$, которое является одновременно и левым, и правым обратным к g в смысле опр. 1.2 и n° 1.5.1 :, т. е.

$$g \circ g^{-1} = \mathrm{Id}_Y \qquad \text{и} \qquad g^{-1} \circ g = \mathrm{Id}_X \tag{1-25}$$

Отображение g^{-1} называется *двусторонним обратным* или просто *обратным* к g.

Предложение 1.4
Следующие условия на отображение $g : X \to Y$ эквивалентны друг другу:

1) g взаимно однозначно

2) у g имеется двустороннее обратное отображение $g' : Y \to X$ со свойствами $g \circ g' = \mathrm{Id}_Y$ и $g' \circ g = \mathrm{Id}_X$

3) g обладает левым и правым обратными отображениями[2].

При выполнении этих условий все левые и правые обратные к g отображения равны друг другу и отображению g^{-1}, описанному перед формулировкой предложения.

Доказательство. Импликация (1) ⇒ (2) уже была установлена. Импликация (2) ⇒ (3) очевидна. Докажем, что (3) ⇒ (2). Если у отображения $g : X \to Y$ есть левое обратное $f : Y \to X$ и правое обратное $h : Y \to X$, то $f = f \circ \mathrm{Id}_Y = f \circ (g \circ h) = (f \circ g) \circ h = \mathrm{Id}_X \circ h = h$ и условие (2) выполняется для $g' = f = h$. Остаётся установить импликацию (2) ⇒ (1) и доказать равенство $g' = g^{-1}$. Поскольку $g(g'(y)) = y$ для любого $y \in Y$, прообраз $g^{-1}(y)$ каждой точки $y \in Y$ содержит

[1] Иными словами, если задано множество попарно непересекающихся множеств, то в каждом из них можно выбрать по элементу.

[2] Обратите внимание, что совпадения левого обратного отображения с правым обратным отображением не требуется.

точку $g'(y)$. С другой стороны, для любого $x \in g^{-1}(y)$ выполнено равенство $x = \mathrm{Id}_X(x) = g'(g(x)) = g'(y)$. Поэтому $f^{-1}(y)$ состоит из единственной точки $g'(y)$, т. е. g — биекция, и $g' = g^{-1}$. □

1.6. Группы преобразований. Непустой набор G взаимно однозначных отображений множества X в себя называется *группой преобразований* множества X, если вместе с каждым отображением $g \in G$ в G лежит и обратное к нему отображение g^{-1}, а вместе с каждыми двумя отображениями $f, g \in G$ в G лежит и их композиция fg. Эти условия гарантируют, что тождественное преобразование Id_X тоже лежит в G, поскольку $\mathrm{Id}_X = g^{-1}g$ для любого $g \in G$. Если группа преобразований G конечна, число элементов в ней обозначается $|G|$ и называется *порядком* группы G. Если подмножество $H \subset G$ тоже является группой, то H называются *подгруппой* группы G.

Пример 1.7 (группы перестановок)
Множество $\mathrm{Aut}(X)$ всех взаимно однозначных отображений $X \to X$ является группой. Эта группа называется *симметрической группой* или *группой перестановок* множества X. Все прочие группы преобразований множества X являются подгруппами этой группы. Группа перестановок n-элементного множества $\{1, 2, \ldots, n\}$ обозначается S_n и называется n-той *симметрической группой*. Согласно предл. 1.2 на стр. 7 порядок $|S_n| = n!$. Перестановки

$$\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$$

принято записывать строчками $\sigma = (\sigma_1, \ldots, \sigma_n)$ их значений $\sigma_i \stackrel{\text{def}}{=} \sigma(i)$, как в прим. 1.1 на стр. 6. Например, перестановки $\sigma = (3, 4, 2, 1)$ и $\tau = (2, 3, 4, 1)$ представляют собою отображения

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 4 & 2 & 1 \end{array} \quad \text{и} \quad \begin{array}{cccc} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 3 & 4 & 1 \end{array}$$

а их композиции записываются как $\sigma\tau = (4, 2, 1, 3)$ и $\tau\sigma = (4, 1, 3, 2)$.

Упражнение 1.15. Составьте таблицу умножения шести элементов группы S_3, аналогичную таблице (1-24) на стр. 14.

Пример 1.8 (абелевы группы)
Группа G, в которой любые два элемента $f, g \in G$ перестановочны, т. е. удовлетворяют соотношению $fg = gf$, называется *коммутативной* или *абелевой*. Примерами абелевых групп являются группы параллельных переносов плоскости или пространства, а также группа SO_2 поворотов плоскости вокруг фиксированной точки. Для каждого натурального $n \geqslant 2$ повороты на углы, кратные $2\pi/n$, образуют в группе SO_2 конечную подгруппу. Она называется *циклической группой порядка* n.

1.7. Частично упорядоченные множества. Бинарное отношение[1] $x \leqslant y$ на множестве Z называется *частичным порядком*, если оно рефлексивно и транзитивно[2], но в отличие от эквивалентности не симметрично, а *кососимметрично*, т. е. из $x \leqslant y$ и $y \leqslant x$ вытекает равенство $x = y$. Если на множестве задан частичный порядок, мы пишем $x < y$, если $x \leqslant y$ и $x \neq y$. Частичный

[1] См. n° 1.4 на стр. 10.
[2] Так же, как и отношение эквивалентности, ср. с опр. 1.1 на стр. 11.

порядок на множестве Z называется *тотальным* (а также *линейным* или просто *порядком*), если любые два элемента сравнимы, т. е. для всех $x, y \in Z$ выполняется одно из трёх альтернативных условий: или $x < y$, или $x = y$, или $y < x$. Например, обычное неравенство между числами является линейным порядком на множестве натуральных чисел $\mathbb{N}$, тогда как отношение делимости $n \mid m$, означающее, что n делит m, задаёт на $\mathbb{N}$ частичный порядок, который не является линейным. Другим важным примером частичного, но не линейного порядка является отношение включения $X \subseteq Y$ на множестве $\mathcal{S}(M)$ всех подмножеств заданного множества M.

Упражнение 1.16 (предпорядок). *Предпорядком* на множестве Z называется любое рефлексивное транзитивное бинарное отношение $x \precsim y$. Убедитесь, что для каждого предпорядка бинарное отношение $x \sim y$, означающее, что одновременно $x \precsim y$ и $y \precsim x$, является отношением эквивалентности и что на факторе $Z/\sim$ бинарное отношение $[x] \leqslant [y]$, означающее, что $x \precsim y$, корректно определено[1] и является частичным порядком. Продумайте, как всё это работает для отношения делимости $n \mid m$ на множестве целых чисел $\mathbb{Z}$.

Множество P с зафиксированным на нём частичным порядком называется *частично упорядоченным множеством*, сокращённо — чумом. Если порядок на P тотальный, мы будем говорить, что чум P *линейно упорядочен*. Всякое подмножество X чума P также является чумом по отношению к частичному порядку, имеющемуся на P. Если этот индуцированный с P порядок на X оказывается линейным, подмножество $X \subset P$ называют *цепью* в чуме P. Элементы x, y чума P называются *сравнимыми*, если $x \leqslant y$ или $y \leqslant x$. Если же ни одно из этих условий не выполняется, то x и y называются *несравнимыми*. Несравнимые элементы автоматически различны. Частичный порядок линеен тогда и только тогда, когда любые два элемента сравнимы.

Отображение $f : M \to N$ между чумами M, N называется *сохраняющим порядок*[2] или *морфизмом* чумов, если для всех $x, y \in M$ соотношение $x \leqslant y$ влечёт соотношение $f(x) \leqslant f(y)$. Два чума M, N называются *изоморфными*, если имеется сохраняющая порядок биекция $M \xrightarrow{\sim} N$. В таком случае мы пишем $M \simeq N$. Отображение f называется *строго возрастающим*, если для всех $x, y \in M$ соотношение $x < y$ влечёт соотношение $f(x) < f(y)$. Всякое сохраняющее порядок вложение является строго возрастающим. Обратное справедливо для возрастающих отображений из линейного упорядоченного множества, однако неверно в общем случае.

Элемент y чума P называется *верхней гранью* подмножества $X \subset P$, если $x \leqslant y$ для всех $x \in X$. Если при этом $y \notin X$, то верхняя грань y называется *внешней*. В таком случае для всех $x \in X$ выполнено строгое неравенство $x < y$.

Элемент $m^* \in X$ называется *максимальным* в подмножестве $X \subset P$, если неравенство $m^* \leqslant x$ для $x \in X$ выполняется только при $x = m^*$. Заметьте, что максимальный элемент не обязан быть сравним со всеми элементами $x \in X$ и, тем самым, может не являться верхней гранью для X. Частично упорядоченное множество может иметь несколько различных максимальных элементов или не иметь их вовсе, как, например, чум $\mathbb{Z}$ с любым из двух указанных выше порядков. Линейно упорядоченный чум имеет не более одного максимального элемента, и если такой элемент существует, то он является верхней гранью.

Симметричным образом, элемент $m_* \in X$ называется *минимальным*, если неравенство $m_* \leqslant x$ выполняется только для $x = m_*$. Аналогично определяются и нижние грани, и всё сказанное выше о максимальных элементах и верхних гранях в равной степени относится и к минимальным элементам и нижним граням.

[1] Т. е. выполнение или невыполнение условия $x \precsim y$ не зависит от выбора представителей x и y в классах $[x]$ и $[y]$.

[2] А также *неубывающим* или *нестрого возрастающим*.

1.8. Вполне упорядоченные множества. Линейно упорядоченное множество W называется *вполне упорядоченным*, если каждое непустое подмножество $S \subset W$ содержит такой элемент $s_* \in S$, что $s_* \leqslant s$ для всех $s \in S$. Этот элемент автоматически единствен и называется *начальным элементом* подмножества S. Например, множество натуральных чисел $\mathbb{N}$ со стандартным отношением неравенства между числами вполне упорядочено, как и любое дизъюнктное объединение вида $\mathbb{N} \sqcup \mathbb{N} \sqcup \mathbb{N} \sqcup \ldots$, в котором все элементы каждой копии множества $\mathbb{N}$ полагаются строго большими всех элементов всех предыдущих копий. Пустое множество тоже вполне упорядочено. Напротив, множество $\mathbb{Q}$ со стандартным отношением неравенства между числами не является вполне упорядоченным.

Вполне упорядоченные множества замечательны тем, что их элементы можно рекурсивно перебрать точно так же, как и элементы множества $\mathbb{N}$. А именно, пусть некоторое зависящее от элемента w вполне упорядоченного множества W утверждение $\Phi(w)$ истинно для начального элемента w_* множества W, и пусть для каждого $w \in W$ истинность утверждения $\Phi(x)$ при всех $x < w$ влечёт за собою истинность утверждения $\Phi(w)$. Тогда $\Phi(w)$ истинно для всех $w \in W$.

Упражнение 1.17. Убедитесь в этом.

Такой способ доказательства утверждения $\Phi(w)$ для всех $w \in W$ называется *трансфинитной индукцией*. Используемые для индуктивного перехода специальные подмножества вида

$$[w) \stackrel{\text{def}}{=} \{x \in W \mid x < w\},$$

состоящие из всех элементов, предшествующих данному элементу w, называются *начальными интервалами* частично упорядоченного множества W. Элемент $w \in W$ называется *точной верхней гранью* начального интервала $[w) \subset W$ и однозначно восстанавливается по интервалу $[w)$ как начальный элемент множества $W \smallsetminus [w)$. Отметим, что начальный элемент $w_* \in W$ является точной верхней гранью пустого начального интервала $[w_*) = \varnothing$.

Упражнение 1.18. Покажите, что собственное подмножество $I \subsetneq W$ тогда и только тогда является начальным интервалом вполне упорядоченного множества W, когда $[x) \subset I$ для каждого $x \in I$, и в этом случае точная верхняя грань интервала I однозначно восстанавливается по I как начальный элемент дополнения $W \smallsetminus I$.

Между вполне упорядоченными множествами имеется отношение порядка $U \leqslant W$, означающее, что U изоморфно с сохранением порядка некоторому начальному интервалу $[w) \subset W$. Если при этом U и W не изоморфны, мы пишем $U < W$. Хорошим упражнением на трансфинитную индукцию является

Упражнение 1.19. Убедитесь, что для любой пары вполне упорядоченных множеств U, W выполнено ровно одно из соотношений: или $U < W$, или $U \simeq W$, или $W < U$.

Классы изоморфных вполне упорядоченных множеств называют *ординалами*[1]. Множество $\mathbb{N}$ можно воспринимать как множество всех конечных ординалов. Все остальные ординалы, включая множество $\mathbb{N}$ со стандартным порядком, называются *трансфинитными*.

1.9. Лемма Цорна. Рассмотрим произвольное частично упорядоченное множество P и обозначим через $\mathcal{W}(P)$ множество всех подмножеств $W \subset P$, которые вполне упорядочены имеющимся на P отношением $x \leqslant y$. Множество $\mathcal{W}(P)$ непусто и содержит пустое подмножество $\varnothing \subset P$, а также все конечные цепи[2] $C \subset P$ и, в частности, все элементы множества P.

[1] Или *кардиналами*.

[2] Т. е. конечные линейно упорядоченные подмножества.

Лемма 1.2
Не существует такого отображения $\varrho : \mathcal{W}(P) \to P$, что $\varrho(W) > w$ для всех $W \in \mathcal{W}(P)$ и $w \in W$.

Доказательство. Пусть такое отображение ϱ существует. Назовём вполне упорядоченное подмножество $W \subset P$ рекурсивным, если $\varrho([w)) = w$ для всех $w \in W$. Например, подмножество

$$\Big\{\varrho(\varnothing),\ \varrho\big(\{\varrho(\varnothing)\}\big),\ \varrho\Big(\big\{\varrho(\varnothing),\ \varrho(\{\varrho(\varnothing)\})\big\}\Big)\Big\}$$

рекурсивно и может неограниченно расширяться вправо. Любые два различных рекурсивных вполне упорядоченных подмножества с общим начальным элементом таковы, что одно из них является начальным интервалом другого.

Упражнение 1.20. Докажите это.

Обозначим через $U \subset P$ объединение всех рекурсивных вполне упорядоченных подмножеств в P с начальным элементом $\varrho(\varnothing)$.

Упражнение 1.21. Убедитесь, что подмножество $U \subset P$ вполне упорядочено и рекурсивно.

Поскольку элемент $\varrho(U)$ строго больше всех элементов из U, он не лежит в U. С другой стороны, множество $W = U \cup \{\varrho(U)\}$ вполне упорядочено, рекурсивно, и его начальным элементом является $\varrho(\varnothing)$. Следовательно, $W \subset U$, откуда $\varrho(U) \in U$. Противоречие. □

Предложение 1.5
Если каждое вполне упорядоченное подмножество чума P имеет верхнюю грань[1], то в P есть максимальный элемент[2] (возможно не единственный).

Доказательство. Если максимального элемента нет, то для любого $p \in P$ имеется такой элемент $p' \in P$, что $p < p'$. Тогда для каждого вполне упорядоченного подмножества $W \subset P$ найдётся такой элемент $w^* \in P$, что $w < w^*$ для всех $w \in W$. Сопоставляя каждому $W \in \mathcal{W}$ один[3] из таких элементов w^*, мы получаем отображение $\varrho : \mathcal{W} \to P$, которого не может быть по лем. 1.2. □

Определение 1.3 (полные чумы)
Частично упорядоченное множество называется *полным*, если каждая его цепь имеет верхнюю грань.

Следствие 1.1 (лемма Цорна)
В каждом полном чуме есть максимальный элемент (возможно не единственный). □

Упражнение 1.22 (лемма Бурбаки – Витта о неподвижной точке). Пусть отображение из полного чума в себя $f : P \to P$ таково, что $f(x) \geqslant x$ для всех $x \in P$. Покажите, что существует такое $p \in P$, что $f(p) = p$.

Упражнение 1.23 (теорема Цермелло). Докажите, что каждое множество можно вполне упорядочить.

Упражнение 1.24 (теорема Хаусдорфа о максимальной цепи). Докажите, что в любом чуме каждая цепь содержится в некоторой максимальной по включению цепи.

[1] Т. е. для любого вполне упорядоченного $W \subset P$ найдётся такой $p \in P$, что $w \leqslant p$ для всех $w \in W$.

[2] Т. е. такой $p^* \in P$, что неравенство $p^* \leqslant x$ выполняется в P только для $x = p^*$, см. последние два абзаца перед n° 1.8 на стр. 18.

[3] Для этого придётся воспользоваться аксиомой выбора из n° 1.5.1 на стр. 14.

§2. Коммутативные кольца и поля

2.1. Определения и примеры. Говоря вольно, поле представляет собою числовую область, где определены четыре стандартные арифметических операции: сложение, вычитание, умножение и деление, которые обладают теми же свойствами, что и соответствующие действия над рациональными числами. Точный перечень этих свойств идёт ниже.

Определение 2.1
Множество $\mathbb{F}$ с двумя операциями $\mathbb{F}\times\mathbb{F}\to\mathbb{F}$: *сложением* $(a,b)\mapsto a+b$ и *умножением* $(a,b)\mapsto ab$ называется *полем*, если выполняются следующие три набора аксиом:

свойства сложения

коммутативность: $a+b=b+a \quad \forall\, a,b\in\mathbb{F}$ (2-1)

ассоциативность: $a+(b+c)=(a+b)+c \quad \forall\, a,b,c\in\mathbb{F}$ (2-2)

наличие нуля: $\exists\, 0\in\mathbb{F}:\quad a+0=a \quad \forall\, a\in\mathbb{F}$ (2-3)

наличие противоположных: $\forall\, a\in\mathbb{F} \quad \exists\,(-a)\in\mathbb{F}:\quad a+(-a)=0$ (2-4)

свойства умножения

коммутативность: $ab=ba \quad \forall\, a,b\in\mathbb{F}$ (2-5)

ассоциативность: $a\,(bc)=(ab)\,c \quad \forall\, a,b,c\in\mathbb{F}$ (2-6)

наличие единицы: $\exists\, 1\in\mathbb{F}:\quad 1\,a=a \quad \forall\, a\in\mathbb{F}$ (2-7)

наличие обратных: $\forall\, a\in\mathbb{F}\smallsetminus 0 \quad \exists\, a^{-1}\in\mathbb{F}:\quad aa^{-1}=1$ (2-8)

свойства, связывающие сложение с умножением

дистрибутивность: $a(b+c)=ab+ac \quad \forall\, a,b,c\in\mathbb{F}$ (2-9)

нетривиальность: $0\neq 1$ (2-10)

Пример 2.1 (поле из двух элементов)
Простейший объект, удовлетворяющий всем аксиомам из опр. 2.1 — это поле $\mathbb{F}_2$, состоящее только из двух элементов 0 и 1, таких что $0+1=1\cdot 1=1$, а все остальные суммы и произведения равны нулю.

Упражнение 2.1. Проверьте, что $\mathbb{F}_2$ действительно является полем.

Элементы этого поля можно воспринимать как классы вычетов по модулю 2, а операции сложения и умножения — как операции сложения и умножения классов вычетов, определённые формулами (1-20) – (1-21) на стр. 12. С другой стороны, элементы поля $\mathbb{F}_2$ могут интерпретироваться как «ложь» $=0$ и «истина» $=1$, сложение — как логическое «исключающее или»[1], а умножение — как логическое «и»[2]. При такой интерпретации алгебраические вычисления в поле $\mathbb{F}_2$ превращаются в логические манипуляции с высказываниями.

Упражнение 2.2. Напишите многочлен от x с коэффициентами из поля $\mathbb{F}_2$, равный «не x», а

[1] Т. е. высказывание $A+B$ истинно тогда и только тогда, когда истинно *ровно одно* из высказываний A, B: $0+1=1+0=1$, но $0+0=1+1=0$.

[2] Т. е. высказывание $A\cdot B$ истинно если и только если истинны *оба* высказывания A и B : $1\cdot 1=1$, но $0\cdot 1=1\cdot 0=0\cdot 0=0$.

также многочлен от x и y, равный «x или[1] y».

Пример 2.2 (рациональные числа)
Напомним, что поле рациональных чисел $\mathbb{Q}$ можно определить как множество дробей a/b, где под «дробью» понимается класс эквивалентности упорядоченной пары (a, b) с $a, b \in \mathbb{Z}$ и $b \neq 0$ по отношению $(a_1, b_1) \sim (a_2, b_2)$ при $a_1 b_2 = a_2 b_1$, которое является минимальным отношением эквивалентности, содержащим все отождествления

$$\frac{a}{b} = \frac{ac}{bc} \quad \forall\, c \neq 0$$

(см. n° 1.4.1). Сложение и умножение дробей определяется формулами

$$\frac{a}{b} + \frac{c}{d} \stackrel{\text{def}}{=} \frac{ad + bc}{bd}\,, \quad \frac{a}{b} \cdot \frac{c}{d} \stackrel{\text{def}}{=} \frac{ac}{bd}\,. \tag{2-11}$$

Упражнение 2.3. Проверьте, что эти операции определены корректно (результат не зависит от выбора представителей в классах) и удовлетворяют аксиомам поля.

Пример 2.3 (вещественные числа)
Множество вещественных чисел $\mathbb{R}$ определяется в курсе анализа несколькими различными способами: как множество классов эквивалентности десятичных[2] дробей, как множество дедекиндовых сечений упорядоченного множества $\mathbb{Q}$, или как множество классов эквивалентности рациональных последовательностей Коши. Мы полагаем, что читатель знаком с этими определениями и понимает, как они связаны друг с другом. Какое бы описание множества $\mathbb{R}$ ни использовалось, задание на нём сложения и умножения и проверка аксиом из опр. 2.1 требуют некоторой умственной работы, традиционно проделываемой в курсе анализа.

2.1.1. Коммутативные кольца. Множество K с операциями сложения и умножения называется *коммутативным кольцом с единицей*, если эти операции обладают всеми свойствами из опр. 2.1 на стр. 20 за исключением свойства (2-8) существования мультипликативно обратного элемента.

Если, кроме существования обратного, из списка аксиом поля исключаются требование существования единицы (2-7) и условие $0 \neq 1$, то множество K с двумя операциями, удовлетворяющими оставшимся аксиомам, называется просто *коммутативным кольцом*.

Примерами отличных от полей колец с единицами являются кольцо целых чисел $\mathbb{Z}$ и кольцо многочленов с коэффициентами в произвольном коммутативном кольце с единицей. Примеры коммутативных колец без единицы доставляют чётные целые числа, многочлены с чётными целыми коэффициентами, многочлены без свободного члена с коэффициентами в любом коммутативном кольце и т. п.

2.1.2. Абелевы группы. Множество A *с одной* операцией $A \times A \to A$, удовлетворяющей первым четырём аксиомам сложения из опр. 2.1, называется *абелевой группой*. Таким образом, всякое коммутативное кольцо K является абелевой группой относительно операции сложения. Эта группа называется *аддитивной группой кольца*. Пример абелевой группы, не являющейся кольцом, доставляют *векторы*.

[1] Здесь имеется в виду обычное, не исключающее «или»: многочлен должен принимать значение 1 тогда и только тогда, когда *хотя бы одна* из переменных равна 1.

[2] Или привязанных к какой-либо другой позиционной системе счисления, например, двоичных.

Пример 2.4 (геометрические векторы)
Будем называть *геометрическим вектором* класс направленного отрезка (на плоскости или в пространстве) по отношению эквивалентности, отождествляющему между собой все отрезки, которые получающиеся друг из друга параллельным переносом. Нулевым вектором назовём класс эквивалентности точки — это единственный вектор, имеющий нулевую длину и не имеющий направления. Сложение векторов определяется стандартным образом: надо выбрать представителей векторов a и b так, чтобы конец a совпал с началом b, и объявить $a+b$ равным вектору с началом в начале a и концом в конце b. Коммутативность и ассоциативность этой операции видны из рис. 2⋄1 и рис. 2⋄2.

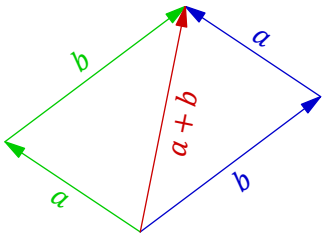


Рис. 2⋄1. Правило параллелограмма.

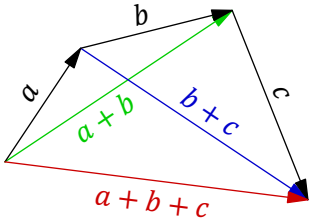


Рис. 2⋄2. Правило четырёхугольника.

Нулевым элементом является нулевой вектор. Вектор $-a$, противоположный вектору a, получается из вектора a изменением его направления на противоположное.

Пример 2.5 (мультипликативная группа поля)
Четыре аксиомы умножения из опр. 2.1 на стр. 20 утверждают, то множество

$$\mathbb{F}^* \stackrel{\text{def}}{=} \mathbb{F} \setminus 0$$

всех *ненулевых* элементов поля $\mathbb{F}$ является абелевой группой относительно операции умножения. Эту группу называют *мультипликативной группой поля*. Роль нуля из аддитивной группы $\mathbb{F}$ в мультипликативной группе $\mathbb{F}^*$ исполняет единица. В абстрактной абелевой группе такой элемент называется *нейтральным*. Мультипликативным аналогом перехода к противоположному элементу является переход к обратному элементу.

Лемма 2.1
В любой абелевой группе A нейтральный элемент единствен, и для каждого $a \in A$ противоположный к a элемент $-a$ определяется по a однозначно. В частности, $-(-a) = a$.

Доказательство. Будем записывать операцию в A аддитивно. Если есть два нулевых элемента 0_1 и 0_2, то $0_1 = 0_1 + 0_2 = 0_2$ (первое равенство выплнено, так как 0_2 является нулевым элементом, второе — поскольку нулевым элементом является 0_1). Если есть два элемента $-a$ и $-a'$, противоположных к a, то $-a = (-a)+0 = (-a)+\big(a+(-a')\big) = \big((-a)+a\big)+(-a') = 0+(-a') = -a'$. □

Лемма 2.2
В любом коммутативном кольце с единицей для любого элемента a выполняются равенства $0 \cdot a = 0$ и $(-1) \cdot a = -a$.

Доказательство. Пусть $a \cdot 0 = b$. Тогда $b + a = a \cdot 0 + a \cdot 1 = a(0+1) = a \cdot 1 = a$. Прибавляя к обеим частям этого равенства $(-a)$, получаем $b = 0$. Второе утверждение проверяется выкладкой $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = ((-1)+1) \cdot a = 0 \cdot a = 0$. □

Замечание 2.1. Аксиома нетривиальности (2-10) в определении поля равносильна требованию $\mathbb{F} \neq 0$, поскольку при $0 = 1$ для каждого $a \in \mathbb{F}$ получалось бы $a = a \cdot 1 = a \cdot 0 = 0$. Образование, состоящее из одного нуля, согласно предыдущим определениям является коммутативным кольцом (без единицы), но не полем.

2.1.3. Вычитание и деление. Из лем. 2.1 вытекает, что в любой абелевой группе корректно определена *разность* любых двух элементов

$$a - b \stackrel{\text{def}}{=} a + (-b)\,. \tag{2-12}$$

В частности, операция вычитания имеется в абелевой группе любого коммутативного кольца. В поле ненулевые элементы образуют абелеву группу по умножению. Поэтому в любом поле имеется ровно один единичный элемент, и для любого ненулевого элемента a обратный к нему элемент a^{-1} однозначно определяется по a. Тем самым, в любом поле помимо сложения, умножения и вычитания (2-12) имеется операция *деления* на любые ненулевые элементы

$$a/b \stackrel{\text{def}}{=} ab^{-1}\,, \quad b \neq 0\,. \tag{2-13}$$

2.2. Делимость в кольце целых чисел. Основным отличием коммутативных колец с единицей от полей является отсутствие обратных элементов к некоторым ненулевым элементам кольца. Элемент a коммутативного кольца K с единицей называется *обратимым*, если в этом кольце существует такой элемент a^{-1}, что $a^{-1}a = 1$. В противном случае элемент a называется *необратимым*.

Например, в кольце $\mathbb{Z}$ обратимыми элементами являются только 1 и -1. В кольце $\mathbb{Q}[x]$ многочленов с рациональными коэффициентами обратимыми элементами являются только ненулевые константы (многочлены степени нуль).

Говорят, что элемент a *делится* на элемент b, если в кольце существует такой элемент q, что $a = bq$. Это записывается как $b|a$ (читается «b делит a») или как $a \vdots b$ (читается «a делится на b»). Отношение делимости тесно связано с решением линейных уравнений.

2.2.1. Уравнение $ax + by = k$ и НОД в кольце $\mathbb{Z}$. Зафиксируем какие-нибудь целые числа a и b и обозначим через

$$(a, b) \stackrel{\text{def}}{=} \{ax + by \mid x, y \in \mathbb{Z}\} \tag{2-14}$$

множество всех целых чисел, представимых в виде $ax + by$ с целыми x, y. Это множество замкнуто относительно сложения и вместе с каждым своим элементом содержит все его целые кратные. Кроме того, все числа из (a, b) нацело делятся на каждый общий делитель чисел a и b, а сами a и b тоже входят в (a, b). Обозначим через d наименьшее положительное число в (a, b). Остаток от деления любого числа $z \in (a, b)$ на d лежит в (a, b), поскольку представляется в виде $z - kd$, а z и $-kd$ лежат в (a, b) при любом k. Так как этот остаток строго меньше d, он равен нулю. Следовательно, (a, b) совпадает с множеством всех чисел, кратных d.

Таким образом, число d является общим делителем чисел $a, b \in (a, b)$, представляется в виде $d = ax + by$ и делится на любой общий делитель чисел a и b. При этом произвольное число $k \in \mathbb{Z}$ представляется в виде $k = ax + by$ если и только если оно делится на d. Число d называется *наибольшим общим делителем* чисел $a, b \in \mathbb{Z}$ и обозначается нод(a, b).

Упражнение 2.4. Обобщите предыдущее рассуждение: для любого конечного набора чисел $a_1, \ldots, a_m \in \mathbb{Z}$ укажите число $d \in \mathbb{Z}$, которое делит все a_i, делится на любой их общий делитель и представляется в виде $d = a_1x_1 + \cdots + a_mx_m$ с целыми x_i. Покажите, что уравнение $n = a_1x_1 + \cdots + a_mx_m$ разрешимо относительно x_i в кольце $\mathbb{Z}$ если и только если $d|n$.

2.2.2. Алгоритм Евклида позволяет явно найти нод(a,b) для данных $a,b\in\mathbb{Z}$ и представить его в виде нод$(a,b)=ax+by$ с целыми x,y. Пусть $a\geqslant b$. Положим

$$E_0=a\,,\ E_1=b\,,\ E_k=\text{остатку от деления } E_{k-2} \text{ на } E_{k-1} \text{ при } k\geqslant 2. \tag{2-15}$$

Числа E_k строго убывают до тех пор, пока очередное число E_r не разделит нацело предыдущее число E_{r-1}, в результате чего E_{r+1} обратится в нуль. Последний ненулевой элемент E_r последовательности E_k и будет наибольшим общим общим делителем нод(a,b).

Упражнение 2.5. Докажите это.

Чтобы получить представление $E_r=x\cdot E_0+y\cdot E_1$ на каждом шаге вычисления надо представлять очередное E_k в виде $E_k=x\cdot E_0+y\cdot E_1$. Например, для чисел $n=10\,203$ и $m=4\,687$ вычисление состоит из восьми шагов:

$$\begin{aligned}
E_0 &= 10\,203 && = && +1\,E_0 && +0\,E_1\\
E_1 &= \ 4\,687 && = && +0\,E_0 && +1\,E_1\\
E_2 &= \ 829 = && E_0-2\,E_1 = && +1\,E_0 && -2\,E_1\\
E_3 &= \ 542 = && E_1-5\,E_2 = && -5\,E_0 && +11\,E_1\\
E_4 &= \ 287 = && E_2-E_3 = && +6\,E_0 && -13\,E_1\\
E_5 &= \ 255 = && E_3-E_4 = && -11\,E_0 && +24\,E_1\\
E_6 &= \ 32 = && E_4-E_5 = && +17\,E_0 && -37\,E_1\\
E_7 &= \ 31 = && E_5-7\,E_6 = && -130\,E_0 && +283\,E_1\\
E_8 &= \ 1 = && E_6-E_7 = && +147\,E_0 && -320\,E_1\\
\Big[E_9 &= \ 0 = && E_7-31\,E_8 = && -4\,687\,E_0 && +10\,203\,E_1\Big]
\end{aligned} \tag{2-16}$$

(взятая в скобки последняя строка служит для проверки). Таким образом,

$$\text{нод}(10\,203,4\,687)=1=147\cdot 10\,203-320\cdot 4\,687\,.$$

Упражнение 2.6. Докажите, что в возникающем на последнем шаге работы алгоритма Евклида представлении нуля в виде $0=E_{r+1}=q_0E_0+q_1E_1$ число $|q_0E_0|=|q_1E_1|$ рано *наименьшему общему кратному* нок(a,b) чисел a и b.

Замечание 2.2. С вычислительной точки зрения отыскание нод(a,b) при помощи алгоритма Евклида *несопоставимо* быстрее разложения чисел a и b на простые множители. Читателю предлагается убедиться в этом, попытавшись вручную разложить на простые множители исходные числа $n=10\,203$ и $m=4\,687$ из проделанного выше вручную вычисления (2-16). Если задано произведение двух *очень* больших простых чисел, то найти по нему сами эти числа за разумное время не под силу даже мощным компьютерам. Это обстоятельство лежит в основе многих популярных систем шифрования данных.

2.3. Взаимная простота. В кольце целых чисел $\mathbb{Z}$ условие нод$(a,b)=1$ равносильно разрешимости в целых числах уравнения $ax+by=1$, и числа a,b, обладающие этими свойствами, называются *взаимно простыми*. В произвольном коммутативном кольце K с единицей из разрешимости уравнения $ax+by=1$ также вытекает отсутствие у элементов a и b необратимых общих

делителей: если $a = d\alpha$, $b = d\beta$, и $ax + by = 1$, то $d(\alpha + \beta) = 1$ и d обратим. Однако, отсутствие у a и b необратимых общих делителей, вообще говоря, не гарантирует разрешимости уравнения $ax + by = 1$. Например, в кольце многочленов от двух переменных $\mathbb{Q}[x, y]$ одночлены x и y не имеют общих делителей, отличных от констант, однако равенство $f(x, y) \cdot x + g(x, y) \cdot y = 1$ невозможно ни при каких $f, g \in \mathbb{Q}[x, y]$.

Упражнение 2.7. Объясните почему.

При этом именно разрешимость уравнения $ax + by = 1$ влечёт за собою наличие у элементов a, b многих приятных свойств, которыми обладают взаимно простые целые числа.

Определение 2.2
Элементы a и b произвольного коммутативного кольца K с единицей называются *взаимно простыми*, если уравнение $ax + by = 1$ разрешимо в K относительно x и y.

Лемма 2.3
В произвольном коммутативном кольце K с единицей для любого $c \in K$ и любых взаимно простых $a, b \in K$ справедливы импликации:

(1) если ac делится на b, то c делится на b

(2) если c делится и на a, и на b, то c делится и на ab.

Кроме того, если $a \in K$ взаимно прост с каждым из элементов $b_1, \dots, b_n$, то он взаимно прост и с их произведением $b_1 \dots b_m$.

Доказательство. Умножая обе части равенства $ax + by = 1$ на c, получаем соотношение

$$c = acx + bcy\,,$$

из которого вытекают обе импликации (1), (2). Если $\forall\, i\ \exists\, x_i, y_i \in K\ :\ ax_i + b_i y_i = 1$, то перемножая все эти равенства и раскрывая скобки, получим в левой части сумму, в которой все слагаемые, кроме $(b_1 \dots b_n) \cdot (y_1 \dots y_n)$, делятся на a. Вынося a за скобку, приходим к соотношению $a \cdot X + (b_1 \dots b_n) \cdot (y_1 \dots y_n) = 1$. □

Упражнение 2.8. Пользуясь лем. 2.3, докажите следующую теорему об однозначности разложения на простые множители в кольце $\mathbb{Z}$: всякое целое число z является произведением конечного числа простых чисел[1], причём любые два таких представления $p_1 \dots p_k = z = q_1 \dots q_m$ имеют одинаковое число сомножителей $k = m$, и эти сомножители можно перенумеровать так, чтобы $p_i = \pm q_i$ для всех i.

Замечание 2.3. (нод в произвольном коммутативном кольце) Если коммутативное кольцо K произвольно, то *наибольшим общим делителем* элементов $a, b \in K$ принято называть любой элемент $d \in K$, который делит a и b и делится на любой общий делитель чисел a и b. Это определение не гарантирует ни существования, ни единственности наибольшего общего делителя, ни его представимости в виде $d = ax + by$.

[1] Напомним, что целое число называется *простым*, если оно не раскладывается в произведение двух чисел, каждое из которых отлично от ± 1.

2.4. Кольцо вычетов $\mathbb{Z}/(n)$. Напомним, что числа $a, b \in \mathbb{Z}$ называются *сравнимыми* по модулю n, что записывается как $a \equiv b \pmod{n}$, если их разность $a - b$ делится на n. Сравнимость по модулю n является отношением эквивалентности[1] и разбивает множество целых чисел на непересекающиеся классы сравнимых по модулю n чисел. Эти классы называются *классами вычетов по модулю* n, а их совокупность обозначается через $\mathbb{Z}/(n)$. Мы будем писать $[a]_n \in \mathbb{Z}/(n)$ для обозначения класса, содержащего число $a \in \mathbb{Z}$. Такое обозначение не однозначно: разные числа $x \in \mathbb{Z}$ и $y \in \mathbb{Z}$ задают один и тот же класс $[x]_n = [y]_n$ если и только если $x = y + dn$ для некоторого $d \in \mathbb{Z}$. Всего в $\mathbb{Z}/(n)$ имеется n различных классов: $[0]_n$, $[1]_n$, … , $[(n-1)]_n$. Сложение и умножение классов вычетов задаётся правилами:

$$[a] + [b] \stackrel{\text{def}}{=} [a + b]\,, \quad [a] \cdot [b] \stackrel{\text{def}}{=} [ab]\,. \tag{2-17}$$

Согласно упр. 1.10 на стр. 12, эти операции определены корректно[2]. Они очевидным образом удовлетворяют аксиомам коммутативного кольца с единицей — формулы (2-17) сводят операции над вычетами к операциям над целыми числами, для которых аксиомы выполнены.

2.4.1. Делители нуля и нильпотенты. В $\mathbb{Z}/(10)$ произведение классов $[2]$ и $[5]$ равно нулю, хотя *каждый* из них отличен от нуля, а в кольце $\mathbb{Z}/(8)$ ненулевой класс $[2]$ имеет нулевой куб $[2]^3 = [8] = [0]$.

В произвольном кольце K элемент $a \in K$ называется *делителем нуля*, если $a \neq 0$ и $ab = 0$ для некоторого ненулевого $b \in K$. Обратимый элемент $a \in K$ не может быть делителем нуля, поскольку, умножая обе части равенства $ab = 0$ на a^{-1}, мы получаем $b = 0$. Поэтому кольцо с делителями нуля не может быть полем. Кольцо с единицей без делителей нуля называется *целостным*.

Ненулевой элемент a кольца K называется *нильпотентом*, если $a^n = 0$ для некоторого $n \in \mathbb{N}$. Всякий нильпотент автоматически является делителем нуля. Кольцо с единицей без нильпотентов называется *приведённым*. Всякое целостное кольцо автоматически приведено.

2.4.2. Обратимые элементы кольца вычетов. Обратимость класса $[m]_n \in \mathbb{Z}/(n)$ означает существование такого класса $[x]_n$, что $[m]_n[x]_n = [mx]_n = [1]_n$. Последнее равенство равносильно наличию таких $x, y \in \mathbb{Z}$, что $mx + ny = 1$ в $\mathbb{Z}$. Тем самым, класс $[m]_n$ обратим в кольце $\mathbb{Z}/(n)$ если и только если $\text{нод}(m, n) = 1$ в кольце $\mathbb{Z}$.

Проверить, обратим ли данный класс $[m]_n$, и если да, то вычислить $[m]_n^{-1}$, можно при помощи алгоритма Евклида[3]. Так, проделанное в форм. (2-16) на стр. 24 вычисление показывает, что класс $[10\,203]$ обратим в $\mathbb{Z}/(4\,687)$ и $10\,203^{-1} = 147 \pmod{4\,687}$, а класс $4\,687$ обратим в $\mathbb{Z}/(10\,203)$ и $4\,687^{-1} = -320 \pmod{10\,203}$.

Обратимые элементы кольца $\mathbb{Z}/(n)$ образуют мультипликативную абелеву группу. Она называется *группой обратимых вычетов* по модулю n и обозначается $\mathbb{Z}/(n)^*$. Порядок этой группы равен количеству натуральных чисел, меньших n и взаимно простых с n. Он обозначается через $\varphi(n) \stackrel{\text{def}}{=} \left|\mathbb{Z}/(n)^*\right|$ и называется *функцией Эйлера* числа $n \in \mathbb{Z}$.

2.4.3. Поля вычетов $\mathbb{F}_p = \mathbb{Z}/(p)$. Из предыдущего вытекает, что кольцо вычетов $\mathbb{Z}/(n)$ является полем тогда и только тогда, когда n является *простым числом*. В самом деле, если $n = mk$ составное, ненулевые классы $[m], [k] \in \mathbb{Z}/(n)$ будут делителями нуля и не могут быть

[1] См. n° 1.4 на стр. 10.

[2] Т. е. не зависят от способа записи классов или, что то же самое — от выбора представителей $a \in [a]$ и $b \in [b]$.

[3] См. n° 2.2.2 на стр. 24.

обратимы. Напротив, если p простое число, то нод$(m, p) = 1$ для всех m, не кратных p, и значит, каждый ненулевой класс $[m] \in \mathbb{Z}/(p)$ обратим. Поле $\mathbb{Z}/(p)$, где p простое, принято обозначать $\mathbb{F}_p$.

Пример 2.6 (бином Ньютона по модулю p)
В поле $\mathbb{F}_p = \mathbb{Z}/(p)$ выполняется замечательное равенство

$$\underbrace{1 + 1 + \cdots + 1}_{p \text{ раз}} = 0\,. \tag{2-18}$$

Из него вытекает, что для любых $a, b \in \mathbb{F}_p$ выполняется равенство

$$(a + b)^p = a^p + b^p\,. \tag{2-19}$$

В самом деле, раскрывая скобки в биноме $(a + b)^p$, мы для каждого k получим $\binom{p}{k}$ одночленов $a^k b^{p-k}$, сумма которых равна $a^k b^{p-k} \cdot (1 + 1 + \cdots + 1)$, где в скобках стоит сумма $\binom{p}{k}$ единиц, равная нулю при $0 < k < p$.

Лемма 2.4
При простом p и любом k в пределах $1 \leqslant k \leqslant (p - 1)$ биномиальный коэффициент $\binom{p}{k}$ делится на p.

Доказательство. Так как число p взаимно просто со всеми числами от 1 до $p-1$, оно по лем. 2.3 взаимно просто с произведением $k!(p-k)!$. Поскольку $p!$ делится на $k!(p-k)!$, из той же лем. 2.3 следует, что $(p-1)!$ делится на $k!(p-k)!$, а значит, $\binom{p}{k} = \frac{p!}{k!(p-k)!}$ делится на p. □

Следствие 2.1 (малая теорема Ферма)
Для любого $a \in \mathbb{Z}$ и любого простого $p \in \mathbb{N}$ выполняется сравнение $a^p \equiv a \pmod p$.

Доказательство. Надо показать, что $[a^p] = [a]$ в поле $\mathbb{F}_p$. Согласно (2-19), имеем

$$[a]^p = \underbrace{\big([1] + [1] + \cdots + [1]\big)}_{a \text{ раз}}^p = \underbrace{[1]^p + [1]^p + \cdots + [1]^p}_{a \text{ раз}} = \underbrace{[1] + [1] + \cdots + [1]}_{a \text{ раз}} = [a]\,.$$

□

Упражнение 2.9. Покажите, что $\binom{mp^n}{p^n} \equiv m \pmod p$ для простого $p \nmid m$.

2.5. Прямые произведения. Прямое произведение

$$\prod_\nu A_\nu = A_1 \times \ldots \times A_\nu = \{(a_1, \ldots, a_m) \mid a_\nu \in A_\nu\ \forall\, \nu\} \tag{2-20}$$

абелевых групп $A_1, \ldots, A_m$ состоит из упорядоченных наборов $(a_1, \ldots, a_m)$ элементов $a_\nu \in A_\nu$ и обладает естественной структурой абелевой группы относительно покомпонентных операций:

$$(a_1, \ldots, a_m) + (b_1, \ldots, b_m) \stackrel{\text{def}}{=} (a_1 + b_1, \ldots, a_m + b_m)\,. \tag{2-21}$$

Упражнение 2.10. Проверьте, что так определённая операция коммутативна и ассоциативна, нулевым элементом для неё является набор нулей $(0, 0, \ldots, 0)$, а противоположным к набору $(a_1, \ldots, a_m)$ является набор $(-a_1, \ldots, -a_m)$.

Абелева группа (2-20) называется *прямым произведением* абелевых групп A_i. Если все группы A_i конечны, прямое произведение (2-20) тоже конечно и имеет порядок

$$\left|\prod A_i\right| = \prod |A_i| .$$

Прямые произведения имеют смысл не только для конечных, но и для любых семейств абелевых групп A_x, занумерованных элементами $x \in X$ произвольного множества X. Соответствующее произведение обозначается в этом случае через $\prod\limits_{x\in X} A_x$.

Аналогичным образом, для любого семейства коммутативных колец $\{K_x\}_{x\in X}$ определено прямое произведение $\prod K_x$, представляющее собою множество семейств элементов $(a_x)_{x\in X}$, в которых каждый элемент a_x лежит в своём кольце K_x. Операции сложения и умножения также определяются покомпонентно:

$$(a_x)_{x\in X} + (b_x)_{x\in X} \stackrel{\text{def}}{=} (a_x + b_x)_{x\in X} \ , \qquad (a_x)_{x\in X} \cdot (b_x)_{x\in X} \stackrel{\text{def}}{=} (a_x \cdot b_x)_{x\in X}$$

Упражнение 2.11. Убедитесь, что $\prod K_x$ является кольцом, причём если все K_x были кольцами с единицей, то $\prod K_x$ также будет кольцом с единицей $(1, 1, \dots, 1)$.

Например, если $X = \mathbb{R}$ и все $K_x = \mathbb{R}$, т. е. перемножается континуальное семейство одинаковых экземпляров поля $\mathbb{R}$, занумерованных действительными числами $x \in \mathbb{R}$, то прямое произведение $\prod\limits_{x\in\mathbb{R}} \mathbb{R}_x$ канонически изоморфно кольцу функций $f : \mathbb{R} \to \mathbb{R}$ с обычными операциями поточечного сложения и умножения значений функций. Этот изоморфизм переводит семейство вещественных чисел $(f_x) \in \prod\limits_{x\in\mathbb{R}} \mathbb{R}_x$, занумерованное вещественным числом x, в функцию $f : \mathbb{R} \to \mathbb{R}$, значение которой в точке $x \in \mathbb{R}$ равно x-тому элементу семейства: $f(x) = f_x$.

В прямом произведении колец любой ненулевой элемент, имеющий хотя бы одну нулевую компоненту, является делителем нуля. Например, $(0,\ 1,\ \dots,\ 1)$ является делителем нуля, так как $(0,\ 1,\ \dots,\ 1)(1,\ 0,\ \dots,\ 0) = (0,\ 0,\ \dots,\ 0) = 0$. Поэтому произведение нескольких колец никогда не является полем. Например, в произведении $\mathbb{F}_p \times \mathbb{F}_q$ конечных полей $\mathbb{F}_p$ и $\mathbb{F}_q$, состоящих из p и q элементов соответственно, имеется ровно $(p-1)(q-1)$ обратимых элементов (a, b), образующих мультипликативную группу $\mathbb{F}_p^* \times \mathbb{F}_q^*$, и $p+q-2$ делителя нуля, имеющих вид $(a, 0)$ и $(0, b)$ с $a, b \neq 0$.

В общем случае элемент $a = (a_1, \dots, a_m) \in K_1 \times \dots \times K_m$ обратим если и только если каждая его компонента $a_\nu \in K_\nu$ обратима в своём кольце K_ν. Поэтому группа обратимых элементов кольца $\prod K_\nu$ является прямым произведением групп обратимых элементов колец K_ν:

$$\left(\prod K_\nu\right)^* = \prod K_\nu^* \tag{2-22}$$

2.6. Гомоморфизмы. Отображение абелевых групп $\varphi : A \to B$ называется *гомоморфизмом*, если для любых $a_1, a_2 \in A$ в кольце B выполнено соотношение

$$\varphi(a_1 + a_2) = \varphi(a_1) + \varphi(a_2) . \tag{2-23}$$

В частности, этим условиям удовлетворяет *нулевой* (или *тривиальный*) гомоморфизм, отображающий все элементы A в нулевой элемент B.

Упражнение 2.12. Убедитесь, что композиция гомоморфизмов — это тоже гомоморфизм.

Любой гомоморфизм $\varphi\colon A \to B$ переводит нулевой элемент группы A в нулевой элемент группы B, так как из равенств $\varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$ вытекает, что $0 = \varphi(0)$. Равенства

$$\varphi(a) + \varphi(-a) = \varphi(a + (-a)) = \varphi(0) = 0$$

показывают, что $\varphi(-a) = -\varphi(a)$. Тем самым, *образ* $\operatorname{im}\varphi = \varphi(A) \subset B$ любого гомоморфизма $\varphi\colon A \to B$ является абелевой подгруппой в B.

2.6.1. Ядро гомоморфизма. Полный прообраз нулевого элемента группы B при гомоморфизме $\varphi\colon A \to B$ называется *ядром* гомоморфизма φ и обозначается

$$\ker\varphi = \varphi^{-1}(0) = \{a \in A \mid \varphi(a) = 0\}\ .$$

Ядро образует в A подгруппу, так как из равенств $\varphi(a_1) = 0$ и $\varphi(a_2) = 0$ вытекает равенство

$$\varphi(a_1 \pm a_2) = \varphi(a_1) \pm \varphi(a_2) = 0 \pm 0 = 0\,.$$

Предложение 2.1
Слой любого гомоморфизма абелевых групп $\varphi\colon A \to B$ над произвольной точкой $b \in B$ либо пуст, либо равен $a + \ker\varphi = \{a + a' \mid a' \in \ker\varphi\}$, где $a \in A$ — произвольно выбранный элемент, переходящий в b. В частности, все непустые слои находятся в биекции с $\ker\varphi$, и инъективность гомоморфизма φ равносильна равенству $\ker\varphi = 0$.

Доказательство. Равенства $\varphi(a_1) = \varphi(a_2)$ и $\varphi(a_1 - a_2) = \varphi(a_1) - \varphi(a_2) = 0$ равносильны. Поэтому элементы $a_1, a_2 \in A$ переходят в один и тот же элемент из B тогда и только тогда, когда $a_1 - a_2 \in \ker(\varphi)$. □

2.6.2. Группа гомоморфизмов. Для абелевых групп A, B через $\operatorname{Hom}(A, B)$ мы обозначаем множество всех *гомоморфизмов* $A \to B$. Это множество является абелевой группой относительно операции поточечного сложения значений:

$$\varphi_1 + \varphi_2\colon a \mapsto \varphi_1(a) + \varphi_2(a)\,.$$

Нулевым элементом группы $\operatorname{Hom}(A, B)$ является *нулевой гомоморфизм*, отображающий все элементы A в нулевой элемент B.

2.6.3. Гомоморфизмы колец. Отображение колец $\varphi\colon A \to B$ называется *гомоморфизмом колец*, если для любых $a_1, a_2 \in A$ в кольце B выполнены соотношения:

$$\begin{aligned} f(a_1 + a_2) &= f(a_1) + f(a_2) \\ f(a_1 a_2) &= f(a_1) f(a_2)\,. \end{aligned} \tag{2-24}$$

Поскольку гомоморфизм колец $\varphi\colon A \to B$ является гомоморфизмом аддитивных абелевых групп, он обладает всеми свойствами гомоморфизмов абелевых групп. В частности, $\varphi(0) = 0$, $\varphi(-a) = -\varphi(a)$, и все непустые слои φ являются сдвигами слоя над нулём: если $\varphi(a) = b$, то $\varphi^{-1}(b) = a + \ker\varphi = \{a + a' \mid a' \in \ker\varphi\}$. Поэтому гомоморфизм φ инъективен тогда и только тогда, когда $\ker\varphi = \{0\}$. Ядро гомоморфизма колец $\varphi\colon A \to B$ вместе с каждым элементом $a \in \ker\varphi$ содержит и все кратные ему элементы aa', поскольку $\varphi(aa') = \varphi(a)\varphi(a') = 0$. В частности, ядро $\ker\varphi$ является подкольцом в A. Образ гомоморфизма колец $\varphi\colon A \to B$ очевидно является подкольцом в B, однако он может не содержать единицы, и $1 \in A$ может не перейти в $1 \in B$.

Упражнение 2.13. Убедитесь, что отображение $\mathbb{Z}/(2) \to \mathbb{Z}/(6)$, $[0] \mapsto [0]$, $[1] \mapsto [3]$, является гомоморфизмом колец.

Предложение 2.2
Любой ненулевой гомоморфизм произвольного кольца с единицей в любое целостное[1] кольцо переводит единицу в единицу.

Доказательство. Из равенств $\varphi(1) = \varphi(1\cdot 1) = \varphi(1)\cdot\varphi(1)$ вытекает равенство

$$\varphi(1)\big(1-\varphi(1)\big) = 0\,.$$

В целостном кольце такое возможно либо при $\varphi(1) = 1$, либо при $\varphi(1) = 0$. Во втором случае $\forall\, a\in A\ \ \varphi(a) = \varphi(1\cdot a) = \varphi(1)\cdot\varphi(a) = 0$. □

2.6.4. Гомоморфизмы полей. Если кольца A и B являются полями, то всякий ненулевой гомоморфизм колец $\varphi : A \to B$ является гомоморфизмом мультипликативных групп этих полей. В частности, $\varphi(a/b) = \varphi(a)/\varphi(b)$ для всех a и всех $b \neq 0$.

Предложение 2.3
Любой ненулевой гомоморфизм из поля в произвольное кольцо является вложением.

Доказательство. Если $\varphi(a) = 0$ для какого-нибудь $a \neq 0$, то для каждого b

$$\varphi(b) = \varphi\left(ba^{-1}a\right) = \varphi\left(ba^{-1}\right)\varphi(a) = 0\,.$$

Поэтому любой ненулевой гомоморфизм из поля имеет нулевое ядро. □

2.7. Китайская теорема об остатках. Пусть целое число $n = n_1\ldots n_m$ является произведением попарно взаимно простых чисел $n_1,\ldots,n_m \in \mathbb{Z}$. Отображение, переводящее вычет $z \pmod n$ в набор вычетов $z \pmod{n_i}$:

$$\begin{aligned}\varphi :\ \mathbb{Z}/(n) &\to \mathbb{Z}/(n_1)\times\cdots\times\mathbb{Z}/(n_m)\\ [z]_n &\mapsto \big([z]_{n_1},\ldots,[z]_{n_m}\big)\,,\end{aligned} \tag{2-25}$$

корректно определено, поскольку при выборе другого представителя $z_1 \equiv z_2 \pmod n$ разность $z_1 - z_2$ делится на произведение $n = n_1\ldots n_m$, и $[z_1]_{n_i} = [z_2]_{n_i}$ при всех i. Легко видеть, что φ перестановочно со сложением:

$$\begin{aligned}\varphi\big([z]_n + [w]_n\big) = \varphi\big([z+w]_n\big) = \big([z+w]_{n_1},\ldots,[z+w]_{n_m}\big) = \big([z]_{n_1}+[w]_{n_1},\ldots,[z]_{n_m}+[w]_{n_m}\big) =\\ = \big([z]_{n_1},\ldots,[z]_{n_m}\big) + \big([w]_{n_1},\ldots,[w]_{n_m}\big) = \varphi\big([z]_n\big) + \varphi\big([w]_n\big)\end{aligned}$$

Аналогично проверяется, что φ перестановочно с умножением, т. е. является гомоморфизмом колец. Если $[z]_n \in \ker\varphi$, то z делится на каждое n_i, а значит, по лем. 2.3 на стр. 25, делится и на их произведение $n = n_1\ldots n_m$, откуда $[z]_n = 0$. Так как гомоморфизм с нулевым ядром инъективен и в кольцах $\mathbb{Z}/(n)$ и $\prod\mathbb{Z}/(n_i)$ одинаковое число элементов $n = n_1\ldots n_m$, отображение (2-25) биективно. Этот факт известен как *китайская теорема об остатках*.

На житейском языке он означает, что для любого набора остатков $r_1,\ldots,r_m$ от деления на попарно взаимно простые числа $n_1,\ldots,n_m$ всегда найдётся число z, имеющее остаток r_i от деления на n_i одновременно для всех i, причём любые два таких числа z_1, z_2 различаются на целое кратное числа $n = n_1\ldots n_k$. Практическое отыскание такого z осуществляется с помощью алгоритма Евклида следующим образом. Из взаимной простоты числа n_i с остальными числами n_ν

[1] Напомню, что *целостным* называется кольцо с единицей без делителей нуля, см. n° 2.4.1 на стр. 26.

вытекает[1], что n_i взаимно просто с произведением $m_i = \prod_{\nu\neq i} n_\nu$. Поэтому для каждого i найдутся такие $x_i, y_i \in \mathbb{Z}$, что $n_i x_i + m_i y_i = 1$. Число $b_i = m_i y_i$ даёт остаток 1 от деления на n_i и делится на все n_ν с $\nu \neq i$. Число $z = r_1 b_1 + \cdots + r_m b_m$ решает задачу.

Пример 2.7
Для демонстрации эффективности предыдущего алгоритма найдём наименьшее натуральное число, имеющее остатки $r_1 = 2$, $r_2 = 7$ и $r_3 = 43$ от деления, соответственно, на $n_1 = 57$, $n_2 = 91$ и $n_3 = 179$. Сначала найдём число, обратное к $91 \cdot 179$ по модулю 57 : замечаем, что $91 \cdot 179 \equiv 34 \cdot 8 \equiv -13 \pmod{57}$, применяем алгоритм Евклида к $E_0 = 57$ и $E_1 = 13$, приходим к равенству $22 \cdot 13 - 5 \cdot 57 = 1$. Таким образом, число

$$b_1 = -22 \cdot 91 \cdot 179 \quad (\equiv 22 \cdot 13 \,(\mathrm{mod}\ 57))$$

даёт при делении на 57, 91 и 179 остатки $(1,\ 0,\ 0)$. Аналогично находим числа

$$b_2 = -33 \cdot 57 \cdot 179 \quad (\equiv 33 \cdot 11 \,(\mathrm{mod}\ 91))$$
$$b_3 = -45 \cdot 57 \cdot 91 \quad (\equiv 45 \cdot 4 \,(\mathrm{mod}\ 179))$$

дающие при делении на 57, 91 и 179 остатки $(0,\ 1,\ 0)$ и $(0,\ 0,\ 1)$ соответственно. Требуемые остатки $(2,\ 7,\ 43)$ имеет число

$$z = 2\,b_1 + 7\,b_2 + 43\,b_3 = -(2 \cdot 22 \cdot 91 \cdot 179 + 7 \cdot 33 \cdot 57 \cdot 179 + 43 \cdot 45 \cdot 57 \cdot 91) =$$
$$= -(716\,716 + 2\,356\,893 + 10\,036\,845) = -13\,110\,454\,,$$

а также все числа, отличаются от него на целые кратные числа $n = 57 \cdot 91 \cdot 179 = 928\,473$. Наименьшим положительным среди них является $z + 15\,n = 816\,641$.

2.8. Характеристика. Для любого кольца K с единицей имеется канонический гомоморфизм колец $\varkappa : \mathbb{Z} \to K$, заданный правилом

$$\varkappa(\pm n) = \pm\underbrace{(1 + 1 + \cdots + 1)}_{n}, \quad \text{где} \quad n \in \mathbb{N}\,. \tag{2-26}$$

Образ $\operatorname{im}\varkappa$ является наименьшим подкольцом в K с единицей, равной единице кольца K. Если гомоморфизм $\varkappa$ инъективен, то говорят, что кольцо K имеет *характеристику нуль*. В противном случае *характеристикой* $\operatorname{char}(K)$ кольца K называют наименьшее $m \in \mathbb{N}$, для которого $\underbrace{1 + 1 + \cdots + 1}_{m} = 0$. Равенство

$$\underbrace{1 + 1 + \cdots + 1}_{mn} = \underbrace{(1 + 1 + \cdots + 1)}_{m} \cdot \underbrace{(1 + 1 + \cdots + 1)}_{n}$$

показывает, что характеристика целостного кольца либо равна нулю, либо является простым числом. Для целостного кольца K характеристики $p > 0$ гомоморфизм $\varkappa$ переводит все числа, кратные p, в нуль и корректно факторизуется до гомоморфизма поля вычетов

$$\varkappa_p : \mathbb{Z}/(p) \to K\,, \quad a\,(\mathrm{mod}\ p) \mapsto \varkappa(a)\,. \tag{2-27}$$

[1] По всё той же лем. 2.3 на стр. 25.

По предл. 2.3 гомоморфизм (2-27) инъективен, и значит, $\operatorname{im}\varkappa = \operatorname{im}\varkappa_p \simeq \mathbb{F}_p$. Таким образом, наименьшее содержащее единицу подкольцо целостного кольца K положительной характеристики является полем, изоморфным полю вычетов $\mathbb{Z}/(p)$ по простому модулю $p \in \mathbb{N}$, равному характеристике $\operatorname{char} K$.

2.8.1. Простое подполе. Пусть теперь $K = \mathbb{F}$ является полем. Его наименьшее по включению подполе называется *простым подполем* в $\mathbb{F}$. В силу своего определения простое подполе содержит образ $\operatorname{im}(\varkappa)$ гомоморфизма (2-26). Если $\operatorname{char}(\mathbb{F}) = p > 0$, то простое подполе совпадает с $\operatorname{im}\varkappa = \operatorname{im}\varkappa_p$ и изоморфно полю вычетов $\mathbb{Z}/(p)$. Если $\operatorname{char}(\mathbb{F}) = 0$, то гомоморфизм $\varkappa$ инъективно вкладывает $\mathbb{Z}$ в $\mathbb{F}$. Так как простое подполе содержит обратные ко всем элементам из $\operatorname{im}\varkappa$, правило $p/q \mapsto \varkappa(p)/\varkappa(q)$ продолжает $\varkappa$ до вложения полей $\varkappa\colon \mathbb{Q} \hookrightarrow \mathbb{F}$, образ которого совпадает с простым подполем. Тем самым, простое подполе поля характеристики нуль изоморфно полю рациональных чисел $\mathbb{Q}$.

Упражнение 2.14. Покажите, что а) каждый ненулевой гомоморфизм из поля в себя тождественно действует на простом подполе б) между полями разной характеристики не существует ненулевых гомоморфизмов.

Пример 2.8 (автоморфизмы поля $\mathbb{R}$)
Покажем, что каждый ненулевой гомоморфизм $\varphi\colon \mathbb{R} \to \mathbb{R}$ тождествен. Поскольку неравенство $x_1 < x_2$ равносильно тому, что $x_2 - x_1 = a^2$ для некоторого $a \neq 0$, мы заключаем, что для всех $x_1 < x_2$ выполняется неравенство $\varphi(x_1) < \varphi(x_2)$, так как $\varphi(x_2) - \varphi(x_1) = \varphi(x_2 - x_1) = \varphi(a^2) = \varphi(a)^2 > 0$. Таким образом, φ является строго монотонной функцией, совпадающей с тождественным отображением $\varphi(x) = x$ на простом подполе $\mathbb{Q} \subset \mathbb{R}$.

Упражнение 2.15 (по анализу). Покажите, что строго монотонная функция $\mathbb{R} \to \mathbb{R}$, совпадающая с функцией $\varphi(x) = x$ на подмножестве $\mathbb{Q} \subset \mathbb{R}$, совпадает с нею всюду.

2.8.2. Гомоморфизм Фробениуса. В поле $\mathbb{F}$ характеристики $\operatorname{char}(\mathbb{F}) = p > 0$ отображение возведения в p-тую степень

$$F_p\colon \mathbb{F} \to \mathbb{F}, \quad x \mapsto x^p, \tag{2-28}$$

является гомоморфизмом, поскольку $\forall\, a, b \in \mathbb{F}$ выполняются равенства $(ab)^p = a^p b^p$ и

$$(a+b)^p = a^p + b^p + \sum_{k=1}^{p-1} \underbrace{(1+1+\cdots+1)}_{\binom{p}{k}} \cdot a^k b^{p-k} = a^p + b^p$$

(ср. с прим. 2.6 и лем. 2.4 на стр. 27). Гомоморфизм (2-28) называется *гомоморфизмом Фробениуса*. Как и всякий ненулевой гомоморфизм из поля в себя, он тождественно действует на простом подполе $\mathbb{F}_p \subset \mathbb{F}$, ср. со сл. 2.1 на стр. 27.

§3. Многочлены и расширения полей

Всюду в этом параграфе мы обозначаем через K произвольное коммутативное кольцо с единицей, а через $\Bbbk$ — произвольное поле.

3.1. Формальные степенные ряды и многочлены. Бесконечное выражение вида

$$A(x) = \sum_{\nu \geqslant 0} a_\nu x^\nu = a_0 + a_1 x + a_2 x^2 + \dots\,, \quad \text{где} \quad a_i \in K\,, \tag{3-1}$$

называется *формальным степенным рядом* от переменной x с коэффициентами в кольце K. Два формальных степенных ряда

$$\begin{aligned} A(x) &= a_0 + a_1 x + a_2 x^2 + \dots \\ B(x) &= b_0 + b_1 x + b_2 x^2 + \dots \end{aligned} \tag{3-2}$$

равны, если $a_i = b_i$ для всех i. Сложение и умножение рядов (3-2) определяется стандартными правилами раскрытия скобок и приведения подобных слагаемых[1]: коэффициенты s_m и p_m рядов $S(X) = A(x) + B(x) = s_0 + s_1 x + s_2 x^2 + \dots$ и $P(x) = A(x)B(x) = p_0 + p_1 x + p_2 x^2 + \dots$ суть

$$\begin{aligned} s_m &= a_m + b_m \\ p_m &= \sum_{\alpha+\beta=m} a_\alpha b_\beta = a_0 b_m + a_1 b_{m-1} + \dots + a_0 b_m \end{aligned} \tag{3-3}$$

Упражнение 3.1. Убедитесь, что эти две операции удовлетворяют аксиомам коммутативного кольца с единицей.

Кольцо формальных степенных рядов от переменной x с коэффициентами в кольце K обозначается через $K[\![x]\!]$. Начальный коэффициент a_0 ряда (3-1) называется *свободным членом* этого ряда. Первый ненулевой коэффициент ряда A называется *младшим* коэффициентом.

Если в кольце K нет делителей нуля, младший коэффициент произведения двух рядов равен произведению младших коэффициентов сомножителей. Поэтому кольцо формальных степенных рядов с коэффициентами из целостного кольца тоже является целостным.

Кольцо формальных степенных рядов от n переменных $K[\![x_1, \dots, x_n]\!]$ определяется по индукции: $K[\![x_1, \dots, x_n]\!] \stackrel{\text{def}}{=} K[\![x_1, \dots, x_{n-1}]\!][\![x_n]\!]$ и представляет собой множество формальных сумм вида

$$F(x) = \sum_{\nu_1, \dots, \nu_n \in \mathbb{Z}_{\geqslant 0}} a_{\nu_1 \dots \nu_n} x_1^{\nu_1} \cdots x_n^{\nu_n}\,.$$

3.1.1. Алгебраические операции над формальными рядами. Назовём *n-арной алгебраической операцией* в $K[\![x]\!]$ правило, сопоставляющее n рядам $f_1, \dots, f_n$ новый ряд f так, что каждый коэффициент ряда f вычисляется по коэффициентам рядов $f_1, \dots, f_n$ при помощи конечного числа[2] сложений и умножений.

[1] Говоря формально, операции, о которых тут идёт речь, являются операциями над *последовательностями* (a_ν) и (b_ν) элементов кольца K. Буква x используется лишь для облегчения восприятия этих операций.

[2] Которое может зависеть от номера коэффициента.

Например, сложение и умножение рядов — это алгебраические операции, а подстановка вместо x численного значения $\alpha \in K$ алгебраической операцией обычно не является[1]. Напротив, подстановка в ряд $f(x)$ вместо x любого ряда $g(x) = b_1x + b_2x^2 + \cdots$ с нулевым свободным членом — это алгебраическая операция, дающая ряд

$$f(g(x)) = \sum a_k(b_1x + b_2x^2 + \cdots)^k = \\ a_0 + a_1(b_1x + b_2x^2 + \cdots) + a_2(b_1x + b_2x^2 + \cdots)^2 + a_3(b_1x + b_2x^2 + \cdots)^3 + \cdots \\ = a_0 + (a_1b_1)\cdot x + (a_1b_2 + a_2b_1^2)\cdot x^2 + (a_1b_3 + 2\,a_2b_1b_2 + a_3b_1^3)\cdot x^3 + \cdots,$$

в котором на коэффициент при x^m влияют лишь начальные члены первых m слагаемых. Ещё одним примером алгебраической операции является обращение рядов.

Предложение 3.1
Ряд $f(x) = a_0 + a_1x + a_2x^2 + \ldots \in K[\![x]\!]$ обратим в $K[\![x]\!]$ если и только если его свободный член a_0 обратим в K, и в этом случае обращение $f \mapsto f^{-1}$ является алгебраической операцией над рядом f.

Доказательство. Если имеется такой ряд $f^{-1}(x) = b_0 + b_1x + b_2x^2 + \ldots$, что

$$f(x)\cdot f^{-1}(x) = \big(a_0 + a_1x + a_2x^2 + \ldots\big)\cdot\big(b_0 + b_1x + b_2x^2 + \ldots\big) = 1, \tag{3-4}$$

то $a_0b_0 = 1$, откуда a_0 обратим. Наоборот, допустим, что $a_0 \in K$ обратим. Приравнивая коэффициенты при одинаковых степенях x в средней и правой части (3-4), мы получаем на коэффициенты b_i бесконечную систему уравнений

$$\begin{aligned} a_0b_0 &= 1\\ a_0b_1 + a_1b_0 &= 0\\ a_0b_2 + a_1b_1 + a_2b_0 &= 0\\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \end{aligned} \tag{3-5}$$

из которой $b_0 = a_0^{-1}$, и $b_k = -a_0^{-1}(a_1b_{k-1} + a_2b_{k-2} + \cdots + a_kb_0)$ при $k \geqslant 1$. □

Упражнение 3.2. Вычислите в $\mathbb{Q}[\![x]\!]$ а) $(1-x)^{-1}$ б) $(1-x^2)^{-1}$ в) $(1-x)^{-2}$.

3.1.2. Многочлены. Ряды с конечным числом ненулевых коэффициентов называются *многочленами*. Многочлены от переменных $x_1, \ldots, x_n$ с коэффициентами в кольце K образуют в кольце всех формальных степенных рядов подкольцо, которое обозначается $K[x_1, \ldots, x_n] \subset K[\![x_1, \ldots, x_n]\!]$. Многочлен от одной переменной x представляет собой формальное выражение вида $f(x) = a_0 + a_1x + \cdots + a_nx^n$. Последний ненулевой коэффициент этого выражения называется *старшим* коэффициентом многочлена f, а его номер называется *степенью* многочлена f и обозначается $\deg f$. Многочлены со старшим коэффициентом 1 называются *приведёнными*. Многочлены степени нуль называются *константами*.

Предложение 3.2
Если кольцо K целостное[2], то для любых многочленов $f_1, f_2 \in K[x]$ выполняется равенство

[1] Очевидным исключением из этого правила служит вычисление значения ряда $f(x)$ при $x = 0$, дающее в качестве результата свободный член этого ряда. Похожий эффект иногда возникает при вычислении значений некоторых очень специальных рядов в некоторых очень специальных точках α. Однако при произвольных α и f вычисление $f(\alpha)$ требует, вообще говоря, выполнения бесконечно большого количества сложений.

[2] Т. е. с единицей и без делителей нуля.

$\deg(f_1 f_2) = \deg(f_1) + \deg(f_2)$. В частности, кольцо $K[x]$ тоже целостное, и его обратимыми элементами являются только обратимые константы.

Доказательство. Все утверждения следуют из того, что старший коэффициент произведения равен произведению старших коэффициентов сомножителей. □

Упражнение 3.3. Покажите, что в кольце $\mathbb{Z}[x, y]$ двучлен $y^n - x^n$ делится нацело на двучлен $y - x$ и найдите частное.

3.1.3. Дифференциальное исчисление. Подставим в степенной ряд

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots$$

вместо x сумму $x + t$, где t — ещё одна переменная. Получится ряд

$$f(x + t) = a_0 + a_1(x + t) + a_2(x + t)^2 + \dots \in K[\![x, t]\!] .$$

Раскроем в нём все скобки и сгруппируем слагаемые по степеням переменной t, обозначив через $f_m(x) \in K[\![x]\!]$ ряд, возникающий как коэффициент при t^m :

$$f(x + t) = f_0(x) + f_1(x) \cdot t + f_2(x) \cdot t^2 + f_3(x) \cdot t^3 + \dots = \sum_{m \geqslant 0} f_m(x) \cdot t^m . \tag{3-6}$$

Упражнение 3.4. Убедитесь, что $f_0(x) = f(x)$ совпадает с исходным рядом f.

Ряд $f_1(x)$ называется *производной* от исходного ряда f и обозначается f' или $\frac{d}{dx} f$. Он однозначно определяется равенством

$$f(x + t) = f(x) + f'(x) \cdot t + (\text{члены, делящиеся на } t^2)$$

и может быть вычислен при помощи упр. 3.3 как значение при $t = 0$ ряда

$$\frac{f(x+t) - f(x)}{t} = a_1 \cdot \frac{(x+t) - t}{t} + a_2 \cdot \frac{(x+t)^2 - t^2}{t} + a_3 \cdot \frac{(x+t)^3 - t^3}{t} + \dots = \\ = \sum_{k \geqslant 1} a_k \cdot \left((x+t)^{k-1} + (x+t)^{k-2} x + (x+t)^{k-3} x^2 + \dots + x^{k-1} \right) .$$

Получаем хорошо известную формулу

$$f'(x) = \sum_{k \geqslant 1} k\, a_k x^{k-1} = a_1 + 2\, a_2 x + 3\, a_3 x^2 + \dots \tag{3-7}$$

Пример 3.1 (ряды с нулевой производной)
Из формулы (3-7) вытекает, что производная от константы равна нулю. Если характеристика[1] $\operatorname{char} K = 0$, то верно и обратное: $f' = 0$ тогда и только тогда, когда $f = \text{const}$. Однако, когда кольцо K имеет положительную характеристику, производная от всех мономов x^m, показатель которых делится на характеристику, обращается в нуль, поскольку согласно проделанному выше вычислению коэффициент m в формуле

$$\frac{d}{dx} x^m = \underbrace{x^{m-1} + \dots + x^{m-1}}_{m} = m \cdot x^{m-1}$$

[1] См. n° 2.8 на стр. 31.

представляет собою сумму m единиц кольца. В частности, над полем $\Bbbk$ характеристики $p > 0$ производная от ряда $f(x)$ равна нулю тогда и только тогда, когда $f(x) = g(x^p)$ для некоторого $g \in \Bbbk[\![x]\!]$.

Упражнение 3.5. Покажите, что при простом $p \in \mathbb{N}$ для любого $g \in \mathbb{F}_p[\![x]\!]$ выполняется равенство $g(x^p) = g(x)^p$.

Предложение 3.3 (правила дифференцирования)
Для любого $\alpha \in K$ и любых $f, g \in K[\![x]\!]$ справедливы равенства

$$(\alpha f)' = \alpha \cdot f', \quad (f+g)' = f' + g', \quad (fg)' = f' \cdot g + f \cdot g'. \tag{3-8}$$

Кроме того, если ряд g не имеет свободного члена, то

$$\big(f(g(x))\big)' = g'(x) \cdot f'\big(g(x)\big), \tag{3-9}$$

а если ряд f обратим, то

$$\frac{d}{dx} f^{-1} = -f'/f^2. \tag{3-10}$$

Доказательство. Первые два равенства в (3-8) вытекают прямо из формулы (3-7). Для доказательства третьего перемножим ряды

$$\begin{aligned} f(x+t) &= f(x) + t \cdot f'(x) + (\text{члены, делящиеся на } t^2) \\ g(x+t) &= g(x) + t \cdot g'(x) + (\text{члены, делящиеся на } t^2). \end{aligned}$$

С точностью до членов, делящихся на t^2, получим

$$f(x+t)g(x+t) = f(x)g(x) + t \cdot (f'(x)g(x) + f(x)g'(x)) + (\text{члены, делящиеся на } t^2),$$

откуда $(fg)' = f' \cdot g + f \cdot g'$. Формула (3-9) доказывается похожим образом: подставляя в $f(x)$ вместо x ряд $g(x+t)$, получаем

$$f\big(g(x+t)\big) = f\big(g(x) + t \cdot g'(x) + (\text{члены, делящиеся на } t^2)\big).$$

Введём ряд $\tau(x,t) \stackrel{\text{def}}{=} g(x+t) - g(x) = t \cdot g'(x) + (\text{члены, делящиеся на } t^2)$ и перепишем правую часть предыдущего разложения как

$$\begin{aligned} f\big(g(x+t)\big) = f\big(g(x) + \tau(x,t)\big) = \\ = f(g(x)) + \tau(x,t) \cdot f'(g(x)) + (\text{члены, делящиеся на } \tau(x,t)^2) = \\ = f(g(x)) + t \cdot g'(x) \cdot f'(g(x)) + (\text{члены, делящиеся на } t^2). \end{aligned}$$

Тем самым, $\big(f(g(x))\big)' = g'(x) \cdot f'\big(g(x)\big)$. Для доказательства формулы (3-10) продифференцируем обе части равенства $f \cdot f^{-1} = 1$. Получим $f' \cdot f^{-1} + f \cdot \big(f^{-1}\big)' = 0$, откуда $\big(f^{-1}\big)' = -f'/f^2$. □

Упражнение 3.6. Покажите, что ряды f_m из разложения (3-6) имеют вид[1]

$$f_m(x) = \frac{1}{m!} \frac{d^m}{dx^m} f(x).$$

[1] Здесь и далее через $\frac{d^m}{dx^m} = \big(\frac{d}{dx}\big)^m$ обозначается m-тая производная, т. е. результат m-кратного применения операции $\frac{d}{dx}$.

3.2. Делимость в кольце многочленов. Известная из школы процедура деления многочленов «уголком» может быть формализована следующим образом.

Предложение 3.4 (деление с остатком)
Пусть K — произвольное коммутативное кольцо с единицей, и многочлен $u \in K[x]$ имеет обратимый старший коэффициент. Тогда для любого многочлена $f \in K[x]$ существуют многочлены $q \in K[x]$ и $r \in K[x]$, такие что $f = u \cdot q + r$ и либо $\deg(r) < \deg(u)$, либо $r = 0$. Если кольцо K целостное, то такие q и r определяются по f и u однозначно.

Доказательство. Пусть

$$f = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$$
$$u = b_0 x^k + b_1 x^{k-1} + \cdots + b_{k-1} x + b_k\,,$$

где b_0 обратим. Если $n < k$, можно взять $q = 0$ и $r = f$. Если $k = 0$, т. е. $u = b_0$, можно взять $r = 0$, $q = b_0^{-1} f$. При $n \geqslant k > 0$ можно по индукции считать, что теорема верна для всех многочленов f степени, строго меньшей, чем n. Поскольку степень многочлена $f - a_0 b_0^{-1} x^{n-k} u$ строго меньше n, он представляется в виде $qu + r$, где $r = 0$ или $\deg r < \deg u$. Тогда

$$f = (q + a_0 b_0^{-1} x^{n-k}) \cdot u + r$$

также представляется в требуемом виде. Если кольцо K целостное, и p, s — другая такая пара многочленов, что $\deg(s) < \deg(u)$ и $up + s = f = uq + r$, то $u(q - p) = r - s$. При $p - q \neq 0$ степень многочлена в левой части не менее $\deg u$, т. е. строго больше, чем степень многочлена в правой части. Следовательно, $p - q = 0$, откуда и $r - s = 0$. □

Определение 3.1
Многочлены q и r, удовлетворяющие условиям предл. 3.4 называются *неполным частным* и *остатком* от деления f на u в $K[x]$.

Следствие 3.1
Для любых многочленов $f, g \in \Bbbk[x]$ с коэффициентами в произвольном поле $\Bbbk$ существует единственная пара многочленов $q, r \in \Bbbk[x]$, таких что $f = g \cdot q + r$ и либо $\deg(r) < \deg(g)$, либо $r = 0$.

Пример 3.2 (вычисление значения многочлена в точке)
Остаток от деления многочлена $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ на линейный двучлен $x - \alpha$ это константа, равная значению $f(\alpha)$ многочлена f при $x = \alpha$, в чём легко убедиться, подставляя $x = \alpha$ в равенство $f(x) = (x - \alpha) \cdot q(x) + r$. Отметим, что «деление уголком» является значительно более быстрым способом вычисления $f(\alpha)$, чем «лобовая» подстановка $x = \alpha$ в $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$.

Упражнение 3.7 (схема Горнера). Убедитесь, что

$$f(\alpha) = a_0 + \alpha \cdot \Big(a_1 + \alpha \cdot \Big(a_2 + \cdots + \alpha \cdot \big(a_{n-2} + \alpha \cdot (a_{n-1} + \alpha \cdot a_n) \big) \cdots \Big) \Big)$$

Предложение 3.5
Пусть $\Bbbk$ — произвольное поле. Для любого набора многочленов $f_1, \dots, f_n \in \Bbbk[x]$ существует единственный приведённый многочлен $d \in \Bbbk[x]$, который делит каждый из многочленов f_i и делится на любой многочлен, делящий каждый из многочленов f_i. Многочлен d представляется в виде

$$f_1 h_1 + \dots + f_n h_n\,, \quad \text{где} \quad h_i \in \Bbbk[x]\,. \tag{3-11}$$

Произвольно взятый многочлен $g \in \Bbbk[x]$ представим в виде (3-11) тогда и только тогда, когда он делится на d.

Доказательство. Единственность очевидна: два многочлена, каждый из которых делится на другой, имеют равные степени и могут различаться лишь постоянным множителем, который равен единице, коль скоро оба многочлена приведены. Существование доказывается тем же рассуждением, что и в n° 2.4.2 на стр. 26. Обозначим множество всех многочленов $g \in \Bbbk[x]$, представимых в виде (3-11), через

$$(f_1, \dots, f_n) \stackrel{\text{def}}{=} \left\{ f_1 h_1 + \dots + f_n h_n \;\middle|\; h_i \in \Bbbk[x] \right\}\,. \tag{3-12}$$

Это подкольцо в $\Bbbk[x]$, содержащее вместе с каждым многочленом g и все кратные ему многочлены hg (с любым $h \in \Bbbk[x]$). Кроме того, $(f_1, \dots, f_n)$ содержит каждый из многочленов f_i, и все многочлены из $(f_1, \dots, f_n)$ делятся на любой общий делитель всех многочленов f_i. Возьмём в качестве d приведённый многочлен наименьшей степени в $(f_1, \dots, f_n)$. Остаток $r = g - qd$ от деления произвольного многочлена $g \in (f_1, \dots, f_n)$ на d лежит в $(f_1, \dots, f_n)$. Так как его степень не может быть строго меньше $\deg d$, он нулевой. Тем самым, все многочлены в $(f_1, \dots, f_n)$ делятся на d. □

Определение 3.2
Многочлен d из предл. 3.5 называется *наибольшим общим делителем* многочленов f_i и обозначается нод$(f_1, \dots, f_n)$.

3.2.1. Взаимная простота. Из предл. 3.5 вытекает, что в кольце $\Bbbk[x]$ многочленов с коэффициентами в поле *взаимная простота* многочленов $f_1, \dots, f_m$, т. е. возможность представить единицу в виде $1 = h_1 f_1 + \dots + h_n f_n$, равносильна равенству нод$(f_1, \dots, f_n) = 1$, т. е. отсутствию у многочленов $f_1, \dots, f_n$ общих делителей положительной степени — точно так же, как это происходит в кольце целых чисел $\mathbb{Z}$.

Определение 3.3
Многочлен $f \in K[x]$ с коэффициентами в целостном[1] кольце K называется *неприводимым*, если из равенства $f = gh$ вытекает, что g или h является обратимой константой.

Упражнение 3.8. Пусть $\Bbbk$ — любое поле. Пользуясь лем. 2.3, докажите следующую теорему об однозначности разложения на простые множители в кольце $\Bbbk[x]$: любой многочлен f является произведением конечного числа неприводимых многочленов, причём любые два таких представления $p_1 p_2 \cdots p_k = f = q_1 q_2 \cdots q_m$ имеют одинаковое число сомножителей $k = m$, и эти сомножители можно перенумеровать так, чтобы $\forall\ i\ p_i = \lambda_i q_i$, где $\lambda_i \in \Bbbk$ — некоторые ненулевые константы.

[1] Т. е. с единицей и без делителей нуля.

3.2.2. Алгоритм Евклида из n° 2.2.2 дословно переносится на многочлены с коэффициентами в произвольном поле $\Bbbk$. А именно, для пары многочленов $f_1, f_2 \in \Bbbk[x]$ с $\deg(f_1) \geqslant \deg(f_2)$ положим $E_0 = f_1$, $E_1 = f_2$, $E_k =$ остатку от деления E_{k-2} на E_{k-1} при $k \geqslant 1$. Степени многочленов E_k строго убывают до тех пор, пока какой-то E_r не разделит нацело предыдущий E_{r-1}, в результате чего E_{r+1} обратится в нуль. Последний ненулевой многочлен $E_r = \text{нод}(f_1, f_2)$.

Упражнение 3.9. Докажите это.

Если при вычислении каждого E_k представлять его в виде $E_k = h_1^{(k)} f_1 + h_2^{(k)} f_2$, то $E_{r+1} = 0$ и $E_r = \text{нод}(f_1, f_2)$ тоже получатся представленными в таком виде, причём в выражении $0 = E_{r+1} = h_1^{(r+1)} f_1 + h_2^{(r+1)} f_2$ многочлены $h_1^{(r+1)}$ и $h_2^{(r+1)}$ будут взаимно простыми множителями, дополняющими f_1 и f_2 до их наименьшего общего кратного

$$\text{нок}(f_1, f_2) = h_1^{(r+1)} f_1 = -h_2^{(r+1)} f_2 \,.$$

Упражнение 3.10. Докажите это.

Вот как выглядит это вычисление для многочленов

$f_1(x) = x^7 + 3x^6 + 4x^5 + x^4 + 5x^2 + 3x^3 + 3x + 4$ и $f_2(x) = x^5 + 5x^4 + 11x^3 + 12x^2 + 7x + 4$:

$$\begin{aligned} E_0 &= x^7 + 3x^6 + 4x^5 + x^4 + 5x^2 + 3x^3 + 3x + 4 \\ E_1 &= x^5 + 5x^4 + 11x^3 + 12x^2 + 7x + 4 \\ E_2 &= -4x^4 - 13x^3 - 21x^2 - 10x - 8 = E_0 - \left(x^2 - 2x + 3\right) E_1 \end{aligned}$$

дальше делить на E_2 удобнее не E_1, а $16E_1$, а потом поделить результат на 16

$$E_3 = \frac{1}{16}\left(x^3 + 5x^2 + 10x + 8\right) = \frac{1}{16}\left(16E_1 + (4x+7)\,E_2\right) = \frac{4x+7}{16}E_0 - \frac{4x^3 - x^2 - 2x + 5}{16}E_1$$

следующий шаг уже даёт наибольший общий делитель

$$E_4 = -16(x^2 + 3x + 4) = E_2 + 16\,(4x - 7)\,E_3 = 16\left(x^2 - 3\right)E_0 - 16\left(x^4 - 2x^3 + 2x - 2\right)E_1$$

поскольку $E_5 = E_3 + (x+2)\cdot E_4 \,/\, 256 = (x^3 + 2x^2 + x + 1)\cdot E_0 - (x^5 + x^2 + 1)\cdot E_1 = 0$. Откуда $\text{нод}(f_1, f_2) = x^2 + 3\,x + 4 = -\left(x^2 - 3\right) f_1(x) + \left(x^4 - 2\,x^3 + 2\,x - 2\right) f_2(x)$, а $\text{нок}(f_1, f_2) = \left(x^3 + 2\,x^2 + x + 1\right) f_1(x) = \left(x^5 + x^2 + 1\right) f_2(x)$.

3.3. Корни многочленов. Элемент $\alpha \in K$ называется *корнем* многочлена $f \in K[x]$, если

$$f(\alpha) = 0\,.$$

Как мы видели в прим. 3.2, это равносильно тому, что $f(x)$ делится в $K[x]$ на $(x - \alpha)$.

Упражнение 3.11. Пусть $\Bbbk$ — поле. Проверьте, что многочлен степени 2 или 3 неприводим в $\Bbbk[x]$ тогда и только тогда, когда у него нет корней в поле $\Bbbk$.

Предложение 3.6
Пусть K — целостное кольцо и $f \in K[x]$ имеет s различных корней $\alpha_1, \ldots, \alpha_s \in K$. Тогда f делится в $K[x]$ на произведение $\prod_i (x - \alpha_i)$. В частности, $\deg(f) \geqslant s$ или $f = 0$.

Доказательство. Так как в K нет делителей нуля и $(\alpha_i - \alpha_1) \neq 0$ при $i \neq 1$, подставляя в равенство $f(x) = (x - \alpha_1) \cdot q(x)$ значения $x = \alpha_2, \alpha_3, \ldots, \alpha_s$, убеждаемся, что $\alpha_2, \alpha_3, \ldots, \alpha_s$ являются корнями многочлена $q(x)$, и применяем индукцию. □

Следствие 3.2
Ненулевой многочлен f с коэффициентами из целостного кольца не может иметь в этом кольце более $\deg(f)$ различных корней.

Следствие 3.3
Пусть кольцо K целостное, и $f, g \in K[x]$ имеют степени, не превосходящие n. Если $f(\alpha_i) = g(\alpha_i)$ для более, чем n попарно разных $\alpha_i \in K$, то $f = g$ в $K[x]$.

Доказательство. Так как $\deg(f - g) \leqslant n$, и многочлен $f - g$ имеет больше n корней, он нулевой. □

Упражнение 3.12 (формула Лагранжа). Покажите, что для любых $n + 1$ попарно разных элементов поля $\Bbbk$ и произвольного набора значений $b_0, b_1, \ldots, b_n \in \Bbbk$ существует единственный такой многочлен $f(x) \in \Bbbk[x]$ степени $\deg f \leqslant n$, что $f(a_i) = b_i$ при всех i, и получите для него явную формулу.

3.3.1. Общие корни нескольких многочленов. Пусть $\Bbbk$ — поле. Число α тогда и только тогда является общим корнем многочленов $f_1, \ldots, f_m \in \Bbbk[x]$, когда α является корнем их наибольшего общего делителя. В самом деле, если $(x-\alpha)$ делит каждый из f_i, то по предл. 3.5 $(x-\alpha)$ делит нод$(f_1, \ldots, f_m)$, и наоборот. Таким образом, отыскание общих корней набора многочленов сводится к отысканию корней их наибольшего общего делителя, что бывает проще, чем отыскание корней любого из f_i в отдельности, т. к. степень нод$(f_1, \ldots, f_m)$ обычно меньше степеней всех f_i. Если многочлены $f_1, \ldots, f_m \in \Bbbk[x]$ взаимно просты, то они не имеют общих корней не только в поле $\Bbbk$, но и ни в каком большем кольце $K \supset \Bbbk$. В самом деле, поскольку существуют многочлены $h_i \in \Bbbk[x]$, такие что $f_1 h_1 + \cdots + f_m h_m = 1$, многочлены f_i не могут одновременно обратиться в нуль ни при каком значении x.

3.3.2. Кратные корни. Пусть $\Bbbk$ — произвольное поле. Число $\alpha \in \Bbbk$ называется m-кратным корнем многочлена $f \in \Bbbk[x]$, если $f(x) = (x - \alpha)^m \cdot g(x)$, где $g(\alpha) \neq 0$. Корни кратности $m \geqslant 2$ называются *кратными*.

Предложение 3.7
Число $\alpha \in \Bbbk$ является кратным корнем многочлена $f \in \Bbbk[x]$ если и только если

$$f(\alpha) = f'(\alpha) = 0\,.$$

Доказательство. Если α — кратный корень многочлена f, то $f(x) = (x - \alpha)^2 g(x)$. Дифференцируя, получаем $f'(x) = (x-\alpha)\big(2g(x)+(x-\alpha)g'(x)\big)$, откуда $f'(\alpha) = 0$. Если α не является кратным корнем, то $f(x) = (x - \alpha)g(x)$, где $g(x) \neq 0$. Тогда $f'(x) = (x - \alpha)g'(x) + g(x)$ и $f'(\alpha) = g(\alpha) \neq 0$. □

Определение 3.4
Многочлен $f \in \Bbbk[x]$ называется *сепарабельным*, если нод$(f, f') = 1$. По предл. 3.7 это означает, что многочлен f не имеет кратных корней ни в каком кольце $K \supset \Bbbk$.

Пример 3.3 (сепарабельность неприводимых многочленов)
Если многочлен $f \in \Bbbk[x]$ неприводим, то он взаимно прост со всеми ненулевыми многочленами меньшей степени. Поэтому нод$(f, f') = 1$, если только f' не обращается тождественно в нуль. Поскольку над полем характеристики нуль производная многочлена положительной степени отлична от нуля, все неприводимые многочлены над полем характеристики нуль сепарабельны. Над конечным полем $\mathbb{F}_p = \mathbb{Z}/(p)$ многочлен $f \in \mathbb{F}_p[x]$ имеет $f' = 0$ если и только если $f(x) = g\left(x^p\right)$ для некоторого $g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m \in \mathbb{F}_p[x]$. Поскольку возведение в p-тую степень является в характеристике p гомоморфизмом[1] и тождественно действует на элементах поля $\mathbb{F}_p$,

$$\begin{aligned} f(x) = g(x^p) = b_0x^{pm} + b_1x^{p(m-1)} + \cdots + b_{m-1}x^p + b_0 = \\ = b_0^px^{pm} + b_1^px^{p(m-1)} + \cdots + b_{m-1}^px^p + b_0^p = \\ = \left(b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m\right)^p = g^p(x)\,. \end{aligned}$$

Таким образом, над полем $\mathbb{F}_p = \mathbb{Z}/(p)$ всякий многочлен с нулевой производной является чистой p-той степенью некоторого другого многочлена. В частности, любой многочлен с нулевой производной приводим, и значит, неприводимые многочлены над полем $\mathbb{F}_p$ тоже сепарабельны. Над бесконечными полями $\Bbbk$ конечной характеристики $\operatorname{char}\Bbbk = p > 0$ неприводимые многочлены уже не обязательно сепарабельны. Например, можно показать, что над полем $\Bbbk = \mathbb{F}_p(t)$ рациональных функций от одной переменной t с коэффициентами в поле $\mathbb{F}_p$ многочлен $f(x) = x^p - t$ неприводим, однако $f' \equiv 0$, т. е. f не сепарабелен.

Предложение 3.8
Если $\operatorname{char}\Bbbk = 0$, то $\alpha \in \Bbbk$ является m-кратным корнем многочлена $f \in \Bbbk[x]$ тогда и только тогда, когда α является корнем f и первых $(m-1)$ производных от f, но не является корнем m-той производной.

Доказательство. Если $f(x) = (x-\alpha)^m g(x)$, то $f'(x) = (x-\alpha)^{m-1}\big(mg(x) + (x-\alpha)g'(x)\big)$. При $g(\alpha) \neq 0$ второй сомножитель в этом равенстве отличен от нуля при $x = \alpha$. Поэтому α является m-кратным корнем f тогда и только тогда, когда α является $(m-1)$-кратным корнем f'. □

3.3.3. Присоединение корней. Кольцо вычетов $\Bbbk[x]/(f)$, где $f \in \Bbbk[x]$, определяется аналогично кольцу[2] $\mathbb{Z}/(n)$. Зафиксируем произвольный отличный от константы многочлен $f \in \Bbbk[x]$ и обозначим через $(f) = \{fh \mid h \in \Bbbk[x]\}$ подкольцо всех многочленов, делящихся на f. Отношение $g_1 \equiv g_2 \pmod f$, означающее по определению, что $g_1 - g_2 \in (f)$, является отношением эквивалентности и разбивает $\Bbbk[x]$ в объединение непересекающихся классов $[g]_f = g + (f) = \{g + fh \mid h \in \Bbbk[x]\}$, которые называются *классами вычетов* по модулю f. Сложение и умножение этих классов задаётся формулами

$$[g] + [h] \stackrel{\text{def}}{=} [g+h]\,, \quad [g]\cdot[h] \stackrel{\text{def}}{=} [gh]\,. \tag{3-13}$$

Упражнение 3.13. Проверьте корректность[3] этого определения, а также выполнение в $\Bbbk[x]/(f)$ всех аксиом коммутативного кольца с единицей.

[1] См. прим. 2.6 на стр. 27.
[2] См. n° 2.4 на стр. 26.
[3] Т. е. независимость классов $[g+h]$ и $[gh]$ от выбора представителей $g \in [g]$ и $h \in [h]$.

Нулём кольца $\Bbbk[x]/(f)$ является класс $[0]_f = (f)$, единицей — класс $[1]_f = 1 + (f)$. Так как константы не делятся на многочлены положительной степени, классы всех констант $c \in \Bbbk$ различны по модулю f. Иначе говоря, поле $\Bbbk$ гомоморфно вкладывается в кольцо $\Bbbk[x]/(f)$ в качестве подполя, образованного классами констант. Поэтому для классов чисел $c \in \Bbbk$ мы всюду далее будем писать c вместо $[c]_f$.

Упражнение 3.14. Покажите, что поле $\Bbbk[x]/(x-\alpha)$ изоморфно полю $\Bbbk$.

Поскольку любой многочлен $g \in \Bbbk[x]$ единственным образом записывается в виде

$$g = fh + r\,, \quad \text{где} \quad \deg r < \deg f\,,$$

в каждом классе $[g]_f$ имеется единственный представитель $r \in [g]_f$ с $\deg(r) < \deg(f)$, т. е. каждый класс $[g]_f \in \Bbbk[x]/(f)$ *однозначно* записывается в виде

$$[a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}]_f = a_0 + a_1\vartheta + \cdots + a_{n-1}\vartheta^{n-1}\,, \quad \text{где} \quad \vartheta = [x]_f \quad \text{и} \quad a_i \in \Bbbk\,.$$

Класс $\vartheta = [x]_f$ удовлетворяет в кольце $\Bbbk[x]/(f)$ уравнению $f(\vartheta) = 0$, т. к.

$$f(\vartheta) = f\left([x]_f\right) = [f(x)]_f = [0]_f\,.$$

Поэтому сложение и умножение классов по правилам (3-13) можно интерпретировать как формальное сложение и умножение записей

$$a_0 + a_1\vartheta + \;\cdots\; + a_{n-1}\vartheta^{n-1}\,, \tag{3-14}$$

по стандартным правилам раскрытия скобок и приведения подобных с учётом того, что символ ϑ удовлетворяет соотношению $f(\vartheta) = 0$. По этой причине кольцо $\Bbbk[x]/(f)$ часто обозначают через $\Bbbk[\vartheta]$, где $f(\vartheta) = 0$, и называют *расширением* поля $\Bbbk$ посредством *присоединения* к нему корня ϑ многочлена $f \in \Bbbk[x]$.

Например, кольцо $\mathbb{Q}[x]/(x^2-2)$ можно воспринимать как множество формальных записей вида $a + b\sqrt{2}$, где $\sqrt{2} \stackrel{\text{def}}{=} [x]$. Сложение и умножение таких записей происходит по стандартным правилам раскрытия скобок с учётом того, что $\left(\sqrt{2}\right)^2 = 2$:

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$$
$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (cb + ad)\sqrt{2}$$

Упражнение 3.15. Проверьте, что $\mathbb{Q}[\sqrt{2}]$ является полем, и выясните, являются ли полями кольца $\mathbb{Q}[\vartheta]$, в которых а) $\vartheta^3 + 1 = 0$ б) $\vartheta^3 + 2 = 0$.

Предложение 3.9
Пусть $\Bbbk$ — произвольное поле. Кольцо $\Bbbk[x]/(f)$ является полем тогда и только тогда, когда многочлен f неприводим в $\Bbbk[x]$.

Доказательство. Если $f = gh$, где оба многочлена f, g имеют строго меньшую, чем f, степень, то ненулевые классы $[g], [h]$ будут делителями нуля в $\Bbbk[x]/(f)$, что невозможно в поле. Если же f неприводим, то для любого $g \notin (f)$ $\text{нод}(f, g) = 1$, а значит, $fh + gq = 1$ для некоторых $h, q \in \Bbbk[x]$, откуда $[q] \cdot [g] = [1]$ в $\Bbbk[x]/(f)$. □

Упражнение 3.16. Напишите явную формулу для вычисления обратного элемента к числу $a_0 + a_1\vartheta$ в поле $\mathbb{Q}[\vartheta]$, где $\vartheta^2 + \vartheta + 1 = 0$.

Теорема 3.1
Для любого поля $\Bbbk$ и любого многочлена $f \in \Bbbk[x]$ существует такое поле $\mathbb{F} \supset \Bbbk$, что f разлагается в $\mathbb{F}[x]$ в произведение $\deg f$ линейных множителей.

Доказательство. Индукция по $n = \deg f$. Пусть для любого поля $\Bbbk$ и для всех многочленов степени $< n$ из $\Bbbk[x]$ мы умеем строить такое поле[1]. Если f приводим: $f = gh$, где $\deg g < n$ и $\deg h < n$, мы можем построить поле $\mathbb{L} \supset \Bbbk$ над которым g полностью разложится на линейные множители, а затем поле $\mathbb{F} \supset \mathbb{L}$ над которым разложится h, а тем самым, и f. Если f неприводим, рассмотрим поле $\mathbb{L} = \Bbbk[x]/(f)$. Оно содержит $\Bbbk$ в качестве классов констант, и многочлен f делится в $\mathbb{L}[x]$ на $(x - \vartheta)$, где $\vartheta = [x] \pmod f$. Частное от этого деления имеет степень $n - 1$ и по индукции раскладывается на линейные множители над некоторым полем $\mathbb{F} \supset \mathbb{L}$. Тогда и f разложится над $\mathbb{F}$. □

Теорема 3.2 (китайская теорема об остатках)
Пусть $\Bbbk$ — произвольное поле, и многочлен $f \in \Bbbk[x]$ является произведением m попарно взаимно простых сомножителей: $f = f_1 f_2 \cdots f_m$, нод$(f_i, f_j) = 1$. Отображение

$$\varphi : \Bbbk[x]/(f) \to \Bbbk[x]/(f_1) \times \Bbbk[x]/(f_2) \times \cdots \times \Bbbk[x]/(f_m)$$
$$\varphi : [g]_f \longmapsto \left([g]_{f_1}, [g]_{f_2}, \ldots, [g]_{f_m}\right)$$

является корректно определённым изоморфизмом колец.

Доказательство. Проверки того, что φ корректно определён[2], является гомоморфизмом и имеет нулевое ядро, дословно повторяют рассуждения из n° 2.7, и мы оставляем их читателю. Покажем, что φ сюръективен. Для этого, как и в n° 2.7, построим для любого заданного набора классов $[r_i]_{f_i} \in \Bbbk[x]/(f_i)$ такой многочлен $g \in \Bbbk[x]$, что $g \equiv r_i \pmod{f_i}$ при всех i. Для каждого i обозначим через $F_i = \prod_{\nu \neq i} f_\nu$ произведение всех многочленов f_ν кроме i-того. Так как f_i взаимно прост с каждым из f_ν с $\nu \neq i$, он по лем. 2.3 взаимно прост с F_i. Поэтому существует такой многочлен[3] $h_i \in \Bbbk[x]$, что $F_i \cdot h_i \equiv 1 \pmod{f_i}$. Так как многочлен $g_i = F_i \cdot h_i$ при этом делится на все f_ν с $\nu \neq i$, многочлен $g = r_1 g_1 + \cdots + r_m g_m \equiv r_i \pmod{f_i}$ при всех i. □

3.4. Поле комплексных чисел $\mathbb{C} \stackrel{\text{def}}{=} \mathbb{R}[t]/(t^2 + 1)$ является расширением поля $\mathbb{R}$ при помощи корня квадратного уравнения $x^2 + 1 = 0$ и состоит из классов $[x + yt] = x + y \cdot i$, где $x, y \in \mathbb{R}$, а класс $i \stackrel{\text{def}}{=} [t]$ удовлетворяет соотношению $i^2 = -1$. $\mathbb{C}$ является полем, так как для $x + yi \neq 0$

$$\frac{1}{x + yi} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} \cdot i\,.$$

Комплексное число $z = x + yi$ удобно изображать на плоскости $\mathbb{R}^2$ с фиксированной прямоугольной системой координат (x, y) *радиус-вектором* $0z$, ведущим из начала координат в точку $z = (x, y)$, как на рис. 3◇1 на стр. 44 ниже. Координаты (x, y) называются при этом *действительной* и *мнимой* частями числа $z \in \mathbb{C}$ и обозначаются через $\mathrm{Re}(z)$ и $\mathrm{Im}(z)$ соответственно, а

[1] Заметим, что при $n = 2$ это так: достаточно взять $\mathbb{F} = \Bbbk$.

[2] Т. е. $\varphi\left([g]_f\right)$ не зависит от выбора представителя $g \in \Bbbk[x]$ в классе $[g]_f \subset \Bbbk[x]$.

[3] Чтобы найти его явно, можно, например, взять остаток R_i от деления F_i на f_i и применить алгоритм Евклида к паре $E_0 = f_i$, $E_1 = R_i$.

длина $|z| = \sqrt{x^2+y^2}$ радиус вектора $0z$ называется *модулем* (или *абсолютной величиной*) комплексного числа z. Множество всех таких $\vartheta \in \mathbb{R}$, что поворот плоскости $\mathbb{C}$ вокруг нуля на угол ϑ совмещает координатный луч $0x$ с лучом $0z$, называется *аргументом* числа z и обозначается $\operatorname{Arg}(z) = \{\alpha + 2\pi k \mid k \in \mathbb{Z}\} \subset \mathbb{R}$, где α — ориентированная длина какой-нибудь дуги[1], идущей по единичной окружности из точки $(1,0)$ в точку $z/|z|$.

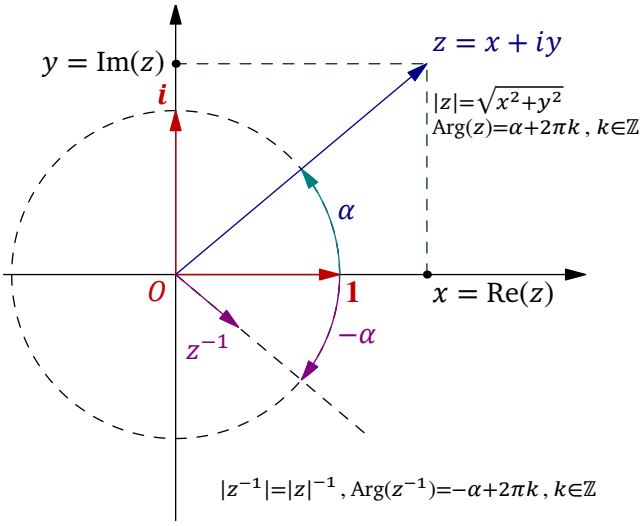


Рис. 3◇1.

Таким образом, $z = x+yi \in \mathbb{C}$ имеет $\operatorname{Re}(z) = |z|\cdot\cos\alpha$, $\operatorname{Im}(z) = |z|\cdot\sin\alpha$ и может быть записан как $z = |z|\cdot(\cos\alpha + i\cdot\sin\alpha)$, где $\alpha \in \operatorname{Arg}(z)$.

Лемма 3.1
Множество радиус-векторов точек z евклидовой координатной плоскости $\mathbb{R}^2$ с операцией сложения векторов и операцией умножения, заключающейся в перемножении модулей и сложении аргументов:

$$|z_1 z_2| \stackrel{\text{def}}{=} |z_1|\cdot|z_2| \tag{3-15}$$

$$\operatorname{Arg}(z_1 z_2) \stackrel{\text{def}}{=} \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2) = \{\vartheta_1 + \vartheta_2 \mid \vartheta_1 \in \operatorname{Arg}(z_1),\ \vartheta_2 \in \operatorname{Arg}(z_2)\}, \tag{3-16}$$

образует поле, изоморфное полю $\mathbb{C}$. Изоморфизм сопоставляет числу $x+iy \in \mathbb{C}$ точку $z = (x,y) \in \mathbb{R}^2$.

Упражнение 3.17. Проверьте, что сложение аргументов (3-16) определено корректно.

Доказательство лем. 3.1. Векторы на плоскости образуют абелеву группу по сложению, а ненулевые векторы — абелеву группу относительно операции умножения, задаваемой правилами (3-15) и (3-16): единицей служит единичный направляющий вектор оси $0x$, а обратным к ненулевому вектору z является вектор z^{-1} с

$$|z^{-1}| = 1/|z|, \quad \operatorname{Arg}(z^{-1}) = -\operatorname{Arg}(z) \tag{3-17}$$

[1] Любые две таких дуги отличаются друг от друга на целое число оборотов, а «ориентированность» означает, что длину дуги следует брать со знаком «+», если движение вдоль этой дуги происходит против часовой стрелки, и со знаком «−», если по часовой стрелке.

(см. рис. 3⋄1). Для проверки дистрибутивности заметим, что отображение $\lambda_a : z \mapsto az$ умножения на фиксированный ненулевой вектор a это *поворотная гомотетия*[1] плоскости $\mathbb{R}^2$ относительно начала координат на угол $\mathrm{Arg}(a)$ с коэффициентом $|a|$. Аксиома дистрибутивности $a(b+c) = ab + ac$ утверждает, что поворотная гомотетия перестановочна со сложением векторов: $\lambda_a(b+c) = \lambda_a(b) + \lambda_a(c)$. Это действительно так, поскольку и повороты и гомотетии переводят параллелограммы в параллелограммы.

Таким образом, векторы на евклидовой координатной плоскости $\mathbb{R}^2$ образуют поле. Векторы, параллельные оси $0x$ образуют в этом поле подполе, изоморфное полю $\mathbb{R}$. Произвольный вектор $z = (x, y) \in \mathbb{R}^2$ записывается в виде $z = x+iy$, где i — единичный направляющий вектор оси $0y$, числа $x, y \in \mathbb{R}$ понимаются как векторы, параллельные оси 0, а сложение и умножение происходят по правилам из условия леммы. При этом $i^2 = -1$ и для любых векторов $z_1 = x_1+iy_1$ и $z_2 = x_2 + iy_2$ выполняются равенства

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

что полностью согласуется с умножением классов вычетов $[x + yt]$ в $\mathbb{R}[t]/(t^2 + 1)$. □

3.4.1. Сопряжение. Числа$z = x + iy$ и $\overline{z} \stackrel{\text{def}}{=} x - iy$ называются *комплексно сопряжёнными*. В терминах комплексного сопряжения формулу для обратного числа можно записать в виде

$$z^{-1} = \overline{z}/|z|^2 .$$

Геометрически, комплексное сопряжение $z \mapsto \overline{z}$ представляет собою симметрию комплексной плоскости относительно вещественной оси $0x$. С алгебраической точки зрения сопряжение является инволютивным[2] автоморфизмом поля $\mathbb{C}$, т. е. $\forall\, z \in \mathbb{C}\ \ \overline{\overline{z}} = z$ и $\forall\ z_1, z_2 \in \mathbb{C}$

$$\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2 \quad \text{и} \quad \overline{z_1 z_2} = \overline{z}_1 \overline{z}_2 .$$

3.4.2. Тригонометрия. Большая часть школьной тригонометрии является не самой удобной для восприятия записью заурядных вычислений с комплексными числами, лежащими на единичной окружности. Например, произведение $z_1 z_2$ двух таких чисел

$$z_1 = \cos\varphi_1 + i\sin\varphi_1 \quad \text{и} \quad z_2 = \cos\varphi_2 + i\sin\varphi_2$$

по лем. 3.1 равно $\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)$, а с другой стороны, пользуясь дистрибутивностью умножения, получаем $z_1 z_2 = \Big(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2\Big) + i\Big(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2\Big)$, откуда $\cos(\varphi_1+\varphi_2) = \cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2$ и $\sin(\varphi_1+\varphi_2) = \cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2$. Тем самым, мы *доказали* тригонометрические формулы сложения аргументов.

[1] *Поворотной гомотетией* относительно точки 0 на угол α с коэффициентом $\varrho > 0$ называется композиция поворота на угол α вокруг точки 0 и растяжения в ϱ раз относительно 0 (поскольку растяжения коммутируют с поворотами, всё равно, в каком порядке эта композиция выполняется).

[2] Отличный от тождественного эндоморфизм $\iota : X \to X$ произвольного множества X называется *инволюцией*, если $\iota \circ \iota = \mathrm{Id}_X$. По предл. 1.4 на стр. 15 всякая инволюция автоматически биективна.

Пример 3.4 (тригонометрические функции кратных углов)
По лем. 3.1 $z = \cos\varphi + i\sin\varphi$ имеет $z^n = \cos(n\varphi) + i\sin(n\varphi)$. Раскрывая в $(\cos\varphi + i\sin\varphi)^n$ скобки по форм. (1-9) на стр. 9, получаем равенство

$$\begin{aligned}\cos(n\varphi) + i\sin(n\varphi) &= (\cos\varphi + i\sin\varphi)^n = \\ &= \cos^n\varphi + i\binom{n}{1}\cos^{n-1}\varphi\sin\varphi - \binom{n}{2}\cos^{n-2}\varphi\sin^2\varphi - i\binom{n}{3}\cos^{n-3}\varphi\sin^3\varphi + \cdots = \\ &= \left(\binom{n}{0}\cos^n\varphi - \binom{n}{2}\cos^{n-2}\varphi\sin^2\varphi + \binom{n}{4}\cos^{n-4}\varphi\sin^4\varphi - \cdots\right) + \\ &\quad + i\cdot\left(\binom{n}{1}\cos^{n-1}\varphi\sin\varphi - \binom{n}{3}\cos^{n-3}\varphi\sin^3\varphi + \binom{n}{5}\cos^{n-5}\varphi\sin^5\varphi - \cdots\right)\end{aligned}$$

заключающее в себе сразу *все* мыслимые формулы для кратных углов:

$$\begin{aligned}\cos(n\varphi) &= \binom{n}{0}\cos^n\varphi - \binom{n}{2}\cos^{n-2}\varphi\sin^2\varphi + \binom{n}{4}\cos^{n-4}\varphi\sin^4\varphi - \cdots \\ \sin(n\varphi) &= \binom{n}{1}\cos^{n-1}\varphi\sin\varphi - \binom{n}{3}\cos^{n-3}\varphi\sin^3\varphi + \binom{n}{5}\cos^{n-5}\varphi\sin^5\varphi - \cdots\end{aligned}$$

Например, $\cos 3\varphi = \cos^3\varphi - 3\cos\varphi\cdot\sin^2\varphi = 4\cos^3\varphi - 3\cos^2\varphi$.

Упражнение 3.18. Выразите $\sin(2\pi/5)$ и $\cos(2\pi/5)$ через радикалы от рациональных чисел.

3.4.3. Корни из единицы и круговые многочлены. Решим в поле $\mathbb{C}$ уравнение

$$z^n = 1\,.$$

Сравнивая модули левой и правой части, получаем $|z^n| = |z|^n = 1$, откуда $|z| = 1$. Сравнивая аргументы, получаем $n\operatorname{Arg}(z) = \operatorname{Arg}(1) = \{2\pi k \mid k\in\mathbb{Z}\}$. Поскольку

$$n\varphi \in \{2\pi k \mid k\in\mathbb{Z}\} \quad\Longleftrightarrow\quad \varphi \in \{2\pi k/n \mid k\in\mathbb{Z}\}\,,$$

имеется ровно n различных классов эквивалентности вещественных чисел по модулю добавления целых кратных 2π, которые при умножении их представителей на n превращаются в класс $\{2\pi k \mid k\in\mathbb{Z}\}$. Это классы n геометрически различных углов $2\pi k/n$ с $0 \leqslant k \leqslant n-1$. Таким образом, уравнение $z^n = 1$ имеет ровно n корней

$$\zeta_k = \cos(2\pi k/n) + i\sin(2\pi k/n) \quad (\text{где } k = 0,\ 1,\ \ldots,\ (n-1))\,,$$

расположенных в вершинах правильного n-угольника, вписанного в единичную окружность так, что вершина ζ_0 находится в точке 1 (см. рис. 3⋄2). Они образуют абелеву группу относительно операции умножения. Эта группа обозначается $\boldsymbol{\mu}_n$ и называется *группой корней n-той степени из* единицы.

Корень $\zeta\in\boldsymbol{\mu}_n$ называются *первообразным корнем* степени n из единицы, если все остальные элементы группы $\boldsymbol{\mu}_n$ представляются в виде ζ^k с $k\in\mathbb{N}$. Например, корень с наименьшим положительным аргументом $\zeta_1 = \cos(2\pi/n) + i\sin(2\pi/n)$ является первообразным. Но есть и другие: скажем, на рис. 3⋄2 все четыре отличных от 1 корня пятой степени из единицы являются первообразными, а в группе $\boldsymbol{\mu}_6$ на рис. 3⋄3 на стр. 47 первообразными являются только ζ_1 и $\zeta_5 = \zeta_1^{-1}$.

Упражнение 3.19. Покажите, что корень $\zeta_1^k = \cos(2\pi k/n) + i\sin(2\pi k/n)$ является первообразным тогда и только тогда, когда $\text{нод}(k,n) = 1$.

Приведённый многочлен, имеющий корнями все первообразные корни степени n из единицы и только их

$$\Phi_n(z) = \prod_{\substack{1 \leqslant k < n : \\ \text{нод}(k,n)=1}} (z - z_1^k), \tag{3-18}$$

называется n-тым *круговым* (или *циклотомическим*) многочленом. Например, пятый и шестой круговые многочлены имеют вид

$$\Phi_5(z) = (z - z_1)(z - z_2)(z - z_3)(z - z_4) = z^4 + z^3 + z^2 + z + 1$$
$$\Phi_6(z) = (z - z_1)(z - z_4) = z^2 - z + 1\,.$$

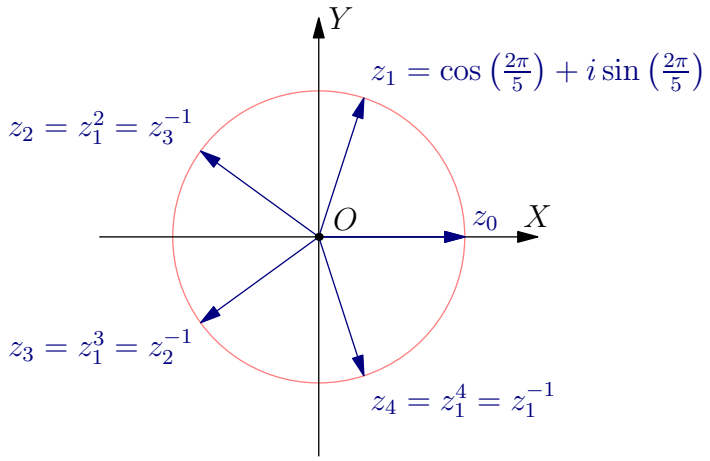


Рис. 3◇2. Корни уравнения $z^5 = 1$.

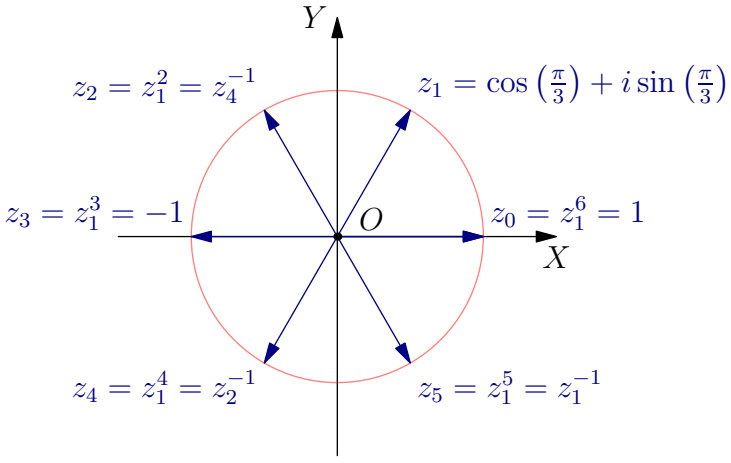


Рис. 3◇3. Корни уравнения $z^6 = 1$.

Упражнение 3.20*. Покажите, что при всех n многочлен Φ_n имеет целые коэффициенты и неприводим[1] в $\mathbb{Q}[x]$.

[1] Т. е. не являются произведениями многочленов строго меньшей степени.

Пример 3.5 (уравнение $z^n = a$)
Корни уравнения $z^n = a$ это числа $z = |z| \cdot (\cos\varphi + i\sin\varphi)$ с $|z|^n = |a|$, а $n\varphi \in \operatorname{Arg}(a)$. При $a = |a| \cdot (\cos\alpha + i\sin\alpha) \neq 0$ имеется ровно n таких чисел

$$z_k = \sqrt[n]{|a|} \cdot \left(\cos\frac{\alpha + 2\pi k}{n} + i \cdot \sin\frac{\alpha + 2\pi k}{n}\right), \quad 0 \leqslant k \leqslant n-1\,.$$

Они располагаются в вершинах правильного n-угольника, вписанного в окружность радиуса $\sqrt[n]{|a|}$ с центром в нуле так, что радиус-вектор одной из его вершин располагается под углом α/n к оси $0x$.

Пример 3.6 (гауссовы числа)
Рассмотрим в $\mathbb{C}$ подкольцо, состоящее из всех чисел с целыми координатами

$$\mathbb{Z}[i] \stackrel{\text{def}}{=} \{z = x + iy \mid x, y \in \mathbb{Z}\}\,.$$

Оно называется кольцом *гауссовых целых чисел* и часто используется в арифметике. Например, классическая задача о представлении натурального числа в виде суммы двух квадратов целых чисел существенно проясняется расширением кольца $\mathbb{Z}$ до кольца $\mathbb{Z}[i]$, в котором $x^2 + y^2 = (x + iy)(x - iy)$, так что разрешимость в кольце $\mathbb{Z}$ уравнения $x^2 + y^2 = n$ равносильна разрешимости в кольце $\mathbb{Z}[i]$ уравнения $n = z \cdot \bar{z}$. Из второго уравнения сразу же видно, что если числа m_1 и m_2 представляются в виде суммы двух квадратов

$$m_1 = a_1^2 + b_1^2 = (a_1 + ib_1)(a_1 - ib_1) = z_1\bar{z}_1$$
$$m_2 = a_2^2 + b_2^2 = (a_2 + ib_2)(a_2 - ib_2) = z_2\bar{z}_2$$

то их произведение $m = m_1 m_2$ также является суммой двух квадратов:

$$m = z_1 z_2 \cdot \overline{z_1 z_2} = |z_1 z_2|^2 = (a_1 b_1 - a_2 b_2)^2 + (a_1 b_2 + a_2 b_1)^2$$

(это соотношение известно как *тождество Эйлера*). В сочетании с теоремой о единственности разложения на простые множители в кольце $\mathbb{Z}[i]$, которую мы докажем в §5, тождество Эйлера сводит вопрос о представимости произвольного натурального числа в виде суммы двух квадратов к анализу представимости простых чисел. Мы ещё вернёмся к этому в прим. 5.6 на стр. 74.

Упражнение 3.21. Покажите, что обратимыми элементами кольца $\mathbb{Z}[i]$ являются четыре числа: ± 1 и $\pm i$.

3.5. Конечные поля. Для конечного поля $\mathbb{F}_p = \mathbb{Z}/(p)$ из p элементов и неприводимого многочлена $f \in \mathbb{F}_p[x]$ степени n поле вычетов $\mathbb{F}_p[x]/(f)$ состоит из p^n элементов вида

$$a_0 + a_1\vartheta + \cdots + a_{n-1}\vartheta^{n-1}, \quad \text{где} \quad a_i \in \mathbb{F}_p \quad \text{и} \quad f(\vartheta) = 0\,.$$

Например, $x^2 + x + 1 \in \mathbb{F}_2[x]$ неприводим, поскольку не имеет корней в $\mathbb{F}_2$. Соответствующее поле $\mathbb{F}_4 = \mathbb{F}_2[x]/(x^2 + x + 1) = \mathbb{F}_2[\omega] \,:\, \omega^2 + \omega + 1 = 0$ состоит из четырёх элементов[1]: 0, 1, $\omega = x \pmod{x^2 + x + 1}$ и $1 + \omega = \omega^2 = \omega^{-1}$.

Упражнение 3.22. Убедитесь, что мультипликативная группа $\mathbb{F}_4^*$ поля $\mathbb{F}_4$ изоморфна циклической группе $\boldsymbol{\mu}_3$.

[1] Отметим, что в силу равенства $-1 = 1$ в поле $\mathbb{F}_2$ можно обходиться без «минусов».

Расширение $\mathbb{F}_2 \subset \mathbb{F}_4$ аналогично расширению $\mathbb{R} \subset \mathbb{C}$, если понимать поле $\mathbb{C}$ как расширение $\mathbb{R}[\omega]$, где $\omega^2+\omega+1=0$, получающееся присоединением к полю $\mathbb{R}$ первообразного комплексного кубического корня из единицы[1]. Аналогом комплексного сопряжения $\mathbb{C} \to \mathbb{C}$, переводящего ω в $\overline{\omega} = \omega^2$, в поле $\mathbb{F}_4$ является гомоморфизм Фробениуса[2] $F_2 : \mathbb{F}_4 \to \mathbb{F}_4$, $a \mapsto a^2$, который тождественно действует на простом подполе $\mathbb{F}_2 = \{0, 1\}$ и переводит корни многочлена $x^2 + x + 1$ друг в друга.

Рассмотрим ещё один пример. Многочлен $x^2 + 1 \in \mathbb{F}_3[x]$ не имеет корней в $\mathbb{F}_3$, и значит, неприводим. Соответствующее поле $\mathbb{F}_9 = \mathbb{F}_3[i]$ состоит из девяти элементов $a + bi$ где $a, b \in \{-1, 0, 1\} = \mathbb{F}_3$, а $i^2 = -1$. Автоморфизм Фробениуса $F_3 : a \mapsto a^3$ переводит элемент $a + bi$ в $a - bi$.

Упражнение 3.23. Составьте для поля $\mathbb{F}_9$ таблицу умножения и таблицу обратных элементов, перечислите все имеющиеся в $\mathbb{F}_9$ квадраты и кубы и выясните, не изоморфна ли мультипликативная группа $\mathbb{F}_9^*$ группе $\boldsymbol{\mu}_8$.

Теорема 3.3
Для каждого $n \in \mathbb{N}$ и простого $p \in \mathbb{N}$ существует конечное поле $\mathbb{F}_q$, состоящее из $q = p^n$ элементов.

Доказательство. Рассмотрим в $\mathbb{F}_p[x]$ многочлен $f(x) = x^q - x$. По теор. 3.1 существует такое поле $\mathbb{F} \supset \mathbb{F}_p$, что f полностью раскладывается в $\mathbb{F}[x]$ в произведение q линейных множителей. Поскольку производная $f'(x) \equiv 1$, все эти множители различны, т. е. в поле $\mathbb{F}$ имеется ровно q различных чисел α, таких что $\alpha^q = \alpha$. Они образуют поле: если $\alpha^q = \alpha$, то $(-\alpha)^q = -\alpha$ и $\left(\alpha^{-1}\right)^q = \alpha^{-1}$, и для любого $\beta = \beta^q$ имеем $\alpha\beta = \alpha^q\beta^q = (\alpha\beta)^q$ и

$$\alpha + \beta = \alpha^{p^n} + \beta^{p^n} = F_p^n(\alpha) + F_p^n(\beta) = F_p^n(\alpha + \beta) = (\alpha + \beta)^q\,,$$

где $F_p : \mathbb{F} \to \mathbb{F}$, $x \mapsto x^p$, это гомоморфизм Фробениуса. □

Упражнение 3.24. Покажите, что число элементов в любом конечном поле является степенью его характеристики.

3.5.1. Конечные мультипликативные подгруппы в поле. Рассмотрим абелеву группу A, операцию в которой будем записывать мультипликативно.

Группа A называется *циклической*, если в ней имеется элемент $a \in A$, такой что все элементы группы A представляются в виде a^n с некоторым $n \in \mathbb{Z}$. Всякий элемент $a \in A$, обладающий этим свойством, называется *образующей* циклической группы A.

Например, группа комплексных корней из единицы $\boldsymbol{\mu}_n \subset \mathbb{C}$, рассматривавшаяся нами в n° 3.4.3, является циклической, а её образующими являются первообразные корни.

Если группа A конечна, то среди степеней любого элемента $b \in A$ будут встречаться одинаковые, скажем $b^k = b^m$ с $k > m$. Домножая обе части этого равенства на b^{-m}, получаем равенство $b^{k-m} = 1$. Таким образом, для каждого элемента $b \in A$ существует показатель $m \in \mathbb{N}$, такой что $b^m = 1$. Наименьший такой показатель называется *порядком* элемента b и обозначается $\operatorname{ord} b$.

Если $\operatorname{ord} b = n$, то элементы $b^0 = 1$, $b^1 = b$, b^2, …, b^{n-1} попарно различны, и любая целая степень b^m совпадает с одним из них: если $m = nq + r$, где r — остаток от деления m на n, то $b^m = (b^n)^q b^r = b^r$.

[1] Т. е. комплексного корня того же самого многочлена $x^2 + x + 1$.

[2] См. n° 2.8.2 на стр. 32.

Предложение 3.10
Любая конечная подгруппа A в мультипликативной группе $\Bbbk^*$ произвольного поля $\Bbbk$ является циклической.

Доказательство. Обозначим через m максимальный из порядков элементов группы A. Достаточно убедиться, что порядок каждого элемента группы A делит m: тогда все элементы группы A будут корнями многочлена $x^m - 1 = 0$, а значит, их не более m и все они исчерпываются степенями имеющегося в A элемента m-того порядка.

Чтобы увидеть, что порядки всех элементов группы являются делителями максимального порядка, достаточно для любых двух элементов $b_1, b_2 \in A$, имеющих порядки m_1, m_2, построить элемент $b \in A$, порядок которого равен нок(m_1, m_2).

Упражнение 3.25. Покажите, что при нод$(m_1, m_2) = 1$ в качестве такого элемента подойдёт $b = b_1 b_2$.

Если m_1 и m_2 не взаимно просты, то, раскладывая их согласно упр. 2.8 в произведение простых чисел, мы можем представить нок(m_1, m_2) в виде произведения $\ell_1 \ell_2$ так, что

$$m_1 = k_1 \ell_1\,, \quad m_2 = k_2 \ell_2 \quad \text{и} \quad (\ell_1, \ell_2) = 1\,.$$

Упражнение 3.26. Убедитесь в этом.

Элементы $b_1' = b_1^{k_1}$ и $b_2' = b_2^{k_2}$ имеют взаимно простые порядки ℓ_1 и ℓ_2, а их произведение $b_1' b_2'$ по упр. 3.25 имеет порядок $\ell_1 \ell_2 =$ нок(m_1, m_2), что и требовалось. □

Теорема 3.4
Всякое конечное поле изоморфно одному из полей $\mathbb{F}_q$, построенных в теор. 3.3.

Доказательство. Если $\operatorname{char} \mathbb{F} = p$, то по упр. 3.24 поле $\mathbb{F}$ состоит из $q = p^n$ элементов для подходящего $n \in \mathbb{N}$, а ненулевые элементы поля $\mathbb{F}$ образуют согласно предл. 3.10 циклическую группу по умножению, порождённую некоторым элементом $\zeta \in \mathbb{F}^*$, так что

$$\mathbb{F} = \{0,\ 1,\ \zeta,\ \zeta^2,\ \ldots\ ,\ \zeta^{q-2}\}\,.$$

Мы построим сейчас ещё одно поле из q элементов, которое будет изоморфно как полю $\mathbb{F}$, так и полю $\mathbb{F}_q$ из теор. 3.3. Для этого обозначим через $g \in \mathbb{F}_p[x]$ приведённый многочлен наименьшей степени, такой что $g(\zeta) = 0$.

Упражнение 3.27. Покажите, что g неприводим в $\mathbb{F}_p[x]$ и нацело делит любой многочлен $f \in \mathbb{F}_p[x]$, для которого $f(\zeta) = 0$.

Из упражнения вытекает, что кольцо вычетов $\mathbb{F}_p[x]/(g)$ является полем, а правило

$$h(x) \,(\mathrm{mod}\ g) \mapsto h(\zeta)$$

корректно задаёт гомоморфизм колец $\mathrm{ev}_\zeta :\ \mathbb{F}_p[x]/(g) \to \mathbb{F}$. Он инъективен, т. к. $\mathbb{F}_p[x]/(g)$ поле, и сюръективен, поскольку его образ содержит все степени ζ^m. Тем самым, $\mathbb{F} \simeq \mathbb{F}_p[x]/(g)$.

С другой стороны, т. к. ζ является корнем многочлена $f(x) = x^q - x$, из упр. 3.27 вытекает, что $f = gu$ для некоторого $u \in \mathbb{F}_p[x]$. Подставляя в это равенство q элементов поля $\mathbb{F}_q$, построенного в теор. 3.3 и состоящего в точности из q корней многочлена f, заключаем, что хотя бы один из них — назовём его $\xi \in \mathbb{F}_q$ — является корнем и для g. Тогда правило $h(x) \,(\mathrm{mod}\ g) \mapsto h(\xi)$ корректно задаёт вложение полей $\mathrm{ev}_\xi :\ \mathbb{F}_p[x]/(g) \hookrightarrow \mathbb{F}_q$, сюръективное, поскольку оба поля состоят из q элементов. Тем самым, $\mathbb{F}_p[x]/(g) \simeq \mathbb{F}_q$. □

3.5.2. Квадратичные вычеты. Зафиксируем целое простое $p > 2$. Ненулевые элементы поля $\mathbb{F}_p$, являющиеся квадратами, называются *квадратичными вычетами* по модулю p. Они образуют в $\mathbb{F}_p^*$ мультипликативную подгруппу — образ мультипликативного гомоморфизма возведения в квадрат $\mathbb{F}_p^* \to \mathbb{F}_p^*$, $x \mapsto x^2$. Ядро этого гомоморфизма состоит из двух элементов, поскольку уравнение $x^2 = 1$ имеет в поле $\mathbb{F}_p$ ровно два корня $x = \pm 1$. Тем самым, квадратичных вычетов имеется ровно $(p-1)/2$. Судить о том, является ли данный элемент $a \in \mathbb{F}_p^*$ квадратом, можно при помощи мультипликативного гомоморфизма возведения в степень $(p-1)/2$:

$$\mathbb{F}_p^* \to \mathbb{F}_p^*, \quad x \mapsto x^{(p-1)/2}. \tag{3-19}$$

По малой теореме Ферма[1], каждый лежащий в образе этого гомоморфизма элемент $x = a^{(p-1)/2}$ удовлетворяет уравнению $x^2 = a^{p-1} = 1$ и стало быть равен ± 1. С другой стороны образ гомоморфизма (3-19) отличен от 1, поскольку уравнение $x^{(p-1)/2} = 1$ имеет не более $(p-1)/2$ корней в поле $\mathbb{F}_p$. Тем самым, ядро гомоморфизма (3-19) состоит ровно из $(p-1)/2$ элементов и совпадает с подгруппой квадратов, т. е. $a \in \mathbb{F}_p^*$ является квадратом если и только если $a^{(p-1)/2} = 1$. Например, -1 является квадратом в $\mathbb{F}_p$ в точности тогда, когда $(p-1)/2$ чётно.

Для произвольного $n \in \mathbb{N}$ и простого $p > 2$ число

$$\left(\frac{n}{p}\right) \stackrel{\text{def}}{=} [n]_p^{(p-1)/2} = \begin{cases} 1 & \text{когда } n \text{ ненулевой квадрат по модулю } p \\ 0 & \text{когда } n \vdots p \\ -1 & \text{когда } n \text{ не является квадратом по модулю } p \end{cases} \tag{3-20}$$

называется *символом Лежандра – Якоби*. Из определения очевидно, что он зависит только от класса $[n]_p \in \mathbb{Z}/(p)$ и мультипликативен по n:

$$\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right) \cdot \left(\frac{n}{p}\right).$$

Упражнение 3.28*. Покажите, что для простого $p > 2$ символ $\left(\frac{2}{p}\right) = 1$ тогда и только тогда, когда $p \equiv \pm 1 \pmod 8$.

В общем случае символ Лежандра – Якоби легко вычисляется благодаря следующей замечательной теореме, открытой Гауссом.

Теорема 3.5 (квадратичный закон взаимности Гаусса)
Для любых простых $p, q > 2$ выполняется равенство

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Два доказательства этой теоремы, предложенные, соответственно, Эйзенштейном и Золотарёвым, намечены в листке $3\frac{1}{2}$. Вот пример того, как эта теорема работает:

$$\left(\frac{57}{179}\right) = \left(\frac{179}{57}\right) = \left(\frac{8}{57}\right) = \left(\frac{2}{57}\right)^3 = 1,$$

т. е. 57 это квадрат по модулю 179.

[1] См. сл. 2.1 на стр. 27.

§4. Рациональные функции и степенные ряды

В этом параграфе мы продолжаем обозначать через K произвольное коммутативное кольцо с единицей, а через $\Bbbk$ — произвольное поле.

4.1. Кольца частных. Конструкция, изготавливающая поле $\mathbb{Q}$ из кольца $\mathbb{Z}$ как множество дробей с целым числителем и целым ненулевым знаменателем[1], имеет смысл в любом коммутативном кольце K с единицей. Будем называть подмножество $S \subset K$ *мультипликативным*, если

$$1 \in S\,, \quad 0 \notin S \quad \text{и} \quad st \in S \text{ для любых } s, t \in S.$$

Например, если элемент $q \in K$ не является нильпотентным, то множество всех его целых неотрицательных степеней q^k мультипликативно[2]. Множество $K^\circ \subset K$, состоящее из всех ненулевых элементов, которые не являются делителями нуля, также мультипликативно. В частности, множество всех ненулевых элементов любого целостного кольца мультипликативно. Свяжем с каждым мультипликативным подмножеством $S \subset K$ наименьшее отношение эквивалентности $\underset{S}{\sim}$ на множестве упорядоченных пар $K \times S$, содержащее все эквивалентности вида $(a, t) \sim (as, ts)$ с произвольными $s \in S$. Будем называть полученные классы эквивалентности *дробями со знаменателями из* S и обозначать a/s. Множество всех дробей со знаменателями в S обозначим KS^{-1} или $K[S^{-1}]$ и назовём *кольцом частных* (или *локализацией*) кольца K со знаменателями в S.

Лемма 4.1
$a/r = b/t$ в $KS^{-1} \iff \exists\, s \in S : ats = brs$ в K.

Доказательство. Будем писать $(a, r) \approx (b, t)$, если $ats = brs$ для некоторого $s \in S$. Двушаговая цепочка отождествлений: $(a, r) \sim (ats, rts) = (brs, rts) \sim (b, t)$ показывает, что отношение $\approx$ содержится в отношении $\underset{S}{\sim}$. Остаётся проверить, что отношение $\approx$ является отношением эквивалентности — тогда оно совпадёт с $\underset{S}{\sim}$ в виду минимальности последнего. Рефлексивность и симметричность очевидны. Докажем транзитивность. Пусть $(a, r) \approx (b, t)$ и $(b, t) \approx (c, u)$, т. е. существуют такие $s_1, s_2 \in S$, что $ats_1 = brs_1$ и $bus_2 = cts_2$. Тогда

$$au(ts_1s_2) = brus_1s_2 = cr(ts_1s_2)\,,$$

т. е. $(a, r) \approx (c, u)$. □

Лемма 4.2
Операции $\frac{a}{r} + \frac{b}{s} \overset{\text{def}}{=} \frac{as+br}{rs}$ и $\frac{a}{r} \cdot \frac{b}{s} \overset{\text{def}}{=} \frac{ab}{rs}$ корректно задают на KS^{-1} структуру коммутативного кольца с единицей $1/1$ и нулём $0/1$.

Доказательство. Поскольку всякое отношение $\underset{S}{\sim}$ представляет собой одно- или двушаговую цепочку элементарных отождествлений $(a, r) \sim (au, ru)$, где $u \in S$, достаточно проверить, что результаты операций не меняются при замене $\frac{a}{r}$ на $\frac{au}{ru}$, а $\frac{b}{s}$ — на $\frac{bw}{sw}$, где $u, w \in S$:

$$\frac{au}{ru} + \frac{bw}{sw} = \frac{ausw + bwru}{rusw} = \frac{(as + br) \cdot wu}{rs \cdot wu} = \frac{as + br}{rs}$$
$$\frac{au}{ru} \cdot \frac{bw}{sw} = \frac{aubw}{rusw} = \frac{(ab) \cdot wu}{rs \cdot wu} = \frac{ab}{rs}\,.$$

[1] См. прим. 1.5 на стр. 13 и прим. 2.2 на стр. 21.
[2] Мы по определению полагаем $q^0 = 1$.

Проверку выполнения в KS^{-1} всех аксиом коммутативного кольца с единицей мы оставляем читателю в качестве упражнения. □

Теорема 4.1
Отображение $\iota_S : K \to KS^{-1}$, переводящее $a \in K$ в дробь $a/1$, является гомоморфизмом колец с ядром $\ker \iota_S = \{a \in K \mid \exists\, s \in S : as = 0\}$. Все элементы $\iota_S(s)$ с $s \in S$ обратимы в KS^{-1}. Для любого гомоморфизма $\varphi : K \to R$ в целостное кольцо R, переводящего все $s \in S$ в обратимые элементы кольца R, существует единственный такой гомоморфизм колец $\varphi_S : KS^{-1} \to R$, что $\varphi = \varphi_S \circ \iota_S$.

Доказательство. Очевидно, что ι_S является гомоморфизмом. Дробь $\iota_S(a) = a/1$ равна $0/1$ если и только если найдётся такой $s \in S$, что $a \cdot 1 \cdot s = 0 \cdot 1 \cdot s = 0$. Обратной к дроби $\iota_S(s) = s/1$ является дробь $1/s$. Остаётся доказать последнее утверждение. Для продолжения гомоморфизма $\varphi : K \to R$ до гомоморфизма $\varphi_S : KS^{-1} \to R$ нет иного выбора как положить $\varphi_S(1/s) = 1/\varphi(s)$, так как в кольце R должны выполняться равенства $\varphi_S(1/s) \cdot \varphi_S(s) = \varphi_S\big(s \cdot (1/s)\big) = \varphi(1) = 1$. Следовательно, искомое продолжение обязано задаваться формулой $\varphi_S(a/r) \stackrel{\text{def}}{=} \varphi(a)/\varphi(r)$. Она корректна, поскольку при замене $\frac{a}{r}$ на $\frac{as}{rs}$ с $s \in S$ имеем $\varphi_S\left(\frac{as}{rs}\right) = \frac{\varphi(as)}{\varphi(rs)} = \frac{\varphi(a)\,\varphi(s)}{\varphi(r)\,\varphi(s)} = \frac{\varphi(a)}{\varphi(r)}$. Проверка того, что построенное отображение φ_S перестановочно со сложением и умножением, столь же бесхитростна, и мы оставляем её читателю. □

Замечание 4.1. Кольцо KS^{-1} и гомоморфизм $\iota_S : K \to KS^{-1}$ однозначно определяются последним свойством из теор. 4.1. В самом деле, пусть гомоморфизм $\iota' : K \to F$ делает все элементы из S обратимыми в F и обладает универсальным свойством из теор. 4.1, т. е. для любого гомоморфизма $\varphi : K \to R$ в целостное кольцо R, делающего все элементы из S обратимыми в R, существует единственный такой гомоморфизм колец $\varphi'_S : F \to R$, что $\varphi = \varphi'_S \circ \iota'$. Тогда существует единственный изоморфизм колец $\psi : KS^{-1} \xrightarrow{\sim} F$, превращающий ι_S в ι' в том смысле, что $\iota' = \psi \circ \iota_S$. Действительно, в силу универсальности гомоморфизма ι_S гомоморфизм ι' единственным образом представляется в виде $\iota' = \psi \circ \iota_S$, а в силу универсальности гомоморфизма ι' гомоморфизм ι_S точно так же единственным образом представляется в виде $\iota_S = \psi' \circ \iota'$. Композиция $\psi' \circ \psi : KS^{-1} \to KS^{-1}$ доставляет разложение самого гомоморфизма $\iota_S : K \to KS^{-1}$ в композицию $\iota_S = (\psi' \circ \psi) \circ \iota_S$. Но такое же разложение можно осуществить при помощи тождественного изоморфизма: $\iota_S = \mathrm{Id}_{KS^{-1}} \circ \iota_S$. Из единственности разложения вытекает равенство $\psi' \circ \psi = \mathrm{Id}_{KS^{-1}}$. По той же причине $\psi \circ \psi' = \mathrm{Id}_F$, т. е. ψ' и ψ являются взаимно обратными изоморфизмами.

Замечание 4.2. Если в определении мультипликативной системы отбросить требование $0 \notin S$, то всё сказанное выше не утратит формального смысла: эквивалентность $\underset{S}{\sim}$ и кольцо KS^{-1} будут по-прежнему определены, а лем. 4.1, лем. 4.2, теор. 4.1 и их доказательства останутся в силе. Однако, если $0 \in S$, кольцо KS^{-1} получится нулевым: любая дробь $a/s = (a \cdot 0)/(s \cdot 0) = 0/0 = (0 \cdot 1)/(0 \cdot 1) = 0/1$ эквивалентна нулю.

Пример 4.1 (поле частных целостного кольца)
Если кольцо K не имеет делителей нуля, его ненулевые элементы образуют мультипликативную систему. Кольцо частных со знаменателями в этой системе является полем и называется *полем частных* целостного кольца K и обозначается Q_K. Гомоморфизм $\iota : K \hookrightarrow Q_K$, $a \mapsto a/1$

в этом случае инъективен, и любой гомоморфизм $\varphi: K \to R$ в целостное кольцо R, переводящий все ненулевые элементы из K в обратимые элементы кольца R, единственным способом продолжается до вложения поля частных $\widetilde{\varphi}: Q_K \hookrightarrow R$.

Пример 4.2 (поле $\mathbb{Q}$)
Полем частных целостного кольца $\mathbb{Z}$ является поле рациональных чисел $\mathbb{Q} = Q_{\mathbb{Z}}$, которое канонически вкладывается в любое поле характеристики нуль в качестве простого подполя[1].

Пример 4.3 (поле рядов Лорана)
Поле частных кольца формальных степенных рядов $\Bbbk[\![x]\!]$ с коэффициентами в произвольном поле $\Bbbk$ называется полем *рядов Лорана* и обозначается $\Bbbk((x)) \stackrel{\text{def}}{=} Q_{\Bbbk[\![x]\!]}$. Название «ряд Лорана» объясняется тем, что каждый элемент $f \in \Bbbk((x))$ можно записать как формальный степенной ряд, в котором допускается конечное число отрицательных степеней переменной x

$$f(x) = \sum_{k \geqslant -m} a_k x^k = x^{-m} h(x), \quad \text{где} \quad h \in \Bbbk[\![x]\!]. \tag{4-1}$$

В самом деле, по определению поля частных $f(x) = p(x)/q(x)$, где $p, q \in \Bbbk[\![x]\!]$ и $q \neq 0$. Если младший член ряда q имеет степень m, то $q = x^m \cdot g(x)$, где $g \in \Bbbk[\![x]\!]$ имеет ненулевой свободный член и, стало быть обратим. Поэтому мы можем записать исходную дробь в виде $f(x) = x^{-m} h(x)$, где $h = p/g \in \Bbbk[\![x]\!]$ является обычным степенным рядом.

4.2. Поле рациональных функций. Поле частных кольца многочленов $\Bbbk[x]$ обозначается через $\Bbbk(x)$ и называется *полем рациональных функций* от одной переменной. Элементы этого поля представляют собой формальные отношения многочленов $f(x) = p(x)/q(x)$ с коэффициентами в поле $\Bbbk$. Деля числитель и знаменатель на $\text{нод}(p, q)$ и на старший коэффициент знаменателя, мы можем записать произвольную дробь в виде отношения двух взаимно простых многочленов с приведённым знаменателем. Мы будем называть такую запись *несократимым представлением* дроби f.

Упражнение 4.1. Покажите, что несократимая запись любой дроби единственна.

Предложение 4.1
Если знаменатель несократимой записи f/g является произведением попарно взаимно простых многочленов $g = g_1 \dots g_m$, то дробь f/g единственным образом представляется в виде суммы

$$\frac{f}{g} = h + \frac{f_1}{g_1} + \frac{f_2}{g_2} + \dots + \frac{f_m}{g_m}, \tag{4-2}$$

в которой $\deg h = \deg f - \deg g$ и $\deg f_i < \deg g_i$.

Доказательство. Поделим f на g с остатком: $f = hg + r$, где $\deg r < \deg g$. Тогда $f/g = h + r/g$. Если $g = g_1 g_2$ и $\text{нод}(g_1, g_2) = 1$, то класс $[g_2]_{g_1}$ многочлена g_2 в кольце вычетов $\Bbbk[x]/(g_1)$ обратим и отношение $[r]_{g_1}/[g_2]_{g_1}$ представляется в $\Bbbk[x]/(g_1)$ классом некоторого многочлена f_1 степени $\deg f_1 < \deg g_1$, т. е. в $\Bbbk[x]$ мы имеем равенство $r = f_1 \cdot g_2 + f_2 \cdot g_1$ для некоторого многочлена f_2, и сравнение степеней показывает, что $\deg f_2 < \deg g_2$, коль скоро $\deg f_1 < \deg g_1$. Таким образом, $r/g = f_1/g_1 + f_2/g_2$, и с каждой из этих дробей можно и далее проделывать

[1] См. n° 2.8.1 на стр. 32.

аналогичные процедуры до тех пор, пока знаменатели раскладываются в произведение взаимно простых многочленов. Это доказывает существование разложения (4-2). Чтобы доказать его единственность, умножим обе части произвольного разложения (4-2) на g. Получим равенство

$$f = hg + f_1 G_1 + \cdots + f_m G_m\,,$$

где $G_i = g\,/\,g_i = g_1 \ldots g_{i-1} g_{i+1} \ldots g_m$ и $\deg(f_1 G_1 + \cdots + f_m G_m) < \deg g$. Тем самым, многочлен h является неполным частным от деления f на g, многочлен $r = f_1 G_1 + \cdots + f_m G_m$ равен остатку от этого деления, а каждый f_i представляет собою единственный многочлен степени $\deg f_i < \deg g_i$, класс которого в кольце вычетов $\Bbbk[x]/(g_i)$ равен $[f]_{g_i} \cdot [G_i]_{g_i}^{-1}$. Таким образом, все ингредиенты формулы (4-2) однозначно определяются многочленами f и $g_1, \ldots, g_n$. □

Предложение 4.2
Любую дробь вида $f\,/\,g^m$, в которой $\deg f < \deg g^m = m \deg g$, можно единственным образом представить в виде суммы

$$\frac{f}{g^m} = \frac{f_1}{g} + \frac{f_2}{g^2} + \cdots + \frac{f_m}{g^m}\,, \tag{4-3}$$

где каждый числитель f_i имеет степень $\deg f_i < \deg g$.

Доказательство. Представление (4-3) равносильно записи f в виде

$$f = f_1 g^{m-1} + f_2 g^{m-2} + \cdots + f_{m-1} g + f_m\,, \tag{4-4}$$

аналогичном записи целого числа f в g-ичной позиционной системе исчисления: f_m является остатком от деления f на g, f_{m-1} — остатком от деления частного $(f - f_m)\,/\,g$ на g, f_{m-2} — остатком от деления частного $\left(\frac{f-f_m}{g} - f_{m-1}\right)/g$ на g и т. д. □

4.2.1. Разложение на простейшие дроби. Из предыдущих двух лемм вытекает, что любая дробь $f\,/\,g \in \Bbbk(x)$ допускает *единственное* представление в виде суммы многочлена степени $\deg f - \deg g$ (неполного частного от деления f на g) и дробей вида $p\,/\,q^m$, где q пробегает множество неприводимых делителей знаменателя, m меняется от 1 до кратности вхождения неприводимого множителя q в разложение многочлена g на неприводимые множители, а каждый числитель p имеет степень $\deg p < \deg q$. Такое представление называется *разложением* дроби $f\,/\,g$ *на простейшие дроби* и часто оказывается полезным при вычислениях с рациональными функциями.

Пример 4.4
Вычислим первообразную[1] и 2013-ю производную от $1/(1+x^2)$. Для этого разложим эту дробь в сумму простейших в поле $\mathbb{C}(x)$:

$$\frac{1}{1+x^2} = \frac{\alpha}{1+ix} + \frac{\beta}{1-ix}\,, \quad \text{где} \quad \alpha, \beta \in \mathbb{C}\,.$$

Подставляя $x = \pm i$ в равенство $1 = \alpha(1-ix) + \beta(1+ix)$, находим $\alpha = \beta = 1/2$, т. е.

$$\frac{1}{1+x^2} = \frac{1}{2}\left(\frac{1}{1+ix} + \frac{1}{1-ix}\right).$$

[1] Точное алгебраическое определение первообразной от степенного ряда см. в n° 4.4 на стр. 58.

Теперь уже легко вычислить как 2013-ю производную:

$$\left(\frac{d}{dx}\right)^{2013}\frac{1}{1+x^2}=\frac{2013!}{2}\left(\frac{(-i)^{2013}}{(1+ix)^{2014}}+\frac{i^{2013}}{(1-ix)^{2014}}\right)=$$
$$=\frac{i}{2}\cdot 2013!\cdot\frac{(1+ix)^{2014}-(1-ix)^{2014}}{(1+x^2)^{2014}}=2013!\cdot\sum_{\nu=0}^{1006}\binom{2014}{2\nu+1}\cdot\frac{x^{2\nu+1}}{(1+x^2)^{2014}},$$

так и первообразную:

$$\int\frac{dx}{1+x^2}=\frac{1}{2}\int\frac{dx}{1+ix}+\frac{1}{2}\int\frac{dx}{1-ix}=\frac{1}{2i}\bigl(\ln(1+ix)-\ln(1-ix)\bigr)=\frac{1}{2i}\ln\frac{1+ix}{1-ix}.$$

Подчеркнём, что все проделанные вычисления корректно определены в кольце $\mathbb{C}[\![x]\!]$ и все написанные равенства суть равенства между элементами этого кольца. О том, что такое логарифм и первообразная в кольце $\mathbb{C}[\![x]\!]$, мы ещё подробно поговорим ниже[1].

4.3. Разложение рациональных функций в степенные ряды. В силу универсального свойства поля частных, поле рациональных функций $\Bbbk(x)$ единственным образом вкладывается в поле рядов Лорана $\Bbbk((x))$ так, что при этом многочлены переходят в многочлены. С практической точки зрения это вложение представляет собою разложение рациональных функций f/g в формальные степенные ряды. Если основное поле $\Bbbk$ алгебраически замкнуто, такое разложение можно описать довольно явными формулами. Пусть $\deg f < \deg g$ и знаменатель дроби f/g имеет вид:

$$g(x)=1+a_1x+a_2x^2+\cdots+a_nx^n=\prod(1-\alpha_ix)^{m_i}, \tag{4-5}$$

где все числа $\alpha_i\in\Bbbk$ попарно различны.

Упражнение 4.2. Убедитесь, что при $a_n\neq 0$ числа α_i из разложения (4-5) суть корни многочлена $t^n+a_1t^{n-1}+\cdots+a_{n-1}t+a_n=\prod(t-\alpha_i)^{m_i}$.

По предл. 4.1 и предл. 4.2 функция f/g является суммой простейших дробей вида

$$\frac{\beta_{ij}}{(1-\alpha_ix)^{k_{ij}}} \tag{4-6}$$

где при каждом i показатели k_{ij} лежат в пределах $1\leqslant k_{ij}\leqslant m_i$, а $\beta_{ij}\in\Bbbk$. Если все кратности $m_i=1$, то константы β_i в получающемся разложении

$$\frac{f(x)}{(1-\alpha_1x)(1-\alpha_2x)\cdots(1-\alpha_nx)}=\frac{\beta_1}{1-\alpha_1x}+\frac{\beta_2}{1-\alpha_2x}+\cdots+\frac{\beta_n}{1-\alpha_nx} \tag{4-7}$$

легко указать явно: умножая обе части (4-7) на знаменатель и беря $x=\alpha_i^{-1}$, получаем

$$\beta_i=\frac{f\left(\alpha_i^{-1}\right)}{\prod\limits_{\nu\neq i}\bigl(1-(\alpha_\nu/\alpha_i)\bigr)}=\frac{\alpha_i^{n-1}f\left(\alpha_i^{-1}\right)}{\prod\limits_{\nu\neq i}\left(\alpha_i-\alpha_\nu\right)}. \tag{4-8}$$

Дробь f/g в этом случае равна сумме геометрических прогрессий (4-7)

$$\frac{f(x)}{g(x)}=\sum\left(\beta_1\alpha_1^k+\beta_2\alpha_2^k+\cdots+\beta_n\alpha_n^k\right)\cdot x^k.$$

[1] См. n° 4.4 на стр. 58. Отметим, что $\frac{1}{2i}\ln\frac{1+ix}{1-ix}=\operatorname{arctg}x$, поскольку $\operatorname{tg}(t)=\frac{\sin t}{\cos t}=\frac{1}{i}\cdot\frac{e^{it}-e^{-it}}{e^{it}+e^{-it}}=\frac{1}{i}\cdot\frac{e^{2it}-1}{e^{2it}+1}$.

Если в простейшей дроби (4-6) показатель $k_{ij} = m > 1$, то она раскладывается в ряд при помощи формулы Ньютона для *бинома с отрицательным показателем*

$$\frac{1}{(1-x)^m} = \sum_{k\geqslant 0} \frac{(k+m-1)(k+m-2)\,\cdots\,(k+1)}{(m-1)!}\cdot x^k = \sum_{k\geqslant 0}\binom{k+m-1}{m-1}\cdot x^k\,, \tag{4-9}$$

которая получается $(m-1)$-кратным дифференцированием обеих частей разложения геометрической прогрессии $(1-x)^{-1} = 1 + x + x^2 + x^3 + x^4 + \ldots$.

Упражнение 4.3. Убедитесь, что $\left(\frac{d}{dx}\right)^m (1-x)^{-1} = m!\,/\,(1-x)^{m+1}$.

Таким образом, разложение простейшей дроби (4-6) имеет вид

$$\frac{\beta}{(1-\alpha_i x)^m} = \beta\sum_{k\geqslant 0}\alpha_i^k\binom{k+m-1}{m-1}\cdot x^k\,. \tag{4-10}$$

4.3.1. Решение линейных рекуррентных уравнений. Предыдущие вычисления можно использовать для отыскания «формулы k-того члена» последовательности z_k, заданной *линейным рекуррентным уравнением n-того порядка*:

$$z_k + a_1 z_{k-1} + a_2 z_{k-2} + \;\cdots\; + a_n z_{k-n} = 0\,, \tag{4-11}$$

где коэффициенты $a_1, \ldots, a_n \in \mathbb{C}$ — некоторые фиксированные заданные числа. При $k \geqslant n$ уравнению (4-11) удовлетворяют коэффициенты z_k степенного ряда

$$\frac{b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}}{1 + a_1 x + a_2 x^2 + \;\cdots\; + a_n x^k} = z_0 + z_1 x + z_2 x^2 + \ldots$$

Если подобрать $b_0, b_1, \ldots, b_{n-1} \in \mathbb{C}$ в числителе левой части так, чтобы первые n коэффициентов справа совпадали с начальным куском последовательности (4-11), и разложить полученную рациональную функцию в ряд, то мы получим явные выражения элементов последовательности z_k через k.

Пример 4.5 (числа Фибоначчи)
Найдём явное выражение через k для элементов последовательности

$$z_0 = 0\,,\quad z_1 = 1\,,\quad z_k = z_{k-1} + z_{k-2}\quad \text{при}\quad k \geqslant 2\,,$$

решающей рекуррентное уравнение $z_k - z_{k-1} - z_{k-2} = 0$ на коэффициенты ряда

$$\frac{b_0 + b_1 x}{1 - x - x^2} = x + z_2 x^2 + z_3 x^3 + \cdots \tag{4-12}$$

(мы подставили в правую часть данные по условию $z_0 = 0$ и $z_1 = 1$). Умножая обе части (4-12) на общий знаменатель и сравнивая коэффициенты при x^0 и x^1, получаем $b_0 = 0$ и $b_1 = 1$. Итак, нас интересуют коэффициенты ряда

$$z(x) = \frac{x}{1-x-x^2} = \frac{\beta_+}{1-\alpha_+ x} + \frac{\beta_-}{1-\alpha_- x}\,,$$

где $\alpha_\pm = (1 \pm \sqrt{5})/2$ суть корни многочлена $t^2 - t - 1$, а числа $\beta_\pm$ находятся по формуле (4-8) с учётом равенств $\alpha_+\alpha_- = -1$, $\alpha_+ + \alpha_- = 1$ и $\alpha_+ - \alpha_- = \sqrt{5}$: $\beta_+ = -\beta_- = 1/(\alpha_+ - \alpha_-) = 1/\sqrt{5}$. Получаем:

$$\frac{x}{1-x-x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-\alpha_+ x} - \frac{1}{1-\alpha_- x}\right) = \sum_{k\geqslant 0} \frac{\alpha_+^k - \alpha_-^k}{\sqrt{5}} \cdot x^k\,,$$

откуда

$$z_k = \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k}{2^k\sqrt{5}}\,.$$

Предложение 4.3
Всякая последовательность z_k, удовлетворяющая при $k \geqslant n$ линейному рекуррентному уравнению n-того порядка

$$z_k + a_1 z_{k-1} + a_2 z_{k-2} + \ \cdots\ + a_n z_{k-n} = 0 \tag{4-13}$$

с постоянными коэффициентами $a_i \in \mathbb{C}$, имеет вид

$$z_k = \alpha_1^k \cdot \varphi_1(k) + \alpha_2^k \cdot \varphi_2(k) + \ \cdots\ + \alpha_r^k \cdot \varphi_r(k)\,,$$

где $\alpha_1, \ldots, \alpha_r$ суть все различные корни многочлена[1]

$$t^n + a_1 t^{n-1} + \cdots + a_{n-1} t + a_n\,, \tag{4-14}$$

а каждая из функций $\varphi_i \in \mathbb{C}[x]$ представляет собою многочлен степени на единицу меньшей, чем кратность соответствующего корня α_i.

Доказательство. Ряд $\sum z_k x^k \in \mathbb{C}[\![x]\!]$, коэффициенты которого решают уравнение (4-13), является суммой дробей вида $\beta \cdot (1-\alpha x)^{-m}$, где α пробегает различные корни многочлена (4-14), показатель m может принимать любое значение от 1 до кратности соответствующего корня α, и для каждой пары α, m комплексное число $\beta = \beta(\alpha, m)$ однозначно вычисляется по α, m и первым n коэффициентам последовательности z_k. Согласно формуле (4-10) коэффициент при x^k у разложения дроби $(1-\alpha x)^{-m}$ в степенной ряд имеет вид $\alpha^k \varphi(k)$, где $\varphi(k) = \binom{k+m-1}{m-1}$ является многочленом степени $m-1$ от k. □

4.4. Логарифм и экспонента. Всюду в этом разделе мы рассматриваем ряды с коэффициентами в поле $\Bbbk$ характеристики $\operatorname{char}\Bbbk = 0$. В этом случае из формулы (3-7) для производной вытекает, что для любого ряда $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots$ существует единственный ряд без свободного члена, производная от которого равна $f(x)$. Этот ряд называется *первообразным рядом* или *интегралом* от f и обозначается

$$\int f(x)\,dx \ \stackrel{\text{def}}{=}\ a_0 x + \frac{a_1}{2}\,x^2 + \frac{a_2}{3}\,x^3 + \ \cdots\ = \sum_{k\geqslant 1} \frac{a_{k-1}}{k}\,x^k\,. \tag{4-15}$$

[1] Он называется *характеристическим многочленом* рекуррентного уравнения (4-11).

Определение 4.1
Первообразный ряд от знакопеременной геометрической прогрессии называется *логарифмом* и обозначается

$$\ln(1+x) \stackrel{\text{def}}{=} \int \frac{dx}{1+x} = \int \left(1 - x + x^2 - x^3 + \cdots\right) dx = \\ = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots = \sum_{k\geqslant 1} \frac{(-1)^{k-1}}{k} x^k . \quad (4\text{-}16)$$

4.4.1. Логарифмирование рядов. Обозначим через $N = x \cdot \Bbbk[\![x]\!] \subset \Bbbk[\![x]\!]$ аддитивную абелеву группу всех рядов без свободного члена, а через $U = 1 + N \subset \Bbbk[\![x]\!]$ — мультипликативную абелеву группу всех рядов с единичным свободным членом. Подстановка в аргумент логарифма вместо $1+x$ произвольного ряда $u(x)$ с единичным свободным членом является алгебраической операцией, поскольку означает подстановку в логарифмический ряд (4-16) вместо переменной x ряда $u(x) - 1$ без свободного члена, а это, как мы видели[1], алгебраическая операция. Таким образом, имеется отображение *логарифмирования*

$$\ln : U \to N , \quad u \mapsto \ln u . \quad (4\text{-}17)$$

Упражнение 4.4 (логарифмическая производная). Убедитесь, что $\frac{d}{dx} \ln u = u' / u$ для всех рядов $u \in U$.

Лемма 4.3
Для рядов $u, w \in U$ равенства $u = w$, $u' = w'$, $\ln(u) = \ln(w)$ и $u'/u = w'/w$ попарно эквивалентны друг другу.

Доказательство. Первое равенство влечёт за собой все остальные. Поскольку ряды с равными свободными членами совпадают если и только если совпадают их производные, первые два равенства и последние два равенства равносильны друг другу. Остаётся показать, что из последнего равенства следует первое. Но последнее равенство утверждает, что $u'/u - w'/w = (u'w - w'u)/uw = (w/u) \cdot (u/w)' = 0$ откуда $(u/w)' = 0$, т. е. $u/w = \text{const} = 1$. □

Упражнение 4.5. Покажите, что $\forall\, u \in U \ \ln(1/u) = -\ln u$.

Определение 4.2
Ряд $e^x \stackrel{\text{def}}{=} \sum\limits_{k\geqslant 0} x^k/k! = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \ldots$ называется *экспонентой*. Это единственный ряд в U, удовлетворяющий дифференциальному уравнению $f'(x) = f(x)$.

4.4.2. Экспоненцирование рядов. Подставляя в экспоненту вместо x любой ряд $\tau(x)$ без свободного члена, мы получаем ряд $e^{\tau(x)}$ со свободным членом 1, который называется *экспонентой* ряда $\tau(x)$. Этим определяется экспоненциальное отображение

$$\exp : N \to U , \quad \tau \mapsto e^{\tau} . \quad (4\text{-}18)$$

[1] См. n° 3.1.1 на стр. 33.

Теорема 4.2
Экспоненциальное и логарифмическое отображения (4-18) и (4-17) являются взаимно обратными изоморфизмами абелевых групп, т. е. для любых рядов u, u_1, u_2 из U и τ, τ_1, τ_2 из N выполняются тождества:

$$\ln e^{\tau} = \tau\,, \quad e^{\ln u} = u\,, \quad \ln(u_1u_2) = \ln(u_1) + \ln(u_2)\,, \quad e^{\tau_1+\tau_2} = e^{\tau_1}e^{\tau_2}\,.$$

Доказательство. Равенство $\ln e^{\tau} = \tau$ проверяется сравнением производных от обеих частей:

$$\bigl(\ln e^{\tau}\bigr)' = \frac{\bigl(e^{\tau}\bigr)'}{e^{\tau}} = \frac{e^{\tau}\tau'}{e^{\tau}} = \tau'\,,$$

а равенство $e^{\ln u} = u$ — сравнением логарифмических производных:

$$\frac{\bigl(e^{\ln u}\bigr)'}{e^{\ln u}} = \frac{e^{\ln u}(\ln u)'}{e^{\ln u}} = \frac{u'}{u}\,.$$

Тем самым, экспоненцирование и логарифмирование являются взаимно обратными биекциями. Ряды $\ln(u_1u_2)$ и $\ln u_1 + \ln u_2$ совпадают, поскольку имеют нулевые свободные члены и равные производные:

$$\bigl(\ln(u_1u_2)\bigr)' = \frac{(u_1u_2)'}{u_1u_2} = \frac{u_1'u_2 + u_1u_2'}{u_1u_2} = \frac{u_1'}{u_1} + \frac{u_2'}{u_2} = \bigl(\ln u_1 + \ln u_2\bigr)'\,.$$

Поэтому логарифмирование — гомоморфизм, а значит, и обратное к нему экспоненцирование — тоже. □

Упражнение 4.6. Докажите в $\Bbbk[\![x, y]\!]$ равенство $e^{x+y} = e^x e^y$ непосредственным сравнением коэффициентов этих двух рядов.

4.5. Степенная функция и бином Ньютона. В этом разделе мы продолжаем считать, что поле $\Bbbk$ имеет характеристику нуль. Для любого числа $\alpha \in \Bbbk$ определим *биномиальный ряд* с показателем α формулой

$$(1+x)^{\alpha} \stackrel{\text{def}}{=} e^{\alpha \ln(1+x)}.$$

Подставляя вместо $1+x$ произвольные ряды $u \in U$, мы для любого числа $\alpha \in \Bbbk$ получаем алгебраическую операцию $U \to U$ *возведения в α-тую степень* $u \mapsto u^{\alpha}$, обладающую всеми интуитивно ожидаемыми от степенной функции свойствами. В частности, для любых рядов $u, v \in U$ и чисел $\alpha, \beta \in \Bbbk$ выполняются равенства

$$u^{\alpha} \cdot u^{\beta} = e^{\alpha \ln u} \cdot e^{\beta \ln u} = e^{\alpha \ln u + \beta \ln u} = e^{(\alpha+\beta)\ln u} = u^{\alpha+\beta} \tag{4-19}$$

$$\bigl(u^{\alpha}\bigr)^{\beta} = e^{\beta \ln(u^{\alpha})} = e^{\beta \ln\left(e^{\alpha \ln u}\right)} = e^{\alpha\beta \ln u} = u^{\alpha\beta} \tag{4-20}$$

$$(uv)^{\alpha} = e^{\alpha \ln(uv)} = e^{\alpha(\ln u + \ln v)} = e^{\alpha \ln u + \alpha \ln v} = e^{\alpha \ln u} \cdot e^{\alpha \ln v} = u^{\alpha} v^{\alpha} \tag{4-21}$$

Например, для любого ряда u с единичным свободным членом ряд $u^{1/n}$ представляет собою $\sqrt[n]{u}$ в том смысле, что $\bigl(u^{1/n}\bigr)^n = u$. Для явного отыскания коэффициентов a_i биномиального ряда

$$(1+x)^{\alpha} = a_0 + a_1 x + a_2 x^2 + \ldots$$

вычислим его логарифмическую производную:

$$\frac{\bigl((1+x)^\alpha\bigr)'}{(1+x)^\alpha} = \bigl(\ln(1+x)^\alpha\bigr)' = \bigl(\alpha\ln(1+x)\bigr)' = \frac{\alpha}{1+x}\,.$$

Приводя левую и правую часть к общему знаменателю, получаем равенство

$$\bigl(a_1 + 2a_2x + 3a_3x^2 + \,\cdots\,\bigr)\cdot(1+x) = \alpha\cdot(1 + a_1x + a_2x^2 + a_3x^3 + \,\cdots\,)\,.$$

Сравнивая коэффициенты при x^{k-1} в правой и левой части, приходим к рекуррентному соотношению $ka_k + (k-1)a_{k-1} = \alpha a_{k-1}$, из которого

$$a_k = \frac{\alpha-(k-1)}{k}\cdot a_{k-1} = \frac{(\alpha-(k-1))(\alpha-(k-2))}{k(k-1)}\cdot a_{k-2} = \;\cdots$$
$$\cdots\; = \frac{(\alpha-(k-1))(\alpha-(k-2))\cdots(\alpha-1)\alpha}{k!}\,.$$

Стоящая в правой части дробь имеет в числителе и знаменателе по k множителей, представляющих собою последовательно уменьшающиеся на единицу числа: в знаменателе — от k до 1, в числителе — от α до $(\alpha-k+1)$. Эта дробь называется *биномиальным коэффициентом* и обозначается

$$\binom{\alpha}{k} \stackrel{\text{def}}{=} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} \tag{4-22}$$

Нами доказано

Предложение 4.4 (формула Ньютона)
Для любого числа $\alpha\in\Bbbk$ имеется разложение

$$(1+x)^\alpha = \sum_{k\geqslant 0}\binom{\alpha}{k}x^k = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{6}x^3 + \,\cdots\,.$$

Пример 4.6 (бином с рациональным показателем)
При натуральном значении показателя $\alpha = n\in\mathbb{N}$ имеется лишь конечное число ненулевых биномиальных коэффициентов, поскольку при $k > n$ в числителе (4-22) образуется нулевой сомножитель. Поэтому разложение бинома в этом случае конечно:

$$(1+x)^n = 1 + n\,x + \frac{n(n-1)}{2}\,x^2 + \,\cdots\, + x^n = \sum_{k=0}^{n}\binom{n}{k}\cdot x^k\,.$$

Оно уже встречалось нам в форм. (1-9) на стр. 9. При целом отрицательном $\alpha = -m$, где $m\in\mathbb{N}$, мы получаем разложение из форм. (4-9) на стр. 57:

$$(1+x)^{-m} = 1 - m\,x + \frac{m(m+1)}{2}\,x^2 - \frac{m(m+1)(m+2)}{6}\,x^3 + \,\cdots = \sum_{k\geqslant 0}(-1)^k\binom{k+m-1}{k}\cdot x^k\,.$$

При $\alpha = 1/n$, где $n\in\mathbb{N}$, формула Ньютона разворачивает в степенной ряд радикал

$$\sqrt[n]{1+x} = 1 + \frac{1}{n}\,x + \frac{\frac{1}{n}\left(\frac{1}{n}-1\right)}{2}\,x^2 + \frac{\frac{1}{n}\left(\frac{1}{n}-1\right)\left(\frac{1}{n}-2\right)}{6}\,x^3 + \cdots\; =$$
$$= 1 + \frac{x}{n} - \frac{n-1}{2}\cdot\frac{x^2}{n^2} + \frac{(n-1)(2n-1)}{2\cdot 3}\cdot\frac{x^3}{n^3} - \frac{(n-1)(2n-1)(3n-1)}{2\cdot 3\cdot 4}\cdot\frac{x^4}{n^4} + \cdots$$

Например, при $n = 2$ в качестве коэффициента при x^k мы получаем дробь вида

$$(-1)^{k-1} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2k-3)}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2k)} = \frac{(-1)^{k-1}}{2k-1} \cdot \frac{(2k)!}{(2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2k))^2} = \\ = \frac{(-1)^{k-1}}{(2k-1) \cdot 4^k} \cdot \binom{2k}{k}.$$

Таким образом,

$$\sqrt{1+x} = \sum_{k \geqslant 0} \frac{(-1)^{k-1}}{2k-1} \cdot \binom{2k}{k} \cdot \frac{x^k}{4^k}. \tag{4-23}$$

Пример 4.7 (числа Каталана)
Воспользуемся разложением (4-23) для получения явной формулы для *чисел Каталана*, часто возникающих в различных комбинаторных задачах. Будем вычислять произведение $n+1$ множителей

$$a_0 a_1 a_2 \cdot \ldots \cdot a_n \qquad \text{(всего } n \text{ умножений)} \tag{4-24}$$

делая за один шаг ровно одно умножение. Если на каждом шагу заключать вычисленное произведение в скобки, то в ходе вычисления мы расставим n пар скобок в выражении (4-24). Количество различных расстановок скобок, возникающих таким образом, называется n-ым *числом Каталана* c_n. При $n = 1$ есть лишь одна расстановка скобок: $(a_1 a_2)$, при $n = 2$ — две:

$$(a_1(a_2 a_3)) \quad \text{и} \quad ((a_1 a_2)a_3),$$

при $n = 3$ — пять:

$$(a_1(a_2(a_3 a_4))),\ (a_1((a_2 a_3)a_4)),\ ((a_1 a_2)(a_3 a_4)),\ ((a_1(a_2 a_3))a_4),\ (((a_1 a_2)a_3)a_4).$$

Множество всех возможных расстановок скобок в произведении (4-24) распадается в дизъюнктное объединение n подмножеств, в которых конфигурации наружных скобок имеют вид

$$(a_0(a_2 \ldots a_n)),\ ((a_0 a_1)(a_2 \ldots a_n)),\ ((a_0 \ldots a_2)(a_3 \ldots a_n)),\ ((a_0 \ldots a_3)(a_4 \ldots a_n)),\ \ldots \\ \ldots,\ ((a_0 \ldots a_{n-2})(a_{n-1} a_n)),\ ((a_0 \ldots a_{n-1})a_n)$$

и которые состоят, соответственно, из c_{n-1}, $c_1 c_{n-2}$, $c_2 c_{n-3}$, $c_3 c_{n-4}$, …, $c_{n-2} c_1$, c_{n-1} элементов. Если добавить к числам Каталана число $c_0 \stackrel{\text{def}}{=} 1$, то мы получим рекурсивное соотношение $c_n = c_0 c_{n-1} + c_1 c_{n-2} + \cdots + c_{n-2} c_1 + c_{n-1} c_0$ на коэффициенты c_n *ряда Каталана* $c(x) = \sum\limits_{k \geqslant 0} c_k x^k = 1 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots \in \mathbb{Z}[\![x]\!]$, означающее, что $c(x)^2 = (c(x) - 1)/x$. Иначе говоря, $t = c(x)$ является лежащим в кольце $\mathbb{Z}[\![x]\!]$ решением квадратного уравнения $x \cdot t^2 - t - 1 = 0$ на неизвестную t. В поле рядов Лорана $\mathbb{Q}((x)) \supset \mathbb{Z}[\![x]\!]$ это квадратное уравнение решается по стандартной школьной формуле, что даёт два корня: $(1 \pm \sqrt{1-4x})/2x$. Так как ряд $1 + \sqrt{1-4x}$ имеет ненулевой свободный член, он не делится на $2x$ в $\mathbb{Z}[\![x]\!]$, и корень $(1 + \sqrt{1-4x})/(2x) \notin \mathbb{Z}[\![x]\!]$. Тем самым, $c(x) = (1 - \sqrt{1-4x})/(2x)$, откуда по формуле (4-23)

$$c_k = \frac{1}{2} \cdot \frac{1}{2k+1} \cdot \binom{2k+2}{k+1} = \frac{1}{k+1} \cdot \binom{2k}{k}.$$

Отметим, что с первого взгляда не вполне понятно, что это число — целое.

4.6. Ряд Тодда и числа Бернулли. Рассмотрим кольцо формальных степенных рядов $\mathbb{Q}[\![x]\!]$ от переменной x и кольцо многочленов $\mathbb{Q}[t]$ от переменной t. Обозначим через

$$D = \frac{d}{dt} : \mathbb{Q}[t] \to \mathbb{Q}[t], \quad g \mapsto g',$$

оператор дифференцирования. Оператор D можно подставить вместо переменной x в любой степенной ряд $\Phi(x) = \sum\limits_{k\geqslant 0} \varphi_k x^k \in \mathbb{Q}[\![x]\!]$. Результатом такой подстановки, по определению, является отображение

$$\Phi(D) : \mathbb{Q}[t] \to \mathbb{Q}[t], \quad f \mapsto \varphi_0 \cdot f + \varphi_1 \cdot f' + \varphi_2 \cdot f'' + \cdots = \sum_{k\geqslant 0} \varphi_k \cdot D^k f. \tag{4-25}$$

Поскольку каждое дифференцирование уменьшает степень многочлена на единицу, все слагаемые в правой части (4-25) обратятся в нуль при $k > \deg f$. Таким образом, для каждого многочлена $f \in \mathbb{Q}[t]$, правая часть (4-25) является корректно определённым многочленом, каждый коэффициент которого вычисляется конечным числом арифметических операций над коэффициентами исходного многочлена f и первыми $\deg(f)$ коэффициентами ряда Φ. Отображение $\Phi(D)$ *линейно* в том смысле, что

$$\forall\, \alpha, \beta \in \mathbb{Q}\ \forall\, f, g \in \mathbb{Q}[t] \quad \Phi(D)(\alpha \cdot f + \beta \cdot g) = \alpha \cdot \Phi(D) f + \beta \cdot \Phi(D) g, \tag{4-26}$$

а в результате подстановки D в произведение рядов $\Phi(x)\Psi(x) \in \mathbb{Q}[\![x]\!]$ получится *композиция* отображений $\Phi(D) \circ \Psi(D) = \Psi(D) \circ \Phi(D)$.

Упражнение 4.7. Убедитесь в этом.

Таким образом, все отображения вида $\Phi(D)$ перестановочны друг с другом, и для биективности отображения вида $\Phi(D)$ необходимо и достаточно, чтобы степенной ряд $\Phi(x)$ был обратим[1] в кольце $\mathbb{Q}[\![x]\!]$. В силу линейности значение отображения $\Phi(D)$ на произвольном многочлене выражается через его значения $\Phi_m(t) \stackrel{\text{def}}{=} \Phi(D)t^m$ на базисных одночленах t^m:

$$\Phi(D)\left(a_0 + a_1 t + \cdots + a_n t^n\right) = a_0 + a_1\Phi_1(t) + a_2\Phi_2(t) + \cdots + a_n\Phi_n(t).$$

Многочлен $\Phi_m(t) \in \mathbb{Q}[t]$ называется m-тым *многочленом Аппеля* ряда Φ. Его степень не превосходит m, а коэффициенты зависит лишь от первых $m+1$ коэффициентов ряда Φ.

Пример 4.8 (операторы сдвига)
Экспонента $e^D = 1 + D + \frac{1}{2}D^2 + \frac{1}{6}D^3 + \ldots$ имеет многочлены Аппеля

$$e^D t^m = \sum_{k\geqslant 0} \frac{1}{k!} D^k t^m = \sum_{k\geqslant 0} \frac{m(m-1)\cdots(m-k+1)}{k!}\, t^{m-k} = \sum_{k=0}^{m} \binom{m}{k} t^{m-k} = (t+1)^m.$$

Следовательно, оператор e^D действует на любой многочлен как *оператор сдвига*:

$$e^D : f(t) \mapsto f(t+1).$$

Так как ряды e^x и e^{-x} обратны друг другу в $\mathbb{Q}[\![x]\!]$, операторы e^D и e^{-D} тоже обратны друг другу, т. е. $e^{-D} f(t) = f(t-1)$.

[1] Т. е. имел ненулевой свободный член, см. предл. 3.1 на стр. 34.

Упражнение 4.8. Убедитесь, что $e^{\alpha D} f(t) = f(t+\alpha)$ при любом $\alpha \in \mathbb{Q}$.

Пример 4.9 (вычисление степенных сумм)
Для произвольно зафиксированного $m \in \mathbb{Z}_{\geqslant 0}$ рассмотрим сумму

$$S_m(n) \stackrel{\text{def}}{=} 0^m + 1^m + 2^m + 3^m + \cdots + n^m = \sum_{k=0}^{n} k^m \tag{4-27}$$

как функцию от n. При $m = 0, 1, 2, 3$ функции $S_m(n)$ достаточно известны:

$$\begin{aligned} S_0(n) &= 1 + 1 + 1 + \cdots + 1 = n \\ S_1(n) &= 1 + 2 + 3 + \cdots + n = n(n+1)/2 \\ S_2(n) &= 1^2 + 2^2 + 3^2 + \cdots + n^2 = n(n+1)(2n+1)/6 \\ S_3(n) &= 1^3 + 2^3 + 3^3 + \cdots + n^3 = n^2(n+1)^2/4 = S_1(n)^2 . \end{aligned} \tag{4-28}$$

Чтобы получить для $S_m(t)$ явное выражение, применим к этой функции *разностный оператор* $\nabla\colon \varphi(t) \mapsto \varphi(t) - \varphi(t-1)$. Получающаяся функция $\nabla S_m(t)$ принимает при всех $t \in \mathbb{Z}_{\geqslant 0}$ те же значения, что и многочлен t^m. Покажем, что существует единственный такой многочлен $S_m(t) \in \mathbb{Q}[t]$ с нулевым свободным членом, что $\nabla S_m(t) = t^m$. Тогда его значения при целых неотрицательных $t = 0, 1, 2, \ldots$ будут рекурсивно определяться, начиная с $S_m(0) = 0$, по формуле $S_m(n) = S_m(n-1) + \nabla S_m(n) = S_m(n-1) + n^m$ и, тем самым, совпадут с суммами (4-27). Из проделанных в прим. 4.8 вычислений вытекает, что

$$\nabla = 1 - e^{-D} = \frac{1 - e^{-D}}{D} \circ D .$$

Ряд $(1 - e^{-x})/x$ имеет свободный член 1 и обратим в $\mathbb{Q}[\![x]\!]$. Обратный ему ряд

$$\operatorname{td}(x) \stackrel{\text{def}}{=} \frac{x}{1 - e^{-x}} \in \mathbb{Q}[\![x]\!]$$

называется *рядом Тодда*. Подставляя $x = D$ в равенство $\operatorname{td}(x) \cdot (1 - e^{-x}) = x$, получаем соотношение $\operatorname{td}(D) \circ \nabla = D$. Стало быть, производная $S'_m(t) = DS_m(t) = \operatorname{td}(D)\nabla S_m(t) = \operatorname{td}(D) t^m$ является многочленом Аппеля $\operatorname{td}_m(t)$ ряда Тодда, а искомый многочлен $S_m(t) = \int \operatorname{td}_m(t)\, dt$ представляет собою его первообразную. Для её вычисления запишем ряд Тодда в «экспоненциальной форме», вынеся из коэффициентов обратные факториалы:

$$\operatorname{td}(x) = \sum_{k \geqslant 0} \frac{a_k}{k!} x^k . \tag{4-29}$$

Тогда сумма m-тых степеней первых t натуральных чисел равна

$$\begin{aligned} S_m(t) &= \int \Bigl(\sum_{k=0}^{m} \frac{a_k}{k!} D^k t^m \Bigr)\, dt = \int \Bigl(\sum_{k=0}^{m} \binom{m}{k} a_k t^{m-k} \Bigr)\, dt = \sum_{k=0}^{m} \binom{m}{k} \frac{a_k t^{m-k+1}}{m-k+1} = \\ &= \frac{1}{m+1} \left(\binom{m+1}{1} a_m t + \binom{m+1}{2} a_{m-1} t^2 + \cdots + \binom{m+1}{m} a_1 t^m + \binom{m+1}{m+1} a_0 t^{m+1} \right) . \end{aligned}$$

Эту формулу часто символически представляют в виде

$$(m+1) \cdot S_m(t) = (a\downarrow + t)^{m+1} - a_{m+1} ,$$

где стрелка у $a\downarrow$ предписывает при раскрытии бинома $(a+t)^{m+1}$ заменять a^k на a_k. Коэффициенты a_k рекурсивно вычисляются из равенства $\mathrm{td}(x)\cdot(1-e^{-x})/x=1$:

$$\Big(1+a_1x+\frac{a_2}{2}x^2+\frac{a_3}{6}x^2+\frac{a_4}{24}\,x^4+\ \cdots\ \Big)\cdot\Big(1-\frac{1}{2}\,x+\frac{1}{6}\,x^2-\frac{1}{24}\,x^3+\frac{1}{120}\,x^4-\ \cdots\ \Big)=1\,.$$

Упражнение 4.9. Найдите первую дюжину чисел a_k, проверьте формулы (4-28), дополните их явными формулами для $S_4(n)$ и $S_5(n)$ и вычислите[1] $S_{10}(1000)$.

Замечание 4.3. (числа Бернулли) Название «ряд Тодда» вошло в обиход во второй половине XX века после работ Хирцебруха и Гротендика, где этот ряд использовался для формулировки и доказательства теоремы Римана – Роха. Во времена Бернулли и Эйлера предпочитали пользоваться рядом $\mathrm{td}(-x)=\frac{x}{e^x-1}$, отличающимся от $\mathrm{td}(x)$ ровно одним членом, ибо

$$\mathrm{td}(-x)-\mathrm{td}(x)=\frac{x}{1-e^{-x}}+\frac{x}{1-e^{x}}=x\cdot\frac{2-e^x-e^{-x}}{(1-e^{-x})\cdot(1-e^x)}=x\,.$$

Это вычисление показывает, что коэффициенты при x в $\mathrm{td}(x)$ и в $\mathrm{td}(-x)$ равны соответственно $+\frac12$ и $-\frac12$, а все прочие коэффициенты при нечётных степенях x^{2k+1} с $k\geqslant 1$ в обоих рядах нулевые. Коэффициенты B_k в экспоненциальном представлении ряда

$$\frac{x}{e^x-1}=\sum_{k\geqslant 0}\frac{B_k}{k!}\,x^k$$

называются *числами Бернулли*. Таким образом, $B_k=a_k$ при $k\neq 1$ и обращаются в нуль при всех нечётных $k\geqslant 3$, а $B_1=-a_1=-\frac12$. Со времён своего открытия числа Бернулли вызывают неослабевающий интерес. Им посвящена обширная литература[2] и даже специальный интернет-ресурс[3], где среди прочего есть программа для быстрого вычисления чисел B_k в виде несократимых рациональных дробей. Однако, не смотря на множество красивых теорем о числах Бернулли, внятных формул, явно выражающих B_n через n нет, и любой содержательный новый взгляд в этом направлении был бы интересен.

[1] Яков Бернулли (1654 – 1705) пользуясь лишь пером и бумагой сложил 10-е степени первой тысячи натуральных чисел примерно за 7 минут, о чём не без гордости написал в своём манускрипте «Ars Conjectandi», изданном в 1713 году уже после его кончины.

[2] Начать знакомство с которой я советую с гл. 15 книги *К. Айрлэнд, М. Роузен*. «Классическое введение в современную теорию чисел» и § 8 гл. V книги *З. И. Боревич, И. Р. Шафаревич*. «Теория чисел».

[3] http://www.bernoulli.org/

§5. Идеалы, фактор кольца и разложение на множители

5.1. Идеалы. Подкольцо I коммутативного кольца K называется *идеалом*, если вместе с каждым своим элементом оно содержит и все его кратные. В n° 2.6.3 мы видели, что этим свойством обладает ядро любого гомоморфизма колец. Множество всех элементов кольца, кратных фиксированному элементу $a \in K$, также является идеалом. Этот идеал обозначается

$$(a) = \{ka \mid k \in K\}, \tag{5-1}$$

и называется *главным* идеалом, порождённым a. Мы встречались с главными идеалами при построении колец вычетов $\mathbb{Z}/(n)$ и $\Bbbk[x]\,/\,(f)$, где они возникали как ядра гомоморфизмов факторизации $\mathbb{Z} \twoheadrightarrow \mathbb{Z}/(n)$, $m \mapsto [m]_n$, и $\Bbbk[x] \twoheadrightarrow \Bbbk[x]/(f)$, $g \mapsto [g]_f$, которые сопоставляют целому числу (соотв. многочлену) его класс вычетов. Среди главных идеалов имеются *тривиальный* идеал (0), состоящий только из нулевого элемента, и *несобственный* идеал (1), совпадающий со всем кольцом. Идеалы, отличные от всего кольца, называются *собственными*.

Упражнение 5.1. Покажите, что следующие условия на идеал I в коммутативном кольце K с единицей эквивалентны: а) $I = K$ б) $1 \in I$ в) I содержит обратимый элемент.

Предложение 5.1
Коммутативное кольцо K с единицей тогда и только тогда является полем, когда в нём нет нетривиальных собственных идеалов.

Доказательство. Из упр. 5.1 вытекает, что в поле таких идеалов нет. Наоборот, если в кольце нет нетривиальных собственных идеалов, то главный идеал (b), состоящий из всех кратных произвольно взятого элемента $b \neq 0$, совпадает со всем кольцом. В частности, он содержит единицу, т. е. $1 = ab$ для некоторого a. Тем самым, любой ненулевой элемент b обратим. □

5.1.1. Нётеровость. Любое подмножество $M \subset K$ порождает идеал $(M) \subset K$, состоящий из всех элементов кольца K, представимых в виде $b_1a_1 + \cdots + b_ma_m$, где $a_1, \ldots, a_m$ — произвольные элементы множества M, а $b_1, \ldots, b_m$ — произвольные элементы кольца K, и число слагаемых $m \in \mathbb{N}$ также произвольно.

Упражнение 5.2. Убедитесь, что $(M) \subset K$ это и в самом деле идеал, совпадающий с пересечением всех идеалов, содержащих множество M.

Любой идеал $I \subset K$ имеет вид (M) для подходящего множества образующих $M \subseteq I$: например, всегда можно положить $M = I$. Идеалы $I = (a_1, \ldots, a_k) = \{b_1a_1 + \cdots + b_ka_k \mid b_i \in K\}$, допускающие конечное множество образующих, называются *конечно порождёнными*. Мы встречались с такими идеалами, когда доказывали существование наибольшего общего делителя в кольцах целых чисел и многочленов с коэффициентами в поле.

Лемма 5.1
Следующие свойства коммутативного кольца K попарно эквивалентны:

1) любое подмножество $M \subset K$ содержит конечный набор элементов $a_1, \ldots, a_k \in M$, порождающий тот же идеал, что и M
2) любой идеал $I \subset K$ конечно порождён
3) любая бесконечная возрастающая цепочка вложенных идеалов $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ в K стабилизируется в том смысле, что найдётся такое $n \in \mathbb{N}$, что $I_\nu = I_n$ для всех $\nu \geqslant n$.

Доказательство. Ясно, что (1) влечёт (2). Чтобы получить (3) из (2), заметим, что объединение $I = \bigcup I_v$ всех идеалов цепочки тоже является идеалом. Согласно (2), идеал I порождён конечным набором элементов. Все они принадлежат некоторому идеалу I_n. Тогда $I_n = I = I_v$ при $v \geqslant n$. Чтобы вывести (1) из (3), будем по индукции строить цепочку идеалов $I_n = (a_1, \ldots, a_n)$, начав с произвольного элемента $a_1 \in M$ и добавляя на k-том шагу очередную образующую $a_k \in M \setminus I_{k-1}$ до тех пор, пока это возможно, т. е. пока $M \not\subset I_k$. Так как $I_{k-1} \subsetneq I_k$, этот процесс не может продолжаться бесконечно, и на каком-то шагу мы получим идеал, содержащий всё множество M, а значит, совпадающий с (M). □

Определение 5.1
Кольцо K, удовлетворяющее условиям лем. 5.1, называется *нётеровым*. Отметим, что любое поле нётерово.

Теорема 5.1
Если кольцо K нётерово, то кольцо многочленов $K[x]$ также нётерово.

Доказательство. Рассмотрим произвольный идеал $I \subset K[x]$ и обозначим через $L_d \subset K$ множество старших коэффициентов всех многочленов степени не выше d из I, а через $L_\infty = \bigcup_d L_d$ — множество старших коэффициентов вообще всех многочленов из I.

Упражнение 5.3. Убедитесь, что все L_d (включая L_∞) являются идеалами в K.

Поскольку кольцо K нётерово, все идеалы L_d конечно порождены. Для каждого d (включая $d = \infty$) обозначим через $f_1^{(d)}, \ldots, f_{m_d}^{(d)} \in K[x]$ многочлены, старшие коэффициенты которых порождают соответствующий идеал $L_d \subset K$. Пусть наибольшая из степеней многочленов $f_i^{(\infty)}$, старшие коэффициенты которых порождают идеал L_∞, равна D. Покажем, что идеал I порождается многочленами $f_i^{(\infty)}$ и $f_j^{(d)}$ с $d < D$.

Каждый многочлен $g \in I$ сравним по модулю многочленов $f_1^{(\infty)}, \ldots, f_{m_\infty}^{(\infty)}$ с многочленом, степень которого строго меньше D. В самом деле, поскольку старший коэффициент многочлена g лежит в идеале L_∞, он имеет вид $\sum \lambda_i a_i$, где $\lambda_i \in K$, а a_i — старшие коэффициенты многочленов $f_i^{(\infty)}$. При $\deg g \geqslant D$ все разности $m_i = \deg g - \deg f_i^{(\infty)} \geqslant 0$, и можно образовать многочлен $h = g - \sum \lambda_i \cdot f_i^{(\infty)}(x) \cdot x_i^{m_i}$, сравнимый с g по модулю I и имеющий $\deg h < \deg g$. Заменим g на h и повторим эту процедуру, пока не получим многочлен $h \equiv g \pmod{(f_1^{(\infty)}, \ldots, f_{m_\infty}^{(\infty)})}$ с $\deg h < D$. Теперь старший коэффициент многочлена h лежит в идеале L_d с $d < D$, и мы можем строго уменьшать его степень, сокращая старший член путём вычитания из h подходящих комбинаций многочленов $f_j^{(d)}$ с $0 \leqslant d < D$. □

Следствие 5.1
Если K нётерово, то кольцо многочленов $K[x_1, \ldots, x_n]$ также нётерово. □

Упражнение 5.4. Покажите, что кольцо формальных степенных рядов над нётеровым кольцом нётерово.

Следствие 5.2
Любая система полиномиальных уравнений с коэффициентами в нётеровом кольце эквивалентна некоторой конечной своей подсистеме.

Доказательство. Если кольцо K нётерово, то кольцо $K[x_1,\dots,x_n]$ тоже нётерово, и в любом множестве многочленов $M \subset K[x_1,\dots,x_n]$ можно указать такой конечный набор многочленов $f_1,\dots,f_m \in M$, что среди многочленов f_ν, что каждый многочлен $g \in M$ представляется в виде $g = h_1f_1 + \dots + h_mf_m$ для некоторых $h_i \in K[x_1,\dots,x_n]$. Поэтому любое уравнение вида $g(x_1,\dots,x_n) = 0$ с $g \in M$ является следствием конечного множества уравнений $f_1(x_1,\dots,x_n) = f_2(x_1,\dots,x_n) = \dots = f_m(x_1,\dots,x_n) = 0$. □

5.1.2. Примеры ненётеровых колец. Кольцо многочленов от счётного множества переменных $\mathbb{Q}[x_1,x_2,x_3,\ \dots]$, элементы которого суть конечные линейные комбинации с рациональными коэффициентами всевозможных мономов вида $x_{\nu_1}^{m_1}x_{\nu_2}^{m_2}\cdots x_{\nu_s}^{m_s}$ не является нётеровым: его идеал $(x_1,x_2,\dots)$, состоящий из всех многочленов без свободного члена, нельзя породить конечным множеством многочленов.

Упражнение 5.5. Докажите это и выясните, является ли конечно порождённым идеал, образованный в кольце бесконечно гладких функций $\mathbb{R}\to\mathbb{R}$ всеми функциями, которые обращаются в нуле в нуль вместе со всеми своими производными.

Предостережение 5.1. Подкольцо нётерова кольца может не быть нётеровым. Например, кольцо формальных степенных рядов $\mathbb{C}[\![z]\!]$ нётерово по упр. 5.4, тогда как его подкольцо образованное рядами, сходящимися всюду в $\mathbb{C}$, нётеровым не является.

Упражнение 5.6. Приведите пример бесконечной возрастающей цепочки строго вложенных идеалов в кольце сходящихся всюду в $\mathbb{C}$ степенных рядов с комплексными коэффициентами.

5.2. Фактор кольца. Пусть на коммутативном кольце K задано отношение эквивалентности, разбивающее K в дизъюнктное объединение классов эквивалентных элементов. Обозначим множество классов через X и рассмотрим сюръективное отображение факторизации

$$\pi\colon K \twoheadrightarrow X, \quad a \mapsto [a], \tag{5-2}$$

переводящее элемент $a \in K$ в его класс эквивалентности $[a] \subset K$, являющийся элементом множества X. Мы хотим задать на множестве X структуру коммутативного кольца, определив сложение и умножение теми же сами правилами

$$[a]+[b] = [a+b], \quad [a]\cdot[b] = [ab], \tag{5-3}$$

которые мы использовали в кольцах вычетов. Если эти правила корректны, то аксиомы коммутативного кольца в X будут автоматически выполнены, как и для колец вычетов, поскольку формулы (5-3) сводят их проверку к проверке аксиом коммутативного кольца в K. В частности, нулевым элементом кольца X будет класс $[0]$. С другой стороны, если формулы (5-3) корректны, то они утверждают, что отображение (5-2) является гомоморфизмом колец. Но если это так, то согласно n° 2.6.3 на стр. 29 класс нуля $[0] = \ker\pi$, служащий ядром этого гомоморфизма, является идеалом в K, а класс $[a] \subset K$ произвольного элемента $a \in K$, служащий прообразом точки $[a] \in X$ при гомоморфизме (5-2), является аддитивным сдвигом ядра на этот элемент:

$$[a] = \pi^{-1}(\pi(a)) = a + \ker\pi = a + [0] = \{a+b \mid b \in [0]\}.$$

Оказывается, что этих необходимых условий на классы также и достаточно для того, чтобы правила (5-3) были корректны, т. е. для любого идеала $I \subset K$ множество классов

$$[a]_I = a + I \stackrel{\text{def}}{=} \{a + b \mid b \in I\} \tag{5-4}$$

образует разбиение кольца K, и правила (5-3) корректно определяют на классах этого разбиения структуру коммутативного кольца с нулевым элементом $[0]_I = I$.

Упражнение 5.7. Убедитесь, что отношение сравнимости по модулю идеала

$$a_1 \equiv a_2 \pmod{I},$$

означающее, что $a_1 - a_2 \in I$, является отношением эквивалентности, и проверьте, что формулы (5-3) корректны.

Определение 5.2
Классы эквивалентности (5-4) называются *классами вычетов* (или *смежными классами*) по модулю идеала I. Множество этих классов с операциями (5-3) называется *фактор кольцом* кольца K по идеалу I и обозначается K/I. Эпиморфизм

$$K \twoheadrightarrow K/I, \quad a \mapsto [a]_I, \tag{5-5}$$

сопоставляющий каждому элементу кольца его класс вычетов, называется *гомоморфизмом факторизации*.

Пример 5.1 (кольца вычетов)
Рассматривавшиеся выше кольца $\mathbb{Z}/(n)$ и $\Bbbk[x]/(f)$ суть фактор кольца кольца целых чисел и кольца многочленов по главным идеалам $(n) \subset \mathbb{Z}$ и $(f) \subset \Bbbk[x]$ соответственно.

Пример 5.2 (образ гомоморфизма)
Согласно n° 2.6.3, для любого гомоморфизма коммутативных колец $\varphi : A \to B$ имеется канонический изоморфизм колец $\overline{\varphi} : A/\ker\varphi \simeq\in \varphi$, $[a]_{\ker\varphi} \mapsto \varphi(a)$, переводящий каждый класс

$$[a]_{\ker\varphi} = a + \ker\varphi = \varphi^{-1}(\varphi(a))$$

в его образ $\varphi(a) = \varphi([a])$ при гомоморфизме φ.

Пример 5.3 (максимальные идеалы и гомоморфизмы вычисления)
Идеал $\mathfrak{m} \subset K$ называется *максимальным*, если фактор кольцо $K/\mathfrak{m}$ является полем. Название связано с тем, что собственный[1] идеал $\mathfrak{m} \subset K$ максимален если и только если он не содержится ни в каком строго большем собственном идеале, т. е. является максимальным элементом в чуме[2] собственных идеалов кольца K, частично упорядоченных отношением нестрогого включения. В самом деле, обратимость всех ненулевых классов $[a]_{\mathfrak{m}}$ в фактор кольце $K/\mathfrak{m}$ означает, что для любого $a \notin \mathfrak{m}$ найдутся такие $b \in K$, $m \in \mathfrak{m}$, что $ab + m = 1$ в K. Последнее равносильно тому, что идеал $(\mathfrak{m}, a) \supsetneq \mathfrak{m}$, порождённый $\mathfrak{m}$ и элементом $a \notin \mathfrak{m}$, содержит 1 и совпадает с K, т. е. что идеал $\mathfrak{m}$ не содержится ни в каком строго большем собственном идеале.

[1] Т. е. отличный от всего кольца.
[2] См. n° 1.7 на стр. 16.

Из леммы Цорна[1] вытекает, что любой собственный идеал произвольного коммутативного кольца с единицей содержится в некотором максимальном идеале. В самом деле, множество всех собственных идеалов, содержащих произвольно заданный идеал $I \subset K$, тоже составляет чум по включению.

Упражнение 5.8. Убедитесь, что он полный, т. е. для любого линейно упорядоченного множества[2] M содержащих I собственных идеалов в K существует собственный идеал J^*, содержащий все идеалы из M.

По лемме Цорна существует такой собственный идеал $\mathfrak{m} \supset I$, который не содержится ни в каком большем собственном идеале, содержащем I. Такой идеал $\mathfrak{m}$ автоматически максимален по включению и в чуме всех собственных идеалов кольца K.

Максимальные идеалы возникают в кольцах функций как ядра *гомоморфизмов вычисления*. А именно, пусть X — произвольное множество, $p \in X$ — любая точка, $\Bbbk$ — любое поле, и K — какое-нибудь подкольцо в кольце всех функций $X \to \Bbbk$, содержащее тождественно единичную функцию 1 и вместе с каждой функцией $f \in K$ содержащее и все пропорциональные ей функции cf, $c \in \Bbbk$. Гомоморфизм вычисления $\mathrm{ev}_p : K \to \Bbbk$ переводит функцию $f \in K$ в её значение $f(p) \in \Bbbk$. Поскольку он сюръективен, его ядро $\ker \mathrm{ev}_p = \{f \in K \mid f(p) = 0\}$ является максимальным идеалом в K.

Упражнение 5.9. Убедитесь, что: а) каждый максимальный идеал кольца $\mathbb{C}[x]$ имеет вид $\ker \mathrm{ev}_p$ для некоторого $p \in \mathbb{C}$ б) в кольце непрерывных функций $[0,1] \to \mathbb{R}$ каждый максимальный идеал имеет вид $\ker \mathrm{ev}_p$ для некоторой точки $p \in [0,1]$. в) Укажите в кольце $\mathbb{R}[x]$ максимальный идеал, отличный от всех идеалов вида $\ker \mathrm{ev}_p$, где $p \in \mathbb{R}$.

Пример 5.4 (простые идеалы и гомоморфизмы в поля)
Идеал $\mathfrak{p} \subset K$ называется *простым*, если в фактор кольце $K/\mathfrak{p}$ нет делителей нуля. Иначе говоря, идеал $\mathfrak{p} \subset K$ прост если и только если из $ab \in \mathfrak{p}$ вытекает, что $a \in \mathfrak{p}$ или $b \in \mathfrak{p}$. Например, главные идеалы $(p) \subset \mathbb{Z}$ и $(q) \subset \Bbbk[x]$, где $\Bbbk$ — поле, просты тогда и только тогда, когда число p просто, а многочлен q неприводим.

Упражнение 5.10. Убедитесь в этом.

Согласно определениям, всякий максимальный идеал прост. Обратное неверно: скажем, главный идеал $(x) \subset \mathbb{Q}[x,y]$ прост, так как кольцо $\mathbb{Q}[x,y]/(x) \simeq \mathbb{Q}[x]$ целостное, но не максимален, поскольку строго содержится в идеале (x,y) многочленов без свободного члена. Простые идеалы кольца K являются ядрами гомоморфизмов из кольца K во всевозможные поля. В самом деле, образ любого такого гомоморфизма, будучи подкольцом в поле, не имеет делителей нуля. Наоборот, фактор кольцо $K/\mathfrak{p}$ по простому идеалу $\mathfrak{p} \subset K$ является подкольцом своего поля частных $Q_{K/\mathfrak{p}}$, и композиция факторизации и вложения $K \twoheadrightarrow K/\mathfrak{p} \hookrightarrow Q_{K/\mathfrak{p}}$ задаёт гомоморфизм из K в поле $Q_{K/\mathfrak{p}}$ с ядром $\mathfrak{p}$.

Упражнение 5.11. Убедитесь, что пересечение конечного множества идеалов содержится в простом идеале $\mathfrak{p}$ только если хотя бы один из пересекаемых идеалов содержится в $\mathfrak{p}$.

[1] См. сл. 1.1 на стр. 19.

[2] В данном случае это означает, что для любых $J_1, J_2 \in M$ выполняется включение $J_1 \subseteq J_2$ или включение $J_2 \subseteq J_1$.

Пример 5.5 (конечно порождённые коммутативные алгебры)
Пусть K — произвольное коммутативное кольцо с единицей. Всякое кольцо вида

$$A = K[x_1, \dots, x_n]/I\,,$$

где $I \subset K[x_1, \dots, x_n]$ — произвольный идеал, называется *конечно порождённой K-алгеброй*[1]. Классы $a_i = [x_i]_I$ называются *образующими* K-алгебры A, а многочлены $f \in I$ — *соотношениями* между этими образующими. Говоря неформально, K-алгебра состоит из всевозможных выражений, которые можно составить из элементов кольца K и коммутирующих букв $a_1, \dots, a_n$ при помощи операций сложения и умножения, производимых с учётом полиномиальных соотношений $f(a_1, \dots, a_n) = 0$ для всех f из I. Из сл. 5.1 и упр. 5.12

Упражнение 5.12. Покажите, что фактор кольцо нётерова кольца тоже нётерово.

мы получаем

Следствие 5.3
Всякая конечно порождённая коммутативная алгебра над нётеровым кольцом нётерова и все соотношения между её образующими являются следствиями конечного числа соотношений.□

5.3. Кольца главных идеалов. Целостное кольцо с единицей называется *кольцом главных идеалов*, если каждый его идеал является главным. Параллелизм между кольцами $\mathbb{Z}$ и $\Bbbk[x]$, где $\Bbbk$ — поле, который мы наблюдали выше, объясняется тем, что оба эти кольца являются кольцами главных идеалов. Мы фактически доказали это, когда строили в этих кольцах наибольший общий делитель. Ниже мы воспроизведём это доказательство ещё раз таким образом, чтобы оно годилось для чуть более широкого класса колец, допускающих *деление с остатком*.

5.3.1. Евклидовы кольца. Целостное кольцо K с единицей называется *евклидовым*, если существует *функция высоты* (или *евклидова норма*) $\nu : K \setminus \{0\} \to \mathbb{N} \cup \{0\}$, сопоставляющая каждому ненулевому элементу $a \in K$ целое неотрицательное число $\nu(a)$ так, что $\forall\, a, b \in K \setminus \{0\}$ выполняется неравенство $\nu(ab) \geqslant \nu(a)$ и существуют такие $q, r \in K$, что

$$a = bq + r\,,\ \text{где } \nu(r) < \nu(b) \text{ или } r = 0. \tag{5-6}$$

Элементы q, r называются *неполным частным* и *остатком* от деления a на b. Подчеркнём, что их единственности (для данных a и b) не предполагается.

Упражнение 5.13. Докажите евклидовость колец: а) $\mathbb{Z}$ с $\nu(z) = |z|$ б) $\Bbbk[x]$ с $\nu(f) = \deg f$
в) $\mathbb{Z}[i] \stackrel{\text{def}}{=} \{\, a + bi \in \mid a, b \in \mathbb{Z},\ i^2 = -1\}$ с $\nu(z) = |z|^2$
г) $\mathbb{Z}[\omega] \stackrel{\text{def}}{=} \{a + b\omega \in \mathbb{C} \mid a, b \in \mathbb{Z},\ \omega^2 + \omega + 1 = 0\,\}$ с $\nu(z) = |z|^2$.

Все четыре кольца из упр. 5.13 являются кольцами главных идеалов в силу следующей теоремы.

Теорема 5.2
Любое евклидово кольцо является кольцом главных идеалов[2].

Доказательство. Пусть $I \subset K$ — идеал, и $d \in I$ — ненулевой элемент наименьшей высоты. Покажем, что каждый элемент $a \in I$ делится на d. Поделим a на d с остатком: $a = dq + r$. Так

[1] Или, более торжественно, *конечно порождённой коммутативной алгеброй* над кольцом K.

[2] Отметим, что обратное неверно, но содержательное обсуждение контрпримеров требует техники, которой мы пока не владеем (см. замечание 3 на стр. 365 книги *Э. Б. Винберг*. «Курс алгебры», М. «Факториал», 1999)

как $a, d \in I$, остаток $r = a - dq \in I$. Поскольку строгое неравенство $\nu(r) < \nu(d)$ невозможно, мы заключаем, что $r = 0$. □

Упражнение 5.14. Покажите, что в любом евклидовом кольце равенство $\nu(ab) = \nu(a)$ для $a, b \neq 0$ равносильно обратимости элемента b.

5.3.2. НОД и взаимная простота. В кольце главных идеалов K у любого набора элементов $a_1, \dots, a_n$ есть наибольший общий делитель — такой элемент $d = \text{нод}(a_1, \dots, a_n) \in K$, который делит все элементы a_i, делится на любой общий делитель элементов a_i и представляется в виде $d = a_1 b_1 + \dots + a_n b_n$ с подходящими $b_i \in K$. Это простая переформулировка того, что порождённый элементами a_i идеал $(a_1, \dots, a_n) = \{x_1 a_1 + \dots + x_n a_n \mid x_i \in K\}$ является главным и имеет вид (d) для некоторого $d \in K$. Отметим, что наибольший общий делитель определён не однозначно, а с точностью до умножения на произвольный обратимый элемент кольца.

Упражнение 5.15. Убедитесь, что в любом целостном[1] коммутативном кольце K главные идеалы (a) и (b) совпадают если и только если $a = sb$ для некоторого обратимого $s \in K$.

Поэтому всюду в дальнейшем обозначение $\text{нод}(a_1, \dots, a_n)$ подразумевает целый класс элементов, получающихся друг из друга умножениями на обратимые константы, и все формулы, которые будут писаться, относятся к произвольно выбранному конкретному представителю этого класса[2]. В частности, равенство $\text{нод}(a_1, \dots, a_n) = 1$ означает, что у элементов a_i нет необратимых общих делителей. Поскольку в этом случае имеется представление $1 = a_1 b_1 + \dots + a_n b_n$ с $b_i \in K$, в кольце главных идеалов отсутствие необратимых общих делителей у элементов a_i равносильно их *взаимной простоте* в смысле опр. 2.2 на стр. 25.

Упражнение 5.16. Проверьте, что идеалы $(x, y) \subset \mathbb{Q}[x, y]$ и $(2, x) \in \mathbb{Z}[x]$ не являются главными.

5.4. Факториальность. Всюду в этом разделе мы по умолчанию обозначаем через K *целостное*[3] кольцо. Ненулевые элементы $a, b \in K$ называются *ассоциированными*, если b делится на a, и a делится на b. Из равенств $a = mb$ и $b = na = nmb$ вытекает равенство $b(1 - nm) = 0$, откуда $mn = 1$. Таким образом, ассоциированность элементов означает, что они получаются друг из друга умножением на обратимый элемент кольца. Например, целые числа a и b ассоциированы в кольце $\mathbb{Z}$ если и только если $a = \pm b$, а многочлены $f(x)$ и $g(x)$ с коэффициентами из поля $\Bbbk$ ассоциированы в $\Bbbk[x]$ если и только если $f(x) = cg(x)$, где $c \in \Bbbk^*$ — ненулевая константа.

5.4.1. Неприводимые элементы. Элемент $q \in K$ называется *неприводимым*, если он необратим, и из равенства $q = mn$ вытекает, что M или n обратим. Другими словами, неприводимость элемента q означает, что главный идеал q не содержится строго ни в каком другом главном идеале, т. е. максимален в множестве главных идеалов. Например, неприводимыми элементами в кольце целых чисел являются простые числа, а в кольце многочленов — неприводимые многочлены.

Отметим, что кольце главных идеалов любые два неприводимых элемента p, q либо взаимно просты[4], либо ассоциированы, поскольку порождённый ими идеал $(p, q) = (d)$ для некоторого $d \in K$, и включения $(p) \subset (d)$ и $(q) \subset (d)$ влекут либо равенство $(d) = (K) = (1)$, либо равенство $(d) = (p) = (q)$. Обратите внимание, что в произвольном целостном кольце два

[1] Т. е. с единицей и без делителей нуля.

[2] Что, конечно же, требует проверки корректности всех таких формул, которую мы, как правило, будем оставлять читателю в качестве упражнения.

[3] См. сноску (1) выше.

[4] В смысле опр. 2.2 на стр. 25, т. е. существуют такие $x, y \in K$, что $px + qy = 1$.

неассоциированных неприводимых элемента могут и не быть взаимно простыми. Например, в $\mathbb{Q}[x, y]$ элементы x и y не взаимно просты и не ассоциированы.

Предложение 5.2
В любом кольце главных идеалов K следующие свойства элемента $p \in K$ попарно эквивалентны друг другу:

1) фактор кольцо $K/(p)$ является полем

2) в фактор кольце $K/(p)$ нет делителей нуля

3) p неприводим, т. е. из равенства $p = ab$ вытекает, что a или b обратим в K.

Доказательство. Импликация (1) $\Rightarrow$ (2) очевидна и имеет место в любом коммутативном кольце с единицей[1]. Покажем, что в любом целостном кольце[2] K справедлива импликация (2) $\Rightarrow$ (3). Из $p = ab$ следует, что $[a][b] = 0$ в $K/(p)$. Так как в $K/(p)$ нет делителей нуля, один из сомножителей, скажем $[a]$, равен $[0]$. Тогда $a = ps = abs$ для некоторого $s \in K$, откуда $a(1 - bs) = 0$. Поскольку в K нет делителей нуля, $bs = 1$, т. е. b обратим. Покажем теперь, что в кольце главных идеалов (3) $\Rightarrow$ (1). Так как каждый собственный идеал в K главный, максимальность идеала (p) в чуме главных идеалов означает его максимальность в чуме всех собственных идеалов. В прим. 5.3 на стр. 69 мы видели, что это равносильно тому, что $K/(p)$ поле. □

Предложение 5.3
В любом нётеровом кольце всякий элемент является произведением конечного числа неприводимых.

Доказательство. Если элемент a неприводим, доказывать нечего. Пусть a приводим. Запишем его в виде произведения необратимых элементов. Каждый приводимый сомножитель этого произведения снова запишем в виде произведения необратимых элементов и т. д. Эта процедура закончится, когда все сомножители станут неприводимы, что и требуется. Если же она никогда не закончится, мы сможем образовать бесконечную последовательность строго вложенных друг в друга главных идеалов $(a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \cdots$, что противоречит нётеровости. □

Определение 5.3
Целостное кольцо K называется *факториальным*, если каждый его необратимый элемент является произведением конечного числа неприводимых, причём любые два таких разложения

$$p_1 p_2 \cdots p_m = q_1 q_2 \cdots q_k$$

состоят из одинакового числа $k = m$ сомножителей, после надлежащей перенумерации которых можно указать такие обратимые элементы $s_\nu \in K$, что $q_\nu = p_\nu s_\nu$ при всех ν.

[1] См. n° 2.4.1 на стр. 26.
[2] Не обязательно являющимся кольцом главных идеалов.

5.4.2. Простые элементы. Элемент $p \in K$ называется *простым*, если порождённый им главный идеал $(p) \subset K$ прост, т. е. в фактор кольце $K/(p)$ нет делителей нуля. Это означает, что для любых $a, b \in K$ из того, что произведение ab делится на p, вытекает, что a или b делится на p. Каждый простой элемент p автоматически неприводим: если $p = xy$, то один из сомножителей, скажем x, делится на p, и тогда $p = pyz$, откуда $yz = 1$ и y обратим. Согласно предл. 5.2 в кольце главных идеалов верно и обратное: все неприводимые элементы кольца главных идеалов просты. Однако в произвольном целостном кольце простота является более сильным свойством, чем неприводимость. Например, в кольце $\mathbb{Z}[\sqrt{5}] = \mathbb{Z}[x]/(x^2-5)$ число 2 неприводимо, но не просто, поскольку в фактор кольце

$$\mathbb{Z}[\sqrt{5}]/(2) \simeq \mathbb{Z}[x]/(2, x^2-5) = \mathbb{Z}[x]/(2, x^2+1) \simeq \mathbb{F}_2[x]/(x^2+1) \simeq \mathbb{F}_2[x]/\big((x+1)^2\big)$$

есть нильпотент — класс $[x+1] \in \mathbb{Z}[x]/(2, x^2+5)$. Среди прочего, это означает, что квадрат $(1+\sqrt{5})^2 = 6+2\sqrt{5}$ делится в кольце $\mathbb{Z}[\sqrt{5}]$ на 2, хотя $1+\sqrt{5}$ не делится на 2, при том что 2 и $\sqrt{5}+1$ неприводимы и не ассоциированы друг с другом в кольце $\mathbb{Z}[\sqrt{5}]$.

Упражнение 5.17. Убедитесь в этом, и покажите, что $2 \cdot 2 = 4 = \big(\sqrt{5}+1\big) \cdot \big(\sqrt{5}-1\big)$ суть два *различных* разложения числа 4 на неприводимые множители в $\mathbb{Z}[\sqrt{5}]$.

Предложение 5.4
Целостное нётерово кольцо K факториально тогда и только тогда, когда все его неприводимые элементы просты.

Доказательство. Покажем сначала, что если K факториально, то любой неприводимый элемент $q \in K$ прост. Пусть произведение ab делится на q. Тогда разложение ab на неприводимые множители содержит множитель, ассоциированный с q. В силу своей единственности, разложение произведения ab на неприводимые множители является произведением таких разложени для a и b. Поэтому q ассоциирован с одним из неприводимых делителей a или b, т. е. a или b делится на q, что и требовалось. Пусть теперь все неприводимые элементы просты. В нётеровом кольце каждый элемент является произведением конечного числа неприводимых и, стало быть, простых элементов. Покажем, что в любом целостном кольце равенство $p_1 \cdots p_k = q_1 \cdots q_m$, в котором все сомножители просты, возможно только если $k = m$ и после надлежащей перенумерации каждый p_i окажется ассоциирован с q_i. Так как произведение $q_1 \cdots q_m$ делится на p_1, один из его сомножителей делится на p_1. Будем считать, что это $q_1 = sp_1$. Поскольку q_1 неприводим, элемент s обратим. Пользуясь целостностью кольца K, сокращаем обе части равенства $p_1 \cdots p_k = q_1 \cdots q_m$ на p_1 и получаем более короткое равенство $p_2 p_3 \cdots p_k = (sq_2)q_3 \cdots q_m$, к которому применимы те же рассуждения. □

Следствие 5.4
Всякое кольцо главных идеалов факториально. □

Пример 5.6 (суммы двух квадратов, продолжение прим. 3.6 на стр. 48)
Согласно упр. 5.13, кольцо гауссовых чисел $\mathbb{Z}[i] \subset \mathbb{C}$ является кольцом главных идеалов, а потому в нём справедлива теорема об однозначности разложения на неприводимые множители. Выясним, какие целые простые числа $p \in \mathbb{Z}$ остаются неприводимыми в кольце гауссовых чисел. В $\mathbb{Z}[i]$ разложение любого целого вещественного числа, будучи инвариантным относительно комплексного сопряжения, содержит вместе с каждым невещественным неприводимым множителем также и сопряжённый ему множитель. Поэтому простое $p \in \mathbb{Z}$, не являющееся простым в $\mathbb{Z}[i]$, представляется в виде $p = (a+ib)(a-ib) = a^2+b^2$ с ненулевыми

$a, b \in \mathbb{Z}$. Таким образом, простое $p \in \mathbb{Z}$ приводимо в $\mathbb{Z}[i]$ если и только если p является суммой двух квадратов. С другой стороны, неприводимость $p \in \mathbb{Z}[i]$ означает, что фактор кольцо $\mathbb{Z}[i]/(p) \simeq \mathbb{Z}[x]/(p, x^2+1) \simeq \mathbb{F}_p[x]/(x^2+1)$ является полем[1], что равносильно неприводимости многочлена x^2+1 над $\mathbb{F}_p$, т. е. отсутствию у него корней в $\mathbb{F}_p$. Мы заключаем, что простое $p \in \mathbb{Z}$ является суммой двух квадратов если и только если -1 квадратичный вычет по модулю p. Как мы видели в n° 3.5.2 на стр. 51, это происходит при $p = 2$ и тех $p > 2$, для которых $(p-1)/2$ чётно, т. е. для $p = 4k+1$.

Упражнение 5.18. Покажите, что натуральное число n тогда и только тогда является квадратом или суммой двух квадратов натуральных чисел, когда в его разложение на простые множители простые числа $p = 4k+3$ входят лишь в чётных степенях.

5.4.3. НОД в факториальном кольце. В факториальном кольце K наибольший общий делитель набора элементов $a_1, \dots, a_m \in K$ допускает следующее описание. Для каждого класса ассоциированных неприводимых элементов $q \in K$ обозначим через m_q максимальное такое целое число, что q^{m_q} делит каждое из чисел a_i. Тогда, с точностью до умножения на обратимые константы,

$$\text{нод}(a_1, \dots, a_m) = \prod_q q^{m_q}.$$

Поскольку любой элемент факториального кольца является произведением конечного количества неприводимых элементов, числа m_q отличны от нуля лишь для конечного числа классов q. Поэтому написанное произведение корректно определено и в силу факториальности K делится на любой общий делитель чисел a_i.

5.5. Многочлены над факториальным кольцом. Пусть K — факториальное кольцо. Обозначим через Q_K его поле частных. Кольцо многочленов $K[x]$ является подкольцом в кольце многочленов $Q_K[x]$. Назовём *содержанием* многочлена $f = a_0 + a_1 x + \dots + a_n x^n \in K[x]$ наибольший общий делитель $\operatorname{cont}(f) \stackrel{\text{def}}{=} \text{нод}(a_0, a_1, \dots, a_n)$ его коэффициентов.

Лемма 5.2
$\operatorname{cont}(fg) = \operatorname{cont}(f) \cdot \operatorname{cont}(g)$ для любых $f, g \in K[x]$.

Доказательство. Достаточно для каждого неприводимого $q \in K$ убедиться в том, что q делит все коэффициенты произведения fg если и только если q делит все коэффициенты одного из многочленов f, g. Поскольку неприводимые элементы факториального кольца просты, фактор кольцо $R = K/(q)$ целостное. Применим к произведению fg гомоморфизм редукции по модулю q: $K[x] \to R[x]$, $a_0 + a_1 x + \dots + a_n x^n \mapsto [a_0]_q + [a_1]_q x + \dots [a_n]_q x^n$, заменяющий все коэффициенты каждого многочлена классами их вычетов по модулю q.

Упражнение 5.19. Проверьте, что это и в самом деле гомоморфизм колец.

Так как кольцо $R[x]$ тоже целостное, произведение $[fg]_q = [f]_q [g]_q$ обращается в нуль если и только если один из сомножителей $[f]_q$, $[g]_q$ равен нулю. □

Лемма 5.3 (редуцированное представление)
Каждый многочлен $f \in Q_K[x]$ представляется в виде $f(x) = (a/b) \cdot f_{\text{red}}(x)$, где $f_{\text{red}} \in K[x]$, $a, b \in K$ и $\operatorname{cont}(f_{\text{red}}) = \text{нод}(a, b) = 1$, причём числа a, b и многочлен f_{red} определяются по f однозначно с точностью до умножения на обратимые элементы кольца K.

[1] См. предл. 5.2 на стр. 73.

Доказательство. Вынесем из коэффициентов f их общий знаменатель, потом вынесем из всех коэффициентов полученного многочлена их наибольший общий делитель. В результате мы получим многочлен содержания 1, умноженный на число из Q_K, которое запишем несократимой дробью a/b. Докажем единственность такого представления. Если $(a/b)\cdot f_{\text{red}}(x) = (c/d)\cdot g_{\text{red}}(x)$ в $Q_K[x]$, то $ad\cdot f_{\text{red}}(x) = bc\cdot g_{\text{red}}(x)$ в $K[x]$. Сравнивая содержание обеих частей, получаем $ad = bc$. В виду отсутствия общих неприводимых множителей у a и b и у c и d, это возможно, только если a ассоциирован с c, а b ассоциирован с d. Но тогда и $f_{\text{red}}(x) = g_{\text{red}}(x)$ с точностью до умножения на обратимую константу. □

Следствие 5.5 (лемма Гаусса)
Многочлен $f \in K[x]$ содержания 1 неприводим в $Q_K[x]$ если и только если он неприводим в $K[x]$.

Доказательство. Пусть $f(x) = g(x)\cdot h(x)$ в $Q_K[x]$. Записывая многочлены g и h в редуцированном виде из лем. 5.3 и сокращая возникающую дробь, приходим к равенству

$$f(x) = \frac{a}{b}\cdot g_{\text{red}}(x)\cdot h_{\text{red}}(x)\,, \tag{5-7}$$

в котором $g_{\text{red}}, h_{\text{red}} \in K[x]$ имеют содержание 1, и нод$(a,b) = 1$. По лем. 5.2

$$\operatorname{cont}(g_{\text{red}}h_{\text{red}}) = \operatorname{cont}(g_{\text{red}})\cdot\operatorname{cont}(h_{\text{red}}) = 1\,,$$

т. е. правая часть в (5-7) является редуцированным представлением многочлена f. В силу единственности редуцированного представления элементы a и b обратимы в K, а $f = g_{\text{red}}h_{\text{red}}$ с точностью до умножения на обратимую константу. □

Теорема 5.3
Кольцо многочленов над факториальным кольцом факториально.

Доказательство. Будучи кольцом главных идеалов, кольцо $Q_K[x]$ факториально, и каждый многочлен $f \in K[x] \subset Q_K[x]$ раскладывается в $Q_K[x]$ в произведение неприводимых множителей $f_\nu \in Q_K[x]$. Записывая их в редуцированном виде из лем. 5.3 и сокращая возникающую при этом числовую дробь, получаем равенство $f = \frac{a}{b}\prod f_{\nu,\text{red}}$, в котором все многочлены $f_{\nu,\text{red}} \in K[x]$ неприводимы в $Q_K[x]$ и имеют содержание 1, а числа $a, b \in K$ взаимно просты. Поскольку $\operatorname{cont}\left(\prod f_{\nu,\text{red}}\right) = 1$, правая часть равенства является редуцированным представлением многочлена $f = \operatorname{cont}(f)\cdot f_{\text{red}}$. В силу единственности редуцированного представления, $b = 1$ и $f = a\prod f_{\nu,\text{red}}$ с точностью до умножения на обратимые константы из K. Раскладывая $a \in K$ в произведение неприводимых констант, получаем разложение f в произведение неприводимых множителей в кольце $K[x]$. Докажем единственность такого разложения. Пусть в $K[x]$

$$a_1a_2\cdots a_k\cdot p_1p_2\cdots p_s = b_1b_2\cdots b_m\cdot q_1q_2\cdots q_r\,,$$

где $a_\alpha, b_\beta \in K$ — неприводимые константы, а $p_\mu, q_\nu \in K[x]$ — неприводимые многочлены. Поскольку неприводимые многочлены имеют содержание 1, сравнивая содержание обеих частей, приходим к равенству $a_1a_2\cdots a_k = b_1b_2\cdots b_m$ в K. В силу факториальности K, имеем $k = m$ и (после надлежащей перенумерации сомножителей) $a_i = s_ib_i$, где s_i обратимы. Следовательно, с точностью до умножения на обратимую константу из K в кольце многочленов $K[x]$ выполняется

равенство $p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_r$. В силу факториальности $Q_K[x]$ и неприводимости многочленов p_i и q_i в $Q_K[x]$, мы заключаем, что $r = s$ и после надлежащей перенумерации сомножителей $p_i = q_i$ с точностью до постоянного множителя из Q_K. Из единственности редуцированного представления[1] вытекает, что эти постоянные множители являются обратимыми константами из K. □

Следствие 5.6
Кольцо многочленов $K[x_1, \dots, x_n]$ над факториальным кольцом[2] K факториально. □

5.6. Разложение многочленов с целыми коэффициентами. Разложение многочлена $f \in \mathbb{Z}[x]$ на множители в $\mathbb{Q}[x]$ разумно начать с отыскания его рациональных корней, что делается за конечное число проб.

Упражнение 5.20. Покажите, что несократимая дробь $p/q \in \mathbb{Q}$ является корнем многочлена $a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{Z}[x]$ только если $p \mid a_0$ и $q \mid a_n$.

Точное знание комплексных корней многочлена f тоже весьма полезно.

Упражнение 5.21. Разложите $x^4 + 4$ в $\mathbb{Z}[x]$ в произведение двух квадратных трёхчленов.

После того, как эти простые соображения будут исчерпаны, следует подключать более трудоёмкие способы.

5.6.1. Редукция коэффициентов многочлена $f \in \mathbb{Z}[x]$ по модулю $m \in \mathbb{Z}$

$$\mathbb{Z}[x] \to \frac{\mathbb{Z}}{(m)}[x], \quad a_0 + a_1 x + \cdots + a_n x^n \mapsto [a_0]_m + [a_1]_m x + \cdots [a_n]_m x^n \tag{5-8}$$

приводит все коэффициенты каждого многочлена по модулю m и является гомоморфизмом колец[3]. Поэтому равенство $f = gh$ в $\mathbb{Z}[x]$ влечёт за собой равенства $[f]_m = [g]_n \cdot [h]_m$ во всех кольцах $\big(\mathbb{Z}/(m)\big)[x]$. Таким образом из неприводимости многочлена $[f]_m$ хотя бы при одном m вытекает его неприводимость в $\mathbb{Z}[x]$. Если число $m = p$ простое, кольцо коэффициентов $\mathbb{Z}/(m) = \mathbb{F}_p$ является полем, и кольцо многочленов $\mathbb{F}_p[x]$ в этом случае факториально. При малых p разложение многочлена небольшой степени на неприводимые множители в $\mathbb{F}_p[x]$ можно осуществить простым перебором, и анализ такого разложения может дать существенную информацию о возможном разложении в $\mathbb{Z}[x]$.

Пример 5.7
Покажем, что многочлен $f(x) = x^5 + x^2 + 1$ неприводим в кольце $\mathbb{Z}[x]$. Поскольку у f нет целых корней, нетривиальное разложение $f = gh$ в $\mathbb{Z}[x]$ возможно только с $\deg(g) = 2$ и $\deg(h) = 3$. Сделаем редукцию по модулю 2. Так как у $[f]_2 = x^5 + x^2 + 1$ нет корней и в $\mathbb{F}_2$, оба многочлена $[g]_2$, $[h]_2$ неприводимы в $\mathbb{F}_2[x]$. Но единственный неприводимый многочлен второй степени в $\mathbb{F}_2[x]$ это $x^2 + x + 1$, и $x^5 + x^2 + 1$ на него не делится. Тем самым, $[f]_2$ неприводим над $\mathbb{F}_2$, а значит, и над $\mathbb{Z}$.

Пример 5.8 (критерий Эйзенштейна)
Пусть все коэффициенты приведённого многочлена $f \in \mathbb{Z}[x]$ делятся на простое число $p \in \mathbb{N}$, а младший коэффициент, делясь на p, не делится при этом на p^2. Покажем, что f неприводим в $\mathbb{Z}[x]$. В силу сделанных об f предположений при редукции по модулю p от f остаётся только

[1] См. лем. 5.3 на стр. 75.
[2] В частности, над полем или над областью главных идеалов.
[3] Мы уже пользовались этим в доказательстве лем. 5.2 на стр. 75, см. упр. 5.19.

старший моном $[f(x)]_p = x^n$. Если $f(x) = g(x)h(x)$ в $\mathbb{Z}[x]$, то в силу единственности разложения на простые множители в $\mathbb{F}_p[x]$ оба сомножителя g, h тоже редуцируются в некоторые степени переменной: $[g]_p = x^k$ и $[h]_p = x^m$. Это означает, что все коэффициенты многочленов g и h кроме старшего делятся на p. Тогда младший коэффициент многочлена f, будучи произведением младших коэффициентов многочленов g и h, должен делиться на p^2, что не так.

Пример 5.9 (неприводимость кругового многочлена Φ_p)
Покажем, что при простом $p \in \mathbb{N}$ круговой многочлен

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1 = \frac{x^p - 1}{x - 1}$$

неприводим в $\mathbb{Z}[x]$. Для этого перепишем его как многочлен от переменной $t = x - 1$

$$f(t) = \Phi_p(t+1) = \frac{(t+1)^p - 1}{t} = t^{p-1} + \binom{p}{1} t^{p-2} + \cdots + \binom{p}{p-2} t + \binom{p}{p-1}.$$

Поскольку при простом p все биномиальные коэффициенты $\binom{p}{k}$ с $1 \leqslant k \leqslant p-1$ делятся[1] на p, а свободный член $\binom{p}{p-1} = p$ не делится на p^2, многочлен $f(t)$ неприводим по критерию Эйзенштейна из прим. 5.8. Поэтому и $\Phi_p(x) = f(x-1)$ неприводим.

5.6.2. Алгоритм Кронекера позволяет путём довольно трудоёмкого, но вполне эффективного конечного вычисления либо явно найти разложение заданного многочлена f с целыми коэффициентами в кольце $\mathbb{Z}[x]$, либо убедиться, что f неприводим в $\mathbb{Z}[x]$ (а значит, по лемме Гаусса, и в $\mathbb{Q}[x]$). Будем для определённости считать, что $\deg f = 2n$ или $\deg f = 2n+1$. Тогда в любом нетривиальном разложении $f = gh$ в $\mathbb{Z}[x]$ степень одного из делителей, назовём его h, не превосходит n. Чтобы выяснить, делится ли f в $\mathbb{Z}[x]$ на какой-нибудь многочлен степени не выше n, подставим в f любые $n+1$ различных чисел $z_0, z_1, \ldots, z_n \in \mathbb{Z}$ и выпишем все возможные наборы чисел $d_0, d_1, \ldots, d_n \in \mathbb{Z}$, в которых каждое d_i делит соответствующее $f(z_i)$. Таких наборов имеется конечное число, и набор значений $h(z_0), \ldots, h(z_n)$ многочлена h на числах z_i, если такой многочлен вообще существует, является одним из выписанных нами наборов $d_0, \ldots, d_n$. Для каждого такого набора в $\mathbb{Q}[x]$ есть ровно один многочлен h степени не выше n с $h(z_i) = d_i$ при всех i — это *интерполяционный многочлен Лагранжа*[2]

$$h(x) = \sum_{i=0}^{n} d_i \cdot \prod_{\nu \neq i} \frac{(x - z_\nu)}{(z_i - z_\nu)}. \tag{5-9}$$

Таким образом, делитель h многочлена f, если он существует, является одним из тех многочленов (5-9), что имеют целые коэффициенты. Остаётся явно разделить f на все такие многочлены и либо убедиться, что они не делят f, либо найти среди них делитель f.

[1] См. сл. 2.1 на стр. 27.
[2] См. упр. 3.12 на стр. 40.

§6. Векторы

Всюду в этом параграфе K по умолчанию обозначает коммутативное кольцо с единицей, а $\Bbbk$ — произвольное поле.

6.1. Модули над коммутативными кольцами. Аддитивная абелева группа[1] M называется *модулем* над коммутативным кольцом K или K*-модулем*, если задана операция $K\times M \to M$, которая переводит пары $(x, v) \in K \times M$ в элементы $x \cdot v \in M$ и обладает известными из курса геометрии свойствами умножения векторов на числа[2]:

$$\forall\, x, y \in K \quad \forall\, v \in M \quad x \cdot (y \cdot v) = (xy) \cdot v \tag{6-1}$$

$$\forall\, x, y \in K \quad \forall\, v \in M \quad (x + y) \cdot v = x \cdot v + y \cdot v \tag{6-2}$$

$$\forall\, x \in K \quad \forall\, u, w \in M \quad x \cdot (v + w) = x \cdot v + x \cdot w\,. \tag{6-3}$$

Если в кольце K есть единица и выполняется дополнительное свойство

$$\forall\, v \in V \quad 1 \cdot v = v\,, \tag{6-4}$$

модуль M называется *унитальным*. Всюду в этом параграфе мы по умолчанию рассматриваем именно такие модули. Унитальные модули над полями принято называть *векторными пространствами*. Я очень рассчитываю на то, что читатель уже имеет некоторый опыт работы с векторными пространствами, полученный в параллельном курсе геометрии[3]. Какой бы ни была природа элементов абелевой группы M и кольца K, продуктивно представлять себе первые именно как «векторы», а вторые — как «скаляры». По этой причине мы часто будем называть элементы модуля M *векторами*, а операцию $K \times M \to M$ — *умножением векторов на скаляры* из K. Часто бывает удобно записывать произведение вектора $v \in M$ на скаляр $x \in K$ не как $x \cdot v$, а как $v \cdot x$. По определению, мы считаем эти две записи эквивалентными обозначениями для одного и того же вектора и, как это обычно принято, будем частенько опускать в произведениях точку, считая по умолчанию, что $xv = vx \stackrel{\text{def}}{=} x \cdot v$.

Упражнение 6.1. Выведите из свойств (6-1) – (6-3), что в любом K-модуле M для всех $v \in M$ и $x \in K$ выполняются равенства $0 \cdot v = 0$ и $x \cdot 0 = 0$, а в унитальном модуле над коммутативным кольцом с единицей — равенство[4] $(-1) \cdot v = -v$.

Аддитивная абелева подгруппа $N \subseteq M$ в K-модуле M называется K*-подмодулем*, если она образует K-модуль относительно имеющейся в M операции умножения векторов на скаляры. Для этого необходимо и достаточно, чтобы $xw \in N$ для всех $x \in K$ и $w \in N$. Подмодули $N \subsetneq M$ называются *собственными*. Собственный подмодуль 0, состоящий из одного нуля, называется *тривиальным*.

Пример 6.1 (кольцо как модуль над собой)
Каждое коммутативное кольцо K является модулем над самим собой, где сложение векторов и умножение векторов на скаляры задаются сложением и умножением в K. Если в K имеется

[1] См. n° 2.1.2 на стр. 21.

[2] В роли векторов выступают элементы модуля M, а в роли чисел — элементы кольца K.

[3] Вариант такого курса см. на http://gorod.bogomolov-lab.ru/ps/stud/geom_ru/1617/list.html. Все необходимые нам факты о векторных пространствах будут собраны в n° 6.4 ниже.

[4] Слева стоит произведение вектора $v \in M$ на скаляр $-1 \in K$, а справа — противоположный к v вектор $-v \in M$.

единица, K-модуль K является унитальным. K-подмодули $I \subset K$ — это в точности идеалы кольца K. В частности, коммутативное кольцо K с единицей является полем если и только если в K-модуле K нет нетривиальных собственных подмодулей[1].

Пример 6.2 (координатный модуль K^r)
Декартово произведение r экземпляров кольца K обозначается $K^r = K \times \cdots \times K$ и состоит из строк $a = (a_1, \ldots, a_r)$, в которых $a_i \in K$. Сложение таких строк и их умножение их на скаляры $x \in K$ происходит покоординатно: для $a = (a_1, \ldots, a_r)$, $b = (b_1, \ldots, b_r)$ и $x \in K$ мы полагаем

$$a + b \stackrel{\text{def}}{=} (a_1 + b_1, \ldots, a_r + b_r) \quad \text{и} \quad xa \stackrel{\text{def}}{=} (xa_1, \ldots, xa_r)\,.$$

Пример 6.3 (абелевы группы как $\mathbb{Z}$-модули)
Каждая аддитивно записываемая абелева группа A может рассматриваться как унитальный $\mathbb{Z}$-модуль, в котором сложение векторов есть сложение в A, а умножение векторов на числа $\pm n$, где $n \in \mathbb{N}$, задаётся правилом $(\pm n) \cdot a \stackrel{\text{def}}{=} \pm(a + \cdots + a)$ с n слагаемыми a в скобках.

Упражнение 6.2. Удостоверьтесь, что эти операции удовлетворяют аксиомам (6-1) – (6-4).

6.1.1. Прямые произведения и прямые суммы. Из любого семейства K-модулей M_ν, занумерованных элементами ν произвольного множества $\mathcal{N}$, можно образовать прямое произведение $\prod_{\nu\in\mathcal{N}} M_\nu$, состоящее из всевозможных семейств $v = (v_\nu)_{\nu\in\mathcal{N}}$ векторов $v_\nu \in M_\nu$, занумерованных элементами $\nu \in \mathcal{N}$, как в n° 2.5 на стр. 27. Такие семейства можно поэлементно складывать и умножать на скаляры точно также, как мы это делали в n° 2.5 в прямых произведениях абелевых групп и коммутативных колец. А именно, сумма $v + w$ семейств $v = (v_\nu)_{\nu\in\mathcal{N}}$ и $w = (w_\nu)_{\nu\in\mathcal{N}}$ имеет ν-тым членом элемент $v_\nu + w_\nu$, а на ν-тым членом произведения xv семейства $v = (v_\nu)_{\nu\in\mathcal{N}}$ на скаляр $x \in K$ является элемент xv_ν. Модуль $\prod_{\nu\in\mathcal{N}} M_\nu$ называется *прямым произведением* модулей M_ν, а его подмодуль $\bigoplus_{\nu\in\mathcal{N}} M_\nu$, состоящий из всех семейств $v = (v_\nu)_{\nu\in\mathcal{N}}$ с конечным числом ненулевых векторов v_ν называется *прямой суммой* модулей M_ν. Для конечных множеств $\mathcal{N}$ прямые суммы совпадают с прямыми произведениями. Так, координатный модуль K^r из прим. 6.2 является прямой суммой и прямым произведением r экземпляров K-модуля K.

Пример 6.4 (многочлены и степенные ряды)
Обозначим через Kt^n множество одночленов вида at^n, где $a \in K$, а t — переменная. Каждое множество Kt^n является K-модулем, изоморфным модулю K. Прямая сумма $\bigoplus_{n\geqslant 0} Kt^n$ изоморфна модулю многочленов $K[t]$, а прямое произведение $\prod_{n\geqslant 0} Kt^n$ — модулю формальных степенных рядов $K[\![t]\!]$.

6.1.2. Пересечения и суммы подмодулей. В произвольном K-модуле M пересечение любого множества подмодулей также является подмодулем в M. Пересечение всех подмодулей, содержащих заданное множество векторов $A \subset M$, называется *K-линейной оболочкой* множества A или K-подмодулем, *порождённым* множеством A, и обозначается $\operatorname{span}(A)$ или $\operatorname{span}_K(A)$, если важно подчеркнуть, из какого кольца берутся константы. Линейная оболочка является наименьшим по включению K-подмодулем в M, содержащим A, и может быть иначе описана как множество всех конечных линейных комбинаций $x_1a_1 + \cdots + x_na_n$ векторов $a_i \in A$ с коэффициентами $x_i \in K$, ибо все такие линейные комбинации образуют подмодуль в M и содержатся во всех подмодулях, содержащих A.

[1] См. предл. 5.1 на стр. 66.

В противоположность пересечениям, объединения подмодулей почти никогда не являются подмодулями.

Упражнение 6.3. Покажите, что объединение двух подгрупп в абелевой группе является подгруппой если и только если одна из подгрупп содержится в другой.

K-линейная оболочка объединения произвольного множества подмодулей $U_\nu \subset M$ называется *суммой* этих подмодулей и обозначается $\sum_\nu U_\nu \stackrel{\text{def}}{=} \operatorname{span} \bigcup_\nu U_\nu$. Таким образом, сумма подмодулей представляет собою множество всевозможных конечных сумм векторов, принадлежащих этим подмодулям. Например,

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1\,,\ u_2 \in U_2\}$$
$$U_1 + U_2 + U_3 = \{u_1 + u_2 + u_3 \mid u_1 \in U_1\,,\ u_2 \in U_2\,,\ u_3 \in U_3\} \quad \text{и т. д.}$$

Если подмодули $U_1, \dots, U_m \subset M$ таковы, что гомоморфизм сложения

$$U_1 \oplus \dots \oplus U_n \to U_1 + \dots + U_n \subset M\,, \quad (u_1, \dots, u_n) \mapsto u_1 + \dots + u_n\,, \tag{6-5}$$

является биекцией между $U_1 \oplus \dots \oplus U_n$ и $U_1 + \dots + U_n$, то сумму $U_1 + \dots + U_n$ называют *прямой* и тоже обозначают $U_1 \oplus \dots \oplus U_n$, как и в n° 6.1.1 выше. Биективность отображения (6-5) эквивалентна тому, что каждый вектор $w \in U_1 + \dots + U_n$ имеет *единственное* разложение $w = u_1 + \dots + u_n$, в котором $u_i \in U_i$ при каждом i.

Предложение 6.1
Сумма подмодулей $U_1, \dots, U_n \subset V$ является прямой если и только если каждый из подмодулей имеет нулевое пересечение с суммой всех остальных. В частности, сумма $U+W$ двух подмодулей прямая тогда и только тогда, когда $U \cap W = 0$.

Доказательство. Обозначим через W_i сумму всех подмодулей U_ν за исключением i-того. Если пересечение $U_i \cap W_i$ содержит ненулевой вектор $u_i = u_1 + \dots + u_{i-1} + u_{i+1} + \dots + u_n$, где $u_i \in U_i$ при всех i, то у этого вектора имеется два различных представления[1]

$$0 + \dots + 0 + u_i + 0 + \dots + 0 = u_1 + \dots + u_{i-1} + 0 + u_{i+1} + \dots + u_n\,.$$

Поэтому такая сумма не прямая. Наоборот, если $U_i \cap W_i = 0$ при всех i, то переписывая равенство

$$u_1 + \dots + u_n = w_1 + \dots + w_n\,,\ \text{где } u_\nu, w_\nu \in U_\nu \text{ при всех } i,$$

как $u_i - w_i = \sum_{\nu \neq i}(w_\nu - u_\nu)$, видим, что этот вектор лежит в $U_i \cap W_i = 0$. Поэтому $u_i = w_i$ для каждого $i = 1, \dots, n$. □

Следствие 6.1
Для того чтобы модуль M распадался в прямую сумму собственных подмодулей $L, N \subset M$ необходимо и достаточно, чтобы $L + N = M$ и $L \cap N = 0$. □

[1] В левом отлично от нуля только i-е слагаемое, а в правом оно нулевое.

6.1.3. Фактор модули. Для любых K-модуля M подмодуля $N \subseteq M$ можно образовать *фактор модуль* M/N, состоящий из классов $[m]_N = m + N = m \pmod N = \{m' \in M \mid m' - m \in N\} \subset M$, представляющих собою аддитивные сдвиги подмодуля N на всевозможные элементы $m \in M$ или, что тоже самое, классы эквивалентности по отношению $m \equiv n \pmod N$ сравнимости по модулю N, означающему, что $m' - m \in N$. Сложение классов и их умножение на элементы кольца определяются обычными формулами $[m_1]_N + [m_2]_N \stackrel{\text{def}}{=} [m_1 + m_2]_N$ и $x \cdot [m]_N \stackrel{\text{def}}{=} [xm]_N$.

Упражнение 6.4. Проверьте, что отношение сравнимости по модулю N является эквивалентностью, а операции корректно определены и удовлетворяют аксиомам (6-1) – (6-4).

В частности, фактор кольцо K/I кольца K по идеалу $I \subset K$ является фактором K-модуля K по его K-подмодулю I, ср. с прим. 6.1 выше.

Упражнение 6.5. Пусть модуль M является прямой суммой $M = L \oplus N$ подмодулей $L, N \subset M$. Покажите, что $M/N \simeq L$ и $M/L \simeq N$.

Пример 6.5 (фактор модуля по идеалу кольца)
Для любого идеала $I \subset K$ и произвольного K-модуля M обозначим через

$$IM \stackrel{\text{def}}{=} \{x_1a_1 + \cdots x_na_n \mid x_i \in I,\ a_i \in M,\ n \in \mathbb{N}\}$$

K-подмодуль, образованный всевозможными линейными комбинациями элементов модуля M с коэффициентами из идеала I.

Упражнение 6.6. Проверьте, что IM действительно является K-подмодулем в M.

Фактор модуль M/IM обладает канонической структурой модуля над фактор кольцом K/I, которая корректно задаётся правилом $[x]_I \cdot [w]_{[IM]} = [xw]_{[IM]}$, где $[x]_I$ и $[a]_{[IM]}$ означают классы элементов $\lambda \in K$ и $w \in M$ соответственно по модулю идеала $I \subset K$ и подмодуля $IM \subset M$.

Упражнение 6.7. Убедитесь, что это правило корректно, и если $M = N_1 \oplus \cdots \oplus N_m$, то

$$IM = IN_1 \oplus \cdots \oplus IN_m \quad \text{и} \quad M/IM = (N_1/IN_1) \oplus \cdots \oplus (N_m/IN_m)$$

для любого идеала $I \subset K$. В частности, $K^n/IK^n = (K/I)^n$.

Пример 6.6 (кручение)
Элемент m модуля M над целостным[1] кольцом K называется *элементом кручения*, если $xm = 0$ для некоторого ненулевого $x \in K$. Например, любой класс $[k] \in \mathbb{Z}/(n)$ является элементом кручения в $\mathbb{Z}$-модуле $\mathbb{Z}/(n)$, поскольку $n[k] = [nk] = [0]$.

Упражнение 6.8. Убедитесь, что элементы кручения составляют подмодуль в M.

Этот подмодуль обозначается $\operatorname{Tors} M \stackrel{\text{def}}{=} \{m \in M \mid \exists\, x \neq 0 : xm = 0\}$ и называется *подмодулем кручения*. Если $\operatorname{Tors} M = 0$, то говорят, что модуль M *не имеет кручения*. Например, любой идеал целостного кольца K и любой подмодуль в координатном модуле K^n над таким кольцом не имеют кручения. Если $\operatorname{Tors} M = M$, то M называется *модулем кручения*. Например, фактор K/I по любому ненулевому идеалу $I \subset K$ является K-модулем кручения, поскольку для любого класса $[a] \in K/I$ и любого ненулегого $x \in I$ класс $x[a] = [xa] = [0]$, так как $xa \in I$.

Предложение 6.2
Для любого модуля M над целостным кольцом K фактор модуль $M/\operatorname{Tors}(M)$ не имеет кручения.

[1] См. n° 2.4.1 на стр. 26.

Доказательство. При ненулевом $x \in K$ равенство $x[m] = [xm] = [0]$ в $M/\operatorname{Tors}(M)$ означает, что $xm \in \operatorname{Tors}(M)$, т. е. $yxm = 0$ для некоторого ненулевого $y \in K$. Поскольку $xy \neq 0$, так как в кольце K нет делителей нуля, $m \in \operatorname{Tors} M$ и $[m] = [0]$. □

6.1.4. Дополнительные подмодули и разложимость. Подмодули $L, N \subset M$ называются *дополнительными*, если $M = L \oplus N$. В этой ситуации модуль M называется *разложимым*, а про подмодули L, N говорят, что они *отщепляются* от M прямыми слагаемыми. Модуль M, не представимый в виде прямой суммы своих собственных подмодулей называется *неразложимым*. Например, $\mathbb{Z}$-модуль $\mathbb{Z}$ неразложим, хотя и имеет собственные $\mathbb{Z}$-подмодули. В самом деле, каждый собственный подмодуль $I \subset \mathbb{Z}$ представляет собою главный идеал $I = (d)$. Согласно упр. 6.5, разложение $\mathbb{Z} = (d) \oplus N$ означает наличие в $\mathbb{Z}$ подмодуля $N \subset \mathbb{Z}$, изоморфного модулю кручения $\mathbb{Z}/(d)$. Но это невозможно, поскольку в $\mathbb{Z}$ нет кручения.

Упражнение 6.9. Рассмотрим $\mathbb{Z}$-подмодуль $N \subset \mathbb{Z}^2$, порождённый векторами $(2,1)$ и $(1,2)$. Покажите, что $N \simeq \mathbb{Z}^2$, $M/N \simeq \mathbb{Z}/(3)$, и не существует подмодуля $L \subset M$, такого что $M = L \oplus N$.

6.2. Гомоморфизмы модулей. Отображение $\varphi : M \to N$ между K-модулями M и N называется *K-линейным* или *гомоморфизмом K-модулей*, если оно перестановочно со сложением векторов и умножением векторов на скаляры, т. е. для всех $x \in K$ и $u, w \in M$

$$\varphi(u + w) = \varphi(u) + \varphi(w) \quad \text{и} \quad \varphi(xu) = x\varphi(u). \tag{6-6}$$

Поскольку K-линейное отображение $\varphi : M \to N$ является гомоморфизмом абелевых групп, оно обладает всеми свойствами из n° 2.6 на стр. 28. В частности, $\varphi(0) = 0$ и $\varphi(-u) = -\varphi(u)$ для всех $u \in M$, а инъективность φ равносильна тому, что ядро

$$\ker \varphi = \varphi^{-1}(0) = \{u \in M \mid \varphi(u) = 0\}$$

состоит из одного нуля. Все непустые слои любого K-линейного гомоморфизма φ являются аддитивными сдвигами его ядра, т. е. $\varphi^{-1}(\varphi(u)) = u + \ker \varphi$ для всех $u \in M$.

Упражнение 6.10. Убедитесь, что ядро и образ K-линейного гомоморфизма $\varphi : M \to N$ являются подмодулями в M и в N соответственно, а сопоставление $[v]_{\ker \varphi} \mapsto \varphi(v)$ корректно задаёт изоморфизм K-модулей $M/\ker \varphi \to \operatorname{im} \varphi$.

Предостережение 6.1. Именуемое в школе «линейной функцией» отображение $\varphi : K \to K$, задаваемое правилом $\varphi(x) = ax + b$, где $a, b \in K$ фиксированы, является K-линейным в смысле предыдущего определения только при $b = 0$. Если же $b \neq 0$, то φ не перестановочно ни со сложением, ни с умножением на числа.

Пример 6.7 (дифференцирование)
Кольцо многочленов $K[x]$ с коэффициентами в коммутативном кольце K можно рассматривать и как K-модуль. Оператор дифференцирования $D = \frac{d}{dx} : K[x] \to K[x]$, $f(x) \mapsto f'(x)$, является гомоморфизмом K-модулей, поскольку перестановочен со сложением многочленов и умножением многочленов на константы, но не является гомоморфизмом колец, так как не перестановочен с умножением многочленов друг на друга.

6.2.1. Модули гомоморфизмов. Отображения $Z \to M$ из любого множества Z в произвольный K-модуль M можно складывать и умножать на числа из K, применяя эти операции к значениям рассматриваемых отображений в каждой точке $z \in Z$. А именно, для любой пары отображений $\varphi, \psi : X \to M$ и числа $x \in K$ сумма $\varphi+\psi : X \to M$ и произведение $x\varphi : X \to M$ действуют на точки $z \in Z$ по правилам

$$\varphi + \psi : z \mapsto \varphi(z) + \psi(z) \quad \text{и} \quad x\varphi : z \mapsto x\varphi(z). \tag{6-7}$$

Эти операции очевидно удовлетворяют аксиомам (6-1) – (6-4), поскольку все эти аксиомы выполняются в модуле M и проверяются отдельно над каждой точкой $z \in Z$. Таким образом, множество M^Z всех отображений $Z \to M$ является K-модулем. Нулевым элементом этого модуля служит *нулевое отображение*, переводящее все элементы множества Z в нуль.

Упражнение 6.11. Убедитесь, что K-модуль M^Z изоморфен прямому произведению[1] $\prod_{z\in Z} M_z$ одинаковых копий $M_z = M$ модуля M, занумерованных элементами $z \in Z$.

Если множество Z тоже является K-модулем, то сумма $\varphi + \psi$ двух K-линейных отображений $\varphi, \psi : N \to M$ и произведение $z\varphi$ гомоморфизма φ с любым скаляром $x \in K$ тоже K-линейны.

Упражнение 6.12. Убедитесь в этом.

Таким образом, K-линейные отображения K-модуля N в K-модуль M составляют в модуле M^N всех отображений из N в M K-подмодуль. Он обозначается $\mathrm{Hom}_K(M, N)$ и называется *модулем K-линейных гомоморфизмов* из M в N.

Упражнение 6.13. Покажите, что композиция K-линейных гомоморфизмов тоже K-линейна.

Пример 6.8 (гомоморфизмы абелевых групп)
Как мы видели в прим. 6.3 на стр. 80, любые две абелевы группы A и B могут рассматриваться как модули над кольцом $\mathbb{Z}$.

Упражнение 6.14. Убедитесь, что отображение множеств $A \to B$ является гомоморфизмом абелевых групп[2] если и только если оно $\mathbb{Z}$-линейно.

В аддитивной абелевой группе вычетов $\mathbb{Z}/(m)$, рассматриваемой как $\mathbb{Z}$-модуль, описанное в прим. 6.3 умножение класса $[k]_m \in \mathbb{Z}/(m)$ на число $z \in \mathbb{Z}$ происходит по правилу $z \cdot [k]_m = [zk]_m$. Тем самым, каждый класс $[k]_m = k \cdot [1]_m$ можно получить, умножая класс $[1]_m$ на подходящее целое число. Поэтому любой $\mathbb{Z}$-линейный гомоморфизм $\varphi : \mathbb{Z}/(m) \to N$ в произвольный $\mathbb{Z}$-модуль N однозначно восстанавливается по вектору $\varphi([1]_m) \in N$: значение φ на произвольном классе $[k]_m$ будет равно $\varphi([k]_m) = \varphi(k \cdot [1]_m) = k\varphi([1]_m)$. При этом вектор $\varphi([1]_m) \in N$ не может быть выбран произвольно: так как в $\mathbb{Z}/(m)$ выполняется соотношение $m \cdot [1]_m = [m]_m = 0$, в модуле N должно выполняться соотношение $m \cdot \varphi([1]_m) = \varphi(m \cdot [1]_m) = \varphi(0) = 0$. В частности, если в модуле N нет ненулевых векторов v с $mv = 0$, то $\varphi([1]_m) = 0$, и это означает, что из $\mathbb{Z}/(m)$ в N нет никаких $\mathbb{Z}$-линейных отображений, кроме нулевого. Например, это так для $N = \mathbb{Z}/(n)$, если $\text{нод}(m, n) = 1$: в этом случае класс $[m]_n$ обратим в кольце $\mathbb{Z}/(n)$ и равенство $[0]_n = m[k]_n = [mk]_n = [m]_n[k]_n$ возможно только при[3] $[k]_n = [0]_n$. Мы заключаем, что $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/(m), \mathbb{Z}/(n)) = 0$ при $\text{нод}(m, n) = 1$, т. е. любой гомоморфизм абелевых групп $\mathbb{Z}/(m) \to \mathbb{Z}/(n)$ имеет при взаимно простых m и n нулевой образ.

[1] См. n° 6.1.1 на стр. 80.
[2] См. n° 2.6 на стр. 28.
[3] Чтобы убедиться в этом, надо умножить левую и правую части равенства в кольце $\mathbb{Z}/(n)$ на класс $[m]_n^{-1}$.

Упражнение 6.15. Покажите, что любой линейный гомоморфизм $\varphi : M \to N$ в свободный от кручения модуль N переводит $\mathrm{Tors}(M)$ в нуль.

Предложение 6.3
Для любых K-модулей M, N и подмодуля $L \subset M$ гомоморфизмы $f : M \to N$, тождественно зануляющиеся на подмодуле L, образуют в K-модуле $\mathrm{Hom}_K(M, N)$ подмодуль, изоморфный K-модулю $\mathrm{Hom}_K(M/L, N)$. Изоморфизм сопоставляет зануляющемуся на L гомоморфизму $f : M \to N$ гомоморфизм $f_L : M/L \to M$, корректно задаваемый правилом $[v]_L \mapsto f(v)$. Обратный изоморфизм сопоставляет K-линейному отображению $g : M/L \to M$ его композицию $f = g\pi_L$ с эпиморфизмом факторизации $\pi_L : M \twoheadrightarrow M/L$.

Доказательство. Если гомоморфизмы $f, g : M \to N$ переводят L в нуль, то любая их K-линейная комбинация $xf + yg$ тоже переводит L в нуль. Следовательно, такие гомоморфизмы образуют K-подмодуль в $\mathrm{Hom}_K(M, N)$. Если $f : M \to N$ переводят L в нуль, то правило $f_L : [v]_L \mapsto f(v)$ корректно задаёт гомоморфизм $f_L : M/L \to M$, поскольку для любого вектора $w = v + u$ с $u \in L$ имеем $f(w) = f(v) + f(u) = f(v)$, ибо $f(u) = 0$. Сопоставление $f \mapsto f_L$ задаёт K-линейный гомоморфизм $\mathrm{Hom}_K(M, N) \to \mathrm{Hom}_K(M/L, N)$ с нулевым ядром. Он сюръективен, поскольку любой K-линейный гомоморфизм $g : M/L \to N$ имеет вид $g = f_L$ для K-линейного гомоморфизма $f : M \to N$, который действует по правилу $f(v) = g([v]_L)$ и является композицией g с гомоморфизмом факторизации $\pi_L : M \twoheadrightarrow M/L$. □

Предложение 6.4
Рассмотрим семейство K-модулей M_μ, занумерованных элементами μ произвольного множества $\mathcal{M}$. Для любого K-модуля N имеется канонический изоморфизмом K-модулей

$$\prod_{\mu\in\mathcal{M}} \mathrm{Hom}_K(M_\mu, N) \xrightarrow{\sim} \mathrm{Hom}_K\Big(\bigoplus_{\mu\in\mathcal{M}} M_\mu,\, N\Big), \tag{6-8}$$

который переводит семейство K-линейных гомоморфизмов $f_\mu : M_\mu \to N$ в гомоморфизм

$$\bigoplus f_\mu : \bigoplus_{\mu\in\mathcal{M}} M_\mu \to N, \tag{6-9}$$

отображающий каждое семейство векторов $(w_\mu)_{\mu\in\mathcal{M}}$ с конечным числом ненулевых членов в сумму $\sum_{\mu\in\mathcal{M}} f_\mu(w_\mu)$ с конечным числом ненулевых слагаемых.

Доказательство. Отображение (6-8) очевидно является K-линейным гомоморфизмом. Обратное к (6-8) отображение переводит каждый K-линейный гомоморфизм $: \bigoplus_{\mu\in\mathcal{M}} M_\mu \to N$ в семейство гомоморфизмов $f_\mu : M_\mu \to N$, где каждый $f_\nu = f\iota_\nu$ является композицией f с вложением $\iota_\nu : M_\nu \hookrightarrow \bigoplus_{\mu\in\mathcal{M}} M_\mu$, отправляющем каждый вектор $u \in M_\nu$ в семейство $(w_\mu)_{\mu\in\mathcal{M}} \in$ $\in \bigoplus_{\mu\in\mathcal{M}} M_\mu$, в котором $w_\nu = u$ и $w_\mu = 0$ при $\mu \neq \nu$. □

Пример 6.9 (продолжение прим. 6.4 на стр. 80)
В прим. 6.4 мы видели, что модуль многочленов $K[t] \simeq \bigoplus_{n\geqslant 0} Kt^n$ можно воспринимать как прямую сумму модулей $Kt^n \simeq K$. Применительно к этому случаю предл. 6.4 утверждает, среди прочего, что каждое K-линейное отображение $f : K[t] \to K$ однозначно задаётся последовательностью K-линейных отображений $f_n = f|_{Kt^n} : Kt^n \to K$ — ограничений отображения f на подмодули $Kt^n \subset K[t]$. Каждое отображение f_n в свою очередь однозначно задаётся

числом $\varphi_n = f_n(t^n) = f(t^n)$ — значением отображения f на базисном мономе t^n. Последовательность чисел $\varphi_n \in K$ может быть любой, и отвечающее такой последовательности K-линейное отображение $f\colon K[t] \to K$ переводит многочлен $a(t) = a_0 + a_1 t + \cdots + a_m t^m$ в число $f(a) = \varphi_0 a_0 + \varphi_1 a_1 + \cdots + \varphi_m a_m$. Таким образом, модуль $\mathrm{Hom}_K(K[t], K)$ изоморфен прямому произведению счётного множества копий модуля K, т. е. модулю формальных степенных рядов $K[\![x]\!]$. Изоморфизм сопоставляет последовательности (φ_n) её *производящую функцию* $\Phi(x) = \sum_{n\geqslant 0} \varphi_n x^n \in K[\![x]\!]$. Например, для любого $\alpha \in K$ *гомоморфизм вычисления*

$$\mathrm{ev}_\alpha\colon K[t] \to K\,, \quad f \mapsto f(\alpha)\,,$$

сопоставляющий многочленам их значения в точке $\alpha \in K$ и действующий на базисные мономы по правилу $t^n \mapsto \alpha^n$, имеет $\varphi_n = \alpha^n$ и задаётся рядом $\sum_{n\geqslant 0} \alpha^n x^n = (1-\alpha x)^{-1} \in K[\![x]\!]$.

6.3. Образующие и соотношения. Говорят, что вектор v из K-модуля M *линейно выражается* над K через векторы $w_1, \ldots, w_m$, если $v = x_1 w_1 + \cdots x_m w_m$ для некоторых $x_1, \ldots, x_m \in K$. Правая часть этой формулы называется *линейной комбинацией* векторов $w_i \in V$ с коэффициентами $x_i \in K$. Линейная комбинация, в которой все коэффициенты $x_i = 0$, называется *тривиальной*.

Мы говорим, что множество $Z \subset M$ *порождает* модуль M, если любой вектор $v \in M$ является линейной комбинацией конечного числа векторов из Z, т. е. $v = x_1 u_1 + \cdots + x_m u_m$ для некоторых $x_i \in K$, $w_i \in G$ и $m \in \mathbb{N}$. Множество векторов $Z \subset M$ называется *линейно зависимым*, если некоторая нетривиальная конечная линейная комбинация векторов из Z обращается в нуль, т. е. существуют такие $k \in \mathbb{N}$, $u_1, \ldots, u_k \in Z$ и $x_1, \ldots, x_k \in K$, что $x_1 u_1 + \cdots + x_k u_k = 0$, но при этом не все x_i равны нулю. Каждая такая такая линейная комбинация называется *линейным соотношением* на векторы из множества Z.

Упражнение 6.16. Покажите, что в модуле без кручения сумма подмодулей $U_1, \ldots, U_m$ прямая[1] если и только если любой набор ненулевых векторов $u_1, \ldots, u_m$, в котором $u_i \in U_i$ при каждом i, линейно независим.

Множество $E \subset M$ называется *базисом* модуля M, если каждый вектор $v \in M$ единственным образом линейно выражается через векторы из E, т. е. $v = \sum_{e\in E} x_e e$, где все $x_e \in K$ и только конечное множество из них отлично от нуля, и равенство $\sum_{e\in E} x_e e = \sum_{e\in E} y_e e$ двух таких сумм с конечным числом ненулевых слагаемых равносильно равенству коэффициентов $x_e = y_e$ при каждом векторе $e \in E$. Коэффициенты x_e единственного линейного выражения вектора v через базисные векторы $e \in E$ называются *координатами* вектора v в базисе E.

Модуль M, обладающий базисом, называется *свободным*. Иначе можно сказать, что свободный модуль с базисом E представляет собою прямую сумму $\bigoplus_{e\in E} Ke$ одинаковых копий $Ke = K$ модуля K, занумерованных элементами $e \in E$. В частности, свободный модуль над целостным кольцом K не имеет кручения[2].

Лемма 6.1
Множество векторов $E \subset M$ тогда и только тогда является базисом K-модуля M, когда оно линейно независимо и порождает M.

Доказательство. Пусть множество векторов E порождает K-модуль M. Если существует линейное соотношение $x_1 e_1 + \cdots + x_n e_n = 0$, в котором $e_i \in E$ и $x_1 \neq 0$, то оно у нулевого вектора

[1] См. n° 6.1.2 на стр. 80.
[2] См. прим. 6.6 на стр. 82.

$0 \in M$ имеется два различных представления в линейной комбинации векторов из E: первое даётся указанным соотношением, второе имеет вид $0 = 0 \cdot e_1$. Наоборот, если множество E линейно независимо и имеется равенство $\sum_{e\in E} x_e e = \sum_{e\in E} y_e e$, в обоих частях которого имеется лишь конечное число ненулевых коэффициентов, то перенося все ненулевые слагаемые в одну часть, получаем конечное линейное соотношение $\sum_{e\in E}(x_e - y_e)\cdot e = 0$, возможное только если все коэффициенты нулевые, т. е. только когда $x_e = y_e$ при всех e. □

Пример 6.10 (примеры несвободных модулей)
Аддитивная группа вычетов $\mathbb{Z}/(m)$, рассматриваемая как $\mathbb{Z}$-модуль в духе прим. 6.8 на стр. 84, не свободна, поскольку в свободном модуле нет кручения. Модуль $\mathbb{Z}/(m)$ порождается над $\mathbb{Z}$ одним вектором $[1]_m$, и этот вектор линейно зависим, поскольку удовлетворяет нетривиальному линейному соотношению $m \cdot [1]_m = 0$.

Упражнение 6.17. Покажите, что класс $[n]_m \in \mathbb{Z}/(m)$ порождает $\mathbb{Z}$-модуль $\mathbb{Z}/(m)$ если и только если $\text{нод}(m, n) = 1$.

Идеал I целостного кольца K, рассматриваемый как K-модуль, свободен если и только если он главный. В самом деле, образующая d главного идеала $I = (d)$ порождает его как K-модуль и линейно независима в силу целостности кольца. Напротив, если идеал $I \subset K$ не является главным, то любой порождающий его набор элементов линейно зависим, поскольку любые два различных элемента $a, b \in K$ линейно зависимы над K, ибо удовлетворяют линейному соотношению $b \cdot a - a \cdot b = 0$. Например, в кольце $K = \mathbb{Q}[x, y]$ многочленов с рациональными коэффициентами идеал $I = (x, y)$, состоящий из многочленов без свободного члена, порождается над $\mathbb{Q}[x, y]$ векторами x и y, которые линейно зависимы над $\mathbb{Q}[x, y]$, ибо $y \cdot x - x \cdot y = 0$, и не может быть порождён одним вектором, поскольку x и y не имеют необратимых общих делителей.

Пример 6.11 (многочлены и ряды, продолжение прим. 6.9 на стр. 85)
Кольцо многочленов $K[t]$ является свободным модулем со счётным базисом из мономов t^n, так как каждый многочлен по определению является конечной K-линейной комбинацией каких мономов, и равенство многочленов означает равенство их коэффициентов при каждом мономе. Иначе говоря, модуль $K[t]$ является прямой суммой модулей $Kt^n \simeq K$. В модуле формальных степенных рядов $K[\![t]\!]$, который является прямым произведением тех же самых модулей Kt^n, мономы t^n базиса уже *не образуют*, поскольку никакой ряд с бесконечным числом ненулевых коэффициентов не является конечной линейной комбинацией мономов.

Упражнение 6.18. Покажите, что при $K \neq 0$ модуль $K[\![t]\!]$ не порождается никаким счётным множеством векторов.

В кольце $\mathbb{R}[\![t]\!]$ несложно предъявить несчётное линейно независимое множество векторов. Например, геометрические прогрессии $(1-\alpha t)^{-1} = 1+\alpha t+\alpha^2 t^2+\ldots$, где α пробегает $\mathbb{R}$, линейно независимы, поскольку равенство $x_1(1-\alpha_1 t)^{-1} + \cdots + x_k(1-\alpha_k t)^{-1} = 0$ в кольце $\mathbb{R}[\![t]\!]$ после приведения к общему знаменателю превращается в равенство

$$x_1 \prod_{\nu\neq 1}(1-\alpha_\nu t) + x_2 \prod_{\nu\neq 2}(1-\alpha_\nu t) + \cdots + x_k \prod_{\nu\neq k}(1-\alpha_\nu t) = 0$$

в кольце $\mathbb{R}[t]$. Последовательно подставляя в него значения $t = 1/\alpha_i$, мы заключаем, что $x_i = 0$ для каждого $i = 1, \ldots, k$.

Пример 6.12 (задание модуля образующими и соотношениями)
Рассмотренный нами в прим. 6.2 на стр. 80 координатный модуль K^n свободен, поскольку каждый вектор $v = (x_1, \dots, x_n)$ единственным образом представляется в виде линейной комбинации $v = x_1e_1 + \cdots + x_ne_n$ стандартных базисных векторов

$$e_i = (0, \dots, 0, 1, 0, \dots, 0), \tag{6-10}$$

единственной ненулевой координатой которых является единица, стоящая у вектора e_i на i-том месте. Если K-модуль M линейно порождается над K векторами $w_1, \dots, w_m$, то имеется K-линейный эпиморфизм $\pi\colon K^m \twoheadrightarrow M$, $(x_1, \dots, x_n) \mapsto x_1w_1 + \cdots + x_mw_m$. Его ядро $R = \ker\pi$ называется *модулем соотношений* между образующими w_i, поскольку оно состоит из всех таких строчек чисел $(x_1, \dots, x_n) \in K^m$, которые задают линейное соотношение $x_1w_1 + \cdots + x_mw_m = 0$ между образующими w_i в модуле M. Таким образом, каждый конечно порождённый K-модуль M имеет вид $M = K^m/R$ для некоторого числа $m \in \mathbb{N}$ и некоторого подмодуля $R \subset K^m$.

6.4. Векторные пространства. В этом разделе собраны необходимые для дальнейшего свойства векторных пространств. Поскольку большинство из них, скорее всего, уже обсуждались в курсе геометрии, обращаться к этому разделу можно лишь по мере необходимости.

Если кольцо скаляров представляет собою поле $\Bbbk$, то наличие $\Bbbk$-линейной зависимости между теми или иными векторами равносильна возможности линейно выразить один этих векторов через остальные. Скажем, если в линейном соотношении $x_1w_1 + \cdots + x_mw_m = 0$ коэффициент $x_m \neq 0$, то

$$v_m = -\frac{x_1}{x_m}\,v_1 - \cdots - \frac{x_{m-1}}{x_m}\,v_{m-1},$$

и аналогичное линейное выражение можно получить для любого вектора, входящего в линейное соотношение с ненулевым коэффициентом. По этой причине каждое векторное пространство над любым полем свободно, т. е. обладает базисом.

Теорема 6.1 (существование базиса)
В каждом отличном от нуля векторном пространстве V для любого[1] линейно независимого множества векторов A и любого[2] линейно порождающего V множества векторов $B \supset A$ существует базис E, содержащий A и содержащийся в B.

Доказательство. Линейно независимые множества векторов $X \subseteq V$ со свойством $A \subseteq X \subseteq B$ образуют частично упорядоченное отношением включения множество, удовлетворяющее лемме Цорна[3]. А именно, в качестве верхней грани линейно упорядоченной цепи вложенных друг в друга линейно независимых множеств можно взять их объединение. Оно линейно независимо, поскольку все векторы в любой конечной линейной комбинации векторов из такого объединения лежат в одном достаточно большом множестве цепочки, а оно линейно независимо. По лемме Цорна существует такое линейно независимое множество E со свойством $A \subseteq E \subseteq B$, что для любого линейно независимого множества X со свойством $A \subseteq X \subseteq B$ включение $E \subseteq X$ влечёт равенство $E = X$. Покажем, что E линейно порождает V. Для этого достаточно убедиться, что каждый вектор $b \in B \smallsetminus E$ линейно выражается через E. Так как множество $E \cup \{b\}$ строго больше E, оно линейно зависимо. Поскольку само множество E линейно независимо, всякая

[1] В том числе пустого.
[2] В том числе совпадающего со всем V.
[3] См. сл. 1.1 на стр. 19.

линейная зависимость между векторами из $E \cup \{b\}$ содержит с ненулевым коэффициентом вектор b. Тем самым, он линейно выражается через векторы из E. □

Следствие 6.2
Каждое ненулевое векторное пространство имеет базис, и любой базис любого подпространства можно дополнить до базиса во всём пространстве. □

6.4.1. Размерность. Все базисы любого векторного пространства V над полем $\mathbb{k}$ равномощны. Для векторного пространства V, которое линейно порождается конечным набором векторов, это вытекает из следующей леммы.

Лемма 6.2 (лемма о замене)
Если векторы $w_1, \dots, w_m$ линейно порождают векторное пространство V над полем $\mathbb{k}$, а векторы $u_1, \dots, u_k \in V$ линейно независимы, то $m \geqslant k$ и векторы w_i можно перенумеровать так, что набор векторов $u_1, \dots, u_k, w_{k+1}, w_{k+2}, \dots, w_m$, полученный заменой первых k векторов w_i векторами u_i, тоже порождает V.

Доказательство. Пусть $u_1 = x_1 w_1 + \dots + x_m w_m$. Так как векторы u_i линейно независимы, $u_1 \neq 0$ и среди коэффициентов x_i есть хоть один ненулевой. Перенумеруем векторы w_i так, чтобы $x_1 \neq 0$. Поскольку вектор w_1 линейно выражается через u_1 и $w_2, \dots, w_m$ как

$$w_1 = \frac{1}{x_1} u_1 - \frac{x_2}{x_1} w_2 - \dots - \frac{x_m}{x_1} w_m ,$$

векторы $u_1, w_2, \dots, w_m$ порождают V. Далее действуем по индукции. Пусть для очередного $i < k$ векторы $u_1, \dots, u_i, w_{i+1}, \dots, w_m$ порождают V. Тогда

$$u_{i+1} = y_1 w_1 + \dots + y_m w_m + x_{i+1} w_{i+1} + \dots + x_m w_m .$$

В силу линейной независимости векторов u_ν вектор u_{i+1} нельзя линейно выразить только через векторы $u_1, \dots, u_k$. Поэтому в предыдущем разложении присутствует с ненулевым коэффициентом хоть один из оставшихся векторов w_j. Следовательно, $m > i$ и мы можем занумеровать оставшиеся w_j так, чтобы $x_{i+1} \neq 0$. Теперь, как и на первом шагу, вектор w_{i+1} линейно выражается через векторы $u_1, \dots, u_{i+1}, w_{i+2}, \dots, w_m$. Тем самым, эти векторы линейно порождают V, что воспроизводит индуктивное предположение. □

Следствие 6.3
Если векторное пространство V обладает базисом из n векторов, то каждый базис пространства V состоит из n векторов, и всякий линейно независимый набор из n векторов, а также всякий порождающий набор из n векторов являются базисами.

Доказательство. Так как каждый базис одновременно линейно независим и порождает[1] V, все базисы состоят из одинакового количества векторов по лем. 6.2. По той же лемме при замене любого базиса любыми n линейно независимыми векторами получится порождающий набор, т. е. тоже базис. По теор. 6.1 любой порождающий набор из n векторов содержит в себе базис. Так как последний тоже состоит из n векторов, он совпадает с исходным набором. □

[1] См. лем. 6.1 на стр. 86.

Определение 6.1
Векторные пространства с конечными базисами называются *конечномерными*. Количество векторов в базисе такого пространства V называется *размерностью* пространства V и обозначается $\dim V$.

Следствие 6.4
В конечномерном пространстве V каждое векторное подпространство $U \subset V$ тоже конечномерно, и $\dim U \leqslant \dim V$, где равенство возможно только при $U = V$. □

Замечание 6.1. (координатные модели конечномерного пространства) Каждое n-мерное векторное пространство V над полем $\Bbbk$ изоморфно координатному пространству $\Bbbk^n$. При этом $\Bbbk$-линейные изоморфизмы $\Bbbk^n \xrightarrow{\sim} V$ взаимно однозначно соответствуют базисам в V, поскольку для любого базиса $v_1, \dots, v_n$ в V отображение

$$f : \Bbbk^n \to V, \quad (x_1, \dots, x_n) \mapsto x_1 v_1 + \dots + x_n v_n, \tag{6-11}$$

линейно и биективно, и наоборот, образы $v_i = f(e_i)$ стандартных базисных векторов[1] $e_i \in \Bbbk^n$ при любом линейном изоморфизме $f : \Bbbk^n \xrightarrow{\sim} V$ составят базис пространства V, причём отображение f действует в этом случае в точности по формуле (6-11).

Упражнение 6.19. Покажите, что векторное пространство бесконечномерно если и только если в нём есть линейно независимый набор из сколь угодно большого числа векторов.

Теорема 6.2 (равномощность базисов)
В каждом векторном пространстве все базисы равномощны.

Доказательство. Пусть базис B строго мощнее базиса E. Так как в конечномерном пространстве это невозможно по сл. 6.3, оба базиса бесконечны. Каждый вектор $e \in E$ является линейной комбинацией конечного множества векторов $B_e \subset B$. Так как множество E бесконечно, объединение $B_E = \bigcup_{e \in E} B_e$ всех множеств B_e равномощно E.

Упражнение 6.20. Убедитесь в этом.

Тем самым, существует вектор $b \in B$, не лежащий в B_e. Линейно выражая b через векторы базиса E, а каждый из входящих в это выражение векторов $e \in E$ — через векторы из B_e, мы получим линейное выражение вектора $b \in B \smallsetminus B_E$ через векторы из B_E. Тем самым, множество B линейно зависимо. □

Следствие 6.5
Всякое более мощное, чем базис, множество векторов линейно зависимо. □

6.4.2. Продолжение линейных отображений. Каждое линейное отображение $f : U \to W$, заданное на каком-либо подпространстве U любого векторного пространства V, может быть продолжено (многими способами) на всё пространство V, т. е. всегда существует такое линейное отображение $g : V \to W$, что $g|_U = f$. Чтобы построить его, выберем произвольный базис B в U, дополним его до базиса $E = B \sqcup C$ в V и рассмотрим любое отображение множеств $g : E \to W$, такое что $g(b) = f(b)$ для всех $b \in B$.

Упражнение 6.21. Убедитесь, что отображение $g : V \to W$, переводящее вектор $v = \sum_{e \in E} x_e e$ в вектор $g(e) = \sum_{e \in E} x_e g(e) \in W$ линейно и совпадает с f на любом векторе $v \in U$.

[1] См. формулу (6-10) на стр. 88.

6.4.3. Размерности конечномерных подпространств и фактор пространств. В этом разделе собраны стандартные факты о размерностях, которые будут повсеместно использоваться в дальнейшем.

Предложение 6.5
Для любых конечномерных подпространств U_1, U_2 в произвольном[1] векторном пространстве V выполняется равенство $\dim(U_1) + \dim(U_2) = \dim(U_1 \cap U_2) + \dim(U_1 + U_2)$.

Доказательство. Выберем какой-нибудь базис $u_1, \dots, u_k$ в $U_1 \cap U_2$ и дополним его векторами $v_1, \dots, v_r$ и $w_1, \dots, w_s$ до базисов в подпространствах U_1 и U_2 соответственно. Достаточно показать, что векторы $u_1, \dots, u_k,\ v_1, \dots, v_r,\ w_1, \dots, w_s$ образуют базис пространства $U_1 + U_2$. Ясно, что они его порождают. Допустим, что они линейно зависимы. Поскольку каждый из наборов $u_1, \dots, u_k, v_1, \dots, v_r$ и $u_1, \dots, u_k, w_1, \dots, w_s$ в отдельности линейно независим, в равенстве

$$x_1u_1 + \dots + x_ku_k + y_1v_1 + \dots + y_rv_r + z_1w_1 + \dots + z_sw_s = 0$$

имеются как векторы v_i, так и векторы w_j. Перенося $w_1, \dots, w_s$ в правую часть, получаем равенство между вектором из U_1 и вектором из U_2, означающее, что этот вектор лежит в пересечении $U_1 \cap U_2$. Но тогда в его разложении по базисам пространств U_1 и U_2 нет векторов v_i и w_j — противоречие. □

Следствие 6.6
Для любых подпространств U_1, U_2 конечномерного векторного пространства V

$$\dim(U_1 \cap U_2) \geqslant \dim(U_1) + \dim(U_2) - \dim(V)\,.$$

В частности, $U_1 \cap U_2 \neq 0$ при $\dim(U_1) + \dim(U_2) > \dim V$.

Доказательство. Это вытекает из предл. 6.5 и неравенства $\dim(U_1 + U_2) \leqslant \dim V$. □

Следствие 6.7 (дополнительные подпространства)
Следующие два свойства векторных подпространств U_1, U_2 в конечномерном векторном пространстве V эквивалентны[2]: (1) $V = U_1 \oplus U_2$ (2) $U_1 \cap U_2 = 0$ и $\dim(U_1) + \dim(U_2) = \dim(V)$.

Доказательство. При $U_1 \cap U_2 = 0$ равенство $\dim(U_1) + \dim(U_2) = \dim(V)$ равносильно равенству $\dim(U_1 + U_2) = \dim V$, означающему, что $U_1 + U_2 = V$. □

Предложение 6.6
Если V конечномерно, то для любого линейного отображения $f : V \to W$

$$\dim \ker f + \dim \operatorname{im} f = \dim V\,. \tag{6-12}$$

Доказательство. Выберем базис $u_1, \dots, u_k \in \ker f$, дополним его векторами $e_1, \dots, e_m$ до базиса в V и покажем, что векторы $f(e_1), \dots, f(e_m)$ образуют базис в $\operatorname{im} f$. Они порождают образ, так как для любого вектора $v = \sum y_iu_i + \sum x_je_j \in V$

$$f(v) = \sum y_i f(u_i) + \sum x_j f(e_j) = \sum x_j f(e_j)\,.$$

[1] Не обязательно конечномерном.
[2] Обладающие этими свойствами подпространства U_1, U_2 называются *дополнительными*.

Они линейно независимы, поскольку равенство $0 = \sum x_i f(e_i) = f(\sum x_i e_i)$ означает, что вектор $\sum x_i e_i$ лежит в $\ker f$, т. е. является линейной комбинацией векторов u_i, что возможно только когда все $x_i = 0$. □

Следствие 6.8
В конечномерном пространстве V для любого подпространства $U \subset V$ выполняется равенство $\dim U + \dim V/U = \dim V$, и если некоторый базис $u_1, \dots, u_n$ подпространства U дополняется до базиса в V векторами $w_1, \dots, w_m$, то их классы $[w_1]_U, \dots, [w_m]_U$ образуют базис в V/U.

Доказательство. Применяем предл. 6.6 и его доказательство к эпиморфизму $V \twoheadrightarrow V/U$. □

Следствие 6.9
Следующие свойства линейного отображения $F : V \to V$ из пространства V в себя эквивалентны друг другу: (1) F изоморфизм (2) $\ker F = 0$ (3) $\operatorname{im} F = V$.

Доказательство. Свойства (2) и (3) равносильны друг другу по предл. 6.6, а их одновременное выполнение равносильно (1), ибо свойство (2) эквивалентно инъективности f. □

Пример 6.13 (интерполяция с кратными узлами)
Зафиксируем несколько различных чисел $a_1, \dots, a_n \in \Bbbk$ и произвольно зададим для каждого числа a_i несколько значений b_{i0}, b_{i1}, … , $b_{im_i} \in \Bbbk$. Пусть общее число заданных значений $(m_1 + 1) + \cdots + (m_n + 1) = m + 1$. Покажем, что существует единственный такой многочлен $g \in \Bbbk[x]$ степени не выше m, что при каждом i сам этот многочлен и первые его m_i производных принимают в точке a_i заданные $m_i + 1$ значений $g(a_i) = b_{i0}$, $g'(a_i) = b_{i1}$, … , $g^{(m_i)}(a_i) = b_{im_i}$, где $g^{(k)}(x) = d^k g(x)/dx^k$ означает k-ю производную от многочлена g. Для этого произвольным образом занумеруем $m + 1$ пар чисел (i, j) с $1 \leqslant i \leqslant n$, $0 \leqslant j \leqslant m_j$ и выпишем их в одну строчку в порядке возрастания номеров. Рассмотрим отображение $F : \Bbbk[x]_{\leqslant m} \to \Bbbk^{m+1}$, переводящее каждый многочлен g степени $\deg g \leqslant m$ в набор значений[1] $g^{(j)}(a_i)$, записанных в строчку согласно зафиксированному только что порядку на множестве индексов (i, j).

Упражнение 6.22. Убедитесь, что отображение F линейно и $\ker F = 0$.

Так как $\dim \operatorname{im} F = \dim \Bbbk[x]_{\leqslant m} = \dim \Bbbk^{m+1}$, мы заключаем, что отображение F биективно, что и требовалось.

6.5. Свободные модули. Свободные модули над произвольным коммутативным кольцом K с единицей имеют много общего с векторными пространствами. Свободный K-модуль F с базисом E является прямой суммой свободных модулей $Ke \simeq K$, порождённых базисными векторами $e \in E$. Каждое K-линейное отображение $f : F \to M$ такого модуля в произвольный K-модуль M однозначно восстанавливается по набору своих значений $w_e = f(e)$ на базисных векторах $e \in E$ и действует на произвольный вектор по правилу[2]

$$f : \sum_{e \in E} x_e e \mapsto \sum_{e \in E} x_e w_e , \tag{6-13}$$

причём формула (6-13) задаёт линейное отображение $f : F \to M$ при любом выборе векторов $w_e \in M$, и это отображение f переводит каждый базисный вектор e в соответствующий вектор w_e. Мы получаем следующий результат, являющийся частным случаем предл. 6.4 на стр. 85.

[1] Где для единообразия обозначений мы полагаем $g^{(0)} \stackrel{\text{def}}{=} g$.

[2] Напомню, что обе суммы в (6-13) имеют лишь конечное число ненулевых коэффициентов x_e.

Теорема 6.3
Для свободного K-модуля F с базисом E и любого K-модуля M сопоставление K-линейному отображению $f : F \to M$ его ограничения на подмножество $E \subset F$ задаёт K-линейный изоморфизм между модулем $\mathrm{Hom}_K(F, M)$ всех K-линейных отображений $F \to M$ и модулем M^E всех отображений[1] множества E в множество M. □

Теорема 6.4
Все базисы свободного модуля M над произвольным коммутативным кольцом K с единицей равномощны.

Доказательство. Рассмотрим произвольный максимальный идеал $\mathfrak{m} \subset K$. Как мы видели в прим. 6.5 на стр. 82, фактор модуль $M/\mathfrak{m}M$ любого K-модуля M по подмодулю $\mathfrak{m}M$, состоящему из всевозможных конечных линейных комбинаций векторв из M с коэффициентами их $\mathfrak{m}$, является векторным пространством над полем $\Bbbk = K/\mathfrak{m}$. Свободный K-модуль F с базисом E является прямой суммой свободных модулей $Ke \simeq K$, порождённых базисными векторами $e \in E$, а его подмодуль $\mathfrak{m}F \subset F$ — прямой суммой их подмодулей $\mathfrak{m}e \subset Ke$. Поэтому[2] фактор $F/\mathfrak{m}F$ является прямой суммой одномерных векторных пространств $(K/\mathfrak{m})[e]$, порождённых классами векторов $e \in E$. Таким образом, мощность множества E совпадает с мощностью базиса векторного пространства $F/\mathfrak{m}F$. Поскольку все базисы векторного пространства равномощны, все базисы в F тоже равномощны. □

Определение 6.2
Свободный модуль F с конечным базисом называется *модулем конечного ранга*, а число элементов в базисе называется *рангом* свободного модуля F и обозначается $\mathrm{rk}\, F$.

Теорема 6.5
Всякий ненулевой подмодуль N свободного модуля M конечного ранга над произвольным кольцом главных идеалов K тоже свободен, и $\mathrm{rk}\, N \leqslant \mathrm{rk}\, M$.

Доказательство. Индукция по $m = \mathrm{rk}\, M$. При $m = 1$ модуль $M \simeq K$ и любой подмодуль $N \subset K$ представляет собою главный идеал $(d) \subset K$, который является свободным K-модулем ранга 1 с базисом d, как мы видели в прим. 6.10 на стр. 87. Пусть теперь $m > 1$. Зафиксируем в M базис $e_1, \dots, e_m$ и будем записывать векторы из M строчками их координат в этом базисе. Первые координаты всевозможных векторов $v \in N$ образуют идеал $(d) \subset K$. Если $d = 0$, подмодуль N содержится в свободном модуле ранга $m-1$ с базисом $e_2, \dots, e_m$. По индукции, такой модуль N свободен и $\mathrm{rk}\, N \leqslant (m-1)$. Если $d \neq 0$, обозначим через $v_1 \in N$ какой-нибудь вектор с первой координатой d. Тогда $N = Kv_1 \oplus N'$, где $N' \subset N$ — подмодуль, состоящий из векторов с нулевой первой координатой. Действительно, $Kv_1 \cap N' = 0$, и любой вектор $v \in N$ представляется в виде $xv_1 + w$, где $x = x_1(v)/d \in K$, а $w = v - xv_1 \in N'$. Модуль Kv_1, порождённый вектором v_1, свободен ранга 1, поскольку в объемлющем свободном модуле M нет кручения. Модуль N' содержится в свободном модуле ранга $m-1$ с базисом $e_2, \dots, e_m$. По индукции N' свободен и $\mathrm{rk}\, N' \leqslant (m-1)$. Поэтому $N = Kv_1 \oplus N'$ тоже свободен и $\mathrm{rk}\, N = 1 + \mathrm{rk}\, N' \leqslant m$. □

[1] См. n° 6.2.1 на стр. 84.
[2] См. упр. 6.7 на стр. 82.

§7. Матрицы

Всюду в этом параграфе K по умолчанию обозначает коммутативное кольцо с единицей, а $\Bbbk$ — произвольное поле.

7.1. Матричный формализм. Таблица из m строк и n столбцов, заполненная элементами множества $\mathcal{A}$, называется $m \times n$ *матрицей с элементами из* $\mathcal{A}$. Множество всех таких матриц обозначается $\mathrm{Mat}_{m\times n}(\mathcal{A})$. Элемент матрицы A, расположенный в i-й строке и j-м столбце, обозначается a_{ij}. Запись $A = (a_{ij})$ означает, что матрица A состоит из таких элементов a_{ij}. Например, матрица $A \in \mathrm{Mat}_{3\times 4}(\mathbb{Z})$ с элементами $a_{ij} = i - j$ имеет вид

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

Если множество $\mathcal{A}$ является модулем над коммутативным кольцом K, то множество матриц $\mathrm{Mat}_{m\times n}(\mathcal{A})$ тоже является K-модулем относительно операций поэлементного сложения таблиц и умножения их на числа: сумма $S = (s_{ij})$ матриц $A = (a_{ij})$ и $B = (b_{ij})$ имеет $s_{ij} = a_{ij} + b_{ij}$, а матрица $P = \lambda A$, где $\lambda \in K$, имеет $p_{ij} = \lambda a_{ij}$. Таким образом, K-модуль $\mathrm{Mat}_{m\times n}(\mathcal{A}) \simeq \mathcal{A}^{mn}$ представляет собою прямую сумму mn экземпляров модуля $\mathcal{A}$, слагаемые которой выписаны не в строку, а в таблицу размера $m \times n$.

7.1.1. Умножение матриц. Пусть элементы K-модулей $\mathcal{A}$ и $\mathcal{B}$ можно билинейно перемножать со значениями в K-модуле $\mathcal{C}$, т. е. имеется отображение $\mathcal{A} \times \mathcal{B} \to \mathcal{C}$, $(a, b) \to ab$, линейное по каждому аргументу в том смысле, что

$$(x_1a_1 + x_2a_2)(y_1b_1 + y_2b_2) = x_1y_1a_1b_1 + x_1y_2a_1b_2 + x_2y_1a_2b_1 + x_2y_2a_2b_2$$

для всех $a_1, a_2 \in \mathcal{A}$, $b_1, b_2 \in \mathcal{B}$, $x_1, x_2, y_1, y_2 \in K$. Тогда для всех $m, s, n \in \mathbb{N}$ имеется *умножение матриц* $\mathrm{Mat}_{m\times s}(\mathcal{A})\times\mathrm{Mat}_{s\times n}(\mathcal{B}) \to \mathrm{Mat}_{m\times n}(\mathcal{C})$, $(A, B) \mapsto AB$. Иными словами, произведение двух матриц определено, когда ширина левой матрицы равна с высоте правой, при этом матрица-произведение имеет столько же строк, сколько в левом сомножителе, и столько же столбцов, сколько в правом. При $m = n = 1$ результатом умножения строки ширины s на столбец высоты s является матрица размера 1×1, т. е. один элемент. Он определяется так:

$$(a_1, \dots, a_s)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{pmatrix} \stackrel{\text{def}}{=} a_1b_1 + \cdots + a_sb_s = \sum_{k=1}^{s} a_kb_k\,. \tag{7-1}$$

Для произвольных m и n элемент c_{ij} матрицы $C = AB$ равен произведению i-й строки из A на j-й столбец из B, посчитанному по формуле (7-1):

$$c_{ij} = \begin{pmatrix} a_{i1}, & a_{i2}, & \dots, & a_{is} \end{pmatrix} \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix} = \sum_{k=1}^{s} a_{ik}b_{kj}\,. \tag{7-2}$$

Иначе можно сказать, что в j-том столбце матрицы AB стоит линейная комбинация s столбцов матрицы A с коэффициентами из j-го столбца матрицы B. Это описание получается, если подставить в формулу (7-1) в качестве элементов b_i — числа из j-го столбца матрицы B, а в качестве

элементов a_j — столбцы матрицы A, интерпретируемые как векторы координатного модуля K^m с координатами, выписанными в столбик.

Упражнение 7.1. Удостоверьтесь, что это описание согласуется с формулой (7-2).

Например, для того, чтобы превратить матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \tag{7-3}$$

в матрицу из четырёх столбцов, равных, соответственно, сумме 1-го столбца матрицы A со 2-м, умноженным на λ, сумме 1-го и 3-го столбцов матрицы A, сумме 3-го столбца матрицы A со 2-м, умноженным на μ, и сумме всех трёх столбцов матрицы A, умноженных на их номера, надо умножить матрицу A справа на матрицу

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ \lambda & 0 & \mu & 2 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

Упражнение 7.2. Проверьте это прямым вычислением по формуле (7-2).

Симметричным образом, если в формуле (7-1) взять в в качестве элементов a_j числа из i-й строки матрицы A, а в качестве b_i — строки матрицы B, являющиеся векторами координатного модуля K^n с координатами, выписанными в строчку, то можно сказать, что i-й строкой матрицы AB является линейная комбинация строк матрицы B с коэффициентами, стоящими в i-й строке матрицы A. Например, если в той же матрице (7-3) хочется поставить вторую строку на место первой, а вместо второй написать её сумму с первой строкой, умноженной на λ, то это достигается умножением слева на матрицу

$$\begin{pmatrix} 0 & 1 \\ \lambda & 1 \end{pmatrix}$$

Упражнение 7.3. Проверьте это прямым вычислением по формуле (7-2).

Обратите внимание, что предыдущие два описания произведения AB получаются друг из друга заменой слова «столбец» на слово «строка» и наоборот с одновременной перестановкой букв A и B местами. Матрица $C^t = (c^t_{ij})$ размера $n \times m$, по строкам которой записаны столбцы $m \times n$ матрицы $C = (c_{ij})$, называется *транспонированной* к матрице C. Её элементы $c^t_{ij} = c_{ji}$ получаются отражением элементов матрицы C относительно биссектрисы левого верхнего угла матрицы.

Упражнение 7.4. Убедитесь, что для матриц с элементами из любого коммутативного кольца K выполняется равенство $(AB)^t = B^t A^t$, т. е. транспонирование обращает порядок сомножителей в произведениях матриц, элементы которых коммутируют друг с другом.

Упражнение 7.5. Убедитесь, что при наличии билинейного умножения K-модулей $\mathcal{A} \times \mathcal{B} \to \mathcal{C}$ произведение матриц $\mathrm{Mat}_{m\times s}(\mathcal{A}) \times \mathrm{Mat}_{s\times n}(\mathcal{B}) \to \mathrm{Mat}_{m\times n}(\mathcal{C})$ тоже билинейно, т. е.

$$(\lambda_1 A_1 + \lambda_2 A_2)(\mu_1 B_1 + \mu_2 B_2) = \lambda_1\mu_1 A_1 B_1 + \lambda_1\mu_2 A_1 B_2 + \lambda_2\mu_1 A_2 B_1 + \lambda_2\mu_2 A_2 B_2$$

для любых $A_1, A_2 \in \mathrm{Mat}_{m\times s}(\mathcal{A})$, $B_1, B_2 \in \mathrm{Mat}_{m\times s}(\mathcal{B})$, $\lambda_1, \lambda_2, \mu_1, \mu_2 \in K$.

Упражнение 7.6. Пусть заданы билинейные умножения K-модулей $\mathcal{A}\times\mathcal{B}\to\mathcal{P}$, $\mathcal{B}\times\mathcal{C}\to\mathcal{Q}$, $\mathcal{P}\times\mathcal{C}\to\mathcal{R}$ и $\mathcal{A}\times\mathcal{Q}\to\mathcal{R}$. Убедитесь, что если они ассоциативны, т. е. $(ab)c=a(bc)$ для всех $a\in\mathcal{A}$, $b\in\mathcal{B}$, $c\in\mathcal{C}$, то и произведение матриц

$$\mathrm{Mat}_{m\times s}(\mathcal{A})\times\mathrm{Mat}_{s\times t}(\mathcal{B})\times\mathrm{Mat}_{t\times n}(\mathcal{C})\to\mathrm{Mat}_{m\times n}(\mathcal{R})$$

ассоциативно, т. е. $(AB)C=A(BC)$ для всех $A\in\mathrm{Mat}_{m\times s}(\mathcal{A})$, $B\in\mathrm{Mat}_{s\times t}(\mathcal{B})$, $C\in\mathrm{Mat}_{t\times n}(\mathcal{R})$.

Предостережение 7.1. Умножение матриц не коммутативно. Например, в $\mathrm{Mat}_{2x2}(\mathbb{Z})$

$$\begin{pmatrix}1&2\\0&3\end{pmatrix}\cdot\begin{pmatrix}3&0\\4&5\end{pmatrix}=\begin{pmatrix}7&10\\12&15\end{pmatrix}$$
$$\begin{pmatrix}3&0\\4&5\end{pmatrix}\cdot\begin{pmatrix}1&2\\0&3\end{pmatrix}=\begin{pmatrix}3&6\\4&23\end{pmatrix}.$$

7.1.2. Матрицы перехода. Пусть в некотором K-модуле M заданы два набора векторов:

$$\boldsymbol{u}=(u_1,\dots,u_n)\quad\text{и}\quad\boldsymbol{w}=(w_1,\dots,w_m),$$

причём первый из них содержится в линейной оболочке второго, т. е. каждый вектор u_j имеет вид $u_j=w_1\cdot c_{1j}+w_2\cdot c_{2j}+\dots+w_m\cdot c_{mj}$, где $c_{ij}\in K$. Эти n равенств удобно собираются в одну матричную формулу $\boldsymbol{u}=\boldsymbol{w}\cdot C_{\boldsymbol{wu}}$, где $\boldsymbol{u}=(u_1,\dots,u_n)$ и $\boldsymbol{w}=(w_1,\dots,w_m)$ суть матрицы-строки с элементами из M, а матрица $C_{\boldsymbol{wu}}=\big(c_{ij}\big)$ получается подстановкой в матрицу $\boldsymbol{u}$ вместо каждого из векторов u_j столбца коэффициентов его линейного выражения через векторы w_i. Матрица $C_{\boldsymbol{wu}}$ называется *матрицей перехода* от векторов $\boldsymbol{u}$ к векторам $\boldsymbol{w}$. Название объясняется тем, что если имеется набор векторов $\boldsymbol{v}=(v_1,\dots,v_k)$, линейно выражающихся через векторы $\boldsymbol{u}$ по формулам $\boldsymbol{v}=\boldsymbol{u}C_{\boldsymbol{uv}}$, то выражение векторов $\boldsymbol{v}$ через векторы $\boldsymbol{w}$ задаётся матрицей

$$C_{\boldsymbol{wv}}=C_{\boldsymbol{wu}}C_{\boldsymbol{uv}}\,,\tag{7-4}$$

которая возникает при подстановке $\boldsymbol{u}=\boldsymbol{w}C_{\boldsymbol{wu}}$ в разложение $\boldsymbol{v}=\boldsymbol{u}C_{\boldsymbol{uv}}$. Таким образом, если записывать линейные выражения $v=u_1x_1+\dots+u_nx_n=w_1y_1+\dots+w_my_m$ произвольного вектора $v\in\mathrm{span}(u_1,\dots,u_n)$ через векторы $\boldsymbol{u}$ и $\boldsymbol{w}$ в виде $v=\boldsymbol{u}x=\boldsymbol{w}y$, где $x=(x_1,\dots,x_n)^t$ и $y=(y_1,\dots,y_m)^t$ суть столбцы коэффициентов, то эти столбцы будут связаны соотношением

$$y=C_{\boldsymbol{wu}}x\,.$$

Отметим, что когда набор векторов $\boldsymbol{w}=(w_1,\dots,w_m)$ линейно зависим, у каждого вектора v из их линейной оболочки имеется много *разных* линейных выражений через векторы w_j. Поэтому обозначение $C_{\boldsymbol{wv}}$ в этой ситуации не корректно в том смысле, что элементы матрицы $C_{\boldsymbol{wv}}$ определяются наборами векторов $\boldsymbol{w}$ и $\boldsymbol{v}$ не однозначно. Тем не менее, равенство (7-4) вполне осмысленно и означает, что имея какие-нибудь линейные выражения $C_{\boldsymbol{wu}}$ и $C_{\boldsymbol{uv}}$ векторов $\boldsymbol{u}$ через $\boldsymbol{w}$ и векторов $\boldsymbol{v}$ через $\boldsymbol{u}$, мы можем явно предъявить одно из линейных выражений $C_{\boldsymbol{wv}}$ векторов $\boldsymbol{v}$ через векторы $\boldsymbol{w}$, перемножив матрицы $C_{\boldsymbol{wu}}$ и $C_{\boldsymbol{uv}}$.

Если же набор векторов $\boldsymbol{e}=(e_1,\dots,e_n)$ является базисом, то матрица перехода $C_{\boldsymbol{ew}}$, выражающая произвольный набор векторов $\boldsymbol{w}=(w_1,\dots,w_m)$ через базис $\boldsymbol{e}$ однозначно определяется по наборам $\boldsymbol{e}$ и $\boldsymbol{w}$, т. е. два набора векторов $\boldsymbol{u}$, $\boldsymbol{w}$ совпадают если и только если выполняется равенство $C_{\boldsymbol{eu}}=C_{\boldsymbol{ew}}$.

7.1.3. Обратимые матрицы. В этом разделе мы рассматриваем квадратные $n \times n$ матрицы с элементами из коммутативного кольца K с единицей. Матрица

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_{n\times n}(K)\,,$$

по диагонали которой стоят единицы, а в остальных местах — нули, называется *единичной*.

Упражнение 7.7. Убедитесь, что $AE = A$ и $EA = A$ всякий раз, когда такие произведения определены.

Матрица $C \in \mathrm{Mat}_{n\times n}(K)$ называется *обратимой*, если существуют такие матрицы A и B, что $AC = E = CB$. В этом случае матрицы A и B автоматически равны друг другу, так как

$$A = AE = A(CB) = (AC)B = EB = B\,.$$

Это вычисление заодно показывает, что матрица $C^{-1} \stackrel{\text{def}}{=} A = B$ однозначно определяется по C свойством $C^{-1}C = CC^{-1} = E$. Матрица C^{-1}, если существует, называется *обратной* к C.

Упражнение 7.8. Докажите, что обратимость матрицы $C \in \mathrm{Mat}_{n\times n}(K)$ равносильна обратимости транспонированной к ней матрицы[1] C^t.

Предложение 7.1
Набор векторов $\boldsymbol{u} = (u_1, \dots, u_n)$ свободного K-модуля с базисом $\boldsymbol{e} = (e_1, \dots, e_n)$ является базисом если и только если матрица перехода[2] $C_{\boldsymbol{eu}} \in \mathrm{Mat}_{n\times n}(K)$ обратима, и тогда $C_{\boldsymbol{eu}}^{-1} = C_{\boldsymbol{ue}}$.

Доказательство. Пусть векторы $\boldsymbol{u}$ образуют базис. Так как каждый набор векторов имеет единственное выражение через базис, $C_{\boldsymbol{ue}}C_{\boldsymbol{eu}} = C_{\boldsymbol{uu}} = E$ и $C_{\boldsymbol{eu}}C_{\boldsymbol{ue}} = C_{\boldsymbol{ee}} = E$ по формуле (7-4). Тем самым, $C_{\boldsymbol{ue}} = C_{\boldsymbol{eu}}^{-1}$. Наоборот, если матрица $C_{\boldsymbol{eu}}$ обратима, то умножая обе части равенства $\boldsymbol{u} = \boldsymbol{e}C_{\boldsymbol{eu}}$ справа на $C_{\boldsymbol{eu}}^{-1}$, мы получаем линейное выражение $\boldsymbol{e} = \boldsymbol{u}C_{\boldsymbol{eu}}^{-1}$ базиса $\boldsymbol{e}$ через векторы $\boldsymbol{u}$ и заключаем, что последние линейно порождают весь модуль. Если существует линейное соотношение $\boldsymbol{u}x = 0$, где $x \in K^n$ — столбец коэффициентов, то $\boldsymbol{e}C_{\boldsymbol{eu}}x = 0$, откуда $C_{\boldsymbol{eu}}x = 0$. Умножая обе части слева на $C_{\boldsymbol{eu}}^{-1}$, получаем $x = 0$. Тем самым, векторы $\boldsymbol{u}$ линейно независимы и образуют базис по лем. 6.1 на стр. 86. □

Пример 7.1 (замена координат при смене базиса)
Пусть набор векторов $\boldsymbol{w} = (w_1, \dots, w_m)$ выражается через базис $\boldsymbol{e} = (e_1, \dots, e_n)$ как $\boldsymbol{w} = \boldsymbol{e}C_{\boldsymbol{ew}}$. Если $\boldsymbol{u} = \boldsymbol{e}C_{\boldsymbol{eu}}$ — другой базис, то выражение векторов $\boldsymbol{w}$ через базис $\boldsymbol{u}$ имеет вид $\boldsymbol{w} = \boldsymbol{e}C_{\boldsymbol{ew}} = \boldsymbol{u}C_{\boldsymbol{eu}}^{-1}C_{\boldsymbol{ew}}$, т. е. $C_{\boldsymbol{uw}} = C_{\boldsymbol{eu}}^{-1}C_{\boldsymbol{ew}}$. В частности, если вектор $v = \boldsymbol{e}x$ имеет в базисе $\boldsymbol{e}$ столбец координат x, то в базисе $\boldsymbol{u} = \boldsymbol{e}C_{\boldsymbol{eu}}$ он имеет столбец координат $y = C_{\boldsymbol{eu}}^{-1}x$

Следствие 7.1
Следующие условия на квадратную матрицу $A \in \mathrm{Mat}_{n\times n}(\Bbbk)$ с элементами из поля $\Bbbk$ эквивалентны:

1) матрица A обратима

[1] См. упр. 7.4 на стр. 95.
[2] См. n° 7.1.2 на стр. 96.

2) столбцы матрицы A линейно независимы

3) столбцы матрицы A линейно порождают координатное пространство $\Bbbk^n$,

и то же самое верно с заменой столбцов на строки.

Доказательство. Обозначим через $a_1, \dots, a_n$ столбцы матрицы A, воспринимаемые как векторы координатного пространства $\Bbbk^n$. Матрица A является матрицей перехода от этих векторов к стандартному базису пространства $\Bbbk^n$. По предл. 7.1 обратимость матрицы A равносильна тому, что векторы a_i образуют в $\Bbbk^n$ базис, что по сл. 6.3 на стр. 89 равносильно каждому из условий (2), (3). Самое последнее утверждение вытекает из упр. 7.8 на стр. 97. □

Пример 7.2 (обратимые 2×2-матрицы над коммутативным кольцом)
Возводя матрицу

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(K)$$

в квадрат, получим

$$C^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{pmatrix} = \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & cb+d^2 \end{pmatrix},$$

откуда

$$(a+d)\cdot C - C^2 = \begin{pmatrix} (ad-bc) & 0 \\ 0 & (ad-bc) \end{pmatrix} = (ad-bc)\cdot E\,. \tag{7-5}$$

Число $\det C \stackrel{\text{def}}{=} ad-bc$ называется *определителем* 2×2-матрицы C.

Упражнение 7.9. Докажите для любых $A, B \in \mathrm{Mat}_{2x2}(K)$ равенство $\det(AB) = \det(A)\cdot\det(B)$.

Из упражнения вытекает, что определитель любой обратимой матрицы обратим в K, поскольку вычисляя определитель обеих частей матричного равенства $C\cdot C^{-1} = E$, получаем

$$\det(C)\cdot\det(C^{-1}) = \det E = 1\,.$$

С другой стороны, если $\det C = ad-bc$ обратим в K, то равенство (7-5) переписывается как

$$C\cdot\big((a+d)E - C\big)\cdot(ad-bc)^{-1} = E\,.$$

Тем самым, матрица C обратима и

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det C}\big((a+d)E - C\big) = (ad-bc)^{-1}\begin{pmatrix} a & -b \\ -c & d \end{pmatrix}. \tag{7-6}$$

Итак, 2×2 матрица обратима если и только если обратим её определитель.

7.1.4. Матрицы линейных отображений. Пусть K-модули N и M линейно порождаются наборами векторов $\boldsymbol{u} = (u_1, \dots, u_n)$ и $\boldsymbol{w} = (w_1, \dots, w_m)$ соответственно. Всякое K-линейное отображение $F: N \to M$ однозначно задаётся набором своих значений

$$F(\boldsymbol{u}) \stackrel{\text{def}}{=} \big(F(u_1), \dots, F(u_n)\big)\,, \tag{7-7}$$

на порождающих векторах и действует на произвольный вектор $v = \boldsymbol{u}x$, где $x \in K^n$ — столбец коэффициентов линейного выражения вектора v через образующие $\boldsymbol{u}$, по правилу

$$F(\boldsymbol{u}x) = F\Big(\sum_{i=1}^{n} u_i x_i\Big) = \sum_{i=1}^{n} F(u_i)\cdot x_i = F(\boldsymbol{u})x\,. \tag{7-8}$$

Матрица перехода от векторов (7-7) к образующим $\boldsymbol{w}$ модуля M обозначается

$$F_{\boldsymbol{wu}} = C_{\boldsymbol{w},F(\boldsymbol{u})} \in \mathrm{Mat}_{m\times n}(K)$$

и называется *матрицей отображения* F в образующих $\boldsymbol{w}$ и $\boldsymbol{u}$. В её j-м стоят коэффициенты линейного выражения вектора $F(u_j)$ через векторы $\boldsymbol{w}$. Согласно (7-8) произвольный вектор $\boldsymbol{u}x$ со столбцом коэффициентов x переводится отображением F в вектор $\boldsymbol{w}F_{\boldsymbol{wu}}x$ со столбцом коэффициентов $F_{\boldsymbol{wu}}x$. Из (7-8) также вытекает, что для любого набора векторов $\boldsymbol{v} = (v_1,\dots,v_k)$ в N, любой матрицы $A \in \mathrm{Mat}_{\ell\times k}(K)$ и любого K-линейного отображения $F\colon N \to M$ выполняется равенство $F(\boldsymbol{v}A) = F(\boldsymbol{v})A$. Если K-модуль L порождается векторами $\boldsymbol{v} = (v_1,\dots,v_\ell)$ и K-линейные отображения $F\colon N \to L$ и $G\colon L \to M$ имеют матрицы $F_{\boldsymbol{vu}}$ и $G_{\boldsymbol{wv}}$, соответственно, в образующих $\boldsymbol{v}$, $\boldsymbol{u}$ и в образующих $\boldsymbol{w}$, $\boldsymbol{v}$, то композиция $H = GF\colon N \to M$ имеет в образующих $\boldsymbol{w}$, $\boldsymbol{u}$ матрицу $H_{\boldsymbol{wu}} = G_{\boldsymbol{wv}}F_{\boldsymbol{vu}}$, ибо $H(\boldsymbol{u}) = G\big(F(\boldsymbol{u})\big) = G\big(\boldsymbol{v}F_{\boldsymbol{vu}}\big) = G(\boldsymbol{v})F_{\boldsymbol{vu}} = \boldsymbol{w}G_{\boldsymbol{wv}}F_{\boldsymbol{vu}}$.

Отметим, что когда образующие $\boldsymbol{w}$ линейно зависимы, то как и в n° 7.1.2, матрица $F_{\boldsymbol{wu}}$ линейного отображения F определяется образующими $\boldsymbol{w}$ и $\boldsymbol{u}$ не однозначно, так как набор векторов $F(\boldsymbol{u})$ имеет много разных линейных выражений через векторы $\boldsymbol{w}$. Предыдущие формулы означают при этом, что если задано какое-то выражение $v = \boldsymbol{u}x$ вектора v через образующие $\boldsymbol{u}$, то столбец коэффициентов $y = F_{\boldsymbol{wu}}x$ даёт одно из возможных линейных выражений $F(v) = \boldsymbol{w}y$ вектора $F(v)$ через образующие $\boldsymbol{w}$, и что получить одну из возможных матриц для композиции отображений можно перемножив какие-нибудь из матриц этих отображений в том же порядке, в каком берётся композиция.

Также важно понимать, что когда образующие $\boldsymbol{u}$ линейно зависимы, матрица $F_{\boldsymbol{wu}}$ не может быть произвольной: для любого линейного соотношения $\boldsymbol{u}x = 0$ между векторами $\boldsymbol{u}$ в N в модуле M должно выполняться соотношение $\boldsymbol{w}F_{\boldsymbol{wu}}x = 0$. Иными словами, если модули $M = K^n/R_M$ и $M = K^m/R_N$ заданы при помощи образующих и соотношений, как в прим. 6.12 на стр. 87, то матрица $A \in \mathrm{Mat}_{m\times n}(K)$ тогда и только тогда является матрицей некоторого линейного отображения $F\colon N \to M$, когда для любого столбца $x \in R_N$ столбец $Ax \in R_M$. Это матричная переформулировка предл. 6.3 на стр. 85 в обозначениях из прим. 6.12.

Если же модули N и M оба свободны и наборы векторов $\boldsymbol{u}$ и $\boldsymbol{w}$ являются их базисами, то сопоставление K-линейному отображению $F\colon N \to M$ его матрицы $F_{\boldsymbol{wu}}$ в этих базисах задаёт K-линейный изоморфизм $\mathrm{Hom}_K(N,M) \xrightarrow{\sim} \mathrm{Mat}_{m\times n}(K)$, $F \mapsto F_{\boldsymbol{wu}}$.

Упражнение 7.10. Убедитесь, что сопоставление отображению его матрицы линейно.

В частности, для свободных K-модулей N и M конечного ранга модуль $\mathrm{Hom}_K(N,M)$ тоже свободен и $\mathrm{rk}\,\mathrm{Hom}_K(N,M) = \mathrm{rk}\,N\cdot\mathrm{rk}\,M$.

Пример 7.3 (замена матрицы линейного отображения при смене базиса)
Если K-линейный гомоморфизм свободных модулей $F\colon N \to M$ имеет в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ матрицу $F_{\boldsymbol{wu}}$, то он переводит векторы $\boldsymbol{e} = \boldsymbol{u}C_{\boldsymbol{ue}}$ любого другого базиса $\boldsymbol{e}$ в N в векторы

$$F(\boldsymbol{e}) = F(\boldsymbol{u}C_{\boldsymbol{ue}}) = F(\boldsymbol{u})C_{\boldsymbol{ue}} = \boldsymbol{w}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}}\,.$$

Если выбрать в M другой базис $\boldsymbol{f}$, через который исходный базис $\boldsymbol{w}$ выражается по формуле $\boldsymbol{w} = \boldsymbol{f}C_{\boldsymbol{fw}}$, и подставить это выражение в предыдущую формулу вместо $\boldsymbol{w}$, мы получим, что $F(\boldsymbol{e}) = \boldsymbol{f}C_{\boldsymbol{fw}}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}}$. Таким образом, матрица $F_{\boldsymbol{fe}}$ отображения F в базисах $\boldsymbol{e}$ и $\boldsymbol{f}$ выражается через матрицу $F_{\boldsymbol{wu}}$ того же отображения в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ по формулам

$$F_{\boldsymbol{fe}} = C_{\boldsymbol{fw}}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}} = C_{\boldsymbol{wf}}^{-1}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}} = C_{\boldsymbol{fw}}F_{\boldsymbol{wu}}C_{\boldsymbol{eu}}^{-1}\,. \tag{7-9}$$

Пример 7.4 (матрица линейного эндоморфизма)
Линейный эндоморфизм $F\colon M \to M$ модуля M, порождённого векторами $\boldsymbol{w} = (w_1, \ldots, w_m)$, обычно принято записывать квадратной матрицей $F_{\boldsymbol{w}} \stackrel{\text{def}}{=} F_{\boldsymbol{ww}}$, в j-м столбце которой стоят коэффициенты линейного выражения вектора $F(w_j)$ через тот же самый набор образующих $\boldsymbol{w}$. Эта матрица называется *матрицей эндоморфизма* F в образующих $\boldsymbol{w}$. Если векторы $\boldsymbol{w}$ составляют базис модуля M, то при переходе к другому базису $\boldsymbol{e} = \boldsymbol{u}C_{\boldsymbol{ue}}$ матрица эндоморфизма поменяется по формулам (7-9):

$$F_{\boldsymbol{e}} = C_{\boldsymbol{ew}}F_{\boldsymbol{w}}C_{\boldsymbol{we}} = C_{\boldsymbol{we}}^{-1}F_{\boldsymbol{w}}C_{\boldsymbol{we}} = C_{\boldsymbol{ew}}F_{\boldsymbol{w}}C_{\boldsymbol{ew}}^{-1}\,. \tag{7-10}$$

7.1.5. Ранг матрицы над полем. В этом разделе мы рассматриваем матрицы с элементами из произвольного поля $\Bbbk$. Размерность линейной оболочки столбцов матрицы $A \in \mathrm{Mat}_{m\times n}(\Bbbk)$ в координатном векторном пространстве $\Bbbk^m$ называется *рангом* матрицы A и обозначается $\operatorname{rk} A$. Каждая матрица A задаёт линейное отображение $F_A\colon \Bbbk^n \to \Bbbk^m$, $x \mapsto Ax$, которое переводит координатный столбец $x \in \Bbbk^n$ в координатный столбец $F_A(x) = Ax \in \Bbbk^m$. В стандартных базисах[1] $\boldsymbol{e}$ и $\boldsymbol{f}$ координатных пространств $\Bbbk^n$ и $\Bbbk^m$ матрица $F_{\boldsymbol{fe}}$ оператора F_A совпадает матрицей A. Поэтому линейная оболочка столбцов матрицы A представляет собою образ оператора F_A. Тем самым, $\operatorname{rk} A = \dim \operatorname{im} F_A$.

Лемма 7.1
Ранг матрицы не меняется при умножении на обратимые матрицы слева или справа.

Доказательство. Если матрицы $D \in \mathrm{Mat}_{n\times n}(\Bbbk)$ и $C \in \mathrm{Mat}_{m\times m}(\Bbbk)$ обратимы, то матрица CAD является матрицей описанного выше оператора $F_A\colon \Bbbk^n \to \Bbbk^m$, $x \mapsto Ax$, но не в стандартных базисах $\boldsymbol{e}$ и $\boldsymbol{f}$ координатных пространств $\Bbbk^n$ и $\Bbbk^m$, а в новых базисах $\boldsymbol{u} = \boldsymbol{e}D$ и $\boldsymbol{w} = \boldsymbol{f}C^{-1}$. В самом деле, $F_{\boldsymbol{wu}} = C_{\boldsymbol{wf}}F_{\boldsymbol{fe}}C_{\boldsymbol{eu}} = CAD$ по формуле (7-9). Тем самым, размерность образа линейного оператора F_A равна размерности линейной оболочки столбцов матрицы CAD. □

Следствие 7.2
Размерность линейной оболочки строк произвольной матрицы A тоже не меняется при умножении матрицы A слева или справа на любые обратимые матрицы.

Доказательство. Применим лем. 7.1 к транспонированной матрице A^t. □

Теорема 7.1 (теорема о ранге матрицы)
Для любой матрицы $A \in \mathrm{Mat}_{m\times n}(\Bbbk)$ выполняется равенство $\operatorname{rk} A = \operatorname{rk} A^t$. Иными словами, линейная оболочка строк матрицы A в координатном пространстве $\Bbbk^n$ и линейная оболочка столбцов матрицы A в координатном пространстве $\Bbbk^m$ имеют равные размерности.

[1] См. прим. 6.12 на стр. 87 и в частности формулу (6-10).

Доказательство. Рассмотрим задаваемое матрицей A линейное отображение

$$F_A : \Bbbk^n \to \Bbbk^m, \quad x \mapsto Ax,$$

выберем в $\Bbbk^n$ базис $\boldsymbol{u} = (u_1, \dots, u_r, u_{r+1}, \dots, u_n)$ так, чтобы его векторы $u_{r+1}, \dots, u_n$ составили базис в $\ker F_A$. В доказательстве предл. 6.6 на стр. 91 мы видели, что векторы $w_j = F_A(u_j)$, где $1 \leqslant j \leqslant r$, образуют в этом случае базис в $\operatorname{im} F_A$, так что $r = \dim \operatorname{im} F_A = \operatorname{rk} A$. Дополним эти векторы w_j до базиса $\boldsymbol{w} = (w_1, \dots, w_m)$ всего пространства $\Bbbk^m$. Матрица $F_{\boldsymbol{wu}} = \left(f_{ij}\right)$ оператора F_A в базисах $\boldsymbol{w}$ и $\boldsymbol{u}$ пространств $\Bbbk^m$ и $\Bbbk^n$ имеет $f_{ii} = 1$ при $1 \leqslant i \leqslant r$ и нули во всех остальных местах. В частности, линейная оболочка её строк в координатном пространстве $\Bbbk^n$ и линейная оболочка её столбцов в координатном пространстве $\Bbbk^m$ имеют одну и ту же размерность r. Согласно прим. 7.3 матрица $F_{\boldsymbol{wu}} = C_{\boldsymbol{wf}} A C_{\boldsymbol{eu}}$ получается из матрицы $A = F_{\boldsymbol{fe}}$ оператора F_A в стандартных базисах $\boldsymbol{e}$ и $\boldsymbol{f}$ пространств $\Bbbk^n$ и $\Bbbk^m$ умножением слева и справа на обратимые матрицы переходов $C_{\boldsymbol{wf}}$ и $C_{\boldsymbol{eu}}$. Согласно лем. 7.1 и сл. 7.2 такое умножение не меняет размерностей линейных оболочек строк и столбцов матрицы A. □

7.2. Ассоциативные алгебры над полем. Векторное пространство A над полем $\Bbbk$ называется *алгеброй* над $\Bbbk$ или $\Bbbk$-*алгеброй*, если на нём имеется такое умножение $A \times A \to A$, что произведение ab линейно по a при фиксированном b, и линейно по b при фиксированном a, т. е.

$$(\lambda_1 a_1 + \mu_1 b_1)(\lambda_2 a_2 + \mu_2 b_2) = \lambda_1\lambda_2 a_1 a_2 + \lambda_1\mu_2 a_1 b_2 + \mu_1\lambda_2 b_1 a_2 + \mu_1\mu_2 b_1 b_2$$

для всех $a_1, a_2, b_1, b_2 \in A$ и $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \Bbbk$. Алгебра A называется *ассоциативной*, если

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

и *коммутативной* — если $ab = ba$ для всех $a, b \in A$. Ассоциативная алгебра, в которой есть нейтральный элемент по отношению к умножению, т. е. такой $e \in A$, что $ea = ae = a$ для всех $a \in A$, называется *алгеброй с единицей*.

Упражнение 7.11. Покажите, что $0 \cdot a = 0$ для всех a в любой алгебре A и что единичный элемент единственен (если существует).

Примерами *коммутативных* ассоциативных алгебр с единицами являются алгебра многочленов $\Bbbk[x_1, \dots, x_n]$ и прочие коммутативные $\Bbbk$-алгебры в смысле прим. 5.5 на стр. 71.

7.2.1. Алгебра матриц. Модельными примерами некоммутативных ассоциативных алгебр с единицами являются алгебры квадратных матриц $\operatorname{Mat}_n(\Bbbk) \stackrel{\text{def}}{=} \operatorname{Mat}_{n\times n}(\Bbbk)$ и алгебры

$$\operatorname{End}_\Bbbk(V) \stackrel{\text{def}}{=} \operatorname{Hom}_\Bbbk(V, V)$$

линейных эндоморфизмов $V \to V$ векторных пространств V над полем $\Bbbk$. Если $\dim V = n$, то каждый базис $\boldsymbol{e} = (e_1, \dots, e_n)$ задаёт линейный изоморфизм $V \xrightarrow{\sim} \Bbbk^n$, переводящий вектор $v = \boldsymbol{e}x \in V$ в столбец $x \in \Bbbk^n$ его координат в базисе $\boldsymbol{e}$, а также изоморфизм алгебр[1]

$$\operatorname{End}_\Bbbk(V) \xrightarrow{\sim} \operatorname{Mat}_n(\Bbbk), \quad F \mapsto F_{\boldsymbol{e}}, \tag{7-11}$$

переводящий линейное отображение $F : V \to V$ в его матрицу[2] $F_{\boldsymbol{e}}$ в базисе $\boldsymbol{e}$. Эти два изоморфизма согласованы в том смысле, что отображение F переводит вектор v со столбцом координат x в вектор Fv со столбцом координат $F_{\boldsymbol{e}}x$.

[1] Т. е. перестановочный с умножением в алгебре $\Bbbk$-линейный изоморфизм векторных пространств.

[2] См. n° 7.1.4 выше, в частности прим. 7.4 на стр. 100.

Стандартный базис матричной алгебры составляют матрицы E_{ij}, единственным ненулевым элементом которых является единица, стоящая в i-й строке и j-м столбце. Произвольная матрица $A = (a_{ij})$ линейно выражается через них по формуле $A = \sum_{i,j} a_{ij}E_{ij}$. Прообразами базисных матриц E_{ij} при изоморфизме (7-11) являются линейные операторы $E_{ij} : V \to V$, которые мы будем обозначать теми же буквами и которые действуют на базисные векторы e_k пространства V по правилам

$$E_{ij}(e_k) = \begin{cases} e_i & \text{при } k = j \\ 0 & \text{при } k \neq j. \end{cases}$$

Отсюда немедленно получается таблица умножения базисных элементов E_{ij}:

$$E_{ik}E_{\ell j} = \begin{cases} E_{ij} & \text{при } k = \ell \\ 0 & \text{при } k \neq \ell, \end{cases} \tag{7-12}$$

из которой лишний раз видно, что алгебра некоммутативна (скажем, $E_{12}E_{21} \neq E_{21}E_{12}$).

Упражнение 7.12. Составьте таблицу *коммутаторов* $[E_{ik}, E_{\ell j}] \stackrel{\text{def}}{=} E_{ik}E_{\ell j} - E_{\ell j}E_{ik}$.

Пример 7.5
Вычислим A^{2020} для матрицы $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Поскольку $A = E + E_{12}$ и матрицы E и E_{12} коммутируют, вычислить $(E + E_{12})^{2020}$ можно по обычной формулой бинома[1]. А так как $E_{12}^n = 0$ при $n > 1$, мы получаем

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{2020} = (E + E_{12})^{2020} = E + 2020\,E_{12} = \begin{pmatrix} 1 & 2020 \\ 0 & 1 \end{pmatrix}.$$

Упражнение 7.13. Покажите, что $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ при всех $n \in \mathbb{Z}$.

7.2.2. Обратимые элементы. Элемент a алгебры A с единицей $e \in A$ называется *обратимым*, если существует такой элемент $a^{-1} \in A$, что $aa^{-1} = a^{-1}a = e$. В ассоциативной алгебре A это требование можно ослабить до существования левого и правого обратных к a элементов $a', a'' \in A$, таких что $a'a = aa'' = e$, ибо они автоматически совпадут друг с другом[2]: $a' = a'e = a'(aa'') = (a'a)a'' = ea'' = a''$. Эта выкладка заодно показывает, что обратный к a элемент a^{-1} однозначно определяется по a равенствами $aa^{-1} = a^{-1}a = e$.

7.2.3. Алгебраические и трансцендентные элементы. Каждый ненулевой элемент a любой ассоциативной $\Bbbk$-алгебры A с единицей задаёт ненулевой гомоморфизм вычисления

$$\mathrm{ev}_a : \Bbbk[t] \to A, \quad f(x) \mapsto f(a), \tag{7-13}$$

переводящий многочлен $f(x) = f_0 + f_1x + \cdots + f_mx^m$ в элемент $f(a) = f_0e + f_1a + \cdots + f_ma^m \in A$ — результат подстановки в f значения[3] $x = a$. Если ядро $\ker \mathrm{ev}_a = 0$, элемент a называется *трансцендентным* над $\Bbbk$. В этом случае гомоморфизм (7-13) инъективен, и все целые неотрицательные степени элемента a линейно независимы над $\Bbbk$. В частности, алгебра A бесконечномерна как векторное пространство над $\Bbbk$.

[1] См. формулу (1-9) на стр. 9.
[2] Ср. с n° 7.1.3 на стр. 97.
[3] При этом мы считаем, что $f_0 = f_0x^0$ и $a^0 \stackrel{\text{def}}{=} e$.

Если ядро гомоморфизма (7-13) ненулевое, элемент a называется *алгебраическим* над $\Bbbk$. В этом случае $\ker \mathrm{ev}_a$ является ненулевым собственным главным идеалом[1] в $\Bbbk[x]$. Порождающий его многочлен со старшим коэффициентом 1 обозначается $\mu_a(x)$ и называется *минимальным многочленом* элемента a. Он однозначно характеризуется как приведённый многочлен наименьшей степени, для которого $\mu_a(a) = 0$, и делит все многочлены, аннулирующие элемент a. Если алгебра A конечномерна как векторное пространство над $\Bbbk$, то все её элементы алгебраичны над $\Bbbk$. В частности, любая квадратная матрица конечного размера и любой линейный эндоморфизм конечномерного векторного пространства удовлетворяют некоторому полиномиальному уравнению.

Пример 7.6 (аннулирующий многочлен матрицы)
Поскольку $\dim_{\Bbbk} \mathrm{Mat}_n(\Bbbk) = n^2$, матрицы A^k, где $0 \leqslant k \leqslant n^2$, линейно зависимы над $\Bbbk$ для любой матрицы $A \in \mathrm{Mat}_n(\Bbbk)$. Это означает, что каждая $n \times n$ матрица удовлетворяет нетривиальному полиномиальному уравнению степени не выше n^2. Вскоре мы увидим[2], что эта априорная оценка степени сильно завышена, и степень минимального многочлена любой $n \times n$ матрицы в действительности не превышает n. Для матриц размера 2×2 это видно из прим. 7.2 на стр. 98: полученная там формула (7-5) утверждает, что матрица

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

удовлетворяет квадратному уравнению $x^2 - (a+b)\,x + (ad - bc) = 0$.

7.2.4. Нильпотентные элементы. Элемент a алгебры A называется *нильпотентным*, если $a^n = 0$ для некоторого $n \in \mathbb{N}$. Каждый нильпотентный элемент a корректно задаёт аналогичный (7-13) гомоморфизм вычисления $\mathrm{ev}_a : \Bbbk[\![x]\!] \to A$, $f(x) \mapsto f(a)$, подставляющий элемент a вместо переменной x в формальные степенные ряды. В частности, для такого элемента a определены элементы e^a, $\ln(1+a)$ и $(1+a)^s$ с произвольным $s \in \Bbbk$, которые удовлетворяют в алгебре A всем алгебраическим соотношениям, что имеются между рядами e^x, $\ln(1+x)$ и $(1+x)^s$ в кольце $\Bbbk[\![x]\!]$. Например, элемент $b = (1+a)^{1/2} \in A$ имеет $b^2 = 1 + a$.

Упражнение 7.14. Предъявите такую рациональную 3×3 матрицу B, что $B^3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

7.3. Некоммутативные кольца. Абелева группа R с операцией умножения $R \times R \to R$ называется *кольцом*, если это умножение ассоциативно, т. е. $f(gh) = (fg)h$ для всех $f, g, h \in R$, двусторонне дистрибутивно по отношению к сложению, т. е. $f(g+h) = fg + fh$ и $(f+g)h = fh + gh$ для всех $f, g, h \in R$, и существует такой элемент $1 \in R$, что $1 \cdot f = f \cdot 1 = f$ для всех $f \in R$. Элемент 1 называется *единицей* кольца R.

Упражнение 7.15. Покажите, что $0 \cdot f = 0$ для всех $f \in R$ и что единица единственна.

Алгебры $\mathrm{Mat}_n(\Bbbk)$ и $\mathrm{End}_{\Bbbk}(V)$ являются примерами некоммутативных колец. Первый из этих примеров допускает значительное обобщение, а именно, квадратные $n \times n$ матрицы с элементами из любого кольца R тоже образуют кольцо $\mathrm{Mat}_n(R)$ относительно операций сложения и умножения матриц, определённых в самом начале этого параграфа[3]. А именно, сумма $S = F + G$ и

[1] Напомню, что $\Bbbk[x]$ является кольцом главных идеалов, см. n° 5.3 на стр. 71.
[2] См. теор. 8.2 на стр. 116.
[3] См. n° 7.1 на стр. 94.

произведение $P = FG$ матриц $F = \left(f_{ij}\right)$ и $G = \left(g_{ij}\right)$ имеют матричными элементами

$$s_{ij} = f_{ij} + g_{ij} \quad \text{и} \quad p_{ij} = \sum_{\nu} f_{i\nu} g_{\nu j}\,.$$

Упражнение 7.16. Убедитесь, что умножение в $\mathrm{Mat}_n(R)$ ассоциативно и дистрибутивно по отношению к сложению, а матрицы E_{ij}, единственным ненулевым элементом которых является единица, стоящая в i-й строке и j-м столбце, перемножаются по форм. (7-12) на стр. 102, и единичная матрица $E = \sum E_{ii}$ является единицей кольца $\mathrm{Mat}_n(R)$.

Вычисления с матрицами, элементы которых лежат в некоммутативном кольце, требуют большей осторожности, чем вычисления с матрицами, элементы которых можно переставлять друг с другом в произведениях. Например, ключевая формула (7-5) из прим. 7.2 на стр. 98 перестаёт быть верной для матриц над некоммутативным кольцом, как и полученные в прим. 7.2 критерий обратимости и формула для обратной матрицы к матрице 2×2.

Упражнение 7.17. Убедитесь в этом.

Пример 7.7 (примеры обратимых матриц 2×2)
Покажем, что матрица

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

с элементами из произвольного[1] кольца R обратима если и только если обратимы её диагональные элементы a и d. Из равенства

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax + bz & ay + bw \\ dz & dw \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

вытекает, что $dw = 1$ и $dz = 0$, откуда d обратим, а $w = d^{-1}$ и $z = 0$. Поэтому $ax = 1$, откуда a обратим, а $x = a^{-1}$. Тогда в правом верхнем углу получаем соотношение $ay + bd^{-1} = 0$, из которого $y = -a^{-1}bd^{-1}$. Таким образом,

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & -a^{-1}bd^{-1} \\ 0 & d^{-1} \end{pmatrix}$$

Аналогичные рассуждения показывают, что обратимость матрицы вида

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$$

равносильна обратимости диагональных элементов a, d, и в этом случае

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & 0 \\ -d^{-1}ca^{-1} & d^{-1} \end{pmatrix}$$

Упражнение 7.18. Покажите, что матрицы $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$ и $\begin{pmatrix} 0 & b \\ c & d \end{pmatrix}$ обратимы если и только если обратимы оба элемента c, b, и в этом случае

$$\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & c^{-1} \\ b^{-1} & -b^{-1}ac^{-1} \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 0 & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} -c^{-1}db^{-1} & c^{-1} \\ b^{-1} & 0 \end{pmatrix}$$

[1] В том числе некоммутативного.

Пример 7.8 (обратимость унитреугольных матриц)
Диагональ, идущая из левого верхнего угла квадратной матрицы в правый нижний, называется *главной*. Если все стоящие под (соотв. над) главной диагональю элементы нулевые, матрица называется *верхней* (соотв. *нижней*) *треугольной*.

Упражнение 7.19. Проверьте, что верхние и нижние треугольные матрицы являются подкольцами[1] в кольце $\mathrm{Mat}_n(R)$ для любого кольца R.

Треугольные матрицы с единицами на главной диагонали называются *унитреугольными*. Покажем, что каждая верхняя унитреугольная матрица $A = \left(a_{ij}\right)$ обратима в кольце $\mathrm{Mat}_n(R)$ для любого кольца R, и обратная к A матрица $B = A^{-1}$ тоже верхняя унитреугольная с наддиагональными элементами

$$b_{ij} = \sum_{s=0}^{j-i-1} (-1)^{s+1} \sum_{i<\nu_1<\ldots<\nu_s<j} a_{i\nu_1} a_{\nu_1\nu_2} a_{\nu_2\nu_3} \ \cdots \ a_{\nu_{s-1}\nu_s} a_{\nu_s j} =$$
$$= -a_{ij} + \sum_{i<k<j} a_{ik}a_{kj} - \sum_{i<k<\ell<j} a_{ik}a_{k\ell}a_{\ell j} + \sum_{i<k<\ell<m<j} a_{ik}a_{k\ell}a_{\ell m}a_{mj} - \cdots . \quad (7\text{-}14)$$

Для этого запишем матрицу A в виде линейной комбинации матриц E_{ij} из упр. 7.16 выше[2]

$$A = E + \sum_{i<j} a_{ij}E_{ij} = E + N\,,$$

где матрица $N = \sum_{i<j} a_{ij}E_{ij}$ представляет собою наддиагональную часть матрицы A. В силу форм. (7-12) на стр. 102 и упр. 7.16, коэффициент при E_{ij} в матрице N^k равен[3] нулю при $j-i < k$, а при $j - i \geqslant k$ представляет собою сумму всевозможных произведений

$$\underbrace{a_{i\nu_1} \cdot a_{\nu_1\nu_2} \cdot \cdots \cdot a_{\nu_{k-2}\nu_{k-1}} \cdot a_{\nu_{k-1}j}}_{k \text{ сомножителей}}, \quad \text{где} \quad i < \nu_1 < \nu_2 < \ldots \nu_{k-1} < j\,.$$

В частности $N^k = 0$ при всех k, больших размера матрицы A. Полагая $x = E$, $y = N$ в равенстве[4]

$$(x+y)(x^{m-1} - x^{m-2}y + \cdots + (-1)^{m-2}xy^{m-2} + (-1)^{m-1}y^{m-1}) = x^m - y^m\,,$$

при достаточно большом m мы получим матричное равенство $A(E - N + N^2 - \cdots) = E$, откуда

$$A^{-1} = E - N + N^2 - N^3 + \ldots\,,$$

что и утверждалось.

[1] Т. е. замкнуты относительно сложения и умножения.

[2] См. также форм. (7-12) на стр. 102.

[3] Продуктивно представлять себе E_{ij} как стрелку, ведущую из числа j в число i на числовой прямой. Произведение k сомножителей E_{ij} отлично от нуля если и только если конец каждой стрелки совпадает с началом предыдущей, и в этом случае такое произведение равно сумме всех перемножаемых стрелок, рассматриваемых как целочисленные векторы на числовой прямой. Таким образом, каждое ненулевое произведение k стрелок имеет длину как минимум k, а разложения элемента E_{ij} в произведение k таких элементов находятся в биекции со всевозможными способами пройти из j в i за k шагов.

[4] Поскольку матрицы E и N коммутируют друг с другом, в результате этой подстановки мы получим верное матричное равенство.

§8. Определители

8.1. Кососимметричные полинейные формы. Функция $M \times \cdots \times M \to K$ от m аргументов из K-модуля M называется *полилинейной формой*[1], если она линейна по каждому своему аргументу при фиксированных остальных, т. е.

$$\omega\,(\dots, \lambda u + \mu w, \dots) = \lambda\,\omega\,(\dots, u, \dots) + \mu\,\omega\,(\dots, w, \dots)\,, \tag{8-1}$$

где обозначенные многоточиями аргументы во всех трёх членах неизменны. Полилинейная форма называется *кососимметричной*, если она обращается в нуль, когда какие-нибудь два аргумента совпадают. Каждая полилинейная кососимметричная форма *знакопеременна* в том смысле, что её значение меняет знак при перестановке любых двух аргументов. В самом деле, если форма ω полилинейна и кососимметрична, то для любых $u, w \in M$

$$\begin{aligned}\omega(\dots, u, \dots, w, \dots) + \omega(\dots, w, \dots, u, \dots) = \\ = \omega(\dots, u, \dots, u, \dots) + \omega(\dots, u, \dots, w, \dots) + \omega(\dots, w, \dots, u, \dots) + \omega(\dots, w, \dots, w, \dots) = \\ = \omega(\dots, u + w, \dots, u + w, \dots) = 0\,,\end{aligned}$$

где обозначенные многоточиями аргументы во всех членах не меняются.

Упражнение 8.1. Покажите, что если $2 = 1 + 1$ не делит нуль в K, то знакопеременность равносильна кососимметричности.

Полилинейные формы образуют K-модуль относительно обычных операций сложения функций и умножения функций на константы, а кососимметричные формы составляют в нём подмодуль.

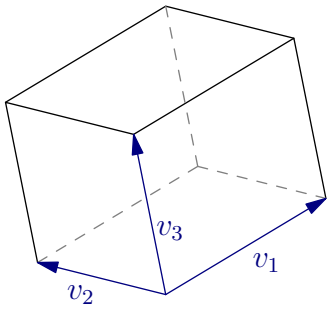


Рис. 8⋄1. Параллелепипед.

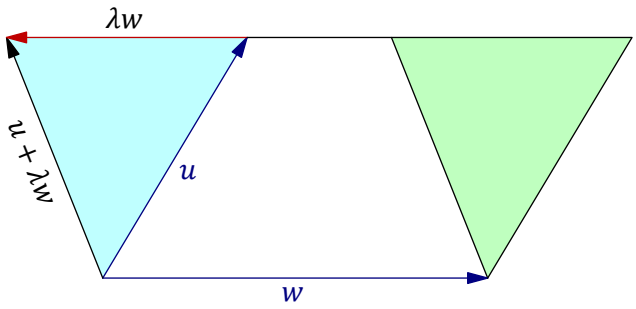


Рис. 8⋄2. Параллельный перекос.

Пример 8.1 (форма объёма на векторном пространстве)
Ненулевая функция от n аргументов $\omega : V \times \cdots \times V \to \Bbbk$ на n-мерном векторном пространстве V над полем $\Bbbk$ называется *объёмом ориентированного n-мерного параллелепипеда* или *формой n-мерного объёма*, если её значение не меняется при добавлений к любому из аргументов произвольной кратности любого другого аргумента, т. е.

$$\omega\,(\dots, u + \lambda w, \dots, w, \dots) = \omega\,(\dots, u, \dots, w, \dots)\,, \tag{8-2}$$

а при умножении любого из аргументов на скаляр её значение умножается на этот скаляр, т. е.

$$\omega(\dots, \lambda v, \dots) = \lambda\,\omega(\dots, v, \dots)\,. \tag{8-3}$$

[1] Или *m-линейной формой* на M, когда важно явно указать количество аргументов.

На геометрическом языке эти свойства означают, что объём параллелепипеда, натянутого на векторы $v_1, \dots, v_n$, как на рис. 8⋄1, умножается на λ при умножении любого ребра на λ, и не меняется при сдвиге двух противоположных $(n-1)$-мерных граней друг относительно друга в направлении какого-нибудь параллельного этим граням ребра (параллельная проекция происходящего на двумерную плоскость, порождённую ребром, вдоль которого делается сдвиг, и ребром, соединяющим сдвигаемые грани, изображена на рис. 8⋄2 выше).

Покажем, что каждая форма n-мерного объёма ω кососимметрична и полилинейна. Первое вытекает из того, что форма объёма обращается в нуль, если один из аргументов линейно выражается через остальные. Скажем, если $v_1 = \lambda_2 v_2 + \dots + \lambda_n v_n$, то

$$\begin{aligned}\omega(v_1, \dots, v_n) = \omega(\lambda_2 v_2 + \dots + \lambda_n v_n,\, v_2, \dots, v_n) = \\ = \omega(0 + \lambda_2 v_2 + \dots + \lambda_n v_n, v_2, \dots, v_n) = \omega(0, v_2, \dots, v_n) = \\ = \omega(0 \cdot 0, v_2, \dots, v_n) = 0 \cdot \omega(0, v_2, \dots, v_n) = 0\,.\end{aligned}$$

Равенство $\omega(\dots, \lambda u + \mu w, \dots) = \lambda\,\omega(\dots, u, \dots) + \mu\,\omega(\dots, w, \dots)$ тривиальным образом выполнено, когда оба набора аргументов в его правой части линейно зависимы: в этом случае набор аргументов в левой части тоже линейно зависим, и обе части нулевые, поскольку линейная зависимость над полем означает, что один из векторов линейно выражается через остальные, и по предыдущему форма объёма обращается на таких векторах в нуль. Поэтому без ограничения общности можно считать, что аргументы первого слагаемого в правой части образуют базис пространства V. Тогда $w = \varrho u + v$, где v является линейной комбинацией остальных $n-1$ аргументов, и левая часть равенства равна

$$\omega(\dots, \lambda u + \mu\varrho u + \mu v, \dots) = \omega(\dots, (\lambda + \mu\varrho)u, \dots) = (\lambda + \mu\varrho)\omega(\dots, u, \dots)\,,$$

а второе слагаемое правой части переписывается как $\mu\omega(\dots, \varrho u + v, \dots) = \mu\varrho \cdot \omega(\dots, u, \dots)$, что и доказывает линейность.

Наоборот, любая n-линейная кососимметричная форма на n-мерном векторном пространстве является формой объёма, поскольку условие (8-3) является составной частью линейности, а условие (8-2) вытекает из линейности и кососимметричности: $\omega(\dots, u + \lambda w, \dots, w, \dots) = \omega(\dots, u, \dots, w, \dots) + \lambda\omega(\dots, w, \dots, w, \dots) = \omega(\dots, u, \dots, w, \dots)$.

8.1.1. Ключевое вычисление. Если модуль $N \simeq K^n$ свободен ранга n, и набор векторов $\boldsymbol{e} = (e_1, \dots, e_n)$ образует базис N над K, то значение произвольной n-линейной кососимметричной формы $\omega\colon M \times \dots \times M \to K$ на любом наборе векторов $(v_1, \dots, v_n) = (e_1, \dots, e_n)\,C$, где в j-том столбце матрицы C стоят координаты вектора v_j в базисе $\boldsymbol{e}$, выражается через значение $\omega(e_1, \dots, e_n)$. В самом деле, поскольку ω линейна по каждому аргументу,

$$\omega(v_1, \dots, v_n) = \omega\Big(\sum_{i_1} e_{i_1} c_{i_1 1}, \dots, \sum_{i_n} e_{i_n} c_{i_n n}\Big) = \sum_{i_1, \dots, i_n} c_{i_1 1} \cdots c_{i_n n}\,\omega\left(e_{i_1}, \dots, e_{i_n}\right).$$

Так как при совпадении каких-либо двух аргументов форма ω зануляется, в последней сумме отличны от нуля только слагаемые с попарно разными индексами $i_1, \dots, i_n$. Каждый такой набор индексов имеет вид $g(1), \dots, g(n)$, где $g\colon \{1, \dots, n\} \xrightarrow{\sim} \{1, \dots, n\}$ — некоторая биекция. Множество всех таких биекций обозначается S_n и называется *группой перестановок n символов* или *n-той симметрической группой*. Перестановка, меняющая местами какие-либо два элемента i, j и

оставляющая все остальные элементы на месте, обозначается σ_{ij} и называется *транспозицией* i-го и j-го элементов.

Упражнение 8.2. Убедитесь, что каждая перестановка $g \in S_n$ является композицией транспозиций.

Разложение перестановки в композицию транспозиций не единственно: например, транспозицию $\sigma_{13} = (3,2,1) \in S_3$ иначе можно записать как $\sigma_{12}\sigma_{23}\sigma_{12}$ или как $\sigma_{23}\sigma_{12}\sigma_{23}$. Тем не менее, чётность количества транспозиций, в композицию которых раскладывается данная перестановка g, не зависит от способа разложения.

8.1.2. Отступление: знак и длина перестановки. Назовём упорядоченную пару $i < j$ элементов множества $\{1, \dots, n\}$ *инверсной* для перестановки $g = (g_1, \dots, g_n) \in S_n$, если $g_i > g_j$. Таким образом, каждая перестановка $g \in S_n$ разбивает множество всех $n(n-1)/2$ упорядоченных пар $i < j$ на два непересекающихся подмножества — инверсные пары и неинверсные пары. Количество $\ell(g)$ инверсных пар перестановки g называется *числом инверсий* или *длиной* перестановки g.

Упражнение 8.3. Найдите $\max \ell(g)$ по всем $g \in S_n$ и укажите все перестановки на которых он достигается.

Число $\operatorname{sgn}(g) \stackrel{\text{def}}{=} (-1)^{\ell(g)}$ называется *знаком* перестановки g. Перестановка g называется *чётной*, если $\operatorname{sgn}(g) = 1$ и *нечётной*, если $\operatorname{sgn}(g) = -1$.

Лемма 8.1
$\operatorname{sgn}(g\sigma_{ij}) = -\operatorname{sgn}(g)$ для любой перестановки $g = (g_1, \dots, g_n)$ и любой транспозиции σ_{ij}.

Доказательство. Перестановки

$$\begin{aligned} g &= \big(g_1, \dots, g_{i-1}, \boldsymbol{g_i}, g_{i+1}, \dots, g_{i-1}, \boldsymbol{g_j}, g_{j+1}, \dots, g_n\big) \\ g\sigma_{ij} &= \big(g_1, \dots, g_{i-1}, \boldsymbol{g_j}, g_{i+1}, \dots, g_{i-1}, \boldsymbol{g_i}, g_{j+1}, \dots, g_n\big) \end{aligned} \tag{8-4}$$

отличаются друг от друга транспозицией элементов g_i и g_j, стоящих на i-том и j-том местах перестановки g. В этих двух перестановках пара (i,j), а также $2(j-i-1)$ пар вида (i,m) и (m,j) с произвольным m из промежутка $i < m < j$ имеют противоположную инверсность, а инверсность всех остальных пар одинакова. □

Следствие 8.1
Если перестановка g является композицией m транспозиций, то $\operatorname{sgn}(g) = (-1)^m$ и чётность перестановки совпадает с чётностью числа m.

Доказательство. Тождественная перестановка не имеет инверсных пар и, стало быть, чётна. В силу леммы, перестановка получающаяся из тождественной умножением на m транспозиций, имеет чётность $(-1)^m$. □

Упражнение 8.4. Убедитесь, что $\operatorname{sgn}(gh) = \operatorname{sgn}(g)\operatorname{sgn}(h)$, т. е. отображение $\operatorname{sgn}\colon S_n \to \{\pm 1\}$ является гомоморфизмом групп.

Пример 8.2 (правило ниточек)
Чётность числа инверсных пар может быть определена следующим наглядным способом, известным как *правило ниточек*[1]. Запишем исходные числа и их перестановку друг под другом, как на рис. 8⋄3, и соединим одинаковые числа нитями так, чтобы ни одна из нитей не вылезала за пределы прямоугольника, образованного четырьмя угловыми числами, и чтобы все точки пересечения нитей были простыми двойными[2]. Тогда чётность числа инверсных пар равна чётности числа точек пересечения нитей.

Упражнение 8.5. Докажите это и найдите при помощи правила ниточек чётность перестановки $(i_1, \dots, i_k, j_1, \dots, j_m)$, в которой $i_1 < i_2 < \dots < i_k$ и $j_1 < j_2 < \dots < j_m$.

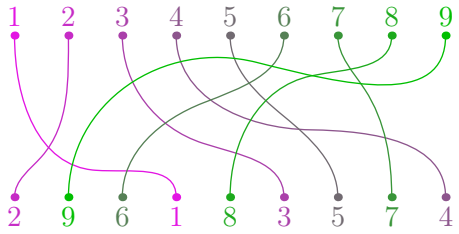


Рис. 8⋄3. sgn(2, 9, 6, 1, 8, 3, 5, 7, 4) = +1 (всего 18 пересечений).

8.1.3. Определитель матрицы. Продолжим вычисление, начатое в n° 8.1.1 выше. В силу знакопеременности формы ω, для каждой перестановки $g \in S_n$ выполняется равенство

$$\omega(e_{g(1)}, \dots, e_{g(n)}) = \operatorname{sgn}(g)\omega(e_1, \dots, e_n),$$

где знак $\operatorname{sgn}(g) = \pm 1$ перестановки g равен $+1$ для чётных перестановок, и -1 для нечётных. Таким образом, для свободного модуля ранга n с базисом $e_1, \dots, e_n$ значение любой n-линейной кососимметричной формы ω на произвольном наборе векторов $(v_1, \dots, v_n) = (e_1, \dots, e_n)\,C$ выражается через её значение на базисе по формуле

$$\omega(v_1, \dots, v_n) = \omega(e_1, \dots, e_n) \cdot \sum_{g \in S_n} \operatorname{sgn}(g) c_{g(1)1} c_{g(2)2} \cdots c_{g(n)n}\,. \tag{8-5}$$

Правая сумма называется *определителем* $n \times n$ матрицы $C = (c_{ij})$ и обозначается

$$\det C \stackrel{\text{def}}{=} \sum_{g \in S_n} \operatorname{sgn}(g) c_{g(1)1} c_{g(2)2} \cdots c_{g(n)n}\,. \tag{8-6}$$

Таким образом, для вычисления определителя следует всеми возможными способами выбирать n элементов в матрице C так, чтобы в каждой строке и в каждом столбце был выбран ровно один элемент. Клетки, где находятся выбранные элементы, задают биекцию $g : j \mapsto g(j)$ из множества столбцов в множество строк матрицы C. Каждую выбранную n-ку элементов следует перемножить и умножить на знак перестановки g, которую она задаёт. Полученные таким образом $n!$ произведений складываются.

[1] Этот способ не слишком эффективен, когда требуется отыскать знак конкретной перестановки длинного набора чисел — обычно быстрее бывает разложить перестановку в композицию непересекающихся циклов и воспользоваться тем, что циклы чётной длины нечётны, а циклы нечётной длины чётны. Однако правило ниточек часто оказывается полезным при анализе абстрактных перестановок.

[2] Т. е. в каждой точке пересечения встречается ровно две нити, причём их касательные в точке пересечения различны.

Пример 8.3
Определители матриц размера 2×2 и 3×3 имеют вид

$$\det\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = c_{11}c_{22} - c_{12}c_{21} \tag{8-7}$$

$$\begin{aligned}\det\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = c_{11}c_{22}c_{33} + c_{13}c_{21}c_{32} + c_{12}c_{23}c_{31} - \\ - c_{11}c_{23}c_{32} - c_{13}c_{22}c_{31} - c_{12}c_{21}c_{33}\,.\end{aligned} \tag{8-8}$$

Во втором равенстве сначала выписаны тождественная и две циклических перестановки, потом — три транспозиции.

Пример 8.4 (определитель треугольной матрицы)
Если матрица C верхнетреугольная[1], т. е. $c_{ij} = 0$ при $i > j$, то единственным ненулевым слагаемым в сумме (8-6) будет произведение диагональных элементов матрицы C, отвечающее тождественной перестановке $g = \mathrm{Id}$. Таким образом, для верхнетреугольной матрицы C определитель $\det C = \prod_i c_{ii}$. В частности, $\det E = 1$.

Предложение 8.1
Для любой квадратной матрицы C выполняется равенство $\det C = \det C^t$.

Доказательство. Суммы (8-6), вычисляющие $\det C$ и $\det C^t$, состоят из одних и тех же произведений всевозможных n-ок элементов матрицы, устанавливающих биекцию $g : j \mapsto g_j$ между номерами столбцов и номерами строк, только в первой из сумм отвечающее такой биекции произведение берётся со знаком $\mathrm{sgn}(g)$, а во второй — со знаком $\mathrm{sgn}(g^{-1})$. Но обратные друг другу перестановки имеют одинаковую чётность: если $g = \sigma_1\sigma_2\cdots\sigma_m$, где σ_i — транспозиции, то $g^{-1} = \sigma_m\sigma_{m-1}\cdots\sigma_1$ в силу равенства $\sigma_i\sigma_i = \mathrm{Id}$. □

Предложение 8.2
Определитель линеен по каждому столбцу матрицы C и обращается в нуль, если какие-то два столбца совпадают.

Доказательство. Первое вытекает из формулы (8-6): так как каждое из суммируемых произведений линейно зависит от каждого столбца, вся сумма тоже линейна по каждому столбцу. Если i-й столбец матрицы C совпадает с j-м, то в сумме (8-6) слагаемое, отвечающее перестановке g сократится со слагаемым, отвечающим перестановке $h = g\sigma_{ij}$, где σ_{ij} меняет местами i и j, а все остальные номера оставляет на месте. В самом деле, $\mathrm{sgn}(h) = -\mathrm{sgn}(g)$, а отвечающие h и g произведения матричных элементов совпадают: $\cdots c_{h(i)i}\cdots c_{h(j)j}\cdots = \cdots c_{g(j)i}\cdots c_{g(i)j}\cdots = {}$
$= \cdots c_{g(j)j}\cdots c_{g(i)i}\cdots = \cdots c_{g(i)i}\cdots c_{g(j)j}\cdots$. □

Следствие 8.2
Определитель $n\times n$-матрицы является n-линейной кососимметричной функцией как столбцов, так и строк.

[1] См. прим. 7.8 на стр. 105.

Следствие 8.3
Модуль n-линейных кососимметричных форм на свободном модуле ранга n с базисом $e_1, \dots, e_n$ свободен и имеет ранг 1. Базисным элементом этого модуля является форма $\omega_{\boldsymbol{e}}$, принимающая на векторах $(v_1, \dots, v_n) = (e_1, \dots, e_n) \cdot C$ значение $\omega_{\boldsymbol{e}}(v_1, \dots, v_n) = \det C$. Координатой произвольной n-линейной кососимметричной формы ω в этом базисе является число $\omega(e_1, \dots, e_n)$.

Доказательство. Форма $\omega_{\boldsymbol{e}}$ полилинейна и кососимметрична по предл. 8.2. Она не является тождественно нулевой, поскольку $\omega_{\boldsymbol{e}}(e_1, \dots, e_n) = \det E = 1$, как мы видели в прим. 8.4. По форм. (8-5) на стр. 109 для любой полилинейной кососимметричной формы ω и любого набора векторов $(v_1, \dots, v_n) = (e_1, \dots, e_n) \cdot C$ выполняется равенство

$$\omega(v_1, \dots, v_n) = \omega(e_1, \dots, e_n) \cdot \det C = \omega(e_1, \dots, e_n)\omega_{\boldsymbol{e}}(v_1, \dots, v_n),$$

означающее, что ω пропорциональна $\omega_{\boldsymbol{e}}$ и коэффициент пропорциональности определяется формой ω однозначно. □

8.1.4. Определитель линейного эндоморфизма. Мы по-прежнему обозначаем через N свободный K-модуль ранга n. Всякое K-линейное отображение $F : N \to N$ задаёт K-линейное отображение модуля n-линейных кососимметричных форм на N в себя, переводящее каждую форму $\omega : N \times \dots \times N \to K$ в форму $\omega_F : N \times \dots \times N \to K$, значения которой вычисляются по правилу

$$\omega_F(v_1, \dots, v_n) \stackrel{\text{def}}{=} \omega\left(Fv_1, \dots, Fv_n\right).$$

Упражнение 8.6. Убедитесь, что форма ω_F полилинейна, кососимметрична и линейно зависит от ω.

Упражнение 8.7. Убедитесь, что всякий линейный эндоморфизм K-модуля, порождённого одним элементом, является умножением на константу.

Таким образом, отображение $\omega \mapsto \omega_F$ умножает все n-линейные кососимметричные формы на одно и то же число. Это число обозначается $\det F$ и называется *определителем* линейного эндоморфизма $F : V \to V$. Поскольку для любого базиса $\boldsymbol{e} = (e_1, \dots, e_n)$ в N векторы $(Fe_1, \dots, Fe_n) = (e_1, \dots, e_n) F_{\boldsymbol{e}}$, где $F_{\boldsymbol{e}}$ — матрица оператора F в базисе $\boldsymbol{e}$, для базисной формы $\omega = \omega_{\boldsymbol{e}}$, построенной по базису $\boldsymbol{e}$ согласно сл. 8.3, имеем

$$\omega_F(e_1, \dots, e_n) = \omega_{\boldsymbol{e}}(Fe_1, Fe_2, \dots, Fe_n) = \omega_{\boldsymbol{e}}(e_1, \dots, e_n) \cdot \det F_{\boldsymbol{e}},$$

откуда $\det(F) = \det F_{\boldsymbol{e}}$. Таким образом, определитель линейного эндоморфизма равен определителю его матрицы в любом базисе и не зависит от выбора базиса.

Поскольку при последовательном выполнении операторов $G : M \to M$ и $F : M \to M$ преобразование $\omega \mapsto \omega_G$ умножает каждую форму ω на $\det G$, а преобразование $\omega \mapsto \omega_F$ умножает каждую форму ω на $\det F$, мы заключаем, что преобразование $\omega \mapsto \omega_{FG}$ умножает каждую форму ω на произведение $\det(F) \cdot \det(G)$. Таким образом, для любых двух линейных эндоморфизмов $F, G : M \to M$ выполняется равенство

$$\det(FG) = \det(F) \det(G) \tag{8-9}$$

В частности, $\det(FG) = \det(GF)$. Применяя это равенство к линейным эндоморфизмам

$$A : x \mapsto Ax \quad \text{и} \quad B : x \mapsto Bx$$

координатного модуля K^n, заданным в его стандартном базисе любыми матрицами A и B, мы заключаем, что для квадратных матриц с элементами из произвольного коммутативного кольца K выполняется равенство

$$\det(AB) = \det(A)\det(B)\,. \tag{8-10}$$

В частности, беря в качестве K кольцо многочленов $\mathbb{Z}[a_{ij}, b_{ij}]$ с целыми коэффициентами от $2n^2$ независимых переменных a_{ij} и b_{ij}, а в качестве $A = \left(a_{ij}\right)$ и $B = \left(b_{ij}\right)$ матрицы, элементами которых являются эти переменные, мы заключаем, что равенство (8-10) представляет собою *формальное тождество* на независимые коммутирующие переменные a_{ij} и b_{ij}.

Следствие 8.4
Если квадратная матрица $A \in \mathrm{Mat}_n(K)$ обратима, то её определитель $\det A$ обратим в K.

Доказательство. Вычисляя определители обеих частей равенства $A \cdot A^{-1} = E$, получаем $\det(A) \cdot \det(A^{-1}) = \det(E) = 1$. □

8.2. Присоединённая матрица и правила Крамера. Для векторов $v_1, \ldots, v_n$ из координатного модуля K^n обозначим через $\det(v_1, \ldots, v_n)$ определитель матрицы, составленной из координат этих векторов. Поскольку определитель не меняется при транспонировании, не имеет значения как записываются координаты — по строкам или по столбцам.

Предложение 8.3 (первое правило Крамера)
Если векторы $v_1, \ldots, v_n$ образуют базис в K^n, то $\det(v_1, \ldots, v_n)$ обратим в K и i-тая координата произвольного вектора $w = x_1 v_1 + \cdots + x_n v_n$ в этом базисе равна

$$x_i = \frac{\det\left(v_1, \ldots, v_{i-1},\, w,\, v_{i+1}, \ldots, v_n\right)}{\det(v_1, \ldots, v_n)}\,. \tag{8-11}$$

Доказательство. Если векторы $v_1, \ldots, v_n \in \Bbbk^n$ образуют базис, то матрица их координат обратима по предл. 7.1 на стр. 97, а значит, $\det(v_1, \ldots, v_n)$ обратим по сл. 8.4. Применяя к обеим частям равенства $w = x_1 e_1 + \cdots + x_n e_n$ линейную функцию

$$K^n \to K\,, \quad u \mapsto \det\left(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_n\right)\,,$$

получаем равенство $\det\left(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n\right) = x_i \cdot \det(v_1, \ldots, v_n)$. □

8.2.1. Присоединённая матрица. Для квадратной матрицы $C = (c_{ij}) \in \mathrm{Mat}_n(K)$ обозначим через C_{ij} подматрицу размера $(n-1) \times (n-1)$, которая получается из C удалением i-й строки и j-го столбца. Число $(-1)^{i+j} \det C_{ij}$ называется *алгебраическим дополнением* к элементу c_{ij} матрицы C. Транспонированная к матрице из алгебраических дополнений матрица

$$C^\vee = \left(c_{ij}^\vee\right)\,, \quad \text{где} \quad c_{ij}^\vee = (-1)^{i+j} \det C_{ji}\,,$$

называется *присоединённой*[1] к матрице C.

Предложение 8.4 (формула для обратной матрицы)
Если матрица $C \in \mathrm{Mat}_n(K)$ обратима, то $C^{-1} = \frac{1}{\det C}\, C^\vee$.

[1] По-английски *adjunct*.

Доказательство. Если матрица C обратима, то её столбцы $v_1, \dots, v_n$ образуют базис $\boldsymbol{v}$ координатного модуля K^n. Стандартный базис $\boldsymbol{e} = (e_1, \dots, e_n)$ в K^n выражается через него по формуле $\boldsymbol{e} = \boldsymbol{v}\,C^{-1}$. Таким образом, i-й элемент j-го столбца матрицы C^{-1} является коэффициентом при v_i в разложении вектора e_j по базису $\boldsymbol{v}$. По правилу Крамера он равен

$$\frac{\det\left(v_1, \dots, v_{i-1}, e_j, v_{i+1}, \dots, v_n\right)}{\det C}.$$

В числителе стоит определитель матрицы, имеющей в i-м столбце ровно один ненулевой элемент — единицу, стоящую в j-й строке. Переставим её в верхний левый угол, сделав $i-1$ транспозиций столбцов и $j-1$ транспозиций строк:

$$\det\left(v_1, \dots, v_{i-1}, e_j, v_{i+1}, \dots, v_n\right) = (-1)^{i-1}\det\left(e_j, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\right) =$$

$$= (-1)^{i+j-2}\det\begin{pmatrix} 1 & c_{j,1} & \cdots & c_{j,i-1} & c_{j,i+1} & \cdots & c_{j,n} \\ 0 & c_{1,2} & \cdots & c_{1,i-1} & c_{1,i+1} & \cdots & c_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & c_{j-1,2} & \cdots & c_{j-1,i-1} & c_{j-1,i+1} & \cdots & c_{j-1,n} \\ 0 & c_{j+1,2} & \cdots & c_{j+1,i-1} & c_{j+1,i+1} & \cdots & c_{j+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & c_{n,1} & \cdots & c_{n,i-1} & c_{n,i+1} & \cdots & c_{n,n} \end{pmatrix}.$$

Ненулевой вклад в этот определитель дают только перестановки, оставляющие 1 на месте. Сумма произведений матричных элементов, отвечающих таким перестановкам, равна определителю $(n-1)\times(n-1)$-матрицы, получающейся удалением j-й строки и i-го столбца из матрицы C. Тем самым, $\det\left(v_1, \dots, v_{i-1}, e_j, v_{i+1}, \dots, v_n\right) = c_{ij}^{\vee}$. □

Пример 8.5

Матрицы размеров 2×2 и 3×3 с определителем 1 обращаются по формулам

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}^{-1} = \begin{pmatrix} (c_{22}c_{33} - c_{23}c_{32}) & -(c_{12}c_{33} - c_{13}c_{31}) & (c_{12}c_{23} - c_{13}c_{22}) \\ -(c_{21}c_{33} - c_{23}c_{31}) & (c_{11}c_{33} - c_{13}c_{31}) & -(c_{11}c_{23} - c_{13}c_{21}) \\ (c_{21}c_{32} - c_{22}c_{31}) & -(c_{11}c_{32} - c_{12}c_{32}) & (c_{11}c_{22} - c_{12}c_{21}) \end{pmatrix}$$

Для матриц с отличным от единицы определителем все матричные элементы в правых частях надо поделить на определитель матрицы из левой части.

Лемма 8.2

Над бесконечным полем $\Bbbk$ многочлен $f(x_1, \dots, x_m) \in \Bbbk[x_1, \dots, x_m]$ принимает нулевое значение в каждой точке аффинного координатного пространства $\Bbbk^m$ если и только если все его коэффициенты нулевые.

Доказательство. Индукция по числу переменных m. При $m = 1$ ненулевой многочлен $f \in \Bbbk[x]$ имеет не более $\deg f$ корней и, тем самым, не может обращаться в нуль во всех точках бесконечной прямой $\Bbbk$. При $m > 1$ запишем $f(x_1, \dots, x_m) = \sum_{k\geqslant 0} f_k(x_1, \dots, x_{m-1})\cdot x_m^k$ как многочлен от x_m с коэффициентами из $\Bbbk[x_1, \dots, x_{m-1}]$. Так как для любой точки $p = (p_1, \dots, p_{m-1}) \in \Bbbk^{m-1}$

многочлен от одной переменной $f_p(x_m) = \sum_{k\geqslant 0} f_k(p)\cdot x_m^k \in \Bbbk[x_m]$, полученный подстановкой координат точки p во все коэффициенты $f_k(x_1,\dots,x_{m-1})$, тождественно зануляется на всей прямой, по уже доказанному все $f_k(p)=0$ для всех $p\in \Bbbk^{m-1}$. По индукции, все коэффициенты всех многочленов $f_k(x_1,\dots,x_{m-1})$ нулевые. Значит и у f все коэффициенты нулевые. □

Теорема 8.1
Обозначим через $K=\mathbb{Z}[c_{ij}]$ кольцо многочленов от n^2 переменных c_{ij}, где $1\leqslant i,j\leqslant n$, а через $C=\left(c_{ij}\right)\in \mathrm{Mat}_n(K)$ матрицу, элементами которой являются эти переменные. В кольце $\mathrm{Mat}_n(K)$ матриц с элементами из K выполняется равенство

$$C\cdot C^\vee = C\cdot C^\vee = \det(C)\cdot E\,. \tag{8-12}$$

Доказательство. Приравнивая соответственные матричные элементы в правой и левой части равенства (8-12), мы получаем набор из n^2 равенств между многочленами с целыми коэффициентами от переменных c_{ij}. Чтобы доказать каждое такое равенство, достаточно проверить, что оно превращается в верное числовое равенство для всех наборов из n^2 численных значений $c_{ij}\in\mathbb{R}$. Более того, поскольку многочлены являются непрерывными функциями $\mathbb{R}^{n^2}\to\mathbb{R}$, численные равенства достаточно проверять не всюду, а на некотором всюду плотном подмножестве в $\mathbb{R}^{n^2}$.

Упражнение 8.8 (по анализу). Убедитесь в этом, а также в том, что для любого ненулевого многочлена $f\in\mathbb{R}[x_1,\dots,x_m]$ множество $\mathcal{D}(f)=\{p\in\mathbb{R}^m \mid f(p)\neq 0\}$ всюду плотно в $\mathbb{R}^m$.

Таким образом, достаточно проверить равенство (8-12) для всех числовых матриц $C\in\mathrm{Mat}_n(\mathbb{R})$, имеющих $\det C\neq 0$. Столбцы такой матрицы линейно независимы, так как если бы один из них линейно выражался через другие, определитель был бы нулевым. Таким образом, столбцы матрицы C образуют базис векторного пространства $\mathbb{R}^n$, а значит, матрица C обратима и для неё выполняется предл. 8.4, а с ним и формула (8-12). □

Следствие 8.5
Квадратная матрица C с элементами в произвольном коммутативном кольце K с единицей обратима если и только если $\det C$ обратим в K, и в этом случае обратная матрица вычисляется согласно предыдущему предл. 8.4. □

Следствие 8.6
Векторы $v_1,\dots,v_n\in K^n$ тогда и только тогда образуют базис в K^n, когда $\det(v_1,\dots,v_n)$ обратим в K, и в этом случае коэффициенты линейного выражения произвольного вектора через этот базис находятся по правилу Крамера из предл. 8.3 на стр. 112. □

Предложение 8.5 (разложение определителя по i-й строке или i-у столбцу)
В кольце $n\times n$ матриц $\mathrm{Mat}_n(K)$ с элементами из кольца $K=\mathbb{Z}[c_{ij}]$ выполняется равенство

$$\det C=\sum_{k=1}^{n}(-1)^{k+i}c_{ik}\det C_{ik}=\sum_{k=1}^{n}(-1)^{k+i}c_{ki}\det C_{ki}\,.$$

Доказательство. Соотношения получаются приравниванием (i,i)-тых диагональных элементов матриц из правой и левой части (8-12). □

Пример 8.6
Раскладывая определитель 3×3 по первому столбцу, получаем

$$\det\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = c_{11}\left(c_{22}c_{33} - c_{23}c_{32}\right) - c_{21}\left(c_{12}c_{33} - c_{13}c_{32}\right) + c_{31}\left(c_{12}c_{23} - c_{13}c_{22}\right).$$

что согласуется с прямым вычислением из прим. 8.3.

Пример 8.7 (однородные системы из n линейных уравнений на $n+1$ неизвестных)
Пространство решений системы из n линейных уравнений

$$\begin{cases} a_{10}x_0 + a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ a_{20}x_0 + a_{21}x_1 + \cdots + a_{2n}x_n = 0 \\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ a_{n0}x_1 + a_{n1}x_1 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{8-13}$$

на $n+1$ неизвестных $(x_0, x_1, \ldots, x_n)$, рассматриваемых как вектор-столбец координатного пространства $\Bbbk^{n+1}$ над произвольным полем $\Bbbk$, является аннулятором линейной оболочки строк матрицы

$$A = \begin{pmatrix} a_{1,0} & a_{1,1} & \cdots & a_{1,n} \\ a_{2,0} & a_{2,1} & \cdots & a_{2,n} \\ \vdots & \cdots & \cdots & \vdots \\ a_{n,0} & a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

в двойственном координатном пространстве $\Bbbk^{n+1*}$. Если строки этой матрицы линейно независимы, пространство решений системы (8-13) одномерно, и базисный вектор в этом подпространстве можно указать явно. Для этого обозначим через

$$A_i \stackrel{\text{def}}{=} (-1)^i \det\begin{pmatrix} a_{1,0} & \cdots & a_{1,i-1} & a_{1,i+1} & \cdots & a_{1,n} \\ a_{2,0} & \cdots & a_{2,i-1} & a_{2,i+1} & \cdots & a_{2,n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{n,0} & \cdots & a_{n,i-1} & a_{n,i+1} & \cdots & a_{n,n} \end{pmatrix} \tag{8-14}$$

определитель $n \times n$ матрицы, получающихся из A выкидыванием i-го столбца. Покажем, что уравнения (8-13) линейно независимы если и только если вектор $a = (A_0, A_1, \ldots, A_n) \neq 0$, и в этом случае вектор a порождает одномерное пространство решений системы (8-13).

Для этого допишем к матрице A сверху ещё одну копию её i-той строки. Определитель получившейся матрицы размера $(n+1) \times (n+1)$ равен нулю. Раскладывая его по верхней строке, получаем $a_{i0}A_0 + a_{i1}A_1 + \cdots + a_{in}A_n = 0$. Тем самым, вектор $a = (A_0, A_1, \ldots, A_n)$ в любом случае является решением системы (8-13). Если строки матрицы A линейно зависимы, то и строки всех матриц (8-14) линейно зависимы с теми же самыми коэффициентами. Поэтому все компоненты вектора A в таком случае нулевые. Если же ковекторы $\alpha_i = (a_{i,0}, a_{i,1}, \ldots, a_{i,n})$ линейно независимы в $\Bbbk^{n+1*}$, то по лемме о замене[1] их можно дополнить до базиса в $\Bbbk^{n+1*}$ одним из стандартных базисных ковекторов e_i^*. Определитель матрицы

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{10} & \cdots & \cdots & a_{1i} & \cdots & \cdots & a_{ni} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n0} & \cdots & \cdots & a_{ni} & \cdots & \cdots & a_{nn} \end{pmatrix},$$

[1] См. лем. 6.2 на стр. 89.

в строки которой записаны координаты базисных ковекторов $e_i^*, \alpha_1, \dots, \alpha_n$, отличен от нуля. Раскладывая его по первой строке, видим, что он равен $(-1)^i A_i$, откуда $A_i \neq 0$.

8.3. Тождество Гамильнотна – Кэли. Для любого коммутативного кольца K с единицей кольцо $n \times n$ матриц $\mathrm{Mat}_n(K[t])$ с элементами из кольца многочленов $K[t]$ совпадает с кольцом многочленов $\mathrm{Mat}_n(K)[t]$ от переменной t с коэффициентами в кольце матриц $\mathrm{Mat}_n(K)$, поскольку каждую матрицу, в клетках которой стоят многочлены от t, можно записать как многочлен от t с матричными коэффициентами и наоборот. Например,

$$\begin{pmatrix} 3t^2+2t & t^3-1 \\ 2t+3 & t^3+t-1 \end{pmatrix} = t^3 \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + t^2 \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} + t \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 3 & -1 \end{pmatrix}.$$

Определение 8.1
Для матрицы $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ многочлен

$$\chi_A(t) \stackrel{\text{def}}{=} \det(tE - A) = t^n - \sigma_1(A) \cdot t^{n-1} + \dots + (-1)^{n-1}\sigma_{n-1}(A) \cdot t + (-1)^n \sigma_n(A) \in K[t]$$

называется *характеристическим многочленом* матрицы A. Коэффициент при t^{n-k} в характеристическом многочлене обозначается через $(-1)^k \sigma_k(A)$.

Упражнение 8.9. Убедитесь, что число $\sigma_k(A) \in K$ равно сумме определителей всех таких $k \times k$ подматриц матрицы A, главная диагональ которых является подмножеством главной диагонали матрицы A. В частности, $\sigma_1(A) = \mathrm{tr}(A)$ и $\sigma_n(A) = \det A$.

Теорема 8.2 (тождество Гамильтона – Кэли)
Пусть, как и выше, $K = \mathbb{Z}[a_{ij}]$ является кольцом многочленов от n^2 переменных a_{ij}. Тогда в кольце матриц $\mathrm{Mat}_n(K)$ для матрицы $A = (a_{ij})$ выполняется равенство $\chi_A(A) = 0$.

Доказательство. Подставляя в форм. (8-12) на стр. 114 вместо C матрицу $tE - A$, где E — единичная матрица размера $n \times n$, заключаем, что в кольце $\mathrm{Mat}_n(K[t])$ выполняется равенство

$$\det(tE - A) \cdot E = (tE - A)(tE - A)^\vee,$$

где $(tE-A)^\vee$ — присоединённая[1] к $(tE-A)$ матрица. Перепишем это равенство в виде равенства между многочленами от t с коэффициентами в кольце матриц $\mathrm{Mat}_n(K)$:

$$t^n \cdot E - \sigma_1(A)\, t^{n-1} \cdot E + \dots + (-1)^n \sigma_n(A) \cdot E = (tE - A)\left(t^m \cdot A_m^\vee + \dots + t \cdot A_1^\vee + A_0^\vee\right),$$

где $A_0^\vee, A_1^\vee, \dots, A_m^\vee \in \mathrm{Mat}_n(K)$ — некоторые матрицы. Подставляя в него $t = A$, получаем в кольце $\mathrm{Mat}_n(K)$ равенство $\chi_A(A) \cdot E = 0$, откуда $\chi_A(A) = 0$. □

8.4. Грассмановы многочлены. Полезным алгебраическим инструментом для работы с кососимметричными формами и определителями является алгебра $\Bbbk\langle \xi_1, \xi_2, \dots, \xi_n \rangle$ *грассмановых многочленов* от переменных $\xi_1, \dots, \xi_n$ с коэффициентами из поля $\Bbbk$. Она определяется точно

[1] См. n° 8.2.1 на стр. 112.

также, как и обычная алгебра многочленов, с той только разницей, что грассмановы переменные ξ_i не коммутируют, но *антикоммутируют* друг с другом, т. е. подчиняются соотношениям[1]

$$\forall\, i,j \quad \xi_i \wedge \xi_j = -\xi_j \wedge \xi_i \quad \text{и} \quad \forall\, i \quad \xi_i \wedge \xi_i = 0\,, \tag{8-15}$$

где символ «$\wedge$» обозначает кососимметричное грассманово умножение, дабы отличать его от обычного коммутативного. Поскольку квадраты грассмановых переменных равны нулю, всякий ненулевой грассманов моном *линеен* по каждой входящей в него переменной. Иначе говоря, для каждого строго возрастающего набора $I = (i_1, \ldots, i_m)$ номеров $i_1 < i_2 < \ldots < i_m$ имеется грассманов моном

$$\xi_I \stackrel{\text{def}}{=} \xi_{i_1} \wedge \ldots \wedge \xi_{i_m}\,, \tag{8-16}$$

который при перестановке $g \in S_m$ переменных $\xi_{i_1}, \ldots, \xi_{i_m}$ меняет знак по правилу

$$\xi_{i_{g(1)}} \wedge \xi_{i_{g(2)}} \wedge \ldots \wedge \xi_{i_{g(m)}} = \operatorname{sgn}(g) \cdot \xi_{i_1} \wedge \ldots \wedge \xi_{i_m}\,. \tag{8-17}$$

Мономы (8-16), занумерованные всевозможными подмножествами $I \subset \{1,\, 2,\, \ldots\,,\, n\}$, составляют базис алгебры $\Bbbk\langle \xi_1, \xi_2, \ldots, \xi_n\rangle$ как векторного пространства над $\Bbbk$ и перемножаются по правилу

$$\xi_I \wedge \xi_J = \begin{cases} \operatorname{sgn}(I,J) \cdot \xi_{I \sqcup J} & \text{если } I \cap J = \varnothing \\ 0 & \text{если } I \cap J \neq \varnothing \end{cases} \tag{8-18}$$

где $\operatorname{sgn}(I,J) = \pm 1$ обозначает знак *тасующей перестановки*, расставляющей в порядке возрастания набор номеров $i_1, \ldots, i_m, j_1, \ldots, j_k$, в котором $i_1 < i_2 < \ldots < i_m$ и $j_1 < j_2 < \ldots < j_k$. Если наборы $I = (i_1, \ldots, i_m)$ и $J = \{1,\, 2,\, \ldots\,,\, n\} \smallsetminus I$ дополнительны друг к другу, то согласно упр. 8.5 на стр. 109 этот знак $\operatorname{sgn}(I,J) = (-1)^{i_1+i_2+\cdots+i_m+m(m+1)/2}$.

Единственный моном старшей степени $\xi_{\text{top}} \stackrel{\text{def}}{=} \xi_1 \wedge \ldots \wedge \xi_n$ аннулируется умножением на любой грассманов многочлен с нулевым свободным членом. Однородные грассмановы многочлены степени k образуют векторное пространство размерности $\binom{n}{k}$, базис в котором составляют мономы (8-16), отвечающие всевозможным k-элементным подмножествам I. Размерность всей грассмановой алгебры $\dim \Bbbk\langle \xi_1, \xi_2, \ldots, \xi_n\rangle = 2^n$.

Два грассмановых монома степеней m и k коммутируют друг с другом по правилу

$$\left(\xi_{i_1} \wedge \ldots \wedge \xi_{i_m}\right) \wedge \left(\xi_{j_1} \wedge \ldots \wedge \xi_{j_k}\right) = \\ = (-1)^{km} \left(\xi_{j_1} \wedge \ldots \wedge \xi_{j_k}\right) \wedge \left(\xi_{i_1} \wedge \ldots \wedge \xi_{i_m}\right)\,,$$

ибо при переносе каждой из k переменных ξ_j через m переменных ξ_i происходит m транспозиций. Поэтому для любых двух однородных грассмановых многочленов η и ω

$$\eta \wedge \omega = (-1)^{\deg\eta \deg\omega} \omega \wedge \eta\,. \tag{8-19}$$

В частности, каждый однородный многочлен чётной степени коммутирует со всеми грассмановыми многочленами.

Упражнение 8.10. Опишите *центр*[2] грассмановой алгебры.

[1] Если $\operatorname{char}\Bbbk \neq 2$ соотношения $\xi_i \wedge \xi_i = 0$ вытекают из соотношений $\xi_i \wedge \xi_j = -\xi_j \wedge \xi_i$ и могут быть опущены. Однако когда $\operatorname{char}\Bbbk = 2$ именно соотношения на квадраты $\xi_i \wedge \xi_i = 0$ отличает грассмановы переменные от обычных коммутативных.

[2] Т. е. подалгебру, состоящую из всех грассмановых многочленов, которые коммутируют со всеми грассмановыми многочленами.

8.4.1. Грассманова алгебра векторного пространства. Если в векторном пространстве V выбран базис $e_1, \dots, e_n$, алгебра грассмановых многочленов $\Bbbk\langle e_1, e_2, \dots, e_n\rangle$ от базисных векторов пространства V обозначается ΛV и называется *грассмановой* (или *внешней*) алгеброй векторного пространства V. Не апеллирующие к выбору базиса название и обозначение вызваны тем, что пространство однородных грассмановых многочленов степени 1 канонически отождествляется с пространством V и, таким образом, не зависит от выбора базиса, а пространство однородных грассмановых многочленов степени k является линейной оболочкой всевозможных произведений $v_1 \wedge \dots \wedge v_k$ из k произвольных векторов $v_i \in V$ и тоже не зависит от выбора базиса. Обозначая пространство однородных грассмановых многочленов степени k через $\Lambda^k V$, мы получаем разложение алгебры ΛV в прямую сумму векторных пространств

$$\Lambda V = \bigoplus_{k=0}^{n} \Lambda^k V\,,$$

где $\Lambda^0 V \stackrel{\text{def}}{=} \Bbbk \cdot 1$ обозначает одномерное пространство констант, тоже не зависящее от базиса.

8.4.2. Линейные замены переменных. Если векторы $\boldsymbol{u} = (u_1, \dots, u_\ell)$ линейно выражены через векторы $\boldsymbol{w} = (w_1, \dots, w_k)$ по формуле $\boldsymbol{u} = \boldsymbol{w}\,C$, где $C = \big(c_{ij}\big) \in \mathrm{Mat}_{k\times\ell}(\Bbbk)$, то их грассмановы произведения $u_J = u_{j_1} \wedge \dots \wedge u_{j_m}$ линейно выражаются через грассмановы произведения $w_I = w_{i_1} \wedge \dots \wedge w_{i_m}$ по формулам

$$u_J = u_{j_1} \wedge \dots \wedge u_{j_m} = \Big(\sum_{i_1} w_{i_1} c_{i_1 j_1}\Big) \wedge \Big(\sum_{i_2} w_{i_2} c_{i_2 j_2}\Big) \wedge \dots \wedge \Big(\sum_{i_m} w_{i_m} c_{i_m j_m}\Big) =$$
$$= \sum_{1\leqslant i_1<i_2<\dots<i_n\leqslant n} w_{i_1} \wedge \dots \wedge w_{i_n} \cdot \sum_{g\in S_m} \operatorname{sgn}(g)\, c_{i_{g(1)} j_1} c_{i_{g(2)} j_2} \dots c_{i_{g(n)} j_n} = \sum_I w_I \cdot c_{IJ}\,,$$

где $c_{IJ} = \det C_{IJ}$ обозначает определитель $m \times m$-подматрицы $C_{IJ} \subset C$, сосредоточенной в пересечениях столбцов с номерами из J и строк с номерами из I, а суммирование происходит по всем наборам $I = (i_1, \dots, i_m)$ из m возрастающих номеров $1 \leqslant i_1 < i_2 < \dots < i_m \leqslant \ell$. Определитель $c_{IJ} = \det C_{IJ}$ называется IJ-тым *минором* m-того порядка в матрице C. Таким образом, IJ-тый элемент матрицы, выражающей грассманов моном u_J через грассмановы мономы w_I равен IJ-тому минору m-того порядка в матрице выражающей векторы $\boldsymbol{u}$ через векторы $\boldsymbol{w}$.

В частности, если наборы векторов $\boldsymbol{e} = (e_1, \dots, e_n)$ и $\boldsymbol{f} = (f_1, \dots, f_n)$ оба являются базисами пространства V, то базисные грассмановы мономы e_J пространства $\Lambda^m V$ выражаются через базисные мономы f_I при помощи матрицы перехода размера $\binom{m}{n} \times \binom{m}{n}$, у которой в позиции IJ стоит IJ-тый минор $\big(c_{IJ}\big)$ матрицы $C_{\boldsymbol{fe}}$, выражающей $\boldsymbol{e}$ через $\boldsymbol{f}$. Эта матрица обозначается $\Lambda^m C_{\boldsymbol{fe}}$ и называется m-той *внешней степенью* матрицы $C_{\boldsymbol{fe}}$.

8.5. Соотношения Лапласа. Для набора возрастающих чисел $J = (j_1, \dots, j_m) \subset \{1, \dots, n\}$ положим $\deg J \stackrel{\text{def}}{=} m$, $|J| \stackrel{\text{def}}{=} j_1 + j_2 + \dots + j_m$ и условимся обозначать через

$$\hat{J} = \big(\hat{j}_1, \hat{j}_2, \dots, \hat{j}_{n-m}\big) = \{1, 2, \dots, n\} \smallsetminus J$$

дополнительный к J набор из $\deg \hat{J} = n - m$ возрастающих номеров.

Рассмотрим произвольную квадратную матрицу $A \in \mathrm{Mat}_{n\times n}(\Bbbk)$, столбцы которой обозначим $\alpha_1, \dots, \alpha_n$ и будем воспринимать как векторы координатного пространства $\Bbbk^n$. Матрица A является матрицей перехода от этих векторов к стандартному базису $e_1, \dots, e_n$ пространства $\Bbbk^n$. Для любых двух мультииндексов I, J одинаковой длины $\deg I = \deg J = m$ грассмановы

мономы $\alpha_J = \alpha_{j_1} \wedge \ldots \wedge \alpha_{j_m}$ и $\alpha_{\hat{I}} = \alpha_{\hat{i}_1} \wedge \ldots \wedge \alpha_{\hat{i}_{n-m}}$ имеют дополнительные степени m и $n-m$ и перемножаются по форм. (8-18) на стр. 117, которая с учётом упр. 8.5 имеет вид:

$$\alpha_J \wedge \alpha_{\hat{I}} = \begin{cases} (-1)^{|J|+\frac{m(m+1)}{2}} \alpha_1 \wedge \ldots \wedge \alpha_n & \text{при } I = J \\ 0 & \text{при } I \neq J\,. \end{cases} \tag{8-20}$$

Выражая мономы α_J и $\alpha_{\hat{I}}$ в левой части (8-20) через базисные мономы e_K, получаем

$$\left(\sum_K e_K a_{KJ}\right) \wedge \left(\sum_L e_L a_{L\hat{I}}\right) = (-1)^{\frac{m(m+1)}{2}} e_1 \wedge \ldots \wedge e_n \sum_K (-1)^{|K|} a_{KJ} a_{\hat{K}\hat{I}}\,,$$

где K пробегает все возрастающие мультииндексы длины $\deg K = m$. Так как правая часть (8-20) при $I = J$ равна $(-1)^{\frac{m(m+1)}{2}+|J|} \det A \cdot e_1 \wedge \ldots \wedge e_n$, для любых двух наборов J, I из m строк произвольной квадратной матрицы A выполняются *соотношения Лапласа*

$$\sum_K (-1)^{|K|+|J|} a_{KJ} a_{\hat{K}\hat{I}} = \begin{cases} \det A & \text{при } I = J \\ 0 & \text{при } I \neq J \end{cases} \tag{8-21}$$

где суммирование идёт по всем наборам K из $m = \deg K$ строк матрицы A.

При $I = J$ соотношение (8-21) даёт формулу для вычисления определителя[1]

$$\det A = \sum_K (-1)^{|K|+|J|} a_{KJ} a_{\hat{K}\hat{J}} \tag{8-22}$$

через всевозможные миноры a_{KJ} порядка m, сосредоточенные в m фиксированных столбцах матрицы A с номерами J, и *дополнительные* к ним миноры $a_{\hat{J}\hat{K}}$ порядка $n-m$, равные определителям матриц, получающихся из A вычёркиванием всех строк и столбцов, которые высекают минор a_{KJ}. Произведение $(-1)^{|K|+|J|} a_{\hat{K}\hat{J}}$ называется *алгебраическим дополнением* к минору a_{KJ} и обозначается $\hat{a}_{KJ}$.

Упражнение 8.11. Для любых матриц $A \in \mathrm{Mat}_n(\Bbbk)$, $C \in \mathrm{Mat}_m(\Bbbk)$, $B \in \mathrm{Mat}_{n\times m}(\Bbbk)$ покажите, что $\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \cdot \det C$.

При $I \neq J$ соотношение (8-21) имеет вид $\sum_K a_{KJ} \hat{a}_{IK} = 0$ и называется *теоремой об умножении на чужие алгебраические дополнения*, поскольку его левая часть отличается от левой части формулы (8-22) тем, что миноры a_{KJ} умножаются не на свои алгебраические дополнения $\hat{a}_{KJ}$, а на дополнения $\hat{a}_{IK}$ к минорам a_{IK}, сосредоточенным в другом наборе столбцов $I \neq J$.

Если согласованно занумеровать все m-элементные подмножества и все $(n-m)$-элементные подмножества в множестве $\{1, 2, \ldots, n\}$ так, чтобы дополнительные подмножества J и $\hat{J}$ имели одинаковые номера, то соотношения Лапласа можно записать одним равенством

$$\Lambda^m A \cdot \Lambda^{n-m} \hat{A}^t = \det A \cdot E \tag{8-23}$$

на матрицы размера $\binom{n}{m} \times \binom{n}{m}$, в котором (IJ)-тый элемент матрицы $\Lambda^{n-m} \hat{A}^t$ равен

$$\hat{a}_{JI} = (-1)^{|J|+|I|} a_{\hat{J}\hat{I}}\,.$$

[1] С геометрической точки зрения эта формула вычисляет объём n-мерного параллелепипеда через объёмы его m-мерных и $(n-m)$-мерных граней.

Упражнение 8.12. Установите транспонированный вариант соотношений Лапласа

$$\sum_K a_{JK}\hat{a}_{IK} = \begin{cases} \det A & \text{при } I = J \\ 0 & \text{при } I \neq J \end{cases} \tag{8-24}$$

Пример 8.8 (соотношения Плюккера)
Рассмотрим 2×4 матрицу $A = \bigl(a_{ij}\bigr) \in \mathrm{Mat}_{2\times 4}(\Bbbk)$ и обозначим через A_{ij} её 2×2 минор, образованный i-м и j-м столбцами. Шесть чисел A_{ij} не могут принимать произвольные значения. Они связаны квадратичным *соотношением Плюккера*

$$A_{12}A_{34} - A_{13}A_{24} + A_{14}A_{23} = 0\,, \tag{8-25}$$

которое получается при раскрытии нулевого определителя 4×4 матрицы $\begin{pmatrix} A \\ A \end{pmatrix}$ по первым двум строкам.

Упражнение 8.13. Убедитесь в этом и для любых шести чисел A_{ij}, удовлетворяющих соотношению (8-25), явно предъявите 2×4 матрицу A с 2×2 минорами A_{ij}.

Пример 8.9 (определитель пучка матриц)
Линейная оболочка пары непропорциональных квадратных матриц $A, B \in \mathrm{Mat}_{n\times n}(\Bbbk)$ называется *пучком матриц* и обозначается (AB). Таким образом, всякая матрица из пучка (AB) имеет вид t_0A+t_1B, где $t_0, t_1 \in \Bbbk$, а её определитель $\det(t_0A+t_1B)$ является однородным многочленом степени n от t_0, t_1. Покажем, что коэффициент этого многочлена при $t_0^k t_1^{n-k}$ равен

$$\sum_{IJ} a_{IJ}\hat{b}_{IJ}\,, \tag{8-26}$$

где суммирование идёт по всем k-элементным подмножествам $I, J \subset \{1, 2, \ldots, n\}$.

Для этого обозначим через $a_1, \ldots, a_n$ и $b_1, \ldots, b_n$ столбцы матриц A и B, понимаемые как векторы координатного пространства $\Bbbk^n$ со стандартным базисом $e_1, \ldots, e_n$. Тогда

$$(t_0a_1 + t_1b_1) \wedge (t_0a_2 + t_1b_2) \wedge \ldots \wedge (t_0a_n + t_1b_n) = \det(t_0A + t_1B)\, e_1 \wedge \ldots \wedge e_n\,.$$

Моном $t_0^k t_1^{n-k}$ возникает в левой части при выборе первого слагаемого в каких-нибудь k из перемножаемых скобок и второго слагаемого в остальных $n-k$ скобках. Если обозначить номера этих k скобок через $I = (i_1, \ldots, i_k)$ то вклад в коэффициент при $t_0^k t_1^{n-k}$ будет равен

$$(-1)^{\frac{k(k+1)}{2}+|I|} a_I \wedge b_{\hat{I}} = (-1)^{\frac{k(k+1)}{2}+|I|}\Bigl(\sum_J e_J a_{JI}\Bigr) \wedge \Bigl(\sum_K e_K b_{K\hat{I}}\Bigr) =$$
$$= (-1)^{\frac{k(k+1)}{2}+|I|} \sum_{JK} e_J \wedge e_K \cdot a_{JI} b_{K\hat{I}} = e_1 \wedge \ldots \wedge e_n \cdot \sum_J (-1)^{|I|+|J|} a_{JI} b_{\hat{J}\hat{I}}$$

Полный коэффициент при $t_0^k t_1^{n-k}$ в $\det(t_0A + t_1B)$ получается суммированием таких подобных слагаемых по всем наборам I из k возрастающих номеров, что и даёт формулу (8-26). В обозначениях из (8-23) её можно переписать в виде

$$\det(t_0A + t_1B) = \sum_{k=0}^{n} \operatorname{tr}\bigl(\Lambda^k A \cdot \Lambda^{n-k}\hat{B}^t\bigr)\, t_0^k t_1^{n-k}\,, \tag{8-27}$$

§9. Конечно порождённые модули над кольцами главных идеалов

Всюду в этом параграфе K по умолчанию означает произвольное кольцо главных идеалов. Все рассматриваемые нами K-модули предполагаются конечно порождёнными. Под свободным K-модулем ранга нуль понимается нулевой K-модуль.

9.1. Метод Гаусса. Рассмотрим произвольную матрицу $A \in \mathrm{Mat}_{m\times n}(K)$ над кольцом главных идеалов K. Элементарным преобразованием строк матрицы A называется замена каких-либо её двух строк a_i и a_j их линейными комбинациями $a_i' = \alpha a_i + \beta a_j$ и $a_j' = \gamma a_i + \delta a_j$ с определителем $\alpha\delta - \beta\gamma = \pm 1$. В этом случае матрица преобразования

$$\begin{pmatrix} a_i \\ a_j \end{pmatrix} \mapsto \begin{pmatrix} a_i' \\ a_j' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a_i \\ a_j \end{pmatrix}$$

обратима, и строки a_i и a_j могут быть выражены обратно через преобразованные строки a_i' и a_j' по формулам $a_i' = \delta a_i - \beta a_j$, $a_j' = -\gamma a_i + \alpha a_j$, если $\alpha\delta - \beta\gamma = 1$, и по формулам $a_i' = -\delta a_i + \beta a_j$, $a_j' = \gamma a_i - \alpha a_j$, если $\alpha\delta - \beta\gamma = -1$.

Упражнение 9.1. Убедитесь в этом.

Таким образом, элементарное преобразование строк матрицы A не меняет линейной оболочки строк матрицы и заключается в умножении матрицы A слева на такую обратимую матрицу $L \in \mathrm{GL}_m(K)$, которая получается из единичной $m \times m$ матрицы E тем же самым элементарным преобразованием строк, которое производится в матрице A.

Симметричным образом, элементарное преобразование столбцов матрицы A заключается в замене каких-либо её двух столбцов a_i и a_j их линейными комбинациями $a_i' = \alpha a_i + \beta a_j$ и $a_j' = \gamma a_i + \delta a_j$ с определителем $\alpha\delta - \beta\gamma = \pm 1$. Такое преобразование не меняет линейной оболочки столбцов матрицы A и заключается в умножении матрицы A справа на обратимую матрицу $R \in \mathrm{GL}_n(K)$, которая получается из единичной $n \times n$ матрицы E тем же самым элементарным преобразованием столбцов, которое производится в матрице A.

Лемма 9.1
Любую пару стоящих в одной строке (соотв. в одном столбце) матрицы A ненулевых элементов (a, b) можно подходящим элементарным преобразованием содержащих их столбцов (соотв. строк) заменить парой $(d, 0)$, где $d = \mathrm{нод}(a, b)$ — наибольший общий делитель[1] a и b.

Доказательство. Запишем $d = \mathrm{нод}(a, b)$ как $d = ax + by$ и пусть $a = da'$, $b = db'$. Тогда $a'x + b'y = 1$ и $a'b - b'a = 0$. Таким образом,

$$(a, b) \cdot \begin{pmatrix} x & -b' \\ y & a' \end{pmatrix} = (d, 0) \quad \text{и} \quad \begin{pmatrix} x & y \\ -b' & a' \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix},$$

где $\det \begin{pmatrix} x & -b' \\ y & a' \end{pmatrix} = \det \begin{pmatrix} x & y \\ -b' & a' \end{pmatrix} = 1$. □

[1] Напомню, что он определён с точностью до умножения на обратимые элементы кольца K, см. n° 5.3.2 на стр. 72.

Теорема 9.1
Любая прямоугольная матрица C над кольцом главных идеалов конечным числом элементарных преобразований строк и столбцов может быть преобразована в такую матрицу D, у которой $d_{ij} = 0$ при $i \neq j$ и $d_{ii} \mid d_{jj}$ при $i < j$, причём эта матрица D не зависит от выбора последовательности элементарных преобразований.

Доказательство. Сначала перестановками строк и столбцов добьёмся того, чтобы $c_{11} \neq 0$. Пусть в матрице C есть элемент a, не делящийся на c_{11}, и пусть $d = \text{нод}\,(a, c_{11})$. Тогда $(c_{11}) \subsetneq (d)$, и если мы перейдём от матрицы C к матрице C' с $c'_{11} = d$, то идеал, порождаемый левым верхним угловым элементом, строго увеличится. Покажем, что это всегда можно сделать элементарными преобразованиями.

Если не делящийся на c_{11} элемент a стоит в первой строке или первом столбце, достаточно заменить пару (c_{11}, a) на $(d, 0)$ согласно лем. 9.1. Если все элементы первой строки и первого столбца делятся на c_{11}, а не делящийся на c_{11} элемент a стоит строго ниже и правее c_{11}, то мы сначала занулим все элементы первой строки и первого столбца за исключением самого c_{11}, добавляя ко всем столбцам подходящие кратные первого столбца, а ко всем строкам — подходящие кратные первой строки. К элементу a при этом будут добавляться числа, кратные c_{11}, и он останется не делящимся на c_{11}. Далее, прибавим ту строку, где стоит a, к первой строке и получим в первой строке копию элемента a. Наконец, заменим пару (c_{11}, a) на $(d, 0)$ по лем. 9.1.

Так как кольцо главных идеалов нётерово, идеал (c_{11}) не может увеличиваться бесконечно долго, и после конечного числа описанных выше переходов мы получим матрицу C, все элементы которой делятся на c_{11}. У этой матрицы, как уже объяснялось выше, можно обнулить все элементы первой строки и первого столбца за исключением c_{11}. Все элементы подматрицы, стоящей в остальных строках и столбцах, при этом останутся делящимися на c_{11}. По индукции, эту подматрицу можно диагонализовать элементарными преобразованиями строк и столбцов. При этом первая строка и первый столбец не поменяются.

Чтобы доказать независимость получающейся в результате диагональной матрицы D от выбора цепочки элементарных преобразований, обозначим через $\Delta_k(C) \in K$ наибольший общий делитель всех $k \times k$-миноров прямоугольной матрицы $C \in \mathrm{Mat}_{m\times n}(K)$. Для матрицы D, ненулевые элементы которой исчерпываются стоящими на главной диагонали числами

$$d_{11} \mid d_{22} \mid \ldots \mid d_{rr}\,,$$

каждое из которых делит все последующие, $\Delta_k(D) = d_{11}d_{22}\ldots d_{kk}$, откуда $d_{kk} = \Delta_k(D)/\Delta_{k-1}(D)$. В силу идущей ниже леммы $\Delta_k(D) = \Delta_k(C)$, поскольку матрица $D = LCR$ получается из матрицы C умножением слева и справа на обратимые матрицы $L \in \mathrm{GL}_m(K)$ и $R \in \mathrm{GL}_n(K)$. Это доказывает независимость итоговых диагональных элементов $d_{kk} = \Delta_k(C)/\Delta_{k-1}(C)$ от выбора преобразований. □

Лемма 9.2
При умножении матрицы C слева или справа на обратимую квадратную матрицу наибольший общий делитель $\Delta_k(C)$ её $k \times k$-миноров не меняется[1].

Доказательство. Поскольку $\Delta_k(C) = \Delta_k(C^t)$ достаточно рассмотреть только левое умножение. Пусть $F = LC$, где L обратима. Тогда каждый $k \times k$ минор матрицы F является K-линейной комбинацией $k \times k$ миноров матрицы C.

Упражнение 9.2. Убедитесь в этом.

[1] С точностью до умножения на обратимые элементы кольца K.

Поэтому $\Delta_k(F)$ делится на $\Delta_k(C)$. Аналогично, из равенства $C = A^{-1}F$ вытекает, что $\Delta_k(C)$ делится на $\Delta_k(F)$. Тем самым, $\Delta_k(C)$ и $\Delta_k(F)$ отличаются обратимым множителем. □

Определение 9.1
Числа $\lambda_k(C) \stackrel{\text{def}}{=} \Delta_k(C)/\Delta_{k-1}(C)$ называются *инвариантными множителями* прямоугольной матрицы $C \in \mathrm{Mat}_{m\times n}(K)$.

Следствие 9.1
Для любой матрицы $C \in \mathrm{Mat}_{m\times n}(K)$ над кольцом главных идеалов K существуют такие обратимые матрицы $L \in \mathrm{GL}_m(K)$ и $R \in \mathrm{GL}_n(K)$, что матрица $D = LCR$ имеет $d_{kk} = \lambda_k = \Delta_k(C)/\Delta_{k-1}(C)$ и $d_{ij} = 0$ при $i \neq j$. □

Замечание 9.1. Обратимые матрицы L и R, преобразующие матрицу C в диагональную матрицу $D = LCR$, представляют собою произведения $L = L_\ell \dots L_2 L_1$ и $R = R_1 R_2 \dots R_r$ обратимых матриц L_i и R_j, осуществляющих последовательные элементарные преобразования строк и столбцов матрицы C. Таким образом, $L = L_\ell \dots L_1 E$ и $R = E R_1 \dots R_r$ получаются применением к единичным матрицам E размеров $m \times m$ и $n \times n$ тех же самых цепочек элементарных преобразований строк и соответственно столбцов, которые осуществляются с матрицей C. Поэтому для явного отыскания матриц L и R следует приписать к матрице $C \in \mathrm{Mat}_{m\times n}(K)$ справа и снизу единичные матрицы размеров $m \times m$ и $n \times n$ соответственно, так что получится Г-образная таблица вида $\left[\begin{array}{c|c} C & E \\ \hline E & \end{array}\right]$, и в процессе приведения матрицы C к диагональному виду осуществлять элементарные преобразования строк и столбцов сразу во всей Г-образной таблице. Тогда на выходе получится Г-образная таблица $\left[\begin{array}{c|c} D & L \\ \hline R & \end{array}\right]$, содержащая наряду с итоговой диагональной матрицей D искомые матрицы L и R, такие что $LCR = D$.

9.2. Теорема об инвариантных множителях. Как мы видели в прим. 6.12 на стр. 87, произвольный K-модуль M, линейно порождённый над K векторами $w_1, \dots, w_m$, является фактором $M \simeq K^m/R_{\boldsymbol{w}}$ координатного модуля K^m по подмодулю $R_{\boldsymbol{w}} \subset K^m$ линейных соотношений между порождающими векторами $\boldsymbol{w}$. Подмодуль $R_{\boldsymbol{w}}$ состоит из всех таких $(x_1, \dots, x_m) \in K^m$, что $x_1 w_1 + \dots + x_m w_m = 0$ в M и представляет собою ядро эпиморфизма

$$\pi_{\boldsymbol{w}} : K^m \twoheadrightarrow M\,, \quad (x_1, \dots, x_m) \mapsto x_1 w_1 + \dots + x_m w_m\,. \tag{9-1}$$

Согласно теор. 6.5 на стр. 93 подмодуль соотношений тоже свободен и имеет $\mathrm{rk}\, R_{\boldsymbol{w}} \leqslant m$. Следующая теорема позволяет выбрать в модуле соотношений особенно удобный базис.

Теорема 9.2 (об инвариантных множителях)
Для любого подмодуля N в свободном модуле F ранга m над кольцом главных идеалов K существует такой базис $\boldsymbol{e} = (e_1, \dots, e_m)$ модуля F над K, что подходящие кратности $\lambda_1 e_1, \dots, \lambda_n e_n$ первых $n \leqslant m$ его базисных векторов составляют базис в N, причём каждый из множителей λ_i делится на все предыдущие множители λ_j с $j < i$. Набор множителей $\lambda_1, \dots, \lambda_n$ с точностью до умножения на обратимые элементы из K не зависит от выбора такого базиса.

Доказательство. Зафиксируем произвольный базис $\boldsymbol{w} = (w_1, \dots, w_m)$ в F и какой-нибудь набор векторов $\boldsymbol{u} = (u_1, \dots, u_k) = \boldsymbol{w}\, C_{\boldsymbol{wu}}$, порождающих подмодуль $N \subset F$. Напомню, что в j-м столбце матрицы $C_{\boldsymbol{wu}}$ стоят координаты образующей u_j в базисе $\boldsymbol{w}$. По сл. 9.1 существуют такие обратимые матрицы $L \in \mathrm{GL}_m(K)$ и $R \in \mathrm{GL}_k(K)$, что матрица $D = L C_{\boldsymbol{wu}} R$ имеет $d_{ij} = 0$

при $i \neq j$, а каждый её диагональный элемент $d_{ii} = \lambda_i$ делится на все предыдущие. Так как матрица L обратима, набор векторов $\boldsymbol{e} = \boldsymbol{w}\,L^{-1}$ является базисом в F. Набор векторов $\boldsymbol{v} = \boldsymbol{u}\,R$ выражается через этот базис по формуле $\boldsymbol{v} = \boldsymbol{u}\,R = \boldsymbol{w}\,C_{\boldsymbol{wu}}R = \boldsymbol{e}\,LC_{\boldsymbol{wu}}R = \boldsymbol{e}\,D$. Тем самым, в наборе $\boldsymbol{v}$ отличны от нуля в точности первые n векторов $v_i = \lambda_i e_i$. Будучи пропорциональны базисным векторам свободного модуля F, они линейно независимы. Исходный набор образующих $\boldsymbol{u}$ подмодуля N линейно выражается через $\boldsymbol{v}$ по формуле $\boldsymbol{u} = \boldsymbol{v}\,L^{-1}$. Тем самым, ненулевые векторы v_i с $1 \leqslant i \leqslant n$ линейно порождают подмодуль N, а значит, образуют в нём базис. Это устанавливает существование базисов с требуемыми свойствами.

Если в F имеются такие базисы $\boldsymbol{e}' = (e'_1, \dots, e'_m)$ и $\boldsymbol{e}'' = (e''_1, \dots, e''_m)$, что некоторые кратности $v'_i = \lambda'_i e'_i$ и $v''_i = \lambda''_i e''_i$ первых их n векторов составляют базисы подмодуля $N \subset F$, а множители $\lambda'_i \mid \lambda'_j$ и $\lambda''_i \mid \lambda''_j$ при $i < j$, то обе диагональные матрицы перехода $C_{\boldsymbol{v''e''}} = C_{\boldsymbol{v''v'}}C_{\boldsymbol{v'e'}}C_{\boldsymbol{e'e''}}$ и $C_{\boldsymbol{v'e'}} = E_n C_{\boldsymbol{v'e'}} E_m$, где E_n и E_m суть единичные $n\times n$ и $m\times m$ матрицы, удовлетворяют условиям сл. 9.1 для одной и той же $n\times m$ матрицы $C = C_{\boldsymbol{v'e'}}$ и, стало быть, совпадают. Это устанавливает независимость инвариантных множителей от выбора взаимных базисов. □

Определение 9.2
Множители $\lambda_1, \dots, \lambda_n$ из теор. 9.2 называются *инвариантными множителями* подмодуля N в свободном модуле F, а построенные в теор. 9.2 базисы $e_1, \dots, e_m$ в F и $\lambda_1 e_1, \dots, \lambda_n e_n$ в N называются *взаимными базисами* свободного модуля F и его подмодуля N.

Пример 9.1 (подрешётки в $\mathbb{Z}^m$)
По теореме об инвариантных множителях для любой абелевой подгруппы $L \subset \mathbb{Z}^m$ существует такой базис $u_1, \dots, u_m$ в $\mathbb{Z}^m$, что подходящие кратности первых ℓ его базисных векторов $m_1 u_1, \dots, m_\ell u_\ell$ составляют базис в L. Тем самым, L тоже является свободным $\mathbb{Z}$-модулем, а фактор модуль

$$\mathbb{Z}^m / L \simeq \frac{\mathbb{Z}}{(m_1)} \oplus \cdots \oplus \frac{\mathbb{Z}}{(m_\ell)} \oplus \mathbb{Z}^{m-\ell}. \tag{9-2}$$

Выясним, скажем, как устроена подгруппа $L \subset \mathbb{Z}^3$, порождённая столбцами матрицы

$$C = \begin{pmatrix} 126 & 51 & 72 & 33 \\ 30 & 15 & 18 & 9 \\ 60 & 30 & 36 & 18 \end{pmatrix} \tag{9-3}$$

Для этого перейдём к взаимным базисам. Заметим, что нод элементов матрицы (9-3) равен 3, и мы можем получить -3 в позиции $(1,4)$, прибавляя к 1-й строке учетверённую 2-ю:

$$\begin{pmatrix} 6 & -9 & 0 & -3 \\ 30 & 15 & 18 & 9 \\ 60 & 30 & 36 & 18 \end{pmatrix}.$$

Умножаем 1-ю строку на -1 и меняем местами первый и последний столбцы

$$\begin{pmatrix} 3 & 9 & 0 & -6 \\ 9 & 15 & 18 & 30 \\ 18 & 30 & 36 & 60 \end{pmatrix}.$$

Теперь мы можем занулить левый столбец и верхнюю строку вне левого углового элемента, отнимая из 2-й и 3-й строк подходящие кратности 1-й строки, а затем из 2-го и 4-го столбцов

подходящие кратности 1-го столбца

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & -12 & 18 & 48 \\ 0 & -24 & 36 & 96 \end{pmatrix}$$

Зануляем 3-ю строку, отнимая из неё удвоенную 2-ю, и видим, что нод элементов второй строки можно получить, прибавляя ко 2-му столбцу 3-й:

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 18 & 48 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Остаётся занулить 3-й и 4-й столбцы, добавляя к ним подходящие кратности второго:

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Таким образом, $L \simeq \mathbb{Z}^2$, а $\mathbb{Z}^3 / L \simeq \mathbb{Z}/(3) \oplus \mathbb{Z}/(6) \oplus \mathbb{Z}$.

Проделанные элементарные преобразования строк состояли в последовательном умножении слева на матрицы

$$\begin{pmatrix} 1 & -4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -6 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 4 & 0 \\ 3 & -11 & 0 \\ 0 & -2 & 1 \end{pmatrix},$$

а преобразования столбцов — в последовательном умножении справа на матрицы

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -8 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & -3 & -8 \\ 0 & 1 & -2 & -8 \\ 1 & -3 & 9 & 26 \end{pmatrix}.$$

Таким образом базис в решётке L составляют векторы $3\,u_1 = c_4$ и $6\,u_2 = c_2 + c_3 - 3\,c_4$, где c_2, c_3, c_4 суть последние три столбца исходной матрицы C, а u_1, u_2 — первые два вектора взаимного с L базиса объемлющей решётки $\mathbb{Z}^3$, образованного столбцами матрицы

$$U = \begin{pmatrix} -1 & 4 & 0 \\ 3 & -11 & 0 \\ 0 & -2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 11 & 4 & 0 \\ 3 & 1 & 0 \\ 6 & 2 & 1 \end{pmatrix}$$

Пример 9.2 (соизмеримые подрешётки)
Из существования взаимных базисов вытекает, что следующие свойства абелевой подгруппы $L \subset \mathbb{Z}^m$, порождённой столбцами матрицы $C \in \mathrm{Mat}_{m\times n}(\mathbb{Z})$, эквивалентны друг другу:

(1) $\mathrm{rk}\, L = m$

(2) фактор группа $\mathbb{Z}^m / L$ конечна

(3) решётка $L \subset \mathbb{Z}^m$ линейно порождает векторное пространство $\mathbb{Q}^m$ над $\mathbb{Q}$

(4) ранг матрицы C над полем $\mathbb{Q}$ равен m.

Решётки $L \subset \mathbb{Z}^m$, удовлетворяющие этим условиям, называются *соизмеримыми* с $\mathbb{Z}^m$. Если решётка $L \subset \mathbb{Z}^m$ задана как $\mathbb{Z}$-линейная оболочка столбцов некоторой матрицы $C \in \mathrm{Mat}_{m\times n}(\mathbb{Z})$, то чтобы убедиться в её соизмеримости с $\mathbb{Z}^m$ достаточно указать в матрице C ненулевой минор порядка m. Для отыскания ранга решётки L достаточно гауссовыми элементарными преобразованиями строк над полем $\mathbb{Q}$ привести матрицу C или[1] C^t к ступенчатому виду с рациональными элементами.

Предложение 9.1
Столбцы матрицы $C \in \mathrm{Mat}_n(\mathbb{Z})$ порождают соизмеримую с $\mathbb{Z}^n$ абелеву подгруппу $L \subset \mathbb{Z}^n$ если и только если $\det C \neq 0$, и в этом случае $|\mathbb{Z}^n / L| = |\det C|$, т. е. число элементов в факторе по соизмеримой подрешётке равно абсолютной величине объёма параллелепипеда, натянутого на любой её базис.

Доказательство. Рассмотрим в $\mathbb{Z}^m$ такой базис $u_1, \dots, u_m$, что векторы $\lambda_1 u_1, \dots, \lambda_\ell u_\ell$ образуют базис в L. Диагональная матрица D, единственными ненулевыми элементами которой являются $d_{ii} = \lambda_i$ с $1 \leqslant i \leqslant \ell$, связана с матрицей C соотношением $D = LCR$, где матрицы $L, R \in \mathrm{GL}_n(\mathbb{Z})$. Поскольку обратимость целочисленной матрицы равносильна тому, что её определитель равен[2] ± 1, мы заключаем, что $|\det D| = \prod_i |d_{ii}| = |\det C|$. Соизмеримость L с $\mathbb{Z}^m$ равносильна тому, что $\ell = m$ или, что то же самое, тому что все $d_{ii} = \lambda_i$ ненулевые. В этом случае $\mathbb{Z}^n / L = \bigoplus\limits_i \mathbb{Z}/(\lambda_i)$ состоит в точности из $\prod_i |\lambda_i| = |\det C|$ элементов. □

9.3. Теорема об элементарных делителях. Вместо упорядоченного набора инвариантных множителей $\lambda_1, \dots, \lambda_n$ иногда бывает удобнее иметь дело с неупорядоченным дизъюнктным объединением всех степеней p^μ неприводимых элементов $p \in K$, входящих в разложения чисел $\lambda_1, \dots, \lambda_n$ на неприводимые множители. Точнее, рассмотрим для каждого $i = 1, \dots, n$ разложение $\lambda_i = p_{i1}^{m_{i1}} \cdots p_{ik_i}^{m_{ik_i}}$, в котором все p_{ij} неприводимы и p_{ij} не ассоциировано с p_{ik} при $j \neq k$. Неупорядоченное дизъюнктное объединение всех степеней[3] $p_{ij}^{m_{ij}}$, входящих в эти разложения при $i = 1, \dots, n$, называется набором *элементарных делителей* набора инвариантных множителей $\lambda_1, \dots, \lambda_n$.

Лемма 9.3
Описанная только что процедура устанавливает биекцию между упорядоченными наборами чисел[4] $\lambda_1, \dots, \lambda_n \in K$, в которых $\lambda_i | \lambda_j$ при $i < j$, и всевозможными неупорядоченными наборами натуральных степеней p^μ неприводимых чисел[5] из K, в которых разрешаются повторяющиеся элементы[6].

[1] Смотря по тому, в какой из двух матриц меньше строк.

[2] См. сл. 7.1 на стр. 97.

[3] Эпитет «дизъюнктное» означает, что степень p^m, входящая в разложение ровно k инвариантных множителей λ_i, присутствует в итоговом неупорядоченном наборе в точности k раз.

[4] Рассматриваемых с точностью до умножения на обратимые элементы кольца K.

[5] Тоже рассматриваемых с точностью до умножения на обратимые элементы кольца K.

[6] Два таких набора считаются одинаковыми, если их можно привести в биективное соответствие друг с другом так, что у соответственных степеней p^μ и q^ν натуральные показатели μ и ν будут равны другу, а простые основания p и q будут ассоциированы друг с другом.

Доказательство. Набор инвариантных множителей $\lambda_1, \dots, \lambda_n$ однозначно восстанавливается по набору элементарных делителей следующим образом. Расставим элементарные делители в клетки диаграммы Юнга так, чтобы в первой строке шли в порядке нестрого убывания степени того простого числа, степеней которого в наборе элементарных делителей имеется больше всего. Во вторую строку поместим в порядке нестрого убывания все степени простого числа, следующего за первым по общему количеству вхождений его степеней в набор элементарных делителей и т. д. Поскольку наибольший инвариантный множитель λ_n делится на все остальные, его разложение на простые множители содержит *все* встречающиеся среди элементарных делителей простые числа, причём каждое из них — с максимальным показателем. Таким образом, λ_n является произведением всех элементарных делителей, стоящих в первом столбце построенной нами диаграммы Юнга. По индукции мы заключаем, что произведения элементарных делителей по столбцам диаграммы образуют прочитанную справа налево последовательность инвариантных множителей. □

Пример 9.3
Набор элементарных делителей

$$\begin{array}{ccccc} 3^2 & 3^2 & 3 & 3 & 3 \\ 2^3 & 2^3 & 2^2 & 2 & \\ 7^2 & 7 & 7 & & \\ 5 & 5 & & & \end{array}$$

возникает из такого набора инвариантных множителей:

$$\lambda_1 = 3\,,\ \lambda_2 = 3\cdot 2\,,\ \lambda_3 = 3\cdot 2^2\cdot 7\,,\ \lambda_4 = 3^2\cdot 2^3\cdot 7\cdot 5\,,\ \lambda_5 = 3^2\cdot 2^3\cdot 7^2\cdot 5\,.$$

9.3.1. Формулировка основной теоремы. Остаток этого раздела будет посвящён доказательству следующего результата.

Теорема 9.3 (об элементарных делителях)
Всякий конечно порождённый модуль над кольцом главных идеалов K изоморфен

$$K^{n_0} \oplus \frac{K}{\left(p_1^{n_1}\right)} \oplus \dots \oplus \frac{K}{\left(p_\alpha^{n_\alpha}\right)} \tag{9-4}$$

где $m_\nu \in \mathbb{N}$, все $p_\nu \in K$ просты, и слагаемые в прямой сумме могут повторяться. Два модуля

$$K^{n_0} \oplus \frac{K}{\left(p_1^{n_1}\right)} \oplus \dots \oplus \frac{K}{\left(p_\alpha^{n_\alpha}\right)} \quad \text{и} \quad K^{m_0} \oplus \frac{K}{\left(q_1^{m_1}\right)} \oplus \dots \oplus \frac{K}{\left(q_\beta^{m_\beta}\right)}$$

изоморфны если и только если $n_0 = m_0$, $\alpha = \beta$ и слагаемые можно перенумеровать так, чтобы $n_\nu = m_\nu$ и $p_\nu = s_\nu q_\nu$, где все $s_\nu \in K$ обратимы.

Определение 9.3
Набор (возможно повторяющихся) степеней $p_i^{n_i}$, по которым происходит факторизация в (9-4), называется *набором элементарных делителей* модуля (9-4).

9.3.2. Существование разложения (9-4)**.** Пусть K-модуль M порождается векторами

$$w_1, \dots, w_m .$$

Тогда $M = K^m/R$, где R — ядро эпиморфизма $K^m \twoheadrightarrow M$, переводящего стандартные базисные векторы $e_i \in K^m$ в образующие $w_i \in M$, как в форм. (9-1) на стр. 123. По теор. 9.2 в K^m существует такой базис $u_1, \dots, u_m$, что некоторые кратности $\lambda_1 u_1, \dots, \lambda_k u_k$ первых K базисных векторов составляют базис в R. Таким образом,

$$M = K^m/R = K/(\lambda_1) \oplus \dots \oplus K/(\lambda_k) \oplus K^{m-k} .$$

Пусть i-й инвариантный множитель $\lambda_i = p_1^{m_1} \cdots p_s^{m_s}$, где $p_j \in K$ — попарно неассоциированные простые элементы. Тогда по китайской теореме об остатках

$$K/(\lambda_i) = K/\left(p_1^{m_1}\right) \oplus \dots \oplus K/\left(p_s^{m_s}\right) ,$$

что и даёт разложение (9-4). Чтобы установить его единственность, мы дадим инвариантное описание всех слагаемых разложения (9-4) во внутренних терминах модуля M.

9.3.3. Отщепление кручения. Сумма $K/\left(p_1^{n_1}\right) \oplus \dots \oplus K/\left(p_\alpha^{n_\alpha}\right)$ в разложении (9-4) совпадает с подмодулем кручения[1] $\operatorname{Tors} M = \{w \in M \mid \exists \lambda \neq 0 : \lambda w = 0\}$, а число n_0 в разложении (9-4) равно рангу свободного модуля $M/\operatorname{Tors} M$ и не зависит от выбора разложения. Из существования разложения (9-4) вытекает

Следствие 9.2
Всякий конечно порождённый модуль над кольцом главных идеалов является прямой суммой свободного модуля и подмодуля кручения. В частности, любой модуль без кручения автоматически свободен. □

9.3.4. Отщепление p-кручения. Для каждого неприводимого $p \in K$ назовём *p-кручением* в K-модуле M подмодуль, образованный всеми векторами, которые аннулируются умножением на какую-нибудь степень числа p, и обозначим этот подмодуль

$$\operatorname{Tors}_p M \stackrel{\text{def}}{=} \{w \in M \mid \exists k > 0 : p^k w = 0\} .$$

Если простое $q \in K$ не ассоциировано с p, то класс p^k обратим в $K/\left(q^m\right)$, и гомоморфизм умножения на $p^k : K/\left(q^m\right) \to K/\left(q^m\right)$, $x \mapsto p^k x$, является изоморфизмом. В частности, он не имеет ядра. Напротив, каждый модуль $K/\left(p^\ell\right)$ полностью аннулируется умножением на достаточно большую степень p. Поэтому прямая сумма всех слагаемых вида $K/\left(p^m\right)$ в разложении (9-4) совпадает с подмодулем p-кручения $\operatorname{Tors}_p M \subset M$ и тоже не зависит от выбора разложения, а из наличия разложения (9-4) вытекает

Следствие 9.3
Всякий конечно порождённый модуль кручения над кольцом главных идеалов является прямой суммой подмодулей p-кручения по всем простым $p \in K$, для которых p-кручение ненулевое. □

Упражнение 9.3. Обозначим через $\varphi : K/\left(p^m\right) \to K/\left(p^m\right)$, $x \mapsto px$, гомоморфизм умножения на p. Покажите, что: а) $\varphi^n = 0$ при $n \geqslant m$ б) $\ker \varphi^n \supset \ker \varphi^{n-1}$ в) $\ker \varphi^n = \operatorname{im} \varphi^{m-n} \simeq K/\left(p^n\right)$ при $0 < n < m$ г) $\ker \varphi^n / \ker \varphi^{n-1}$ нулевой при $n > m$ и изорфен $K/(p)$ при $1 \leqslant n \leqslant m$.

[1] См. прим. 6.6 на стр. 82.

9.3.5. Инвариантность показателей p-кручения. Для завершения доказательства теор. 9.3 остаётся проверить, что если при простом $p \in K$ слагаемые прямого разложения

$$M = \frac{K}{(p^{\nu_1})} \oplus \cdots \oplus \frac{K}{(p^{\nu_k})} \tag{9-5}$$

выписаны в порядке нестрого убывания показателей $\nu_1 \geqslant \nu_2 \geqslant \cdots \geqslant \nu_k$, то этот набор показателей не зависит от выбора разложения и однозначно определяется модулем M. Для этого рассмотрим диаграмму Юнга ν, строки которой имеют длины $\nu_1, \ldots, \nu_k$, и дадим инвариантное описание длинам столбцов этой диаграммы. Обозначим через $\varphi : M \to M$, $x \mapsto px$, гомоморфизм умножения на p, как в упр. 9.3. Согласно этому упражнению, применённому к правой части разложения (9-5), при каждом $i = 1, 2, 3, \ldots$ фактор модуль $\ker \varphi^i / \ker \varphi^{i-1}$ изоморфен прямой сумме одинаковых слагаемых $K/(p)$ в количестве, равном числу строк диаграммы ν, длина которых не меньше i, т. е. высоте i-го столбца диаграммы ν. С другой стороны, фактор модуль $\ker \varphi^i / \ker \varphi^{i-1}$ никак не зависит от разложения (9-5) и является векторным пространством над полем $K/(p)$: умножение на класс $[x]_p \in K/(p)$ переводит класс $[z] \in \ker \varphi^i / \ker \varphi^{i-1}$ в класс $[xz] \in \ker \ker \varphi^i / \ker \varphi^{i-1}$.

Упражнение 9.4. Убедитесь, что это правило корректно и удовлетворяет аксиомам векторного пространства.

Мы заключаем, что высота i-того столбца диаграммы ν равна размерности векторного пространства $\ker \varphi^i / \ker \varphi^{i-1}$ над полем $K/(p)$. Теорема об элементарных делителях полностью доказана.

9.4. Строение конечно порождённых абелевых групп. При $K = \mathbb{Z}$ теорема об элементарных делителях даёт полную классификацию конечно порождённых абелевых групп.

Теорема 9.4
Всякая конечно порождённая абелева группа изоморфна прямой сумме аддитивных групп

$$\mathbb{Z}^r \oplus \frac{\mathbb{Z}}{\left(p_1^{n_1}\right)} \oplus \cdots \oplus \frac{\mathbb{Z}}{\left(p_\alpha^{n_\alpha}\right)} \tag{9-6}$$

где $p_\nu \in \mathbb{N}$ — простые числа (не обязательно различные). Две аддитивных группы

$$\mathbb{Z}^r \oplus \frac{\mathbb{Z}}{\left(p_1^{n_1}\right)} \oplus \cdots \oplus \frac{\mathbb{Z}}{\left(p_\alpha^{n_\alpha}\right)} \quad \text{и} \quad \mathbb{Z}^s \oplus \frac{\mathbb{Z}}{\left(q_1^{m_1}\right)} \oplus \cdots \oplus \frac{\mathbb{Z}}{\left(q_\beta^{m_\beta}\right)}$$

изоморфны тогда и только тогда, когда $r = s$, $\alpha = \beta$ и после надлежащей перестановки слагаемых $n_\nu = m_\nu$ и $p_\nu = q_\nu$ при всех ν. □

Определение 9.4
Единственное представление заданной конечно порождённой абелевой группы A в виде прямой суммы аддитивных групп (9-6) называется её *каноническим представлением*.

Пример 9.4 (группы, заданные образующими и соотношениями)
На практике конечно порождённые абелевы группы часто задаются описанием вроде: абелева

группа A, порождённая элементами $a_1, \dots, a_n$, которые связаны соотношениями

$$\begin{cases} \mu_{11}a_1 + \mu_{12}a_2 + \cdots + \mu_{1n}a_n = 0 \\ \mu_{21}a_1 + \mu_{22}a_2 + \cdots + \mu_{2n}a_n = 0 \\ \mu_{31}a_1 + \mu_{32}a_2 + \cdots + \mu_{3n}a_n = 0 \\ \qquad \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ \mu_{\mu 1}a_1 + \mu_{\mu 2}a_2 + \cdots + \mu_{mn}a_n = 0\,, \end{cases} \tag{9-7}$$

где $\mu_{ij} \in \mathbb{Z}$. По определению, это означает, что $A = \mathbb{Z}^n/R$, где $R \subset \mathbb{Z}^n$ — подмодуль, порождённый строками $\mu_1, \dots, \mu_m$ матрицы $M = \left(\mu_{ij}\right)$. В каноническом разложении (9-6) группы A ранг r свободного слагаемого равен $n - \operatorname{rk} M$, а степени $p_i^{n_i}$ суть элементарные делители подмодуля $R \subset \mathbb{Z}^n$. Про конкретный элемент $w = x_1a_1 + \cdots + x_na_n$ часто бывает нужно знать, отличен он от нуля в A или нет, и если нет, то каков его порядок[1] $\operatorname{ord}(w)$.

Выяснить первое можно посредством вычислений в векторном пространстве $\mathbb{Q}^n \supset \mathbb{Z}^n$ над полем $\mathbb{Q}$. Если w не лежит в $\mathbb{Q}$-линейной оболочке строк матрицы M, то никакое его целое кратное mw не лежит в R, т. е. $w \neq 0$ в A и $\operatorname{ord} w = \infty$. Если же w лежит в $\mathbb{Q}$-линейной оболочке строк матрицы M, то подходящее целое кратное mw этого элемента лежит в R и класс w в группе $A = \mathbb{Z}^n/R$ имеет конечный порядок. Оценить этот порядок сверху тоже можно при помощи вычислений над полем $\mathbb{Q}$. Если строки $\mu_{i_1}, \dots, \mu_{i_k}$, где $k = \operatorname{rk} M = n - r$, образуют базис в $\mathbb{Q}$-линейной оболочке строк матрицы M, то $\mathbb{Z}$-линейная оболочка этих строк соизмерима[2] с R. Если $w = x_{i_1}\mu_{i_1} + \cdots + x_{i_k}\mu_{i_k}$, где $x_j = p_j/q_j \in \mathbb{Q}$ несократимы, то вектор mw с $m = \text{нок}(q_{i_1}, \dots, q_{i_k})$ лежит в $\mathbb{Z}$-линейной оболочке строк $\mu_{i_1}, \dots, \mu_{i_k}$, а значит, и в R. Поэтому $\operatorname{ord} w \leqslant m$ в группе A. Для точного отыскания порядка $\operatorname{ord} w$ вычислений над $\mathbb{Q}$ уже не достаточно, и требуется явный базис $e_1, \dots, e_k$ модуля R над $\mathbb{Z}$. Если такой базис найден[3], и $w = \sum x_ie_i$, где $x_i = p_i/q_i \in \mathbb{Q}$ несократимы, то $\operatorname{ord} w = \text{нок}(q_1, \dots, q_k)$. В частности, если все $q_i = 1$, то $w = 0$ в $A = \mathbb{Z}^n/R$.

[1] Напомню, что *порядком* $\operatorname{ord}(w)$ элемента w в аддитивной абелевой группе называется наименьшее такое $n \in \mathbb{N}$, что $nw = 0$, или же $\operatorname{ord}(w) = \infty$, если такого n нет (см. n° 3.5.1 на стр. 49).

[2] Т. е. является подгруппой конечного индекса в R, см. прим. 9.2 на стр. 125.

[3] Например, методом Гаусса, как это объяснялось в прим. 9.1 на стр. 124.

§10. Пространство с оператором

10.1. Классификация пространств с оператором. Пусть $\Bbbk$ — произвольное поле, V — конечномерное векторное пространство над $\Bbbk$, а $F: V \to V$ — линейный эндоморфизм пространства V. Мы будем называть пару (F, V) *пространством с оператором* или просто *оператором* над $\Bbbk$. Линейное отображение $C: U_1 \to U_2$ между пространствами с операторами (F_1, U_1) и (F_2, U_2) называется *гомоморфизмом*, если $F_2 \circ C = C \circ F_1$. В этом случае говорят, что диаграмма

$$\begin{array}{ccc} U_1 & \xrightarrow{C} & U_2 \\ {\scriptstyle F_1}\uparrow & & \uparrow{\scriptstyle F_2} \\ U_1 & \xrightarrow[C]{} & U_2 \end{array}$$

коммутативна[1]. Если гомоморфизм C биективен, операторы $F_1: U_1 \to U_1$ и $F_2: U_2 \to U_2$ называются *изоморфными* или *подобными*. Поскольку в этом случае $F_2 = CF_1C^{-1}$, то говорят, что оператор F_2 получается из F_1 *сопряжением* посредством изоморфизма C.

Подпространство $U \subset V$ называется *F-инвариантным*, если $F(U) \subset U$. В этом случае пара $(F|_U, U)$ тоже является пространством с оператором и вложение $U \hookrightarrow V$ представляет собою гомоморфизмом пространств с операторами. Оператор, не имеющий инвариантных подпространств, отличных от нуля и всего пространства, называется *неприводимым* или *простым*.

Упражнение 10.1. Покажите, что оператор умножения на класс $[t]$ в фактор кольце $\mathbb{R}[t]/(t^2+1)$ неприводим.

Оператор $F: V \to V$ называется *разложимым*, если V раскладывается в прямую сумму двух ненулевых F-инвариантных подпространств, и *неразложимым* — в противном случае. Все простые операторы неразложимы.

Упражнение 10.2. Покажите, что оператор умножения на класс $[t]$ в фактор кольце $\Bbbk[t]/(t^n)$ при всех $n > 1$ приводим, но неразложим.

Таким образом, над любым полем $\Bbbk$ имеются неразложимые пространства с оператором любой размерности. Очевидно, что всякое пространство с оператором является прямой суммой неразложимых.

10.1.1. Пространство с оператором как $\Bbbk[t]$-модуль. Задание на пространстве V линейного оператора $F: V \to V$ эквивалентно заданию на V структуры модуля над кольцом многочленов $\Bbbk[t]$. В самом деле, структура $\Bbbk[t]$-модуля включает в себя операцию умножения векторов на переменную t: $v \mapsto tv$, которая является линейным отображением $V \to V$. Если обозначить его буквой F, то умножение векторов на произвольный многочлен $f(t) = a_0 + a_1 t + \cdots + a_m t^m$ происходит по правилу $f(t)\, v = a_0 v + a_1 Fv + \ldots + a_m F^m v = f(F)\, v$, где

$$f(F) = a_0 \mathrm{Id}_V + a_1 F + \ldots + a_m F^m$$

есть результат вычисления многочлена f на элементе F в $\Bbbk$-алгебре $\mathrm{End}(V)$. Наоборот, каждый линейный оператор $F: V \to V$ задаёт на V структуру $\Bbbk[t]$-модуля, в котором умножение вектора $v \in V$ на многочлен $f(t) \in \Bbbk[t]$ происходит по формуле $f(t)\, v \stackrel{\text{def}}{=} f(F)\, v$. Мы будем обозначать такой $\Bbbk[t]$-модуль через V_F.

[1] произвольная диаграмма отображений называется *коммутативной*, если композиции отображений вдоль любых двух путей с общим началом и концом одинаковы

Гомоморфизм $\Bbbk[t]$-модулей $C : V_F \to W_G$, построенных по операторам $F : V \to V$ и $G : W \to W$ — это линейное отображение $C : V \to W$, перестановочное с умножением векторов на t, т. е. такое что $C \circ F = F \circ C$. Поэтому операторы F и G изоморфны тогда и только тогда, когда изоморфны $\Bbbk[t]$-модули V_F и W_G.

Векторное подпространство $U \subset V$ является $\Bbbk[t]$-подмодулем в модуле V_F если и только если оператор умножения на t переводит U в себя, т. е. тогда и только тогда, когда это подпространство F-инвариантно. Аналогично, разложимость V в прямую сумму инвариантных подпространств означает разложимость $\Bbbk[t]$-модуля V_F в прямую сумму $\Bbbk[t]$-подмодулей.

Если векторное пространство V конечномерно над $\Bbbk$, то $\Bbbk[t]$-модуль V_F является конечно порождённым модулем кручения. В самом деле, любой базис пространства V над $\Bbbk$ линейно порождает модуль V_F над $\Bbbk[t]$, и в каноническом разложении модуля V_F в прямую сумму свободного модуля и модуля кручения[1] свободное слагаемое отсутствует, поскольку оно бесконечномерно над $\Bbbk$. Из теоремы об элементарных делителях[2] вытекает

Теорема 10.1
Любой линейный оператор в конечномерном векторном пространстве над произвольным полем $\Bbbk$ подобен оператору умножения на класс $[t]$ в прямой сумме фактор колец

$$\frac{\Bbbk[t]}{(p_1^{m_1}(t))} \oplus \cdots \oplus \frac{\Bbbk[t]}{(p_k^{m_k}(t))}, \tag{10-1}$$

где все многочлены $p_\nu(t) \in \Bbbk[t]$ приведены и неприводимы, и слагаемые могут повторяться. Операторы умножения на класс $[t]$, действующие в суммах

$$\frac{\Bbbk[t]}{(p_i^{m_1}(t))} \oplus \cdots \oplus \frac{\Bbbk[t]}{(p_k^{m_k}(t))} \quad \text{и} \quad \frac{\Bbbk[t]}{(q_i^{n_1}(t))} \oplus \cdots \oplus \frac{\Bbbk[t]}{(q_\ell^{n_\ell}(t))}$$

изоморфны если и только если $k = \ell$ и прямые слагаемые можно переставить так, чтобы $p_\nu = q_\nu$ и $m_\nu = n_\nu$ при всех ν. □

Определение 10.1 (элементарные делители линейного оператора)
Дизъюнктное объединение[3] всех многочленов $p_\nu^{m_\nu}$, стоящих в правой части разложения (10-1), называется *набором элементарных делителей* оператора $F : V \to V$ и обозначается через $\mathcal{El}(F)$.

Следствие 10.1
Линейные операторы F и G подобны тогда и только тогда, когда $\mathcal{El}(F) = \mathcal{El}(G)$. □

Следствие 10.2
Линейный оператор неразложим тогда и только тогда, когда он подобен оператору умножения на класс $[t]$ в фактор кольце $\Bbbk[t]/(p^m)$, где $p \in \Bbbk[t]$ неприводим и приведён. Неразложимый оператор неприводим если и только если $m = 1$. □

Следствие 10.3
Многочлен $f \in \Bbbk[t]$ тогда и только тогда аннулирует оператор $F : V \to V$, когда он делится на все элементарные делители оператора F. □

[1] См. сл. 9.2 на стр. 128.

[2] См. теор. 9.3 на стр. 127.

[3] Каждый элементарный делитель p^m входит в него ровно столько раз, сколько прямых слагаемых вида $\Bbbk[t]/(p^m)$ имеется в разложении (10-1).

Упражнение 10.3. Пусть пространство с оператором (F, V) разлагается в прямую сумму F-инвариантных подпространств U_i. Покажите, что $\mathcal{E}\ell(F) = \bigsqcup_i \mathcal{E}\ell(F|_{U_i})$.

10.1.2. Характеристический многочлен. Пусть оператор $F : V \to V$ имеет в некотором базисе $\boldsymbol{v}$ пространства V матрицу $F_{\boldsymbol{v}}$. Характеристический многочлен $\det(tE - F_{\boldsymbol{v}})$ этой матрицы не меняется при переходе к любому другому базису $\boldsymbol{w} = \boldsymbol{v}\,C$, поскольку[1] $F_{\boldsymbol{w}} = C^{-1}F_{\boldsymbol{v}}C$ и

$$\det(tE - F_{\boldsymbol{w}}) = \det(tC^{-1}EC - C^{-1}F_{\boldsymbol{v}}C) = \det\bigl(C^{-1}(tE - F_{\boldsymbol{v}})C\bigr) = \\ = \det C^{-1} \cdot \det(tE - F_{\boldsymbol{v}}) \cdot \det C = \det(tE - F_{\boldsymbol{v}})\,.$$

Многочлен $\chi_F(t) \stackrel{\text{def}}{=} \det(tE - F_{\boldsymbol{v}})$ называется *характеристическим многочленом* оператора F. Предыдущее вычисление показывает, что подобные операторы имеют равные характеристические многочлены.

Упражнение 10.4. Пусть пространство с оператором (F, W) распадается в прямую сумму пространств с операторами (G, U) и (H, V). Убедитесь, что $\chi_F(t) = \chi_G(t) \cdot \chi_H(t)$ в $\Bbbk[t]$.

Упражнение 10.5. Убедитесь, что для любого приведённого многочлена $f \in \Bbbk[t]$ характеристический многочлен оператора умножения на класс $[t]$ в фактор кольце $\Bbbk[t]/(f)$ равен f.

Из этих упражнений и теор. 10.1 мы получаем

Предложение 10.1
Характеристический многочлен равен произведению всех элементарных делителей. □

Упражнение 10.6. Выведите из предл. 10.1 новое доказательство теоремы Гамильтона – Кэли.

10.1.3. Минимальный многочлен. Для каждого неприводимого приведённого многочлена $p \in \Bbbk[t]$ обозначим через $m_p(F)$ максимальный показатель m, с которым p^m присутствует в наборе $\mathcal{E}\ell(F)$ элементарных делителей оператора F, а для тех неприводимых приведённых многочленов $p \in \Bbbk[x]$, степени которых не представлены в $\mathcal{E}\ell\,F$, положим $m_p(F) = 0$. Таким образом, $m_p(F) = 0$ для всех неприводимых приведённых $p \in \Bbbk[x]$ кроме конечного числа. Из теор. 10.1 вытекает, что приведённый многочлен $\mu_F(t)$ наименьшей возможной степени, аннулирующий оператор F, равен

$$\mu_F(t) = \prod_p p^{m_p(F)}\,,$$

где произведение берётся по всем приведённым неприводимым $p \in \Bbbk[t]$. Многочлен $\mu_F(t)$ называется *минимальным многочленом* оператора F. Напомню, что минимальный многочлен порождает ядро гомоморфизма вычисления[2] многочленов на операторе F

$$\mathrm{ev}_F : \Bbbk[t] \to \mathrm{End}_{\Bbbk}(V)\,, \quad f(t) \mapsto f(F)\,,$$

и делит в $\Bbbk[t]$ все многочлены, аннулирующие оператор F, включая и характеристический многочлен $\chi_F(t) = \det(t\,\mathrm{Id}_V - F)$.

Пример 10.1 (операторы над алгебраически замкнутым полем)
Если основное поле $\Bbbk$ алгебраически замкнуто, то неприводимые приведённые многочлены в $\Bbbk[t]$ исчерпываются линейными двучленами $(t - \lambda)$, $\lambda \in \Bbbk$. Оператор умножения на класс

[1] См. прим. 7.3 на стр. 99.
[2] См. n° 7.2.3 на стр. 102.

$[t] = [\lambda] + [t-\lambda]$ в фактор кольце $\Bbbk[t]/\big((t-\lambda)^m\big)$ является суммой скалярного оператора $\lambda\,\mathrm{Id}: [g] \mapsto \lambda[g]$, умножающего все векторы на λ, и оператора умножения на класс $(t-\lambda)$, который действует на состоящий из векторов $e_i = \big[(t-\lambda)^{m-i}\big]$, $1 \leqslant i \leqslant m$, базис пространства $\Bbbk[t]/\big((t-\lambda)^m\big)$ по правилу

$$0 \leftarrow e_1 \leftarrow e_2 \leftarrow e_3 \leftarrow \cdots \leftarrow e_{m-1} \leftarrow e_m\,. \tag{10-2}$$

Таким образом, умножение на класс $[t]$ задаётся в базисе $e_1, \ldots, e_n$ матрицей

$$J_m(\lambda) \stackrel{\text{def}}{=} \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}, \tag{10-3}$$

которая называется *жордановой клеткой* размера m с *собственным числом* λ. По теор. 10.1 каждый линейный оператор F над алгебраически замкнутым полем подобен оператору умножения на класс $[t]$ в прямой сумме фактор колец вида $\Bbbk[t]/\big((t-\lambda)^m\big)$, и два таких оператора подобны если и только если прямые суммы отличаются друг от друга перестановкой слагаемых. При этом характеристический многочлен оператора F равен произведению всех $(t-\lambda)^m$, встречающихся в прямой сумме, а минимальный многочлен оператора F равен произведению максимальных для данного $\lambda \in \Bbbk$ степеней $(t-\lambda)^m$, взятому по всем различным λ, встречающимся в прямой сумме. Таким образом, характеристический и минимальный многочлены имеют одинаковый набор корней. Он обозначается $\operatorname{Spec} F$ и называется *спектром* оператора F, а сами корни $\lambda \in \operatorname{Spec} F$ называются *собственными числами* или *собственными значениями* оператора F. Кратность корня $\lambda \in \operatorname{Spec} F$ в минимальном многочлене $\mu_F(t)$ равна максимальному такому m, что $(t-\lambda)^m \in \mathcal{E}\ell(F)$, а кратность корня $\lambda \in \operatorname{Spec} F$ в характеристическом многочлене $\chi_F(t)$ равна сумме всех таких m, что $(t-\lambda)^m \in \mathcal{E}\ell(F)$.

На языке матриц сказанное означает, что любая квадратная матрица A над алгебраически замкнутым полем $\Bbbk$ сопряжена блочно диагональной матрице, по главной диагонали которой располагаются жордановы клетки (10-3), причём эта блочно диагональная матрица однозначно с точностью до перестановки клеток определяется матрицей A. Она называется *жордановой нормальной формой* матрицы A. Две матрицы сопряжены если и только если у них одинаковые с точностью до перестановки клеток жордановы нормальные формы. Числа λ, встречающиеся в клетках жордановой нормальной формы матрицы A суть корни характеристического многочлена $\chi_A(t) = \det(tE - A)$, и кратность каждого корня λ равна сумме размеров всех жордановых клеток с собственным числом λ. Минимальный многочлен $\mu_A = \prod_{\lambda \in \operatorname{Spec} A} (t-\lambda)^{m_\lambda}$ равен взятому по всем корням λ характеристического многочлена матрицы A одночленов $(t-\lambda)$ в степенях, равных максимальным размерам жордановых клеток с собственным числом λ.

Упражнение 10.7. Как действует умножение на класс $[t]$ в фактор кольце $\Bbbk[t]/(t-\lambda)$ и в прямой сумме конечного множества таких фактор колец?

10.1.4. Отыскание элементарных делителей. Зафиксируем в пространстве V какой-нибудь базис $\boldsymbol{v} = (v_1, \ldots, v_n)$ над полем $\Bbbk$ и обозначим через $F_{\boldsymbol{v}} \in \operatorname{Mat}_n(\Bbbk)$ матрицу оператора $F: V \to V$ в этом базисе. Поскольку векторы v_i линейно порождают пространство V над $\Bbbk$, они тем более порождают модуль V_F над $\Bbbk[t]$, и $V_F = \Bbbk[t]^n/R_{\boldsymbol{v}}$, где подмодуль $R_{\boldsymbol{v}} = \ker \pi_{\boldsymbol{v}} \subset \Bbbk[t]^n$

является ядром эпиморфизма[1] $\pi_{\boldsymbol{v}}\colon \Bbbk[t]^n \to V_F$, переводящего стандартный базисный вектор $e_i \in \Bbbk[t]^n$ в вектор $v_i \in V$, и состоит из всех $\Bbbk[t]$-линейных соотношений между векторами $\boldsymbol{v}$ в V_F. Таким образом, множество $\mathcal{El}(F)$ элементарных делителей оператора F представляет собою множество элементарных делителей, ассоциированное с набором инвариантных множителей подмодуля соотношений $R_{\boldsymbol{v}}$ в свободном координатном модуле $\Bbbk[t]^n$.

Лемма 10.1
Если записывать элементы свободного модуля $\Bbbk[t]^n$ в виде координатных столбцов с элементами из $\Bbbk[t]$, то подмодуль соотношений $\ker \pi_{\boldsymbol{v}} \subset \Bbbk[t]^n$ линейно порождается над $\Bbbk[t]$ столбцами матрицы $tE - F_{\boldsymbol{v}}$.

Доказательство. Пусть $F_{\boldsymbol{v}} = \big(f_{ij}\big)$. Тогда j-й столбец матрицы $tE - F_{\boldsymbol{v}}$ выражается через стандартный базис $\boldsymbol{e}$ модуля $\Bbbk[t]^n$ как $te_j - \sum_{i=1}^n e_i f_{ij}$. Применяя к этому вектору гомоморфизм $\pi_{\boldsymbol{v}}$, получаем $\pi_{\boldsymbol{v}}\Big(te_j - \sum_{i=1}^n e_i f_{ij}\Big) = tv_j - \sum_{i=1}^n v_i f_{ij} = Fv_j - \sum_{i=1}^n v_i f_{ij} = 0$. Тем самым, все столбцы матрицы $tE - F_{\boldsymbol{v}}$ лежат в $\ker \pi_{\boldsymbol{v}}$. Рассмотрим теперь произвольный вектор $h \in \ker \pi_{\boldsymbol{v}} \subset \Bbbk[t]^n$ и запишем его в виде многочлена от t с коэффициентами в $\Bbbk^n$ (ср. с n° 8.3 на стр. 116):

$$h = t^m h_m + t^{m-1} h_{m-1} + \cdots + t h_1 + h_0\,, \quad \text{где} \quad h_i \in \Bbbk^n\,.$$

Этот многочлен можно поделить с остатком слева на многочлен $tE - F_{\boldsymbol{v}}$ точно также, как делят «уголком» обычные полиномы с постоянными коэффициентами[2]. В результате получим равенство вида $t^m h_m + \ldots + t h_1 + h_0 = (tE - F_{\boldsymbol{v}}) \cdot (t^{m-1} g_{m-1} + \ldots + t g_1 + g_0) + r$ с $g_i, r \in \Bbbk^n$.

Упражнение 10.8. Убедитесь в этом.

Иными словами, вычитая из столбца $h \in \Bbbk[t]^n$ подходящую $\Bbbk[t]$-линейную комбинацию столбцов матрицы $tE - F_{\boldsymbol{v}}$, можно получить вектор $r \in \Bbbk^n$, т. е. $\Bbbk$-линейную комбинацию $r = \sum \lambda_i e_i$ стандартных базисных векторов e_i модуля $\Bbbk[t]^n$. Так как столбцы матрицы $tE - F_{\boldsymbol{v}}$ лежат в ядре гомоморфизма $\pi_{\boldsymbol{v}}$, а векторы $v_i \in V$ линейно независимы над $\Bbbk$, вектор $\pi_{\boldsymbol{v}}(h) = \pi_{\boldsymbol{v}}(r) = \sum \lambda_i v_i$ обращается в нуль если и только если все $\lambda_i = 0$. Следовательно, $r = 0$ и столбец h лежит в $\Bbbk[t]$-линейной оболочке столбцов матрицы $tE - F_{\boldsymbol{v}}$. □

Следствие 10.4
Множество $\mathcal{El}(F)$ является дизъюнктным объединением степеней p^m неприводимых приведённых многочленов, встречающихся в разложениях инвариантных множителей[3]

$$f_i(t) = \Delta_i(tE - F_{\boldsymbol{v}}) / \Delta_{i-1}(tE - F_{\boldsymbol{v}})$$

матрицы $tE - F_{\boldsymbol{v}}$ на простые множители в $\Bbbk[t]$. Инвариантные множители $f_i(t) = d_{ii}$ совпадают с диагональными элементами матрицы D, которая получается в результате приведения матрицы $tE - F_{\boldsymbol{v}}$ к диагональному виду элементарными преобразованиями строк и столбцов над кольцом $\Bbbk[t]$. □

10.2. Специальные классы операторов. В этом разделе мы подробно остановимся на свойствах нескольких специальных классов операторов, играющих важную роль в различных задачах их самых разных областей математики.

[1] См. n° 9.2 на стр. 123.
[2] См. n° 3.2 на стр. 37.
[3] Напомню, что Δ_i означает нод всех $k \times k$ миноров матрицы, см. сл. 9.1 на стр. 123.

10.2.1. Нильпотентные операторы. Линейный оператор $F : V \to V$ называется *нильпотентным*, если $F^m = 0$ для некоторого $m \in \mathbb{N}$. Поскольку нильпотентный оператор аннулируется многочленом t^m, все его элементарные делители являются степенями t. В частности минимальный многочлен тоже является степенью t, и поскольку минимальный многочлен делит характеристический многочлен, степень которого равна $\dim V$, в определении нильпотентного оператора можно без ограничения общности считать, что $m \leqslant \dim V$. По теор. 10.1 нильпотентный оператор изоморфен оператору умножения на класс $[t]$ в прямой сумме фактор колец вида

$$\frac{\Bbbk[t]}{(t^{\nu_1})} \oplus \cdots \oplus \frac{\Bbbk[t]}{(t^{\nu_k})} \tag{10-4}$$

и два таких оператора изоморфны друг другу если и только если выписанные в порядке нестрого убывания наборы показателей $\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_k$ у них одинаковы. Таким образом, нильпотентные операторы над произвольным полем $\Bbbk$ взаимно однозначно соответствуют диаграммам Юнга ν. Диаграмма $\nu(F)$, характеризующая нильпотентный оператор F, называется его *цикловым типом*.

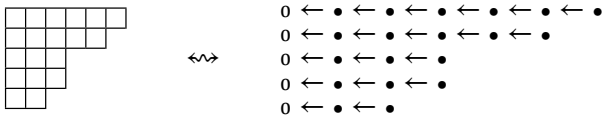
Умножение на класс $[t]$ действует на состоящий из векторов $e_i = [t^{m-i}]$ базис пространства $\Bbbk[t]/(t^m)$ по правилу[1]

$$0 \leftarrow\!\shortmid e_1 \leftarrow\!\shortmid e_2 \leftarrow\!\shortmid e_3 \leftarrow\!\shortmid \cdots \leftarrow\!\shortmid e_{m-1} \leftarrow\!\shortmid e_m$$

и задаётся в этом базисе матрицей

$$J_m(0) \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

которая называется *нильпотентной жордановой клеткой* размера m. Тем самым, для нильпотентного оператора F циклового типа $\nu(F)$ в пространстве V имеется базис, векторы которого размещаются по клеткам диаграммы $\nu(F)$ так, что F переводит каждый из них в левый соседний, а все векторы самого левого столбца — в нуль:



(10-5)

Базис такого вида называется *циклическим* или *жордановым* базисом нильпотентного оператора F, а наборы базисных векторов, стоящие по строкам диаграммы, называются *жордановыми цепочками*. Так как сумма длин первых m столбцов диаграммы $\nu(F)$ равна $\dim \ker F^m$, длина m-того столбца диаграммы $\nu(F)$ равна $\dim \ker F^m - \dim \ker F^{m-1}$.

Упражнение 10.9. В условиях прим. 10.1 на стр. 133 покажите, что для отыскания жордановой нормальной формы оператора F над алгебраически замкнутым полем достаточно разложить характеристический многочлен $\chi_F(t)$ на линейные множители:

$$\chi_F(t) = \prod_{\lambda \in \operatorname{Spec} F} (t - \lambda)^{m_\lambda}$$

[1] См. формулу (10-2) на стр. 134.

и для каждого $\lambda \in \operatorname{Spec} F$ и натурального k в пределах $1 \leqslant k \leqslant m_\lambda$, где m_λ — кратность корня λ, вычислить[1] $\dim \ker(\lambda \operatorname{Id} - F)^k$, после чего построить диаграмму Юнга ν, в которой k-й столбец имеет длину $\dim \ker(\lambda \operatorname{Id} - F)^k - \dim \ker(\lambda \operatorname{Id} - F)^{k-1}$. Количество жордановых клеток размера m с заданным собственным значением λ в жордановой нормальной форме оператора F равно количеству строк длины m в диаграмме Юнга ν.

10.2.2. Полупростые операторы. Прямая сумма простых[2] пространств с операторами называется *полупростым* или *вполне приводимым* пространством с оператором.

Предложение 10.2
Следующие свойства оператора $F : V \to V$ эквивалентны друг другу:

1) V является прямой суммой неприводимых F-инвариантных подпространств

2) V линейно порождается неприводимыми F-инвариантными подпространствами

3) для каждого ненулевого F-инвариантного подпространства $U \subsetneq V$ существует такое F-инвариантное подпространство $W \subset V$, что $V = U \oplus W$

4) оператор F подобен умножению на класс $[t]$ в прямой сумме фактор колец

$$\Bbbk[t]/(p_1) \oplus \Bbbk[t]/(p_2) \oplus \cdots \oplus \Bbbk[t]/(p_r),$$

где $p_i \in \Bbbk[t]$ приведены и неприводимы[3] (но не обязательно различны).

Доказательство. Импликация (1) ⇒ (2) очевидна. Покажем, что (2) ⇒ (3). Индукция по $\dim V$. При $\dim V = 1$ доказывать нечего. Пусть $\dim V > 1$. Для каждого неприводимого F-инвариантного подпространства $L \subset V$ пересечение $L \cap U$, будучи F-инвариантным подпространством в L, либо нулевое, либо совпадает с L. Если все неприводимые инвариантные подпространства $L \subset V$ лежат в U, то $U = V$ в силу (2), и доказывать нечего. Если есть ненулевое неприводимое F-инвариантное подпространство $L \subset V$ с $L \cap U = 0$, рассмотрим фактор $V' = V/L$ и проекцию $\pi : V \twoheadrightarrow V'$ с ядром L. Она инъективно отображает подпространство $U \subset V$ на ненулевое F-инвариантное подпространство $\pi(U) \subset V'$. Поскольку $\dim V' < \dim V$, по индукции найдётся такое F-инвариантное подпространство $W' \subset V'$, что $V' = W' \oplus \pi(U)$ (при $\pi(U) = V$ мы полагаем $W' = 0$). Пусть $W = \pi^{-1}(W') \subset V$. Проверим, что $V = U + W$. Проекция любого $v \in V$ на V' представляется в виде $\pi(v) = \pi(u) + w'$ с $u \in U$, $w' \in W'$, и разность $w = v - u \in W$, поскольку $\pi(w) = \pi(v) - \pi(u) = w' \in W'$. Тем самым, $v = w + u$ с $w \in W$, $u \in U$. Если вектор $v \in U \cap W$, то $\pi(v) \in \pi(U) \cap W' = 0$, откуда $v \in \ker \pi = L$. Так как $L \cap U = 0$, мы заключаем, что $U \cap W = 0$ и $V = W \oplus U$.

Чтобы доказать импликацию (3) ⇒ (4), покажем сначала, что если свойство (3) выполнено для пространства V, то оно выполнено и для каждого F-инвариантного подпространства $H \subset V$. Рассмотрим любое инвариантное подпространство $U \subset H$ и отыщем в V такие инвариантные

[1] Причём это вычисление достаточно продолжать только до тех пор, пока $\dim \ker(\lambda \operatorname{Id} - F)^k$ строго увеличивается с ростом k. Если при очередном k размерность останется такой же, как при предыдущем k, то она будет оставаться такой и для всех последующих k.

[2] Или — в другой терминологии — неприводимых, см. начало n° 10.1 на стр. 131.

[3] Иными словами, в прямой сумме (10-1) из теор. 10.1 все показатели степеней $m_i = 1$.

подпространства Q и R, что $V = H \oplus Q = U \oplus Q \oplus R$. Рассмотрим проекцию $\pi : V \twoheadrightarrow H$ с ядром Q и положим $W = \pi(R)$.

Упражнение 10.10. Проверьте, что $H = U \oplus W$.

Итак, если свойство (3) выполнено для прямой суммы фактор колец (10-1) из теор. 10.1, то оно выполнено и для каждого слагаемого этой суммы. Однако по сл. 10.2 при $m > 1$ пространство $\Bbbk[t]/(p^m)$ приводимо, но неразложимо.

Импликация (4) ⇒ (1) также немедленно вытекает из сл. 10.2. □

Следствие 10.5 (из доказательства предл. 10.2)
Ограничение полупростого оператора на инвариантное подпространство также является полупростым оператором.

10.2.3. Циклические векторы. Вектор $v \in V$ называется *циклическим вектором* линейного оператора $F : V \to V$, если его F-орбита v, Fv, F^2v, F^3v, , … линейно порождает пространство V над полем $\Bbbk$. Иначе можно сказать, что v порождает модуль V_F над $\Bbbk[t]$.

Предложение 10.3
Следующие свойства оператора $F : V \to V$ эквивалентны друг другу:

1) F обладает циклическим вектором

2) F подобен умножению на класс $[t]$ в фактор кольце $\Bbbk[t]/(f)$, где $f \in \Bbbk[t]$ — какой-либо приведённый многочлен

3) каждый неприводимый $p \in \Bbbk[t]$ встречается в $\mathcal{E}\ell\, F$ не более одного раза

4) минимальный многочлен оператора F совпадает с характеристическим.

Доказательство. Условия (3) и (4) эквивалентны в силу предл. 10.1 и означают, что оператор F подобен умножению на t в прямой сумме фактор колец

$$\Bbbk[t]/(p_1^{m_1}) \oplus \Bbbk[t]/(p_2^{m_2}) \oplus \cdots \oplus \Bbbk[t]/(p_r^{m_r}),$$

в которой все неприводимые приведённые многочлены $p_1, \dots, p_r$ попарно различны. По китайской теореме об остатках, эта сумма изоморфна $\Bbbk[t]/(f)$, где

$$f = \chi_F = \mu_F = \prod_{i=1}^{r} p_i^{m_i}.$$

Тем самым, (2) равносильно (3) и (4). Импликация (2) ⇒ (1) очевидна: в качестве циклического вектора для оператора умножения на t в фактор кольце $\Bbbk[t]/(f)$ можно взять $v = [1]$. Наоборот, если модуль V_F порождается над $\Bbbk[t]$ одним вектором v, то $V_F = \Bbbk[t]/R$, где $R = \ker \pi$ — ядро $\Bbbk[t]$-линейного эпиморфизма $\Bbbk[t] \twoheadrightarrow V_F$, преводящего 1 в v. Поскольку $\Bbbk[t]$ — кольцо главных идеалов, модмодуль $R \subset \Bbbk[t]$ имеет вид (f), где f-приведённый многочлен наименьшей степени со свойством $f(F)v = 0$. Тем самым, $V = \Bbbk[t]/(f)$. □

10.2.4. Собственные подпространства и собственные числа. Максимальное по включению ненулевое подпространство в V, на котором оператор $F\colon V \to V$ действует как умножение на скаляр $\lambda \in \Bbbk$, называется *собственным подпространством* оператора F с *собственным числом* или *собственным значением* λ и обозначается

$$V_\lambda \stackrel{\text{def}}{=} \{v \in V \mid F(v) = \lambda v\} = \ker(\lambda \operatorname{Id}_V - F)\,.$$

Ненулевые векторы $v \in V_\lambda$ называются *собственными векторами* оператора F с собственным числом[1] λ.

Предложение 10.4
Любой набор собственных векторов с попарно различными собственными числами линейно независим.

Доказательство. Пусть собственные векторы $v_1, \ldots, v_m$ имеют попарно разные собственные числа $\lambda_1, \ldots, \lambda_m$ и линейно зависимы. Рассмотрим линейное соотношение между ними, в котором задействовано минимально возможное число векторов. Пусть это будут векторы $e_1, \ldots, e_k$. Тогда $k \geqslant 2$ и $e_k = x_1 e_1 + \ldots + x_{k-1} e_{k-1}$, где все $x_i \in \Bbbk$ отличны от нуля. При этом $\lambda_k e_k = F(e_k) = \sum x_i F(e_i) = \sum x_i \lambda_i e_i$. Вычитая из этого равенства предыдущее, умноженное на λ_k, получаем более короткую линейную зависимость

$$0 = x_1(\lambda_1 - \lambda_k)\cdot e_1 + x_1(\lambda_1 - \lambda_k)\cdot e_2 + \cdots + x_{k-1}(\lambda_{k-1} - \lambda_k)\cdot e_{k-1}$$

с ненулевыми коэффициентами. □

Следствие 10.6
Сумма ненулевых собственных подпространств с попарно разными собственными числами является прямой. □

10.2.5. Спектр. Множество собственных чисел линейного оператора $F\colon V \to V$, т. е. всех таких $\lambda \in \Bbbk$, для которых существует ненулевое собственное подпространство $V_\lambda = \ker(\lambda \operatorname{Id}_V - F)$, называется *спектром*[2] оператора F в поле $\Bbbk$ и обозначается

$$\operatorname{Spec} F = \{\lambda \in \Bbbk \mid \ker(\lambda \operatorname{Id}_V - F) \neq 0\} = \{\lambda \in \Bbbk \mid \det(tE - F) = 0\}\,.$$

Поскольку $\ker(\lambda \operatorname{Id}_V - F) \neq 0$ если и только если $\det(tE - F) = 0$, спектр совпадает с множеством корней характеристического многочлена $\chi_F(t) = \det(tE - F)$ в поле $\Bbbk$. В частности, количество различных собственных чисел не превосходит $\deg \chi_F = \dim V$, что также вытекает из сл. 10.6, согласно которому

$$\sum_{\lambda \in \operatorname{Spec} F} \dim V_\lambda \leqslant \dim V\,. \tag{10-6}$$

Упражнение 10.11. Покажите, что $\operatorname{Spec} F$ содержится в множестве корней любого многочлена, аннулирующего F.

Если известен спектр F, отыскание собственных подпространств сводится к решению систем линейных однородных уравнений $(\lambda \operatorname{Id}_V - F)\, v = 0$, которые гарантированно имеют ненулевые решения при $\lambda \in \operatorname{Spec} F$. Если основное поле $\Bbbk$ алгебраически замкнуто, спектр любого оператора гарантированно не пуст, поскольку характеристический многочлен $\chi_F(t)$ обязательно имеет корень в поле $\Bbbk$.

[1] Или собственным значением.
[2] Ср. с прим. 10.1 на стр. 133.

Предложение 10.5
Над алгебраически замкнутым полем $\Bbbk$ любой оператор обладает хотя бы одним ненулевым собственным подпространством. □

Упражнение 10.12. Покажите, что над алгебраически замкнутым полем $\Bbbk$ оператор F нильпотентен если и только если когда $\operatorname{Spec} F = \{0\}$, и приведите пример оператора, для которого неравенство (10-6) строгое.

10.2.6. Диагонализуемые операторы. Оператор $F : V \to V$ называется *диагонализуемым*, если в V имеется базис, в котором F записывается диагональной матрицей. Такой базис состоит из собственных векторов оператора F, а элементы диагональной матрицы суть собственные числа F, причём каждое собственное число $\lambda \in \operatorname{Spec} F$ встречается на диагонали ровно столько раз, какова кратность корня $t = \lambda$ в характеристическом многочлене $\chi_F(t)$ и какова размерность собственного подпространства V_λ. Иначе можно сказать, что диагонализуемый оператор F подобен оператору умножения на класс $[t]$ в прямой сумме фактор колец[1] $\Bbbk[t]/(t-\lambda) \simeq \Bbbk$, где λ пробегает $\operatorname{Spec} F$, и каждое такое прямое слагаемое представлено в сумме ровно $\dim V_\lambda$ раз.

Предложение 10.6
Следующие свойства линейного оператора $F : V \to V$ эквивалентны:

1) F диагонализуем

2) пространство V линейно порождается собственными векторами оператора F

3) характеристический многочлен $\chi_F(t) = \det(tE - F)$ полностью раскладывается в $\Bbbk[t]$ на линейные множители, и кратность каждого его корня λ равна размерности собственного подпространства V_λ

4) все элементарные делители F имеют вид $(t - \lambda)$, $\lambda \in \Bbbk$

5) оператор F аннулируется многочленом f, раскладывающимся в $\Bbbk[t]$ в произведение попарно различных линейных множителей.

Доказательство. Эквивалентности (2) $\iff$ (1) $\iff$ (4) и импликация (1) $\Rightarrow$ (3) очевидны из предваряющего предл. 10.6 обсуждения. Эквивалентность (4) $\iff$ (5) следует из сл. 10.3. Из (3) вытекает, что $\sum \dim V_\lambda = \deg \chi_F = \dim V$. Поэтому прямая по сл. 10.6 сумма всех различных собственных подпространств V_λ совпадает с V, что даёт импликацию (3) $\Rightarrow$ (1). □

Следствие 10.7
Если оператор $F : V \to V$ диагонализуем, то его ограничение на любое инвариантное подпространство тоже диагонализуемо на этом подпространстве.

Доказательство. Это вытекает из свойства (5) предл. 10.6. □

Упражнение 10.13. Убедитесь, что над алгебраически замкнутым полем диагонализуемость равносильна полупростоте.

[1] Ср. с упр. 10.7 на стр. 134.

10.2.7. Перестановочные операторы. Если линейные операторы $F, G : V \to V$ на векторном пространстве V над произвольным полем $\Bbbk$ коммутируют друг с другом, то ядро и образ любого многочлена от оператора F переводятся оператором G в себя, поскольку

$$\begin{aligned} f(F)\,v = 0 \quad &\Rightarrow \quad f(F)\,G\,v = G\,f(F)\,v = 0 \\ v = f(F)\,w \quad &\Rightarrow \quad Gv = G\,f(F)\,w = f(F)\,Gw\,. \end{aligned}$$

В частности, все собственные подпространства $V_\lambda = \ker(F - \lambda E)$ инвариантны относительно любого перестановочного с F оператора G.

Предложение 10.7
В конечномерном векторном пространстве V над алгебраически замкнутым полем $\Bbbk$ любое множество коммутирующих друг с другом операторов обладает общим для всех операторов собственным вектором. Над произвольным полем $\Bbbk$ любое множество коммутирующих друг с другом диагонализуемых операторов на V можно одновременно диагонализовать в одном общем для всех операторов базисе.

Доказательство. Индукция по $\dim V$. Если все операторы скалярны (что так при $\dim V = 1$), то доказывать нечего — подойдут, соответственно, любой ненулевой вектор и любой базис. Если среди операторов есть хоть один нескалярный оператор F, то над замкнутым полем у него есть собственное подпространство строго меньшей размерности, чем V, а в диагонализуемом случае V является прямой суммой таких собственных подпространств. Каждое собственное подпространство оператора F инвариантно для всех операторов, причём если операторы диагонализуемы на всём пространстве, то их ограничения на собственные подпространства оператора F останутся диагонализуемы по сл. 10.7. Применяя к собственным подпространствам оператора F предположение индукции, получаем требуемое. □

Пример 10.2 (конечные группы операторов)
Если m линейных операторов на конечномерном пространстве V над алгебраически замкнутым полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk > m$ образуют группу G, то каждый из этих операторов аннулируется многочленом $t^m - 1$, который раскладывается в произведение m попарно различных линейных множителей[1]. Поэтому каждый оператор в группе G диагонализуем. Все операторы из группы G одновременно диагонализуются в одном общем базисе если и только если группа G абелева.

10.2.8. Аннулирующие многочлены. Если задан многочлен $f \in \Bbbk[x]$, аннулирующий линейный оператор[2] $F : V \to V$, и известно, как f раскладывается в $\Bbbk[t]$ на простые множители, то в силу сл. 10.3 это оставляет лишь конечное число возможностей для набора элементарных делителей $\mathcal{El}(F)$ и часто позволяет явно описать разложение V в прямую сумму F-инвариантных подпространств во внутренних терминах действия F на пространстве V.

Пример 10.3 (инволюции)
Линейный оператор $\sigma : V \to V$ называется *инволюцией*, если он удовлетворяет соотношению $\sigma^2 = \mathrm{Id}_V$, т. е. аннулируется многочленом $t^2 - 1$. Тождественная инволюция $\sigma = \mathrm{Id}_V$ называется

[1] Поскольку производная mt^{m-1} многочлена $t^m - 1$ отлична от нуля и взаимно проста с этим многочленом, она не имеет с ним общих корней. Следовательно, у многочлена нет кратных корней.

[2] В силу тождества Гамильтона–Кэли по крайней мере один такой многочлен, а именно — характеристический многочлен $\chi_F(t) = \det(tE - F)$, всегда можно явно предъявить.

тривиальной. Так как $t^2 - 1 = (t+1)(t-1) = 0$ является произведением различных линейных множителей, все инволюции диагонализуемы, причём спектр любой инволюции исчерпывается числами ± 1. Пространство V с инволюцией σ распадается в прямую сумму собственных подпространств $V = V_+ \oplus V_-$ с собственными значениями ± 1, и любой вектор $v \in V$ однозначно представим в виде $v = v_+ + v_-$, где $v_+ = (v + Fv)/2 \in V_+ = \ker(\sigma - \mathrm{Id}_V) = \mathrm{im}(\sigma + \mathrm{Id}_V)$ и $v_- = (v - Fv)/2 \in V_- = \ker(\sigma + \mathrm{Id}_V) = \mathrm{im}(\sigma - \mathrm{Id}_V)$.

Теорема 10.2 (теорема о разложении)
Пусть линейный оператор $F : V \to V$ на произвольном[1] векторном пространстве V над любым полем $\Bbbk$ аннулируется многочленом $q \in \Bbbk[t]$, который раскладывается в $\Bbbk[t]$ в произведение $q = q_1 \cdot q_2 \cdot \dots \cdot q_r$ попарно взаимно простых многочленов $q_i \in \Bbbk[t]$. Положим $Q_j = q/q_j$. Тогда $\ker q_j(F) = \mathrm{im}\, Q_j(F)$ для каждого j, все эти подпространства F-инвариантны, и пространство V является прямой суммой тех из них, что отличны от нуля.

Доказательство. Так как $q(F) = q_i(F) \circ Q_j(F) = 0$, имеем включение $\mathrm{im}\, Q_i(F) \subset \ker q_i(F)$. Поэтому достаточно показать, что V линейно порождается образами операторов $Q_i(F)$, а сумма ядер $\ker q_i(F)$ прямая[2], т. е. $\ker q_i(F) \cap \sum_{j \neq i} \ker q_j(F) = 0$ для всех i. Первое вытекает из того, что $\mathrm{нод}(Q_1, \dots, Q_r) = 1$, а значит, существуют такие $h_1, \dots, h_r \in \Bbbk[t]$, что $1 = \sum Q_j(t) h_j(t)$. Подставляя в это равенство $t = F$ и применяя обе части к произвольному вектору $v \in V$, получаем разложение $v = Ev = \sum Q_j(F) h_j(F) v \in \sum \mathrm{im}\, Q_j(F)$. Второе вытекает из взаимной простоты q_i и Q_i, в силу которой существуют такие $g, h \in \Bbbk[t]$, что $1 = g(t) \cdot q_i(t) + h(t) \cdot Q_i(t)$. Подставим сюда $t = F$ и применим обе части полученного равенства $E = g(F)\, q_i(F) + h(F) \circ Q_i(F)$ к произвольному вектору $v \in \ker q_i(F) \cap \sum_{j \neq i} \ker q_j$. Так как $\ker q_j(F) \subset \ker Q_i(F)$ при всех $j \neq i$, получим $v = Ev = g(F)\, q_i(F)\, v + h(F)\, Q_i(F)\, v = 0$, что и требовалось. □

Пример 10.4 (проекторы)
Линейный оператор $\pi : V \to V$ называется *идемпотентом* или *проектором*, если он аннулируется многочленом $t^2 - t = t(t-1)$, т. е. удовлетворяет соотношению $\pi^2 = \pi$. По теор. 10.2 образ любого идемпотента $\pi : V \to V$ совпадает с подпространством его неподвижных векторов: $\mathrm{im}\, \pi = \ker(\pi - \mathrm{Id}_V) = \{v \mid \pi(v) = v\}$, и всё пространство распадается в прямую сумму $V = \ker \pi \oplus \mathrm{im}\, \pi$. Тем самым, оператор π проектирует V на $\mathrm{im}\, \pi$ вдоль $\ker \pi$. Отметим, что оператор $\mathrm{Id}_V - \pi$ тоже является идемпотентом и проектирует V на $\ker \pi$ вдоль $\mathrm{im}\, \pi$. Таким образом, задание прямого разложения $V = U \oplus W$ равносильно заданию пары идемпотентных эндоморфизмов $\pi_1 = \pi_1^2$ и $\pi_2 = \pi_2^2$ пространства V, связанных соотношениями $\pi_1 + \pi_2 = 1$ и $\pi_1 \pi_2 = \pi_2 \pi_1 = 0$.

Упражнение 10.14. Выведите из этих соотношений, что $\ker \pi_1 = \mathrm{im}\, \pi_2$ и $\mathrm{im}\, \pi_1 = \ker \pi_2$.

Предложение 10.8
Над полем вещественных чисел $\mathbb{R}$ любой оператор обладает одномерным или двумерным инвариантным подпространством.

Доказательство. Пусть $\chi_F = q_1 \dots q_m$, где $q_i \in \mathbb{R}[t]$ — неприводимые приведённые линейные или квадратичные многочлены, не обязательно различные. Применим нулевой оператор $0 =$

[1] Возможно даже бесконечномерном.
[2] См. предл. 6.1 на стр. 81.

$q_1(F)\circ q_2(F)\circ\dots\circ q_m(F)$ к какому-нибудь ненулевому вектору $v\in V$. Тогда при некотором $i\geqslant 0$ мы получим такой ненулевой вектор $w = q_{i+1}(F)\circ\dots\circ q_m(F)\,v$, что $q_i(F)\,w = 0$. Если $q_i(t) = t-\lambda$ линеен, то $F(w) = \lambda w$, и мы имеем 1-мерное F-инвариантное подпространство $\Bbbk\cdot w$. Если $q_i(t) = t^2-\alpha t-\beta$ квадратичен, то $F(Fw) = \alpha F(w)+\beta w$ лежит в линейной оболочке векторов w и Fw, которая тем самым является F-инвариантным подпространством, и её размерность не превышает 2. □

10.3. Корневое разложение и функции от операторов. Всюду в этом разделе мы предполагаем, что линейный оператор $F: V\to V$ аннулируется многочленом, который полностью разлагается над полем $\Bbbk$ на линейные множители. Для этого необходимо и достаточно, чтобы полностью разлагался на линейные множители минимальный или характеристический многочлен оператора F. Спектр такого оператора F исчерпывается степенями линейных двучленов $(t-\lambda)^m$ с $\lambda\in\operatorname{Spec}F$ и произвольными $m\in\mathbb{N}$, причём как числа λ, так и числа m могут повторяться. Подмодуль $(t-\lambda)$-кручения в $\Bbbk[t]$-модуле V_F называется *корневым подпространством* оператора F, отвечающим собственному числу $\lambda\in\operatorname{Spec}F$, и обозначается

$$K_\lambda = \{v\in V\mid \exists\, m\in\mathbb{N}: (\lambda\operatorname{Id}-F)^m v = 0\} = \bigcup_{m\geqslant 1}\ker(\lambda\operatorname{Id}-F)^m = \ker(\lambda\operatorname{Id}-F)^{m_\lambda}, \tag{10-7}$$

где m_λ — максимальный из показателей степеней элементарных делителей оператора F вида $(t-\lambda)^m$. Каждое корневое подпространство K_λ содержит ненулевое собственное подпространство V_λ и тем самым отлично от нуля. Разложение $\Bbbk[t]$-модуля V_F в прямую сумму $\Bbbk[t]$-подмодулей $(t-\lambda)$-кручения из сл. 9.3 на стр. 128 имеет вид $V = \bigoplus\limits_{\lambda\in\operatorname{Spec}F} K_\lambda$ и называется *корневым разложением* оператора F.

Следствие 10.8 (теорема о корневом разложении)
Пусть характеристический многочлен $\chi_F(t)$ линейного оператора $F: V\to V$ на конечномерном векторном пространстве V над полем $\Bbbk$ полностью разлагается в $\Bbbk[t]$ на линейные множители: $\chi_F(t) = \prod_{\lambda\in\operatorname{Spec}F}(t-\lambda)^{m_\lambda}$. Тогда $V = \bigoplus_{\lambda\in\operatorname{Spec}F} K_\lambda$, причём $K_\lambda = \ker(\lambda\operatorname{Id}-F)^{m_\lambda}$ для всех $\lambda\in\operatorname{Spec}F$. □

Упражнение 10.15. Выведите существование корневого разложения из теор. 10.2 и тождества Гамильтона – Кэли без использования сл. 9.3 и теоремы об элементарных делителях.

10.3.1. Функции от операторов. Пусть линейный оператор F действует на конечномерном векторном пространстве V над полем $\mathbb{R}$ или $\mathbb{C}$, которое мы обозначим через $\mathbb{K}$. Всюду далее мы предполагаем, что F аннулируется многочленом $\alpha(t)\in\mathbb{K}[t]$, который полностью разлагается над $\mathbb{K}$ на линейные множители, т. е.

$$\alpha(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\cdots(t-\lambda_s)^{m_s}, \tag{10-8}$$

где $\lambda_i\neq\lambda_j$ при $i\neq j$ и все $m_i\in\mathbb{N}$. В этом случае характеристический и минимальный многочлены оператора F тоже полностью разлагаются на линейные множители в $\mathbb{K}[t]$, и можно взять в качестве $\alpha(t)$ один из них. Мы полагаем $m = \deg\alpha = m_1+\dots+m_s$. Алгебра $\mathcal{A}$, состоящая из функций $U\to\mathbb{K}$, заданных на каком-нибудь подмножестве $U\subset\mathbb{K}$, содержащем все корни многочлена (10-8), называется *алгебраически вычислимой* на операторе F, если $\mathbb{K}[t]\subset\mathcal{A}$ и для каждого корня λ кратности k многочлена (10-8) все функции $f\in\mathcal{A}$ определены в точке $\lambda\in\mathbb{K}$

вместе с первыми $k-1$ производными $f^{(\nu)} = \frac{d^\nu f}{dt^\nu}$ и допускают разложение вида

$$f(t) = f(\lambda) + \frac{f'(\lambda)}{1!}(t-\lambda) + \cdots + \frac{f^{(k-1)}(\lambda)}{(k-1)!}(t-\lambda)^{k-1} + g_\lambda(t)\cdot(t-\lambda)^k, \tag{10-9}$$

где функция $g_\lambda(t)$ тоже лежит в алгебре $\mathcal{A}$.

Например, алгебра $\mathcal{A}$ всех функций, определённых в ε-окрестности каждого собственного числа $\lambda \in \operatorname{Spec} F$ и представимых в ней суммой абсолютно сходящегося степенного ряда от $(t-\lambda)$, алгебраически вычислима на операторе F. Подалгебра в $\mathcal{A}$, состоящая из всех аналитических функций[1] $\mathbb{K} \to \mathbb{K}$, алгебраически вычислима на всех линейных операторах $F \in \operatorname{End}(V)$, характеристические многочлены которых полностью разлагаются на линейные множители в $\mathbb{K}[t]$.

Теорема 10.3
В сделанных выше предположениях каждая алгебраически вычислимая на операторе $F: V \to V$ алгебра функций $\mathcal{A}$ допускает единственный такой гомоморфизм $\mathbb{K}$-алгебр $\operatorname{ev}_F: \mathcal{A} \to \operatorname{End} V$, что $\operatorname{ev}_F(p) = p(F)$ для всех многочленов $p \in \mathbb{K}[t] \subset \mathcal{A}$.

Определение 10.2 (гомоморфизм вычисления)
Гомоморфизм $\operatorname{ev}_F: \mathcal{A} \to \operatorname{End} V$ из теор. 10.3 называется *вычислением* функций $f \in \mathcal{A}$ на операторе F. Линейный оператор $\operatorname{ev}_F(f): V \to V$, в который переходит функция $f \in \mathcal{A}$ при гомоморфизме вычисления, обозначается $f(F)$ и называется *функцией f от оператора F*.

Замечание 10.1. (как относиться к функциям от операторов) Из теор. 10.3 вытекает, что если характеристический многочлен линейного оператора $F: V \to V$ полностью разлагается на линейные множители в $\mathbb{K}[t]$, то на пространстве V определены такие линейные операторы, как e^F или $\sin F$, а если $F \in \operatorname{GL}(V)$, то и такие задаваемые аналитическими вне нуля функциями операторы, как $\ln F$ или $\sqrt{F}$, причём алгебраические свойства всех этих операторов точно такие же, как у числовых функций e^t, $\sin t$, $\ln t$ и $\sqrt{t}$. В частности, все эти функции от оператора F коммутируют друг с другом и с F, а также удовлетворяют соотношениям вроде $\ln F^2 = 2\ln F$ и $\sqrt{F}\sqrt{F} = F$. Таким образом, функции от операторов можно использовать для отыскания операторов с предписанными свойствами, например, для извлечения корней из невырожденных операторов.

Доказательство теор. 10.3. Пусть оператор F аннулируется многочленом $\alpha(t) = \prod_\lambda (t-\lambda)^{m_\lambda}$, где $\lambda = \lambda_1, \ldots, \lambda_r$ пробегает все различные корни этого многочлена, и пусть искомый гомоморфизм $\operatorname{ev}_F: \mathcal{A} \to \mathbb{K}$ существует. По теореме о разложении[2] пространство V является прямой суммой F-инвариантных подпространств $K_\lambda = \ker(F-\lambda\mathrm{Id})^{m_\lambda}$, и согласно формуле (10-9) оператор

$$f(F) = f(\lambda)\cdot E + f'(\lambda)\cdot(F-\lambda E) + \cdots + \frac{f^{(m_\lambda-1)}(\lambda)}{(m_\lambda-1)!}(F-\lambda E)^{m_\lambda-1} + g_\lambda(F)(F-\lambda E)^{m_\lambda} \tag{10-10}$$

действует на каждом подпространстве K_λ точно так же, как результат подстановки оператора F в многочлен

$$j_\lambda^{m_\lambda-1} f(t) \stackrel{\text{def}}{=} f(\lambda) + f'(\lambda)\cdot(t-\lambda) + \cdots + f^{(m_\lambda-1)}(\lambda)\cdot(t-\lambda)^{m_\lambda-1}/(m_\lambda-1)!,$$

[1] Т. е. функций, задаваемых сходящимися всюду в $\mathbb{K}$ степенными рядами.
[2] См. теор. 10.2 на стр. 142.

класс которого в фактор кольце $\mathbb{K}[t]/\left((t-\lambda)^{m_\lambda}\right)$ называется $(m_\lambda-1)$-*струёй* функции $f\in\mathcal{A}$ в точке $\lambda\in\mathbb{K}$. По китайской теореме об остатках существует единственный такой многочлен $p_{f(F)}(t)\in\mathbb{K}[t]$ степени меньшей $\deg\alpha(t)$, что

$$p_{f(F)}(t)\equiv j_\lambda^{m_\lambda-1}f(t)\,(\mathrm{mod}\,\alpha(t))$$

для всех корней λ многочлена α. Поскольку операторы $p_{f(F)}(F)$ и $f(F)$ одинаково действуют на каждом подпространстве K_λ, мы имеем равенство $f(F)=p_{f(F)}(F)$. Таким образом гомоморфизм вычисления единствен. Остаётся убедиться, что отображение $f\mapsto p_{f(F)}(F)$ действительно является гомоморфизмом $\mathbb{K}$-алгебр. Проверим сначала, что отображение

$$\begin{gathered}J:\mathcal{A}\to\frac{\mathbb{K}[t]}{\left((t-\lambda_1)^{m_1}\right)}\times\cdots\times\frac{\mathbb{K}[t]}{\left((t-\lambda_r)^{m_r}\right)}\simeq\frac{\mathbb{K}[t]}{(\alpha)}\\ f\mapsto\left(j_{\lambda_1}^{m_1-1}f,\,\ldots\,,j_{\lambda_s}^{m_r-1}f\right),\end{gathered}\tag{10-11}$$

сопоставляющее функции $f\in\mathcal{A}$ набор её струй[1] во всех корнях многочлена α, является гомоморфизмом $\mathbb{K}$-алгебр, т. е. $\mathbb{K}$-линейно и удовлетворяет равенству $J(fg)=J(f)J(g)$. Первое очевидно, второе достаточно установить для каждой струи j_λ^{m-1} отдельно. Используя правило Лейбница: $(fg)^{(k)}=\sum_{\nu=0}^k\binom{k}{\nu}f^{(\nu)}g^{(k-\nu)}$, получаем следующие равенства по модулю $(t-\lambda)^m$:

$$\begin{aligned}j_\lambda^{m-1}(fg)&=\sum_{k=0}^{m-1}\frac{(t-\lambda)^k}{k!}\sum_{\nu+\mu=k}\frac{k!}{\nu!\mu!}f^{(\nu)}(\lambda)g^{(\mu)}(\lambda)=\\&=\sum_{k=0}^{m-1}\sum_{\nu+\mu=k}\frac{f^{(\nu)}(\lambda)}{\nu!}(t-\lambda)^\nu\cdot\frac{g^{(\mu)}(\lambda)}{\mu!}(t-\lambda)^\mu\equiv j_\lambda^{m-1}(f)j_\lambda^{m-1}(g)\,.\end{aligned}$$

Отображение $f\mapsto P_{f(F)}(F)$ является композицией гомоморфизма (10-11) с гомоморфизмом вычисления многочленов $\mathrm{ev}_F:\mathbb{K}[t]\to\operatorname{End}V$, $p\mapsto p(F)$, который корректно пропускается через фактор $\mathbb{K}[t]/(\alpha)$, так как $\alpha(F)=0$. □

Определение 10.3 (интерполяционный многочлен)
Многочлен $p_{f(F)}(t)\in\mathbb{K}[t]$, принимающий на операторе F то же самое значение, что и функция $f\in\mathcal{A}$, называется *интерполяционным многочленом* для вычисления $f(F)$. Он однозначно определяется тем, что в каждом корне λ кратности m аннулирующего оператор f многочлена α многочлен $p_{f(F)}(t)$ и первые его $m-1$ производные принимают те же значения, что и функция f и её $m-1$ производные, т. е. многочлен $p_{f(F)}(t)$ решает интерполяционную задачу с кратными узлами из прим. 6.13 на стр. 92. Если $\deg\alpha=n$, отыскание коэффициентов интерполяционного многочлена $p_{f(F)}$ сводится к решению системы из n линейных уравнений на n неизвестных.

Пример 10.5 (степенная функция и рекуррентные уравнения)
Задача отыскания n-того члена a_n числовой последовательности $z:\mathbb{Z}\to\mathbb{K}$, $n\mapsto z_n$, решающей рекуррентное уравнение $z_n=\alpha_1z_{n-1}+\alpha_2z_{n-2}+\cdots+\alpha_mz_{n-m}$ с начальным условием

[1] Мы рассматриваем этот набор как элемент прямого произведения соответствующих колец вычетов, которое по китайской теореме об остатках изоморфно фактор кольцу $\mathbb{K}[t]/(\alpha)$.

$(z_0, \dots, z_{n-1}) = (a_0, \dots, a_{n-1}) \in \mathbb{K}^n$, сводится вычислению n-той степени *матрицы сдвига*

$$S = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_m \\ 1 & 0 & \ddots & \vdots & \alpha_{m-1} \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & \alpha_2 \\ 0 & \cdots & 0 & 1 & \alpha_1 \end{pmatrix}$$

смещающей каждый фрагмент из m последовательных элементов на один шаг вправо:

$$(z_{k+1},\, z_{k+2},\, \dots\, ,\, z_{k+m}) \cdot S = (z_{k+2},\, z_{k+3},\, \dots\, ,\, z_{k+m+1}),$$

так что член a_n оказывается равным первой координате вектора

$$(a_n,\, a_{n+1},\, \dots\, ,\, a_{n+m-1}) = (a_0, a_1, \dots, a_{m-1}) \cdot S^n .$$

Матрица $S^n = p_{S^n}(S)$ является результатом подстановки матрицы S в интерполяционный многочлен $p_{S^n}(t) \in \mathbb{K}[t]$ для вычисления на матрице S *степенной функции* $f(t) = t^n$. Обратите внимание, что $\deg p_{S^n} < m$, и коэффициенты многочлена p_{S^n} находятся решением системы из m линейных уравнений на m неизвестных.

Например, для *уравнения Фиббоначчи* $a_n = a_{n-1} + a_{n-2}$ матрица сдвига

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} .$$

Интерполяционный многочлен для вычисления степенной функции t^n на этой матрице линеен. Записывая его в виде $p_{S^n}(t) = at + b$ с неопределёнными коэффициентами a и b, получаем

$$S^n = a\,S + b\,E = \begin{pmatrix} b & a \\ a & a+b \end{pmatrix} .$$

В частности, n-тое *число Фиббоначчи*, решающее уравнение Фиббоначчи с начальным условием $(a_0, a_1) = (0, 1)$, равно первой координате вектора $\left(a_n, a_{n+1}\right) = (0,1) \cdot S^n = (a, a+b)$. Матрица S аннулируется своим характеристическим многочленом

$$\chi_S(t) = t^2 - t\ \operatorname{tr} S + \det S = t^2 - t - 1 = (t - \lambda_+)(t - \lambda_-)$$

с однократными корнями $\lambda_\pm = (1 \pm \sqrt{5})/2$. Функция t^n принимает на них значения $\lambda_\pm^n$. Коэффициенты a и b находятся из системы

$$\begin{cases} a\,\lambda_+ + b = \lambda_+^n \\ a\,\lambda_- + b = \lambda_-^n , \end{cases}$$

и по правилу Крамера первый из них $a = \left(\lambda_+^n - \lambda_-^n\right) / \left(\lambda_+ - \lambda_-\right)$. Тем самым,

$$a_n = a = \frac{\left((1+\sqrt{5})/2\right)^n - \left((1-\sqrt{5})/2\right)^n}{\sqrt{5}} .$$

Замечание 10.2. Имеются различные аналитические способы продолжения гомоморфизма вычисления многочленов на матрице $F \in \mathrm{Mat}_n(\mathbb{C})$ с алгебры $\mathbb{C}[z]$ на большие алгебры функций $\mathcal{C} \supset \mathbb{C}[z]$. А именно, пространства $\mathbb{C}[z]$ и $\mathrm{Mat}_n(\mathbb{C})$ наделяются той или иной топологией, и функция $f \in \mathcal{C}$ представляется в виде предела $f = \lim_{n\to\infty} f_n$ какой-нибудь последовательности многочленов (f_n). Матрица $f(F)$ полагается равной пределу последовательности матриц $f_n(F) \in \mathrm{Mat}_n(\mathbb{C})$. Разумеется, при этом необходимо проверять, что предел $\lim_{n\to\infty} f_n(F) \in \mathrm{Mat}_n(\mathbb{C})$ существует и зависит только от функции f, а не от выбора сходящейся к f последовательности многочленов (f_n). Отдельно необходимо проверить, что возникающее таким образом отображение $\mathrm{ev}_F : \mathcal{C} \to \mathrm{Mat}_n(\mathbb{C})$, $f \mapsto f(F)$, является гомоморфизмом алгебр[1]. Однако, как бы ни определялась сходимость в пространстве функций и какой бы ни была сходящаяся к функции f последовательность многочленов (f_n), последовательность матриц $f_n(F)$ всегда лежит в конечномерном векторном пространстве, линейно порождённом над $\mathbb{C}$ степенями F^m с $0 \leqslant m < \dim n$, и если переход к пределу в пространстве матриц перестановочен со сложением и умножением на константы[2], то предел последовательности матриц $(f_n(F))$ неминуемо является *многочленом* от F степени, строго меньшей n. Это означает, что какая бы аналитическая процедура не применялась для построения гомоморфизма $\mathrm{ev}_F : \mathcal{C} \to \mathrm{Mat}_n(\mathbb{C})$, значение этого гомоморфизма на заданной функции $f \in \mathcal{C}$ *a priori* вычисляется по указанному нами рецепту. Отметим также, что если матрицы F и G подобны, т. е. $G = CFC^{-1}$ для некоторой матрицы $C \in \mathrm{GL}_n(\mathbb{C})$, то и аналитически определённые функции от них подобны: поскольку равенство $f_n(G) = Cf_n(F)C^{-1}$ выполнено для всех многочленов, приближающих функцию f, оно останется выполненным и для предельной функции: $f(G) = Cf(F)C^{-1}$, при условии, что топология на пространстве $\mathrm{Mat}_n(\mathbb{C})$ такова, что все $\mathbb{C}$-линейные отображения $\mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ непрерывны.

10.4. Разложение Жордана. Этот раздел является уточнением прим. 10.1 на стр. 133. Всюду далее речь идёт об операторах на конечномерном векторном пространстве V над алгебраически замкнутым полем $\Bbbk$.

Теорема 10.4 (разложение Жордана)
Для каждого оператора F на конечномерном векторном пространстве V над алгебраически замкнутым полем $\Bbbk$ существует единственная пара таких операторов F_d и F_n, что F_n нильпотентен, F_d диагонализуем, $F_dF_n = F_nF_d$ и $F = F_d + F_n$. Кроме того, операторы F_d и F_n являются многочленами от оператора F с нулевыми свободными членами.

Доказательство. Пусть $\operatorname{Spec} F = \{\lambda_1, \dots, \lambda_r\}$. В силу алгебраической замкнутости поля $\Bbbk$, характеристический многочлен оператора F полностью разлагается на линейные множители: $\chi_F(t) = \prod_i (t - \lambda_i)^{m_i}$, а пространство V является прямой суммой корневых подпространств: $V = \bigoplus_i K_i$, где $K_i = \ker(F - \lambda_i \mathrm{Id})^{m_i}$. В качестве диагонализуемого оператора F_d можно взять оператор, действующий на каждом корневом подпространстве K_λ умножением на λ, а в качестве нильпотентного оператора F_n взять разность $F_n = F - F_d$, которая действует на каждом

[1] В качестве упражнения по анализу читателю настоятельно рекомендуется попробовать самостоятельно реализовать намеченную программу, используя на пространстве функций топологию, в которой сходимость последовательности функций означает равномерную сходимость в каждом круге в $\mathbb{C}$, а на пространстве $\mathrm{Mat}_n(\mathbb{C})$ — стандартную топологию пространства $\mathbb{C}^{n^2}$, где сходимость определяется покоординатно.

[2] Т. е. $\lim\limits_{n\to\infty}(\lambda F_n + \mu G_n) = \lambda \lim\limits_{n\to\infty} F_n + \mu \lim\limits_{n\to\infty} G_n$.

корневом подпространстве K_λ нильпотентным оператором $F - \lambda\,\mathrm{Id}$. Покажем, что оба эти оператора являются многочленами без свободного члена от F. Для этого достаточно представить в таком виде оператор F_d.

Так как многочлены $(t-\lambda_i)^{m_i}$ попарно взаимно просты, по китайской теореме об остатках существуют такие многочлены $f_1, \dots, f_r \in \Bbbk[t]$, что

$$f_i(t) \equiv \begin{cases} 1 \mod \big(t-\lambda_i\big)^{m_i} \\ 0 \mod \big(t-\lambda_j\big)^{m_j} \text{ при } j \neq i\,. \end{cases}$$

Если $\lambda_i \neq 0$, то многочлен t обратим по модулю $(t-\lambda_i)^{m_i}$. Поэтому найдётся такой многочлен $g_i(t)$, что $t \cdot g_i(t) \equiv \lambda_i \mod \big(t-\lambda_i\big)^{m_i}$. Если $\lambda_i = 0$, то положим $g_i(t) = 0$. Тогда при каждом i многочлен $p_s(t) \stackrel{\text{def}}{=} t\sum\limits_{j=1}^{r} g_j(t)f_j(t) \equiv \lambda_i \mod \big(t-\lambda_i\big)^{m_i}$ и не имеет свободного члена. Из этих сравнений вытекает, что оператор $F_d \stackrel{\text{def}}{=} p_s(F)$ действует на каждом корневом подпространстве $K_i = \ker(F-\lambda_i\mathrm{Id})^{m_i}$ как умножение на λ_i и, стало быть, равен F_d. Будучи многочленами от F, операторы F_d и $F_n = F - F_d$ перестановочны между собою и с F. Это доказывает существование операторов F_d и F_n с требуемыми свойствами, включая последнее утверждение предложения.

Докажем их единственность. Пусть есть ещё одно разложение $F = F_s' + F_n'$, в котором F_d' диагонализуем, F_n' нильпотентен и $F_d'F_n' = F_n'F_d'$. Из последнего равенства вытекает, что F_d' и F_n' перестановочны с любым многочленом от $F = F_s' + F_n'$ и, в частности, с построенными выше F_d и F_n. Поэтому каждое собственное подпространство V_λ оператора F_d переводится оператором F_d' в себя[1], причём F_d' диагонализуем[2] на каждом V_λ. Если бы оператор F_d' имел на V_λ собственный вектор с собственным значением $\mu \neq \lambda$, то этот вектор был бы собственным для оператора $F_n - F_n' = F_d - F_d'$ с собственным значением $\lambda - \mu \neq 0$, что невозможно, так как оператор $F_n - F_n'$ нильпотентен.

Упражнение 10.16. Докажите, что разность двух перестановочных нильпотентных операторов нильпотентна.

Следовательно, оператор F_s' действует на каждом собственном подпространстве V_λ оператора F_d как умножение на λ, откуда $F_d' = F_d$. Тогда и $F_n' = F - F_s' = F - F_d = F_n$. □

Определение 10.4
Операторы F_d и F_n из теор. 10.4 называются, соответственно, *диагонализуемой* и *нильпотентной* составляющими оператора F.

Замечание 10.3. Поскольку операторы F_d и F_n являются многочленами от F, каждое F-инвариантное подпространство $U \subset V$ является инвариантным для F_d и F_n.

Предложение 10.9
В условиях теор. 10.3 на стр. 144 для любой функции f из алгебраически вычислимой на операторе F алгебры функций $\mathcal{A}$ спектр оператора $f(F)$ состоит из чисел $f(\lambda)$, где $\lambda \in \operatorname{Spec} F$. Если $f'(\lambda) \neq 0$, то элементарные делители $(t-\lambda)^m \in \mathcal{El}(F)$ биективно соответствуют элементарным делителям $\big(t-f(\lambda)\big)^m \in \mathcal{El}\big(f(F)\big)$. Если $f'(\lambda) = 0$, то элементарные делители вида $(t-\lambda)^m \in \mathcal{El}(F)$, имеющие $m > 1$, распадаются в объединения элементарных делителей $\big(t-f(\lambda)\big)^\ell \in \mathcal{El}\big(f(F)\big)$, имеющих $\ell < m$.

[1] См. n° 10.2.7 на стр. 141.
[2] См. сл. 10.7 на стр. 140.

Доказательство. Реализуем F как оператор умножения на класс $[t]$ в прямой сумме фактор колец

$$V = \frac{\mathbb{C}[t]}{\left((t-\lambda_1)^{s_1}\right)} \oplus \cdots \oplus \frac{\mathbb{C}[t]}{\left((t-\lambda_r)^{s_r}\right)}.$$

Из доказательства теор. 10.3 вытекает, что диагональная и нильпотентная составляющие ограничения оператора $f(F)$ на корневое подпространство K_λ суть $f_s(F) = f(\lambda)\cdot \mathrm{Id}$ и

$$f_n(F) = f'(\lambda)\cdot\eta + \frac{1}{2}f''(\lambda)\cdot\eta^2 + \cdots,$$

где η обозначает нильпотентный оператор умножения на класс $(t-\lambda)$. На каждом слагаемом $\mathbb{C}[t]/\left((t-\lambda)^k\right)$ оператор η имеет ровно одну жорданову цепочку максимальной длины k. Если $f'(\lambda)\neq 0$, то $f_n^{k-1}(F) = f'(\lambda)^{k-1}\cdot\eta^{k-1}\neq 0$. Поэтому $f_n(F)$ тоже имеет ровно одну жорданову цепочку длины k. При $f'(l)=0$ и $m>1$ равенство $f_n^m(F)=0$ наступит при $m<k$. Поэтому цикловой тип ограничения $f_n(F)$ на $\mathbb{C}[t]/\left((t-\lambda)^k\right)$ состоит из нескольких цепочек длины $<k$. □

Упражнение 10.17. Покажите, что матрица $J_n^{-1}(\lambda)$, обратная к жордановой клетке размера $n\times n$ с собственным числом λ, подобна матрице $J_n(\lambda^{-1})$.

§11. Группы

11.1. Группы, подгруппы, циклы. Множество G называется *группой*, если на нём задана операция *композиции* $G \times G \to G$, $(g_1, g_2) \mapsto g_1 g_2$ со свойствами

$$\text{ассоциативность}: \quad \forall\, f, g, h \in G \quad (fg)h = f(gh) \tag{11-1}$$

$$\text{наличие единицы}: \quad \exists\, e \in G : \forall\, g \in G \quad eg = g \tag{11-2}$$

$$\text{наличие обратных}: \quad \forall\, g \in G \quad \exists\, g^{-1} \in G : g^{-1} g = e \tag{11-3}$$

Группа называется *коммутативной* или *абелевой*, если дополнительно имеет место

$$\text{коммутативность}: \quad \forall\, f, g \in G \quad fg = gf\,. \tag{11-4}$$

Левый обратный к g элемент g^{-1} из (11-3) является также и правым обратным, т. е. $gg^{-1} = e$, что устанавливается умножением правой и левой части в $g^{-1}gg^{-1} = eg^{-1} = g^{-1}$ слева на левый обратный к g^{-1} элемент.

Упражнение 11.1. Убедитесь, что обратный к g элемент g^{-1} однозначно определяется элементом g и что $\bigl(g_1 \cdots g_k\bigr)^{-1} = g_k^{-1} \cdots g_1^{-1}$.

Для единицы e из (11-2) при любом $g \in G$ выполнятся также и равенство $ge = g$, поскольку $ge = g(g^{-1}g) = (gg^{-1})g = eg = g$.

Упражнение 11.2. Убедитесь, что единичный элемент $e \in G$ единствен.

Если группа G конечна, число элементов в ней обозначается $|G|$ и называется *порядком* группы G. Подмножество $H \subset G$ называется *подгруппой*, если оно образует группу относительно имеющейся в G композиции. Для этого достаточно, чтобы вместе с каждым элементом $h \in H$ в H лежал и обратный к нему элемент h^{-1}, а вместе с каждой парой элементов $h_1, h_2 \in H$ — их произведение $h_1 h_2$. Единичный элемент $e \in G$ автоматически окажется в H, т. к. $e = hh^{-1}$ для произвольного $h \in H$.

Упражнение 11.3. Проверьте, что пересечение любого множества подгрупп является подгруппой.

Пример 11.1 (группы преобразований)
Модельными примерами групп являются *группы преобразований*, обсуждавшиеся нами в n° 1.6. Все взаимно однозначные отображения произвольного множества X в себя очевидно образуют группу. Она обозначается $\operatorname{Aut} X$ и называется *группой автоморфизмов* множества X. Подгруппы $G \subset \operatorname{Aut} X$ называются *группами преобразований* множества X. Для $g \in G$ и $x \in X$ мы часто будем сокращать обозначение $g(x)$ до gx. Группа всех автоморфизмов n-элементного множества $X = \{1, \ldots, n\}$ называется n-той *симметрической группой* и обозначается S_n. Порядок $|S_n| = n!$. Чётные перестановки образуют в S_n подгруппу, обозначаемую A_n и часто называемую *знакопеременной группой*. Порядок $|A_n| = n!/2$.

11.1.1. Циклические группы и подгруппы. Наименьшая по включению подгруппа в G, содержащая заданный элемент $g \in G$, состоит из всевозможных целых степеней g^m элемента g, где мы, как обычно, полагаем $g^0 \stackrel{\text{def}}{=} e$ и $g^{-n} \stackrel{\text{def}}{=} \bigl(g^{-1}\bigr)^n$. Она называется *циклической подгруппой*, порождённой g, и обозначается $\langle g \rangle$. Группа $\langle g \rangle$ абелева и является образом сюръективного гомоморфизма абелевых групп $\varphi_g : \mathbb{Z} \twoheadrightarrow \langle g \rangle$, $m \mapsto g^m$, который переводит сложение в композицию. Если $\ker \varphi_g \neq 0$, то $\ker \varphi_g = (n)$ и $\langle g \rangle \simeq \mathbb{Z}/(n)$, где $n \in \mathbb{N}$ — наименьшая степень, для которой $g^n = e$. Она называется *порядком* элемента g и обозначается $\operatorname{ord}(g)$. В этом случае

группа $\langle g\rangle$ имеет порядок[1] $n = \operatorname{ord} g$ и состоит из элементов $e = g^0$, $g = g^1$, g^2, … , g^{n-1}. Если $\ker \varphi_g = 0$, то $\varphi_g : \mathbb{Z} \xrightarrow{\sim} \langle g\rangle$ является изоморфизмом и все степени g^m попарно различны. В этом случае говорят, что g имеет *бесконечный порядок* и пишут $\operatorname{ord} g = \infty$.

Напомним[2], что группа G называется *циклической*, если в ней существует элемент $g \in G$ такой, что все элементы группы являются его целыми степенями, т. е. $G = \langle g\rangle$. Элемент g называется в этом случае *образующей* циклической группы G. Например, аддитивная группа целых чисел $\mathbb{Z}$ является циклической, и в качестве образующего элемента можно взять любой из двух элементов ± 1. В предл. 3.10 на стр. 50 мы видели, что всякая конечная подгруппа в мультипликативной группе любого поля является циклической. Аддитивная группа вычетов $\mathbb{Z}/(10)$ также является циклической, и в качестве её образующего элемента можно взять любой из четырёх классов[3] $[\pm 1]_6$, $[\pm 3]_6$.

Упражнение 11.4. Укажите необходимые и достаточные условия для того, чтобы конечно порождённая абелева группа[4] $G = \mathbb{Z}^r \oplus \mathbb{Z}/(p_1^{n_1}) \oplus \cdots \oplus \mathbb{Z}/(p_\alpha^{n_\alpha})$ была циклической.

Лемма 11.1
Элемент $h = g^k$ тогда и только тогда является образующей циклической группы $\langle g\rangle$ порядка n, когда $\text{нод}(k, n) = 1$.

Доказательство. Так как $\langle h\rangle \subset \langle g\rangle$, равенство $\langle h\rangle = \langle g\rangle$ равносильно неравенству $\operatorname{ord} h \geqslant n$. Но $h^m = g^{mk} = e$ если и только если $mk \vdots n$. При $\text{нод}(n, k) = 1$ такое возможно только когда $m|n$, и в этом случае $\operatorname{ord} h \geqslant n$. Если же $n = n_1 d$ и $k = k_1 d$, где $d > 1$, то $h^{n_1} = g^{kn_1} = g^{nk_1} = e$ и $\operatorname{ord} h \leqslant n_1 < n$. □

11.1.2. Разложение перестановок в композиции циклов. Перестановка $\tau \in S_n$ по кругу переводящая друг в друга какие-нибудь m различных элементов[5]

$$i_1 \mapsto i_2 \mapsto \cdots \mapsto i_{m-1} \mapsto i_m \mapsto i_1 \tag{11-5}$$

и оставляющая на месте все остальные элементы, называется *циклом* длины m.

Упражнение 11.5. Покажите, что k-тая степень цикла длины m является циклом тогда и только тогда, когда $\text{нод}(k, m) = 1$.

Цикл (11-5) часто бывает удобно обозначать $\tau = \big|i_1, \dots, i_m\rangle$, не смотря на то, что один и тот же цикл (11-5) допускает m *различных* таких записей, получающихся друг из друга циклическими перестановками элементов.

Упражнение 11.6. Сколько имеется в S_n различных циклов длины k?

Теорема 11.1
Каждая перестановка $g \in S_n$ является композицией $g = \tau_1 \cdots \tau_k$ непересекающихся перестановочных циклов $\tau_i\tau_j = \tau_j\tau_i$, и такое разложение единственно с точностью до перестановки циклов.

[1] Таким образом, порядок элемента равен порядку порождённой им циклической подгруппы.
[2] См. n° 3.5.1 на стр. 49.
[3] Обратите внимание, что остальные 6 классов не являются образующими.
[4] См. теор. 9.4 на стр. 129.
[5] Числа $i_1, \dots, i_m$ могут быть любыми, не обязательно соседними или возрастающими.

Доказательство. Поскольку множество $X = \{1, 2, \dots, n\}$ конечно, в последовательности

$$x \overset{g}{\longmapsto} g(x) \overset{g}{\longmapsto} g^2(x) \overset{g}{\longmapsto} g^3(x) \overset{g}{\longmapsto} \cdots, \tag{11-6}$$

возникающей при применении g к произвольной точке $x \in X$, случится повтор. Так как преобразование $g : X \xrightarrow{\sim} X$ биективно, первым повторившимся элементом будет стартовый элемент x. Таким образом, каждая точка $x \in X$ под действием g движется по циклу. В силу биективности g два таких цикла, проходящие через различные точки x и y, либо не пересекаются, либо совпадают. Таким образом, перестановка g является произведением непересекающихся циклов, очевидно, перестановочных друг с другом. □

Упражнение 11.7. Покажите, что два цикла $\tau_1, \tau_2 \in S_n$ перестановочны ровно в двух случаях: либо когда они не пересекаются, либо когда $\tau_2 = \tau_1^s$ и оба цикла имеют равную длину, взаимно простую с s.

Определение 11.1 (цикловой тип перестановки)
Написанный в порядке нестрогого убывания набор длин непересекающихся циклов[1], в которые раскладывается перестановка $g \in S_n$, называется *цикловым типом* перестановки g и обозначается $\lambda(g)$.

Цикловой тип перестановки $g \in S_n$ удобно изображать n-клеточной диаграммой Юнга, а сами циклы записывать по строкам этой диаграммы. Например, перестановка

$$g = (6, 5, 4, 1, 8, 3, 9, 2, 7) = |1, 6, 3, 4\rangle\,|2, 5, 8\rangle\,|7, 9\rangle = \begin{array}{|c|c|c|c|}\hline 1&6&3&4\\\hline 2&5&8\\\cline{1-3} 7&9\\\cline{1-2}\end{array}$$

имеет цикловой тип $\begin{array}{|c|c|c|c|}\hline &&&\\\hline &&&\\\cline{1-3} &&\\\cline{1-2}\end{array}$, т. е. $\lambda(6, 5, 4, 1, 8, 3, 9, 2, 7) = (4, 3, 2)$. Единственной перестановкой циклового типа $\lambda = (1, \dots, 1)$ (один столбец высоты n) является тождественная перестановка Id. Диаграмму $\lambda = (n)$ (одна строка длины n) имеют $(n-1)!$ циклов максимальной длины n.

Упражнение 11.8. Сколько перестановок в симметрической группе S_n имеют заданный цикловой тип, содержащий для каждого $i = 1, \dots, n$ ровно m_i циклов длины i?

Пример 11.2 (вычисление порядка и знака перестановки)
Порядок перестановки $g \in S_n$ равен наименьшему общему кратному длин непересекающихся циклов, из которых она состоит. Например, порядок перестановки

$$(3, 12, 7, 9, 10, 4, 11, 1, 6, 2, 8, 5) = |1, 3, 7, 11, 8\rangle\,|2, 12, 5, 10\rangle\,|4, 9, 6\rangle \in S_{12}$$

равен $5 \cdot 4 \cdot 3 = 60$. По правилу ниточек из прим. 8.2 на стр. 108 знак цикла длины ℓ равен $(-1)^{\ell-1}$. Поэтому перестановка чётна тогда и только тогда, когда у неё чётное число циклов чётной длины.

Упражнение 11.9. Найдите чётность $g = (6, 5, 4, 1, 8, 3, 9, 2, 7) \in S_9$ и вычислите g^{15}.

[1] Включая циклы длины один, отвечающие элементам, которые перестановка оставляет на месте.

11.2. Группы фигур. Для любой фигуры Φ в евклидовом[1] пространстве $\mathbb{R}^n$ биективные отображения $\Phi \to \Phi$ индуцированные ортогональными[2] линейными преобразованиями пространства $\mathbb{R}^n$, переводящими фигуру Φ в себя, образуют группу преобразований фигуры Φ. Эта группа называется *полной группой фигуры* Φ и обозначается O_Φ. Подгруппу $\mathrm{SO}_\Phi \subset \mathrm{O}_\Phi$, состоящую из биекций, индуцированных собственными[3] ортогональными операторами $\mathbb{R}^n \to \mathbb{R}^n$, мы будем называть *собственной группой фигуры* Φ. Если фигура $\Phi \subset \mathbb{R}^n$ содержится в некоторой гиперплоскости $\Pi \subset \mathbb{R}^n$, то собственная группа фигуры Φ совпадает с полной: беря композицию любого несобственного движения из группы фигуры с отражением в плоскости Π, мы получаем собственное движение, которое действует на фигуру Φ точно также, как и исходное несобственное движение.

Упражнение 11.10. Изготовьте модели пяти *платоновых тел* — тетраэдра, октаэдра, куба, додекаэдра и икосаэдра (см. рис. 11◇5 – рис. 11◇8 на стр. 156).

Пример 11.3 (группы диэдров D_n)
Группа правильного плоского n-угольника, лежащего в пространстве $\mathbb{R}^3$ так, что его центр находится в нуле, обозначается D_n и называется *n-той группой диэдра*. Простейший диэдр — *двуугольник* — возникает при $n = 2$. Его можно представлять себе как вытянутую симметричную луночку с двумя сторонами, изображённую на рис. 11◇1. Группа D_2 такой луночки совпадает с группами описанного вокруг неё прямоугольника и вписанного в неё ромба[4]. Она состоит из тождественного отображения и трёх поворотов на 180° вокруг перпендикулярных друг другу осей, одна из которых проходит через вершины луночки, другая — через середины её сторон, а третья перпендикулярна плоскости луночки и проходит её центр.

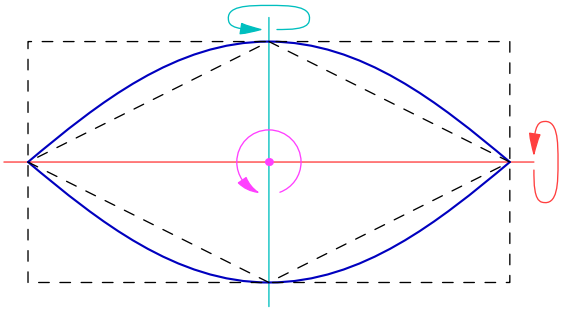


Рис. 11◇1. Двуугольник D_2.

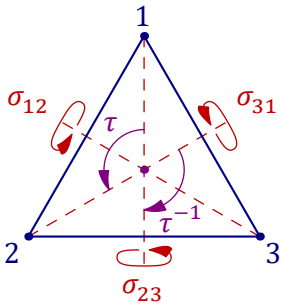


Рис. 11◇2. Группа треугольника.

Упражнение 11.11. Убедитесь, что $D_2 \simeq \mathbb{Z}/(2) \oplus \mathbb{Z}/(2)$.

Следующая диэдральная группа — *группа треугольника* D_3 — состоит из шести движений: тождественного, двух поворотов τ, τ^{-1} на $\pm 120^\circ$ вокруг центра треугольника и трёх осевых симметрий σ_{ij} относительно его медиан (см. рис. 11◇2). Так как движение плоскости однозначно

[1]Напомню, что *евклидовость* означает фиксацию в векторном пространстве $\mathbb{R}^n$ симметричного билинейного положительного скалярного произведения $V \times V \to \mathbb{R}$, обозначаемого (v, w).

[2]Линейный оператор $F\colon \mathbb{R}^n \to \mathbb{R}^n$ на евклидовом пространстве $\mathbb{R}^n$ называется *ортогональным*, если он сохраняет скалярное произведение, т. е. $\forall v, w \in \mathbb{R}^n$ $(Fv, Fw) = (v, w)$ (достаточно, чтобы это равенство выполнялось при $v = w$).

[3]Т. е. ортогональными операторами, сохраняющими ориентацию или, что то же самое, с определителем 1.

[4]Мы предполагаем, что луночка такова, что оба они не квадраты.

задаётся своим действием на вершины треугольника, группа треугольника D_3 изоморфна группе перестановок S_3 его вершин. При этом повороты на $\pm 120^\circ$ отождествляются с циклическими перестановками $(2,\ 3,\ 1)$, $(3,\ 1,\ 2)$, а осевые симметрии — с транспозициями $\sigma_{23} = (1,\ 3,\ 2)$, $\sigma_{13} = (3,\ 2,\ 1)$, $\sigma_{12} = (2,\ 1,\ 3)$. Поскольку движение плоскости, переводящее в себя правильный n-угольник, однозначно определяется своим действием на аффинный репер, образованный какой-нибудь вершиной и примыкающей к ней парой сторон, группа диэдра D_n при каждом $n \geqslant 2$ состоит из $2n$ движений: выбранную вершину можно перевести в любую из n вершин, после чего одним из двух возможных способов совместить рёбра. Эти $2n$ движений суть n поворотов вокруг центра многоугольника на углы[1] $2\pi k/n$ с $k = 0, 1, \ldots, (n-1)$ и n осевых симметрий[2] относительно прямых, проходящих при нечётном n через вершину и середину противоположной стороны, а при чётном n — через пары противоположных вершин и через середины противоположных сторон (см. рис. 11⋄3).

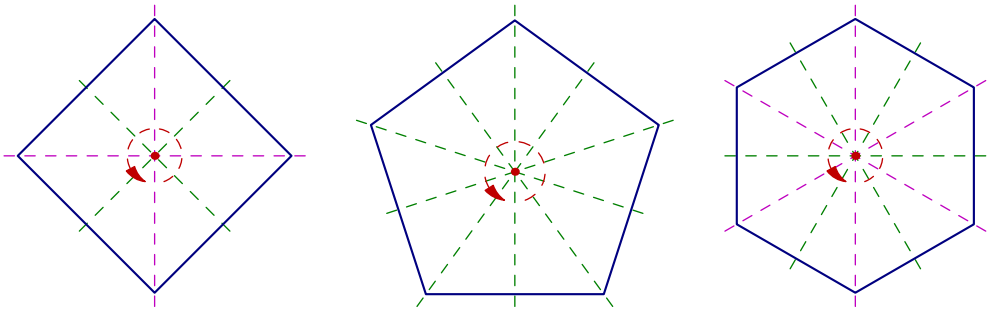


Рис. 11⋄3. Оси диэдров D_4, D_5 и D_6.

Упражнение 11.12. Составьте таблицы умножения в группах D_3, D_4 и D_5, аналогичные таблице из форм. (1-24) на стр. 14.

Пример 11.4 (группа тетраэдра)
Поскольку каждое движение трёхмерного евклидова пространства $\mathbb{R}^3$ однозначно задаётся своим действием на вершины правильного тетраэдра и это действие может быть произвольным, полная группа правильного тетраэдра с центром в нуле изоморфна группе S_4 перестановок его вершин и состоит из 24 движений. Собственная группа состоит из $12 = 4 \cdot 3$ движений: поворот тетраэдра однозначно задаётся своим действием на аффинный репер, образованный какой-нибудь вершиной и тремя выходящими из неё рёбрами, и может переводить эту вершину в любую из четырёх вершин, после чего остаются ровно три возможности для совмещения рёбер, сохраняющего ориентацию пространства. Полный список всех собственных движений тетраэдра таков (см. рис. 11⋄4): тождественное, $4 \cdot 2 = 8$ поворотов на углы $\pm 120^\circ$ вокруг прямых, проходящих через вершину и центр противоположной грани, а

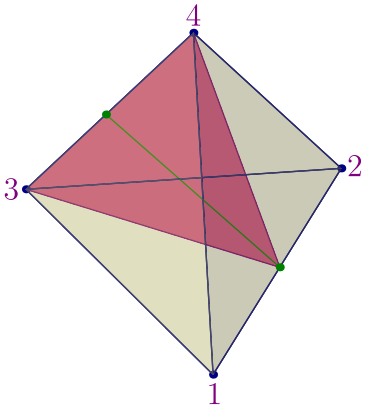


Рис. 11⋄4. Плоскость симметрии σ_{12} и ось поворота на 180°.

[1] При $k = 0$ получается тождественное преобразование.
[2] Или, что то же самое, поворотов на 180° в пространстве.

также 3 поворота на 180° вокруг прямых, проходящих через середины противоположных рёбер. В несобственной группе, помимо перечисленных поворотов, имеется 6 отражений σ_{ij} в плоскостях, проходящих через середину ребра $[i,j]$ и противоположное ребро. При изоморфизме с S_4 отражение σ_{ij} переходит в транспозицию букв i и j, повороты на $\pm120°$, представляющие собой всевозможные композиции $\sigma_{ij}\sigma_{jk}$ с попарно различными i, j, k, переходят в циклические перестановки букв i, j, k, три вращения на $\pm180°$ относительно осей, соединяющих середины противоположных рёбер, — в одновременные транспозиции непересекающихся пар букв: $\sigma_{12}\sigma_{34} = (2,\ 1\,,4,\ 3)$, $\sigma_{13}\sigma_{24} = (3,\ 4\,,1,\ 2)$, $\sigma_{14}\sigma_{23} = (4,\ 3\,,2,\ 1)$.

Упражнение 11.13. Убедитесь, что вместе с тождественным преобразованием эти три поворота образуют группу двуугольника D_2.

Оставшиеся шесть несобственных преобразований тетраэдра отвечают шести циклическим перестановкам вершин $|1234\rangle$, $|1243\rangle$, $|1324\rangle$, $|1342\rangle$, $|1423\rangle$, $|1432\rangle$ и реализуются поворотами на $\pm90°$ относительно прямых, проходящих через середины противоположных рёбер с последующим отражением в плоскости, проходящей через центр тетраэдра и перпендикулярной оси поворота.

Упражнение 11.14. Выразите эти 6 движений через отражения σ_{ij}.

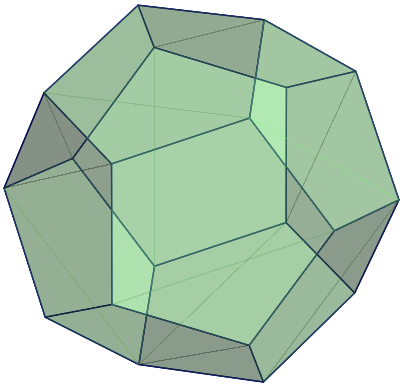


Рис. 11⋄5. Додекаэдр.

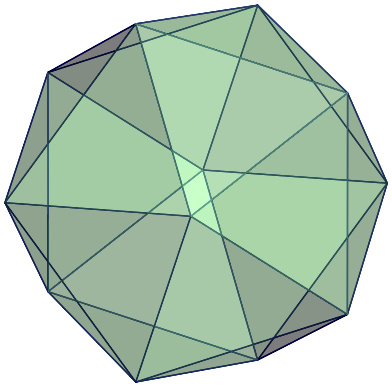


Рис. 11⋄6. Икосаэдр.

Пример 11.5 (группа додекаэдра)

Как и для тетраэдра, всякое вращение додекаэдра однозначно задаётся своим действием на аффинный репер, образованный вершиной и тремя выходящими из неё рёбрами, и может переводить эту вершину в любую из 20 вершин, а затем тремя способами совмещать рёбра с сохранением ориентации. Поэтому собственная группа додекаэдра (см. рис. 11⋄5 на стр. 155) состоит из $20 \cdot 3 = 60$ движений: $6 \cdot 4 = 24$ поворотов на углы $2\pi k/5$, $1 \leqslant k \leqslant 4$, вокруг осей, проходящих через центры противоположных граней додекаэдра, $10 \cdot 2 = 20$ поворотов на углы $\pm2\pi/3$ вокруг осей, проходящих через противоположные вершины, 15 поворотов на 180° вокруг осей, проходящих через середины противоположных рёбер, и тождественного преобразования. Полная группа додекаэдра состоит из $20 \cdot 6 = 120$ движений и помимо перечисленных 60 поворотов содержит их композиции с центральной симметрией относительно центра додекаэдра.

Упражнение 11.15. Покажите что полные группы куба, октаэдра и икосаэдра состоят, соответственно из 48, 48 и 120 движений, а собственные — из 24, 24 и 60 поворотов.

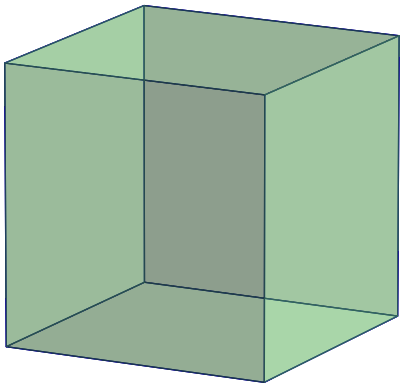


Рис. 11◇7. Куб.

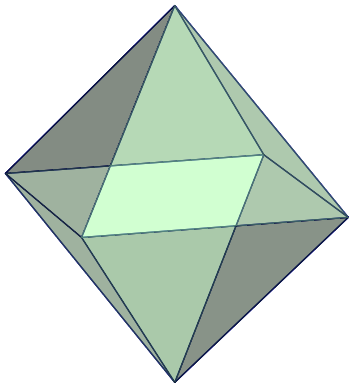


Рис. 11◇8. Октаэдр.

11.3. Гомоморфизмы групп. Отображение групп $\varphi\colon G_1 \to G_2$ называется *гомоморфизмом*, если оно переводит композицию в композицию, т. е. для любых $g, h \in G_1$ в группе G_2 выполняется соотношение $\varphi(gh) = \varphi(g)\varphi(h)$. Термины *эпиморфизм*, *мономорфизм* и *изоморфизм* применительно к отображению групп далее по умолчанию будут подразумевать, что это отображение является *гомоморфизмом* групп.

Упражнение 11.16. Убедитесь, что композиция гомоморфизмов тоже является гомоморфизмом.

Каждый гомоморфизм групп $\varphi\colon G_1 \to G_2$ переводит единицу e_1 группы G_1 в единицу e_2 группы G_2: равенство $\varphi(e_1) = e_2$ получается из равенств $\varphi(e_1)\varphi(e_1) = \varphi(e_1e_1) = \varphi(e_1)$ умножением правой и левой части на $\varphi(e_1)^{-1}$. Кроме того, для любого $g \in G$ выполняется равенство $\varphi(g^{-1}) = \varphi(g)^{-1}$, поскольку $\varphi(g^{-1})\varphi(g) = \varphi(g^{-1}g) = \varphi(e_1) = e_2$. Поэтому образ

$$\operatorname{im}\varphi \stackrel{\text{def}}{=} \varphi(G_1) \subset G_2$$

гомоморфизма групп является *подгруппой* группы G_2. Полный прообраз единицы $e_2 \in G_2$

$$\ker\varphi \stackrel{\text{def}}{=} \varphi^{-1}\left(e_2\right) = \left\{ g \in G_1 \;\middle|\; \varphi(g_1) = e_2 \right\}\ .$$

называется *ядром* гомоморфизма φ и является подгруппой в G_1, ибо из равенств $\varphi(g) = e_2$, $\varphi(h) = e_2$ вытекает равенство $\varphi(gh) = \varphi(g)\varphi(h) = e_2e_2 = e_2$, а из равенства $\varphi(g) = e_2$ — равенство $\varphi(g^{-1}) = \varphi(g)^{-1} = e_2^{-1} = e_2$.

Предложение 11.1

Все непустые слои произвольного гомоморфизма групп $\varphi\colon G_1 \to G_2$ находится во взаимно однозначном соответствии его ядром $\ker\varphi$, причём $\varphi^{-1}\big(\varphi(g)\big) = g(\ker\varphi) = (\ker\varphi)g$, где

$$g(\ker\varphi) \stackrel{\text{def}}{=} \{gh \mid h \in \ker\varphi\} \quad \text{и} \quad (\ker\varphi)g \stackrel{\text{def}}{=} \{hg \mid h \in \ker\varphi\}\,.$$

Доказательство. Если $\varphi(t) = \varphi(g)$, то $\varphi(tg^{-1}) = \varphi(t)\varphi(g)^{-1} = e$ и $\varphi(g^{-1}t) = \varphi(g)^{-1}\varphi(t) = e$, т. е. $tg^{-1} \in \ker\varphi$ и $g^{-1}t \in \ker\varphi$. Поэтому $t \in (\ker\varphi)g$ и $t \in g(\ker\varphi)$. Наоборот, для всех $h \in \ker\varphi$ выполняются равенства $\varphi(hg) = \varphi(h)\varphi(g) = \varphi(g)$ и $\varphi(gh) = \varphi(g)\varphi(h) = \varphi(g)$. Тем самым, полный прообраз $\varphi^{-1}\big(\varphi(g)\big)$ элемента $\varphi(g)$ совпадает и с $(\ker\varphi)g$, и с $g(\ker\varphi)$, а $(\ker\varphi)g$ и $g(\ker\varphi)$ совпадают друг с другом. Взаимно обратные биекции

$$\ker\varphi \underset{g^{-1}t \leftarrow t}{\overset{h \mapsto gh}{\rightleftarrows}} g(\ker\varphi)$$

между ядром и слоем $\varphi^{-1}\big(\varphi(g)\big) = g(\ker\varphi)$ задаются левым умножением элементов ядра на g, а элементов слоя — на g^{-1}. □

Следствие 11.1
Для того, чтобы гомоморфизм групп $\varphi: G_1 \to G_2$ был инъективен, необходимо и достаточно, чтобы его ядро исчерпывалось единичным элементом. □

Следствие 11.2
Для любого гомоморфизма конечных групп $\varphi: G_1 \to G_2$ выполнено равенство

$$|\operatorname{im}(\varphi)| = |G_1|/|\ker(\varphi)|\,. \tag{11-7}$$

В частности, $|\ker\varphi|$ и $|\operatorname{im}\varphi|$ делят $|G_1|$. □

Пример 11.6 (знакопеременные группы)
Согласно упр. 8.4 на стр. 108 имеется мультипликативный гомоморфизм $\operatorname{sgn}: S_n \to \{\pm 1\}$, сопоставляющий перестановке её знак. Ядро этого гомоморфизма обозначается $A_n = \ker\operatorname{sgn}$ и называется *знакопеременной группой* или группой чётных перестановок. Порядок $|A_n| = n!/2$.

Пример 11.7 (линейные группы)
Все линейные автоморфизмы произвольного векторного пространства V над произвольным полем $\Bbbk$ образуют *полную линейную группу* $\mathrm{GL}(V)$. В n° 8.1.4 на стр. 111 мы построили гомоморфизм полной линейной группы в мультипликативную группу $\Bbbk^*$ поля $\Bbbk$, сопоставляющий невырожденному линейному оператору $F: V \xrightarrow{\sim} V$ его определитель:

$$\det: \ \mathrm{GL}(V) \to \Bbbk^*, \quad F \mapsto \det F\,. \tag{11-8}$$

Ядро этого гомоморфизма называется *специальной линейной группой* и обозначается

$$\mathrm{SL}(V) = \ker\det = \{F: V \xrightarrow{\sim} V \ \ \det F = 1\}.$$

Если $\dim V = n$ и поле $\Bbbk = \mathbb{F}_q$ состоит из q элементов, полная линейная группа конечна и

$$\big|\mathrm{GL}_n(\mathbb{F}_q)\big| = (q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{n-1})\,,$$

поскольку элементы $\mathrm{GL}(V) \simeq \mathrm{GL}_n(\mathbb{F}_q)$ взаимно однозначно соответствуют базисам пространства V.

Упражнение 11.17. Убедитесь в этом.

Поскольку гомоморфизм (11-8) сюръективен[1] порядок специальной линейной группы

$$\big|\mathrm{SL}_n(\mathbb{F}_q)\big| = \big|\mathrm{GL}_n(\mathbb{F}_q)\big|/|\Bbbk^*|(q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{n-1})/(q-1)$$

Пример 11.8 (проективные группы)
Напомню[2], что с каждым векторным пространством V ассоциировано *проективное пространство* $\mathbb{P}(V)$, точками которого являются одномерные векторные подпространства в V или, что

[1] Диагональный оператор F с собственными значениями $(\lambda, 1, 1, \ldots, 1)$ имеет $\det F = \lambda$.

[2] Мы предполагаем, что читатель знаком с проективными пространствами и проективными преобразованиями по курсу геометрии.

то же самое, классы пропорциональности ненулевых векторов в V. Каждый линейный оператор $F \in \mathrm{GL}(V)$ корректно задаёт биекцию $\overline{F}\colon \mathbb{P}(V) \to \mathbb{P}(V)$, переводящую класс вектора $v \neq 0$ в класс вектора $F(v)$. Таким образом возникает гомоморфизм $F \mapsto \overline{F}$ группы $\mathrm{GL}(V)$ в группу биективных преобразований проективного пространства $\mathbb{P}(V)$. Образ этого гомоморфизма обозначается $\mathrm{PGL}(V)$ и называется *проективной линейной группой* пространства V. Из курса геометрии известно, что два оператора $F, G \in \mathrm{GL}(V)$ тогда и только тогда задают одинаковые преобразования $\overline{F} = \overline{G}$ проективного пространства $\mathbb{P}(V)$, когда они пропорциональны, т. е. $F = \lambda G$ для некоторого $\lambda \in \Bbbk^*$. Поэтому ядром эпиморфизма групп

$$\pi\colon\ \mathrm{GL}(V) \twoheadrightarrow \mathrm{PGL}(V)\,, \quad F \mapsto \overline{F} \tag{11-9}$$

является *подгруппа гомотетий* $\Gamma \simeq \Bbbk^*$, состоящая из диагональных скалярных операторов $v \mapsto \lambda v$, $\lambda \in \Bbbk^*$. Таким образом, группа $\mathrm{PGL}(V)$ образована классами пропорциональности линейных операторов. Классы пропорциональности операторов с единичным определителем образуют в ней подгруппу, обозначаемую $\mathrm{PSL}(V) \subset \mathrm{PGL}(V)$. Ограничение эпиморфизма (11-9) на подгруппу $\mathrm{SL}(V) \subset \mathrm{GL}(V)$ доставляет эпиморфизм

$$\pi'\colon\ \mathrm{SL}(V) \twoheadrightarrow \mathrm{PSL}(V)\,, \quad F \mapsto \overline{F} \tag{11-10}$$

ядром которого является конечная мультипликативная подгруппа $\boldsymbol{\mu}_n(\Bbbk) \subset \Bbbk^*$ содержащихся в поле $\Bbbk$ корней n-той степени из единицы, где[1] $n = \dim V = \dim \mathbb{P}(V) + 1$.

Пример 11.9 (эпиморфизм $S_4 \twoheadrightarrow S_3$)
На проективной плоскости $\mathbb{P}_2$ над любым полем $\Bbbk$ с каждой четвёркой точек a, b, c, d, никакие три из которых не коллинеарны связана фигура, образованная тремя парами проходящих через эти точки прямых[2]

$$(ab) \text{ и } (cd), \quad (ac) \text{ и } (bd), \quad (ad) \text{ и } (bc) \tag{11-11}$$

и называемая *четырёхвершинником* (см. рис. 11⋄9). Пары прямых (11-11) называются *противоположными сторонами* четырёхвершинника. С четырёхвершинником $abcd$ ассоциирован треугольник xyz с вершинами в точках пересечения пар противоположных сторон

$$x = (ab)\cap(cd),\ y = (ac)\cap(bd),\ z = (ad)\cap(bc) \tag{11-12}$$

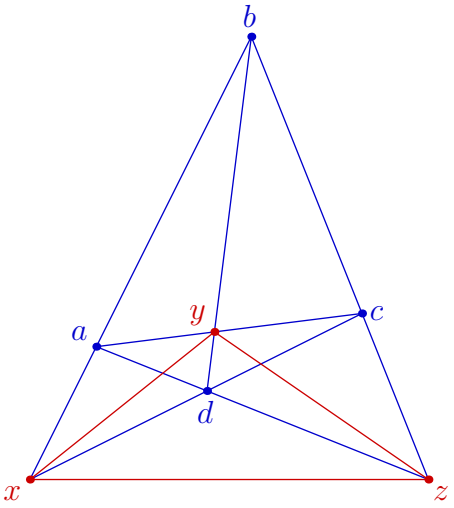


Рис. 11⋄9. Четырёхвешинник и ассоциированный треугольник.

Каждая перестановка вершин a, b, c, d однозначно определяет линейное проективное преобразование[3] плоскости, что даёт вложение $S_4 \hookrightarrow \mathrm{PGL}_3(\Bbbk)$. Преобразования из S_4 переводят ассоциированный треугольник xyz в себя, переставляя его вершины x, y, z согласно формулам (11-12). Например, 3-цикл $(b, c, a, d) \in S_4$ задаёт циклическую перестановку

[1] Напомню, что по определению $\dim \mathbb{P}(V) \stackrel{\text{def}}{=} \dim V - 1$.

[2] Они отвечают трём возможным способам разбить точки a, b, c, d на две пары.

[3] Напомню, что каждое линейное проективное преобразование $\overline{F} \in \mathrm{PGL}(V)$ однозначно определяется своим действием на любые $\dim V + 1$ точек пространства $\mathbb{P}(V)$, никакие $\dim V$ из которых не лежат в одной гиперплоскости.

(y,z,x), а транспозиции (b,a,c,d), (a,c,b,d) и (c,b,a,d) дают транспозиции (x,z,y), (y,x,z) и (z,y,x) соответственно. Таким образом, мы получаем сюръективный гомоморфизм $S_4 \twoheadrightarrow S_3$. Его ядро имеет порядок $4!/3! = 4$ и состоит из тождественной перестановки и трёх пар независимых транспозиций (b,a,d,c), (c,d,a,b), (d,c,b,a).

Пример 11.10 (S_4 и собственная группа куба)
Линейные преобразования евклидова пространства $\mathbb{R}_3$, составляющие собственную группу куба с центром в нуле, действуют на четырёх прямых a, b, c, d, соединяющих противоположные вершины куба, а также на трёх прямых x, y, z, соединяющих центры его противоположных граней, см. рис. 11⋄10. На проективной плоскости $\mathbb{P}_2 = \mathbb{P}(\mathbb{R}^3)$ эти 7 прямых становятся вершинами четырёхвершинника $abcd$ и ассоциированного с ним треугольника xyz, как на рис. 11⋄9. Поворот на 180° вокруг оси, соединяющей середины противоположных рёбер куба, меняет местами примыкающие к этому ребру диагонали и переводит в себя каждую их двух оставшихся диагоналей. Тем самым, вращения куба осуществляют транспозиции любых двух соседних диагоналей, и мы имеем сюръективный гомоморфизм $\mathrm{SO}_{\text{куб}} \to S_4$. Так как обе группы имеют порядок 24, это изоморфизм. Он переводит 6 поворотов на $\pm 90°$ вокруг прямых x, y, z в 6 циклов длины 4 циклового типа $\square\square\square\square$, 3 поворота на 180° вокруг тех же прямых — в 3 пары независимых транспозиций циклового типа $\begin{smallmatrix}\square\square\\\square\square\end{smallmatrix}$, 8 поворотов на $\pm 120°$ вокруг прямых a, b, c, d — в 8 циклов длины 3 циклового типа $\begin{smallmatrix}\square\square\square\\\square\end{smallmatrix}$, а 6 поворотов на 180° вокруг осей, проходящих через середины противоположных рёбер — в 6 простых транспозиций циклового типа $\begin{smallmatrix}\square\square\\\square\\\square\end{smallmatrix}$. Гомоморфизм $\mathrm{SO}_{\text{куб}} \to S_3$, возникающий из действия группы куба на прямых x, y, z, согласован с изоморфизмом $\mathrm{SO}_{\text{куб}} \xrightarrow{\sim} S_4$ и эпиморфизмом $S_4 \twoheadrightarrow S_3$ из предыдущего прим. 11.9. Его ядро состоит из собственных ортогональных преобразований евклидова пространства $\mathbb{R}^3$, переводящих в себя каждую из декартовых координатных осей x, y, z в $\mathbb{R}^3$, и совпадает, таким образом, с группой двуугольника D_2 с осями x, y, z. В таком контексте эту группу иногда называют *четвертной группой Клейна* и обозначают V_4. Изоморфизм $\mathrm{SO}_{\text{куб}} \xrightarrow{\sim} S_4$ переводит её в ядро эпиморфизма $S_4 \twoheadrightarrow S_3$ из прим. 11.9.

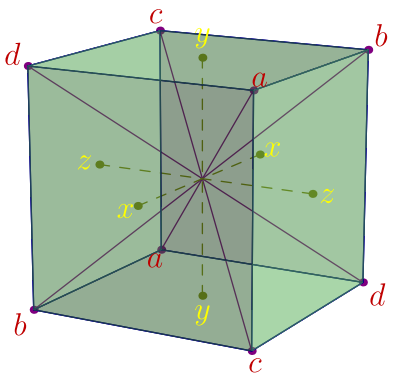


Рис. 11⋄10. От куба к четырёхвершиннику.

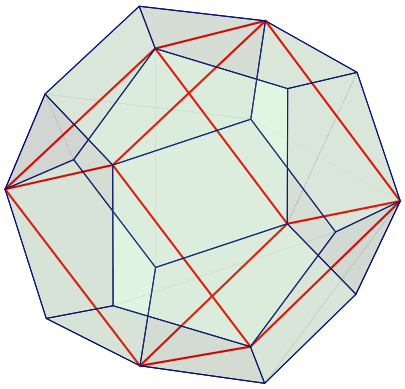


Рис. 11⋄11. Один из пяти кубов на додекаэдре.

Пример 11.11 (собственная группа додекаэдра и A_5)
Любая диагональ любой грани додекаэдра единственным образом достраивается до лежащего на поверхности додекаэдра куба, образованного диагоналями граней так, что в каждой грани

рисуется ровно одна диагональ[1], как на рис. 11⋄11. Всего на поверхности додекаэдра имеется ровно 5 таких кубов — они биективно соответствуют пяти диагоналям какой-либо фиксированной грани. Собственная группа додекаэдра переставляет эти кубы друг с другом, что даёт гомоморфизм собственной группы додекаэдра в симметрическую группу S_5:

$$\psi_{\text{дод}}\colon\ \mathrm{SO}_{\text{дод}} \to S_5 \tag{11-13}$$

Глядя на модель додекаэдра, легко видеть, что образами $20 \cdot 3 = 60$ поворотов, из которых состоит группа $\mathrm{SO}_{\text{дод}}$ будут в точности 60 чётных перестановок: $6 \cdot 4 = 24$ поворота на углы $2\pi k/5$, $1 \leqslant k \leqslant 4$, вокруг осей, проходящих через центры противоположных граней, переходят во всевозможные циклы длины 5, т. е. в 24 перестановки циклового типа $\square\square\square\square\square$); $10 \cdot 2 = 20$ поворотов на углы $\pm 2\pi/3$ вокруг осей, проходящих через противоположные вершины додекаэдра, переходят во всевозможные циклы длины 3, т. е. в 20 перестановок циклового типа $\begin{smallmatrix}\square\square\square\\\square\\\square\end{smallmatrix}$; 15 поворотов на 180° вокруг осей, проходящих через середины противоположных рёбер додекаэдра, переходят во всевозможные пары независимых транспозиций, т. е. в 10 перестановок циклового типа $\begin{smallmatrix}\square\square\\\square\square\\\square\end{smallmatrix}$. Оставшееся неучтённым тождественное преобразование додекаэдра задаёт тождественную перестановку кубов. Таким образом, гомоморфизм (11-13) является изоморфизмом собственной группы додекаэдра со знакопеременной подгруппой $A_5 \subset S_5$. В отличие от прим. 11.4 переход от собственной группы додекаэдра к полной не добавляет новых перестановок кубов, поскольку каждое несобственное движение является композицией собственного движения и центральной симметрии, которая переводит каждый из кубов в себя.

Упражнение 11.18. Покажите, что симметрическая группа S_5 не изоморфна полной группе додекаэдра.

11.4. Действие группы на множестве. Пусть G — группа, а X — множество. Обозначим через $\mathrm{Aut}(X)$ группу всех взаимно однозначных отображений из X в себя. Гомоморфизм

$$\varphi\colon\ G \to \mathrm{Aut}(X)$$

называется *действием* группы G на множестве X или *представлением* группы G автоморфизмами множества X. Отображение $\varphi(g)\colon\ X \to X$, отвечающее элементу $g \in G$ при действии φ часто бывает удобно обозначать через $\varphi_g\colon\ X \to X$. Тот факт, что сопоставление $g \mapsto \varphi_g$ является гомоморфизмом групп, означает, что $\varphi_{gh} = \varphi_g \circ \varphi_h$ для всех $g, h \in G$. Если понятно, о каком действии идёт речь, мы часто будем сокращать $\varphi_g(x)$ до gx. При наличии действия группы G на множестве X мы пишем $G\ :\ X$. Действие называется *транзитивным*, если любую точку множества X можно перевести в любую другую точку каким-нибудь преобразованием из группы G, т. е. $\forall x, y \in X\ \exists g \in G\ :\ gx = y$. Более общим образом, действие называется *m-транзитивным*, если любые два упорядоченных набора из m различных точек множества X можно перевести друг в друга подходящими преобразованиями из G. Действие называется *свободным*, если каждый отличный от единицы элемент группы действует на X без неподвижных точек, т. е. $\forall\, g \in G\ \forall\, x \in X\ gx = x \Rightarrow g = e$. Действие $\varphi\colon\ G \to \mathrm{Aut}\, X$ называется *точным* (или

[1]Проще всего это увидеть на модели додекаэдра, которую мы ещё раз настоятельно рекомендуем изготовить.

эффективным), если каждый отличный от единицы элемент группы действует на X не тождественно, т. е. когда $\ker\varphi = e$. Точное представление отождествляет G с группой преобразований $\varphi(G) \subset \mathrm{Aut}(X)$ множества X. Отметим, что любое свободное действие точно.

Если группа G действует на множестве X, то она действует и на подмножествах множества X: элемент $g \in G$ переводит подмножество $M \subset X$ в подмножество $gM = \{gm \mid m \in M\}$. При этом отображение $g : M \to gM$, $x \mapsto gx$ биективно, и обратным к нему является отображение $g^{-1} : gM \to g$, $\mapsto g^{-1}y$, ибо $g^{-1}gx = ex = x$. Говорят, что элемент $g \in G$ *нормализует*[1] подмножество $M \subset X$, если $gM = M$, т. е. $gx \in M$ для каждого $x \in M$. Каждый такой элемент задаёт биекцию $g|_M : M \to M$. Если эта биекция тождественна, т. е. $gx = x$ для всех $x \in M$, то говорят, что элемент g *централизует* подмножество M. Множество всех элементов $g \in G$, нормализующих (соотв. централизующих) данное подмножество $M \subset X$ обозначается $N(M)$ (соотв. $Z(M)$) и называется *нормализатором* (соотв. *централизатором*) подмножества $M \subset X$ при заданном действии группы G на X.

Упражнение 11.19. Убедитесь, что $N(M)$ и $Z(M)$ являются подгруппами в G.

Пример 11.12 (регулярные действия)
Обозначим через X множество элементов группы G, а через $\mathrm{Aut}(X)$ — группу автоморфизмов этого *множества*[2]. Отображение $\lambda : G \to \mathrm{Aut}\, X$, переводящее элемент $g \in G$ в преобразование[3] $\lambda_g : x \mapsto gx$ левого умножения на g является гомоморфизмом групп, поскольку

$$\lambda_{gh}(x) = ghx = \lambda_g(hx) = \lambda_g\left(\lambda_h(x)\right) = \lambda_g \circ \lambda_h\,(x)\,.$$

Оно называется *левым регулярным действием* группы G на себе. Так как равенство $gh = h$ в группе G влечёт равенство $g = e$, левое регулярное действие свободно и, в частности, точно. Симметричным образом, *правое регулярное действие* $\varrho_g : G \to \mathrm{Aut}(X)$ сопоставляет элементу $g \in G$ преобразование $x \mapsto xg^{-1}$ правого умножения на обратный[4] к g элемент.

Упражнение 11.20. Убедитесь, что ϱ_g является свободным действием.

Тем самым, любая абстрактная группа G может быть реализована как группа преобразований некоторого множества. Например, левые регулярные представления числовых групп реализуют аддитивную группу $\mathbb{R}$ группой сдвигов $\lambda_v : x \mapsto x + v$ числовой прямой , а мультипликативную группу $\mathbb{R}^*$ — группой гомотетий $\lambda_c : x \mapsto cx$ проколотой прямой $\mathbb{R}^* = \mathbb{R} \smallsetminus \{0\}$.

Пример 11.13 (присоединённое действие)
Отображение $\mathrm{Ad} : G \to \mathrm{Aut}(G)$, сопоставляющее элементу $g \in G$ автоморфизм *сопряжения* этим элементом

$$\mathrm{Ad}_g : G \to G\,, \quad h \mapsto ghg^{-1}\,, \tag{11-14}$$

называется *присоединённым действием* группы G на себе.

Упражнение 11.21. Убедитесь, что $\forall\, g \in G$ сопряжение (11-14) является *гомоморфизмом* из G в G и что отображение $g \mapsto \mathrm{Ad}_g$ является гомоморфизмом из G в $\mathrm{Aut}\, G$.

[1] В этом случае также говорят, что подмножество $M \subset X$ является g-инвариантным.

[2] Возможно, не перестановочных с имеющейся в G композицией, т. е. не обязательно являющихся автоморфизмами *группы* G.

[3] Обратите внимание, что это преобразование множества X не является гомоморфизмом группы G, поскольку равенство $g(h_1h_2) = (gh_1)(gh_2)$, вообще говоря, не выполняется.

[4] Появление g^{-1} не случайно: проверьте, что сопоставление элементу $g \in G$ отображения правого умножения на g является не гомоморфизмом, а антигомоморфизмом (т. е. оборачивает порядок сомножителей в произведениях).

Образ присоединённого действия $\mathrm{Ad}(G) \subset \mathrm{Aut}\, G$ обозначается $\mathrm{Int}(G)$ и называется группой *внутренних автоморфизмов* группы G. Не лежащие в $\mathrm{Int}(G)$ автоморфизмы группы G называются *внешними*. В отличие от левого и правого регулярных действий присоединённое действие, вообще говоря, не свободно и не точно. Например, если группа G абелева, все внутренние автоморфизмы (11-14) тождественные, и ядро присоединённого действия в этом случае совпадает со всей группой. В общем случае $\ker(\mathrm{Ad})$ образовано такими $g \in G$, что $ghg^{-1} = h$ для всех $h \in G$. Последнее равенство равносильно равенству $gh = hg$ и означает, что g *коммутирует* со всеми элементами группы. Подгруппа элементов, перестановочных со всеми элементами группы G называется *центром* группы G и обозначается

$$Z(G) = \ker(\mathrm{Ad}) = \{g \in G \mid \forall\, h \in G\ gh = hg\}\,.$$

Стабилизатор заданного элемента $g \in G$ в присоединённом действии состоит из всех элементов группы, коммутирующих с g. Он называется *централизатором* элемента g и обозначается

$$C_g = \mathrm{Stab}_{\mathrm{Int}(G)}(g) = \{h \in G \mid hg = gh\}\,.$$

11.4.1. Орбиты. Со всякой группой преобразований G множества X связано бинарное отношение $y \sim x$ на X, означающее, что $y = gx$ для некоторого $g \in G$. Это отношение рефлексивно, ибо $x = ex$, симметрично, поскольку $y = gx \iff x = g^{-1}y$, и транзитивно, т. к. из равенств $y = gx$ и $z = hy$ вытекает равенство $z = (hg)x$. Таким образом, это отношение является эквивалентностью. Класс эквивалентности точки $x \in X$ состоит из всех точек, которые можно получить из x, применяя всевозможные преобразования из группы G. Он обозначается $Gx = \{gx \mid g \in G\}$ и называется *орбитой* x под действием G. Согласно n° 1.4 на стр. 10 множество X распадается в дизъюнктное объединение орбит. Множество всех орбит называется *фактором* множества X по действию группы G и обозначается X/G. С каждой орбитой Gx связано сюръективное отображение[1] множеств $\mathrm{ev}_x : G \twoheadrightarrow Gx$, $g \mapsto gx$, слой которого над точкой $y \in Gx$ состоит из всех преобразований группы G, переводящих x в y. Он называется *транспортёром* x в y и обозначается $G_{yx} = \{g \in G \mid gx = y\}$. Слой над самой точкой x состоит из всех преобразований, оставляющих x на месте. Он называется *стабилизатором* точки x в группе G и обозначается $\mathrm{Stab}_G(x) = G_{xx} = \{g \in G \mid gx = x\}$ или просто $\mathrm{Stab}(x)$, если понятно, о какой группе G идёт речь.

Упражнение 11.22. Убедитесь, что $\mathrm{Stab}_G(x)$ является подгруппой в группе G.

Если $y = gx$ и $z = hx$, то для любого $s \in \mathrm{Stab}(x)$ преобразование $hsg^{-1} \in G_{zy}$. Наоборот, если $fy = z$, то $h^{-1}fg \in \mathrm{Stab}(x)$. Таким образом, мы имеем обратные друг другу отображения множеств:

$$\mathrm{Stab}(x) \underset{h^{-1}fg \mapsfrom f}{\overset{s \mapsto hsg^{-1}}{\rightleftarrows}} G_{zy}\,, \tag{11-15}$$

и стало быть, для любых трёх точек x, y, z из одной G-орбиты имеется биекция между G_{zy} и $\mathrm{Stab}(x)$.

Предложение 11.2 (формула для длины орбиты)
Длина орбиты произвольной точки x при действии на неё конечной группы преобразований G равна $|Gx| = |G| : |\mathrm{Stab}_G(x)|$. В частности, длины всех орбит и порядки стабилизаторов всех точек являются делителями порядка группы.

[1] При желании его можно воспринимать как «некоммутативное»*отображения вычисления*.

Доказательство. Группа G является дизъюнктным объединением множеств G_{yx} по всем $y \in Gx$ и согласно предыдущему все эти множества состоят из $|\mathrm{Stab}(x)|$ элементов. □

Предложение 11.3
Стабилизаторы всех точек, лежащих в одной орбите конечной группы, сопряжены:

$$y = gx \;\Rightarrow\; \mathrm{Stab}(y) = g\,\mathrm{Stab}(x)\,g^{-1} = \{ghg^{-1} \mid h \in \mathrm{Stab}(x)\}\,.$$

В частности, все они имеют одинаковый порядок.

Доказательство. Это сразу следует из диаграммы (11-15). □

Пример 11.14 (действие перестановок букв на словах)
Зафиксируем какой-нибудь k-буквенный алфавит $A = \{a_1, \dots, a_k\}$ и рассмотрим множество X всех n-буквенных слов w, которые можно написать с его помощью. Иначе X можно воспринимать как множество всех отображений $w : \{1, 2, \dots, n\} \to A$. Сопоставим каждой перестановке $\sigma \in S_n$ преобразование $w \mapsto w\sigma^{-1}$, которое переставляет буквы в словах так, как предписывает[1] σ. Таким образом, мы получили действие симметрической группы S_n на множестве слов. Орбита слова $w \in X$ под действием этой группы состоит из всех слов, где каждая буква алфавита встречается столько же раз, сколько в слове w. Стабилизатор $\mathrm{Stab}(w)$ слова w, в котором буква a_i встречается m_i раз (для каждого $i = 1, \dots, k$), состоит из перестановок между собою одинаковых букв и имеет порядок $|\mathrm{Stab}(w)| = m_1! \cdot m_2! \cdot \dots \cdot m_k!$. Тем самым, длина орбиты такого слова равна мультиномиальному коэффициенту

$$|S_n w| = \frac{|S_n|}{|\mathrm{Stab}(w)|} = \frac{n!}{m_1! \cdot m_2! \cdot \dots \cdot m_k!} = \binom{n}{m_1 \dots m_k}.$$

Этот пример показывает, что разные орбиты могут иметь разную длину, и порядки стабилизаторов точек из разных орбит могут быть разными.

Упражнение 11.23. Для каждого из пяти платоновых тел рассмотрите действие группы этого тела на его гранях и по формуле для длины орбиты найдите порядок собственной и несобственной группы каждого из платоновых тел.

Пример 11.15 (классы сопряжённости в симметрической группе)
Перестановка $\mathrm{Ad}_g(\sigma) = g\sigma g^{-1}$, сопряжённая перестановке $\sigma = (\sigma_1, \dots, \sigma_n) \in S_n$, для каждого $i = 1, 2, \dots, n$ переводит элемент $g(i)$ в элемент $g(\sigma_i)$. Поэтому при сопряжении цикла $\tau = |i_1, \dots, i_k\rangle \in S_n$ перестановкой $g = (g_1, \dots, g_n)$ получится цикл

$$g\tau g^{-1} = \left|g_{i_1}, \dots, g_{i_k}\right\rangle.$$

Если перестановка $\sigma \in S_n$ имеет цикловой тип λ и является произведением независимых циклов, записанных по строкам диаграммы λ, то действие на такую перестановку внутреннего автоморфизма Ad_g заключается в применении отображения g к заполнению диаграммы λ, т. е. в замене каждого числа i числом g_i.

[1] Т. е. переводит слово $w = a_{\nu_1} \dots a_{\nu_n}$ в слово $a_{\nu_{\sigma^{-1}(1)}} a_{\nu_{\sigma^{-1}(2)}} \dots a_{\nu_{\sigma^{-1}(n)}}$, на i-том месте которого стоит та буква, номер которой в исходном слове w переводится перестановкой σ в номер i.

Таким образом, орбиты присоединённого действия симметрической группы S_n на себе взаимно однозначно соответствуют n-клеточным диаграммам Юнга, и орбита, отвечающая диаграмме λ, состоит из всех перестановок циклового типа λ. Если диаграмма λ имеет m_i строк длины i для каждого $i = 1, 2, \dots, n$, то централизатор любой перестановки σ циклового типа λ состоит из таких перестановок элементов заполнения диаграммы λ независимыми циклами перестановки σ, которые не меняют σ, т. е. циклически переставляют элементы вдоль строк или произвольным образом переставляют строки одинаковой длины между собой как единое целое. Тем самым, порядок стабилизатора перестановки циклового типа λ зависит только от λ и равен

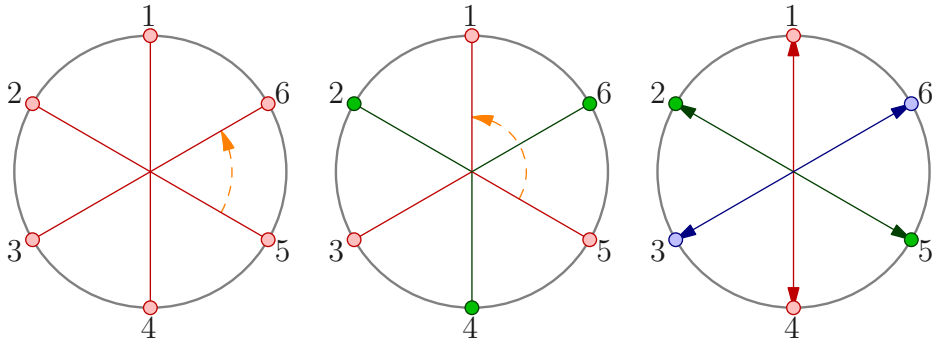
$$z_\lambda = 1^{m_1} \cdot m_1! \cdot 2^{m_2} \cdot m_2! \cdot \ \cdots \ \cdot n^{m_n} \cdot m_n! = \prod_{\alpha=1}^{n} m_\alpha! \alpha^{m_\alpha}.$$

Количество перестановок циклового типа λ, т. е. длина соответствующей орбиты присоединённого действия, равна $n!/z_\lambda$.

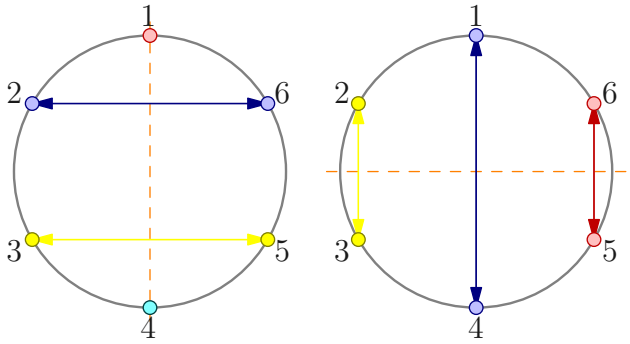
11.4.2. Перечисление орбит. Подсчёт числа элементов в факторе X/G конечного множества X по действию конечной группы G наталкивается на очевидную трудность: поскольку длины у орбит могут быть разные, число орбит «разного типа» придётся подсчитывать по отдельности, заодно уточняя по ходу дела, что именно имеется в виду под «типом орбиты». Разом преодолеть обе эти трудности позволяет

Теорема 11.2 (формула Полиа – Бернсайда)
Пусть конечная группа G действует на конечном множестве X. Для каждого $g \in G$ обозначим через $X^g = \{x \in X \mid gx = x\} = \{x \in X \mid g \in \mathrm{Stab}(x)\}$ множество неподвижных точек преобразования g. Тогда $|X/G| = |G|^{-1} \sum_{g\in G} |X^g|$.

Доказательство. Обозначим через $F \subset G \times X$ множество всех таких пар (g, x), что $gx = x$. Иначе F можно описать как $F = \bigsqcup_{x\in X} \mathrm{Stab}(x) = \bigsqcup_{g\in G} X^g$. Первое из этих описаний получается из рассмотрения проекции $F \twoheadrightarrow X$, второе — из рассмотрения проекции $F \twoheadrightarrow G$. Согласно второму описанию, $|F| = \sum_{g\in G} |X^g|$. С другой стороны, из первого описания мы заключаем, что $|F| = |G| \cdot |X/G|$. В самом деле, стабилизаторы всех точек, принадлежащих одной орбите, имеют одинаковый порядок, и сумма этих порядков по всем точкам орбиты равна произведению порядка стабилизатора на длину орбиты, т. е. $|G|$. Складывая по всем $|X/G|$ орбитам, получаем требуемое. □



τ-инваринтные бусы τ^2-инваринтные бусы τ^3-инваринтные бусы



σ_{14}-инваринтные бусы $\overline{\sigma}_{14}$-инваринтные бусы

Рис. 11⋄12. Симметричные ожерелья из шести бусин.

Пример 11.16 (ожерелья)
Пусть имеется неограниченный запас одинаковых по форме бусин n различных цветов. Сколько различных ожерелий можно сделать из 6 бусин? Ответом на этот вопрос является количество орбит группы диэдра D_6 на множестве всех раскрасок вершин правильного шестиугольника в n цветов. Группа D_6 состоит из 12 элементов: тождественного преобразования e, двух поворотов $\tau^{\pm 1}$ на $\pm 60^\circ$, двух поворотов $\tau^{\pm 2}$ на $\pm 120^\circ$, центральной симметрии τ^3, трёх отражений σ_{14}, σ_{23}, σ_{36} относительно больших диагоналей и трёх отражений $\overline{\sigma}_{14}$, $\overline{\sigma}_{23}$, $\overline{\sigma}_{36}$ относительно срединных перпендикуляров к сторонам. Единица оставляет на месте все n^6 раскрасок. Раскраски, симметричные относительно остальных преобразований, показаны на рис. 11⋄12. Беря на этих рисунках все допустимые сочетания цветов, получаем, соответственно, n, n^2, n^3, n^4 и n^3 раскрасок. По теор. 11.2 искомое число 6-бусинных ожерелий равно $\bigl(n^6 + 3\,n^4 + 4\,n^3 + 2\,n^2 + 2\,n\bigr)/12$.

Упражнение 11.24. Подсчитайте количество ожерелий из 7, 8, 9, и 10 бусин.

11.5. Смежные классы и факторизация. Каждая подгруппа $H \subset G$ задаёт на группе G два отношения эквивалентности, происходящие из левого и правого регулярного действия подгруппы H на группе G. Левое действие $\lambda_h : g \mapsto hg$ приводит к эквивалентности

$$g_1 \underset{L}{\sim} g_2 \iff g_1 = hg_2 \text{ для некоторого } h \in H, \tag{11-16}$$

разбивающей группу G в дизъюнктное объединение орбит вида $Hg \stackrel{\text{def}}{=} \{hg \mid h \in H\}$, называемых *правыми смежными классами* (или *правыми сдвигами*) подгруппы H в группе G. Множество правых смежных классов обозначается $H \backslash G$.

Упражнение 11.25. Покажите, что равенство $Hg_1 = Hg_2$ равносильно любому из эквивалентных друг другу включений $g_1^{-1}g_2 \in H$, $g_2^{-1}g_1 \in H$.

С правым действием $\varrho_h : g \mapsto gh^{-1}$ связано отношение эквивалентности

$$g_1 \underset{R}{\sim} g_2 \iff g_1 = g_2h \text{ для некоторого } h \in H, \tag{11-17}$$

разбивающее группу G в дизъюнктное объединение орбит $gH \stackrel{\text{def}}{=} \{gh \mid h \in H\}$, которые называются *левыми смежными классами* (или *левыми сдвигами*) подгруппы H в группе G. Множество левых смежных классов обозначается G/H.

Поскольку и левое и правое действия подгруппы H на группе G свободны, все орбиты каждого из них состоят из $|H|$ элементов. Тем самым, число орбит в обоих действиях одинаково и равно $|G|\,/\,|H|$. Это число называется *индексом* подгруппы H в группе G и обозначается $[G\,:\,H] \stackrel{\text{def}}{=} |G\,/\,H|$. Нами установлена

Теорема 11.3 (теорема Лагранжа об индексе подгруппы)
Порядок и индекс любой подгруппы H в произвольной конечной группе G нацело делят порядок G и $[G\,:\,H] = |G|\,:\,|H|$.

Следствие 11.3
Порядок любого элемента конечной группы нацело делит порядок группы.

Доказательство. Порядок элемента $g \in G$ равен порядку порождённой им циклической подгруппы $\langle g\rangle \subset G$. □

11.5.1. Нормальные погруппы. Подгруппа $H \subset G$ называется *нормальной* (или *инвариантной*) , если для любого $g \in G$ выполняется равенство $gHg^{-1} = H$ или, что то же самое, $gH = Hg$. Иначе можно сказать, что подгруппа $H \subset G$ нормальна тогда и только тогда, когда левая и правая эквивалентности (11-16) и (11-17) совпадают друг с другом и, в частности, $H\setminus G = G\,/\,H$. Если подгуппа $H \subset G$ нормальна, мы пишем $H \lhd G$.

Пример 11.17 (ядра гомоморфизмов)
Ядро любого гомоморфизма групп $\varphi\colon G_1 \to G_2$ является нормальной подгруппой в G_1, поскольку при $\varphi(h) = e$ для любого $g \in G$ имеем равенство $\varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = e$, означающее, что $g\,(\ker\varphi)\,g^{-1} \subset \ker\varphi$.

Упражнение 11.26. Покажите, что если для любого $g \in G$ есть включение $gHg^{-1} \subset H$, то все эти включения — равенства.

Отметим, что совпадение правых и левых смежных классов ядра $g\,(\ker\varphi) = (\ker\varphi)\,g$ уже было установлено нами ранее в предл. 11.1.

Пример 11.18 ($V_4 \lhd S_4$)
Подгруппа Клейна $V_4 \subset S_4$ состоящая из перестановок циклового типа  и тождественной перестановки нормальна.

Пример 11.19 (внутренние автоморфизмы)
Подгруппа внутренних автоморфизмов $\mathrm{Int}(G) = \mathrm{Ad}(G)$ нормальна в группе $\mathrm{Aut}(G)$ всех автоморфизмов группы G, поскольку сопрягая внутренний автоморфизм $\mathrm{Ad}_g\colon h \mapsto ghg^{-1}$ произвольным автоморфизмом $\varphi\colon G \xrightarrow{\sim} G$, мы получаем внутренний автоморфизм

$$\varphi \circ \mathrm{Ad}_g \circ \varphi^{-1} = \mathrm{Ad}_{\varphi(g)}\ .$$

Упражнение 11.27. Убедитесь в этом.

Пример 11.20 (параллельные переносы)
Подгруппа параллельных переносов нормальна в группе $\mathrm{Aff}(\mathbb{A}^n)$ всех биективных аффинных преобразований аффинного пространства $\mathbb{A}^n$, т. к. сопрягая параллельный перенос τ_v на век-

тор v любым аффинным преобразованием $\varphi\colon \mathbb{A}^n \to \mathbb{A}^n$, получаем перенос[1] $\tau_{D_\varphi(v)}$ на вектор $D_\varphi(v)$.

Упражнение 11.28. Убедитесь в этом.

Пример 11.21 (нормализатор и централизатор, ср. с упр. 11.19 на стр. 161)
Пусть группа G действует на множестве X и $M \subset X$ — произвольное подмножество. Напомню[2], что подгруппы

$$N(M) \stackrel{\text{def}}{=} \{g \in G \mid \forall\, x \in M\ gx \in M\}$$
$$Z(M) \stackrel{\text{def}}{=} \{g \in G \mid \forall\, x \in M\ gx = x\}$$

называются соответственно *нормализатором* и *централизатором* подмножества M. Поскольку для любых $g \in N(M)$, $h \in Z(M)$ и $x \in M$ выполняется равенство $ghg^{-1}x = gg^{-1}x = x$, ибо $h(g^{-1}x) = g^{-1}x$, так как $g^{-1}x \in M$, централизатор является нормальной подгруппой в нормализаторе.

11.5.2. Фактор группы. Попытка определить умножение на множестве левых смежных классов G/H неабелевой группы G формулой

$$(g_1H)\cdot(g_2H) \stackrel{\text{def}}{=} (g_1g_2)H\,, \tag{11-18}$$

вообще говоря, некорректна: различные записи $g_1H = f_1H$ и $g_2H = f_2H$ одних и тех же классов могут приводить к *различным* классам $(g_1g_2)H \neq (f_1f_2)H$.

Упражнение 11.29. Убедитесь, что для группы $G = S_3$ и подгруппы второго порядка $H \subset G$, порождённой транспозицией σ_{12}, формула (11-18) некорректна.

Предложение 11.4
Для того, чтобы правило $g_1H \cdot g_2H = (g_1g_2)H$ корректно определяло на G/H структуру группы, необходимо и достаточно, чтобы подгруппа H была нормальна в G.

Доказательство. Если формула (11-18) корректна, то она задаёт на множестве смежных левых классов G/H групповую структуру: ассоциативность композиции наследуется из[3] G, единицей служит класс $eH = H$, обратным к классу gH — класс $g^{-1}H$. Факторизация $G \twoheadrightarrow G/H$, $g \mapsto gH$, является гомоморфизмом групп с ядром H. Поэтому подгруппа H нормальна в силу прим. 11.17. Наоборот, пусть H нормальна и пусть $f_1H = g_1H$ и $f_2H = g_2H$. Мы должны убедиться, что $(f_1f_2)H = (g_1g_2)H$. Так как левый смежный класс $f_2H = g_2H$ совпадает с правым классом Hg_2, каждый элемент вида f_1f_2h можно переписать как $f_1h_1g_2$ с подходящими $h_1 \in H$. Аналогично, $f_1h_1 = h_2g_1$ для подходящего $h_2 \in H$ в виду равенств $f_1H = g_1H = Hg_1$. Наконец из равенства $H(g_1g_2) = (g_1g_2)H$ мы заключаем, что $f_1f_2h = h_2g_1g_2 = g_1g_2h_3$ для некоторого $h_3 \in H$, откуда $(f_1f_2)H \subset (g_1g_2)H$. Противоположное включение доказывается аналогично. □

[1] Напомню, что преобразование $\varphi\colon \mathbb{A}(V) \to \mathbb{A}(V)$ аффинного пространства $\mathbb{A}(V)$, ассоциированного с векторным пространством V, называется *аффинным*, если отображение $D_\varphi\colon \overrightarrow{pq} \mapsto \overrightarrow{\varphi(p)\varphi(q)}$ является корректно определённым линейным преобразованием векторного пространства V (оно называется *дифференциалом* отображения φ).

[2] См. n° 11.4 на стр. 160.

[3] $(g_1H\cdot g_2H)\cdot g_3H = (g_1g_2)H\cdot g_3H = \big((g_1g_2)g_3\big)H = \big(g_1(g_2g_3)\big)H = g_1H\cdot(g_2g_3)H = g_1H\cdot(g_2H\cdot g_3H)$.

Определение 11.2
Множество смежных классов G/H нормальной подгруппы $H \lhd G$ с операцией

$$g_1H \cdot g_2H \stackrel{\text{def}}{=} (g_1g_2)H$$

называется *фактором* (или *фактор группой*) группы G по нормальной подгруппе H. Гомоморфизм групп $G \twoheadrightarrow G/H$, $g \mapsto gH$, называется *гомоморфизмом факторизации*.

Следствие 11.4
Каждый гомоморфизм групп $\varphi\colon G_1 \to G_2$ является композицией эпиморфизма факторизации $G_1 \twoheadrightarrow G_1/\ker\varphi$ и мономорфизма $G_1/\ker\varphi \hookrightarrow G_2$, переводящего смежный класс $g\ker\varphi \in G_1/\ker\varphi$ в элемент $\varphi(g) \in G_2$. В частности, $\operatorname{im}\varphi \simeq G/\ker\varphi$.

Доказательство. Следствие утверждает, что слой $\varphi^{-1}\bigl(\varphi(g)\bigr)$ гомоморфизма φ над каждой точкой $\varphi(g) \in \operatorname{im}\varphi \subset G_2$ является левым сдвигом ядра $\ker\varphi$ на элемент g, что мы уже видели в предл. 11.1 на стр. 156. □

Предложение 11.5
Если подгруппа $H \subset G$ нормализует[1] подгруппу $N \subset G$, то множества $HN = \{hn \mid h \in H,\ n \in N\}$ и $NH = \{nh \mid n \in N,\ h \in H\}$ совпадают друг с другом и являются подгруппой в G, причём $N \lhd HN$, $H \cap N \lhd H$ и $HN/N \simeq H/(H \cap N)$.

Доказательство. $NH = HN$ ибо $nh = h(h^{-1}nh) \in HN$ и $hn = (hnh^{-1})h \in NH$ для всех $n \in N$, $h \in H$. Это подгруппа, так как $(nh)^{-1} = h^{-1}n^{-1} \in HN = NH$ и

$$(n_1h_1)(n_2h_2) = n_1(h_1n_2)h_2 = n_1(n_3h_3)h_2 = (n_1n_3)(h_3h_2) \in NH$$

(существование таких $n_3 \in N$ и $h_3 \in H$, что $h_1n_2 = n_3h_3$, вытекает из равенства $HN = NH$). Подгруппы $H \cap N \lhd H$ и $N \lhd HN$ нормальны, так как по условию $hNh^{-1} \subset N$ для всех $h \in H$. Отображение $\varphi\colon HN \to H/(H \cap N)$, переводящее произведение hn в смежный класс $h \cdot (H \cap N)$, определено корректно, поскольку при $h_1n_1 = h_2n_2$ элемент $h_1^{-1}h_2 = n_1n_2^{-1} \in H \cap N$, откуда $h_1 \cdot (H \cap N) = h_1 \cdot (h_1^{-1}h_2) \cdot (H \cap N) = h_2 \cdot (H \cap N)$. Оно сюръективно и является гомоморфизмом, поскольку $\varphi(h_1n_1h_2n_2) = \varphi(h_1h_2(h_2^{-1}n_1h_2)n_2) = h_1h_2 \cdot (H \cap N)$. Так как $\ker\varphi = eN = N$, по сл. 11.4 имеем $H/(H \cap N) = \operatorname{im}\varphi \simeq HN/\ker\varphi = HN/N$. □

Упражнение 11.30. Пусть $\varphi\colon G_1 \twoheadrightarrow G_2$ — сюръективный гомоморфизм групп. Покажите, что полный прообраз $N_1 = \varphi^{-1}(N_2)$ любой нормальной подгруппы $N_2 \lhd G_2$ является нормальной подгруппой в G_1 и $G_1/N_1 \simeq G_2/N_2$.

11.5.3. Геометрический смысл нормальности. Согласно предл. 11.4 и прим. 11.17 нормальность подгруппы $H \subset G$ равносильна наличию гомоморфизма $\varphi\colon G \to G'$ с ядром $H = \ker\varphi$. Если группа G' представлена как группа преобразований[2] какого-либо множества X, то возникает такое действие $G \to \operatorname{Aut} X$ исходной группы G на X, что H состоит из всех преобразований группы G, оставляющих на месте каждую точку X. Таким образом, нормальность подгруппы H означает наличие действия группы G на некоем множестве X с ядром H. Например, четвертная подгруппа Клейна $V_4 \subset S_4$ является ядром действия собственной группы куба на трёх отрезках, соединяющих центры противоположных граней.

[1] Т. е. $hNh^{-1} = N$ для всех $h \in H$.

[2] Как мы видели в прим. 11.12, такое представление всегда возможно.

§12. О строении групп

12.1. Свободные группы и соотношения. С любым множеством M можно связать группу F_M, которая называется *свободной группой*, порождённой множеством M. Она состоит из классов эквивалентных слов, которые можно написать буквами x и x^{-1}, где $x \in M$, по наименьшему отношению эквивалентности, отождествляющему между собою слова, отличающиеся друг от друга вставкой или удалением[1] двубуквенного фрагмента xx^{-1} или $x^{-1}x$. Композиция определяется как приписывание одного слова к другому. Единицей служит пустое слово. Обратным к классу слова $w = x_1 \dots x_M$ является класс слова $w^{-1} = x_m^{-1} \dots x_1^{-1}$, где каждая из букв x_i равна x или x^{-1}, где $x \in M$, и $\left(x^{-1}\right)^{-1} \stackrel{\text{def}}{=} x$.

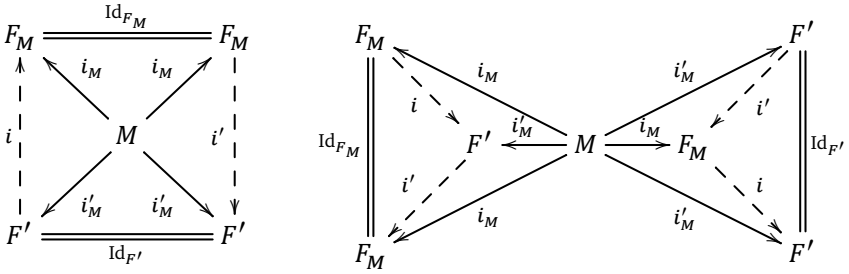
Упражнение 12.1. Убедитесь, что композиция корректно определена на классах эквивалентности слов и что в каждом классе содержится ровно одно *несократимое*[2] слово, которое одновременно является и самым коротким словом в своём классе.

Элементы множества M называются *образующими* свободной группы F_M. Свободная группа с k образующими обозначается F_k. Группа $F_1 \simeq \mathbb{Z}$ — это циклическая группа бесконечного порядка. Группа F_2 классов слов на четырёхбуквенном алфавите x, y, x^{-1}, y^{-1} уже трудно обозрима.

Упражнение 12.2. Постройте инъективный гомоморфизм групп $F_{\mathbb{N}} \hookrightarrow F_2$.

Предложение 12.1 (универсальное свойство свободных групп)
Отображение $i_M : M \to F_M$, переводящее элемент $x \in M$ в класс однобуквенного слова $x \in F_M$, обладает следующим свойством: для любых группы G и отображения множеств $\varphi_M : M \to G$ существует единственный такой гомоморфизм групп $\varphi : F_M \to G$, что $\varphi_M = \varphi \circ i_M$. Для любого обладающего этим свойством отображения $i'_M : M \to F'$ множества M в группу F' имеется единственный такой изоморфизм групп $i' : F_M \xrightarrow{\sim} F'$, что $i'_M = i' \circ i_M$.

Доказательство. Гомоморфизм φ единствен, так как обязан переводить слово $x_1^{\varepsilon_1} \dots x_m^{\varepsilon_m} \in F_M$, где $x_\nu \in M$, $\varepsilon_\nu = \pm 1$, в произведение $\varphi_M(x_1)^{\varepsilon_1} \dots \varphi_M(x_m)^{\varepsilon_m} \in G$. С другой стороны, это правило корректно задаёт гомоморфизм групп, что доказывает первое утверждение. Если отображение $i' : M \to F'$ множества M в группу F' обладает универсальным свойством из предл. 12.1, то существуют единственные гомоморфизмы $i' : F_M \to F'$ и $i : F' \to F_M$, встраивающиеся в коммутативные диаграммы



Разложения вида $i_M = \varphi \circ i_M$, $i'_M = \psi \circ i'_M$ в силу их единственности возможны только с $\varphi = \mathrm{Id}_{F_M}$, $\psi = \mathrm{Id}_{F'}$. Поэтому $i' \circ i = \mathrm{Id}_{F'}$, $i \circ i' = \mathrm{Id}_{F_M}$. □

[1] В начале, в конце, или же между произвольными двумя последовательными буквами слова.
[2] Т. е. не содержащее двубуквенных фрагментов xx^{-1} и $x^{-1}x$.

12.1.1. Задание групп образующими и соотношениями. Если гомоморфизм групп

$$\varphi : F_M \twoheadrightarrow G\,, \tag{12-1}$$

заданный отображением $\varphi_M : M \to G$ множества M в группу G, является *сюръективным*, то говорят, что группа G *порождается* элементами $g_m = \varphi_M(m)$, $m \in M$, а сами элементы g_m называются *образующими* группы G. В этом случае G исчерпывается всевозможными произведениями $g_1^{\varepsilon_1} g_2^{\varepsilon_2} \dots g_k^{\varepsilon_k}$, $\varepsilon = \pm 1$, образующих и обратных к ним элементов. Группа G называется *конечно порождённой*, если она допускает конечное множество образующих. Ядро $\ker\varphi \rtimes F_M$ эпиморфизма (12-1) называется *группой соотношений* между образующими g_m. Набор слов $R \subset \ker\varphi$ называется набором *определяющих соотношений*, если $\ker\varphi$ — это наименьшая нормальная подгруппа в F_M, содержащая R. Это означает, что любое соотношение можно получить из слов множества R конечным числом умножений, обращений и сопряжений произвольными элементами из свободной группы F_M. Группа, допускающая конечное число образующих с конечным набором определяющих соотношений называется *конечно определённой*.

Всякую группу можно задать образующими и соотношениями, например, взяв в качестве M множество всех элементов группы. Удачный выбор образующих с простыми определяющими соотношениями может значительно прояснить устройство группы и её гомоморфизмов в другие группы. Однако в общем случае выяснить, изоморфны ли две группы, заданные своими образующими и определяющими соотношениями, или даже определить, отлична ли группа, заданная образующими и соотношениями, от тривиальной группы $\{e\}$, бывает очень непросто. Более того, обе эти задачи являются *алгоритмически неразрешимыми*[1] даже в классе конечно определённых групп.

Предложение 12.2
Пусть группа G_1 задана множеством образующих M и набором определяющих соотношений R, а G_2 — произвольная группа. Отображение $\varphi : M \to G_2$ тогда и только тогда корректно задаёт гомоморфизм групп $G_1 \to G_2$ правилом $x_1^{\varepsilon_1}\dots x_m^{\varepsilon_m} \mapsto \varphi(x_1)^{\varepsilon_1}\dots\varphi(x_m)^{\varepsilon_m}$, когда для каждого слова $y_1^{\varepsilon_1}\dots y_m^{\varepsilon_m} \in R$ в группе G_2 выполняется соотношение $\varphi(y_1)^{\varepsilon_1}\dots\varphi(y_m)^{\varepsilon_m} = 1$.

Доказательство. Отображения множеств $\varphi_M : M \to G_2$ биективно соответствуют гомоморфизмам групп $\varphi : F_M \to G_2$. Такой гомоморфизм φ факторизуется до гомоморфизма из группы $G_1 = F_M/N_R$, где $N_R \rtimes F_M$ — наименьшая нормальная подгруппа, содержащая R, тогда и только тогда, когда $N_R \subset \ker\psi$. Так как $\ker\psi \rtimes F_M$, для этого необходимио и достаточно включения $R \subset \ker\psi$. □

Пример 12.1 (образующие и соотношения группы диэдра)
Покажем, что группа диэдра D_n задаётся двумя образующими x_1, x_2 и соотношениями

$$x_1^2 = x_2^2 = (x_1x_2)^n = e\,. \tag{12-2}$$

Оси симметрии правильного n-угольника разбивают его на $2n$ конгруэнтных прямоугольных треугольников как на рис. 12⋄1 ниже. Обозначим один из них через e. Поскольку любое движение плоскости однозначно задаётся своим действием на треугольник e, треугольники разбиения находятся в биекции с движениями $g \in D_n$, и каждый из них можно однозначно пометить

[1] В формальном смысле, принятом в математической логике.

тем единственным преобразованием g, которое переводит треугольник e в этот треугольник. При этом каждое преобразование $h \in D_n$ переводит каждый треугольник g в треугольник hg.

Упражнение 12.3. Для любого движения F евклидова пространства $\mathbb{R}^n$ и отражения σ_π в произвольной гиперплоскости $\pi \subset \mathbb{R}^n$ докажите соотношения

$$\sigma_{F(\pi)} = F \circ \sigma_\pi \circ F^{-1} \quad \text{и} \quad \sigma_{F(\pi)} \circ F = F \circ \sigma_\pi\,. \tag{12-3}$$

Обозначим через ℓ_1 и ℓ_2 боковые стороны треугольника e, а отражения плоскости в этих сторонах обозначим через $\sigma_1 = \sigma_{\ell_1}$ и $\sigma_2 = \sigma_{\ell_2}$. Тогда по второму из равенств (12-3) треугольники, получающиеся из e последовательными отражениями в направлении часовой стрелки пометятся элементами

$$\begin{aligned}
\sigma_{\ell_1} &= \sigma_1\,,\\
\sigma_{\sigma_1(\ell_2)}\sigma_1 &= \sigma_1\sigma_2\,,\\
\sigma_{\sigma_1\sigma_2(\ell_1)}\sigma_1\sigma_2 &= \sigma_1\sigma_2\sigma_1\,,\\
\sigma_{\sigma_1\sigma_2\sigma_1(\ell_2)}\sigma_1\sigma_2\sigma_1 &= \sigma_1\sigma_2\sigma_1\sigma_2\,,\ \ldots
\end{aligned}$$

а треугольники, получающиеся из e последовательными отражениями против часовой стрелки пометятся элементами

$$\begin{aligned}
\sigma_{\ell_2} &= \sigma_2\,,\\
\sigma_{\sigma_2(\ell_1)}\sigma_2 &= \sigma_2\sigma_1\,,\\
\sigma_{\sigma_2\sigma_1(\ell_2)}\sigma_2\sigma_1 &= \sigma_2\sigma_1\sigma_2\,,\\
\sigma_{\sigma_2\sigma_1\sigma_2(\ell_1)}\sigma_2\sigma_1\sigma_2 &= \sigma_2\sigma_1\sigma_2\sigma_1\,,\ \ldots
\end{aligned}$$

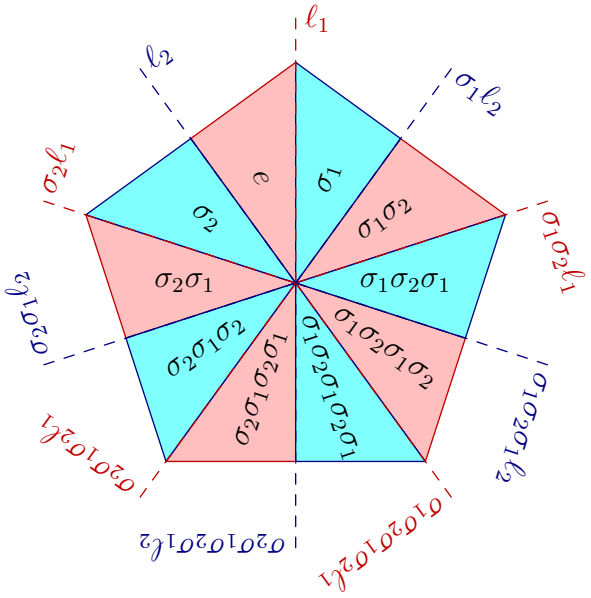


Рис. 12◇1. Образующие группы диэдра.

В результате каждый треугольник пометится словом вида $\sigma_1\sigma_2\sigma_1\sigma_2\ldots$ или $\sigma_2\sigma_1\sigma_2\sigma_1\ldots$ Так как композиция $\sigma_1 \circ \sigma_2$ является поворотом на угол $2\pi/n$, в группе D_n имеются соотношения

$$\sigma_1^2 = \sigma_2^2 = (\sigma_1\sigma_2)^n = e\,. \tag{12-4}$$

Последнее из них равносильно вытекающему из рис. 12◇1 равенству

$$\underbrace{\sigma_1\sigma_2\sigma_1\ldots}_{k} = \underbrace{\sigma_2\sigma_1\sigma_2\ldots}_{2n-k}\,. \tag{12-5}$$

Из сказанного вытекает, что правило $x_1 \mapsto \sigma_1$, $x_2 \mapsto \sigma_2$ корректно задаёт сюръективный гомоморфизм $\varphi\colon F_2/H \twoheadrightarrow D_n$ из фактора свободной группы F_2 с образующими x_1, x_2 по наименьшей нормальной подгруппе $H \rtimes F_2$, содержащей слова x_1^2, x_2^2 и $(x_1x_2)^n$. Каждое слово в алфавите $\{x_1, x_2\}$ по модулю соотношений (12-2) записывается содержащим меньше $2n$ букв словом $x_1x_2x_1\ldots$ или $x_2x_1x_2\ldots$, и два таких слова переводятся гомоморфизмом φ в один и тот же элемент $g \in D_n$ если и только если выполняется равенство (12-5), т. е. при

$$\underbrace{x_1x_2x_1\ldots}_{k} = \underbrace{x_2x_1x_2\ldots}_{2n-k}\,, \tag{12-6}$$

а это тождество является следствием тождества $(x_1x_2)^n = e$. Мы заключаем, что гомоморфизм $\varphi : F_2/H \twoheadrightarrow D_n$ биективен.

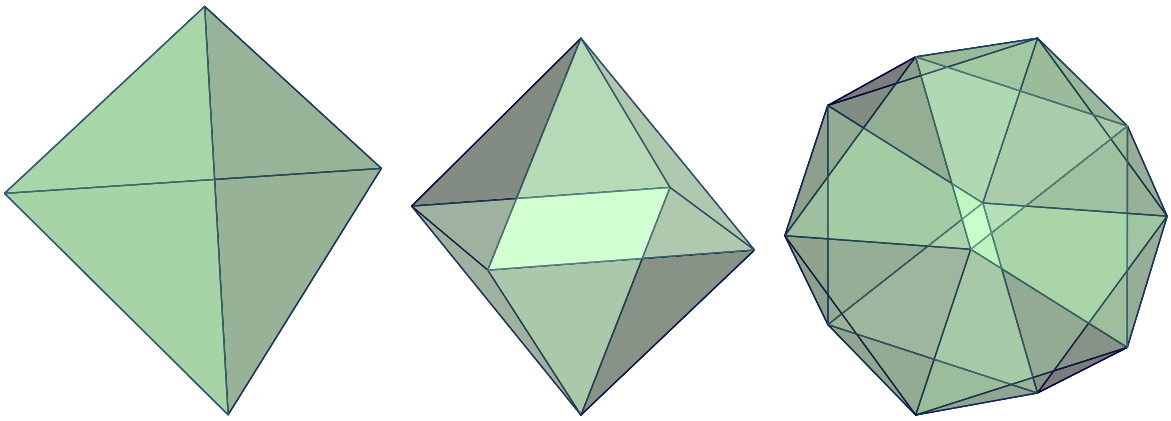


Рис. 12⋄2. Тетраэдр, октаэдр и икосаэдр.

Пример 12.2 (группы тетраэдра, октаэдра и икосаэдра)
Обозначим через M платоново тело с треугольными гранями, т. е. правильный *тетраэдр*, *октаэдр* или *икосаэдр* (см. рис. 12⋄2). Плоскости симметрии многогранника M задают *барицентрическое разбиение* каждой грани на 6 треугольников с вершинами в вершине M, в середине примыкающего к этой вершине ребра и центре примыкающей к этому ребру грани, как на рис. 12⋄3. Все эти треугольники конгруэнтны друг другу и сходятся по $2m_1 = 6$ штук в центрах граней, по $2m_2 = 4$ штуки в серединах рёбер и по $2m_3$ штук в вершинах, где числа m_i, а также число γ граней у M и общее число треугольников $N = 6\gamma$ представлены в таблице[1]:

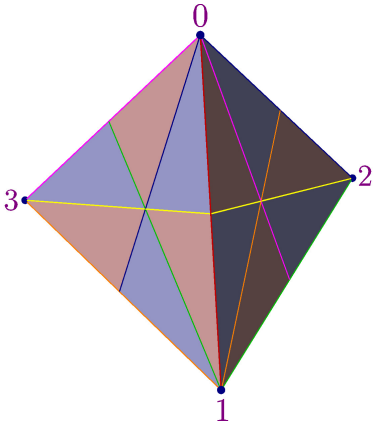


Рис. 12⋄3. Барицентрическое разбиение тетраэдра плоскостями симметрии.

M	m_1	m_2	m_3	γ	N
тетраэдр	3	2	3	4	24
октаэдр	3	2	4	8	48
икосаэдр	3	2	5	20	120.

Пометим один из этих треугольников буквой e и назовём высекающие его плоскости симметрии буквами π_1, π_2, π_3 так, чтобы для всех циклических перестановок (i, j, k) тройки индексов $(1, 2, 3)$ двугранный угол между плоскостями π_i и π_j равнялся π/m_k, и обозначим через σ_i отражение в плоскости π_i. Так как каждое преобразование из группы O_M однозначно определяется своим действием на тройку векторов с концами в углах треугольника e, каждый треугольник триангуляции является образом треугольника e при одном и ровно одном преобразовании $g \in \mathrm{O}_M$. Надпишем каждый треугольник тем преобразованием $g \in \mathrm{O}_M$, которое переводит в него треугольник e, и надпишем стороны треугольника g, высекаемые плоскостями $g(\pi_1)$, $g(\pi_2)$, $g(\pi_3)$ соответствующими номерами 1, 2, 3. Отметим, что каждое преобразование $h \in \mathrm{O}_M$ переводит каждый треугольник g в треугольник hg.

[1] Обратите внимание, что помещённый в пространство n-угольный диэдр из прим. 12.1 тоже можно включить в этот список со значениями $m_1 = n$, $m_2 = 2$, $m_3 = 2$, $\gamma = 2$ и $N = 4n$, если условиться, что плоский диэдр имеет две двумерные грани: «верхнюю» и «нижнюю».

На рис. 12⋄4 изображена стереографическая проекция картинки, которую 24 трёхгранных угла барицентрического разбиения тетраэдра с рис. 12⋄3 высекают на описанной около этого тетраэдра сфере. На каждом сферическом треугольнике написана композиция отражений $\sigma_1, \sigma_2, \sigma_3$, переводящая треугольник e в этот треугольник. Стороны треугольников, помеченные номерами 1, 2 и 3, изображены на рисунке в синим, зелёным и лиловым цветом.

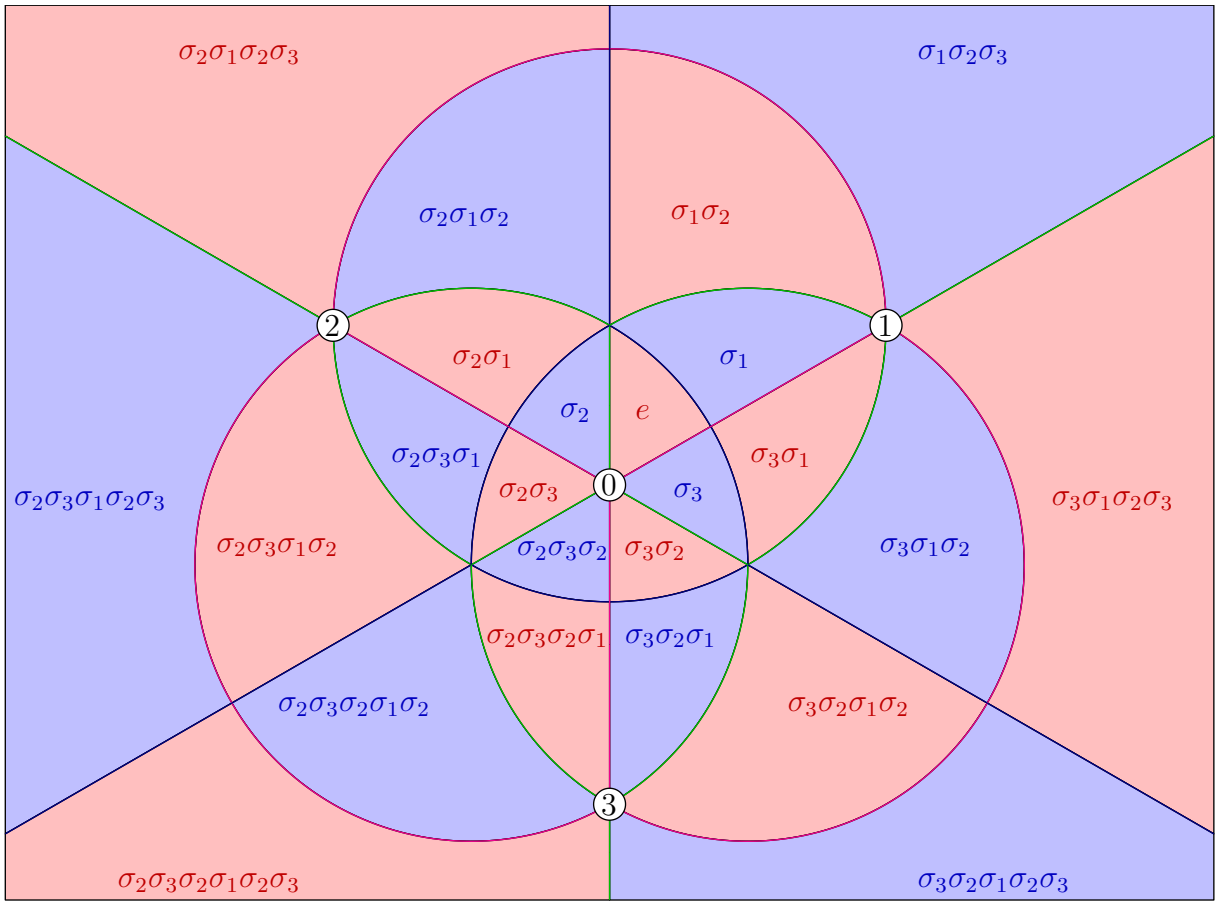


Рис. 12⋄4. Триангуляция описанной сферы плоскостями симметрии тетраэдра в стереографической проекции из диаметрально противоположного к вершине «0» полюса сферы на экваториальную плоскость, параллельную грани «123».

Чтобы явно написать композицию отражений $\sigma_1, \sigma_2, \sigma_3$, переводящую треугольник e в треугольник g, выберем внутри опирающихся на эти треугольники трёхгранных углов векторы u и w с концами на описанной вокруг M сфере так, чтобы $w \neq -u$ и натянутая на них плоскость Π_{uw} не содержала линий пересечения плоскостей симметрии многогранника M, и пройдём из u в w по кратчайшей дуге окружности, высекаемой на описанной сфере плоскостью Π_{uw}. Пусть мы при этом последовательно побываем в треугольниках $g_1 = e, g_2, g_3, \dots, g_{m+1} = g$. Обозначим через $\nu_i \in \{1, 2, 3\}$ номер, надписанный на той стороне треугольника g_i, сквозь которую осуществляется проход из g_i в g_{i+1}. Это означает, что общая сторона треугольников g_i и g_{i+1} высекается плоскостью $g_i(\pi_{\nu_i})$, т. е. образом плоскости π_{ν_i} при отображении g_i. По второму из равенств форм. (12-3) на стр. 171, $g_2 = \sigma_{\nu_1}$, $g_3 = \sigma_{g_2(\pi_{\nu_2})} g_2 = \sigma_{\nu_1}\sigma_{\nu_2}$, $g_4 = \sigma_{g_3(\pi_{\nu_3})} g_3 = \sigma_{\nu_1}\sigma_{\nu_2}\sigma_{\nu_3}$ и т. д. Таким образом, последовательность индексов $\nu_i \in \{1, 2, 3\}$ в разложении $g = \sigma_{\nu_1} \dots \sigma_{\nu_m}$

состоит из выписанных по порядку номеров сторон, которые приходится пересекать по пути из $e = g_1$ в $g = g_{m+1}$ по дуге uw, как на рис. 12⋄5, где стороны с номерами 1, 2, 3 изображены соответственно красным, зелёным и жёлтым цветами. Отметим, что полученное нами разложение элемента $g \in \mathrm{O}_M$ в композицию отражений $\sigma_1, \sigma_2, \sigma_3$ не единственно и зависит от выбора векторов u и w внутри трёхгранных углов e и g. При изменении любого из этих векторов последовательность $\nu_1, \ldots, \nu_m$ номеров зеркал, пересекаемых по дороге из u в w, не меняется до тех пор, пока натянутая на эти векторы плоскость Π_{uw} не натолкнётся на линию пересечения зеркал, а в момент пересечения такой линии в последовательности $\nu_1, \ldots, \nu_m$ некоторый фрагмент вида $\sigma_i\sigma_j\sigma_i\sigma_j \ldots$ длины m_k заменяется симметричным фрагментом $\sigma_j\sigma_i\sigma_j\sigma_i \ldots$ той же самой длины m_k, как показано на рис. 12⋄5.

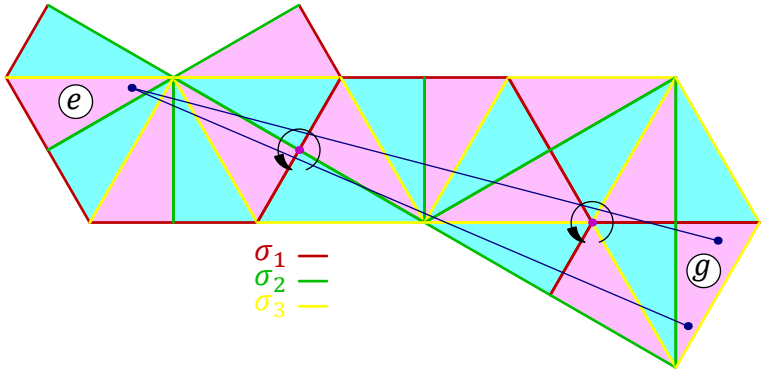


Рис. 12⋄5. $\sigma_2\sigma_3\sigma_2\sigma_3\boldsymbol{\sigma_1\sigma_2}\sigma_3\sigma_2\sigma_3\sigma_2\boldsymbol{\sigma_3\sigma_1\sigma_3}\sigma_2 = g = \sigma_2\sigma_3\sigma_2\sigma_3\boldsymbol{\sigma_2\sigma_1}\sigma_3\sigma_2\sigma_3\sigma_2\boldsymbol{\sigma_1\sigma_3\sigma_1}\sigma_2$.

Разложения, отвечающие верхней и нижней траекториям на рис. 12⋄5 отличаются друг от друга тем, что линии пересечения зеркал обходятся в противоположных направлениях. Композиции возникающих при этом отражений удовлетворяют соотношениям

$$\sigma_1\sigma_2 = \sigma_2\sigma_1 \quad \text{и} \quad \sigma_1\sigma_3\sigma_1 = \sigma_3\sigma_1\sigma_3$$

той же самой природы, что соотношения (12-4) в группе диэдра: так как композиция отражений $\sigma_i \circ \sigma_j$ является поворотом вокруг прямой $\pi_i \cap \pi_j$ на угол $2\pi/m_k$, равный удвоенному углу между плоскостями π_i и π_j, в группе O_M выполняются соотношения $\sigma_i^2 = e$ и $(\sigma_i\sigma_j)^{m_k} = e$, где $i = 1, 2, 3$, а тройка (i, j, k) пробегает три циклические перестановки номеров $(1, 2, 3)$. Отсюда вытекает, во-первых, что длина представления $g = \sigma_{\nu_1} \cdots \sigma_{\nu_m}$, считанного вдоль кратчайшей из двух дуг, соединяющих векторы u и w, не зависит от выбора этих векторов внутри трёхгранных углов, опирающихся на треугольники e и g, при условии, что плоскость Π_{uw} не проходит через линии пересечения зеркал, а во-вторых, что правило $x_i \mapsto \sigma_i$ задаёт сюръективный гомоморфизм $\varphi : F_3/H \twoheadrightarrow \mathrm{O}_M$ из фактора свободной группы F_3 на алфавите $\{x_1, x_2, x_3\}$ по наименьшей нормальной подгруппе $H \rtimes F_3$, содержащей шесть слов

$$x_i^2 \qquad \text{и} \qquad (x_ix_j)^{m_k}\,. \tag{12-7}$$

Для проверки того, что этот гомоморфизм является изоморфизмом, достаточно показать, что кратчайшее по модулю соотношений (12-7) представление каждого элемента $w \in F_3/H$ в виде $w = x_{\nu_1} \ldots x_{\nu_k}$ имеет в качестве набора индексов $\nu_1, \ldots \nu_k$ одну из возможных последовательностей номеров сторон, которые придётся пересечь, идя из треугольника e в треугольник $g = \sigma_{\nu_1} \ldots \sigma_{\nu_k}$ по дуге $[u, w]$, где $u \in e$, $w \in g$, так, как это объяснялось выше. Сделаем

это индукцией по длине k кратчайшего по модулю соотношений (12-7) слова $x_{\nu_1} \dots x_{\nu_k}$, представляющего данный элемент $y \in F_3 / H$. Для однобуквенных слов $y = x_1, x_2, x_3$ утверждение очевидно. Пусть оно верно для всех $y \in F_3 / H$, представимых словами из $\leqslant k$ букв. Рассмотрим произвольный такой y и проверим утверждение для всех элементов yx_j, $j = 1, 2, 3$, которые нельзя по модулю соотношений (12-7) записать словом из $\leqslant k$ букв. Пусть $g = \varphi(y)$ и $h = \varphi(yx_j) = g\sigma_j$. Рассмотрим плоскость $H = g(\pi_j)$.

Если треугольники e и g лежат по одну сторону от плоскости H, как на рис. 12⋄6, выберем векторы $u \in e$ и $w \in g$ так, чтобы продолжение дуги $[u, w]$ дальше за точку w уходило из треугольника g сквозь высекаемую плоскостью H сторону с номером j, и обозначим через v какой-нибудь вектор, лежащий в пересечении трёхгранного угла над треугольником h с продолжением дуги $[u, w]$. По предположению индукции в кратчайшем по модулю соотношений (12-7) представлении $y = x_{\nu_1} \dots x_{\nu_m}$ число букв $m \leqslant k$ и $\nu_1, \dots, \nu_m$ суть номера рёбер, которые приходится пересекать по пути из u в w по дуге $[u, w]$. При этом $h = \varphi(yx_j) = g\sigma_j = \sigma_{i_1} \dots \sigma_{i_m} \sigma_j$, и представление $yx_j = x_{\nu_1} \dots x_{\nu_m} x_j$ по нашему предположению состоит, как минимум, из $k + 1$ букв. Мы заключаем, что $m = k$, представление $yx_j = x_{\nu_1} \dots x_{\nu_k} x_j$ является одним из кратчайших для элемента yx_j и считывается с дуги $[u, v]$, как и требуется.

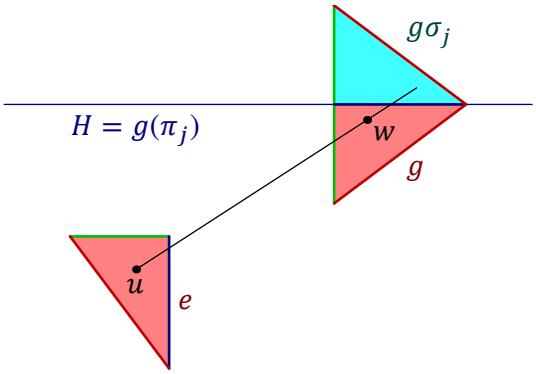


Рис. 12⋄6. H не разделяет e и g.

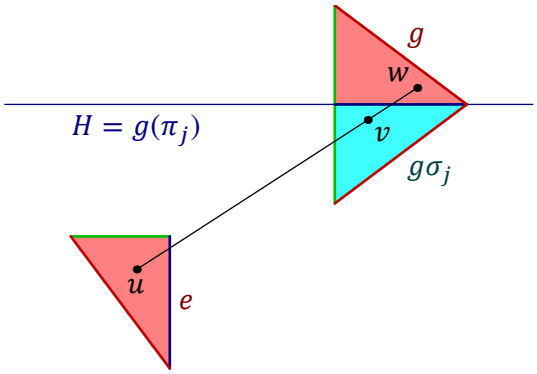


Рис. 12⋄7. H разделяет e и g.

Если треугольники e и g лежат по разные стороны от плоскости H, как на рис. 12⋄7, выберем вектор u в трёхгранном угле над e и вектор w в трёхгранном угле над g так, чтобы дуга $[u, w]$ входила в трёхгранный угол над треугольником g сквозь плоскость H, и обозначим через v какую-нибудь точку этой дуги, лежащую в трёхгранном угле над предыдущим треугольником $\sigma_{g(\pi_j)} g = g\sigma_j g^{-1} g = g\sigma_j = h$. По предположению индукции в кратчайшем по модулю соотношений (12-7) представлении $y = x_{\nu_1} \dots x_{\nu_m}$ число букв $m \leqslant k$ и $\nu_1, \dots, \nu_m$ суть номера рёбер, которые приходится пересекать по пути из u в w по дуге $[u, w]$. В частности, последняя буква $x_{\nu_m} = x_j$. Поэтому элемент $yx_j = x_{\nu_1} \dots x_{\nu_{m-1}}$ записывается более коротким словом, чем y, и утверждение для него верно по индуктивному предположению.

Итак, группа O_M платонова тела M с треугольными гранями порождается тремя элементами x_1, x_2, x_3, связанными шестью образующими соотношениями (12-7).

12.1.2. Образующие и соотношения симметрической группы S_{n+1}. Обозначим числами от 0 до n концы стандартных базисных векторов $e_0, e_1, \dots, e_n$ в $\mathbb{R}^{n+1}$ и рассмотрим n-мерный правильный симплекс $\Delta \subset \mathbb{R}^{n+1}$ с вершинами в этих точках. Поскольку каждое аффинное преобразование n-мерной гиперплоскости $x_0 + x_1 + \dots + x_n = 1$, в которой лежит симплекс Δ,

однозначно задаётся своим действием на вершины симплекса Δ, полная группа O_Δ симплекса Δ изоморфна симметрической группе S_{n+1} перестановок его вершин $0, 1, \dots, n$. Каждая k-мерная грань симплекса Δ является правильным k-мерным симплексом и представляет собою выпуклую оболочку каких-либо $k+1$ вершин симплекса Δ, и наоборот, выпуклая оболочка $[i_0, i_1, \dots, i_k]$ любых $k+1$ различных вершин $\{i_0, i_1, \dots, i_k\} \subset \{0, 1, \dots, n\}$ является k-мерной гранью симплекса Δ. Симплекс Δ симметричен относительно $n(n+1)/2$ гиперплоскостей π_{ij}, проходящих через середину ребра $[i, j]$ и противолежащую этому ребру грань коразмерности 2 с вершинами $\{0, 1, \dots, n\} \setminus \{i, j\}$. Гиперплоскость π_{ij} перпендикулярна вектору $e_i - e_j$ и отражение $\sigma_{ij} \in \mathrm{O}_\Delta$ в этой гиперплоскости отвечает транспозиции элементов i и j в симметрической группе S_{n+1}.

Упражнение 12.4. Убедитесь, что гиперплоскости π_{ij} и π_{km} с $\{i, j\} \cap \{k, m\} = \varnothing$ ортогональны, а плоскости π_{ij} и π_{jk} с различными i, j, k пересекаются под углом $\pi/3 = 60^\circ$.

Плоскости π_{ij} осуществляют *барицентрическое разбиение* симплекса Δ на $(n+1)!$ меньших симплексов с вершинами в центрах граней симплекса Δ и в центре самого симплекса. Если обозначить через $\langle i_0 i_1 \dots i_m \rangle$ центр m-мерной грани с вершинами в $i_0, i_1, \dots, i_m$, то каждый симплекс барицентрического разбиения будет иметь одну из вершин в какой-либо вершине $\langle i_0 \rangle$ симплекса Δ, следующую вершину — в центре $\langle i_0 i_1 \rangle$ какого-либо примыкающего к вершине i_0 ребра $[i_0, i_1]$, следующую вершину — в центре $\langle i_0 i_1 i_2 \rangle$ какой-либо примыкающей к ребру $[i_0, i_1]$ двумерной треугольной грани $[i_0, i_1, i_2]$ и т. д. вплоть до центра $\langle i_0 i_1 \dots i_n \rangle$ самого симплекса Δ. Таким образом, симплексы барицентрического разбиения симплекса Δ, осуществляемого гиперплоскостями π_{ij}, находятся в естественной биекции с перестановками $g \in S_{n+1}$: перестановке $g = (g_0, g_1, \dots, g_n) \in S_{n+1}$ отвечает симплекс с вершинами[1]

$$\langle g_0 \rangle, \langle g_0, g_1 \rangle, \langle g_0, g_1, g_2 \rangle, \dots, \langle g_0 g_1 \dots g_{n-1} \rangle, \langle g_0 g_1 \dots g_n \rangle . \tag{12-8}$$

Этот симплекс является образом начального симплекса

$$e = \big[\langle 0 \rangle, \langle 01 \rangle, \langle 012 \rangle, \dots, \langle 0, 1, \dots, n-1 \rangle, \langle 0, 1, \dots, n \rangle\big] \tag{12-9}$$

под действием ортогонального преобразования $g \in S_{n+1} = \mathrm{O}_M$. Как и выше, пометим каждый симплекс (12-8) соответствующим преобразованием g и спроектируем поверхность симплекса Δ из его центра на описанную сферу. Мы получим разбиение $(n-1)$-мерной сферы S^{n-1} на $(n+1)!$ надписанных элементами $g \in S_{n+1}$ попарно конгруэнтных $(n-1)$-мерных симплексов, грани которых высекаются из сферы гиперплоскостями π_{ij}. При $n = 3$ получится представленная на рис. 12⋄4 на стр. 173 триангуляция двумерной сферы S^2 двадцатью четырьмя сферическими треугольниками с углами $\pi/3$, $\pi/3$ и $\pi/2$. Помеченному тождественным преобразованием e начальному симплексу (12-9) отвечает сферический симплекс, высекаемый из сферы n гиперплоскостями $\pi_i \stackrel{\text{def}}{=} \pi_{i-1,i}$ с $1 \leqslant i \leqslant n$. Обозначим через $\sigma_i = \sigma_{i-1,i}$ отражения в этих гиперплоскостях. В симметрической группе S_{n+1} эти отражения суть транспозиции $|i-1, i\rangle$ пар соседних элементов. В силу упр. 12.4 они удовлетворяют соотношениям[2]

$$\sigma_i^2 = e\,, \quad \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \quad \text{и} \quad \sigma_i \sigma_j = \sigma_j \sigma_i\,, \quad \text{где} \quad |i-j| \geqslant 2\,. \tag{12-10}$$

[1] Первой вершиной служит вершина g_0 симплекса Δ, второй — середина выходящего из g_0 ребра $[g_0, g_1]$, третьей — центр примыкающей к этому ребру треугольной грани $[g_0, g_1, g_2]$ и т. д. вплоть до последней вершины, расположенной в центре симплекса Δ.

[2] Соотношение $\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$ является более употребительной в данном контексте записью циклического соотношения $(\sigma_i \sigma_{i+1})^3 = e$ на поворот $\sigma_i \sigma_{i+1}$ на 120° вокруг $(n-2)$-мерного подпространства $\pi_i \cap \pi_{i+1}$.

Упражнение 12.5. Убедитесь напрямую, что транспозиции $\sigma_i = |i-1, i\rangle \in S_{n+1}$ удовлетворяют соотношениям (12-10).

В силу этих соотношений, гомоморфизм свободой группы на алфавите $\{x_1, \dots, x_n\}$, переводящий x_i в σ_i, факторизуется до гомоморфизма $\varphi : F_n / H \to S_{n+1}$, где $H \rtimes F_n$ — наименьшая нормальная подгруппа, содержащая слова

$$x_i^2\,, \quad (x_i x_{i+1})^3 \quad \text{и} \quad (x_i x_j)^2\,, \quad \text{где} \quad |i-j| \geqslant 2\,. \tag{12-11}$$

Чтобы убедиться в его сюръективности, выберем в симплексах e и g точки a и b так, чтобы они не были диаметрально противоположны и соединяющая их геодезическая[1] не пересекала граней коразмерности[2] 2. Пройдя из a в b по этой геодезической, мы получим разложение

$$g = \sigma_{i_1} \dots \sigma_{i_m}\,, \tag{12-12}$$

в котором каждое $i_\nu \in \{1, \dots, n\}$ равно номеру того зеркала $g_\nu(\pi_\nu)$, через которое осуществляется переход из ν-того встреченного по дороге симплекса $g_\nu = \sigma_1 \dots \sigma_{\nu-1}$ в следующий симплекс $g_{\nu+1} = \sigma_{g_\nu(\pi_{i_\nu})} g_\nu = g_\nu \sigma_{i_\nu}$. Дословно также как и в прим. 12.2 проверяется, что длина представления (12-12), полученного с помощью дуги $[a, b]$ не зависит от выбора её концов $a \in e$ и $b \in g$ при условии, что они не диаметрально противоположны и плоскость π_{ab} не проходит через пересечения зеркал π_{ij}: если при перемещении точек a и b внутри симплексов e и g дуга $[a, b]$ пройдёт через грань коразмерности 2 вида $g_k(\pi_i \cap \pi_j)$ с $|i-j| \geqslant 2$, вдоль которой пересекаются перпендикулярные гиперграни $g_k(\pi_i)$, $g_k(\pi_j)$, или через грань вида $g_k(\pi_i \cap \pi_{i+1})$, вдоль которой под углом 60° пересекаются гиперграни $g_k(\pi_i)$, $g_k(\pi_{i+1})$, то в представлении $g = \sigma_1 \dots \sigma_m$ стоящий на k-том месте фрагмент $\sigma_i\sigma_j$ или $\sigma_i\sigma_{i+1}\sigma_i$ заменится, соответственно, равным ему в группе O_Δ фрагментом $\sigma_j\sigma_i$ или $\sigma_{i+1}\sigma_i\sigma_{i+1}$. В ортогональной проекции вдоль $(n-2)$-мерного подпространства $g_k(\pi_i \cap \pi_j)$ или $g_k(\pi_i \cap \pi_{i+1})$ на ортогональную ему двумерную плоскость мы при этом увидим картину вроде показанной на рис. 12⋄5 на стр. 174. Как и в прим. 12.2, индукция по длине кратчайшего представления элемента $w \in F_n / H$ показывает, что последовательность индексов $i_1, \dots, i_m$ в каждом кратчайшем по модулю соотношений (12-11) представлении $w = x_{i_1} \dots x_{i_m}$ совпадает с последовательностью индексов в представлении (12-12) элемента $g = \varphi(w) \in \mathrm{O}_\Delta$, полученном при помощи подходящей дуги $[a, b]$ с $a \in e$, $b \in g$. Таким образом, симметрическая группа S_{n+1} задаётся n образующими x_i, $1 \leqslant i \leqslant n$, связанными соотношениями (12-11).

Разумеется, эту геометрическую картину можно выхолостить до сугубо комбинаторного рассуждения, что мы сделаем в n° 12.1.3 ниже.

Упражнение 12.6. Покажите, что знакопеременная группа A_{n+1} порождается а) парами непересекающихтранспозиций б) 3-циклами $|k-2, k-1, k\rangle$, где $2 \leqslant k \leqslant n$.

12.1.3. Порядок Брюа на S_{n+1}. Будем называть количество всех инверсных пар[3] в перестановке $g = (g_0, g_1, \dots, g_n) \in S_{n+1}$ *длиной* перестановки g и обозначать его $\ell(g)$.

Упражнение 12.7. Убедитесь, что $0 \leqslant \ell(g) \leqslant n(n+1)/2$ для всех $g \in S_{n+1}$, причём имеется ровно по одной перестановке длин 0 и $n(n+1)/2$. Что это за перестановки?

[1] Кратчайшая из двух дуг ab большой окружности, высекаемой из сферы двумерной плоскостью, проходящей через точки a, b и центр сферы.

[2] Т. е. пересечений всевозможных пар зеркал π_{ij}.

[3] Напомню, пара (i, j), где $1 \leqslant i < j \leqslant n$ называется *инверсной парой* перестановки $g \in S_n$, если $g_i = g(i) > g(j) = g_j$, см. n° 8.1.2 на стр. 108.

Правое умножение перестановки g на транспозицию $\sigma_i = |i-1, i\rangle$ приводит к перестановке $g\sigma_i$, отличающейся от g транспозицией $(i-1)$-того и i-го символов g_{i-1} и g_i:

$$\big(g_0, \dots, g_{i-2}, \boldsymbol{g_{i-1}}, \boldsymbol{g_i}, g_{i+1}, \dots, g_n\big) \circ \sigma_i = \big(g_0, \dots, g_{i-2}, \boldsymbol{g_i}, \boldsymbol{g_{i-1}}, g_{i+1}, \dots, g_n\big)\,,$$

причём $\ell(g\sigma_i) = \ell(g) + 1$, если $g_{i-1} < g_i$, и $\ell(g\sigma_i) = \ell(g) - 1$, если $g_{i-1} > g_i$. Поэтому любая перестановка g длины $\ell(g) = m$ может быть записана словом $g = \sigma_{i_1} \cdots \sigma_{i_m}$, в котором каждый переход от перестановки $h = \sigma_{i_1} \cdots \sigma_{i_{k-1}} = (h_0, \dots, h_n)$ к перестановке $h\sigma_{i_k}$ заключается в транспозиции пары соседних возрастающих элементов $h_{i_\nu - 1} < h_{i_\nu}$. Частичный порядок на S_{n+1}, в котором $g < h$, если h получается из g увеличивающими длину транспозициями соседних элементов, называется *порядком Брюа*. Слово $w = x_{i_1} \cdots x_{i_m}$ в свободной группе F_n с образующими $x_1, \dots, x_n$ называется *минимальным словом* перестановки $g \in S_{n+1}$, если $m = \ell(g)$ и $g = \sigma_{i_1} \cdots \sigma_{i_m}$. Начальные фрагменты минимального слова задают строго возрастающую в смысле порядка Брюа последовательность элементов $h_\nu = \sigma_{i_1}\sigma_{i_2} \cdots \sigma_{i_\nu} \in S_{n+1}$. Перестановка g может иметь много разных минимальных слов, однако не может быть записана никаким более коротким словом.

Предложение 12.3
При гомоморфизме $\varphi : F_n \to S_{n+1}$, $x_i \mapsto \sigma_i$, каждое слово $w \in F_n$ эквивалентно минимальному слову перестановки $\varphi(w) \in S_{n+1}$ по модулю соотношений

$$x_i^2 = e\,, \quad x_i x_{i+1} x_i = x_{i+1} x_i x_{i+1} \quad \text{и} \quad x_i x_j = x_j x_i \text{ при } |i-j| \geqslant 2\,,$$

а все минимальные слова перестановки $\varphi(w)$ эквивалентны между собой.

Доказательство. Индукция по количеству букв в слове $w \in F_{n-1}$. Для $w = \varnothing$ утверждение очевидно. Пусть для всех слов из $\leqslant m$ букв предложение доказано. Достаточно для каждого m-буквенного слова w и каждой буквы x_ν проверить предложение для слова wx_ν. Если слово w не является минимальным словом элемента $g = \varphi(w)$, то по индукции оно эквивалентно более короткому минимальному слову. Тогда и wx_ν эквивалентно более короткому слову, и предложение справедливо по индукции. Поэтому мы будем далее считать, что слово w является минимальным словом элемента $g = \varphi(w) = (g_0, g_1, \dots, g_n)$. Возможны два случая: либо $g_{\nu-1} > g_\nu$, либо $g_{\nu-1} < g_\nu$. В первом случае у перестановки g есть минимальное слово вида ux_ν, по предположению индукции эквивалентное слову w. Тогда $wx_\nu \sim ux_\nu x_\nu \sim u$ и элемент $\varphi(wx_\nu) = \varphi(u)$ является образом более короткого, чем w слова u, эквивалентного слову wx_ν. По индукции, слово u эквивалентно минимальному слову элемента $\varphi(wx_\nu)$ и все такие слова эквивалентны друг другу. Поэтому то же верно и для эквивалентного u слова wx_ν.

Остаётся рассмотреть случай $g_{\nu-1} < g_\nu$. Здесь $\ell(g\sigma_\nu) = \ell(g) + 1$ и слово wx_ν является минимальным словом для элемента $\varphi(wx_\nu)$. Мы должны показать, что любое другое минимальное слово w' этого элемента эквивалентно wx_ν. Для самой правой буквы слова w' есть 3 возможности: либо она равна x_ν, либо она равна $x_{\nu\pm1}$ либо она равна x_μ с $|\mu - \nu| \geqslant 2$. В пером случае $w' = ux_\nu$, где u, как и w, является минимальным словом элемента g. По индукции $u \sim w$, а значит, и $w' = ux_k \sim wx_k$.

Пусть теперь $w' = ux_{\nu+1}$ — ситуация, когда $w' = ux_{\nu-1}$, полностью симметрична. Поскольку оба слова wx_ν и $ux_{\nu+1}$ минимальны для перестановки $h = \varphi(wx_\nu) = \varphi(ux_{\nu+1})$, в перестановке h на местах с номерами $\nu - 1$, ν, $\nu + 1$ стоят числа $g_\nu > g_{\nu-1} > g_{\nu+1}$, а в перестановке $g = (g_0, g_1, \dots, g_n) = \varphi(w)$ на этих же местах — числа $g_{\nu-1} < g_\nu > g_{\nu+1}$ с $g_{\nu-1} > g_{\nu+1}$. Поэтому

у перестановки h имеется минимальное слово вида $sx_{\nu+1}x_\nu x_{\nu+1}$, а у перестановки g — минимальное слово вида $tx_\nu x_{\nu+1}$. Перестановка $h' = \varphi(s) = \varphi(t)$ отличается от h тем, что числа на местах с номерами $\nu-1$, ν, $\nu+1$ в ней возрастают и равны $g_{\nu+1} < g_{\nu-1} < g_\nu$. Поскольку $\ell(h') = \ell(h)-3 = \ell(g)-2$, оба слова t и s минимальны для h' и по индукции эквивалентны. Кроме того, по индукции w эквивалентно $tx_\nu x_{\nu+1}$. Поэтому $wx_\nu \sim tx_\nu x_{\nu+1}x_\nu \sim sx_\nu x_{\nu+1}x_\nu \sim sx_{\nu+1}x_\nu x_{\nu+1}$. Но $sx_{\nu+1}x_\nu \sim u$, поскольку оба слова минимальны для одной и той же перестановки[1] длины $m = \ell(h) - 1$. Таким образом, $wx_\nu \sim ux_{\nu+1}$.

Наконец, пусть $h = \varphi(wx_\nu) = \varphi(ux_\mu)$, где $|\mu - \nu| \geqslant 2$. Тогда в h есть два непересекающихся фрагмента $g_{\nu-1} > g_\nu$ и $g_{\mu-1} > g_\mu$. Поэтому у h есть минимальные слова вида $tx_\mu x_\nu$ и вида $sx_\nu x_\mu$, где t и s являются минимальными словами для перестановки $\varphi(t) = \varphi(s)$, отличающейся от h тем, что расматриваемые 2 фрагмента в ней имеют вид $g_\nu < g_{\nu-1}$ и $g_\mu < g_{\mu-1}$. Так как длина этой перестановки равна $\ell(h) - 2 = m - 1$, по индукции $t \sim s$. Поскольку tx_μ — минимальное слово для g, по индукции $w \sim tx_\mu$. Аналогично, т. к. sx_ν и u — минимальные слова для перестановки $\varphi(sx_\nu) = \varphi(u)$, отличающейся от h' транспозицией первого из двух фрагментов и потому имеющей длину $\ell(h) - 1 = m$, по индукции $sx_\nu \sim u$. Таким образом, $wx_\nu \sim tx_\mu x_\nu \sim sx_\mu x_\nu \sim sx_\nu x_\mu \sim ux_\mu$, что и требовалось. □

12.2. Простые группы и композиционные факторы. Группа G называется *простой*, если она не содержит нормальных подгрупп, отличных от $\{e\}$ и G. Например, любая группа простого порядка проста, поскольку по теореме Лагранжа вообще не содержит никаких подгрупп кроме $\{e\}$ и G. Согласно сл. 11.1 на стр. 157 простота группы G равносильна тому, что всякий гомоморфизм $G \to G'$ либо является вложением, либо отображает всю группу G в единицу.

Определение 12.1 (композиционный ряд)
Конечная строго убывающая последовательность подгрупп

$$G = G_0 \supsetneq G_1 \supsetneq G_2 \supsetneq \cdots \supsetneq G_{n-1} \supsetneq G_n = \{e\} \tag{12-13}$$

называется *композиционным рядом* или *рядом Жордана–Гёльдера* группы G, если при каждом i подгруппа G_{i+1} нормальна в G_i и фактор G_i / G_{i+1} прост. В этой ситуации неупорядоченный набор простых групп G_i / G_{i+1} (в котором возможны повторения) называется набором *композиционных факторов* (или *факторов Жордана–Гёльдера*) группы G. Число n называется *длиной* композиционного ряда (12-13).

Пример 12.3 (композиционные факторы S_4)
Выше мы видели, что симметрическая группа S_4 имеет композиционный ряд

$$S_4 \rhd A_4 \rhd V_4 \rhd \mathbb{Z}/(2) \rhd \{e\}\,,$$

в котором $A_4 \rtimes S_4$ — подгруппа чётных перестановок, $V_4 \rtimes A_4$ — подгруппа Клейна, состоящая из тождественной перестановки и трёх перестановок циклового типа $\square\!\square$, а

$$\mathbb{Z}/(2) \rtimes V_4 \simeq \mathbb{Z}/(2) \oplus \mathbb{Z}/(2)$$

любая из трёх циклических подгрупп второго порядка, порождённых неединичными элементами. Таким образом, симметрическая группа S_4 имеет композиционные факторы $\mathbb{Z}/(2) = S_4/A_4$, $\mathbb{Z}/(3) = A_4/V_4$, $\mathbb{Z}/(2) = V_4/\big(\mathbb{Z}/(2)\big)$ и $\mathbb{Z}/(2) = \mathbb{Z}/(2)/\{e\}$.

[1] Она отличается от g, h и h' тем, что числа в позициях с номерами $\nu-1$, ν, $\nu+1$ в ней упорядочены как $g_\nu > g_{\nu+1} < g_{\nu-1}$, где $g_\nu > g_{\nu-1}$.

Упражнение 12.8. Убедитесь, что $A_4/V_4 \simeq \mathbb{Z}/(3)$.

Теорема 12.1 (теорема Жордана – Гёльдера)
Если группа G имеет конечный композиционный ряд, то неупорядоченный набор его композиционных факторов не зависит от выбора композиционного ряда. В частности, все композиционные ряды имеют одинаковую длину.

Доказательство. Пусть у группы G есть два композиционных ряда

$$G = P_0 \supsetneq P_1 \supsetneq P_2 \supsetneq \cdots \supsetneq P_{n-1} \supsetneq P_n = \{e\} \tag{12-14}$$

$$G = Q_0 \supsetneq Q_1 \supsetneq Q_2 \supsetneq \cdots \supsetneq Q_{m-1} \supsetneq Q_m = \{e\}\,. \tag{12-15}$$

Мы собираемся вставить между последовательными членами этих рядов дополнительные цепочки нестрого убывающих подгрупп так, чтобы получившиеся удлинённые последовательности состояли из одинакового числа элементов, и построить между последовательными факторами полученных цепочек такую биекцию, что соответствующие друг другу факторы будут изоморфны. Применяя предл. 11.5 на стр. 168 к нормальной подгруппе $P_{i+1} \lhd P_i$ и подгруппам $Q_\nu \cap P_i \subset P_i$, мы для каждого i получаем цепочку

$$P_i \supseteq (Q_1 \cap P_i)P_{i+1} \supseteq (Q_2 \cap P_i)P_{i+1} \supseteq \cdots \supseteq (Q_{m-1} \cap P_i)P_{i+1} \supseteq P_{i+1}\,, \tag{12-16}$$

которая начинается с P_i, кончается в P_{i+1} и имеет $(Q_{k+1} \cap P_i)P_{i+1} \lhd (Q_k \cap P_i)P_{i+1}$ с

$$\frac{(Q_k \cap P_i)P_{i+1}}{(Q_{k+1} \cap P_i)P_{i+1}} \simeq \frac{(Q_k \cap P_i)}{(Q_{k+1} \cap P_i)(Q_k \cap P_{i+1})}\,. \tag{12-17}$$

Упражнение 12.9. Для любой четвёрки подгрупп A, B, C, D, в которой $A \lhd B$ и $C \lhd D$, постройте изоморфизм $(B \cap D)C/(A \cap D)C \simeq (B \cap D)/(A \cap D)(B \cap C)$.

Группа P_{i+1} является нормальной подгруппой во всех группах цепочки (12-16). Факторизуя по ней, получаем цепочку

$$\frac{P_i}{P_{i+1}} \supseteq \frac{(Q_1 \cap P_i)P_{i+1}}{P_{i+1}} \supseteq \frac{(Q_2 \cap P_i)P_{i+1}}{P_{i+1}} \supseteq \cdots \supseteq \frac{(Q_{m-1} \cap P_i)P_{i+1}}{P_{i+1}} \supseteq \{e\}\,, \tag{12-18}$$

в которой каждая подгруппа нормальна в предыдущей, а последовательные факторы

$$\frac{(Q_k \cap P_i)P_{i+1}/P_{i+1}}{(Q_{k+1} \cap P_i)P_{i+1}/P_{i+1}} \simeq \frac{(Q_k \cap P_i)P_{i+1}}{(Q_{k+1} \cap P_i)P_{i+1}} \simeq \frac{(Q_k \cap P_i)}{(Q_{k+1} \cap P_i)(Q_k \cap P_{i+1})}$$

совпадают с (12-17). Так как группа P_i/P_{i+1} проста, мы заключаем, что в цепочке (12-18) имеется ровно одно нестрогое включение, а все остальные включения — равенства. Тем самым, ровно один из факторов (12-17) отличен от единицы и изоморфен P_i/P_{i+1}.

Те же самые рассуждения с заменой P на Q позволяют вставить между последовательными группами $Q_k \rhd Q_{k+1}$ композиционного ряда (12-15) убывающую цепочку подгрупп

$$Q_k \supseteq (P_1 \cap Q_k)Q_{k+1} \supseteq (P_2 \cap Q_k)Q_{k+1} \supseteq \cdots \supseteq (P_{n-1} \cap Q_k)Q_{k+1} \supseteq Q_{k+1}\,, \tag{12-19}$$

каждая из которых нормальна в предыдущей, а последовательные факторы имеют вид

$$\frac{(P_i \cap Q_k)Q_{k+1}}{(P_{i+1} \cap Q_k)Q_{k+1}} \simeq \frac{(Q_k \cap P_i)}{(Q_{k+1} \cap P_i)(Q_k \cap P_{i+1})} \tag{12-20}$$

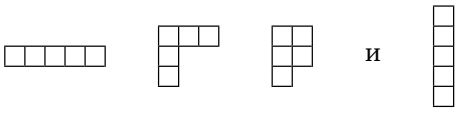
и изоморфны соответствующим факторам (12-17). Таким образом, вставляя между последовательными элементами композиционного ряда (12-14) цепочки (12-16), а между последовательными элементами ряда (12-15) — цепочки (12-19), мы получим цепочки одинаковой длины, в которых не все включения строгие, однако факторы которых биективно соответствуют друг другу так, что соответственные факторы (12-20) и (12-17) изоморфны. Остаётся заметить, что группа Q_{k+1} является нормальной подгруппой во всех группах цепочки (12-19), и то же рассуждение, что и с подгруппой P_{i+1} для цепочки (12-16), показывает, что при фиксированном k среди факторов (12-20) имеется ровно один отличный от единицы, и он изоморфен Q_k/Q_{k+1}. □

Замечание 12.1. Непростая группа может иметь несколько разных композиционных рядов с одинаковым набором факторов, а группы с одинаковыми наборами факторов Жордана-Гёльдера не обязательно изоморфны.

12.2.1. Конечные простые группы. Одним из крупных достижений математики XX века было создание полного списка всех конечных простых групп. Этот список состоит из нескольких бесконечных серий и 26 так называемых *спорадических групп*, не входящих в серии. Бесконечные серии делятся на три семейства: циклические группы $\mathbb{Z}/(p)$ простого порядка, знакопеременные группы A_n с[1] $n \geqslant 5$ и простые линейные алгебраические группы над конечными полями[2], такие как $\mathrm{PSL}_n(\mathbb{F}_q)$, $\mathrm{PSO}_n(\mathbb{F}_q)$, $\mathrm{PSp}_n(\mathbb{F}_q)$ и т. п. Эта классификация является итогом сотен работ десятков авторов по множеству напрямую несвязанных друг с другом направлений. Последние пробелы в ней, как принято считать, были устранены лишь в 2008 году. Какая-либо универсальная концепция, позволяющая единообразно классифицировать все конечные простые группы до сих пор не известна. Далее мы обсудим простоту знакопеременных групп.

Лемма 12.1
Знакопеременная группа A_5 проста.

Доказательство. В симметрической группе две перестановки сопряжены тогда и только тогда, когда у них одинаковый цикловой тип. Цикловые типы чётных перестановок из S_5 изображаются диаграммами



(12-21)

(5-циклы, 3-циклы, пары независимых транспозиций и тождественное преобразование). Эти классы сопряжённости в S_5 имеют мощность

$$5!/5 = 24 \qquad 5!/(3\cdot 2) = 20 \qquad 5!/(2^2\cdot 2) = 15 \qquad \text{и} \qquad 1\,.$$

Если перестановка относится к одному из последних трёх типов (12-21), то её централизатор содержит транспозицию пары неподвижных элементов или пары элементов, составляющих цикл длины 2. Поэтому две такие перестановки, сопряжённые в S_5, сопряжены и в A_5. Стало быть, перестановки каждого из трёх последних типов (12-21) образуют один класс сопряжённости

[1] Группа $A_3 \simeq \mathbb{Z}/(3)$ тоже проста.

[2] Описание и классификация таких групп даются в курсах линейных алгебраических и арифметических групп; представление о них можно получить по книге *Дж. Хамфри. Линейные алгебраические группы*. М., «Наука», 1980.

также и в A_5. Циклы длины 5 разбиваются в A_5 на два класса сопряжённости: 12 циклов, сопряжённых $|1,2,3,4,5\rangle$, и 12 циклов, сопряжённых $|2,1,3,4,5\rangle$. Поскольку любая нормальная подгруппа $H \rtimes A_5$ вместе с каждой перестановкой содержит и все ей сопряжённые,

$$|H| = 12\varepsilon_1 + 12\varepsilon_2 + 20\varepsilon_3 + 15\varepsilon_4 + 1\,,$$

где каждый из коэффициентов ε_k равен либо 1, либо 0. С другой стороны, $|H|$ является делителем $|A_5| = 60 = 3\cdot 4\cdot 5$.

Упражнение 12.10. Убедитесь, что такое возможно ровно в двух случаях: когда все $\varepsilon_k = 1$ или когда все $\varepsilon_k = 0$.

Таким образом, нормальные подгруппы в A_5 исчерпываются единичной подгруппой и всей группой A_5. □

Теорема 12.2
Все знакопеременные группы A_n с $n > 5$ тоже просты.

Доказательство. Индукция по n. Стабилизатор $\mathrm{Stab}_{A_n}(k)$ любого элемента $k \in \{1, 2, \dots, n\}$ изоморфен A_{n-1}. Если $N \rtimes A_n$, то пересечение $N \cap \mathrm{Stab}_{A_n}(k) \rtimes \mathrm{Stab}_{A_n}(k)$ по индукции либо совпадает со $\mathrm{Stab}_{A_n}(k)$ либо равно $\{e\}$. Поскольку стабилизаторы всех элементов сопряжены, подгруппа N либо содержит стабилизаторы всех элементов $1, 2, \dots, n$, либо тривиально пересекается с каждым из них. В первом случае N содержит все пары транспозиций и, стало быть, совпадает с A_n по упр. 12.6. Во втором случае если в N есть хоть одна перестановка, переводящая некое i в $j \neq i$, то в силу тривиальности $\mathrm{Stab}_N(j)$ эта перестановка является *единственной* в N перестановкой, переводящей i в j. Но при $n \geqslant 6$ у любой перестановки $g \in A_n$, переводящей i в j и не имеющей неподвижных точек, есть сопряжённые ей в A_n и отличные от неё перестановки, также переводящие i в j.

Упражнение 12.11. Убедитесь в этом.

Поскольку N нормальна, все эти перестановки тоже лежат в N. Противоречие. □

12.3. Полупрямые произведения. Для пары подгрупп N, H группы G положим

$$NH = \{xh \mid x \in N\,,\ h \in H\}\,.$$

Отображение $N \times H \to NH$, $(x,h) \mapsto xh$, биективно если и только если $N \cap H = \{e\}$. В самом деле, при $x_1h_1 = x_2h_2$ элемент $x_2^{-1}x_1 = h_2h_1^{-1} \in N \cap H$, и если $N \cap H = \{e\}$, то $x_2 = x_1$ и $h_2 = h_1$, а если в $N \cap H$ есть элемент $z \neq e$, то разные пары (e,e), $(z,z^{-1}) \in N \times H$ перейдут в один и тот же элемент $e \in NH$.

Будем называть подгруппы $N, H \subset G$ *дополнительными*, если $N \cap H = \{e\}$ и $NH = G$. В этом случае группа G как множество находится в биекции с прямым произведением $N \times H$. Если подгруппа $N \rtimes G$ при этом нормальна, то композиция элементов $g_1 = x_1h_1$ и $g_2 = x_2h_2$ может быть выражена в терминах пар (x_1,h_1), $(x_2,h_2) \in N \times H$. А именно, так как

$$g_1g_2 = x_1h_1x_2h_2 = x_1(h_1x_2h_1^{-1})\cdot h_1h_2 \quad \text{и} \quad h_1x_2h_1^{-1} \in N\,,$$

группу G можно *описать* как множество $N \times H$ с операцией

$$(x_1,h_1)\cdot(x_2,h_2) = (x_1\,\mathrm{Ad}_{h_1}(x_2), h_1h_2)\,, \tag{12-22}$$

где через $\mathrm{Ad}_h : N \xrightarrow{\sim} N$, $x \mapsto hxh^{-1}$, обозначено присоединённое действие элемента h на нормальной подгруппе N. В этой ситуации говорят, что группа G является *полупрямым произведением* нормальной подгруппы $N \rtimes G$ и дополнительной к ней подгруппы $H \subset G$ и пишут $G = N \rtimes H$. Если сопряжение элементами из подгруппы H действует на подгруппе N тривиально, что равносильно перестановочности $xh = xh$ любых двух элементов $x \in N$ и $h \in H$, то полупрямое произведение называется *прямым*. В этом случае

$$(x_1, h_1) \cdot (x_2, h_2) = (x_1 x_2, h_1 h_2)$$

для любых пар (x_1, h_1), $(x_2, h_2) \in N \times H$.

Пример 12.4 ($D_n = \mathbb{Z}/(n) \rtimes \mathbb{Z}/(2)$)
Группа диэдра D_n содержит нормальную подгруппу поворотов, изоморфную аддитивной группе $\mathbb{Z}/(n)$. Подгруппа второго порядка, порождённая любым отражением, дополнительна к группе поворотов и изоморфна аддитивной группе $\mathbb{Z}/(2)$. Присоединённое действие отражения на группе поворотов меняет знак у угла поворота. При отождествлении группы поворотов с $\mathbb{Z}/(n)$ это действие превращается в умножение на -1. Таким образом, $D_n = \mathbb{Z}/(n) \rtimes \mathbb{Z}/(2)$ и в терминах пар $(x, y) \in \mathbb{Z}/(n) \times \mathbb{Z}/(2)$ композиция на группе диэдра задаётся правилом

$$\big(x_1, y_1\big) \cdot \big(x_2, y_2\big) = \big(x_1 + (-1)^{y_1} x_2, y_1 + y_2\big)\,, \quad x_1, x_2 \in \mathbb{Z}/(n)\,, \quad y_1, y_2 \in \mathbb{Z}/(2)\,.$$

Пример 12.5 ($\mathrm{Aff}(V) = V \rtimes \mathrm{GL}(V)$, продолжение прим. 11.20 на стр. 166)
Аффинная группа[1] $\mathrm{Aff}(V)$ содержит нормальную подгруппу параллельных переносов, которая изоморфна аддитивной группе векторного пространства V и является ядром сюръективного гомоморфизма групп

$$D : \mathrm{Aff}(V) \twoheadrightarrow \mathrm{GL}(V)\,, \quad \varphi \mapsto D_\varphi\,, \tag{12-23}$$

сопоставляющего аффинному преобразованию $\varphi : \mathbb{A}(V) \to \mathbb{A}(V)$ его дифференциал

$$D_\varphi : V \to V\,, \quad \overrightarrow{pq} \mapsto \overrightarrow{\varphi(p)\varphi(q)}\,.$$

Если зафиксировать в $\mathbb{A}(V)$ какую-нибудь точку p, то ограничение гомоморфизма (12-23) на стабилизатор $\mathrm{Stab}_p \subset \mathrm{Aff}(V)$ задаст изоморфизм $D_p : \mathrm{Stab}_p \xrightarrow{\sim} \mathrm{GL}(V)$. Обратный изоморфизм сопоставляет линейному оператору $f : V \xrightarrow{\sim} V$ аффинное преобразование

$$\varphi_f : \mathbb{A}(V) \to \mathbb{A}(V)\,, \quad x \mapsto p + f(\overrightarrow{px})\,,$$

оставляющее на месте точку p. Поскольку каждое преобразование $\varphi \in \mathrm{Aff}(V)$ раскладывается в композицию $\varphi = \tau_v \circ (\tau_{-v} \circ \varphi)$ параллельного переноса τ_v на вектор $v = \overrightarrow{p\varphi(p)}$ и преобразования $\tau_{-v} \circ \varphi \in \mathrm{Stab}(p)$, группа $\mathrm{Aff}(V) = V \rtimes \mathrm{Stab}_p \simeq V \rtimes \mathrm{GL}(V)$. Согласно прим. 11.20 на стр. 166, композиция в группе $V \rtimes \mathrm{GL}(V)$ задаётся правилом $(u, f) \cdot (w, g) = (u + f(w),\, fg)$.

12.3.1. Полупрямое произведение групп. Предыдущую конструкцию можно применить к двум абстрактным группам N и H как только задано действие группы H на группе N, т. е. гомоморфизм группы H в группу автоморфизмов группы N:

$$\psi : H \to \mathrm{Aut}\, N\,, \quad h \mapsto \psi_h : N \xrightarrow{\sim} N\,, \tag{12-24}$$

[1] См. прим. 11.20 на стр. 166.

По аналогии с форм. (12-22) на стр. 182 зададим на множестве $N \times H$ операцию правилом

$$(x_1, h_1) \cdot (x_2, h_2) \stackrel{\text{def}}{=} (x_1 \psi_{h_1}(x_2), h_1 h_2)\,. \tag{12-25}$$

Упражнение 12.12. Проверьте, что формула (12-25) задаёт на $N \times H$ структуру группы с единицей (e, e) и обращением $(x, h)^{-1} = \left(\psi_h^{-1}(x^{-1}), h^{-1}\right)$, где $\psi_h^{-1} = \psi_{h^{-1}}$ — автоморфизм, обратный к $\psi_h : N \xrightarrow{\sim} N$.

Полученная таким образом группа называется *полупрямым произведением* групп N и H по действию $\psi\colon N \to \operatorname{Aut} H$ и обозначается $N \rtimes_\psi H$. Подчеркнём, что результат зависит от выбора действия ψ. Если действие тривиально, т. е. $\psi_h = \mathrm{Id}_N$ для всех $h \in H$, мы получаем прямое произведение $N \times H$ с покомпонентными операциями.

Упражнение 12.13. Убедитесь, что подмножество $N' \stackrel{\text{def}}{=} \{(x, e) \mid x \in N\}$ является изоморфной группе N нормальной подгруппой в $G = N \rtimes_\psi H$ и фактор $G / N' \simeq H$, а подмножество $H' \stackrel{\text{def}}{=} \{(e, h) \mid h \in H\}$ (e, h) является изоморфной H и дополнительной к N' подгруппой в G, причём $G = N' \rtimes H'$ является полупрямым произведением своих подгрупп N' и H'.

12.4. p-группы и теоремы Силова. Группа порядка p^n, где $p \in \mathbb{N}$ — простое, называется *p-группой*. Поскольку все подгруппы p-группы также являются p-группами, длина любой орбиты p-группы при любом её действии на любом множестве либо делится на p, либо равна единице. Мы получаем простое, но полезное

Предложение 12.4
Пусть p-группа G действует на конечном множестве X, число элементов в котором не делится на p. Тогда G имеет на X неподвижную точку. □

Предложение 12.5
Любая p-группа имеет нетривиальный центр.

Доказательство. Рассмотрим присоединённое действие группы на себе. Центр группы представляет собой множество неподвижных точек этого действия. Поскольку и число элементов в группе, и длины всех орбит, содержащих более одной точки, делятся на p, кроме одноточечной орбиты e должны быть и другие одноточечные орбиты. □

Упражнение 12.14. Покажите, что любая группа G порядка p^2 (где p простое) абелева.

Определение 12.2 (силовские подгруппы)
Пусть G — произвольная конечная группа. Запишем её порядок в виде $|G| = p^n m$, где p — простое, $n \geqslant 1$, и m взаимно просто с p. Всякая подгруппа $S \subset G$ порядка $|S| = p^n$ называется *силовской p-подгруппой* в G. Количество силовских p-подгрупп в G обозначается через $N_p(G)$.

Теорема 12.3 (теорема Силова)
Для любого простого p, делящего $|G|$, силовские p-подгруппы в G существуют. Все они сопряжены друг другу, и любая p-подгруппа в G содержится в некоторой силовской p-подгруппе.

Доказательство. Пусть $|G| = p^n m$, где m взаимно просто с p. Обозначим через $\mathcal{E}$ множество p^n-элементных подмножеств в G и рассмотрим действие G на $\mathcal{E}$, индуцированное левым регулярным действием G на себе. Стабилизатор точки $F \in \mathcal{E}$ состоит из всех элементов $g \in G$, левое умножение на которые переводит множество $F \subset G$ в себя: $\operatorname{Stab}(F) = \{g \in G \mid gF \subset F\}$. Так

как $g_1 x \neq g_2 x$ при $g_1 \neq g_2$ в группе G, группа $\mathrm{Stab}(F)$ свободно действует на множестве F и все орбиты этого действия состоят из $|\mathrm{Stab}(F)|$ точек. Поэтому $|F| = p^n$ делится на $|\mathrm{Stab}(F)|$ и имеется следующая альтернатива: либо длина G-орбиты элемента $F \in \mathcal{E}$ делится на p, либо G-орбита элемента $F \in \mathcal{E}$ состоит из m элементов и $|\mathrm{Stab}(F)| = p^n$, т. е. подгруппа $\mathrm{Stab}(F) \subset G$ силовская. Во втором случае согласно предл. 12.4 каждая p-подгруппа $H \subset G$ (в частности, каждая силовская подгруппа), имеет на G-орбите элемента F неподвижную точку gF, а значит, содержится в силовской подгруппе $\mathrm{Stab}(gF) = g\,\mathrm{Stab}(F)\,g^{-1}$, сопряжённой к $\mathrm{Stab}(F)$ (и совпадает с ней, если H силовская). Таким образом, для доказательства теоремы остаётся убедиться, что в множестве $\mathcal{E}$ есть G-орбита, длина которой не делится на p. Это вытекает из следующей ниже леммы. □

Лемма 12.2
$|\mathcal{E}| = \binom{p^n m}{p^n} \equiv m \pmod p$ не делится на p.

Доказательство. Класс вычетов $\binom{p^n m}{p^n} \pmod p$ равен коэффициенту при x^{p^n}, возникающему при раскрытии бинома $(1+x)^{p^n m}$ над полем $\mathbb{F}_p = \mathbb{Z}/(p)$. Так как возведение в p-тую степень над $\mathbb{F}_p$ является аддитивным гомоморфизмом, $(1+x)^{p^n} = 1+x^{p^n}$, откуда $(1+x)^{p^n m} = \left(1 + x^{p^n}\right)^m = 1 + m x^{p^n} +$ старшие степени. □

Следствие 12.1 (дополнение к теореме Силова)
В условиях теоремы Силова число N_p силовских p-подгрупп в G делит m и сравнимо с единицей по модулю p.

Доказательство. Обозначим множество силовских p-подгрупп в G через $\mathcal{S}$ и рассмотрим действие G на $\mathcal{S}$, индуцированное присоединённым действием G на себе. По теореме Силова это действие транзитивно, откуда $|\mathcal{S}| = |G|/|\mathrm{Stab}(P)|$, где $P \in \mathcal{S}$ — произвольно взятая силовская p-подгруппа. Поскольку $P \subset \mathrm{Stab}(P)$, порядок $|\mathrm{Stab}(P)|$ делится на $|P| = p^n$, а значит $|\mathcal{S}|$ делит $|G|/p^n = m$, что доказывает первое утверждение.

Для доказательства второго утверждения достаточно проверить, что P, действуя сопряжениями на $\mathcal{S}$, имеет там ровно одну неподвижную точку, а именно, саму себя. Тогда порядки всех остальных P-орбит будут делиться на p, и мы получим $|\mathcal{S}| \equiv 1 \pmod p$.

Пусть силовская подгруппа $H \in \mathcal{S}$ неподвижна при сопряжении подгруппой P. Это означает, что $P \subset \mathrm{Stab}(H) = \{g \in G \mid gHg^{-1} \subset H\}$. Поскольку $H \subset \mathrm{Stab}(H) \subset G$, порядок $|\mathrm{Stab}(H)| = p^n m'$, где $m'|m$ взаимно просто с p. Таким образом, и P, и H являются силовскими p-подгруппами в $\mathrm{Stab}(H)$, причём H нормальна в $\mathrm{Stab}(H)$. Так как все силовские подгруппы сопряжены, мы заключаем, что $H = P$, что и требовалось. □

Пример 12.6 (группы порядка pq с простыми $p > q$)
Пусть $|G| = pq$, где $p > q$ простые. Тогда в G есть ровно одна силовская p-подгруппа $H_p \simeq \mathbb{Z}/(p)$, автоматически нормальная. Рассмотрим любую силовскую q-подгруппу $H_q \simeq \mathbb{Z}/(q)$. Поскольку H_p и H_q просты, $H_p \cap H_q = e$ и $G = H_p H_q$. Согласно n° 12.3 $G = \mathbb{Z}/(p) \rtimes_\psi \mathbb{Z}/(q)$ для некоторого гомоморфизма $\psi : \mathbb{Z}/(q) \to \mathrm{Aut}(\mathbb{Z}/(p))$.

Упражнение 12.15. Убедитесь, что $\mathrm{Aut}(\mathbb{Z}/(p)) \simeq \mathbb{F}_p^* \simeq \mathbb{Z}/(p-1)$.

Гомоморфизм $\psi : \mathbb{Z}/(q) \to \mathrm{Aut}(\mathbb{Z}/(p))$ однозначно задаётся своим значением на образующей $[1]_q$, которая является элементом порядка q. Поэтому элемент $\eta = \psi([1]_q) \in \mathrm{Aut}(\mathbb{Z}/(p))$ либо единичный, либо имеет порядок q. По упр. 12.15 последнее возможно только при $q \mid (p-1)$,

и в этом случае элементы q-го порядка образуют в $\mathbb{F}_p^*$ циклическую мультипликативную подгруппу порядка q.

Упражнение 12.16. Убедитесь в этом.

Обозначим через $\eta \in \mathbb{F}_p^*$ одну из образующих этой подгруппы. Гомоморфизм

$$\psi : \mathbb{Z}/(p) \to \mathrm{Aut}(\mathbb{Z}/(p)), \quad [1]_q \mapsto \eta, \tag{12-26}$$

сопоставляет каждому элементу $[y]_q \in \mathbb{Z}/(q)$ автоморфизм $\psi_y : \mathbb{Z}/(p) \xrightarrow{\sim} \mathbb{Z}/(p)$, $[x]_p \mapsto [\eta^y x]_p$, и задаёт полупрямое произведение $\mathbb{Z}/(p) \rtimes_\psi \mathbb{Z}/(q)$ с операцией

$$\bigl([x_1]_p, [y_1]_q\bigr) \cdot \bigl([x_2]_p, [y_2]_q\bigr) = \bigl([x_1 + \eta^{y_1} x_2]_p, [y_1 + y_2]_q\bigr). \tag{12-27}$$

Любой другой гомоморфизм $\varphi : \mathbb{Z}/(q) \to \mathrm{Aut}(\mathbb{Z}/(p))$, $[1]_q \mapsto \eta^m$, с $1 \leqslant m \leqslant q-1$ является композицией гомоморфизма (12-26) и умножения на $m : \mathbb{Z}/(q) \xrightarrow{\sim} \mathbb{Z}/(q)$, $[y]_q \mapsto [my]_q$. Согласно упр. 12.17 ниже, полупрямые произведения $\mathbb{Z}/(p) \rtimes_\varphi \mathbb{Z}/(q)$ и $\mathbb{Z}/(p) \rtimes_\psi \mathbb{Z}/(q)$ изоморфны.

Упражнение 12.17. Для гомоморфизма $\psi : H \to \mathrm{Aut}(N)$, $h \mapsto \psi_h$, и автоморфизмов $\alpha : H \xrightarrow{\sim} H$ и $\beta : N \xrightarrow{\sim} N$ убедитесь, что отображения $(n, h) \mapsto (n, \alpha^{-1}h)$ и $(n, h) \mapsto (\beta n, h)$ задают, соответственно, изоморфизмы полупрямых произведений

$$N \rtimes_\psi H \xrightarrow{\sim} N \rtimes_{\psi\circ\alpha} H \quad \text{и} \quad N \rtimes_\psi H \xrightarrow{\sim} N \rtimes_{\mathrm{Ad}_\beta(\psi)} H,$$

где $\mathrm{Ad}_\beta(\psi) : H \to \mathrm{Aut}(N)$, $h \mapsto \beta\psi_h\beta^{-1}$.

Мы заключаем, что для простых $p > q$ при $q \nmid (p-1)$ группа порядка pq изоморфна $\mathbb{Z}/(p) \oplus \mathbb{Z}/(q)$, а при $q \mid (p-1)$ кроме абелевой есть ровно одна неабелева группа $\mathbb{Z}/(p) \rtimes \mathbb{Z}/(q)$ с операцией (12-27). В частности, для простого $p > 2$ имеется единственная с точностью до изоморфизма неабелева группа порядка $2p$, а именно, группа правильного p-угольника из прим. 12.4 на стр. 183.

§13. Пространство с билинейной формой

13.1. Билинейные формы. Отображение $\beta : V \times V \to \Bbbk$ называется *билинейной формой* на векторном пространстве V, если оно линейно по каждому из двух своих аргументов при фиксированном другом, т. е. удовлетворяет равенству

$$\beta(x_1u_1 + x_2u_2, y_1w_1 + y_2w_2) = \sum_{i,j=1}^{2} x_iy_j\beta(u_i, w_j) \tag{13-1}$$

при всех $u_1, u_2, w_1, w_2 \in V$ и $x_1, x_2, y_1, y_2 \in \Bbbk$.

Упражнение 13.1. Убедитесь, что билинейные формы образуют векторное подпространство в пространстве всех функций $V \times V \to \Bbbk$.

Если форма β на пространстве V зафиксирована, то её значение $\beta(u, w) \in \Bbbk$ на паре векторов $u, w \in V$ иногда бывает удобно записывать в виде *скалярного произведения* $u \cdot w$, принимающего значения в поле $\Bbbk$ и, вообще говоря, некоммутативного. В таких обозначениях формула (13-1) утверждает, что это произведение дистрибутивно по отношению к линейным комбинациям векторов, т. е. подчиняется стандартным правилам раскрытия скобок:

$$(x_1u_1 + x_2u_2) \cdot (y_1w_1 + y_2w_2) = \sum_{i,j=1}^{2} x_iy_j\, u_i \cdot w_j\,.$$

13.1.1. Матрицы Грама. В пространстве с билинейной формой с любыми двумя наборами векторов $\boldsymbol{u} = (u_1, \dots, u_n)$, $\boldsymbol{w} = (w_1, \dots w_m)$, где все $u_i, w_j \in V$, связана матрица их попарных скалярных произведений $B_{\boldsymbol{uw}} \stackrel{\text{def}}{=} \boldsymbol{u}^t \cdot \boldsymbol{w} \in \mathrm{Mat}_{n\times m}(\Bbbk)$ с элементами $b_{ij} = v_i \cdot w_j = \beta(u_i, w_j)$. Она называется *матрицей Грама* наборов $\boldsymbol{u}$, $\boldsymbol{w}$ и формы β. Когда наборы $\boldsymbol{u} = \boldsymbol{w}$ совпадают, вместо $B_{\boldsymbol{uu}}$ пишут просто $B_{\boldsymbol{u}}$. В этом случае $\det B_{\boldsymbol{u}} \in \Bbbk$ называется *определителем Грама* формы β и набора векторов $\boldsymbol{u}$. Если наборы векторов $\boldsymbol{u}$ и $\boldsymbol{w}$ линейно выражаются через наборы $\boldsymbol{e}$ и $\boldsymbol{f}$ по формулам $\boldsymbol{u} = \boldsymbol{e}\, C_{\boldsymbol{eu}}$ и $\boldsymbol{w} = \boldsymbol{f}\, C_{\boldsymbol{fw}}$, то $B_{\boldsymbol{uw}} = \boldsymbol{u}^t\boldsymbol{w} = \big(\boldsymbol{e}C_{\boldsymbol{eu}}\big)^t \big(\boldsymbol{f}C_{\boldsymbol{fw}}\big) = C_{\boldsymbol{eu}}^t\boldsymbol{e}^t\boldsymbol{f}C_{\boldsymbol{fw}} = C_{\boldsymbol{eu}}^t B_{\boldsymbol{ef}} C_{\boldsymbol{fw}}$. В частности, если $\boldsymbol{u} = \boldsymbol{w}\, C_{\boldsymbol{wu}}$, то

$$B_{\boldsymbol{u}} = C_{\boldsymbol{wu}}^t B_{\boldsymbol{w}} C_{\boldsymbol{wu}}\,. \tag{13-2}$$

Например, если векторы $\boldsymbol{e} = (e_1, \dots e_n)$ образуют базис в V, а векторы $u = \boldsymbol{e}\, x$ и $w = \boldsymbol{e}\, y$ заданы столбцами $x, y \in \Bbbk^n$ своих координат в этом базисе, то

$$\beta(u, w) = u^t \cdot w = x^t\boldsymbol{e}^t \cdot \boldsymbol{e}y = x^t B_{\boldsymbol{e}} y\,. \tag{13-3}$$

Так как любая квадратная матрица $B_{\boldsymbol{e}} \in \mathrm{Mat}_n(\Bbbk)$ задаёт по этой формуле билинейную форму на пространстве V, сопоставление билинейной форме её матрицы Грама в произвольно зафиксированном базисе устанавливает биекцию между пространством билинейных форм на n-мерном векторном пространстве V и пространством матриц размера $n \times n$.

Упражнение 13.2. Убедитесь, что эта биекция линейна.

13.1.2. Корреляции. Задание билинейной формы $\beta : V \times V \to \Bbbk$ эквивалентно заданию линейного отображения *правой корреляции* $\beta^\wedge : V \to V^*$, сопоставляющего каждому вектору $v \in V$ линейный функционал $\beta^\wedge v : V \to \Bbbk$, который задаётся правым скалярным умножением на вектор v и является ограничением билинейного отображения $\beta : V \times V \to \Bbbk$ на подмножество $V \times \{v\} \subset V \times V$:

$$\beta^\wedge v : V \to \Bbbk\,, \quad u \mapsto u \cdot v = \beta(u, v)\,. \tag{13-4}$$

Упражнение 13.3. Убедитесь, что для каждого $v \in V$ функционал (13-4) линеен и линейно зависит от v.

Форма $\beta : V \times V \to \Bbbk$ однозначно восстанавливается по правой корреляции $\beta^\wedge : V \to V^*$ как

$$\beta(u, w) = \beta^\wedge w(u)\,.$$

Если зафиксировать в V и V^* двойственные базисы $\boldsymbol{e} = (e_1, \dots e_n)$ и $\boldsymbol{e}^* = (e_1^*, \dots e_n^*)$, то в этих базисах матрица $B^\wedge_{\boldsymbol{e}^*\boldsymbol{e}}$ линейного отображения $\beta^\wedge : V \to V^*$ имеет в клетке (i, j) значение i-той координаты функционала $\beta^\wedge e_j : u \mapsto \beta(u, e_j)$ в базисе $\boldsymbol{e}^*$, которая равна значению этого функционала на базисном векторе e_i, т. е. скалярному произведению $\beta(e_i, e_j)$. Таким образом, матрица правой корреляции $B^\wedge_{\boldsymbol{e}^*\boldsymbol{e}} = B_{\boldsymbol{e}}$ совпадёт с матрицей Грама формы β в базисе $\boldsymbol{e}$. Мы заключаем, что сопоставление билинейной форме β её правой корреляции $\beta^\wedge$ устанавливает линейный изоморфизм пространства билинейных форм на V с пространством линейных отображений $V \to V^*$.

Симметричным образом, задание билинейной формы $\beta : V \times V \to \Bbbk$ эквивалентно заданию *левой корреляции* ${}^\wedge\beta : V \to V^*$, которая переводит каждый вектор $v \in V$ в линейный функционал ${}^\wedge\beta v : V \to \Bbbk$, получающийся ограничением отображения $\beta : V \times V \to \Bbbk$ на подмножество $\{v\} \times V \subset V \times V$ и задаваемый левым скалярным умножением на вектор v:

$${}^\wedge\beta v : V \to \Bbbk\,, \quad u \mapsto \beta(v, u) = v \cdot u\,. \tag{13-5}$$

Иначе можно сказать, что левая корреляция билинейной формы β является правой корреляцией для *транспонированной* формы $\beta^t(u, w) \stackrel{\text{def}}{=} \beta(w, u)$, матрица Грама которой транспонирована к матрице Грама формы β. Поэтому матрица левой корреляции билинейной формы β в двойственных базисах $\boldsymbol{e}$ и $\boldsymbol{e}^*$ пространств V и V^* равна транспонированной матрице Грама $B^t_{\boldsymbol{e}}$ формы β в базисе $\boldsymbol{e}$.

13.1.3. Ядра, ранг и коранг. Векторные пространства

$$\begin{aligned} V^\perp &= \ker \beta^\wedge = \{u \in V \mid \forall\, v \in V\ \beta(v, u) = 0\} \\ {}^\perp V &= \ker {}^\wedge\beta = \{u \in V \mid \forall\, v \in V\ \beta(u, v) = 0\} \end{aligned} \tag{13-6}$$

называются соответственно *правым* и *левым* ядром билинейной формы β. Если форма β не является симметричной или кососимметричной[1], то подпространства $V^\perp$ и ${}^\perp V$, вообще говоря, различны. Тем не менее, их размерности всегда одинаковы и равны

$$\dim V^\perp = \dim {}^\perp V = \dim V - \operatorname{rk} B_{\boldsymbol{e}}\,, \tag{13-7}$$

где $B_{\boldsymbol{e}}$ — матрица Грама формы β в произвольном базисе $\boldsymbol{e}$ пространства V. В самом деле, так как транспонированные матрицы имеют одинаковый ранг, размерности образов $\operatorname{im}\beta^\wedge$ и $\operatorname{im}{}^\wedge\beta$ операторов правой и левой корреляций, равные, соответственно, $\operatorname{rk} B_{\boldsymbol{e}}$ и $\operatorname{rk} B^t_{\boldsymbol{e}}$, совпадают, а значит, совпадают и $\dim\ker\beta^\wedge = \dim V - \dim\operatorname{im}\beta^\wedge$ и $\dim\ker{}^\wedge\beta = \dim V - \dim\operatorname{im}{}^\wedge\beta$. Из сказанного вытекает, что ранг матрицы Грама $B_{\boldsymbol{e}}$, равный размерности образа каждой из корреляций, не зависит от выбора базиса. Он называется *рангом* билинейной формы β и обозначается $\operatorname{rk}\beta$, а разность $\dim V - \operatorname{rk}\beta$ из формулы (13-7) называется *корангом* формы β и обозначается $\operatorname{cork}\beta$.

[1] Т. е. не удовлетворяет соотношениям $\beta^t = \pm\beta$. Мы подробнее поговорим о таких формах в n° 13.4 на стр. 196 ниже.

13.1.4. Изометрии. Линейное отображение $f : U \to W$ между векторными пространствами U и W, на которых заданы билинейные формы β и γ, называется *изометрическим* или *гомоморфизмом пространств с билинейными формами*, если для любых векторов $u_1, u_2 \in U$ выполняется равенство $\beta(u_1, u_2) = \gamma\big(f(u_1), f(u_2)\big)$. Билинейные формы β и γ называются *изоморфными*, если между пространствами U и W имеется изометрический линейный изоморфизм.

Если произвольно зафиксировать в U и W базисы $\boldsymbol{u} = (u_1, \dots, u_n)$ и $\boldsymbol{w} = (w_1, \dots, w_m)$, то отображение f с матрицей $F_{\boldsymbol{wu}}$ в этих базисах является изометрическим если и только если матрица Грама набора векторов $f(\boldsymbol{u}) = (f(u_1), \dots, f(u_n)) = \boldsymbol{w}\, F_{\boldsymbol{wu}}$ равна матрице Грама базиса $\boldsymbol{u}$. По форм. (13-2) на стр. 187 это равносильно матричному равенству

$$F_{\boldsymbol{wu}}^t B_{\boldsymbol{w}} F_{\boldsymbol{wu}} = B_{\boldsymbol{u}}\,. \tag{13-8}$$

13.2. Невырожденные формы. Билинейная форма β называется *невырожденной*[1], если она удовлетворяет условиям следующего ниже предл. 13.1. Формы, не удовлетворяющие этим условиям, называются *вырожденными* или *особыми*.

Предложение 13.1 (критерии невырожденности)
Следующие свойства билинейной формы β на конечномерном векторном пространстве V равносильны друг другу:

1) в V существует базис с ненулевым определителем Грама
2) любой базис в V имеет ненулевой определитель Грама
3) левая корреляция $^\wedge\beta : V \xrightarrow{\sim} V^*$ является изоморфизмом
4) правая корреляция $\beta^\wedge : V \xrightarrow{\sim} V^*$ является изоморфизмом
5) для любого ненулевого вектора $v \in V$ существует такой вектор $u \in V$, что $\beta(v, u) \neq 0$
6) для любого ненулевого вектора $v \in V$ существует такой вектор $u \in V$, что $\beta(u, v) \neq 0$
7) для любой линейной функции $\varphi : V \to \Bbbk$ существует такой вектор $v \in V$, что
$$\varphi(u) = \beta(v, u) \quad \text{для всех } u \in V$$
8) для любой линейной функции $\varphi : V \to \Bbbk$ существует такой вектор $v \in V$, что
$$\varphi(u) = \beta(u, v) \quad \text{для всех } u \in V\,,$$

причём при выполнении этих условий вектор v в последних двух пунктах определяется формой φ однозначно.

Доказательство. Поскольку $\dim V = \dim V^*$, биективность, инъективность и сюръективность линейного отображения $V \to V^*$ равносильны друг другу и тому, что это отображение задаётся невырожденной матрицей в каких-нибудь базисах. Поэтому условия (3), (5), (7) и условия (4), (6), (8), утверждающие, соответственно, биективность, обращение в нуль ядра и сюръективность для операторов $^\wedge\beta$ и $\beta^\wedge$, равносильны между собой и условию (1), означающему, что

[1] А также *неособой* или *регулярной*.

транспонированные друг другу матрицы этих операторов обратимы. Условие (1) равносильно условию (2) в силу форм. (13-2) на стр. 187, из которой вытекает, что определители Грама двух базисов $\boldsymbol{e}$ и $\boldsymbol{f}$ связаны друг с другом по формуле $\det B_{\boldsymbol{e}} = \det B_{\boldsymbol{f}} \cdot \det^2 C_{\boldsymbol{fe}}$, где $C_{\boldsymbol{fe}}$ — матрица перехода[1] от базиса $\boldsymbol{e}$ к базису $\boldsymbol{f}$. □

Пример 13.1 (евклидова форма)
Симметричная билинейная форма на координатном пространстсве $\Bbbk^n$ с единичной матрицей Грама E в стандартном базисе называется *евклидовой*. Эта форма невырождена и над полем $\Bbbk = \mathbb{R}$ задаёт евклидову структуру на пространстве $\mathbb{R}^n$. Однако над отличными от $\mathbb{R}$ полями свойства этой формы могут существенно отличаться от интуитивно привычных свойств евклидовой структуры. Например, над полем $\mathbb{C}$ ненулевой вектор $e_1 - ie_2 \in \mathbb{C}^2$ имеет нулевой скалярный квадрат.

Упражнение 13.4. Приведите пример n-мерного подпространства в $\mathbb{C}^{2n}$, на которое евклидова форма ограничивается в тождественно нулевую форму.

Базисы, в которых матрица Грама евклидовой формы равна E называются *ортонормальными*. Ниже[2] мы увидим, что над алгебраически замкнутым полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ любая невырожденная симметричная билинейная форма изометрически изоморфна евклидовой.

Пример 13.2 (гиперболическое пространство H_{2n})
Симметричная билинейная форма h на чётномерном координатном пространстве $H_{2n} = \Bbbk^{2n}$, матрица Грама которой в стандартном базисе равна

$$H = \begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}, \tag{13-9}$$

где E — единичная матрица размера $n \times n$, называется *гиперболической*. Она невырождена и над алгебраически замкнутым полем изометрически изоморфна евклидовой форме: ортонормальный базис гиперболической формы состоит из векторов

$$\varepsilon_{2\nu-1} = \bigl(e_\nu - e_{n+\nu}\bigr)/\sqrt{-2} \quad \text{и} \quad \varepsilon_{2\nu} = \bigl(e_\nu + e_{n+\nu}\bigr)/\sqrt{2}, \quad 1 \leqslant \nu \leqslant n.$$

Над полями $\mathbb{R}$ и $\mathbb{Q}$ гиперболическая форма не изоморфна евклидовой, поскольку евклидовы скалярные квадраты всех ненулевых векторов положительны, тогда как ограничение гиперболической формы на линейную оболочку первых n базисных векторов тождественно нулевое. Базис, в котором матрица Грама гиперболической формы имеет вид (13-9), называется *гиперболическим базисом*.

Пример 13.3 (симплектическое пространство Ω_{2n})
Кососимметричная форма на чётномерном координатном пространстве $\Omega_{2n} = \Bbbk^{2n}$, матрица Грама которой в стандартном базисе равна

$$J = \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}, \tag{13-10}$$

где E — единичная матрица размера $n \times n$, называется *симплектической*. Матрица J вида (13-10) называется *симплектической единицей*. Она имеет $J^2 = -E$ и $\det J = 1$. Таким образом, симплектическая форма невырождена. Базис, в котором матрица Грама кососимметричной формы равна J, называется *симплектическим базисом*. Ниже[3] мы покажем, что всякая невырожденная

[1] См. n° 7.1.2 на стр. 96.
[2] См. сл. 13.1 на стр. 197.
[3] См. теор. 13.5 на стр. 199.

кососимметричная билинейная форма над *любым* полем изометрически изоморфна симплектической. Это означает, в частности, что размерность пространства с невырожденной кососимметричной формой обязательно чётна.

Упражнение 13.5. Убедитесь в том, что все кососимметричные квадратные матрицы нечётного размера над полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ вырождены.

13.2.1. Левый и правый двойственный базис. Если билинейная форма β на пространстве V невырождена, то у любого базиса $\boldsymbol{e} = (e_1, \dots e_n)$ в V есть *правый* и *левый* двойственные базисы $\boldsymbol{e}^\vee = (e_1^\vee, \dots, e_n^\vee)$ и ${}^\vee\boldsymbol{e} = ({}^\vee e_1, \dots, {}^\vee e_n)$, состоящие из прообразов векторов двойственного к $\boldsymbol{e}$ базиса $\boldsymbol{e}^* = (e_1^*, \dots, e_n^*)$ в V^* относительно изоморфизмов правой и левой корреляций соответственно. Они однозначно характеризуются соотношениями ортогональности

$$\beta\left(e_i, e_j^\vee\right) = \beta\left({}^\vee e_i, e_j\right) = \delta_{ij}\,, \tag{13-11}$$

которые на матричном языке означают, что взаимные матрицы Грама двойственных относительно формы β базисов единичные: $B_{\boldsymbol{e}\boldsymbol{e}^\vee} = B_{{}^\vee\boldsymbol{e}\boldsymbol{e}} = E$. Согласно формулам из n° 13.1.1 матрицы переходов $C_{\boldsymbol{e},\boldsymbol{e}^\vee}$ и $C_{\boldsymbol{e},{}^\vee\boldsymbol{e}}$, в j-тых столбцах которых стоят координаты векторов $e_j^\vee$ и ${}^\vee e_j$ в базисе $\boldsymbol{e}$, удовлетворяют соотношениям $B_{\boldsymbol{e}}C_{\boldsymbol{e},\boldsymbol{e}^\vee} = B_{\boldsymbol{e},\boldsymbol{e}^\vee} = E$ и $C^t_{\boldsymbol{e},{}^\vee\boldsymbol{e}}B_{\boldsymbol{e}} = B_{{}^\vee\boldsymbol{e},\boldsymbol{e}} = E$, откуда

$$C_{\boldsymbol{e},\boldsymbol{e}^\vee} = B_{\boldsymbol{e}}^{-1} \quad \text{и} \quad C_{\boldsymbol{e}\,{}^\vee\boldsymbol{e}} = \left(B_{\boldsymbol{e}}^t\right)^{-1}\,.$$

Знание двойственного к базису $\boldsymbol{e}$ относительно билинейной формы β базиса позволяет находить коэффициенты разложения любого вектора $v \in V$ по каждому из двойственных базисов как взятые с надлежащей стороны скалярные произведения вектора v с соответствующими элементами двойственного базиса:

$$v = \sum_i \beta({}^\vee e_i, v)\, e_i = \sum_i \beta(v, e_i^\vee)\, e_i = \sum_i \beta(v, e_i)\, {}^\vee e_i = \sum_i \beta(e_i, v)\, e_i^\vee\,. \tag{13-12}$$

Упражнение 13.6. Убедитесь в этом.

13.2.2. Изотропные подпространства. Подпространство $U \subset V$ называется *изотропным* для билинейной формы β, если эта форма ограничивается на него в тождественно нулевую форму, т. е. когда $\beta(u, w) = 0$ для всех $u, w \in U$. Например, каждое одномерное подпространство является изотропным для любой кососимметричной формы, а линейные оболочки первых n и последних n базисных векторов пространства $\Bbbk^{2n}$ изотропны для гиперболической формы из прим. 13.2 и симплектической формы из прим. 13.3.

Предложение 13.2
Размерность изотропного подпространства невырожденной билинейной формы на пространстве V не превосходит $\dim V/2$.

Доказательство. Изотропность подпространства $U \subset V$ означает, что корреляция $\beta^\wedge : V \xrightarrow{\sim} V^*$ отображает U внутрь $\operatorname{Ann} U \subset V^*$. Так как корреляция невырожденной формы инъективна, $\dim U \leqslant \dim \operatorname{Ann} U = \dim V - \dim U$, откуда $2\dim U \leqslant \dim V$. □

Замечание 13.1. Примеры гиперболической и симплектической форм показывают, что оценка из предл. 13.2 в общем случае неулучшаема.

13.2.3. Группа изометрий. Как мы видели в n° 13.1.4 на стр. 189, линейный эндоморфизм $f : V \to V$ является изометрическим для билинейной формы β на пространстве V если и только если его матрица $F_{\boldsymbol{e}}$ в произвольном базисе $\boldsymbol{e}$ пространства V связана с матрицей Грама $B_{\boldsymbol{e}}$ этого базиса соотношением[1] $F_{\boldsymbol{e}}^t B_{\boldsymbol{e}} F_{\boldsymbol{e}} = B_{\boldsymbol{e}}$. Если форма β невырождена, то беря определители обеих частей, заключаем, что $\det^2 F_{\boldsymbol{e}} = 1$, откуда $\det F_{\boldsymbol{e}} = \pm 1$. Поэтому любая изометрия конечномерного пространства с невырожденной билинейной формой обратима и с точностью до знака сохраняет объём. Так как композиция изометрий и обратное к изометрии отображение тоже являются изометриями, изометрические преобразования пространства V образуют группу. Она обозначается $\mathrm{O}_\beta(V)$ и называется *группой изометрий*[2] невырожденной билинейной формы β. Изометрии определителя 1 называются *специальными* или *собственными* и образуют в группе всех изометрий подгруппу, обозначаемую $\mathrm{SO}_\beta(V)$.

Из форм. (13-8) на стр. 189 вытекает, что обратная к изометрии f изометрия имеет матрицу

$$F_{\boldsymbol{e}}^{-1} = B_{\boldsymbol{e}}^{-1} F_{\boldsymbol{e}}^t B_{\boldsymbol{e}} \,. \tag{13-13}$$

Пример 13.4 (изометрии вещественной гиперболической плоскости)
Оператор $f : H_2 \to H_2$, имеющий в стандартном гиперболическом базисе e_1, $e_2 \in H_2$ матрицу

$$F = \begin{pmatrix} a & b \\ c & d \end{pmatrix} ,$$

является изометрическим тогда и только тогда, когда

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ,$$

что равносильно уравнениям $ac = bd = 0$ и $ad + bc = 1$, имеющим два семейства решений:

$$F_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \quad \text{и} \quad \tilde{F}_\lambda = \begin{pmatrix} 0 & \lambda \\ \lambda^{-1} & 0 \end{pmatrix} , \quad \text{где } \lambda \in \Bbbk^* = \Bbbk \smallsetminus \{0\}. \tag{13-14}$$

Над полем $\mathbb{R}$ оператор F_λ является собственным, и при $\lambda > 0$ называется *гиперболическим поворотом*, т. к. каждый вектор $v = (x, y)$, обе координаты которого ненулевые, движется при действии на него операторов F_λ с $\lambda \in (0, \infty)$ по гиперболе $xy = \mathrm{const}$. Если положить $\lambda = e^t$ и перейти к ортогональному базису из векторов $p = (e_1 + e_2)/\sqrt{2}$, $q = (e_1 - e_2)/\sqrt{2}$, то оператор F_λ запишется в нём матрицей, похожей на матрицу евклидова поворота

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \operatorname{ch} t & \operatorname{sh} t \\ \operatorname{sh} t & \operatorname{ch} t \end{pmatrix} ,$$

где $\operatorname{ch} t \stackrel{\text{def}}{=} (e^t + e^{-t})/2$ и $\operatorname{sh} t \stackrel{\text{def}}{=} (e^t - e^{-t})/2$ называются *гиперболическими косинусом* и *синусом* вещественного числа t. Оператор F_λ с $\lambda < 0$ является композицией гиперболического поворота и центральной симметрии. Несобственный оператор $\widetilde{F}_\lambda$ является композицией гиперболического поворота с отражением относительно той оси гиперболы, которая пересекается с её ветвями.

[1] См. формулу (13-8) на стр. 189.
[2] А также *ортогональной группой* или *группой автоморфизмов*.

13.2.4. Биекция между формами и операторами. На пространстве V с билинейной формой $\beta : V \times V \to \Bbbk$ каждому линейному оператору $f : V \to V$ можно сопоставить билинейную форму $\beta_f(u,w) \stackrel{\text{def}}{=} \beta(u, fw)$ с матрицей Грама $\boldsymbol{e}^t \cdot f(\boldsymbol{e}) = \boldsymbol{e}^t \cdot \boldsymbol{e}\, F_{\boldsymbol{e}} = B_{\boldsymbol{e}} F_{\boldsymbol{e}}$ в произвольно выбранном базисе $\boldsymbol{e}$ пространства V. Поскольку на языке матриц отображение $f \mapsto \beta_f$ заключается в левом умножении матрицы оператора на матрицу Грама: $F_{\boldsymbol{e}} \mapsto B_{\boldsymbol{e}} F_{\boldsymbol{e}}$, оно линейно и обратимо, если форма β невырождена. Обратное отображение задаётся умножением матрицы оператора слева на обратную к матрице Грама матрицу. Поэтому каждая билинейная форма

$$\alpha : V \times V \to \Bbbk$$

на конечномерном векторном пространстве V с фиксированной невырожденной билинейной формой β имеет вид $\alpha(u,w) = \beta(u, f_\alpha w)$ для некоторого линейного оператора $f_\alpha : V \to V$, однозначно определяемого формой α. Матрица $F_{\boldsymbol{e}}$ оператора f_α выражается через матрицы Грама $B_{\boldsymbol{e}}$ и $A_{\boldsymbol{e}}$ форм β и α по формуле $F_{\boldsymbol{e}} = B_{\boldsymbol{e}}^{-1} A_{\boldsymbol{e}}$.

Пример 13.5 (канонический оператор)
Задаваемая невырожденной билинейной формой β биекция между формами и операторами сопоставляет транспонированной к β форме $\beta^t(u,w) \stackrel{\text{def}}{=} \beta(w,u)$ оператор $\varkappa : V \to V$, который называется *каноническим оператором* невырожденной билинейной формы β и однозначно характеризуется свойством

$$\forall\, u, w \in V \quad \beta(w,u) = \beta(u, \varkappa w)\,. \tag{13-15}$$

Матрица $K_{\boldsymbol{e}}$ канонического оператора в произвольном базисе $\boldsymbol{e}$ пространства V выражается через матрицу Грама $B_{\boldsymbol{e}}$ формы β по формуле $K_{\boldsymbol{e}} = B_{\boldsymbol{e}}^{-1} B_{\boldsymbol{e}}^t$.

Упражнение 13.7. Убедитесь, что при замене матрицы Грама по правилу $B \mapsto C^t B C$, где $C \in \mathrm{GL}_n(\Bbbk)$, матрица $K = B^{-1}B^t$ меняется по правилу $K \mapsto C^{-1} K C$, т. е. канонические операторы изоморфных билинейных форм подобны.

Так как $\beta(u,w) = \beta(w, \varkappa u) = \beta(\varkappa u, \varkappa w)$ для всех $u, w \in V$, канонический оператор является изометрическим.

Теорема 13.1
Над алгебраически замкнутым полем характеристики, отличной от двух, две невырожденные билинейные формы изометрически изоморфны если и только если их канонические операторы подобны.

Доказательство. Импликация «только если» вытекает из упр. 13.7 и имеет место над любым полем. Докажем обратную импликацию. Пусть невырожденные билинейные формы α и β имеют подобные канонические операторы $\varkappa_\alpha$ и $\varkappa_\beta = g^{-1}\varkappa_\alpha g$. Тогда форма $\alpha'(u,w) = \alpha(gu, gw)$ изометрически изоморфна форме α и имеет канонический оператор $g^{-1}\varkappa_\alpha g = \varkappa_\beta$, поскольку $\alpha'(u,w) = \alpha(gu, gw) = \alpha(gw, \varkappa_\alpha gu) = \alpha'(w, g^{-1}\varkappa_\alpha g u)$ для всех u, w. Таким образом, заменяя форму α на форму α', мы без ограничения общности можем считать, что формы α и β имеют один и тот же канонический оператор $\varkappa$. Линейный оператор f, однозначно определяемый равенством $\beta(u,w) = \alpha(u, fw)$ для всех u, w, обратим в силу невырожденности форм α, β и *самосопряжён* относительно α в том смысле, что для всех u, w выполняется равенство

$$\alpha(fu, w) = \alpha(\varkappa^{-1} w, fu) = \beta(\varkappa^{-1} w, u) = \beta(u,w) = \alpha(u, fw)\,.$$

Любой многочлен от оператора f тоже самосопряжён относительно формы α. В силу идущей ниже лем. 13.1, над алгебраически замкнутым полем $\Bbbk$ с $\operatorname{char}\Bbbk \neq 2$ существует такой многочлен $P(t) \in \Bbbk[t]$, что оператор $h = P(f)$ удовлетворяет равенству $h^2 = f$. Такой оператор h биективен и самосопряжён относительно α. Поэтому форма

$$\beta(u,w) = \alpha(u, fw) = \alpha(u, h^2 w) = \alpha(hu, hw)$$

изометрически изоморфна форме α. □

Лемма 13.1
Над алгебраически замкнутым полем $\Bbbk$ с $\operatorname{char}\Bbbk \neq 2$ из любого биективного линейного оператора f на конечномерном векторном пространстве V можно извлечь квадратный корень, являющийся многочленом от оператора f.

Доказательство. Поскольку при всех целых $k \geqslant 0$ биномиальный коэффициент $\binom{2k}{k}$ нацело делится на[1] $(k+1)$, над любым полем $\Bbbk$ с $\operatorname{char}\Bbbk \neq 2$ корректно определён биномиальный степенной ряд

$$\sqrt{1+x} = \sum_{k\geqslant 0} \binom{1/2}{k} x^k = \sum_{k\geqslant 0} \frac{(-1)^{k-1}}{2^k k!} \cdot 3 \cdot 5 \cdot \cdots \cdot (2k-3) \cdot x^k = \\ = 1 + \frac{1}{2} \sum_{k\geqslant 1} \frac{(-1)^{k-1}}{4^{k-1}} \binom{2k-2}{k-1} \frac{x^k}{k}. \quad (13\text{-}16)$$

Упражнение 13.8. Убедитесь в том, что квадрат многочлена, равного сумме первых $n+1$ членов этого ряда, сравним в $\Bbbk[x]$ с $1+x$ по модулю x^{n+1}.

Если поле $\Bbbk$ алгебраически замкнуто, характеристический многочлен $\chi_f(t)$ оператора f разлагается на взаимно простые множители $(t-\lambda)^{m_\lambda}$, где $\lambda \in \operatorname{Spec}(f)$, и пространство V распадается в прямую сумму f-инвариантных корневых подпространств[2] $K_\lambda = \ker(f - \lambda\mathrm{Id})^{m_\lambda}$. Так как f биективен, в этом разложении все λ отличны от нуля, и для каждого λ корректно определён многочлен $p_\lambda(t) \in \Bbbk[t]$, равный сумме первых m_λ членов формального разложения Тэйлора функции $\sqrt{t}$ в точке λ, которое получается из биномиальной формулы (13-16) заменой переменных:

$$\sqrt{t} = \sqrt{\lambda + (t-\lambda)} = \sqrt{\lambda} \cdot \bigl(1 + \lambda^{-1/2}(t-\lambda)\bigr)^{1/2} = \\ = \lambda^{1/2} + \frac{1}{2}(t-\lambda) - \frac{\lambda^{-1/2}}{8}(t-\lambda)^2 + \frac{\lambda^{-1}}{16}(t-\lambda)^3 - \cdots.$$

Согласно упр. 13.8, $p_\lambda^2(t) \equiv t \bmod (t-\lambda)^{m_\lambda}$. По китайской теореме об остатках существует многочлен $p(t)$, сравнимый с $p_\lambda(t)$ по модулю $(t-\lambda)^{m_\lambda}$ сразу для всех $\lambda \in \operatorname{Spec}(f)$. Его квадрат

$$p^2(t) \equiv t \bmod (t-\lambda)^{m_\lambda} \quad \forall\, \lambda \in \operatorname{Spec}(f).$$

Поэтому квадрат оператора $p(f)$ действует на каждом корневом подпространстве K_λ точно также, как f. Тем самым, $p^2(f) = f$. □

[1] См. прим. 4.7 на стр. 62.
[2] См. n° 10.3 на стр. 143.

13.3. Ортогоналы и ортогональные проекции. С каждым подпространством U векторного пространства V с билинейной формой $\beta : V \times V \to \Bbbk$ связаны *левый* и *правый ортогоналы*

$$\begin{aligned} {}^{\perp}U &= \{v \in V \mid \forall\, u \in U\ \beta(v,u) = 0\}\,, \\ U^{\perp} &= \{v \in V \mid \forall\, u \in U\ \beta(u,v) = 0\}\,. \end{aligned} \tag{13-17}$$

Вообще говоря, это два разных подпространства в V.

Предложение 13.3
Если билинейная форма β на конечномерном пространстве V невырождена, то для всех подпространств $U \subset V$ выполняются равенства

$$\dim {}^{\perp}U = \dim V - \dim U = \dim U^{\perp} \quad \text{и} \quad ({}^{\perp}U)^{\perp} = U = {}^{\perp}(U^{\perp})\,.$$

Доказательство. Первые два равенства верны, так как ортогоналы (13-17) суть прообразы подпространства $\operatorname{Ann} U \subset V^*$ при изоморфизмах ${}^{\wedge}\beta, \beta^{\wedge} : V \xrightarrow{\sim} V^*$, и $\dim \operatorname{Ann} U = \dim V - \dim U$. Вторые два равенства вытекают из первых, поскольку оба подпространства $({}^{\perp}U)^{\perp}$ и ${}^{\perp}(U^{\perp})$ содержат U и имеют размерность $\dim U$. □

Предложение 13.4
Пусть билинейная форма β на произвольном[1] векторном пространстве V ограничивается на конечномерное подпространство $U \subset V$ в невырожденную на этом подпространстве билинейную форму $\beta|_U : U \times U \to \Bbbk$. Тогда $V = U \oplus U^{\perp}$, и проекция $v_U \in U$ каждого вектора $v \in V$ на подпространство U вдоль $U^{\perp}$ однозначно определяется тем, что $\beta(u,v) = \beta(u,v_U)$ для всех $u \in U$. Вектор v_U выражается через произвольный базис $u_1, \dots, u_n$ пространства U по формуле

$$v_U = \sum_{i=1}^{n} \beta\left({}^{\vee}u_i, v\right)\, u_i = \sum_{i=1}^{n} \beta\left(u_i, v\right)\, u_i^{\vee}\,, \tag{13-18}$$

где ${}^{\vee}u_1, \dots, {}^{\vee}u_n$ и $u_1^{\vee}, \dots, u_n^{\vee}$ суть левый и правый двойственные к $u_1, \dots, u_n$ относительно формы β базисы[2] в U.

Доказательство. Так как ограничение формы β на U невырождено, для любого вектора $v \in V$ существует единственный такой вектор $v_U \in U$, что линейная функция $u \mapsto \beta(u,v)$ на пространстве U задаётся правым скалярным умножением векторов из U на этот вектор v_U, т. е. для всех $u \in U$ выполняется равенство $\beta(u,v) = \beta(u,v_U)$. Поэтому разность $v - v_U \in U^{\perp}$. Таким образом, каждый вектор $v \in V$ представляется в виде суммы $v = v_U + (v - v_U)$ с $v \in U$ и $v - v_U \in U^{\perp}$. Поскольку в любом разложении $v = v'_U + w$ с $v'_U \in U$ и $w \in U^{\perp}$ для всех $u \in U$ выполняется равенство $\beta(u,v) = \beta(u,v'_U)$, имеем равенство $v'_U = v_U$, а значит и равенство $w = v - v'_U = v - v_U$, что доказывает первые два утверждения предложения. Последнее утверждение вытекает из форм. (13-12) на стр. 191: $v_U = \sum_i \beta({}^{\vee}u_i, v_U)\, u_i = \sum_i \beta({}^{\vee}u_i, v)\, u_i$. □

Упражнение 13.9. Докажите симметричное утверждение: $V = {}^{\perp}U \oplus U$ если и только если билинейная форма β ограничивается на конечномерное подпространство $U \subset V$ в невырожденную на этом подпространстве билинейную форму $\beta|_U : U \times U \to \Bbbk$; при этом проекция ${}_U v$ каждого вектора $v \in V$ на U вдоль ${}^{\perp}U$ однозначно определяется тем, что $\beta(v,u) = \beta({}_U v, u)$ для всех $u \in U$ и находится по формуле ${}_U v = \sum \beta\left(v, u_i^{\vee}\right)\, u_i = \sum \beta\left(v, u_i\right)\, {}^{\vee}u_i$.

[1] Возможно даже бесконечномерном.
[2] См. n° 13.2.1 на стр. 191.

13.4. Симметричные и кососимметричные формы. Билинейная форма β называется *симметричной*, если $\beta(u,w) = \beta(w,u)$ для всех $u, w \in V$, и *кососимметричной* — если $\beta(v,v) = 0$ для всех $v \in V$. В последнем случае для любых $u, w \in V$ выполняется равенство

$$0 = \beta(u+w, u+w) = \beta(u,w) + \beta(w,u)\,,$$

откуда $\beta(u,w) = -\beta(w,u)$.

Упражнение 13.10. Убедитесь, что при $\operatorname{char}\Bbbk \neq 2$ равенство $\beta(u,w) = -\beta(w,u)$ всех $u, w \in V$ равносильно равенству $\beta(v,v) = 0$ для всех $v \in V$ и что формы $\beta(u,w)$ и $\beta^t(u,w) = \beta(w,u)$ пропорциональны ровно в двух случаях: когда $\beta^t = \pm\beta$.

Если $\operatorname{char}\Bbbk = 2$, каждая кососимметричная форма автоматически симметрична, но не наоборот. Если $\operatorname{char}\Bbbk \neq 2$, пространства симметричных и кососимметричных билинейных форм имеют нулевое пересечение, и каждая билинейная форма β однозначно раскладывается в сумму $\beta = \beta_+ + \beta_-$ симметричной и кососимметричной форм

$$\beta_+(v,w) = \frac{\beta(v,w) + \beta(w,v)}{2} \quad \text{и} \quad \beta_-(v,w) = \frac{\beta(v,w) - \beta(w,v)}{2}\,.$$

13.4.1. Левая и правая корреляции симметричной билинейной формы совпадают друг с другом, и мы будем в этом случае обозначать оператор $\beta^\wedge = {}^\wedge\beta$ через $\widehat{\beta}: V \to V^*$ и называть просто *корреляцией* симметричной формы β. Напомню, корреляция переводит вектор $v \in V$, в линейную функцию

$$\widehat{\beta}v: V \to \Bbbk\,, \quad u \mapsto \beta(u,v) = \beta(v,u)\,.$$

Для кососимметричной формы левая и правая корреляции различаются знаком: $\beta^\wedge = -{}^\wedge\beta$.

13.4.2. Ядро. Левое и правое ядро (косо)симметричной формы β совпадают друг с другом и называются просто *ядром* этой формы. Поэтом для (косо)симметричной формы β пространство $\ker{}^\wedge\beta = \ker\beta^\wedge$ обозначается просто $\ker\beta \stackrel{\text{def}}{=} \{w \in V \mid \forall\, v \in V\ \beta(v,w) = \pm\beta(w,v) = 0\}$.

Предложение 13.5
Ограничение (косо) симметричной билинейной формы β на любое дополнительное к ядру $\ker\beta$ подпространство $U \subset V$ невырождено.

Доказательство. Пусть подпространство $U \subset V$ таково, что $V = \ker\beta \oplus U$, а вектор $w \in U$ удовлетворяет для всех $u \in U$ соотношению $\beta(u,w) = 0$. Записывая произвольный вектор $v \in V$ в виде $v = e + u$, где $e \in \ker\beta$ и $u \in U$, получаем $\beta(v,w) = \beta(e,w) + \beta(u,w) = 0$, откуда $w \in U \cap \ker\beta = 0$. □

Предложение 13.6
Любая (косо) симметричная билинейная форма β на пространстве V корректно определяет на фактор пространстве $V/\ker\beta$ невырожденную билинейную форму $\overline{\beta}$ по формуле

$$\overline{\beta}([u],[w]) \stackrel{\text{def}}{=} \beta(u,w)\,. \tag{13-19}$$

Доказательство. Если $[u] = [u']$, а $[w] = [w']$, то векторы $u - u'$ и $w - w'$ лежат в $\ker\beta$ и имеют нулевые левые и правые скалярные произведения с любым вектором. Поэтому

$$\overline{\beta}([u'],[w']) = \beta(u',w') = \beta\big(u + (u'-u), w + (w'-w)\big) = \beta(u,w) = \overline{\beta}([u],[w])\,,$$

что доказывает корректность формулы (13-19). Пусть класс $[u] \in V/\ker\beta$ имеет нулевое скалярное произведение $\overline{\beta}([u],[w]) = 0$ со всеми классами $[w] \in V/\ker\beta$. По определению формы $\overline{\beta}$ это означает, что $\beta(u,w) = 0$ для всех $w \in U$, откуда $u \in \ker\beta$ и $[u] = 0$. □

Предостережение 13.1. Для произвольной билинейной формы, которая не является симметричной или кососимметричной, левое и правое ядра $\ker({}^\vee\beta)$ и $\ker(\beta^\vee)$ могут быть различны, и в этом случае предл. 13.5 и предл. 13.6, вообще говоря, неверны.

13.4.3. Ортогоналы и проекции. Если форма β на пространстве V (косо) симметрична, то левый и правый ортогоналы к любому подпространству $U \subset V$ совпадают друг с другом и обозначаются через $U^\perp$. Если (косо) симметричная форма β ограничивается на подпространство $U \subset V$ в невырожденную на этом подпространстве форму, то $V = U \oplus U^\perp$ по предл. 13.4. В этом случае подпространство $U^\perp$ называется *ортогональным дополнением* к подпространству U. Проекция v_U вектора $v \in V$ на U вдоль $U^\perp$ называется *ортогональной проекцией* на U относительно формы β. Вектор v_U однозначно характеризуется тем, что его левое и правое скалярное произведение со всеми векторами из U такие же, как и у вектора v.

Если форма β невырождена на всём пространстве V, то по предл. 13.4

$$\dim U^\perp = \dim V - \dim U \quad \text{и} \quad U^{\perp\perp} = U$$

для всех подпространств $U \subset V$. В этом случае ограничение формы β на подпространство $U \subset V$ невырождено если и только если невырождено её ограничение на $U^\perp$.

Теорема 13.2 (теорема Лагранжа)
Каждое конечномерное векторное пространство с симметричной билинейной формой β над любым полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ обладает базисом с диагональной матрицей Грама[1].

Доказательство. Если $\dim V = 1$ или форма β нулевая, то матрица Грама любого базиса диагональна. Если форма β ненулевая, то найдётся вектор $e \in V$ с $\beta(e,e) \neq 0$, ибо в противном случае $2\beta(u,w) = \beta(u+w,u+w) - \beta(u,u) - \beta(w,w)) = 0$ для всех $u,w \in V$. Возьмём такой вектор e в качестве первого вектора искомого базиса. Поскольку ограничение формы β на одномерное подпространство $U = \Bbbk\cdot e$ невырождено, пространство V распадается в прямую ортогональную сумму $U \oplus U^\perp$. По индукции, в $U^\perp$ есть базис с диагональной матрицей Грама. Добавляя к нему e, получаем искомый базис в V. □

Следствие 13.1
Над алгебраически замкнутым полем $\Bbbk$ характеристики $\operatorname{char}(\Bbbk) \neq 2$ две симметричных билинейных формы изометрически изоморфны если и только если их матрицы Грама имеют одинаковый ранг.

Доказательство. Над алгебраически замкнутым полем каждый ненулевой диагональный элемент матрицы Грама ортогонального базиса можно сделать единичным, заменив соответствующий ему базисный вектор e_i на $e_i/\sqrt{\beta(e_i,e_i)}$. □

[1] Такие базисы называются *ортогональными*.

Пример 13.6 (ортогональный базис гиперболического пространства)
В гиперболическом пространстве[1] H_{2n} с гиперболическим базисом $e_1, \dots, e_n, e_{n+1}, \dots, e_{2n}$ над произвольным полем $\Bbbk$ характеристики $\operatorname{char}(\Bbbk) \neq 2$ в качестве ортогонального базиса можно взять, например, векторы $p_i = e_i + e_{n+i}$ и $q_i = e_i - e_{n+i}$ со скалярными квадратами $h(p_i, p_i) = 2$ и $h(q_i, q_i) = -2$.

Теорема 13.3
Каждое изотропное подпространство U в пространстве V с невырожденной симметричной билинейной формой β содержится в некотором гиперболическом подпространстве $W \subset V$ размерности $\dim W = 2 \dim U$. При этом любой базис подпространства U дополняется до гиперболического базиса пространства W.

Доказательство. Рассмотрим произвольный базис $u_1, \dots, u_m$ в U, дополним его до базиса в V и обозначим через $u_1^\vee, \dots, u_m^\vee$ первые m векторов двойственного относительно формы β базиса в V. Тогда

$$\beta(u_i, u_j^\vee) = \begin{cases} 1 & \text{при } i = j \\ 0 & \text{при } i \neq j\,, \end{cases} \tag{13-20}$$

и эти соотношения ортогональности не нарушаются при добавлении к любому из векторов $u_j^\vee$ произвольной линейной комбинации векторов u_i. Заменим каждый из векторов $u_j^\vee$ на вектор

$$w_j = u_j^\vee - \frac{1}{2} \sum_{\nu=1}^{m} \beta(u_j^\vee, u_\nu^\vee) \cdot u_\nu\,.$$

Векторы $w_1, \dots, w_m$ по-прежнему удовлетворяют соотношениям (14-1) и вдобавок

$$\beta(w_i, w_j) = \beta(u_i^\vee, u_j^\vee) - \frac{1}{2}\beta(u_i^\vee, u_j^\vee) - \frac{1}{2}\beta(u_j^\vee, u_i^\vee) = 0\,,$$

т. е. $2m$ векторов u_i, w_j, $1 \leqslant i, j \leqslant m$, образуют гиперболический базис в своей линейной оболочке, которую мы и возьмём в качестве W. □

Следствие 13.2
Следующие свойства пространства V с невырожденной симметричной билинейной формой эквивалентны:

1) V изометрически изоморфно гиперболическому пространству
2) V является прямой суммой двух изотропных подпространств
3) $\dim V$ чётна, и в V имеется изотропное подпространство половинной размерности.

Доказательство. Импликация (1)⇒(2) очевидна. Пусть выполнено (2). По предл. 13.2 размерность каждого из из двух изотропных прямых слагаемых не превышает половины размерности V, что возможно только если обе эти размерности равны $\frac{1}{2}\dim V$. Тем самым, (2)⇒(3). По предл. 14.1 на стр. 201 каждое изотропное подпространство размерности $\frac{1}{2}\dim V$ содержится в гиперболическом подпространстве размерности $\dim V$, которое таким образом совпадает со всем пространством V, что даёт импликацию (3)⇒(1). □

[1] См. прим. 13.2 на стр. 190.

Теорема 13.4 (кососимметричная версия предл. 14.1)
Каждое изотропное подпространство U невырожденной кососимметричной формы ω на пространстве V содержится в некотором симплектическом подпространстве $W \subset V$ размерности $\dim W = 2\dim U$. При этом любой базис в U дополняется до симплектического базиса в W.

Доказательство. Действуя как в предл. 14.1, для произвольного базиса $u_1, \dots, u_m$ в U построим векторы $u_1^\vee, \dots, u_m^\vee$ с

$$\omega\left(u_i, u_j^\vee\right) = \begin{cases} 1 & \text{при } i = j \\ 0 & \text{при } i \neq j\,, \end{cases} \tag{13-21}$$

и положим

$$w_j = u_j^\vee - \sum_{\nu<j} \omega\left(u_j^\vee, u_\nu^\vee\right) \cdot u_\nu\,. \tag{13-22}$$

Векторы $w_1, \dots, w_m$ также удовлетворят равенствам (13-21) и для всех $i < j$

$$\omega(w_i, w_j) = \omega(u_i^\vee, u_j^\vee) - \omega(u_j^\vee, u_i^\vee) \cdot \omega(u_i^\vee, u_i) = 0\,.$$

Тем самым, векторы u_i и w_j с $1 \leqslant i, j \leqslant m$ составляют симплектический базис в своей линейной оболочке, которую мы и возьмём в качестве W. □

Теорема 13.5 (теорема Дарбу)
Над произвольным полем $\Bbbk$ любой характеристики для каждой кососимметричной билинейной формы ω на конечномерном векторном пространстве V имеется базис с матрицей Грама, ненулевые элементы которой сосредоточены в расположенных на главной диагонали 2×2 блоках вида

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{13-23}$$

В частности, $\operatorname{rk}\omega$ всегда чётен.

Доказательство. Если форма тождественно нулевая, то доказывать нечего. Если $\omega(u, w) \neq 0$ для каких-то u, w, положим $e_1 = u, e_2 = w/\omega(u, w)$. Так как матрица Грама векторов e_1, e_2 имеет вид (13-23), эти векторы не пропорциональны и порождают двумерное подпространство $U \subset V$, на которое форма ω ограничивается невырождено. Поэтому $V = U \oplus U^\perp$. Применяя индукцию по $\dim V$, можно считать, что в подпространстве $U^\perp$ требуемый базис есть. Добавляя к нему e_1, e_2, получаем искомый базис в V. □

Следствие 13.3
Над произвольным полем $\Bbbk$ любой характеристики каждое пространство с невырожденной кососимметричной формой изометрически изоморфно симплектическому пространству Ω_{2n} из прим. 13.3 на стр. 190.

Доказательство. Согласно теор. 13.5 все ненулевые элементы матрицы Грама формы в подходящем базисе сосредоточатся в расположенных на главной диагонали 2×2 блоках (13-23). Чтобы получить из такого базиса симплектический, надо лишь переставить базисные векторы: сначала написать подряд все векторы с нечётными номерами, а потом — с чётными. □

Следствие 13.4
Группа изометрий невырожденной кососимметричной формы транзитивно действует на изотропных и на симплектических подпространствах любой фиксированной размерности.

Доказательство. Если подпространства $W_1, W_2 \subset \Omega_{2n}$ оба изометрически изоморфны Ω_{2k}, то их ортогоналы $W_1^\perp, W_2^\perp \subset \Omega_{2n}$ оба изометрически изоморфны $\Omega_{2(n-k)}$. Прямая сумма любых двух изометрических изоморфизмов $W_1 \xrightarrow{\sim} W_2$ и $W_1^\perp \xrightarrow{\sim} W_2^\perp$ даёт изометрический изоморфизм $\Omega_{2n} = W_1 \oplus W_1^\perp \xrightarrow{\sim} W_2 \oplus W_2^\perp = \Omega_{2n}$, переводящий W_1 в W_2. Если k-мерные подпространства $U_1, U_2 \subset \Omega_{2n}$ изотропны, то любой базис $\boldsymbol{u_1}$ в U_1 и любой базис $\boldsymbol{u_2}$ в U_2 дополняются по теор. 13.4 до состоящих из $2k$ векторов наборов $\boldsymbol{w_1}$ и $\boldsymbol{w_2}$, являющихся симплектическими базисами в своих линейных оболочках W_1 и W_2. Отображая первый набор во второй, мы получаем изометрический изоморфизм $W_1 \xrightarrow{\sim} W_2$, переводящий U_1 в U_2. Беря, как и выше, прямую сумму этого автоморфизма с любым изометрическим изоморфизмом $W_1^\perp \xrightarrow{\sim} W_2^\perp$, получаем изометрический автоморфизм пространства Ω_{2n}, переводящий U_1 в U_2. □

§14. Квадратичные формы

В этом параграфе мы по умолчанию считаем, что основное поле $\Bbbk$ имеет $\operatorname{char}(\Bbbk) \neq 2$.

14.1. Пространства со скалярным произведением. Будем называть *пространством со скалярным произведением* конечномерное векторное пространство V над произвольным полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ с зафиксированной на нём невырожденной[1] симметричной билинейной формой $\beta : V \times V \to \Bbbk$. В этом и следующем разделах буква V по умолчанию обозначает именно такое пространство.

14.1.1. Ортогональные прямые суммы. Из двух пространств V_1, V_2 со скалярными произведениями β_1, β_2 можно изготовить пространство $V_1 \oplus V_2$ со скалярным произведением $\beta_1 \dotplus \beta_2$, относительно которого слагаемые ортогональны друг другу и которое ограничивается на V_1 и V_2 в β_1 и β_2. Это скалярное произведение задаётся формулой

$$\big[\beta_1 \dotplus \beta_2\big]\big((u_1,u_2),\,(w_1,w_2)\big) \stackrel{\text{def}}{=} \beta_1(u_1,u_2) + \beta_2(w_1,w_2)\,.$$

Его матрица Грама в любом базисе, первые $\dim V_1$ векторов которого образуют базис в V_1 с матрицей Грама B_1, а последние $\dim V_2$ векторов — базис в V_2 с матрицей Грама B_2, имеет блочный вид

$$\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}.$$

Пространство $V_1 \oplus V_2$ со скалярным произведением $\beta_1 \dotplus \beta_2$ обозначается $V_1 \dotplus V_2$ и называется *ортогональной прямой суммой* пространств V_1 и V_2.

Упражнение 14.1. Обозначим через H_{2n} гиперболическое пространство[2] размерности $2n$. Постройте изометрический изоморфизм[3] $H_{2m} \dotplus H_{2k} \xrightarrow{\sim} H_{2(m+k)}$.

14.1.2. Изотропные и анизотропные подпространства. Вектор $v \in V$ называется *изотропным*, если $\beta(v,v) = 0$. Подпространство $U \subset V$, целиком состоящее из изотропных векторов, изотропно в смысле n° 13.2.2 на стр. 191, т. е. $\beta(u,w) = 0$ для всех $u, w \in U$, поскольку

$$2\beta(u,w) = \beta(u+w,u+w) - \beta(u,u) - \beta(w,w) = 0\,.$$

Подпространство $U \subset V$ называется *анизотропным*, если в нём нет ненулевых изотропных векторов. Если анизотропно всё пространство V, то говорят, что скалярное произведение на V *анизотропно*. Например, евклидово скалярное произведение на вещественном векторном пространстве анизотропно. Так как анизотропная форма обладает свойствами (5,6) из предл. 13.1 на стр. 189, каждая анизотропная форма невырождена. Поэтому для любого анизотропного подпространства $U \subset V$ имеет место ортогональное разложение $V = U \oplus U^{\perp}$ из предл. 13.4 на стр. 195.

Предложение 14.1
Каждое изотропное подпространство U в пространстве V со скалярным произведением β содержится в некотором гиперболическом подпространстве $W \subset V$ размерности $\dim W = 2\dim U$. При этом любой базис подпространства U дополняется до гиперболического базиса пространства W.

[1] См. предл. 13.1 на стр. 189.
[2] См. прим. 13.2 на стр. 190.
[3] См. n° 13.1.4 на стр. 189.

Доказательство. Рассмотрим произвольный базис $u_1, \dots, u_m$ в U, дополним его до базиса в V и обозначим через $u_1^\vee, \dots, u_m^\vee$ первые m векторов ортогонально двойственного базиса. Тогда

$$\beta(u_i, u_j^\vee) = \begin{cases} 1 & \text{при } i = j \\ 0 & \text{при } i \neq j\,, \end{cases} \tag{14-1}$$

и эти соотношения ортогональности не нарушаются при добавлении к любому из векторов $u_j^\vee$ произвольной линейной комбинации векторов u_i. Заменим каждый из векторов $u_j^\vee$ на вектор

$$w_j = u_j^\vee - \frac{1}{2}\sum_{\nu=1}^{m} \beta(u_j^\vee, u_\nu^\vee)\cdot u_\nu\,.$$

Векторы $w_1, \dots, w_m$ по-прежнему удовлетворяют соотношениям (14-1) и вдобавок

$$\beta(w_i, w_j) = \beta(u_i^\vee, u_j^\vee) - \frac{1}{2}\beta(u_i^\vee, u_j^\vee) - \frac{1}{2}\beta(u_j^\vee, u_i^\vee) = 0\,,$$

т. е. $2m$ векторов u_i, w_j, $1 \leqslant i, j \leqslant m$, образуют гиперболический базис в своей линейной оболочке, которую мы и возьмём в качестве W. □

Теорема 14.1
Каждое пространство V со скалярным произведением распадается в прямую ортогональную сумму $V \simeq H_{2k} \dotplus A$, первое слагаемое которой гиперболическое и может быть нулевым или совпадать со всем пространством V, а второе слагаемое $A = H_{2k}^\perp$ анизотропно.

Доказательство. Индукция по $\dim V$. Если V анизотропно (что так при $\dim V = 1$), доказывать нечего. Если существует ненулевой изотропный вектор $e \in V$, то по предл. 14.1 он лежит в некоторой гиперболической плоскости $H_2 \subset V$, и $V = H_2 \oplus H_2^\perp$ согласно предл. 13.4. По индукции, $H_2^\perp = H_{2m} \oplus A$, где $A = H_{2m}^\perp$ анизотропно. Поэтому $V = H_{2m+2} \oplus A$ и $A = H_{2m+2}^\perp$. □

Замечание 14.1. Ниже, в теор. 14.4 на стр. 205, мы увидим, что разложение из теор. 14.1 единственно в следующем смысле: если $V \simeq H_{2k} \dotplus U \simeq H_{2m} \dotplus W$, где U и W анизотропны, то $k = m$ и существует изометрический изоморфизм $U \xrightarrow{\sim} W$.

Следствие 14.1
Следующие свойства пространства V со скалярным произведением эквивалентны:

1) V изометрически изоморфно гиперболическому пространству
2) V является прямой суммой двух изотропных подпространств
3) $\dim V$ чётна, и в V имеется изотропное подпространство половинной размерности.

Доказательство. Импликация (1)⇒(2) очевидна. Пусть выполнено (2). По предл. 13.2 размерность каждого из из двух изотропных прямых слагаемых не превышает половины размерности V, что возможно только если обе эти размерности равны $\frac{1}{2}\dim V$. Тем самым, (2)⇒(3). По предл. 14.1 на стр. 201 каждое изотропное подпространство размерности $\frac{1}{2}\dim V$ содержится в гиперболическом подпространстве размерности $\dim V$, которое таким образом совпадает со всем пространством V, что даёт импликацию (3)⇒(1). □

14.2. Изометрии и отражения. Всякий анизотропный вектор $e \in V$ задаёт разложение пространства V в прямую ортогональную сумму $V = \Bbbk \cdot e \oplus e^\perp$. Линейный оператор $\sigma_e : V \to V$, тождественно действующий на гиперплоскости $e^\perp$ и переводящий вектор e в $-e$, называется *отражением* в гиперплоскости $e^\perp$, см. рис. 14⋄1. Произвольный вектор $v = v_e + v_{e^\perp} \in V$, где $v_e = e\,\beta(e,v)/\beta(e,e)$ — проекция вектора v на одномерное подпространство[1] $\Bbbk \cdot e$ вдоль гиперплоскости $e^\perp$, а $v_{e^\perp} = v - v_e \in e^\perp$, переходит при этом в вектор

$$\sigma_e(v) = -v_e + v_{e^\perp} = v - 2v_e = v - 2\,\frac{\beta(e,v)}{\beta(e,e)} \cdot e\,. \tag{14-2}$$

Упражнение 14.2. Убедитесь, что $\sigma_e \in \mathrm{O}_\beta(V)$ и $\sigma_e^2 = \mathrm{Id}_V$, и докажите для любых изометрии $f \in \mathrm{O}(V)$ и анизотропного вектора $e \in V$ равенство $f \circ \sigma_e \circ f^{-1} = \sigma_{f(e)}$.

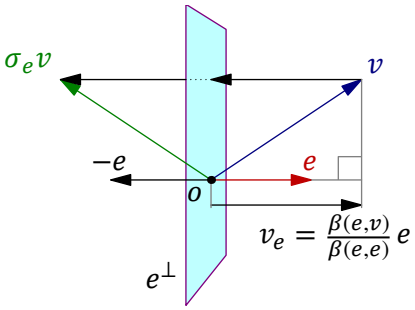


Рис. 14⋄1. Отражение σ_e.

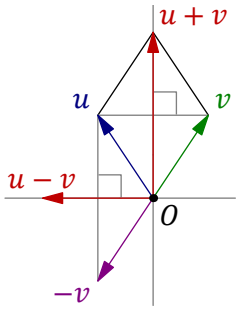


Рис. 14⋄2. Отражения в ромбе.

Лемма 14.1
В любом пространстве V со скалярным произведением β для каждой пары различных анизотропных векторов u, v с равными скалярными квадратами $\beta(u,u) = \beta(v,v) \neq 0$ существует отражение, переводящее u либо в v, либо в $-v$.

Доказательство. Если u и v коллинеарны, то искомым отражением является $\sigma_v = \sigma_u$. Если u и v не коллинеарны, то хотя бы одна из двух диагоналей $u+v$, $u-v$ натянутого на них ромба (см. рис. 14⋄2) анизотропна, поскольку эти диагонали ортогональны:

$$\beta(u+v, u-v) = \beta(u,u) - \beta(v,v) = 0\,,$$

и их линейная оболочка содержит анизотропные векторы u, v. Тем самым, хотя бы одно из отражений σ_{u-v}, σ_{u+v} определено. При этом $\sigma_{u-v}(u) = v$, а $\sigma_{u+v}(u) = -v$. □

Упражнение 14.3. Проверьте, последние два равенства.

Теорема 14.2
Всякая изометрия n-мерного пространства со скалярным произведением является композицией не более чем $2n$ отражений.

[1] Мы пользуемся тем, что $e^\vee = e\,/\,\beta(e,e)$ является двойственным к e относительно формы β базисным вектором одномерного пространства $\Bbbk\, e$ и по форм. (13-18) на стр. 195 ортогональная проекция произвольного вектора v на это подпространство равна $v_e = \beta(e,v)\,e^\vee$.

Доказательство. Индукция по n. Ортогональная группа одномерного пространства состоит из тождественного оператора E и отражения $-E$. Пусть $n > 1$ и $f : V \to V$ — изометрия. Выберем в V какой-нибудь анизотропный вектор v и обозначим через σ отражение, переводящее $f(v)$ в v или в $-v$. Композиция σf переводит v в $\pm v$, а значит, переводит в себя $(n-1)$-мерную гиперплоскость $v^\perp$. По индукции, действие σf на $v^\perp$ является композицией не более $2n-2$ отражений в гиперплоскостях внутри $v^\perp$. Продолжим их до отражений всего пространства V, добавив в зеркало каждого отражения вектор v. Композиция полученных отражений совпадает с σf на гиперплоскости $v^\perp$, а её действие на v либо такое же, как у σf (при $\sigma f(v) = v$), либо отличается от него знаком (при $\sigma f(v) = -v$). Поэтому σf, как оператор на всём пространстве V, есть композиция построенных $2n-2$ отражений и, возможно, ещё одного отражения в гиперплоскости $v^\perp$. Следовательно, $f = \sigma\sigma f$ это композиция не более $2n$ отражений. □

Упражнение 14.4. Покажите, что в анизотропном пространстве V в условиях лем. 14.1 всегда найдётся отражение, переводящее u в точности в v, и выведите отсюда, что любая изометрия n-мерного анизотропного пространства является композицией не более n отражений.

Теорема 14.3 (лемма Витта)
Пусть четыре пространства U_1, W_1, U_2, W_2 со скалярными произведениями таковы, что некоторые два из трёх пространств U_1, $U_1 \dotplus W_1$, W_1 изометрически изоморфны соответствующей паре пространств из тройки U_2, $U_2 \dotplus W_2$, W_2. Тогда оставшиеся третьи элементы троек тоже изометрически изоморфны.

Доказательство. Если есть изометрические изоморфизмы $f : U_1 \xrightarrow{\sim} U_2$ и $g : W_1 \xrightarrow{\sim} W_2$, то их прямая сумма $f \oplus g : U_1 \dotplus W_1 \to U_2 \dotplus W_2, (u, w) \mapsto (f(u), g(w))$, является требуемым изометрическим изоморфизмом. Оставшиеся два случая симметричны, и мы разберём один из них. Пусть имеются изометрические изоморфизмы

$$f : U_1 \xrightarrow{\sim} U_2 \quad \text{и} \quad h : U_1 \dotplus W_1 \xrightarrow{\sim} U_2 \dotplus W_2 .$$

Изометрический изоморфизм $g : W_1 \xrightarrow{\sim} W_2$ строится индукцией по $\dim U_1 = \dim U_2$. Если пространство U_1 одномерно с базисом u, то вектор u анизотропен. Поэтому векторы $f(u)$ и $h(u,0)$ тоже анизотропны и имеют одинаковые скалярные квадраты. Обозначим через σ отражение пространства $U_2 \dotplus W_2$, переводящее $h(u,0)$ в $\big(\pm f(u), 0\big)$. Композиция

$$\sigma h : U_1 \dotplus W_1 \xrightarrow{\sim} U_2 \dotplus W_2$$

изометрично отображает одномерное подпространство U_1 первой суммы на одномерное подпространство U_2 второй, а значит, изометрично отображает ортогональное дополнение к U_1 в первой сумме на ортогональное дополнение к U_2 во второй, что и даёт требуемый изоморфизм $\sigma h|_{W_1} : W_1 \xrightarrow{\sim} W_2$. Пусть теперь $\dim U_1 > 1$. Выберем в U_1 любой анизотропный вектор u и рассмотрим ортогональные разложения

$$U_1 \dotplus W_1 = \Bbbk \cdot u \dotplus u^\perp \dotplus W_1 \quad \text{и} \quad U_2 \dotplus W_2 = \Bbbk \cdot f(u) \dotplus f(u)^\perp \dotplus W_2 ,$$

в которых $u^\perp \subset U_1$ и $f(u)^\perp \subset U_2$ означают ортогональные дополнения к анизотропным векторам u и $f(u)$ внутри U_1 и U_2 соответственно. Так как пространства $\Bbbk \cdot u$ и $\Bbbk \cdot f(u)$ изометрически изоморфны, по уже доказанному существуют изометрии

$$f' : u^\perp \xrightarrow{\sim} f(u)^\perp \quad \text{и} \quad h' : u^\perp \dotplus W_1 \xrightarrow{\sim} f(u)^\perp \dotplus W_2 ,$$

к которым применимо индуктивное предположение. □

Теорема 14.4
Построенное в теор. 14.1 разложение пространства V со скалярным произведением в прямую ортогональную сумму гиперболического и анизотропного подпространств единственно в том смысле, что для любых двух таких разложений $V = H_{2k} \dotplus U = H_{2m} \dotplus W$ имеет место равенство $k = m$ и существует изометрический изоморфизм $U \simeq W$.

Доказательство. Пусть $m \geqslant k$, так что $H_{2m} = H_{2k} \dotplus H_{2(m-k)}$. Тождественное отображение $\mathrm{Id} : V \to V$ задаёт изометрический изоморфизм $H_{2k} \dotplus U \xrightarrow{\sim} H_{2k} \dotplus H_{2(m-k)} \dotplus W$. По лемме Витта существует изометрический изоморфизм $U \xrightarrow{\sim} H_{2(m-k)} \dotplus W$. Так как U анизотропно, $H_{2(m-k)} = 0$ (иначе в U будет ненулевой изотропный вектор), откуда $k = m$ и $U \simeq W$. □

Теорема 14.5
Если скалярное произведение на пространстве V невырожденно ограничивается на подпространства $U, W \subset V$ и существует изометрический изоморфизм $\varphi : U \xrightarrow{\sim} W$, то он продолжается (неоднозначно) до такого изометрического автоморфизма $f : V \xrightarrow{\sim} V$, что $f|_U = \varphi$.

Доказательство. Если есть хоть какой-нибудь изометрический изоморфизм $\psi : U^\perp \xrightarrow{\sim} W^\perp$, то изометрия $f = \varphi \oplus \psi : U \oplus U^\perp \xrightarrow{\sim} W \oplus W^\perp$, $(u, u') \mapsto \big(\varphi(h'), \psi(u')\big)$ является требуемым автоморфизмом пространства V. В силу сделанных предположений имеются изометрические изоморфизмы $\eta : U \dotplus U^\perp \xrightarrow{\sim} V$, $(u, u') \mapsto u + u'$, и $\zeta : U \dotplus W^\perp \xrightarrow{\sim} V$, $(u, w') \mapsto \varphi(u) + w'$. Композиция $\zeta^{-1}\eta : U \dotplus U^\perp \xrightarrow{\sim} U \dotplus W^\perp$ тоже изометрический изоморфизм. Так что по лемме Витта[1] ортогоналы $U^\perp$ и $W^\perp$ изометрически изоморфны. □

Следствие 14.2
Для каждого натурального числа k в диапазоне $1 \leqslant k \leqslant \dim V / 2$ группа изометрий $\mathrm{O}(V)$ транзитивно действует на k-мерных изотропных и $2k$-мерных гиперболических подпространствах в V.

Доказательство. Утверждение про гиперболические подпространства вытекает непосредственно из теор. 14.5, а про изотропные — получается из него применением предл. 14.1. □

14.3. Поляризация квадратичных форм. Функция $q : V \to \Bbbk$ на n-мерном векторном пространстве V над полем $\Bbbk$ называется *квадратичной формой*, если она является однородным многочленом степени 2 от координат, т. е. существуют такие базис $\boldsymbol{e} = (e_1, \dots, e_n)$ в V и однородный многочлен второй степени $q_{\boldsymbol{e}} \in \Bbbk[x_1, \dots, x_n]$, что $q(\lambda_1 e_1 + \dots + \lambda_n e_n) = q_{\boldsymbol{e}}(\lambda_1, \dots, \lambda_n)$ для всех $(\lambda_1, \dots, \lambda_n) \in \Bbbk^n$. Если $\mathrm{char}(\Bbbk) \neq 2$, то многочлен $q_{\boldsymbol{e}}$ можно записать в виде

$$q_{\boldsymbol{e}}(x_1, \dots, x_n) = \sum_{i,j=1}^{n} q_{ij} x_i x_j\,, \tag{14-3}$$

где суммирование происходит по всем парам индексов $1 \leqslant i, j \leqslant n$, а коэффициенты q_{ij} симметричны по i и j, т. е. при $i \neq j$ число $q_{ji} = q_{ij}$ равно *половине*[2] фактического коэффициента при $x_i x_j$ в многочлене $q_{\boldsymbol{e}}$, получающегося после приведения подобных слагаемых в (14-3). Если организовать числа q_{ij} в симметричную матрицу $Q_{\boldsymbol{e}} = \big(q_{ij}\big)$, которую мы будем называть

[1] См. теор. 14.3 на стр. 204.
[2] Обратите внимание, что над полем характеристики 2 многочлен $x_1 x_2$ не записывается в виде (14-3).

матрицей Грама многочлена $q_{\boldsymbol{e}}$, и обозначить через x и $x^t = (x_1, \dots, x_n)$ столбец и строку, составленные из переменных, то (14-3) можно переписать в виде

$$q(x_1, \dots, x_n) = \sum_{i,j=1}^{n} x_i q_{ij} x_j = x^t\, Q_{\boldsymbol{e}}\, x\,. \tag{14-4}$$

Сравнивая это с форм. (13-3) на стр. 187, мы заключаем, что $q(v) = \tilde{q}(v, v)$, где $\tilde{q} : V \times V \to \Bbbk$ — симметричная билинейная форма с матрицей Грама $Q_{\boldsymbol{e}}$ в базисе $\boldsymbol{e}$. Поскольку

$$q(u + w) - q(u) - q(w) = \tilde{q}(u + w, u + w) - \tilde{q}(u, u) - \tilde{q}(w, w) = 2\tilde{q}(u, w)\,,$$

симметричная билинейная форма $\tilde{q}$ со свойством $\tilde{q}(v, v) = q(v)$ однозначно определяется квадратичной формой q, если $\operatorname{char}\Bbbk \neq 2$. Симметричная билинейная форма $\tilde{q}$ называется *поляризацией* квадратичной формы q. Обратите внимание, что взаимно однозначное соответствие между квадратичными и симметричными билинейными формами

$$\begin{aligned} \tilde{q}(u, w) &\mapsto q(v) = \tilde{q}(v, v) \\ q(v) &\mapsto \tilde{q}(u, w) = \frac{1}{2}\big(q(u + w) - q(u) - q(w)\big) \end{aligned} \tag{14-5}$$

не зависят от базиса $\boldsymbol{e}$ в V. В частности, для любого базиса $\boldsymbol{f} = \boldsymbol{e}\, C_{\boldsymbol{ef}}$ в V значение $q(v)$ является однородным многочленом второй степени $q_{\boldsymbol{f}}$ от координат вектора v в базисе $\boldsymbol{f}$, причём матрица Грама этого многочлена, равная матрице Грама билинейной формы $\tilde{q}$ в базисе $\boldsymbol{f}$, будет равна[1] $Q_{\boldsymbol{f}} = C_{\boldsymbol{ef}}^t Q_{\boldsymbol{e}}\, C_{\boldsymbol{ef}}$.

Поскольку при переходе от базиса к базису определитель Грама умножается на квадрат определителя матрицы перехода, класс числа $\det Q_{\boldsymbol{e}} \in \Bbbk$ по модулю умножения на ненулевые квадраты из поля $\Bbbk$ не зависит от выбора базиса $\boldsymbol{e}$. Мы будем обозначать этот класс $\det q \in \Bbbk/\Bbbk^{*2}$ и называть его *определителем Грама* квадратичной формы q. Квадратичная форма q называется *вырожденной*, если $\det q = 0$. Формы с $\det q \neq 0$ называются *невырожденными*. Таким образом, невырожденность квадратичной формы q означает в точности то же, что невырожденность её поляризации[2] $\tilde{q}$. Под *рангом* квадратичной формы q мы понимаем ранг её поляризации $\tilde{q}$, равный рангу матрицы Грама $Q_{\boldsymbol{e}}$ в любом базисе $\boldsymbol{e}$. Также, как и для симметричных билинейных форм, мы будем называть ненулевой вектор $v \in V$ *изотропным* для квадратичной формы q, если $q(v) = 0$. Квадратичная форма называется *анизотропной*, если $q(v) \neq 0$ при $v \neq 0$.

Из доказанных выше результатов про симметричные билинейные формы немедленно получаются аналогичные результаты про квадратичные формы.

Следствие 14.3 (из теор. 14.1 на стр. 202)
Всякая квадратичная форма q над произвольным полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ в подходящих координатах записывается в виде $x_1x_{i+1} + x_2x_{i+2} + \cdots + x_ix_{2i} + \alpha(x_{2i+1}, x_{2i+2} \dots, x_r)$, где $r = \operatorname{rk}(q)$ и $\alpha(x) \neq 0$ при $x \neq 0$. □

Следствие 14.4 (из теор. 13.2 на стр. 197)
Всякая квадратичная форма над произвольным полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ линейной обратимой заменой переменных приводится к виду $\sum a_i x_i^2$. □

[1] См. формулу (13-2) на стр. 187.
[2] См. предл. 13.1 на стр. 189.

Следствие 14.5 (из сл. 13.1 на стр. 197)
Два однородных многочлена второй степени $f, g \in \Bbbk[x_1, \dots, x_n]$ над алгебраически замкнутым полем $\Bbbk$ характеристики $\operatorname{char}(\Bbbk) \neq 2$ тогда и только тогда переводятся друг в друга линейными обратимыми заменами переменных, когда задаваемые им квадратичные формы $f, g : \Bbbk^n \to \Bbbk$ имеют одинаковый ранг. □

Пример 14.1 (квадратичные формы от двух переменных)
Согласно сл. 14.4, ненулевая квадратичная форма от двух переменных

$$q(\boldsymbol{x}) = a\,x_1^2 + 2\,b\,x_1x_2 + c\,x_2^2 = (x_1, x_2)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{14-6}$$

подходящей линейной заменой координат приводятся либо к виду αt^2 с $\alpha \neq 0$, либо к виду

$$\alpha t_1^2 + \beta\, t_2^2\,, \quad \text{где} \quad \alpha\beta \neq 0\,.$$

Условимся писать $\xi \sim \eta$ для чисел $\xi, \eta \in \Bbbk$, если $\xi = \lambda^2\eta$ для какого-нибудь ненулевого $\lambda \in \Bbbk$. Тогда в первом случае $ac - b^2 \sim \det q \sim \alpha \cdot 0 = 0$, т. е. форма q вырождена, а во втором случае $ac - b^2 \sim \det q \sim \alpha\beta \neq 0$ и форма q невырождена. Тем самым, вырожденность ненулевой квадратичной формы (14-6) означает, что с точностью до постоянного множителя она является полным квадратом линейной формы $t \in V^*$. Такая форма q зануляется вдоль одномерного подпространства $\operatorname{Ann}(t) \subset V$ и отлична от нуля на всех остальных векторах.

Если форма (14-6) невырождена, и у неё есть ненулевой изотропный вектор $v = (\vartheta_1, \vartheta_2)$, то из равенства $\alpha\vartheta_1^2 + \beta\vartheta_2^2 = 0$ вытекает, что $\vartheta_2 \neq 0$ и $-\det q \sim -\alpha\beta \sim -\beta/\alpha = (\vartheta_1/\vartheta_2)^2$ является квадратом в поле $\Bbbk$. В этом случае многочлен

$$\alpha t_1^2 + \beta\, t_2^2 = \alpha\left(t_1 + \frac{\vartheta_1}{\vartheta_2}t_2\right)\left(t_1 - \frac{\vartheta_1}{\vartheta_2}t_2\right)$$

раскладывается над полем $\Bbbk$ в произведение двух непропорциональных линейных форм. Поэтому квадратичная форма q, у которой $-\det q$ является ненулевым квадратом, тождественно зануляется на двух одномерных подпространствах и отлична от нуля на всех прочих векторах. Мы будем называть такие формы *гиперболическими*[1]. Если же $-\det q$ не квадрат, то форма q анизотропна. Число $-\det(q) = b^2 - ac$ часто обозначают через $D/4$ и называют D *дискриминантом* квадратичной формы (14-6).

14.4. Квадратичные формы над конечными полями. Из курса алгебры известно[2], что для каждого простого $p \in \mathbb{N}$ любого $m \in \mathbb{N}$ существует единственное с точностью до изоморфизма поле $\mathbb{F}_q$ из $q = p^m$ элементов, и каждое конечное поле изоморфно одному и только одному из полей $\mathbb{F}_q$. Следуя принятому в начале этой лекции соглашению, всюду далее мы считаем, что $p = \operatorname{char}\mathbb{F}_q > 2$. Зафиксируем какой-нибудь элемент $\varepsilon \in \mathbb{F}_q$, не являющийся квадратом.

Упражнение 14.5. Убедитесь, что ненулевые квадраты образуют в мультипликативной группе $\mathbb{F}_q^*$ поля $\mathbb{F}_q$ подгруппу индекса 2. В частности, нужный нам элемент ε существует, и любой ненулевой элемент поля $\mathbb{F}_q$ умножением на подходящий ненулевой квадрат можно сделать равным либо 1, либо ε.

[1] Поскольку поляризация такой формы является гиперболическим скалярным произведением.
[2] См. раздел 3.5 на стр. 45 лекции http://gorod.bogomolov-lab.ru/ps/stud/algebra-1/1314/lec-03.pdf.

Лемма 14.2
При любых $a_1, a_2 \in \mathbb{F}_q^*$ квадратичная форма $a_1x_1^2+a_2x_2^2$ на двумерном координатном пространстве $\mathbb{F}_q^2$ принимает все значения из поля $\mathbb{F}_q$.

Доказательство. В силу упр. 14.5 при любых фиксированных $a_1, a_2 \in \mathbb{F}_q^*$ и $b \in \mathbb{F}_q$ чисел вида $a_1x_1^2$ и чисел вида $b - a_2x_2^2$, где x_1, x_2 независимо пробегают $\mathbb{F}_q$, имеется ровно по

$$1 + \frac{q-1}{2} = \frac{q+1}{2}$$

штук. Следовательно эти два множества чисел имеют общий элемент $a_1x_1^2 = b - a_2x_2^2$. Тем самым, $f(x_1, x_2) = b$. □

Предложение 14.2
Каждая квадратичная форма q ранга r над полем $\mathbb{F}_q$ в подходящих координатах записывается как $x_1^2+\cdots+x_{r-1}^2+x_r^2$ или как $x_1^2+\cdots+x_{r-1}^2+\varepsilon x_r^2$, и эти две формы изометрически не изоморфны.

Доказательство. По теор. 13.2 форма q в подходящих координатах записывается в виде

$$a_1x_1^2 + \cdots + a_rx_r^2\,, \quad \text{где все } a_i \neq 0.$$

Согласно упр. 14.5, умножая базисные векторы на подходящие ненулевые константы, мы можем считать, что каждое a_i равно либо 1, либо ε. Если $a_i = a_j = \varepsilon$ при каких-то $i \neq j$, то в линейной оболочке U базисных векторов e_i, e_j по лем. 14.2 найдётся вектор v_i с $q(v_i) = 1$. Ортогональное дополнение к v_i в плоскости U одномерно, и форма q ограничивается на него невырожденно. Поэтому там найдётся вектор v_j с $q(v_2)$, равным 1 или ε. Заменяя e_i, e_j на v_i, v_j, мы сохраняем вид формы, но получаем $a_i = 1$, строго уменьшая тем самым число коэффициентов, равных ε. Эту процедуру можно повторять, пока таких коэффициентов останется не более одного. Формы $q = x_1^2+\cdots+x_{r-1}^2+x_r^2$ и $q' = x_1^2+\cdots+x_{r-1}^2+\varepsilon x_r^2$ изометрически не изоморфны, поскольку индуцированные ими невырожденные квадратичные формы q_{red} и q'_{red} на факторах $V/\ker\tilde{q}$ и $V/\ker\tilde{q}'$ исходного пространства V, где были заданы формы, по ядрам этих форм[1], имеют разные определи Грама: $\det q_{\text{red}} = 1$ является квадратом, а $\det q'_{\text{red}} = \varepsilon$ — нет. □

Предложение 14.3
Всякая квадратичная форма на пространстве размерности $\geqslant 3$ над полем $\mathbb{F}_q$ имеет ненулевой изотропный вектор.

Доказательство. По теор. 13.2 форма записывается в подходящем базисе как

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + \cdots\,.$$

Если $a_1 = 0$ или $a_2 = 0$, то вектор $(1, 0, 0, \ldots)$ или вектор $(0, 1, 0, \ldots)$ изотропен. Если $a_1a_2 \neq 0$, то по лем. 14.2 найдутся такие $\lambda, \mu \in \mathbb{F}_q$, что $a_1\lambda^2 + a_2\mu^2 = -a_3$. Тогда вектор $(\lambda, \mu, 1, 0, \ldots)$ изотропен. □

Предложение 14.4 (перечисление анизотропных форм)
Анизотропные формы над полем $\mathbb{F}_q$, где $q = p^m$ и $p > 2$, имеются только в размерностях 1 и 2. В размерности 2 квадратичная форма $x_1^2 + x_2^2$ анизотропна если и только если $q \equiv -1 \pmod 4$, а форма $x_1^2 + \varepsilon x_2^2$ анизотропна если и только если $q \equiv 1 \pmod 4$.

[1] См. предл. 13.6 на стр. 196.

Доказательство. Из прим. 14.1 на стр. 207 вытекает, что форма $x_1^2 + x_2^2$ имеет изотропный вектор если и только если её $D/4 = -1$ является квадратом в $\mathbb{F}_q$. В этом случае вторая форма $x_1^2 + \varepsilon x_2^2$ имеет $D/4 = -\varepsilon$, не являющееся квадратом, и тем самым анизотропна. Наоборот, если -1 не квадрат, то $-\varepsilon$ квадрат, и форма $x_1^2 + \varepsilon x_2^2$ имеет изотропный вектор. Остаётся убедиться, что -1 является квадратом в $\mathbb{F}_q$ если и только если $q \equiv 1 \pmod 4$. Для этого рассмотрим гомоморфизм мультипликативных групп $\gamma : \mathbb{F}_q^* \to \mathbb{F}_q^*$, $x \mapsto x^{\frac{q-1}{2}}$. Поскольку порядок $|\mathbb{F}_q^*| = q - 1$, для каждого $x \in \mathbb{F}_q^*$ выполняется равенство $x^{q-1} = 1$, из которого вытекает, что все ненулевые квадраты лежат в $\ker\gamma$, а все $x \in \operatorname{im}\gamma$ имеют $x^2 = 1$, откуда $\operatorname{im}\gamma \subset \{\pm 1\}$. Так как у уравнения $x^{\frac{q-1}{2}} = 1$ не более $(q-1)/2$ корней в поле $\mathbb{F}_q$, образ γ имеет порядок 2, а $\ker\gamma \subset \mathbb{F}_q^*$ имеет индекс 2 и совпадает с группой квадратов, т. е. $x \in \mathbb{F}_q^*$ является квадратом тогда и только тогда, когда $x^{\frac{q-1}{2}} = 1$. В частности, -1 квадрат если и только если $(q-1)/2$ чётно. □

14.5. Вещественные квадратичные формы. Из сл. 14.4 вытекает, что любая квадратичная форма на вещественном вектором пространстве V в подходящем базисе записывается в виде

$$q(x) = x_1^2 + x_2^2 + \cdots + x_p^2 - x_{p+1}^2 - x_{p+2}^2 - \cdots - x_{p+m}^2 . \tag{14-7}$$

Для этого надо перейти к базису с диагональной матрицей Грама и поделить каждый базисный вектор e_i с $q(e_i) \neq 0$ на $\sqrt{|q(e_i)|}$. Числа p и m в представлении (14-7) называются *положительным* и *отрицательным индексами инерции*, упорядоченная пара (p, m) — *сигнатурой*, а разность $p - m$ — просто *индексом* вещественной квадратичной формы q.

Теорема 14.6
Числа p и m в представлении (14-7) не зависят от выбора базиса, в котором квадратичная форма имеет вид (14-7).

Доказательство. Будем считать, что $p \geqslant m$, поскольку противоположный случай сводится к этому заменой q на $-q$. Сумма $p + m = \operatorname{rk} q$ равна рангу билинейной формы $\tilde{q}$ и не зависит от выбора базиса. Линейная оболочка базисных векторов e_k с номерами $k > p + m$ является ядром билинейной формы $\tilde{q}$. Классы $[e_i]$ остальных базисных векторов по модулю $\ker\tilde{q}$ образуют базис фактор пространства $W = V/\ker\tilde{q}$. По предл. 13.6 на стр. 196 форма $\tilde{q}$ корректно задаёт на W невырожденную симметричную билинейную форму $\tilde{q}_{\text{red}}([u],[w]) = \tilde{q}(u,w)$, которая в базисе из классов $[e_i]$ с $1 \leqslant i \leqslant p+m$ записывается той же самой формулой (14-7). Каждая пара базисных векторов $[e_i]$, $[e_{p+i}]$ порождает гиперболическую плоскость с гиперболическим базисом из векторов $\big([e_i] \pm [e_{p+i}]\big)/\sqrt{2}$. Поэтому форма $\tilde{q}_{\text{red}}$ является прямой ортогональной суммой гиперболического пространства H_{2m}, натянутого на классы $[e_i]$, $[e_{p+i}]$ с $1 \leqslant i \leqslant m$, и анизотропного пространства размерности $p - m$, натянутого на оставшиеся классы $[e_j]$ с $m < j \leqslant p$. По теор. 14.4 на стр. 205 размерности гиперболического и анизотропного слагаемых не зависят от выбора разложения пространства со скалярным произведением в ортогональную сумму гиперболического и анизотропного. Поэтому индекс $p - m$ и отрицательный индекс инерции m не зависят от выбора базиса, в котором форма q имеет вид (14-7). □

Следствие 14.6 (из доказательства теор. 14.6)
Для каждого n на пространстве $\mathbb{R}^n$ с точностью до изометрического изоморфизма имеются ровно два анизотропных скалярных произведения — евклидово и *антиевклидово*, получающееся

из евклидова сменой знака. Вещественные квадратичные формы положительного индекса имеют ненулевое евклидово анизотропное слагаемое, а формы отрицательного индекса — ненулевое антиевклидово анизотропное слагаемое, размерности которых равны абсолютной величине индекса. Гиперболичность невырожденной вещественной квадратичной формы равносильна тому, что её индекс равен нулю. □

Следствие 14.7
Два однородных многочлена второй степени $f, g \in \mathbb{R}[x_1, \dots, x_n]$ тогда и только тогда переводятся друг в друга линейными обратимыми заменами переменных, когда задаваемые ими квадратичные формы $f, g : \mathbb{R}^n \to \mathbb{R}$ имеют одинаковый ранг и индекс. □

14.5.1. Вычисление сигнатуры методом Якоби – Сильвестра. Обозначим через $V_k \subset \mathbb{R}^n$ линейную оболочку первых k базисных векторов $e_1, \dots, e_k$, а через Δ_k их определитель Грама, т. е. рассматриваемый с точностью до умножения на ненулевые положительные числа[1] главный угловой $k \times k$ минор матрицы Грама формы, сосредоточенный в первых k строках и столбцах. Если ограничение формы на подпространство V_k неособо, то знак $\operatorname{sgn} \Delta_k = (-1)^{m_k}$, где показатель m_k равен отрицательному индексу инерции ограничения формы на V_k. Таким образом, когда все $\Delta_i \neq 0$, соседние миноры Δ_k, Δ_{k+1} различаются знаком если и только если отрицательный индекс инерции $m_{k+1} = m_k + 1$. Поэтому полный отрицательный индекс инерции $m = m_n$ в этом случае равен числу перемен знака в последовательности $1, \Delta_1, \dots, \Delta_n$.

Если некоторый $\Delta_k = 0$, но при этом Δ_{k-1} и Δ_{k+1} оба ненулевые, то ограничения формы на подпространства V_{k+1} и V_{k-1}, а также на двумерное ортогональное дополнение W к подпространству V_{k-1} внутри V_{k+1} невырождены, и в W имеется изотропный вектор, порождающий ядро ограничения формы на подпространство V_k, где она вырождена. Тем самым, $W \simeq H_2$ является гиперболической плоскостью с сигнатурой $(1,1)$, и из ортогонального разложения $V_{k+1} = V_{k-1} \dot{+} W$ вытекает равенство $(p_{k+1}, m_{k+1}) = (p_{k-1} + 1, m_{k-1} + 1)$. Обратите внимание, что в этом случае Δ_{k-1} и Δ_{k+1} имеют противоположные знаки, т. е. при $\Delta_k = 0$ неравенство $\Delta_{k-1}\Delta_{k+1} > 0$ невозможно.

Если $\Delta_k = \Delta_{k+1} = 0$, но при этом $\Delta_{k-1}\Delta_{k+2} \neq 0$, то $V_{k+2} = V_{k-1} \dot{+} W$, где W — трёхмерное ортогональное дополнение к V_{k-1} внутри V_{k+2}. Как и выше, ограничение формы на W невырождено, и в W есть изотропный вектор. Поэтому W имеет сигнатуру $(2,1)$ или $(1,2)$ и

$$(p_{k+2}, m_{k+2}) = (p_{k-1} + 2, m_{k-1} + 1), \quad \text{если } \Delta_{k-1}\Delta_{k+2} < 0,$$
$$(p_{k+2}, m_{k+2}) = (p_{k-1} + 1, m_{k-1} + 2), \quad \text{если } \Delta_{k-1}\Delta_{k+2} > 0.$$

Итак, когда в последовательности $1, \Delta_1, \dots, \Delta_n$ не встречается более двух нулей подряд, прочтение её слева направо позволяет проследить за изменением сигнатуры (p_i, m_i) ограничения формы на пространства V_i с ненулевыми Δ_i и найти индекс.

Скажем, пусть $\Delta_1 < 0$, $\Delta_2 = 0$, $\Delta_3 > 0$, $\Delta_4 = 0$, $\Delta_5 = 0$, $\Delta_6 < 0$. Тогда

$$(p_1, m_1) = (0,1), \quad (p_3, m_3) = (2,1), \quad (p_6, m_6) = (3,3).$$

[1] Т. е. на ненулевые квадраты поля $\mathbb{R}$.

Примером такой формы является форма с матрицей Грама

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

14.5.2. Вычисление сигнатуры методом Гаусса. Над любым полем $\Bbbk$ перейти от произвольного базиса $e_1, \ldots, e_n$ к ортогональному базису заданной симметричной билинейной формы $\tilde{q}$ можно при помощи гауссовых элементарных преобразований базисных векторов[1]: перестановок каких-нибудь двух векторов e_i, e_j местами и замен одного из базисных векторов e_i на вектор $e_i' = e_i + \lambda e_j$, где $j \neq i$, а $\lambda \in \Bbbk$ произвольно, или на вектор $e_i' = \lambda e_i$, где $\lambda \in \Bbbk^*$ отлично от нуля. При перестановке местами векторов e_i, e_j в матрице Грама формы $\tilde{q}$ одновременно переставляются друг с другом i-я и j-я строки, а также i-й и j-й столбцы. Обратите внимание, что диагональные элементы $\tilde{q}(e_i, e_i)$ и $\tilde{q}(e_j, e_j)$ при этом переставятся друг с другом, а элементы $\tilde{q}(e_i, e_j) = \tilde{q}(e_j, e_i)$ останутся без изменения. Например, перестановка первого и третьего базисного вектора действует на симметричную 3×3 матрицу так:

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mapsto \begin{pmatrix} f & e & c \\ e & d & b \\ c & b & a \end{pmatrix}.$$

При замене вектора e_i вектором λe_i i-я строка и i-й столбец матрицы Грама одновременно умножаются на λ. Обратите внимание, что диагональный элемент $\tilde{q}(e_i, e_i)$ при этом умножится на λ^2. Например, замена e_2 на $2e_2$ подействует на предыдущую матрицу так:

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mapsto \begin{pmatrix} a & 2b & c \\ 2b & 4d & 2e \\ c & 2e & f \end{pmatrix}.$$

Наконец, замена e_i на $e_i' = e_i + \lambda e_j$ преобразует стоящие в i-й строке и i-м столбце недиагональные элементы $q_{ik} = \tilde{q}(e_i, e_k)$ и $q_{ki} = \tilde{q}(e_k, e_i)$ с $k \neq i$ в элементы $q_{ik}' = q_{ik} + \lambda q_{jk}$ и $q_{ki}' = q_{ki} + \lambda q_{kj}$ соответственно, а диагональный элемент $q_{ii} = \tilde{q}(e_i, e_i)$ — в

$$q_{ii}' = q_{ii} + \lambda q_{ij} + \lambda q_{ji} + \lambda^2 q_{jj}.$$

Иными словами, в матрице Грама к i-й строке прибавится j-я, умноженная на λ, и одновременно к i-у столбцу прибавится j-й, умноженный на λ, после чего к диагональному элементу в позиции[2] (i, i) добавится ещё диагональный элемент из позиции (j, j), умноженный на λ^2. Например, замена e_3 на $e_3 + 3e_2$ подействует на предыдущую матрицу так:

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mapsto \begin{pmatrix} a & b & c + 3b \\ b & d & e + 3d \\ c + 3b & e + 3d & f + 6e + 9d \end{pmatrix}.$$

[1] См. **??** на стр. **??**.

[2] Обратите внимание, что в текущий момент этот элемент уже увеличился на $\lambda q_{ij} + \lambda q_{ji} = 2\lambda q_{ij}$.

Метод Гаусса заключается в том, чтобы при помощи описанных трёх типов преобразований матрицы Грама превратить заданную симметричную матрицу в диагональную. Для вещественной формы количества положительных и отрицательных чисел на диагонали итоговой матрицы — это в точности положительный и отрицательный индексы инерции.

Для иллюстрации вычислим методом Гаусса сигнатуру вещественной квадратичной формы с матрицей Грама

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

Сначала обнулим 1-ю строку и 1-й столбец вне диагонали, добавляя к векторам e_2, e_4 соответственно векторы $2e_1$ и $-3e_1$:

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 6 & -1 & -6 \\ 0 & -1 & 0 & -2 \\ 0 & -6 & -2 & 9 \end{pmatrix}.$$

Теперь обнулим вне диагонали 2-ю строку и 2-й столбец, добавляя к текущим векторам e_3, e_4 соответственно текущие векторы $e_1/6$ и e_1:

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & -\frac{1}{6} & -3 \\ 0 & 0 & -3 & 3 \end{pmatrix}.$$

Наконец, обнулим вне диагонали 3-ю строку и 3-й столбец, добавляя к текущему вектору e_4 текущий вектор $-18e_3$:

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 57 \end{pmatrix}.$$

Таким образом, форма имеет сигнатуру $(2, 2)$.

14.6. Самосопряжённые операторы. Пусть на векторном пространстве V над произвольным полем $\Bbbk$ задана невырожденная симметричная билинейная форма

$$(*, *) \colon V \times V \to \Bbbk, \quad u, w \mapsto (u, w). \tag{14-8}$$

Будем называть линейный оператор $f : V \to V$ *самосопряжённым* относительно скалярного произведения (14-8), если $(fu, w) = (u, fw)$ при всех $u, w \in V$. Самосопряжённость оператора f равносильна тому, что при биекции между формами и операторами, которая задаётся скалярным произведением[1] (14-8), отвечающая оператору f билинейная форма $\beta_f(u, w) = (u, fw)$ является симметричной.

Упражнение 14.6. Убедитесь в этом.

На матричном языке самосопряжённость оператора f означает, что его матрица F в любом базисе пространства V связана с матрицей Грама G скалярного произведения (14-8) в том же базисе соотношением $F^t G = GF$.

[1] См. n° 13.2.4 на стр. 193.

Лемма 14.3
Пусть линейный оператор $f : V \to V$ самосопряжён. Тогда любые два собственных вектора оператора f с разными собственными значениями ортогональны друг другу, и для любого f-инвариантного подпространства $U \subset V$ ортогонал $U^\perp$ тоже f-инвариантен.

Доказательство. Если $fu = \lambda u$ и $fw = \mu w$, то из равенства $(fu, w) = (u, fw)$ вытекает равенство $(\lambda - \mu) \cdot (u, w) = 0$, откуда (u, v) при $\lambda \neq \mu$. Пусть $w \in U^\perp$, т. е. $(u, w) = 0$ для всех $u \in U$. Тогда $(u, fw) = (fu, w) = 0$ для всех $u \in U$, ибо $fu \in U$. Тем самым, $fw \in U^\perp$. □

Предложение 14.5
Если характеристический многочлен самосопряжённого линейного оператора $f : V \to V$ полностью раскладывается в поле $\Bbbk$ на линейные множители и все ненулевые собственные векторы оператора f анизотропны, то в пространстве V имеется ортогональный базис из собственных векторов оператора f.

Доказательство. Индукция по $\dim V$. Если оператор f является умножением на скаляр (что имеет место при $\dim V = 1$), то подойдёт любой ортогональный базис пространства V. Допустим, что $\dim V > 1$ и оператор f не скалярен. Поскольку характеристический многочлен $\det(tE - F)$ имеет корни в поле $\Bbbk$, у оператора F есть ненулевое собственное подпространство

$$V_\lambda = \{v \in V \mid fv = \lambda v\} \subsetneq V\,.$$

По условию леммы, оно анизотропно, и значит, скалярное произведение ограничивается на него невырождено. Поэтому $V = V_\lambda \oplus V_\lambda^\perp$, и ограничение скалярного произведения на $V_\lambda^\perp$ тоже невырождено. По лем. 14.3 оператор f переводит подпространство $V_\lambda^\perp$ в себя. Тем самым, характеристический многочлен оператора f является произведением характеристических многочленов ограничений $f|_{V_\lambda}$ и $f|_{V_\lambda^\perp}$. В силу единственности разложения на множители в кольце $\Bbbk[t]$ и предположения леммы, каждый из этих двух характеристических многочленов полностью раскладываются на линейные множители в поле $\Bbbk$. По индуктивному предположению, в подпространстве $V_\lambda^\perp$ есть ортогональный базис из собственных векторов оператора f. Добавляя к нему любой ортогональный базис собственного пространства V_λ, получаем нужный базис в V. □

14.7. Грассмановы квадратичные формы. Покажем, что каждый ненулевой однородный грассманов многочлен[1] второй степени $\omega \in \Lambda^2 V$ на конечномерном пространстве V над любым полем $\Bbbk$ в подходящем базисе $\boldsymbol{e}$ пространства V может быть записан в *нормальном виде Дарбу*

$$e_1 \wedge e_2 + e_3 \wedge e_4 + \cdots + e_{2r-1} \wedge e_{2r}\,. \tag{14-9}$$

Для этого рассмотрим произвольный базис $\boldsymbol{u}$ и перенумеруем его векторы так, чтобы

$$\omega = u_1 \wedge (\alpha_2 u_2 + \cdots + \alpha_n u_n) + u_2 \wedge (\beta_3 u_3 + \cdots + \beta_n u_n) + (\text{члены без } u_1 \text{ и } u_2)\,,$$

где коэффициент $\alpha_2 \neq 0$ и вектор $v_2 \stackrel{\text{def}}{=} \alpha_2 u_2 + \cdots + \alpha_n u_n \neq 0$. Перейдём к новому базису $\boldsymbol{v}$ из векторов $v_i = u_i$ при $i \neq 2$ и вектора v_2.

Упражнение 14.7. Убедитесь, что это действительно базис.

[1] См. n° 8.4 на стр. 116.

Подставляя в предыдущую формулу $u_2 = \bigl(v_2 - \alpha_3 v_3 - \cdots - \alpha_n v_n\bigr)/\alpha_2$, получаем

$$\omega = v_1 \wedge v_2 + v_2 \wedge (\gamma_3 v_3 + \cdots + \gamma_n v_n) + (\text{члены без } v_1 \text{ и } v_2) = \\ = (v_1 - \gamma_3 v_3 - \cdots - \gamma_n v_n) \wedge v_2 + (\text{члены без } v_1 \text{ и } v_2)$$

для некоторых $\gamma_3, \ldots, \gamma_n \in \Bbbk$. Переходя к базису $\boldsymbol{w}$ из векторов $w_1 = v_1 - \gamma_3 v_3 - \cdots - \gamma_n v_n$ и $w_i = v_i$ при $i \neq 1$, получаем $\omega = w_1 \wedge w_2 +$ (члены без w_1 и w_2), после чего процесс может быть продолжен по индукции.

Следствие 14.8
Над полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk \neq 2$ однородный грассманов многочлен $\omega \in \Lambda^2 V$ тогда и только тогда разложим в произведение $u \wedge w$ двух векторов $u, w \in V$, когда $\omega \wedge \omega = 0$.

Доказательство. Если $\omega = u \wedge w$, то $\omega \wedge \omega = u \wedge w \wedge u \wedge w = 0$. Чтобы получить обратное, выберем в V базис $\boldsymbol{e}$, в котором $\omega = e_1 \wedge e_2 + e_3 \wedge e_4 + \cdots$. Если в этой сумме есть хотя бы два слагаемых, то базисный моном $e_1 \wedge e_2 \wedge e_3 \wedge e_4$ войдёт в $\omega \wedge \omega$ с ненулевым коэффициентом 2, а значит, $\omega \wedge \omega \neq 0$. Таким образом, равенство $\omega \wedge \omega = 0$ влечёт равенство $\omega = e_1 \wedge e_2$. □

14.7.1. Поляризация грассмановой квадратичной формы. Напомню[1], что с каждым базисом $\boldsymbol{e} = (e_1, \ldots, e_n)$ пространства V связан базис в $\Lambda^2 V$, состоящий из $n(n-1)/2$ грассмановых мономов $e_{ij} = e_i \wedge e_j$ с $i < j$, и каждый однородный грассманов многочлен второй степени $\omega \in \Lambda^2 V$ однозначно представляется в виде

$$\omega = \sum_{i<j} \omega_{ij}\, e_{ij}\,, \quad \text{где} \quad \omega_{ij} \in \Bbbk\,, \tag{14-10}$$

и суммирование происходит по всем $1 \leqslant i < j \leqslant n$. Если $\operatorname{char}\Bbbk \neq 2$, то подобно тому, как это делалось для коммутативных квадратичных форм[2], каждое слагаемое в (14-10) можно переписать в виде $\omega_{ij} e_{ij} = \omega'_{ij} e_i \wedge e_j + \omega'_{ji} e_j \wedge e_i$, где $\omega'_{ij} = -\omega'_{ji} = \omega_{ij}/2$. Составленная из чисел ω'_{ij} кососимметричная квадратная матрица $\Omega_{\boldsymbol{e}} = (\omega'_{ij}) \in \operatorname{Mat}_n(\Bbbk)$ называется *матрицей Грама* грассмановой квадратичной формы ω в базисе $\boldsymbol{e}$. В терминах матрицы Грама форма ω записывается в виде

$$\omega = \sum_{i,j=1}^{n} \omega'_{ij} e_i \wedge e_j = \bigl(\boldsymbol{e}\,\Omega_{\boldsymbol{e}}\bigr) \wedge \boldsymbol{e}^t\,, \tag{14-11}$$

где в отличие от (14-10) суммирование происходит по всем n^2 парам индексов i, j, а обозначение $A \wedge B$ для матриц A, B, элементами которых являются векторы, предписывает перемножить эти матрицы по обычному правилу, используя в качестве произведения матричных элементов грассманово произведение соответствующих векторов, т. е. в (i,j)-й позиции матрицы $A \wedge B$ стоит вектор $a_{i1} \wedge b_{1j} + a_{i2} \wedge b_{2j} + \cdots + a_{in} \wedge b_{nj}$. При выборе в V другого базиса $\boldsymbol{f}$, через который базис $\boldsymbol{e}$ выражается по формуле $\boldsymbol{e} = \boldsymbol{f}\, C_{\boldsymbol{fe}}$, грассманова квадратичная форма ω перепишется в терминах новых базисных векторов в виде $\omega = \bigl(\boldsymbol{e}\,\Omega_{\boldsymbol{e}}\bigr) \wedge \boldsymbol{e}^t = \bigl(\boldsymbol{f}\, C_{\boldsymbol{fe}} \Omega_{\boldsymbol{e}}\bigr) \wedge \bigl(C^t_{\boldsymbol{fe}} \boldsymbol{f}^t\bigr) = \bigl(\boldsymbol{f}\, C_{\boldsymbol{fe}} \Omega_{\boldsymbol{e}} C^t_{\boldsymbol{fe}}\bigr) \wedge \boldsymbol{f}^t$. В частности, матрица Грамма $\Omega_{\boldsymbol{f}}$ формы ω в базисе $\boldsymbol{f}$ выразится через матрицу Грама $\Omega_{\boldsymbol{e}}$ как

$$\Omega_{\boldsymbol{e}} = C_{\boldsymbol{fe}} \Omega_{\boldsymbol{e}} C^t_{\boldsymbol{fe}}\,. \tag{14-12}$$

[1] См. 8-16 на стр. 117.
[2] Ср. с n° 14.3 на стр. 205.

Пример 14.2 (нормальная форма Дарбу)
Если $\operatorname{char}\Bbbk \neq 2$, то существование базиса $\boldsymbol{e}$, в котором заданная грассманова квадратичная форма $\omega \in \Lambda^2 V$ имеет вид (14-9), вытекает из теоремы о приведении кососимметричной билинейной формы к нормальному виду Дарбу[1]. Действительно, доказывая эту теорему, мы установили, что для любой кососимметричной матрицы Ω существует такая обратимая матрица C, что все ненулевые элементы матрицы $C\Omega C^t$ сосредоточены в расположенных на главной диагонали 2×2-блоках вида $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Поэтому грассманова квадратичная форма, имеющая матрицу Грама Ω в некотором базисе $\boldsymbol{f}$, запишется в базисе $\boldsymbol{g} = \boldsymbol{f}\,C$ как $\omega = 2g_1\wedge g_2 + 2g_3\wedge g_4 + \ldots$ Искомый базис $\boldsymbol{e}$ получается из $\boldsymbol{g}$ удвоением векторов с нечётными номерами: $e_{2i+1} = 2g_{2i+1}$, $e_{2i} = g_{2i}$.

14.7.2. Пфаффиан. Рассмотрим кососимметричную матрицу $B = \left(b_{ij}\right)$ размера $2n\times 2n$, наддиагональные элементы b_{ij} с $i<j$ которой будем считать независимыми переменными, и обозначим через $\mathbb{Z}[b_{ij}]$ кольцо многочленов с целыми коэффициентами от этих $2n^2-n$ переменных. Сопоставим, как и выше, матрице B грассманову квадратичную форму с коэффициентами в кольце $\mathbb{Z}[b_{ij}]$ от $2n$ грассмановых переменных $\xi = (\xi_1,\ldots,\xi_{2n})$:

$$\beta_B(\xi) \stackrel{\text{def}}{=} (\xi B)\wedge\xi^t = \sum_{ij} b_{ij}\xi_i\wedge\xi_j\,.$$

Поскольку чётные мономы $\xi_i\wedge\xi_j$ лежат в центре грассмановой алгебры, n-тая грассманова степень этой формы имеет вид

$$\begin{aligned}\beta_B(\xi)^{\wedge n} &= \beta_B(\xi)\wedge\cdots\wedge\beta_B(\xi) = \\ &= \Big(\sum_{i_1 j_1} b_{i_1 j_1}\xi_{i_1}\wedge\xi_{j_1}\Big)\wedge\Big(\sum_{i_2 j_2} b_{i_2 j_2}\xi_{i_2}\wedge\xi_{j_2}\Big)\wedge\cdots\wedge\Big(\sum_{i_n j_n} b_{i_n j_n}\xi_{i_n}\wedge\xi_{j_n}\Big) = \\ &= 2^n n!\sum_{\substack{\{i_1,j_1\}\sqcup\cdots\sqcup\{i_n,j_n\}=\\ =\{1,2,\ldots,2n\}}} \operatorname{sgn}\big(i_1 j_1 i_2 j_2\ldots i_n j_n\big) b_{i_1 j_1} b_{i_2 j_2}\cdots b_{i_n j_n}\xi_1\wedge\xi_2\wedge\cdots\wedge\xi_{2n}\,,\end{aligned} \tag{14-13}$$

где в суммирование происходит по всем разбиениям множества $\{1,\,2,\,\ldots\,,\,2n\}$ в объединение n неупорядоченных непересекающихся двухэлементных множеств $\{i_\nu, j_\nu\}$, порядок внутри которых тоже не существен, а sgn означает знак указанной в его аргументе перестановки из симметрической группы S_{2n}.

Упражнение 14.8. Убедитесь, что этот знак не меняется при перестановках пар друг с другом, а вся правая часть формулы (14-13) не меняется при перестановке элементов внутри любой из пар.

Сумма из правой части формулы (14-13) называется *пфаффианом* кососимметричной матрицы B и обозначается

$$\operatorname{Pf}(B) \stackrel{\text{def}}{=} \sum_{\substack{\{i_1,j_1\}\sqcup\cdots\sqcup\{i_n,j_n\}=\\ =\{1,2,\ldots,2n\}}} \operatorname{sgn}\big(i_1 j_1 i_2 j_2\ldots i_n j_n\big) b_{i_1 j_1} b_{i_2 j_2}\cdots b_{i_n j_n} \in \mathbb{Z}[b_{ij}]\,. \tag{14-14}$$

[1] См. теор. 13.5 на стр. 199.

Например,

$$\operatorname{Pf}\begin{pmatrix} 0 & b_{12} \\ -b_{12} & 0 \end{pmatrix} = b_{12}\,, \quad \operatorname{Pf}\begin{pmatrix} 0 & b_{12} & b_{13} & b_{14} \\ -b_{12} & 0 & b_{23} & b_{24} \\ -b_{13} & -b_{23} & 0 & b_{34} \\ -b_{14} & -b_{24} & -b_{34} & 0 \end{pmatrix} = b_{12}b_{23} - b_{13}b_{24} + b_{14}b_{23}\,.$$

Упражнение 14.9. Проверьте, что в обоих случаях $(\operatorname{Pf} B)^2 = \det B$.

Если в левой и правой части формулы (14-13) заменить грассмановы переменные ξ на новые грассмановы переменные η по формуле $\xi = \eta\, C$, где $C = (c_{ij})$ — квадратная матрица размера $2n \times 2n$ все элементы которой мы будем считать независимыми переменными, лежащими в поле $\mathbb{K} = \mathbb{Q}(c_{ij})$ рациональных функций с рациональными коэффициентами от $4n^2$ переменных c_{ij}, то в правой части (14-13) мы получим $2^n n!\operatorname{Pf}(B)\ \det C\ \eta_1 \wedge \eta_2 \wedge \cdots \wedge \eta_{2n}$. Квадратичная форма $\beta_B(\xi)$ в левой части (14-13) заменится на форму

$$\beta_B(\xi) = (\xi\, B) \wedge \xi^t = (\eta\, CB) \wedge (\eta\, C)^t = (\eta\, CBC^t) \wedge \eta^t = \beta_{CBC^t}(\eta)\,,$$

с n-той грассмановой степенью $\beta_{CBC^t}(\eta)^{\wedge n} = 2^n n!\ \operatorname{Pf}(CBC^t)\,\eta_1 \wedge \eta_2 \wedge \cdots \wedge \eta_{2n}$. Таким образом, для любой матрицы $C \in \mathrm{GL}_{2n}(\mathbb{K})$ в кольце многочленов $\mathbb{K}[b_{ij}]$ от переменных b_{ij} с коэффициентами в поле $\mathbb{K}$ выполняется равенство

$$\operatorname{Pf}(CBC^t) = \operatorname{Pf}(B)\ \det C\,. \tag{14-15}$$

Теорема 14.7
Обозначим через J' блочно диагональную $2n \times 2n$ матрицу из n расположенных на главной диагонали 2×2-блоков

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

и нулями в остальных местах, через $\mathbb{Z}[a_{ij}]$ — кольцо многочленов с целыми коэффициентами от $2n^2 - n$ переменных a_{ij} с $1 \leqslant i < j \leqslant 2n$, а через $A = (a_{ij})$ — кососимметричную матрицу размера $(2n) \times (2n)$, наддиагональными элементами которой являются переменные a_{ij}. Тогда $\operatorname{Pf}(A)$ — это единственный такой многочлен в $\mathbb{Z}[a_{ij}]$, что $\operatorname{Pf}^2(A) = \det(A)$ и $\operatorname{Pf}(J') = 1$.

Доказательство. Обозначим через $\mathbb{F} = \mathbb{Q}(a_{ij})$ поле рациональных функций от переменных a_{ij} с коэффициентами в поле $\mathbb{Q}$ и рассмотрим на координатном векторном пространстве $\mathbb{F}^{2n}$ невырожденную кососимметричную форму с матрицей Грама A в стандартном базисе. По теореме Дарбу[1] в K^{2n} есть базис, где эта форма имеет матрицу Грама J'. Поэтому $A = MJ'M^t$ для некоторой матрицы $M \in \mathrm{GL}_{2n}(\mathbb{F})$. Поскольку $\det J' = 1$, определитель $\det(A) = \det^2(M)$. С другой стороны, подставляя $C = M$ и $B = J'$ в равенство (14-15), заключаем, что $\det M = \operatorname{Pf}(A)$.

Упражнение 14.10. Убедитесь, что $\operatorname{Pf}(J') = 1$.

Единственность вытекает из того, что многочлен $x^2 - \det A = \big(x - \operatorname{Pf}(A)\big)\big(x + \operatorname{Pf}(A)\big)$ имеет в целостном кольце $\mathbb{Z}[a_{ij}]$ ровно два корня $x = \pm\operatorname{Pf}(A)$, и требование $\operatorname{Pf}(J') = 1$ однозначно фиксирует ровно один из них. □

[1] См. теор. 13.5 на стр. 199.

§15. Эрмитовы пространства

Всюду в этом параграфе речь идёт про конечномерные векторные пространства над полем $\mathbb{C}$.

15.1. Эрмитова геометрия. Векторное пространство W над полем комплексных чисел $\mathbb{C}$ называется *эрмитовым* (или *унитарным*), если на нём задано билинейное над подполем $\mathbb{R} \subset \mathbb{C}$ скалярное произведение $W \times W \to \mathbb{C}$, обозначаемое (w_1, w_2) или $w_1 \cdot w_2$ и обладающее следующими тремя свойствами:

$$\begin{aligned} \forall\, w_1, w_2,\ \forall\, z \quad (zw_1, w_2) = z\,(w_1, w_2) = (w_1, \overline{z}w_2) \quad & \text{(полуторалинейность)} \\ \forall\, w_1, w_2 \quad (w_1, w_2) = \overline{(w_2, w_1)} \quad & \text{(эрмитова симметричность)} \\ \forall\, w \neq 0 \quad (w, w) > 0 \quad & \text{(положительность)}\,, \end{aligned} \tag{15-1}$$

В силу второго свойства, скалярный квадрат $(w, w) = \overline{(w, w)}$ любого вектора $w \in W$ является вещественным числом, а последнее свойство утверждает, что это вещественное число положительно для всех ненулевых векторов. Скалярное произведение со свойствами (15-1) называется *эрмитовой* (или *унитарной*) *структурой* на комплексном векторном пространстве W.

Пример 15.1 (стандартная эрмитова структура на $\mathbb{C}^n$)
Координатное пространство $\mathbb{C}^n$ имеет *стандартную эрмитову структуру*, в которой строки $z = (z_1, \dots, z_n)$ и $w = (w_1, \dots, w_n)$ перемножаются по формуле

$$(z, w) = z_1\overline{w}_1 + \cdots + z_n\overline{w}_n\,. \tag{15-2}$$

Пример 15.2 (эрмитова структура на пространстве комплекснозначных функций)
На пространстве непрерывных функций $[a, b] \to \mathbb{C}$ на отрезке $[a, b] \subset \mathbb{R}$ имеется эрмитово скалярное произведение

$$(f, g) = \int_a^b f(x)\overline{g}(x)\,dx\,, \tag{15-3}$$

где под интегралом от комплекснозначной функции f по определению понимается комплексное число, действительная и мнимая части которого равны интегралам от вещественной и мнимой частей функции f, которые являются обычными вещественными функциями:

$$\int_a^b f\,dx = \int_a^b \mathrm{Re}(f)\,dx + i \cdot \int_a^b \mathrm{Im}(f)\,dx\,.$$

Разумеется, вместо отрезка можно рассматривать любое другое пространство, по которому можно интегрировать вещественные функции, например диск или какую-нибудь кривую в $\mathbb{C}$.

15.1.1. Эрмитова норма вектора. Пользуясь тем, что скалярный квадрат любого вектора в эрмитовом пространстве вещественен и положителен, определим *эрмитову норму* (или *длину*) вектора $w \in W$ формулой[1]

$$\|w\| \stackrel{\text{def}}{=} \sqrt{(w, w)} \in \mathbb{R}_{\geqslant 0}\,. \tag{15-4}$$

[1]Мы используем обозначение $\|w\|$, чтобы отличать нормы *векторов* $w \in W$ от модулей комплексных *чисел* $z \in \mathbb{C}$, которые будем обозначать, как и раньше, через $|z| = \sqrt{z \cdot \overline{z}}$.

Эрмитова структура однозначно восстанавливается по норме, так как из равенств

$$(w_1 + w_2, w_1 + w_2) = \|w_1\|^2 + \|w_2\|^2 + 2\,\mathrm{Re}(w_1, w_2)$$
$$(w_1 + iw_2, w_1 + iw_2) = \|w_1\|^2 + \|w_2\|^2 + 2\,\mathrm{Im}(w_1, w_2)\,,$$

вытекает равенство

$$2(w_1, w_2) = \|w_1 + w_2\|^2 - \|w_1\|^2 - \|w_2\|^2 + i(\|w_1 + iw_2\|^2 - \|w_1\|^2 - \|w_2\|^2)\,. \tag{15-5}$$

15.1.2. Ортогонализация Грама – Шмидта. Базис $e_1, \ldots, e_n$ эрмитова пространства W называется *ортонормальным*, если $\|e_i\| = 1$ при всех i и $(e_i, e_j) = 0$ при всех $i \neq j$. Так же, как и в евклидовом пространстве, из любого базиса $w_1, \ldots, w_n$ в W можно изготовить такой ортонормальный базис $e_1, \ldots, e_n$, линейная оболочка первых первых k базисных векторов которого при каждом k совпадает с линейной оболочкой первых первых k базисных векторов исходного базиса. Векторы e_i находятся по рекурсивным формулам

$$\begin{gathered} e_1 = w_1/\sqrt{(w_1, w_1)} \quad \text{и} \quad e_k = u_k/\sqrt{(u_k, u_k)} \text{ при } k > 1\,, \\ \text{где } u_k = w_k - \sum_{v=1}^{k-1} (w_k, e_v)\,. \end{gathered} \tag{15-6}$$

Упражнение 15.1. Убедитесь, что $\mathrm{span}(u_1, \ldots, u_k) = \mathrm{span}(w_1, \ldots, w_k)$ при каждом k (в частности, все $u_k \neq 0$) и векторы $e_1, \ldots, e_n$ действительно образуют ортонормальный базис.

15.1.3. Матрицы Грама. Эрмитова симметричность скалярного произведения означает, что матрица Грама $G_{\boldsymbol{w}} = \big((w_i, w_j)\big) = \boldsymbol{w}^t \cdot \boldsymbol{w}$ любого набора векторов $\boldsymbol{w} = (w_1, \ldots, w_m)$ пространства W *эрмитово симметрична*, т. е. комплексно сопряжена транспонированной матрице:

$$G_{\boldsymbol{w}}^t = \overline{G}_{\boldsymbol{w}}\,.$$

В силу полуторалинейности эрмитова скалярного произведения по второму аргументу, при линейной замене набора векторов по формуле $\boldsymbol{w} = \boldsymbol{u}\,C_{\boldsymbol{uw}}$ матрица Грама меняется по правилу

$$G_{\boldsymbol{w}} = \boldsymbol{w}^t \cdot \boldsymbol{w} = (C_{\boldsymbol{uw}}^t \boldsymbol{u}^t) \cdot (\boldsymbol{u}\,C_{\boldsymbol{uw}}) = C_{\boldsymbol{uw}}^t (\boldsymbol{u}^t \cdot \boldsymbol{u})\,\overline{C}_{\boldsymbol{uw}} = C_{\boldsymbol{uw}}^t G_{\boldsymbol{u}} \overline{C}_{\boldsymbol{uw}}\,.$$

Ортонормальность набора векторов $\boldsymbol{e}$ означает, что его матрица Грама $G_{\boldsymbol{e}} = E$ единичная.

Лемма 15.1
Определитель Грама $\det G_{\boldsymbol{w}}$ любого набора векторов $\boldsymbol{w} = (w_1, \ldots, w_m)$ является вещественным неотрицательным числом и обращается в нуль если и только если этот набор линейно зависим.

Доказательство. Пусть $\boldsymbol{w} = \boldsymbol{e}\,C_{\boldsymbol{ew}}$, где набор векторов $\boldsymbol{e} = (e_1, \ldots, e_n)$ составляет ортонормальный базис в линейной оболочке набора векторов $\boldsymbol{w}$. Тогда $G_{\boldsymbol{w}} = C_{\boldsymbol{ew}}^t \overline{C}_{\boldsymbol{ew}}$. Если набор w линейно зависим, то $n < m$ и $\mathrm{rk}\, G_{\boldsymbol{w}} \leqslant \min(\mathrm{rk}\, C_{\boldsymbol{ew}}^t, \mathrm{rk}\, \overline{C}_{\boldsymbol{ew}}) \leqslant n$ строго меньше размера матрицы, т. е. $\det G_{\boldsymbol{w}} = 0$. Если векторы $\boldsymbol{w}$ составляют базис своей линейной оболочки, то матрица $C_{\boldsymbol{ew}}$ невырождена, и $\det G_{\boldsymbol{w}} = \det C_{\boldsymbol{ew}}\,\overline{\det C_{\boldsymbol{ew}}} = |\det C_{\boldsymbol{ew}}|^2$ вещественен и положителен. □

15.1.4. Эрмитово двойственный базис. Для любого базиса $\boldsymbol{u} = (u_1, \dots, u_n)$ эрмитова пространства W существует единственный *эрмитово двойственный* базис $\boldsymbol{u}^\times = (u_1^\times, \dots, u_n^\times)$ со свойством

$$(u_i, u_j^\times) = \begin{cases} 0 & \text{при } i \neq j \\ 1 & \text{при } i = j\,. \end{cases} \tag{15-7}$$

На матричном языке эти соотношения означают, что взаимная матрица Грама

$$G_{\boldsymbol{u}\boldsymbol{u}^\times} = \boldsymbol{u}^t \cdot \boldsymbol{u}^\times = E\,.$$

Подставляя сюда $\boldsymbol{u}^\times = \boldsymbol{u}\, C_{\boldsymbol{u}\boldsymbol{u}^\times}$, где в j-м столбце матрицы $C_{\boldsymbol{u}\boldsymbol{u}^\times}$ стоят координаты вектора $u_j^\times$ в базисе $\boldsymbol{u}$, и пользуясь полулинейностью скалярного произведения по второму аргументу, получаем $E = \boldsymbol{u}^t \cdot \boldsymbol{u}^\times = (\boldsymbol{u}^t \cdot \boldsymbol{u})\, \overline{C}_{\boldsymbol{u}\boldsymbol{u}^\times} = G_{\boldsymbol{u}} \overline{C}_{\boldsymbol{u}\boldsymbol{u}^\times}$, откуда

$$(u_1^\times, \dots, u_n^\times) = (u_1, \dots, u_n)\, \overline{G}_{\boldsymbol{u}}^{-1}\,. \tag{15-8}$$

Упражнение 15.2. Убедитесь, что $\boldsymbol{u}^{\times\times} = \boldsymbol{u}$.

Из соотношений ортогональности (15-7) вытекает, что в разложении $w = \sum z_i u_i$ произвольного вектора $w \in W$ по базису $u_1, \dots, u_n$ коэффициент $z_i = (w, u_i^\times)$, в чём легко удостовериться, скалярно умножив обе части разложения справа на $u_i^\times$. Таким образом,

$$w = \sum_i u_i \cdot (w, e_i^\times)\,. \tag{15-9}$$

Обратите внимание, что ортонормальность базиса $\boldsymbol{e}$ равносильна равенству $\boldsymbol{e}^\times = \boldsymbol{e}$.

15.1.5. Неравенства Коши – Буняковского – Шварца и треугольника. Применяя лем. 15.1 к набору из двух векторов u, w, мы заключаем, что

$$\det \begin{pmatrix} (u,u) & (u,w) \\ (w,u) & (w,w) \end{pmatrix} = \|u\|^2 \|w\|^2 - (u,w)\,\overline{(u,w)} \geqslant 0\,,$$

где равенство равносильно пропорциональности векторов u и w над полем $\mathbb{C}$. Таким образом, в эрмитовом пространстве выполняется *неравенство Коши – Буняковского – Шварца*

$$|(u,w)| \leqslant \|u\| \cdot \|w\|\,, \tag{15-10}$$

равенство в котором равносильно комплексной пропорциональности u и w. Далее,

$$\|u+w\|^2 = (u+w, u+w) = \|u\|^2 + \|w\|^2 + 2\mathrm{Re}\,(u,w) \leqslant$$
$$\leqslant \|u\|^2 + \|w\|^2 + 2|(u,w)| \leqslant \|u\|^2 + \|w\|^2 + 2\,\|u\|\,\|w\| = \big(\|w_1\| + \|w_2\|\big)^2\,,$$

где второе неравенство — это неравенство Коши – Буняковского – Шварца, а первое неравенство $2\mathrm{Re}\,(u,w) \leqslant |(v,w)|$ для комплексно пропорциональных векторов $u = zw$ обращается в равенство если и только если коэффициент пропорциональности $z \in \mathbb{R}$ и неотрицателен, ибо $\mathrm{Re}(zw, w) = (w,w)\,\mathrm{Re}(z)$, т. к. $(w,w) \in \mathbb{R}$ и неотрицательно. Мы заключаем, что для любых двух векторов u, w эрмитова пространства выполняется *неравенство треугольника*

$$\|u\| + \|w\| \geqslant \|u+w\|\,, \tag{15-11}$$

становящееся равенством если и только если $u = \lambda w$ с вещественным неотрицательным λ.

15.1.6. Угол между комплексными прямыми. Напомню, что на вещественной евклидовой плоскости угол φ, равный меньшему из двух смежных углов между прямыми, параллельными векторам u и w, имеет

$$\cos\varphi = \frac{|(u,w)|}{\|u\|\cdot\|w\|} = |(u/\|u\|, w/\|w\|)|\,. \tag{15-12}$$

На геометрическом языке, векторы $u/\|u\|$ и $w/\|w\|$ являются единичными направляющими векторами рассматриваемых прямых, и каждый из них определяется этим свойством однозначно с точностью до умножения на ±1. Выбор знаков определяет выбор одного из четырёх углов, на которые эти прямые разбивают плоскость, и меньший из углов получается при таком выборе знаков, что скалярное произведение неотрицательно.

Двумерное комплексное пространство, натянутое на непропорциональные векторы u, w эрмитова пространства W с вещественной точки зрения представляет собой четырёхмерное пространство $\mathbb{R}^4$, в котором комплексные прямые $\mathbb{C}u$ и $\mathbb{C}w$ образуют пару трансверсальных двумерных вещественных плоскостей с нулевым пересечением. Объемлющее пространство $\mathbb{R}^4$ не разбивается этими плоскостями ни на какие связные компоненты, и в каждой из плоскостей концы векторов единичной длины пробегают единичную окружность. Эти две окружности не пересекаются и лежат на компактной трёхмерной сфере

$$S^3 = \{\, u \in \mathbb{R}^4 = \mathbb{C}u \oplus \mathbb{C}w \mid \|u\| = 1 \,\},$$

состоящей из векторов единичной длины в $\mathbb{R}^4$. Поэтому длины больших дуг[1], соединяющих точку на одной из окружностей с точкой на другой, ограничены снизу и достигают своего минимального значения. Иначе говоря, угол между вещественными прямыми $\mathbb{R}\cdot e_1$ и $\mathbb{R}\cdot e_2$, параллельными всевозможным векторам $e_1 \in \mathbb{C}u$ и $e_2 \in \mathbb{C}w$ единичной длины, достигает на некоторой паре векторов своего минимума. Такой минимальный угол φ и называется углом между (комплексными) одномерными подпространствами $\mathbb{C}u$ и $\mathbb{C}w$ в эрмитовом пространстве W.

Предложение 15.1
Косинус угла φ между натянутыми на векторы u и w одномерными подпространствами эрмитова пространства W вычисляется по той же формуле (15-12), что и в евклидовом пространстве:

$$\cos\varphi = \frac{|(u,w)|}{\|u\|\cdot\|w\|}\,. \tag{15-13}$$

Доказательство. Запишем эрмитово скалярное произведение на пространстве W в виде

$$(u,w) = g(u,w) + i\omega(u,w)\,,$$

где $g(u,w) \stackrel{\text{def}}{=} \operatorname{Re}(u,w)$ и $\omega(u,w) \stackrel{\text{def}}{=} \operatorname{Im}(u,w)$ суть вещественная и мнимая части комплексного числа (u,w). Форма $g\colon W\times W \to \mathbb{R}$ вещественно билинейна, симметрична и положительна. Она задаёт на вещественном пространстве $\mathbb{R}^4 = \mathbb{C}u \oplus \mathbb{C}w$ евклидову структуру, в которой евклидова длина каждого вектора совпадает с его эрмитовой длиной, ибо $g(u,u) = (u,u)$ для всех $u \in W$. Форма $\omega(u_1,u_2)$ вещественно билинейна, кососимметрична (ибо $\omega(u,u) = 0$ для всех $u \in W$), и невырождена, так как $\omega(iu,u) = g(u,u) = \|u\|^2 \neq 0$ для любого $u \neq 0$. Когда вектор e_1 пробегает окружность векторов единичной длины на вещественной плоскости $\mathbb{C}u$, а

[1] Т. е. дуг, высекаемых на сфере S^3 всевозможными двумерными вещественными плоскостями, проходящими через центр этой сферы.

вектор e_2 — такую же окружность на плоскости $\mathbb{C}w$, сумма $g^2(e_1,e_2)+\omega^2(e_1,e_2)=|(e_1,e_2)|^2$ не меняется, так как для всех $t,s\in\mathbb{C}$ с $|t|=|s|=1$ имеем $|(te_1,se_2)|=|t\overline{s}\,(e_1,e_2)|=|(e_1,e_2)|$. Минимальный из евклидовых углов φ между векторами e_1 и e_2 имеет максимально возможный $\cos^2\varphi = g^2(e_1,e_2)$. *A priori* максимальным значением для $g^2(e_1,e_2)$ является константа $|(e_1,e_2)|^2$. Это значение достигается векторах e_1, e_2 если и только если $\omega(e_1,e_2)=0$. Так как форма ω невырождена, ω-ортогонал $v_\omega^\perp=\{v'\in\mathbb{R}^4\mid \omega(v,v')=0\}$ к любому ненулевому вектору $v\in\mathbb{R}^4$ является трёхмерной вещественной гиперплоскостью в $\mathbb{R}^4$ и имеет ненулевое пересечение с двумерным вещественным подпространством $\mathbb{C}w$. Мы заключаем, что для каждого единичного вектора $e_1\in\mathbb{C}u$ евклидов угол между векторами e_1 и e_2 достигает своего минимального значения на некотором единичном векторе $e_2\in\mathbb{C}w$, и косинус такого минимального угла равен равен $|(e_1,e_2)|=|(u/\|u\|,w/\|w\|)|$. □

Замечание 15.1. В силу неравенства Коши – Буняковского – Шварца правая часть в (15-13) принимает значения на отрезке $[0,1]$, откуда $0\leqslant\varphi\leqslant\pi/2$.

15.1.7. Унитарная группа. Линейный оператор $f\colon W\to W$ на эрмитовом пространстве W называется *унитарным*, если $\|fw\|=\|w\|$ для всех $w\in W$. Согласно формуле (15-5), каждый такой оператор f сохраняет скалярное произведение: $(fu,fw)=(u,w)$ для всех $u,w\in W$. Тем самым, матрица F унитарного оператора f в любом базисе связана с матрицей Грама G этого базиса соотношением

$$F^t\cdot G\cdot\overline{F}=G\,. \tag{15-14}$$

Беря определители и сокращая на $\det G\neq 0$, получаем $\det^2 F=1$, откуда $|\det F|=1$. В частности, каждый унитарный оператор f обратим, причём матрица обратного оператора в любом базисе выражается через F и G по формуле $F^{-1}=\overline{G}^{-1}\overline{F}^t\overline{G}=(G^t)^{-1}\overline{F}^tG^t$, которая в ортонормальном базисе с $G=E$ редуцируется до $F^{-1}=\overline{F}^t$.

Унитарные операторы составляют *унитарную группу* пространства W, которая обозначается $\mathrm{U}(W)$. Запись унитарных операторов матрицами в каком-нибудь ортонормальном базисе $e_1,\dots,e_n$ отождествляет эту группу с группой *унитарных матриц*

$$\mathrm{U}_n\stackrel{\text{def}}{=}\{F\in\mathrm{GL}_n(\mathbb{C})\mid F^{-1}=\overline{F}^t\}\,.$$

Подчеркнём, что в отличие от вещественных ортогональных матриц определители унитарных матриц могут принимать не только значения ±1, но любые значения на единичной окружности в $\mathbb{C}$, которая является ни чем иным, как унитарной группой $\mathrm{U}_1=\{z\in\mathbb{C}\mid z\overline{z}=1\}$. Поэтому в эрмитовом пространстве нет понятия *ориентации*: в n° 15.5.1 на стр. 230 ниже мы увидим, что группа U_n представляет собою компактное *линейно связное* подмножество в пространстве комплексных матриц.

Подгруппа $\mathrm{SU}_n=\{F\in\mathrm{U}_n\mid \det F=1\}$ унитарных матриц определителя 1 называется *специальной унитарной* группой.

15.1.8. Эрмитов объём. Выберем какой-нибудь ортонормальный базис $e_1,\dots,e_n$ эрмитова пространства W в качестве базиса единичного объёма и определим *эрмитов объём* n-мерного параллелепипеда, натянутого на векторы $v=e\,C_{ev}$ формулой $\mathrm{Vol}(v_1,\dots,v_n)=|\det C|$. Поскольку модуль определителя матрицы перехода между ортонормальными базисами равен единице, эрмитов объём не зависит от выбора эталонного ортонормального базиса, и квадрат эрмитова объёма, как и в евклидовом случае, равен определителю Грама:

$$\mathrm{Vol}^2(v_1,\dots,v_n)=|\det C_{ev}|^2=\det C_{ev}^t\cdot\overline{\det C}_{ev}=\det G_v\,.$$

15.1.9. Ортогональное проектирование. В эрмитовом пространстве W для любого подпространства $U \subset W$ и любого вектора $w \in W$ имеется единственный вектор $\pi_U(w) \in U$, обладающий следующими эквивалентными друг другу свойствами:

(1) $w - \pi_U(w) \in U^\perp \stackrel{\text{def}}{=} \{v \in W \mid \forall\, u \in U\ (u, v) = 0\}$

(2) $(u, w) = (u, \pi_U(w))$ для всех $u \in U$

(3) вектор $\pi_U(w)$ является ближайшим к w вектором из U в том смысле, что для всех отличных от него векторов $u \in U$ выполняется строгое неравенство $\|w - \pi_U(w)\| < \|w - u\|$.

Свойства (1) и (2) равносильны, поскольку равенства $(u, w) = (u, \pi_U(w))$ и $(u, w - \pi_U(w)) = 0$ равносильны друг другу. Если вектор $\pi_U(w)$ обладает свойствами (1) и (2), то он обладает и свойством (3), так как для любого ненулевого вектора $u \in U$

$$\|w - (\pi_U(w) + u)\|^2 = \|(w - \pi_U(w)) - u\|^2 = \|w - \pi_U(w)\|^2 + \|u\|^2 > \|w - \pi_U(w)\|^2 .$$

С другой стороны, вектор обладающий свойством (3), очевидно, единствен, а обладающий свойствами (1) и (2) вектор $\pi_U(w)$ можно предъявить явно. Для этого выберем в U произвольный базис $u_1, \ldots, u_n$, рассмотрим эрмитово двойственный к нему базис[1] $u_1^\times, \ldots, u_n^\times$ и положим

$$\pi_U(w) = \sum_i (u_i, w)\, u_i^\times . \tag{15-15}$$

Так как равенство из свойства (2) линейно по u, его достаточно проверить только на базисных векторах $u = u_k$, и в этом случае $(u_k, \pi_U(w)) = (u_k, \sum_i (w, u_i)\, u_i^\times) = \sum_i \overline{(w, u_i)}(u_k, u_i^\times) = (u_k, w)$, как и требуется.

Упражнение 15.3. Покажите, что $\pi_U(w) = \sum_i (u_i^\times, w)\, u_i$.

Итак, каждый вектор $w \in W$ допускает единственное разложение $w = \pi_U(w) + \pi_{U^\perp}(w)$, где $\pi_U(w) \in U$, а $\pi_{U^\perp}(w) = w - \pi_U(w) \in U^\perp$. Это означает, в частности, что $W = U \oplus U^\perp$. Подпространство $U^\perp$ называется *ортогональным дополнением* к U, а линейный оператор

$$\pi_U : W \twoheadrightarrow U , \quad w \mapsto \pi_U(w) ,$$

проектирующий W на U вдоль $U^\perp$, называется *ортогональной проекцией* на U. Явно вычислить ортогональную проекцию можно по формуле (15-15) или двойственной формуле из упр. 15.3.

15.2. Сопряжение линейных отображений. Линейные отображения $f : U \to W$ и $f^\times : W \to U$ между эрмитовыми пространствами U и W называются *сопряжёнными*, если для всех $u \in U$ и $w \in W$ выполняется равенство $(fu, w) = (u, f^\times w)$. Это эквивалентно требованию, чтобы для произвольно выбранных базисов $\boldsymbol{u} = (u_1, \ldots, u_n)$ и $\boldsymbol{w} = (w_1, \ldots, w_m)$ при всех i, j выполнялись равенства $(fu_i, w_j) = (u_i, f^\times w_j)$, что равносильно соотношению

$$F_{\boldsymbol{wu}}^t G_{\boldsymbol{w}} = G_{\boldsymbol{u}} \overline{F}_{\boldsymbol{uw}}^\times$$

на матрицы $F_{\boldsymbol{wu}}$ и $F_{\boldsymbol{uw}}^\times$ операторов f и $f^\times$ в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ и матрицы Грама $G_{\boldsymbol{u}}$ и $G_{\boldsymbol{w}}$ этих базисов.

Упражнение 15.4. Убедитесь в этом.

[1] См. n° 15.1.4 на стр. 219.

Таким образом, матрица сопряжённого оператора выражается через матрицу исходного оператора и матрицы Грама по формуле

$$F^{\times}_{\boldsymbol{uw}} = \overline{G^{-1}_{\boldsymbol{u}} F^{t}_{\boldsymbol{wu}} G_{\boldsymbol{w}}}\,. \tag{15-16}$$

В ортонормальных базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ эта формула упрощается до $F^{\times}_{\boldsymbol{uw}} = \overline{F}^{t}_{\boldsymbol{wu}}$. Мы заключаем, что у каждого оператора имеется ровно один сопряжённый, и сопряжение операторов

$$\operatorname{Hom}(U,W) \xrightarrow{\sim} \operatorname{Hom}(W,U)\,, \quad f \mapsto f^{\times}\,,$$

является вещественно линейным комплексно полулинейным изоморфизмом комплексных векторных пространств, т. е. $(\lambda f + \mu g)^{\times} = \overline{\lambda} f^{\times} + \overline{\mu} g^{\times}$ для всех $f, g \in \operatorname{Hom}(U,W)$ и $\lambda, \mu \in \mathbb{C}$.

Предложение 15.2
Для любого линейного отображения $f : U \to W$ выполняются равенства

$$f^{\times\times} = f\,, \quad \ker f^{\times} = (\operatorname{im} f)^{\perp}\,, \quad \operatorname{im} f^{\times} = (\ker f)^{\perp}\,,$$

а для пары линейных отображений $f : U \to V$, $g : V \to W$ — равенство $(gf)^{\times} = f^{\times} g^{\times}$.

Доказательство. Первое равенство $f^{\times\times} = f$ проверяется выкладкой

$$(f^{\times} w, u) = \overline{(u, f^{\times} w)} = \overline{(fu, w)} = (w, fu)\,,$$

а последнее равенство $(gf)^{\times} = f^{\times} g^{\times}$ — выкладкой $(gfu, w) = (fu, g^{\times} w) = (u, f^{\times} g^{\times} w)$. Вектор $w \in \ker f^{\times}$ если и только если для всех $u \in U$ выполняется равенство $0 = (u, f^{\times} w) = (fu, w)$, означающее, что $w \in \operatorname{im} f^{\perp}$. Тем самым, $\ker f^{\times} = (\operatorname{im} f)^{\perp}$. Написав это равенство для оператора $f^{\times}$ в роли f и взяв ортогоналы к обеим частям, получаем $(\ker f)^{\perp} = \operatorname{im} f^{\times}$. □

15.2.1. Сопряжение эндоморфизмов. Если пространство U совпадает с W, сопряжение операторов задаёт вещественно линейный комплексно полулинейный инволютивный[1] антиавтоморфизм[2] алгебры $\operatorname{End}(W)$ комплексно линейных операторов $W \to W$. Согласно прим. 10.3 на стр. 141 пространство $\operatorname{End}(W)$ распадается над полем вещественных чисел в прямую сумму собственных подпространств инволюции $f \mapsto f^{\times}$ с собственными числами ± 1. Они обозначаются

$$\operatorname{End}_{+}(W) = \{f : W \to W \mid f^{\times} = f\} \tag{15-17}$$

$$\operatorname{End}_{-}(W) = \{f : W \to W \mid f^{\times} = -f\} \tag{15-18}$$

и называются пространствами *самосопряжённых* (или *эрмитовых*) и *антисамосопряжённых* (или *косоэрмитовых*) операторов соответственно. Таким образом, для эрмитова оператора f при всех $u, w \in W$ выполняется равенство $(fu, w) = (u, fw)$, а для косоэрмитова — равенство $(fu, w) = -(u, fw)$. В ортонормированном базисе самосопряжённые операторы задаются *эрмитово симметричными* матрицами $F^{t} = \overline{F}$, а антисамосопряжённые — *эрмитово кососимметричными* матрицами $F^{t} = -\overline{F}$.

Эрмитова и антиэрмитова компоненты $f_{+} \in \operatorname{End}_{+}(W)$ и $f_{-} \in \operatorname{End}_{-}(W)$ произвольного оператора $f : W \to W$ в разложении

$$\operatorname{End}(W) = \operatorname{End}_{+}(W) \oplus \operatorname{End}_{-}(W) \tag{15-19}$$

[1] Т. е. обратный самому себе.
[2] Т. е. оборачивающий порядок сомножителей в произведениях.

находятся по формулам $f_+ = (f + f^\times)/2$ и $f_- = (f - f^\times)/2$.

Подчеркнём, что разложение (15-19) определено над полем $\mathbb{R}$, и его компоненты $\mathrm{End}_\pm(W)$ являются *вещественными*, но *не комплексными* векторными подпространствами комплексного векторного пространства $\mathrm{End}(W)$. Умножение на комплексное число i биективно переводит компоненты $\mathrm{End}_\pm(W)$ друг в друга, устанавливая между ними $\mathbb{R}$-линейный изоморфизм.

Упражнение 15.5. Убедитесь, что оператор f эрмитов если и только если оператор if косоэрмитов.

Пример 15.3 (унитарные операторы)
Так как каждый унитарный оператор[1] биективен, всякий вектор $w \in W$ можно записать в виде $f^{-1}v$ для некоторого $v \in W$. Поэтому выполнение для всех u, w равенства $(fu, fw) = (u, w)$ равносильно выполнению для всех u и $v = fw$ равенства $(fu, v) = (u, f^{-1}v)$. Таким образом, унитарную группу пространства W можно охарактеризовать как множество обратимых операторов, сопряжённых своим обратным:

$$\mathrm{U}(W) = \left\{ f \in \mathrm{End}(W) \mid \forall\, u \in W \ \|fu\| = \|u\| \right\} = \left\{ f \in \mathrm{GL}(W) \mid f^\times = f^{-1} \right\}.$$

15.2.2. Сопряжение операторов в евклидовом пространстве. На алгебре $\mathbb{R}$-линейных эндоморфизмов $\mathrm{End}(V)$ вещественного евклидова пространства V требование

$$\forall\, u, w \in V \quad (fu, w) = (u, f^\times w)$$

также корректно определяет операцию сопряжения $f \leftrightarrow f^\times$. Эта операция является инволютивным антиавтоморфизмом $\mathbb{R}$-алгебры $\mathrm{End}(V)$. Матрица $F^\times$ сопряжённого оператора в произвольном базисе связана с матрицей Грама G этого базиса и матрицей F исходного оператора по формуле $F^\times = G^{-1}F^tG$, которая в ортонормальном базисе упрощается до $F^\times = F^t$. Пространство $\mathbb{R}$-линейных эндоморфизмов евклидова пространства V также раскладывается в прямую сумму $\mathrm{End}(V) = \mathrm{End}_+(V) \oplus \mathrm{End}_-(V)$ подпространств (анти) самосопряжённых операторов

$$\mathrm{End}_\pm(V) = \{ f \mid f^\times = \pm f \}.$$

В ортонормальном базисе пространства V (анти) самосопряжённые операторы имеют в (косо) симметричные матрицы. При этом ортогональные[2] операторы на евклидовом пространстве характеризуются как операторы, сопряжённые к своим обратным.

Пример 15.4 (сопряжение дифференциальных операторов)
Обозначим через V пространство бесконечно дифференцируемых функций $f : [a, b] \to \mathbb{R}$, которые обращаются на концах отрезка в нуль вместе со всеми своими производными, и введём на V евклидово скалярное произведение

$$(f, g) = \int_a^b f(t)g(t)\,dt\,.$$

[1] См. n° 15.1.7 на стр. 221.

[2] Т. е. сохраняющие евклидову длину векторов или, что равносильно, евклидово скалярное произведение.

Интегрирование по частям показывает, что дифференцирование $\frac{d}{dt} : f \mapsto f'$ является антисамосопряжённым линейным оператором:

$$\left(\frac{d}{dt} f\,,\, g\right) = \int_a^b f' g\, dt = -\int_a^b f g'\, dt = \left(f\,,\, -\frac{d}{dt} g\right).$$

Умножение на любую заданную функцию $g : f \mapsto gf$ является самосопряжённым оператором. Поскольку сопряжение является антигомоморфизмом по отношению к композиции, оператор, сопряжённый к линейному дифференциальному оператору вида

$$t^3 \frac{d^2}{dt^2} : f(t) \mapsto t^3 f''(t),$$

переводит функцию f в функцию $\left(t^3 f\right)'' = 6tf + 6t^2 f' + t^3 f''$, т. е. имеет вид

$$\left[t^3 \frac{d^2}{dt^2}\right]^* = t^3 \frac{d^2}{dt^2} + 6t^2 \frac{d}{dt} + 6t\,.$$

Упражнение 15.6. Вычислите оператор, сопряжённый к оператору

$$L = a(t)\frac{d^2}{dt^2} + b(t)\frac{d}{dt} + c(t) : f \mapsto af'' + bf + c\,,$$

где $a, b, c \in V$.

15.3. Ортогональная диагонализация нормальных операторов. Оператор $f : W \to W$ на эрмитовом пространстве W называется *нормальным*, если он перестановочен со своим сопряжённым, т. е. $f^\times f = f f^\times$. Например, все эрмитовы, косоэрмитовы и унитарные операторы нормальны, так как сопряжённый к такому оператору f оператор равен f, $-f$ и f^{-1} соответственно.

Теорема 15.1
Оператор f на конечномерном эрмитовом пространстве W нормален если и только если он диагонализуем в ортонормальном базисе пространства W. При этом диагональная матрица для f с точностью до перестановки диагональных элементов не зависит от выбора такого базиса.

Доказательство. Если f диагонализуем в ортонормальном базисе, то сопряжённый к f оператор имеет в этом базисе сопряжённую диагональную матрицу, которая коммутирует с диагональной матрицей оператора f. Поэтому f нормален. Так как диагональные элементы любой диагональной матрицы, задающей оператор f, представляют собою собственные числа оператора f, и каждое из них присутствует на диагонали столько раз, какова размерность отвечающего ему собственного подпространства, диагональные элементы с точностью до перестановки не зависят от выбора базиса, в котором матрица диагональна.

Диагонализуемость нормального оператора $f : W \to W$ в ортонормальном базисе доказывается индукцией по $\dim W$. Если оператор f скалярен (что так при $\dim W = 1$), то он диагонален в любом базисе. Если $\dim W > 1$ и оператор f не скалярен, то у него есть ненулевое собственное подпространство $V_\lambda \subsetneq W$, и $W = V_\lambda \oplus V_\lambda^\perp$. Поскольку оператор $f^\times$ перестановочен с f, он переводит собственное подпространство V_λ в себя[1]. Поэтому для всех $u \in V_\lambda$ и любого

[1] См. n° 10.2.7 на стр. 141.

$w \in V_\lambda^\perp$ выполняется равенство $(fw, u) = (w, f^\times u) = 0$, означающее, что $fw \in V_\lambda^\perp$. Таким образом, оператор f переводит подпространство $V_\lambda^\perp$ в себя. По индукции, ограничение f на $V_\lambda^\perp$ диагонализуемо в некотором ортонормальном базисе пространства $V_\lambda^\perp$. Добавляя к этому базису любой ортонормальный базис собственного подпространства V_λ, получаем ортонормальный базис пространства W, в котором матрица оператора f диагональна. □

Следствие 15.1
Оператор самосопряжён если и только если он диагонализуем в ортонормальном базисе и все его собственные числа вещественны.

Следствие 15.2
Оператор антисамосопряжён если и только если он диагонализуем в ортонормальном базисе и все его собственные числа чисто мнимы.

Следствие 15.3
Оператор унитарен если и только если он диагонализуем в ортонормальном базисе и все его собственные числа лежат на единичной окружности в $\mathbb{C}$.

Упражнение 15.7. Покажите, что унитарная группа U_n является компактным линейно связным подмножеством пространства $\mathrm{Mat}_n(\mathbb{C})$.

15.4. Сингулярные числа и сингулярные направления. В этом разделе мы покажем, что каждое линейное отображение $f : U \to W$ между эрмитовыми пространствами U и W однозначно раскладывается в композицию $f = gh\pi$, где $\pi : U \twoheadrightarrow V$ — ортогональная проекция на ортогональное дополнение $V \stackrel{\text{def}}{=} (\ker F)^\perp$ к ядру оператора F, оператор $h : V \to V$ самосопряжён и имеет положительные собственные числа[1], а $g : V \hookrightarrow W$ — унитарное вложение, сохраняющее скалярное произведение. Попарно перпендикулярные собственные векторы, вдоль которых растягивает подпространство $V \subset U$ самосопряжённый оператор $h : V \to V$, и положительные вещественные коэффициенты этих растяжений называются *сингулярными направлениями* и *сингулярными числами* линейного отображения f.

Лемма 15.2
Для любого линейного отображения $f : U \to W$ между эрмитовыми пространствами U, W обе композиции $ff^\times \in \mathrm{End}(W)$, $f^\times f \in \mathrm{End}(U)$ являются самосопряжёнными линейными операторами с неотрицательными собственными числами. Отображение f сюръективно (соотв. инъективно) если и только если все собственные числа оператора $ff^\times$ (соотв. $f^\times f$) строго положительны.

Доказательство. Каждый из операторов $ff^\times$ и $f^\times f$ очевидно самосопряжён и следовательно диагонализуем по сл. 15.1 на стр. 226. Если для некоторого ненулевого вектора $w \in W$ выполняется равенство $ff^\times w = \lambda w$, то $(f^\times w, f^\times w) = (ff^\times w, w) = \lambda \cdot (w, w)$ и либо $w \in \ker f^\times$ и $\lambda = 0$, либо $\lambda = (f^\times w, f^\times w)/(w, w) > 0$. Аналогично, если $f^\times f u = \mu u$ для ненулевого $u \in U$, то либо $\mu = 0$ и $u \in \ker f$, либо $\mu = (fu, fu)/(u, u) > 0$. Поэтому все ненулевые собственные числа каждого из операторов положительны. Если $\operatorname{im} f = W$, то[2] $\ker f^\times = (\operatorname{im} f)^\perp = 0$, откуда все собственные числа оператора $ff^\times$ положительны. Наоборот, если $\operatorname{im} f \neq W$, то $\ker ff^\times \supset$

[1] Т. е. по сл. 15.1 является растяжением с вещественными положительными коэффициентами во взаимно перпендикулярных комплексных направлениях.

[2] См. предл. 15.2 на стр. 223.

$\ker f^\times = (\operatorname{im} f)^\perp \neq 0$. Аналогично, если $\ker f = 0$, то все собственные числа оператора $f^\times f$ строго положительны, и наоборот, если $\ker f \neq 0$, то и $\ker f^\times f \supset \ker f \neq 0$. □

Теорема 15.2
Каждое линейное отображение $f : U \to W$ между эрмитовыми пространствами U, W единственным образом раскладывается в композицию $f = g_f \circ h_f \circ \pi_f$ ортогональной проекции $\pi_f : U \twoheadrightarrow V$ на ортогональное дополнение $V \stackrel{\text{def}}{=} \ker^\perp f$ к ядру $\ker f \subset U$, невырожденного самосопряжённого оператора $h_f : V \xrightarrow{\sim} V$ с положительными собственными числами $\alpha_1, \dots \alpha_r$, где $r = \operatorname{rk} f = \dim \operatorname{im} f$, и унитарного[1] вложения $g_f : V \hookrightarrow W$. При этом набор $\alpha_1^2, \dots \alpha_r^2$ квадратов собственных чисел оператора h_f является набором всех (с учётом кратностей) ненулевых собственных чисел оператора $f^\times f : U \to U$.

Доказательство. Согласно сл. 15.1 на стр. 226 в эрмитовом пространстве U имеется ортонормальный базис, состоящий из собственных векторов $u_1, \dots, u_n$ самосопряжённого линейного оператора $f^\times f : U \to U$, причём по лем. 15.2 все собственные значения этого оператора неотрицательны, т. е. $f^\times f u_i = \alpha_i^2 u_i$ для некоторых вещественных $\alpha_i \geqslant 0$. Перенумеруем базис так, чтобы $\alpha_i \neq 0$ при $1 \leqslant i \leqslant r$ и $\alpha_i = 0$ при $i > r$. Тогда, как мы видели в доказательстве лем. 15.2, все векторы u_i с $i > r$ лежат в ядре отображения f. Напротив, при $1 \leqslant i, j \leqslant r$ равенства

$$(fu_i, fu_j) = (f^\times f u_i, u_j) = \alpha_i^2 (u_i, u_j) = \begin{cases} \alpha_i^2 > 0 & \text{при } i = j \\ 0 & \text{при } i \neq j \end{cases}$$

показывают, что векторы $w_i = fu_i / \alpha_i$ образуют в пространстве W ортонормальную систему. В частности, они линейно независимы. Так как $f(u_j) = 0$ при $j > r$, для любого $u = \sum x_i u_i \in U$ выполняется равенство $f(u) = \alpha_1 x_1 w_1 + \dots + \alpha_r x_r w_r$, т. е. векторы w_i с $1 \leqslant i \leqslant r$ составляют ортонормальный базис в $\operatorname{im} f$, а векторы u_i с $1 \leqslant i \leqslant r$ — ортонормальный базис в ортогональном дополнении V к ядру $\ker f$. Оператор f является композицией изометрического изоморфизма $g_f : V \xrightarrow{\sim} \operatorname{im} f$, $u_i \mapsto w_i$, диагонального оператора $h_f : V \to V$, $u_i \mapsto \alpha_i u_i$, и ортогональной проекции $\pi_f : U \twoheadrightarrow V$ вдоль $\ker f$.

Пусть имеется какое-либо ещё разложение $f = ghp_f$, где $\pi_f : U \twoheadrightarrow V$ — ортогональная проекция вдоль $\ker f$. Из предыдущего рассуждения вытекает, что пространство $V = (\ker f)^\perp$ является прямой ортогональной суммой всех собственных подпространств V_i оператора $f^\times f$, отвечающих ненулевым собственным значениям α_i^2 этого оператора, и композиция $gh : V \xrightarrow{\sim} \operatorname{im} f$ совпадает с ограничением $f|_V$. Поскольку $h^\times = h$ как операторы $V \to V$, а $g^\times = g^{-1}$ как унитарные операторы $\operatorname{im} f \xrightarrow{\sim} V$, мы заключаем, что ограничение $f^\times f|_V = h^2$. Так как оператор h^2 диагонализуется в том же самом базисе, что и h, мы заключаем, что самосопряжённый оператор h действует на каждом подпространстве V_i умножением на α_i. Тем самым, h определяется по f однозначно. А тогда и $g = h^{-1} \circ f|_V : V \to W$ определяется однозначно. □

Упражнение 15.8. Убедитесь, что оператор $f^\times : W \to V$ действует на построенные в доказательстве теор. 15.2 векторы $w_1, \dots, w_r \in W$ по правилу $w_i \mapsto \alpha_i u_i$ и аннулирует ортогональное дополнение к их линейной оболочке. Выведите отсюда, что множества всех (с учётом кратностей) ненулевых собственных чисел у операторов $f^\times f$ и $ff^\times$ одинаковы.

[1] Т. е. сохраняющего скалярное произведение: $(g_f u_1, g_f u_2) = (u_1, u_2)$ для всех $u_1, u_2 \in U$.

Определение 15.1 (сингулярные числа и сингулярные направления)
В условиях теор. 15.2 набор из $\dim U$ неотрицательных квадратных корней α_i из собственных значений самосопряжённого оператора $f^\times f : U \to U$ называется *набором сингулярных чисел* линейного отображения $f : U \to W$ между эрмитовыми пространствами U, W. Ровно $\operatorname{rk} f$ из них строго положительны. Одномерные инвариантные подпространства[1] оператора $f^\times f$ называются *сингулярными направлениями* отображения f.

Следствие 15.4 (SVD-разложение[2])
Каждая комплексная прямоугольная матрица $F \in \mathrm{Mat}_{m\times n}(\mathbb{C})$ раскладывается в произведение $F = T_m D T_n$, в котором матрицы $T_m \in \mathrm{U}_m$ и $T_n \in \mathrm{U}_n$ унитарны, а $m \times n$-матрица $D = \bigl(d_{ij}\bigr)$ диагональна, вещественна и неотрицательна в том смысле, что $d_{ij} = 0$ при $i \neq j$, а все $d_{ii} \in \mathbb{R}_{\geqslant 0}$. При этом ровно $\operatorname{rk} F$ диагональных элементов матрицы D отлично от нуля, и они с точностью до перестановки диагональных элементов не зависят от выбора указанного разложения.

Доказательство. Будем воспринимать $F = F_{\boldsymbol{mn}}$ как записанную в стандартных базисах $\boldsymbol{n}$ и $\boldsymbol{m}$ эрмитовых пространств $U = \mathbb{C}^n$ и $W = \mathbb{C}^m$ матрицу линейного оператора $F : \mathbb{C}^n \to \mathbb{C}^m$. Обозначим через $\boldsymbol{u} = (u_1, \dots, u_n)$ ортонормальный базис пространства U, построенный в доказательстве теор. 15.2, а через $\boldsymbol{w} = (w_1, \dots, w_m)$ — любой ортонормальный базис пространства W, содержащий ортонормальный набор векторов $w_i = F(u_i)\,/\,\alpha_i$, $1 \leqslant i \leqslant r$, из доказательства теор. 15.2. Оператор $F : u_i \mapsto \alpha_i w_i$ задаётся в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ диагональной матрицей $D = F_{\boldsymbol{wu}}$, ненулевые диагональные элементы которой суть сингулярные числа $\alpha_1, \dots, \alpha_n$ оператора F. Поэтому $F = F_{\boldsymbol{mn}} = C_{\boldsymbol{mw}} F_{\boldsymbol{wu}} C_{\boldsymbol{un}}$, где $C_{\boldsymbol{mw}}$ — унитарная матрица перехода от базиса $\boldsymbol{w}$ к стандартному базису $\boldsymbol{m}$ в $\mathbb{C}^m$, а $C_{\boldsymbol{un}} = C_{\boldsymbol{nu}}^{-1} = C_{\boldsymbol{nu}}^t$ — унитарная матрица перехода от стандартного базиса $\boldsymbol{n}$ в $\mathbb{C}^n$ к базису $\boldsymbol{u}$. Для любого другого разложения $F = T_m A T_n$ с унитарными T_n, T_m и диагональной матрицей A имеем $F^t F = T_n^{-1} A^t A T_n$. Поскольку собственные числа подобных матриц одинаковы, стоящие на диагонали диагональной матрицы $A^t A$ квадраты диагональных элементов матрицы A суть собственные числа матрицы $F^t F$. □

15.5. Полярное разложение. Каждое комплексное число $z \in \mathbb{C}^* = \mathrm{GL}_1(\mathbb{C})$ имеет вид

$$z = \varrho\, e^{i\vartheta}, \tag{15-20}$$

где $\varrho = |z|$ вещественно и положительно, а $e^{i\vartheta} = \cos\vartheta + i\sin\vartheta = \operatorname{Arg} z \in \mathrm{U}_1$. Если воспринимать z как оператор умножения $w \mapsto zw$ на одномерном эрмитовом координатном пространстве $\mathbb{C}$, то формула (15-20) даёт разложение такого оператора в композицию самосопряжённого оператора $w \mapsto \varrho w$ с положительным собственным числом $\varrho = \sqrt{\bar z z} = \sqrt{z^\times z}$ и унитарного оператора $w \mapsto e^{i\vartheta} w$ с собственным числом $e^{i\vartheta} = z\varrho^{-1}$. Непосредственным обобщением этого на старшие размерности является

Следствие 15.5 (полярное разложение)
Каждое биективное линейное преобразование $f \in \mathrm{GL}(W)$ эрмитова пространства W допускает единственное разложение $f = g_f h_f$, в котором оператор $g_f \in \mathrm{U}(W)$ унитарен, а $h_f \in \mathrm{GL}(W)$ самосопряжён и имеет положительные собственные числа, квадраты которых являются собственными числами оператора $f^\times f$.

[1] Т. е. одномерные подпространства, порождённые ненулевыми собственными векторами.
[2] «SVD» является аббревиатурой от английского *singular values decomposition*.

Доказательство. Поскольку оператор f биективен, проекция π_f в его каноническом разложении $f = g_f \circ h_f \circ \pi_f$ из теор. 15.2 является тождественным отображением, а самосопряжённый оператор h_f не имеет ядра. Следовательно все собственные числа оператора h_f строго положительны. □

Замечание 15.2. (явные формулы для g_f и h_f) Компоненты $g_f \in \mathrm{U}(W)$ и h_f полярного разложения $f = g_f \circ h_f$ однозначно находятся из условий $g_f^\times g_f = \mathrm{Id}_W$ и $h_f^\times = h_f$. А именно,

$$f^\times f = h_f^\times g_f^\times g_f h_f = h_f^2\,,$$

откуда $h_f = \sqrt{f^\times f}$ и $g_f = f h_f^{-1}$. Так как $0 \notin \mathrm{Spec}(f^\times f)$, аналитическая вне нуля функция $\sqrt{t}$ алгебраически вычислима на операторе $f^\times f$ при помощи стандартной интерполяционной процедуры[1] из n° 10.3.1 на стр. 143.

Упражнение 15.9. Покажите, что каждый невырожденный линейный оператор f на эрмитовом пространства W также допускает единственное разложение $f = hr$, в котором оператор $r \in \mathrm{U}(W)$, а оператор r самосопряжён и имеет положительные собственные значения, квадраты которых равны собственным числам оператора $ff^\times$.

Пример 15.5
Найдём полярное разложение $f = gh$ для оператора $f : \mathbb{C}^3 \to \mathbb{C}^3$ с матрицей

$$F = \begin{pmatrix} 22/15 & -4/3 & 4/15 \\ 4/15 & 2/3 & 28/15 \\ 2/3 & 2/3 & -1/3 \end{pmatrix}.$$

Так как $\det F = -4$, оператор f невырожден. Самосопряжённый оператор $f^\times f$ имеет матрицу

$$C = F^t F = \begin{pmatrix} 22/15 & 4/15 & 2/3 \\ -4/3 & 2/3 & 2/3 \\ 4/15 & 28/15 & -1/3 \end{pmatrix}\begin{pmatrix} 22/15 & -4/3 & 4/15 \\ 4/15 & 2/3 & 28/15 \\ 2/3 & 2/3 & -1/3 \end{pmatrix} = \begin{pmatrix} 8/3 & -4/3 & 2/3 \\ -4/3 & 8/3 & 2/3 \\ 2/3 & 2/3 & 11/3 \end{pmatrix}$$

у которой след $\mathrm{tr}(C) = 9$, сумма главных 2×2-миноров

$$\det\begin{pmatrix} 8/3 & -4/3 \\ -4/3 & 8/3 \end{pmatrix} = 16/3\,, \quad \det\begin{pmatrix} 8/3 & 2/3 \\ 2/3 & 11/3 \end{pmatrix} = 28/3\,, \quad \det\begin{pmatrix} 8/3 & 2/3 \\ 2/3 & 11/3 \end{pmatrix} = 28/3$$

равна 24, определитель $\det(C) = \det^2 F = 16$ и характеристический многочлен

$$\det(tE - C) = t^3 - 9\,t^2 + 24\,t - 16 = (t-1)(t-4)^2\,.$$

Так как оператор $f^\times f$ диагонализуем, он аннулируется многочленом[2] $(t-1)(t-4)$. Следовательно, матрица $H = \sqrt{C}$ самосопряжённого сомножителя h полярного разложения $f = gh$ имеет вид[3] $aE + bC$, где интерполяционный многочлен $p(t) = a + bt$ для вычисления функции $\sqrt{t}$ на

[1] См. опр. 10.3 на стр. 145.
[2] См. предл. 10.6 на стр. 140.
[3] См. n° 10.3.1 на стр. 143.

матрице C однозначно определяется тем, что $p(1)=\sqrt{1}=1$ и $p(4)=\sqrt{4}=2$, т. е. $a+b=1$ и $a+4b=2$, откуда $a=2/3$, $b=1/3$. Таким образом, самосопряжённая матрица $H=\sqrt{C}$ равна

$$\begin{pmatrix}2/3&0&0\\0&2/3&0\\0&0&2/3\end{pmatrix}+\begin{pmatrix}8/9&-4/9&2/9\\-4/9&8/9&2/9\\2/9&2/9&11/9\end{pmatrix}=\begin{pmatrix}14/9&-4/9&2/9\\-4/9&14/9&2/9\\2/9&2/9&17/9\end{pmatrix}$$

а унитарная матрица $G=FH^{-1}$ равна

$$\begin{pmatrix}22/15&-4/3&4/15\\4/15&2/3&28/15\\2/3&2/3&-1/3\end{pmatrix}\begin{pmatrix}13/18&2/9&-1/9\\2/9&13/18&-1/9\\-1/9&-1/9&5/9\end{pmatrix}=\begin{pmatrix}11/15&-2/3&2/15\\2/15&1/3&14/15\\2/3&2/3&-1/3\end{pmatrix}$$

Упражнение 15.10. Убедитесь, что $G^tG=E$.

15.5.1. Экспоненциальное накрытие унитарной группы. Алгебра $\mathcal{A}\subset\mathbb{C}[\![z]\!]$, состоящая из абсолютно сходящихся всюду в $\mathbb{C}$ степенных рядов, алгебраически вычислима[1] на любом линейном операторе $F\colon\mathbb{C}^n\to\mathbb{C}^n$. В частности, у любого оператора F имеется экспонента $e^F\colon\mathbb{C}^n\to\mathbb{C}^n$. Если оператор F антисамосопряжён относительно стандартной эрмитовой структуры на $\mathbb{C}^n$, набор элементарных делителей $\mathcal{El}(F)$ состоит из n двучленов $t-ia$, где $a\in\mathbb{R}$, $ia\in\operatorname{Spec}F$. По предл. 10.9 на стр. 148 $\mathcal{El}(e^F)$ состоит из n двучленов $t-e^{ia}$ биективно соответствующих (с учётом кратностей) собственным числам ia оператора F. Поэтому экспонента $e^F\colon\mathbb{C}^n\to\mathbb{C}^n$ является унитарным оператором с собственными числами $e^{ia}=\cos a+i\sin a$ находящимися в биекции (с учётом кратностей) с собственными числами ia оператора F. Если разложить $\mathbb{C}^n$ в прямую ортогональную сумму одномерных F-инвариантных подпространств, то каждое из них будет и e^F-инвариантно, и каждый собственный вектор оператора F с собственным значением ia будет собственным вектором оператора e^F с собственным значением e^{ia}. Поскольку любой унитарный оператор является прямой ортогональной суммой унитарных операторов, действующих на одномерных подпространствах и имеющих собственные числа вида e^{ia} с $a\in\mathbb{R}$, мы заключаем, что *экспоненциальное отображение*

$$\operatorname{End}_-(\mathbb{C}^n)\twoheadrightarrow \mathrm{U}_n\,,\quad F\mapsto e^F\,, \tag{15-21}$$

сюръективно. Иначе говоря, каждый унитарный оператор имеет вид $G=e^{iT}$ для некоторого самосопряжённого оператора T. В частности, полярное разложение оператора $F\in\mathrm{GL}_n(\mathbb{C})$ можно переписать в виде $F=e^{iT}H$, где $T,H\in\operatorname{End}_+(\mathbb{C}^n)$ и все собственные числа оператора H положительны. Однако в отличие от унитарного оператора G в представлении $F=GH$ из сл. 15.5 самосопряжённый оператор T определяется оператором F (или, что то же самое, оператором G) уже не однозначно, поскольку экспоненциальное отображение не инъективно.

Упражнение 15.11. Убедитесь, что $e^{2\pi i\operatorname{Id}}=\operatorname{Id}$.

Предостережение 15.1. Экспоненциальное отображение (15-21) не является гомоморфизмом аддитивной группы в мультипликативную, поскольку $e^{A+B}\neq e^Ae^B$ если матрицы A и B не перестановочны. Композиция e^Ae^B является экспонентой от бесконечного *ряда Кэмпбела–Хаусдорфа*, составленного из итерированных коммутаторов операторов A и B. Прочитать об этом можно в книге *Серр Ж. П. Алгебры Ли и группы Ли*. М. «Мир» 1969, гл. IV, §§ 7, 8.

[1] См. n° 10.3.1 на стр. 143.

Ответы и указания к некоторым упражнениям

Упр. 1.1. Ответ: 2^n.

Упр. 1.2. Ответ на второй вопрос — нет. Пусть $X = \{1, 2\}$, $Y = \{2\}$. Все их парные пересечения и объединения суть $X \cap Y = Y \cap Y = Y \cup Y = Y$ и $X \cup Y = X \cup X = X \cap X = X$, и любая формула, составленная из X, Y, $\cap$, $\cup$, даст на выходе или $X = \{1, 2\}$, или $Y = \{2\}$, тогда как $X \setminus Y = \{1\}$.

Упр. 1.3. В первом случае имеется 6 наложений и ни одного вложения, во втором — 6 вложений и ни одного наложения.

Упр. 1.5. Если X конечно, то инъективное или сюръективное отображение $X \to X$ автоматически биективно. Если X бесконечно, то в X есть подмножество, изоморфное $\mathbb{N}$. Инъекция $\mathbb{N} \hookrightarrow \mathbb{N}$, $n \mapsto (n + 1)$, и сюръекция $\mathbb{N} \twoheadrightarrow \mathbb{N}$, $n \mapsto \max(1, (n - 1))$, обе не биективны и продолжаются до точно таких же отображений $X \to X$ тождественным действием на $X \setminus \mathbb{N}$.

Упр. 1.6. Ответ: нет. Воспользуйтесь «диагональным трюком» Кантора: пусть все биекции $\mathbb{N} \to \mathbb{N}$ занумерованы натуральными числами; глядя на этот список, постройте биекцию, которая при каждом $k = 1, 2, 3, \ldots$ отображает некоторое число $n_k \in \mathbb{N}$ не туда, куда его отображает k-тая биекция из списка.

Упр. 1.7. Обозначим через $\sigma : X \xrightarrow{\sim} X$ биекцию, переставляющую между собою точки x и x' и тождественно действующую на остальные точки. Искомое отображение переводит биекцию $f : X \xrightarrow{\sim} X$ в композицию $\sigma \circ f : z \mapsto \sigma\big(f(z)\big)$.

Упр. 1.8. Ответ: $\binom{n+m-1}{m-1} = \binom{n+m-1}{n} = \frac{(n+m-1)!}{n!(m-1)!}$. Указание: слагаемых столько же, сколько имеется упорядоченных наборов неотрицательных целых чисел $(k_1, \ldots, k_m)$ с суммой $\sum k_i = n$. Такой набор можно закодировать словом, составленным из $(m - 1)$ букв 0 и n букв 1: сначала пишем k_1 единиц, потом нуль, потом k_2 единиц, потом нуль, и т. д. (слово кончится k_m единицами, стоящими следом за последним, $(m - 1)$-м нулём) .

Упр. 1.9. Ответ: $\binom{n+k}{k}$. Каждая такая диаграмма представляет собою ломаную, ведущую из левого нижнего угла прямоугольника в правый верхний. В такой ломаной ровно n горизонтальных звеньев и ровно k вертикальных.

Упр. 1.10. Пусть $[x']_n = [x]_n$ и $[y']_n = [y]_n$, т. е. $x' = x + nk$, $y' = y + n\ell$ с некоторыми $k, \ell \in \mathbb{Z}$. Тогда $x' + y' = x + y + n(k + \ell)$ и $x'y' = xy + n(\ell x + ky + k\ell n)$ сравнимы по модулю n с $x + y$ и xy соответственно, т. е. $[x' + y']_n = [x + y]_n$ и $[x'y']_n = [xy]_n$.

Упр. 1.11. Положим $x \sim y$, если существует конечная последовательность точек

$$x = z_0,\ z_1,\ z_2,\ \ldots\ ,\ z_n = y$$

как в условии задачи. Проверьте, что это отношение эквивалентности и что оно содержится в любой эквивалентности $S \subset X \times X$, содержащей R.

Упр. 1.12. Рефлексивность и симметричность очевидны. Транзитивность: если $(p, q) \sim (r, s)$ и $(r, s) \sim (u, w)$, т. е. $ps - rq = 0 = us - rw$, то $psw - rqw = 0 = usq - rwq$, откуда $s(pw - uq) = 0$, и $pw = uq$, т. е. $(p, q) \sim (u, w)$.

Упр. 1.13. Если прямые ℓ_1 и ℓ_2 пересекаются в точке O под углом $0 < \alpha \leqslant \pi / 2$, то отражение относительно ℓ_1, за которым следует отражение относительно ℓ_2, это поворот вокруг точки O на угол 2α в направлении от первой прямой ко второй. Таким образом, отражения относительно пересекающихся прямых коммутируют тогда и только тогда, когда прямые перпендикулярны.

Упр. 1.15. Таблица композиций gf в симметрической группе S_3:

$g \diagdown f$	$(1,2,3)$	$(1,3,2)$	$(3,2,1)$	$(2,1,3)$	$(2,3,1)$	$(3,1,2)$
$(1,2,3)$	$(1,2,3)$	$(1,3,2)$	$(3,2,1)$	$(2,1,3)$	$(2,3,1)$	$(3,1,2)$
$(1,3,2)$	$(1,3,2)$	$(1,2,3)$	$(3,1,2)$	$(2,3,1)$	$(2,1,3)$	$(3,2,1)$
$(3,2,1)$	$(3,2,1)$	$(2,3,1)$	$(1,2,3)$	$(3,1,2)$	$(1,3,2)$	$(2,1,3)$
$(2,1,3)$	$(2,1,3)$	$(3,1,2)$	$(2,3,1)$	$(1,2,3)$	$(3,2,1)$	$(1,3,2)$
$(2,3,1)$	$(2,3,1)$	$(3,2,1)$	$(2,1,3)$	$(1,3,2)$	$(3,1,2)$	$(1,2,3)$
$(3,1,2)$	$(3,1,2)$	$(2,1,3)$	$(1,3,2)$	$(3,2,1)$	$(1,2,3)$	$(2,3,1)$

Упр. 1.16. Отношение $n \mid m$ на множестве $\mathbb{Z}$ не кососимметрично: $n \mid m$ и $m \mid n$ если и только если $|m| = |n| \neq 0$. Фактор множества $\mathbb{Z}$ по этому отношению эквивалентности можно отождествить с множеством $\mathbb{Z}_{\geqslant 0}$ неотрицательных целых чисел, на котором отношение $n \mid m$ является частичным порядком (обратите внимание, что нуль является нижней гранью этого множества, т. е. делит все элементы.)

Упр. 1.17. Пусть множество $S \subset W$ состоит из всех таких элементов $z \in W$, что утверждение $\Phi(z)$ ложно. Если $S \neq \varnothing$, то в нём есть начальный элемент $s_* \in S$. Поскольку утверждение $\Phi(w)$ истинно для всех $w < s_*$, утверждение $\Psi(s_*)$ тоже истинно, т. е. $s_* \notin S$. Противоречие.

Упр. 1.18. Обозначим через x_I начальный элемент дополнения $W \smallsetminus I$. Начальный интервал $[x_I) \subset W$ является объединением начальных интервалов $[y) \subset W$ по всем $y < x$. Так как I содержит все интервалы $[y)$ с $y < x_I$, мы заключаем, что $I \supseteq [x_I)$, откуда $I = [x_I)$.

Упр. 1.19. Пусть соотношение $U \geqslant W$ не выполняется. Покажем, что любой начальный отрезок $[u) \subset U$ изоморфен некоторому начальному отрезку $[w) \subset W$, где $w = w(u)$ однозначно восстанавливается по u. Это верно для пустого начального отрезка $\varnothing = [u_*]$, где $u_* \in U$ — минимальный элемент. Пусть это верно для всех начальных отрезков $[y) \subset U$ с $y < u$. Тогда $[y) = \bigcup_{y<u}[y)$ изоморфен объединению вложенных отрезков $\bigcup_{y<u}[w(y)) \subset W$. Если это объединение исчерпывает всё множество W, то $W \simeq [y)$, т. е. $W \leqslant U$ вопреки предположению. Положим $w(u) \in W$ равным минимальному элементу, не содержащемуся в $\bigcup_{y<u}[w(y))$. Проверьте, что $\bigcup_{y<u}[w(y)) = [w(u))$ и что отображение $u \mapsto w(u)$ устанавливает изоморфизм множества U либо со всем множеством W, либо с некоторым его начальным отрезком.

Упр. 1.20. Рассмотрим подмножество $Z \subseteq W_1$, состоящее из всех таких $z \in W_1$, что начальный интервал $[z)_1$ в множестве W_1 является одновременно начальным интервалом $[z)_2$ множества W_2. Множество Z не пусто, поскольку содержит общий начальный элемент множеств W_1 и W_2. Если $Z \subsetneq W_1$ и $Z \subsetneq W_2$, то по упр. 1.18 на стр. 18 подмножество Z является начальным интервалом как в W_1, так и в W_2, что невозможно, поскольку точные верхние границы этих интервалов в W_1 и W_2, с одной стороны, не лежат в Z и, стало быть, различны, а с другой стороны в силу рекурсивности множеств W_1 и W_2 обе они равны $\varrho(Z)$, то есть совпадают. Тем самым, $Z = W_1$ или $Z = W_2$. По упр. 1.18 в первом случае W_1 является начальным интервалом в W_2, а во втором — W_2 является начальным интервалом в W_1.

Упр. 1.21. Каждое подмножество $S \subset U$ имеет непустое пересечение с каким-нибудь рекурсивным вполне упорядоченным подмножеством $W \subset P$ с начальным элементом $\varrho(\varnothing)$. По упр. 1.20 подмножество W является начальным интервалом всех содержащих W рекурсивных вполне упорядоченных подмножеств с начальным элементом $\varrho(\varnothing)$. Поэтому начальный элемент пересечения $S \cap W$ не зависит от выбора W с $W \cap S \neq \varnothing$ и является начальным элементом подмножества S.

Каждый начальный интервал $[u) \subset U$ является начальным интервалом любого содержащего u множества W из цепи. В силу рекурсивности W элемент $\varrho[u) = u$.

Упр. 1.22. Пользуясь аксиомой выбора, зафиксируем для каждого $W \in \mathcal{W}(P)$ какую-нибудь верхнюю грань $b(W) \in P$. Если $f(x) > x$ для всех $x \in P$, то отображение $\beta : \mathcal{W}(P) \to P, W \mapsto f(b(W))$ противоречит лем. 1.2 на стр. 19.

Упр. 1.23. Обозначим через $\mathcal{S}(X)$ множество всех непустых подмножеств данного множества X, включая само X. При помощи аксиомы выбора постройте такое отображение $\mu : \mathcal{S}(X) \to X$, что $\mu(Z) \in Z$ для всех $Z \in \mathcal{S}(X)$. Обозначим через $\mathcal{W}(X)$ множество всех $W \in \mathcal{S}(X)$, которые можно вполне упорядочить так, что $\mu(X \setminus [w)) = w$ для всех $w \in W$. Вдохновляясь лем. 1.2 на стр. 19 покажите, что $\mathcal{W}(X) \neq \varnothing$, и убедитесь, что $X \in \mathcal{W}(X)$.

Упр. 1.24. Убедитесь, что множество P всех цепей, содержащих данную цепь, является полным чумом относительно частичного порядка, задаваемого включением, и примените лемму Цорна.

Упр. 2.2. Ответы: $1 + x$ и $xy + x + y$.

Упр. 2.3. При умножении числителя и знаменателя любой из дробей в левых частях равенств форм. (2-11) на стр. 21 на одно и то же число c, числитель и знаменатель дроби в правой части соответствующего равенства также умножатся на c. Отсюда следует корректность. Проверка выполнения аксиом бесхитростна.

Упр. 2.5. Возрастающая индукция по k, начинающаяся с $k = 0$, показывает, что все числа E_k лежат в (a, b), в частности, делятся на нод(a, b). С другой стороны, убывающая индукция по k, начинающаяся с $k = r + 1$, показывает, что все числа E_k (в том числе $E_0 = a$ и $E_1 = b$) делятся на E_r. Поэтому и нод$(a, b) = ax + by$ делится E_r.

Упр. 2.8. Существование. Если число n простое, то оно само и будет своим разложением. Если n составное, представим его в виде произведения строго меньших по абсолютной величине чисел, каждое из которых в свою очередь или просто или является произведением строго меньших по абсолютной величине чисел и т. д. Поскольку модуль целого числа нельзя бесконечно долго уменьшать, мы в конце концов получим требуемое разложение.

Единственность. Для любого простого числа p и любого целого числа z выполняется следующая альтернатива: либо нод$(z, p) = |p|$, и тогда z делится на p, либо нод$(z, p) = 1$, и тогда z взаимно прост с p. Пусть в равенстве $p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_m$ все сомножители просты. Поскольку $\prod q_i$ делится на p_1, число p_1, в силу лем. 2.3, не может быть взаимно просто с каждым q_i. Согласно упомянутой выше альтернативе, хотя бы один из множителей q_i (можно считать, что q_1) делится на p_1. Поскольку q_1 прост, $q_1 = \pm p_1$. Сокращаем первый множитель и повторяем рассуждение.

Упр. 2.9. Класс $\binom{mp^n}{p^n} \pmod p$ равен коэффициенту при x^{p^n}, возникающему после раскрытия скобок и приведения подобных слагаемых в биноме $(1 + x)^{mp^n}$ над полем $\mathbb{F}_p$. Последовательно применяя формулу форм. (2-19) на стр. 27, получаем

$$(1+x)^{p^n m} = \left((1+x)^p\right)^{p^{n-1}m} = \left(1+x^p\right)^{p^{n-1}m} = \left((1+x^p)^p\right)^{p^{n-2}m} = \left(1+x^{p^2}\right)^{p^{n-2}m} = \cdots$$
$$\cdots = \left(1+x^{p^n}\right)^m = 1 + mx^{p^n} + \text{ старшие степени}$$

Упр. 2.14. Ненулевой гомоморфизм полей инъективен, переводит единицу в единицу и перестановочен со сложением, вычитанием, умножением и делением[1]. Простое подполе состоит из эле-

[1] См. n° 2.6.4 на стр. 30.

ментов вида $\pm(1+\dots+1)/(1+\dots+1)$, каждый из которых остаётся на месте. Если имеется ненулевой гомоморфизм $\Bbbk \to \mathbb{F}$, то равенство или неравенство нулю суммы некоторого количества единиц в поле $\Bbbk$ влечёт точно такое же равенство или неравенство в поле $\mathbb{F}$, откуда $\operatorname{char}\Bbbk = \operatorname{char}\mathbb{F}$.

Упр. 2.15. Воспользуйтесь тем, что $\mathbb{R}$ является множеством дедекиндовых сечений линейно упорядоченного множества $\mathbb{Q}$.

Упр. 3.3. Ответ: $(y^n - x^n)/(y-x) = y^{n-1} + y^{n-2}x + y^{n-3}x^2 + \dots + yx^{n-2} + x^{n-1}$.

Упр. 3.5. $\left(a_0 + a_1x + a_2x^2 + \dots\right)^p = a_0^p + a_1^p x^p + a_2^p x^{2p} + \dots = a_0 + a_1x^p + a_2x^{2p} + \dots$ (первое равенство справедливо, поскольку возведение в p-тую степень перестановочно со сложением, второе — по малой теореме Ферма).

Упр. 3.6. Если $f(x) = \sum a_k x^k$, то $f(x+t) = \sum\limits_{k,\nu} a_k \binom{k}{\nu} \cdot x^{k-\nu}t^\nu = \sum\limits_\nu t^\nu \cdot f_\nu(x)$, где

$$f_\nu(x) = \sum_{k\geqslant\nu} a_k \binom{k}{\nu} \cdot x^{k-\nu} = \frac{1}{\nu!}\frac{d^k}{dx^k}\sum_{k\geqslant 0} a_k x^k\,.$$

Упр. 3.8. Годятся дословно те же аргументы, что и в упр. 2.8.

Существование. если f неприводим, то он сам и будет своим разложением, если f приводим, то он является произведением многочленов строго меньшей степени, которые в свою очередь или неприводимы или являются произведениями многочленов строго меньшей степени и т. д. Поскольку степень не может бесконечно уменьшаться, мы в конце концов получим требуемое разложение.

Единственность. Для любого приведённого неприводимого многочлена p и любого многочлена g выполняется следующая альтернатива: либо $\text{нод}(p,g) = p$, и тогда g делится на p, либо $\text{нод}(p,g) = 1$, и тогда g взаимно прост с p. Пусть в равенстве

$$p_1p_2\cdots p_k = q_1q_2\cdots q_m$$

все сомножители неприводимы. Деля p_1 на старший коэффициент, мы можем считать, что он приведён. Поскольку $\prod q_i$ делится на p_1, многочлен p_1, в силу лем. 2.3, не может быть взаимно прост с каждым q_i. Согласно упомянутой выше альтернативе, найдётся q_i (скажем, q_1), который делится на p_1. Так как q_1 неприводим, $q_1 = \lambda p_1$, где λ — ненулевая константа. Сокращаем первый множитель и повторяем рассуждение.

Упр. 3.11. Если многочлен степени $\leqslant 3$ приводим, то у него есть делитель степени один, корень которого будет корнем исходного многочлена.

Упр. 3.12. Единственность вытекает из сл. 3.3. Для отыскания такого многочлена заметим, что многочлен $\prod\limits_{\nu\neq i}(x-a_\nu)$ зануляется во всех точках a_ν кроме i-той, где он принимает ненулевое значение. Деля его на это значение, получаем такой многочлен $f_i(x) = \prod\limits_{\nu\neq i}(x-a_\nu)/\prod\limits_{\nu\neq i}(a_i-a_\nu)$, что

$$f_i(a_\nu) = \begin{cases} 1, \text{ при } \nu = i \\ 0, \text{ при } \nu \neq i\,. \end{cases}$$

Искомый многочлен равен $\sum\limits_{i=0}^n b_i\cdot f_i(x) = \sum\limits_{i=0}^n b_i \prod\limits_{\nu\neq i}(x-a_\nu)/(a_i-a_\nu)$.

Упр. 3.13. См. упр. 1.10 на стр. 12.

Упр. 3.14. Вложение $\varphi : \Bbbk \hookrightarrow \Bbbk[x]/(x-\alpha)$ в качестве констант сюръективно, поскольку число $\alpha \in \Bbbk$ переходит в класс $[x]$, и значит, для любого $g \in \Bbbk[x]$ число $g(\alpha)$ переходит в класс $[g]$.

Упр. 3.15. Обратным элементом к произвольному ненулевому $a+b\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$ является $\frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2}$. Кольцо в (а) содержит делители нуля: $[t+1]\cdot[t^2-t+1] = [0]$ и, тем самым, не является полем. Кольцо в (б) является полем: многочлен $p = \vartheta^3 + 2$ не имеет корней в $\mathbb{Q}$, и значит, не делится в $\mathbb{Q}[x]$ ни на какой многочлен первой или второй степени; следовательно, p взаимно прост со всеми $g \in \mathbb{Q}[x]$, не делящимися на p, т. е. для любого $[g] \neq [0]$ существуют $h_1, h_2 \in \mathbb{Q}[x]$, такие что $h_1 g + h_2 p = 1$; тем самым, $[h_1] = [g]^{-1}$.

Упр. 3.16. Указание: достаточно найти обратные ко всем элементам $\vartheta - a$, что делается по алгоритму Евклида[1] — класс $h(\vartheta)$, обратный к классу $\vartheta - a$, задаётся таким многочленом $h \in \mathbb{Q}[x]$, что

$$h(x)(x-a) + g(x)(x^2+x+1) = 1$$

для некоторого $g \in \mathbb{Q}[x]$. Поскольку остаток от деления x^2+x+1 на $x-a$ равен a^2+a+1, алгоритм Евклида остановится уже на втором шагу.

Упр. 3.18. Число $\zeta = \cos(2\pi/5) + i\cdot\sin(2\pi/5)$ является корнем многочлена

$$z^5 - 1 = (z-1)(z^4+z^3+z^2+z+1)\,.$$

Уравнение $z^4+z^3+z^2+z+1 = 0$ можно решить в радикалах, деля обе части на z^2 и вводя новую переменную $t = z + z^{-1}$.

Упр. 3.19. Пусть $\zeta = \zeta_1 = \cos(2\pi/n) + i\sin(2\pi/n)$ — первообразный корень с наименьшим положительным аргументом, и $\xi = \zeta^k$. Докажите более сильное утверждение: среди целых степеней корня ξ встречаются те и только те степени первообразного корня ζ, которые делятся на $\text{нод}(k,n)$, ибо равенство $\zeta^m = \xi^x$ означает, что $m = kx + ny$ для некоторого $y \in \mathbb{Z}$.

Упр. 3.20. См. листок № $3\frac{2}{3}$.

Упр. 3.21. Из равенства $z_1 z_2 = 1$ вытекает равенство $|z_1|\cdot|z_2| = 1$ на длины. Поскольку гауссово число $z \neq 0$ имеет $|z|^2 \in \mathbb{N}$, обратимым может быть только z с $|z| = 1$. Таких чисел в $\mathbb{Z}[i]$ ровно четыре: ± 1 и $\pm i$, и все они обратимы.

Упр. 3.24. Это сразу следует из теоремы теор. 6.1 на стр. 88 о существовании базиса в конечномерном векторном пространстве: если $\operatorname{char}\mathbb{F} = p$, то $\mathbb{F} \supset \mathbb{F}_p$ и является конечномерным векторным пространством над $\mathbb{F}_p$. Выбирая в нём базис $e_1, \ldots, e_n$, заключаем, что $\mathbb{F}$ состоит из p^n векторов $x_1 e_1 + \cdots + x_n e_n$, где каждый коэффициент x_i независимо пробегает $\mathbb{F}_p$ (см. **??** на стр. **??**). Менее геометрическое решение заключается в том, чтобы получить конечное поле $\mathbb{F}$ последовательными расширениями простого подполя $\mathbb{F}_p \subset \mathbb{F}$. Каждый шаг этого построения заключается в присоединении к очередному, уже построенному полю $\mathbb{L}$, такому что $\mathbb{F}_p \subset \mathbb{L} \subset \mathbb{F}$, какого-нибудь элемента $\zeta \in \mathbb{F} \smallsetminus \mathbb{L}$. Число элементов в получающемся поле $\mathbb{F}[\zeta] \supset \mathbb{L}$ является n-той степенью числа элементов в поле $\mathbb{L}$, откуда нужное утверждение следует по индукции.

Упр. 3.25. Равенство $(b_1 b_2)^k = 1$ равносильно равенству $b_1^k = b_2^{m_2-k}$. Тогда

$$b_2^{m_1(m_2-k)} = b_1^{m_1 k} = 1\,,$$

[1] См. n° 3.2.2 на стр. 39.

откуда $m_1(m_2-k)$ делится на m_2, а значит, k делится на m_2. В силу симметрии между b_1 и b_2, показатель k делится также и на m_1. А так как m_1 и m_2 взаимно просты, k делится на m_1m_2. Поскольку $(b_1b_2)^{m_1m_2}=1$, $\operatorname{ord}(b_1b_2)=m_1m_2$.

Упр. 3.26. Надо отправить в ℓ_1 все простые делители числа m_1, входящие в разложение числа m_1 в большей степени, чем в разложение числа m_2.

Упр. 3.27. Если $g(x)=h_1(x)\cdot h_2(x)$, то $h_1(\zeta)=0$ или $h_2(\zeta)=0$, поэтому степень одного из сомножителей не меньше, чем $\deg g$. Если $f(\zeta)=0$, то деля f на g с остатком: $f=gh+r$, и вычисляя при $x=\zeta$, получаем $r(\zeta)=0$. Так как $\deg r<\deg g$, заключаем, что $r=0$.

Упр. 3.28. Запишите элементы поля $\mathbb{F}_p$ в строку вида:

$$-[(p-1)/2]\,,\ \dots\ ,\ -[1]\,,\ [0]\,,\ [1]\,,\ \dots\ ,\ [(p-1)/2]$$

и покажите, что[1] $a\in\mathbb{F}_p^*$ тогда и только тогда является квадратом, когда число «положительных» чисел этой записи, становящихся «отрицательными» от умножения на a, чётно, после чего примените это к $a=2$.

Упр. 4.1. Равенство несократимых записей $p/q=r/s$ означает равенство $ps=qr$, в котором p взаимно прост с q, S взаимно прост с R, и q и S приведены. Из лем. 2.3 следует, что тогда $p=rf$, а $q=sg$ для некоторых $f,g\in\Bbbk[x]$. Равенство $frs=grs$ влечёт $f=g$. Поскольку $\text{нод}(p,q)=\text{нод}(rg,sg)=1$, многочлен g — обратимая константа, а т. к. q и S приведены, $g=1$.

Упр. 4.3. Согласно правилу дифференцирования композиции $\left(f^m\right)'=m\cdot f^{m-1}\cdot f'$, имеем $\frac{d}{dx}(1-x)^{-m}=\left(\left(\frac{1}{1-x}\right)^m\right)'=m(1-x)^{-m-1}$, откуда нужная формула легко получается по индукции.

Упр. 4.4. Воспользуйтесь форм. (3-9) на стр. 36 для производной композиции.

Упр. 4.5. Продифференцируйте обе части.

Упр. 4.9. Ответы: $a_1=\frac{1}{2}$, $a_2=\frac{1}{6}$, $a_3=0$, $a_4=-\frac{1}{30}$, $a_5=0$, $a_6=\frac{1}{42}$, $a_7=0$, $a_8=-\frac{1}{30}$, $a_9=0$, $a_{10}=\frac{5}{66}$, $a_{11}=0$, $a_{12}=-\frac{691}{2730}$,

$$S_4(n)=n(n+1)(2n+1)(3n^2+3n-1)/30$$
$$S_5(n)=n^2(n+1)^2(2n+1)(2n^2+2n-1)/12$$
$$S_{10}(1000)=91\,409\,924\,241\,424\,243\,424\,241\,924\,242\,500\,.$$

Упр. 5.1. Импликации (а)⇒(б)⇒(в) очевидны. Если I содержит обратимый элемент, то среди его кратных есть единица, кратные которой исчерпывают всё кольцо.

Упр. 5.2. Первое утверждение очевидно, второе вытекает из того, что все суммы вида $b_1a_1+\dots+b_ma_m$, где $a_1,\dots,a_m\in M$, $b_1,\dots,b_m\in K$, лежат во всех идеалах„ содержащих множество M.

Упр. 5.3. Если a и b являются старшими коэффициентами многочленов $f(x)$ и $g(x)$ из идеала I, причём $\deg f=m$ и $\deg g=n$, где $m\geqslant n$, то $a+b$ либо нуль, т. е. является старшим коэффициентом нулевого многочлена, либо является старшим коэффициентом многочлена $f(x)+x^{m-n}\cdot g(x)\in I$ степени m. Аналогично, для любого $\alpha\in K$ произведение αa является старшим коэффициентом многочлена $\alpha f(x)\in I$ степени m.

[1] Это утверждение известно как *лемма Гаусса о квадратичных вычетах*.

Упр. 5.4. Повторите доказательство теор. 5.1, следя за младшими коэффициентами вместо старших.

Упр. 5.6. Обозначим через I_0 идеал, образованный всеми аналитическими функциями[1], обращающимися в нуль на множестве $\mathbb{Z} \subset \mathbb{C}$, а через I_k — идеал всех функций, обращающихся в нуль на множестве $\mathbb{Z} \smallsetminus \{1,\ 2,\ \ldots\ ,\ k\}$. Убедитесь, что $\sin(2\pi z)/\prod_{\alpha=1}^{k}(z-\alpha) \in I_k \smallsetminus I_{k-1}$, откуда $I_k \subsetneq I_{k+1}$.

Упр. 5.7. Из того, что I является абелевой подгруппой в K немедленно вытекает, что отношение $a_1 \equiv a_2 \pmod{I}$ рефлексивно, транзитивно и симметрично. Корректность операций проверяется так же, как в упр. 1.10: если $[a']_I = [a]_I$ и $[b']_I = [b]_I$, т. е. $a' = a+x$, $b' = b+y$ с некоторыми $x, y \in I$, то $a'+b' = a+b+(x+y)$ и $a'b' = ab+(ay+bx+xy)$ сравнимы по модулю I с $a+b$ и ab соответственно, поскольку суммы в скобках лежат в I (именно в этот момент мы пользуемся тем, что идеал вместе с каждым элементом содержит и все его кратные); таким образом, $[a'+b']I = [a+b]_I$ и $[a'b']_I = [ab]_I$.

Упр. 5.8. Возьмите в качестве J^* объединение всех идеалов из M.

Упр. 5.9. В (а) всякий идеал в $\mathbb{C}[x]$ является главным. Если фактор кольцо $\mathbb{C}[x]/(f)$ не имеет делителей нуля, то многочлен f неприводим. Над полем $\mathbb{C}$ неприводимые многочлены исчерпываются линейными, поэтому $f(x) = x-p$ для некоторого $p \in \mathbb{C}$ и $(f) = (x-p) = \ker \mathrm{ev}_p$. В (б) с помощью леммы о конечном покрытии докажите, что для любого идеала I в кольце непрерывных функций $[0,1] \to \mathbb{R}$ найдётся точка $p \in [0,1]$, в которой все функции из I обращаются в нуль, что даст включение $I \subset \ker \mathrm{ev}_p$. В (в) подойдёт главный идеал $\mathfrak{m} = (x^2+1)$.

Упр. 5.11. Если в каждом идеале I_k есть элемент $x_k \in I_k \smallsetminus \mathfrak{p}$, то произведение этих элементов $x_1 \ldots x_m \in \bigcap I_k \subset \mathfrak{p}$, что противоречит простоте $\mathfrak{p}$.

Упр. 5.12. Рассмотрим эпиморфизм факторизации $\pi\colon K \twoheadrightarrow K/I$. Полный прообраз $\pi^{-1}(J)$ любого идеала $J \subset K/I$ является идеалом в K. Классы элементов, порождающих этот идеал в K порождают идеал J в K/I.

Упр. 5.13. Для колец (в) и (г) свойство (**??**) очевидно, поскольку модули всех ненулевых элементов не меньше 1, а свойство (5-6) вытекает из того, что для любого $z \in \mathbb{C}$ существует такой элемент кольца w, что $|z-w| < 1$. Беря такой w для $z = a/b$, получаем $|a-bw| < |b|$, так что можно положить $q = w$ и $r = a-bw$.

Упр. 5.14. Если $\exists\ b^{-1}$, то $\nu(ab) \leqslant \nu(abb^{-1}) = \nu(a)$. Наоборот, если $\nu(ab) = \nu(a)$, то деля a на ab с остатком, получаем $a = abq + r$, где либо $\nu(r) < \nu(ab) = \nu(a)$, либо $r = 0$. Из равенства $r = a(1-bq)$ вытекает, что либо $\nu(r) \geqslant \nu(a)$, либо $1-bq = 0$. С учётом предыдущего, такое возможно только при $1-bq = 0$ или $r = 0$. Во втором случае $a(1-bq) = 0$, что тоже влечёт $1-bq = 0$. Следовательно $bq = 1$ и b обратим.

Упр. 5.15. Если $b = ax$ и $a = by = axy$, то $a(1-xy) = 0$, откуда $xy = 1$.

Упр. 5.16. Многочлены x и y не имеют в $\mathbb{Q}[x,y]$ никаких общих делителей, кроме констант. Общими делителями элементов 2 и x в $\mathbb{Z}[x]$ являются только ± 1.

Упр. 5.17. По аналогии с комплексными числами, назовём *сопряжённым* к числу $\vartheta = a+b\sqrt{5}$ число $\overline{\vartheta} = a-b\sqrt{5}$, а целое число $||\vartheta|| \stackrel{\text{def}}{=} \vartheta\cdot\overline{\vartheta} = a^2-5b^2$ назовём *нормой* числа ϑ. Легко видеть, что $\overline{\vartheta_1\vartheta_2} = \overline{\vartheta}_1\cdot\overline{\vartheta}_2$, откуда $||\vartheta_1\vartheta_2|| = \vartheta_1\vartheta_2\overline{\vartheta}_1\overline{\vartheta}_2 = ||\vartheta_1||\cdot||\vartheta_2||$. Поэтому $\vartheta \in \mathbb{Z}[\sqrt{5}]$ обратим тогда и

[1] Функция $\mathbb{C} \to \mathbb{C}$ называется *аналитической*, если она задаётся сходящимся всюду в $\mathbb{C}$ степенным рядом из $\mathbb{C}[\![z]\!]$.

только тогда, когда $||\vartheta|| = \pm 1$, и в этом случае $\vartheta^{-1} = \pm\overline{\vartheta}$. Поскольку $||2|| = 4$, а $||1 \pm \sqrt{5}|| = -4$, разложение этих элементов в произведение xy с необратимыми x и y возможно только при $||x|| = ||y|| = \pm 2$. Но элементов нормы ± 2 в $\mathbb{Z}[\sqrt{5}]$ нет, так как равенство $a^2 - 5b^2 = \pm 2$ при редукции по модулю 5 превращается в равенство $a^2 = \pm 2$ в поле $\mathbb{F}_5$, где числа ± 2 не являются квадратами.

Упр. 5.20. Это следует из равенства $a_0 q^n + a_1 q^{n-1} p + \cdots + a_{n-1} q p^{n-1} + a_n p^n = 0$

Упр. 5.21. Ответ: $(x^2 - 2x + 2)(x^2 + 2x + 2)$.

Упр. 6.1. Пусть $0 \cdot v = w$. Тогда $w + v = 0 \cdot v + 1 \cdot v = (0 + 1) \cdot v = 1 \cdot v = v$. Прибавляя к обеим частям этого равенства $-v$, получаем $w = 0$. Из равенства $0 \cdot v = 0$ вытекает, что $x \cdot 0 = x(0 \cdot v) = (x \cdot 0) \cdot v = 0 \cdot v = 0$. Наконец, равенство $(-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = ((-1) + 1) \cdot v = 0 \cdot v = 0$ означает, что $(-1) \cdot v = -v$.

Упр. 6.3. Пусть $A \not\subseteq B$ — две подгруппы в абелевой группе. Выберем $a \in A \smallsetminus B$. Если $A \cup B$ является подгруппой, то $\forall\, b \in B\ a + b \in A \cup B$, но $a + b \notin B$, поскольку $a \notin B$. Следовательно, $a + b \in A$, откуда $b \in A$, т. е. $B \subseteq A$.

Упр. 6.4. Все проверки проводятся дословно также, как для классов вычетов по модулю идеала коммутативного кольца (ср. с упр. 5.7 на стр. 69).

Упр. 6.7. Если $\lambda' = \lambda + x$ и $a' = a + v$, где $x \in I$, $v \in IM$, то $\lambda' a' = \lambda a + (xa + \lambda v + xv)$, где взятая в скобки сумма лежит в IM.

Упр. 6.8. Если $x_1 m_1 = 0$ и $x_2 m_2 = 0$ для ненулевых $x_1, x_2 \in K$, то $x_1 x_2 (m_1 \pm m_2) = 0$ и $x_1 x_2 \neq 0$, так как в K нет делителей нуля. Кроме того, $\forall\, y \in K\ x_1(y m_1) = x_2(y m_2) = 0$.

Упр. 6.10. Ядро и образ любого гомоморфизма абелевых групп являются абелевыми подгруппами согласно n° 2.6 на стр. 28. Если гомоморфизм K-линеен, то обе эти подгруппы выдерживают умножение на элементы из K, поскольку $x\varphi(u) = \varphi(xu)$ и $\varphi(u) = 0 \Rightarrow \varphi(xu) = x\varphi(u) = 0$. Последнее утверждение является переформулировкой того, что $\varphi(v_1) = \varphi(v_2) \iff v_1 - v_2 \in \ker\varphi$. Убедитесь, однако, что отображение $[v] \mapsto \varphi(v)$ K-линейно.

Упр. 6.11. Сопоставьте отображению $\varphi : X \to M$ семейство его значений $\big(\varphi(x)\big)_{x \in X} \in \prod_{x \in X} M_x$.

Упр. 6.13. Прямая проверка: если f и g оба K-линейны, то $fg(xa + yb) = f(xg(a) + yg(b)) = xfg(a) + yfg(b)$ для любых скаляров x, y и векторов a, b.

Упр. 6.15. Если $x \in K \smallsetminus 0$ и $m \in M$ таковы, что $xm = 0$, то $x\varphi(m) = \varphi(mx) = \varphi(0) = 0$.

Упр. 6.16. Если векторы u_i ненулевые, то векторы $x_i u_i$ тоже ненулевые при любых $x_i \neq 0$. Поэтому любое нетривиальное линейное соотношение между векторами u_i является нетривиальным линейным разложением нулевого вектора в сумму векторов из подмодулей U_i.

Упр. 6.18. Множество всевозможных конечных K-линейных комбинаций счётного множества векторов равномощно $K \times \mathbb{N}$, т. е. дизъюнктному объединению счётного множества одинаковых копий множества K, тогда как множество $K[\![t]\!]$ равномощно множеству $K^{\mathbb{N}}$ всевозможных отображений $\mathbb{N} \to K$, которое строго мощнее, чем $K \times \mathbb{N}$ (используйте рассуждение Кантора).

Упр. 6.20. Очевидно, что E вкладывается в B_E, а B_E вкладывается в дизъюнктное объединение

$$\bigsqcup_{n \geqslant 1} \underbrace{E \sqcup \ldots \sqcup E}_{n}$$

счётного множества копий множества E, которое в силу того, что множество E бесконечно, равномощно E. Тем самым, B_E вкладывается в E. Остаётся применить теорему Кантора – Бернштейна.

Упр. 6.22. Линейность F вытекает из того, что отображение дифференцирования

$$d/dx\colon \Bbbk[x] \to \Bbbk[x], \quad g \mapsto g',$$

и все отображения вычисления $\mathrm{ev}_a\colon \Bbbk[x] \to \Bbbk$, $g \mapsto g(a)$, где $a \in \Bbbk$, линейны и композиция линейных отображений тоже линейна. Если $g \in \ker F$, то каждое число $a_i \in \Bbbk$ является как минимум $(m_i{+}1)$-кратным корнем многочлена g, и g делится на $\prod_i (x{-}a_i)^{m_i+1}$, что невозможно при $g \neq 0$, поскольку степень этого произведения равна $m + 1 > \deg g$.

Упр. 7.4. Пусть $AB = C$, $B^tA^t = D$, тогда $c_{ij} = \sum\limits_k a_{ik}b_{kj} = \sum\limits_k a^t_{ki}b^t_{jk} = \sum\limits_k b^t_{jk}a^t_{ki} = d_{ji}$.

Упр. 7.6. Пусть $AB = P$, $BC = Q$, тогда (i,j)-е элементы произведений PC и AQ равны друг другу: $\sum_k p_{ik}c_{kj} = \sum_k \sum_\ell (a_{i\ell}b_{\ell k})c_{kj} = \sum_{k\ell} a_{i\ell}(b_{\ell k}c_{kj}) = \sum_\ell a_{i\ell} \sum_k b_{\ell k}c_{kj} = \sum_\ell a_{i\ell}q_{\ell j}$.

Упр. 7.8. Поскольку $(AB)^t = B^tA^t$, матрица B обратна матрице A если и только если матрица B^t обратна матрице A^t.

Упр. 7.9. Прямое вычисление:

$$(a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21}) = \\ = (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21}).$$

Упр. 7.11. Первое доказывается выкладкой $0 \cdot a = (b + (-1) \cdot b)a = ba + (-1)ba = 0$, второе — выкладкой $e' = e' \cdot e'' = e''$.

Упр. 7.12. Ответ:

$$[E_{ij}, E_{k\ell}] \stackrel{\text{def}}{=} E_{ij}E_{k\ell} - E_{k\ell}E_{ij} = \begin{cases} E_{ii} - E_{jj} & \text{при } j = k \text{ и } i = \ell \\ E_{i\ell} & \text{при } j = k \text{ и } i \neq \ell \\ -E_{kj} & \text{при } j \neq k \text{ и } i = \ell \\ 0 & \text{в остальных случаях.} \end{cases}$$

Упр. 7.15. См. указания к упр. 7.11

Упр. 7.18. Можно воспользоваться тем, что

$$\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} b & a \\ 0 & c \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = \begin{pmatrix} b & 0 \\ d & c \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Упр. 8.1. В силу знакопеременности $\omega(\dots, u, \dots, u, \dots) = -\omega(\dots, u, \dots, u, \dots)$, откуда $2\omega(\dots, u, \dots, u, \dots) = 0$, что возможно только если $\omega(\dots, u, \dots, u, \dots) = 0$.

Упр. 8.2. Индукция по n. Каждая перестановка $g = (g_1, \dots, g_n)$ является композицией $g = \sigma \circ g'$ транспозиции σ, переставляющей между собою элементы n и g_n множества $\{1, 2, \dots, n\}$, и перестановки $g' = \sigma \circ g$, оставляющей на месте элемент n. По индукции, g' раскладывается в композицию транспозиций, не затрагивающих элемента n.

Упр. 8.3. $\max \ell(g) = n(n-1)/2$ достигается на единственной перестановке $(n, n-1, \dots, 1)$.

Упр. 8.5. Если все точки пересечения двойные и трансверсальные, две нити, выходящие из элементов i и j пересекаются между собою нечётное число раз если и только если (i,j) инверсна[1].

[1] В действительности картинку всегда можно нарисовать так, чтобы в этом случае была ровно одна точка пересечения.

Знак тасующей перестановки $(i_1, \dots, i_k, j_1, \dots, j_m)$ равен $(-1)^{|I|+\frac{1}{2}k(k+1)}$, где *вес* $|I| \stackrel{\text{def}}{=} \sum_\nu i_\nu$. Действительно, нити, выходящие из чисел $i_1, \dots, i_k$ верхней строчки не пересекаются между собою и пересекают, соответственно, $i_1 - 1$, $i_2 - 2$, ..., $i_k - k$ начинающихся левее нитей, выходящих из j-точек и тоже между собою не пересекающихся.

Упр. 8.7. Пусть модуль M порождается вектором e и $F : M \to M$ переводит эту образующую в $F(e) = \lambda e$, где $\lambda \in K$. Тогда для любого вектора $v = xe$ имеем $F(xe) = x\,F(e) = \lambda xe = \lambda v$.

Упр. 8.8. Всюду плотность множества $\mathcal{D}(f)$ означает, что в любой ε-окрестности[1] каждой точки $p \in \mathbb{R}^m$ найдётся точка $r \neq p$, в которой $f(r) \neq 0$. Так как многочлен f ненулевой, имеется точка $q \in \mathbb{R}^m$ с $f(q) \neq 0$. Ограничение f на прямую (pq), будучи ненулевым многочленом от одной переменной, обращается в нуль лишь в конечном числе точек.

Упр. 8.10. При чётном n центр алгебры $\Bbbk\langle\xi_1, \xi_2, \dots, \xi_n\rangle$ линейно порождается мономами чётных степеней, при нечётном n — мономами чётных степеней и старшим мономом $\xi_1 \wedge \dots \wedge \xi_n$, степень которого нечётна.

Упр. 8.11. Разложите определитель по первым n столбцам.

Упр. 8.12. Это сразу следует из равенства $\det A = \det A^t$.

Упр. 8.13. Если $A_{12} \neq 0$, то можно взять

$$A = \begin{pmatrix} 1 & 0 & -A_{23}/A_{12} & -A_{24}/A_{12} \\ 0 & A_{12} & A_{13} & A_{14} \end{pmatrix}.$$

Равенство

$$A_{34} = \det \begin{pmatrix} -A_{23}/A_{12} & -A_{24}/A_{12} \\ A_{13} & A_{14} \end{pmatrix}$$

эквивалентно квадратичному соотношению Плюккера[2].

Упр. 9.1. $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} = \frac{1}{\alpha\delta - \beta\gamma} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$.

Упр. 9.2. Рассмотрим в грассмановой алгебре $K\langle\xi_1, \xi_2, \dots, \xi_m\rangle$ два набора линейных форм $\eta = \xi \cdot A$ и $\zeta = \eta \cdot C = \xi \cdot F$, где $F = AC$. Тогда грассмановы мономы степени k от η и ζ суть $\eta_I = \sum_J \xi_J a_{JI}$ и $\zeta_K = \sum_L \xi_L f_{LK}$. Поскольку $\zeta_I = \sum_J \eta_J c_{JI}$, мы получаем $f_{LK} = \sum_J a_{LJ} c_{JK}$.

Упр. 9.3. Утверждения (а) и (б) очевидны. Пусть $0 < n < m$, как в (в). Если $\varphi^n(x) = 0$, то $p^n x = p^m y$ для некоторого $y \in K$, откуда $x = p^{m-n} y$, т. к. в K нет делителей нуля. Наоборот, если $x = p^{m-n} y$, то $p^n x = 0$ в $K/(p^m)$. Тем самым, $\ker \varphi^n = \operatorname{im} \varphi^{m-n}$. Правило $[x]_{p^n} \mapsto [p^{m-n}x]_{p^m}$ корректно задаёт инъективный гомоморфизмом K-модулей $\psi : K/(p^n) \hookrightarrow K/(p^m)$, который изоморфно отображает $K/(p^n)$ на $\operatorname{im} \varphi^{m-n} = \ker \varphi^n \subset K/(p^m)$. Это доказывает (в). Изоморфизм

$$\frac{\ker \varphi^n}{\ker \varphi^{n-1}} = \frac{p^{m-n}K/(p^m)}{p^{m-n+1}K/(p^m)} \xrightarrow{\sim} \frac{K}{p}$$

из (г) сопоставляет классу элемента $p^{n-m}x$ по модулю элементов вида $p^{n-m+1}y$ класс элемента x по модулю (p).

[1] Под ε-окрестностью точки $p \in \mathbb{R}^m$ мы понимаем m-мерный куб с центром в точке p и стороной 2ε.

[2] См. формулу (8-25) на стр. 120.

Упр. 9.4. Если $z' = z + q$, где $p^{i-1}q = 0$, а $x' = x + py$, то $x'z' = xz + q(x + py) + pyz$ и $p^{i-1}\big(q(x + py) + pyz\big) = 0$, поскольку $p^i z = 0$.

Упр. 10.1. Если отождествить $\mathbb{R}[t]/(t^2+1)$ с полем $\mathbb{C}$, отправив классы $[1]$ и $[t]$ в 1 и i соответственно, умножение на класс $[t]$ превратится в умножение на i, т. е. в поворот на угол $\pi/2$, который не переводит никакое одномерное векторное подпространство в себя.

Упр. 10.2. Пусть $\Bbbk[t]/\big(t^n\big) = U \oplus W$, где U и W переводятся в себя умножением на $[t]$. Оба этих подпространства не могут целиком содержаться в образе оператора умножения на $[t]$, так как иначе их сумма тоже бы в нём содержалась. Поэтому в одном из них, пусть это будет U, имеется класс $[g]$ многочлена g с ненулевым свободным членом. Тогда классы $[t^{n-1}g], \dots, [tg], [g] \in U$ выражаются через базис $[1], [t], \dots, [t^{n-1}]$ пространства $\Bbbk[t]/(t^n)$ при помощи верхнетреугольной матрицы, на диагонаи которой всюду стоит ненулевой свободный член многочлена g. Следовательно, эти классы тоже образуют базис в $\Bbbk[t]/(t^n)$, и значит, содержащее их подпространство U совпадает со всем пространством $\Bbbk[t]/(t^n)$.

Упр. 10.3. Разложите каждое пространство $(F|_{U_i}, U_i)$ по форм. (10-1) на стр. 132. В силу единственности такого разложения прямая сумма полученных разложений является разложением исходного пространства (F, V).

Упр. 10.4. В согласованном с разложением в прямую сумму базисе матрица $tE - F$ имеет блочно диагональный вид $\begin{pmatrix} tE - G & 0 \\ 0 & tE - H \end{pmatrix}$. С другой стороны, для любых матриц $A \in \mathrm{Mat}_n(\Bbbk)$, $C \in \mathrm{Mat}_m(\Bbbk)$, $B \in \mathrm{Mat}_{n\times m}(\Bbbk)$ определитель $\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \cdot \det C$ согласно формуле для разложения определителя по первым n столбцам.

Упр. 10.5. Пусть $f = t^n + a_1 t^{n-1} + \dots + a_{n-1}t + a_n$. Напишите матрицу F оператора умножения на класс $[t]$ в фактор кольце $\Bbbk[x]/(f)$ в базисе $[t^{n-1}], [t^{n-2}], \dots, [t], [1]$ и разложите $\det(tE - F)$ по первому столбцу.

Упр. 10.6. Поскольку умножение на произведение всех элементарных делителей полностью аннулирует прямую сумму форм. (10-1) на стр. 132, оператор $\chi_F(F)$ нулевой для любого оператора F над любым полем $\Bbbk$. Поскольку теорема Гамильтона-Кэли для матрицы A представляет собою набор тождеств между многочленами с целыми коэффициентами от элементов матрицы A, достаточно убедиться в её справедливости для всех матриц с рациональными элементами, т. е. для любого оператора над полем $\mathbb{Q}$.

Упр. 10.8. Векторы $g_i \in \Bbbk^n$ вычисляются рекурсивно по формулам $g_{m-1} = h_m$, $g_{i-1} = h_i + Ag_i$ при $i \leqslant m - 1$. Остаток $r = h_0 + F_v g_0 = h_0 + F_v(h_1 + F_v g_1) = h_0 + F_v(h_1 + A(h_2 + F_v g_2)) = \dots =$ $= h_0 + h_1 F_v + \dots + h_m F_v^m$ имеет степень 0 по t и тоже лежит в $\Bbbk^n$.

Упр. 10.10. Так как любой вектор $h \in H$ представляется в V как $h = u + q + r$ с $u \in U$, $q \in Q$, $r \in R$, в U выполняется равенство $h = \pi(h) = \pi(u) + \pi(r)$, в котором $\pi(u) = u \in U$ и $\pi(r) \in W$, т. е. $U + W = H$. Если $u \in U \cap W$, то $u = \pi(r)$ для некоторого $r \in R$, и $\pi(u - r) = \pi(u) - \pi(r) = u - u = 0$, откуда $u - r \in \ker \pi = Q$, что возможно только при $u = r = 0$. Поэтому $U \cap W = 0$.

Упр. 10.11. Если $\lambda \in \operatorname{Spec} F$ и $g(\lambda) \neq 0$, то $g(F)$ действует на ненулевом собственном подпространстве V_λ умножением на ненулевое число $g(\lambda)$. Тем самым, $g(F) \neq 0$.

Упр. 10.12. Над алгебраически замкнутым полем всякий многочлен имеющий только один корень 0 равен t^m. Поэтому $\chi_F(t) = t^m$ и по теореме Гамильтона – Кэли $F^m = 0$.

Упр. 10.15. Разложение характеристического многочлена оператора F в виде произведения степеней попарно разных линейных форм $\chi_F(t) = \prod_{\lambda\in\operatorname{Spec} F} (t-\lambda)^{N_\lambda}$ удовлетворяет условиям **??** с $q_i = (t-\lambda)^{N_\lambda}$, а корневые подпространства $K_\lambda = \ker(\lambda\operatorname{Id} - F)^{N_\lambda}$.

Упр. 10.16. Если $a^n = 0$, $b^m = 0$ и $ab = ba$, то $(a-b)^{m+n-1} = 0$ по формуле Ньютона.

Упр. 10.17. Над полем $\mathbb{C}$ можно применить предл. 10.9. Над произвольным полем $\Bbbk$ оператор F с матрицей $J_n(\lambda)$ имеет вид $\lambda\mathrm{Id} + N$, где $N^n = 0$, но $N^{n-1} \neq 0$. Обратный оператор

$$F^{-1} = (\lambda\mathrm{Id} + N)^{-1} = \lambda^{-1}(\mathrm{Id} + N/\lambda)^{-1} = \lambda^{-1} - \lambda^{-2}N + \lambda^{-3}N^2 - \cdots + (-1)^{n-1}\lambda^{-n}N^{n-1}$$

имеет вид $\lambda^{-1}\mathrm{Id} + M$, где оператор $M = -\lambda^{-2}N(1 - \lambda^{-1}N + \cdots)$ тоже имеет $M^n = 0$, а $M^{n-1} = \lambda^{2(1-n)}N^{n-1} \neq 0$. Таким образом, ЖНФ оператора F^{-1} это одна клетка $J_n\left(\lambda^{-1}\right)$.

Упр. 11.1. Если $fg = e$ и $gh = e$, то $f = fe = f(gh) = (fg)h = eh = h$.

Упр. 11.2. Для двух единичных элементов e' и e'' выполнены равенства $e' = e'e'' = e''$.

Упр. 11.4. Ответ: либо $r = 1$ и $\operatorname{Tors}(G) = 0$ (т. е. $G \simeq \mathbb{Z}$), либо $r = 0$ (т. е. G конечна) и каждое простое число $p \in \mathbb{N}$ присутствует в каноническом разложении

$$G = \frac{\mathbb{Z}}{\left(p_1^{n_1}\right)} \oplus \cdots \oplus \frac{\mathbb{Z}}{\left(p_\alpha^{n_\alpha}\right)}$$

не более одного раза. Доказательство аналогично доказательству предл. 10.3 на стр. 138.

Упр. 11.5. Пусть $k = dr$, $m = \operatorname{ord}(\tau) = ds$, где $\operatorname{нод}(r,s) = 1$. Если $d > 1$, то τ^d является произведением d независимых циклов длины s, и $\tau^k = \left(\tau^d\right)^r$ будет произведением s-тых степеней этих циклов. Остаётся показать, что когда $\operatorname{ord}(\tau) = m$ взаимно прост с k, то τ^k тоже цикл длины m. Если для какого-то элемента a цикла τ выполняется равенство $\left(\tau^k\right)^r(a) = a$, то kr делится на m, что при $\operatorname{нод}(k,m) = 1$ возможно только когда r делится на m. Поэтому $r \geqslant m$, т. е. длина содержащего a цикла перестановки τ^k не меньше m.

Упр. 11.6. Ответ: $n(n-1)\cdots(n-k+1)/k$ (в числителе дроби k сомножителей).

Упр. 11.7. Непересекающиеся циклы очевидно коммутируют. Если коммутирующие циклы τ_1 и τ_2 пересекаются по элементу a, то $\tau_1(a)$ является элементом цикла τ_2, поскольку в противном случае $\tau_2\tau_1(a) = \tau_1(a)$, а $\tau_1\tau_2(a) \neq \tau_1(a)$, так как $\tau_2(a) \neq a$. По той же причине $\tau_2(a)$ является элементом цикла τ_1, и значит, оба цикла состоят из одних и тех же элементов. Пусть $\tau_1(a) = \tau_2^s(a)$. Любой элемент b, на который оба цикла реально действуют имеет вид $b = \tau_2^r(a)$, и цикл τ_1 действует на него как τ_2^s:

$$\tau_1(b) = \tau_1\tau_2^r(a) = \tau_2^r\tau_1(a) = \tau_2^r\tau_2^s(a) = \tau_2^s\tau_2^r(a) = \tau_2^s(b)\,.$$

Второе утверждение следует из упр. 11.5.

Упр. 11.8. Ответ: $n!/\prod_{i=1}^{n} i^{m_i}m_i!$ (ср. с форм. (1-12) на стр. 10). Решение: сопоставим каждому заполнению диаграммы циклов λ неповторяющимися числами от 1 до n произведение независимых циклов, циклически переставляющих элементы каждой строки слева направо; получаем сюръективное отображение множества заполнений на множество всех перестановок циклового типа λ; прообраз каждой перестановки состоит из $\prod_{i=1}^{n} i^{m_i}m_i!$ заполнений, получающихся друг

из друга независимыми циклическими перестановками элементов в каждой строке и произвольными перестановками строк одинаковой длины между собою как единого целого.

Упр. 11.9. $|1,6,3,4\rangle^{15}\cdot|2,5,8\rangle^{15}\cdot|7,9\rangle^{15}=|1,6,3,4\rangle^{-1}\cdot|7,9\rangle=(4,\ 2,\ 6,\ 3,\ 5,\ 1,\ 9,\ 8,\ 7)$

Упр. 11.14. Ответ: $|1,2,3,4\rangle=\sigma_{12}\sigma_{23}\sigma_{34}$, $|1,2,4,3\rangle=\sigma_{12}\sigma_{24}\sigma_{34}$, $|1,3,2,4\rangle=\sigma_{13}\sigma_{23}\sigma_{24}$, $|1,3,4,2\rangle=\sigma_{13}\sigma_{34}\sigma_{24}$, $|1,4,2,3\rangle=\sigma_{24}\sigma_{23}\sigma_{13}$, $|1,4,3,2\rangle=\sigma_{34}\sigma_{23}\sigma_{12}$.

Упр. 11.15. Подсчёт для группы куба дословно тот же, что и для группы додекаэдра. Группы октаэдра и икосаэдра изоморфны группам куба и додекадра с вершинами в центрах граней октаэдра и икосаэдра соответственно.

Упр. 11.17. Зафиксируем в V какой-либо базис и сопоставим оператору $F\in\mathrm{GL}(V)$ базис, состоящий из векторов $f_i=F(e_i)$. Для выбора первого базисного вектора f_1 имеется $|V|-1=q^n-1$ возможностей, для выбора второго — $|V|-|\Bbbk\cdot f_1|=q^n-q$ возможностей, для выбора третьего — $|V|-|\Bbbk\cdot f_1\oplus\Bbbk\cdot f_2|=q^n-q^2$ возможностейи т. д.

Упр. 11.18. Подсказка: центральная симметрия коммутирует со всеми элементами полной группы додекаэдра; покажите, что единственная перестановка в S_5, коммутирующая со всеми перестановками из S_5 — это тождественное преобразование.

Упр. 11.23. Проиллюстрируем рассуждение на примере икосаэдра. И собственная и полная группы транзитивно действуют на 20 его треугольных гранях. Стабилизатор грани в собственной и полной группах представляет собой собственную и полную группу треугольника на плоскости, состоящую, соответственно из 3 и из 6 преобразований. По формуле для длины орбиты получаем $|\mathrm{SO}_{\text{ико}}|=20\cdot3=60$ и $|\mathrm{O}_{\text{ико}}|=20\cdot6=120$.

Упр. 11.25. Равенство $h_1g_1=h_2g_2$ влечёт равенства $g_2g_1^{-1}=h_2^{-1}h_1\in H$ и $g_1g_2^{-1}=h_1^{-1}h_2\in H$. С другой стороны, если один из обратных друг другу элементов $g_1^{-1}g_2$ и $g_2^{-1}g_1$ лежит в H, то в H лежит и второй, и $Hg_1=H(g_2g_1^{-1})g_2=Hg_2$.

Упр. 11.26. Включение $gHg^{-1}\subset H$ влечёт включение $H\subset g^{-1}Hg$. Если это так для всех $g\in G$, то заменяя g на g^{-1} мы получаем обратное к исходному включение $gHg^{-1}\supset H$.

Упр. 11.27. $\varphi\circ\mathrm{Ad}_g\circ\varphi^{-1}\colon\ h\mapsto\varphi\bigl(g\,\varphi^{-1}(h)\,g^{-1}\bigr)=\varphi(g)\,h\,\varphi(g)^{-1}$.

Упр. 11.28. Для любой точки $x\in\mathbb{R}^n$ положим $p=\varphi^{-1}(x)$. Так как $\varphi\colon\ \mathbb{R}^n\to\mathbb{R}^n$ аффинно, $\varphi(p+v)=x+D_\varphi(v)$. Поэтому $\varphi\circ\tau_v\circ\varphi^{-1}\colon\ x\mapsto\varphi(p+v)=x+D_\varphi(v)$.

Упр. 11.30. Если $\varphi(x)\in N_2$, то $\varphi(gxg^{-1})=\varphi(g)\varphi(x)\varphi(g)^{-1}\in N_2$ в силу нормальности $N_2\lhd G_2$. Поэтому $N_1=\varphi^{-1}(N_2)\lhd G_1$. Композиция сюръективных гомоморфизмов $G_1\twoheadrightarrow G_2\twoheadrightarrow G_2/N_2$ является сюръективным гомоморфизмом с ядром N_1.

Упр. 12.1. Первое очевидно, второе вытекает из того, что при вставке фрагментаЁ$x^\varepsilon x^{-\varepsilon}$ в произвольное слово w получится такое слово, в котором сокращение любого фрагмента вида $y^\varepsilon y^{-\varepsilon}$ приведёт либо обратно[1] к слову w, либо к слову, получающемуся из w сначала сокращением *того же самого фрагмента* $y^\varepsilon y^{-\varepsilon}$, а уже затем вставкой $x^\varepsilon x^{-\varepsilon}$ в *то же самое место*, что и в w.

Упр. 12.2. Отобразите $n\in\mathbb{N}$ в $x^nyx^n\in F_2$ и воспользуйтесь предл. 12.1 на стр. 169.

Упр. 12.3. Поскольку отображение $F\colon\ \mathbb{R}^n\to\mathbb{R}^n$ биективно, достаточно убедиться, что отображения $\sigma_{F(\pi)}$ и $F\circ\sigma_\pi\circ F^{-1}$ одинаково действуют на точку вида $F(p)$ с произвольным $p\in\mathbb{R}^n$.

Упр. 12.4. Обозначим через v_i вектор, идущий из центра симплекса Δ в вершину i. Вектор $n_{ij}=v_i-v_j$ ортогонален гиперплоскости π_{ij}, поскольку для любого $k\neq i,j$ скалярное произведение

[1] Обратите внимание, что такое происходит *не только* при сокращении того же самого фрагмента $x^\varepsilon x^{-\varepsilon}$, который был перед этим вставлен, но и при сокращении одной из букв $x^{\pm\varepsilon}$ с её соседкой.

$\bigl(n_{ij}, v_k - (v_i + v_j)/2\bigr) = (v_i, v_k) - (v_j, v_k) + (v_i, v_i)/2 - (v_j, v_j)/2 = 0$, т. к. все произведения (v_i, v_j) с $i \neq j$ и все скалярные квадраты (v_i, v_i) одинаковы. Аналогичная выкладка показывает, что при $\{i, j\} \cap \{k, m\} = \varnothing$ векторы n_{ij} и n_{km} ортогональны. Векторы $v_i - v_k$ и $v_k - v_j$ образуют в натянутой на них двумерной плоскости стороны правильного треугольника с вершинами в концах векторов v_i, v_j и v_k, и угол между ними равен 60°.

Упр. 12.8. При эпиморфизме S_4 на группу треугольника из прим. 11.9 подгруппа чётных перестановок $A_4 \subset S_4$ переходит в группу вращений треугольника.

Упр. 12.9. Примените изоморфизм $HN/N \simeq H/H \cap N$ из предл. 11.5 на стр. 168 для $G = D$, $H = B \cap D$ и $N = (A \cap D)C$ и воспользуйтесь тем, что $HN = (B \cap D)(A \cap D)C = (B \cap D)C$ и $H \cap N = (B \cap D) \cap (A \cap D) = (A \cap D)(B \cap C)$ (последнее равенство вытекает из того, что любой элемент $d = ac \in (B \cap D) \cap (A \cap D)$ с $d \in B \cap D$, $a \in A \cap D$, и $c \in C$ имеет $c = a^{-1}d \in C \cap B$).

Упр. 12.10. Правая часть равенства $|H| = 12\varepsilon_1 + 12\varepsilon_2 + 20\varepsilon_3 + 15\varepsilon_4 + 1$, приведённая по модулям 3, 4 и 5, равна, соответственно, $1 - \varepsilon_3$, $1 - \varepsilon_4$ и $1 + 2(\varepsilon_1 + \varepsilon_2)$. Она может делиться на 3 или на 4 только если $\varepsilon_3 = 1$ или $\varepsilon_4 = 1$. В обоих случаях $|H| \geqslant 16$, так что $|H|$ не может быть ни 3, ни 4, ни $3 \cdot 4$, ни $3 \cdot 5$. Если $|H|$ делится на 5, то $\varepsilon_1 = \varepsilon_2 = 1$ и $|H| \geqslant 25$, так что $|H|$ не может быть ни 5, ни $4 \cdot 5$. Остаются ровно две возможности: $|H| = 1$ и $|H| = 3 \cdot 4 \cdot 5$.

Упр. 12.11. Рассмотрим любое $k \notin i, j, g^{-1}(i)$. Тогда $g(k) = m \notin \{i, j, k\}$. При $n \geqslant 6$ найдётся чётная перестановка h, оставляющая на месте i, j, k и переводящая m в $\ell \neq m$. Тогда hgh^{-1} переводит i в j, а k — в $\ell \neq m$.

Упр. 12.12. Проверка ассоциативности:

$$\bigl((x_1, h_1) \cdot (x_2, h_2)\bigr) \cdot (x_3, h_3) = \bigl(x_1\psi_{h_1}(x_2),\ h_1h_2\bigr) \cdot (x_3, h_3) = \bigl(x_1\psi_{h_1}(x_2)\psi_{h_1h_2}(x_3),\ h_1h_2h_3\bigr)$$
$$(x_1, h_1) \cdot \bigl((x_2, h_2) \cdot (x_3, h_3)\bigr) = (x_1, h_1) \cdot \bigl(x_2\psi_{h_2}(x_3),\ h_2h_3\bigr) = \Bigl(x_1\psi_{h_1}\bigl(x_2\psi_{h_2}(x_3)\bigr),\ h_1h_2h_3\Bigr).$$

Но $\psi_{h_1}\bigl(x_2\psi_{h_2}(x_3)\bigr) = \psi_{h_1}(x_2)\psi_{h_1} \circ \psi_{h_2}(x_3) = \psi_{h_1}(x_2)\psi_{h_1h_2}(x_3)$. Существование единицы: $(x, h) \cdot (e, e) = (x, \psi_h(e), he) = (x, h)$, поскольку $\psi_h(e) = e$ в силу того, что ψ_h гомоморфизм. Существование обратного: $\bigl(\psi_h^{-1}(x^{-1}),\ h^{-1}\bigr) \cdot (x, h) = \bigl(\psi_h^{-1}(x^{-1})\psi_h^{-1}(x^{-1}),\ h^{-1}h\bigr) = (e, e)$.

Упр. 12.13. Так как $\psi\colon H \to \operatorname{Aut} N$ — гомоморфизм, $\psi_e = \mathrm{Id}_N$ и

$$(x_1, e) \cdot (x_2, e) = (x_1\psi_e(x_2), e) = (x_1x_2, e),$$

т. е. элементы (x, e) образуют подгруппу, изоморфную N. Она нормальна, поскольку

$$(y, h) \cdot (x, e) \cdot \bigl(\psi_h^{-1}(y^{-1}), h^{-1}\bigr) = (y\psi_h(x), h) \cdot \bigl(\psi_h^{-1}(y^{-1}), h^{-1}\bigr) = \bigl(y\psi_h(x)y^{-1}, e\bigr)\ .$$

Элементы (e, h) очевидно образуют дополнительную подгруппу, изоморфную H, и

$$\mathrm{Ad}_{(e,h)}(x, e) = (\psi_h(x), e)\,.$$

Упр. 12.14. Пусть центр $Z(G) = C$. Если $|C| = p$, то $C \simeq \mathbb{Z}/(p) \simeq G/C$. Пусть $a \in C$ — образующая центра, $b \in G$ — такой элемент, что смежный класс bC является образующей в G/C. Тогда любой элемент группы имеет вид $b^k a^m$. Так как a централен, любые два таких элемента коммутируют.

Упр. 12.15. Аддитивные автоморфизмы группы $\mathbb{Z}/(p)$ суть линейные автоморфизмы одномерного векторного пространства над полем $\mathbb{F}_p$. Они образуют группу $\mathrm{GL}_1(\mathbb{F}_p) \simeq \mathbb{F}_p^*$ ненулевых элементов поля $\mathbb{F}_p$ по умножению. Как и всякая конечная мультипликативная подгруппа поля, она циклическая.

Упр. 12.16. Корни многочлена $x^q - 1$ образуют в поле $\mathbb{F}_p$ мультипликативную подгруппу из $\leqslant q$ элементов, автоматически циклическую[1]. При $q \mid (p-1)$ многочлен $x^q - 1$ имеет ровно q корней $\eta = \zeta^{\alpha k}$, где $0 \leqslant \alpha \leqslant q-1$, а $\zeta \in \mathbb{F}_p^*$ — любая образующая циклической мультипликативной группы $\mathbb{F}_p^*$.

Упр. 12.17. Отображение $(n, h) \mapsto (n, \alpha^{-1}h)$ переводит сомножители из левой части равенства $(n_1, h_1)(n_2, h_2) = (n_1\psi_{h_1}n_2, h_1h_2)$ в $(n_1, \alpha^{-1}h_1)$ и $(n_2, \alpha^{-1}h_2)$, произведение которых в $N \rtimes_{\psi\circ\alpha} H$ равно $(n_1\psi_{h_1}n_2, \alpha^{-1}(h_1h_2))$. Отображение $(n, h) \mapsto (\beta n, h)$ переводит те же самые сомножители в $(\beta n_1, h_1)$ и $(\beta n_2, h_2)$. Их произведение в $N \rtimes_{\mathrm{Ad}_\beta(\psi)} H$ равно $(\beta(n_1\psi_{h_1}n_2), h_1h_2)$.

Упр. 13.3. Это переформулировка того, что форма $\beta : V \times V \to \Bbbk$ билинейна.

Упр. 13.4. Линейная оболочка векторов $e_\nu + ie_{n+\nu}$ с $1 \leqslant \nu \leqslant n$.

Упр. 13.5. Если матрица $B \in \mathrm{Mat}_n(\Bbbk)$ кососимметрична, то при нечётном n

$$\det B = \det B^t = \det(-B) = (-1)^n \det B = -\det B\,,$$

откуда $\det B = 0$ если $\mathrm{char}\,\Bbbk \neq 2$.

Упр. 13.6. Пусть $v = \sum x_i e_i$. Скалярно умножая v слева на ${}^\vee e_i$, получаем $\beta({}^\vee e_i, v) = x_i$. Скалярно умножая v справа на $e_i^\vee$, получаем $\beta(v, e_i^\vee) = x_i$, и т. д.

Упр. 13.8. В $\Bbbk[\![x]\!]$ квадрат ряда $\sqrt{(1+x)}$ равен $1+x$, а коэффициенты при x^k для $0 \leqslant k \leqslant n$ у квадрата ряда $\sqrt{(1+x)}$ такие же, как и у квадрата многочлена из условия.

Упр. 14.5. Ненулевые квадраты составляют образ гомоморфизма мультипликативных групп

$$\mathbb{F}_q^* \to \mathbb{F}_q^*, \quad x \mapsto x^2\,.$$

Так как уравнение $x^2 = 1$ имеет в поле $\mathbb{F}_q$ ровно два корня $x = \pm 1$, ядро этого гомоморфизма состоит из двух элементов, а значит, образ является подгруппой порядка $(q-1)/2$.

Упр. 14.6. Если оператор f самосопряжён, то $\beta_f(u, w) = (u, fw) = (fu, w) = (w, fu) = \beta_f(w, u)$. Если билинейная форма β_f симметрична, то $(fu, w) = (w, fu) = \beta_f(w, u) = \beta_f(u, w) = (u, fw)$

Упр. 14.8. Перестановка одной пары с другой как единого целого чётная (это пара транспозиций). Перестановка между собою элементов из ν-й пары меняет знак $\mathrm{sgn}\big(i_1j_1i_2j_2\ldots i_nj_n\big)$, но одновременно заменяет матричный элемент $a_{i_\nu j_\nu}$ элементом $a_{j_\nu i_\nu} = -a_{i_\nu j_\nu}$.

Упр. 15.2. Так как правые части форм. (15-7) на стр. 219 вещественны, $(u_i^\times, u_j) = \overline{(u_j, u_i^\times)} = \delta_{ij}$, что и означает равенство $u_j^{\times\times} = u_j$.

Упр. 15.3. Так как равенство из свойства (2) линейно по u, его достаточно проверить только на базисных векторах $u = u_k^\times$. Сделайте это.

Упр. 15.4. Равенства $(fu_i, w_j) = (u_i, f^\times w_j)$ означают равенство матрицы Грама $G_{f(\boldsymbol{u}),\boldsymbol{w}} = f(\boldsymbol{u})^t \cdot \boldsymbol{w}$ наборов векторов $f(\boldsymbol{u}) = \boldsymbol{w}\,F_{\boldsymbol{wu}}$ и $\boldsymbol{w}$ и матрицы Грама $G_{\boldsymbol{u},f^\times(\boldsymbol{w})} = \boldsymbol{u}^t \cdot f^\times(\boldsymbol{w})$ наборов векторов $\boldsymbol{u}$ и $f^x(\boldsymbol{w}) = \boldsymbol{u}\,F_{\boldsymbol{wu}}^\times$.

Упр. 15.6. Ответ: $a(t)\cdot\left(\frac{d}{dt}\right)^2 - (b(t) - 2a'(t))\cdot\frac{d}{dt} + (c(t) - b'(t) + a''(t))$.

Упр. 15.7. Рассмотрим $\mathrm{Mat}_n(\mathbb{C})$ как вещественное n^2-мерное векторное пространство с базисом E_{ij} и iE_{ij}, где E_{ij} — матрица с единицей в i-той строке j-того столбца и нулями в остальных местах. В координатах (x_{ij}, y_{ij}) относительно этого базиса матричное уравнение $f^t \cdot \overline{f} = E$, задающее

[1] См. предл. 3.10 на стр. 50.

унитарные матрицы $(f_{ij}) = (x_{ij}) + i \cdot (y_{ij})$, запишется системой системой квадратичных уравнений $\sum_\nu (x_{\nu i}^2 + y_{\nu i}^2) = 1$ (для каждого $i = 1, \ldots n$) и $\sum_\nu (x_{\nu i} x_{\nu j} + y_{\nu i} y_{\nu j}) = \sum_\nu (y_{\nu i} x_{\nu j} - x_{\nu i} y_{\nu j}) = 0$ (для всех $1 \leqslant i < j \leqslant n$). Поэтому множество U_n замкнуто. Складывая все уравнения первого типа, видим, что U_n находятся внутри единичного шара радиуса $\sqrt{n}$ с центром в начале координат, и значит, компактно. Диагональная матрица D с диагональными элементами вида $e^{i\vartheta}$ очевидно соединяется с единичной матрицей гладким путём $\gamma : [0,1] \to \mathrm{U}_n$, образ которого целиком состоит из диагональных матриц того же вида (надо просто согласованно устремить все ϑ к нулю). Поскольку произвольная унитарная матрица f записывается как $f = CDC^{-1}$ для некоторого $C \in \mathrm{U}_n$, путь $t \mapsto C \cdot \gamma(t) \cdot C^{-1}$ будет целиком лежать в U_n и соединять f с E.

Упр. 15.9. Так как оператор $ff^\times$ самосопряжён и биективен, все его собственные числа строго положительны. Поэтому имеется единственный самосопряжённый оператор h с положительными собственными значениями, квадрат которого равен[1] $ff^\times$. Тогда $f = hr$, где $r = h^{-1}f$ унитарен, поскольку $r^\times r = r^\times h^{-2} f = f^x (ff^\times)^{-1} f = \mathrm{Id}_W$.

[1] Так как h и h^2 диагонализуются в одном базисе, оператор h обязан действовать на каждом собственном подпространстве V_λ оператора h^2 умножением на положительный $\sqrt{\lambda}$.